

Fractals in Geophysics

Edited by
Christopher H. Scholz
Benoît B. Mandelbrot

Springer Basel AG

B

Fractals in Geophysics

Edited by
Christopher H. Scholz
Benoît B. Mandelbrot

1989

Springer Basel AG

Reprint from Pure and Applied Geophysics
(PAGEOPH), Volume 131 (1989), No. 1/2

Editors' addresses:

Christopher H. Scholz
Lamont-Doherty Geological Observatory
Columbia University
Palisades, NY 10964
USA

Benoît B. Mandelbrot
IBM, T. J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598
USA

Library of Congress Cataloging in Publication Data

Fractals in geophysics / edited by Christopher H. Scholz, Benoît B. Mandelbrot.

p. cm.

»Reprint from Pure and applied geophysics (PAGEOPH), volume 131
(1989), no. 1/2« - - T. p. verso.

Bibliography: p.

ISBN 978-3-0348-6391-9 (U.S.: alk. paper)

1. Fractals. 2. Geology - - Mathematics. I. Scholz, C. H.
(Christopher H.) II. Mandelbrot, Benoît B. III. Pure and applied
geophysics.

QE 33.2.M3F73 1989 551 - - dc20 89-31640

CIP-Titelaufnahme der Deutschen Bibliothek

Fractals in geophysics / ed. by Christopher H. Scholz; Benoît
B. Mandelbrot. – Reprint. – Basel ; Boston ; Berlin :
Birkhäuser, 1989

Aus: Pure and applied geophysics ; Vol. 131

ISBN 978-3-0348-6391-9 (Basel ...) brosch.

ISBN 978-3-0348-6391-9 (Boston) brosch.

NE: Scholz, Christopher H. [Hrsg.]

This work is subject to copyright. All rights are reserved, whether the whole or part
of the material is concerned, specifically those of translation, reprinting, re-use of
illustrations, broadcasting, reproduction by photocopying machine or similar means,
and storage in data banks. Under § 54 of the German Copyright Law where copies are
made for other than private use a fee is payable to »Verwertungsgesellschaft Wort«, Munich.

Printed on acid-free paper

© Springer Basel AG 1989

Originally published by Birkhäuser Verlag Basel in 1989

ISBN 978-3-0348-6391-9 ISBN 978-3-0348-6389-6 (eBook)

DOI 10.1007/978-3-0348-6389-6

Contents

- 1 Introduction, *C. H. Scholz and B. B. Mandelbrot*
- 5 Multifractal measures, especially for the geophysicist, *B. B. Mandelbrot*
- 43 The fractal geometry of interfaces and the multifractal distribution of dissipation in fully turbulent flows, *K. R. Sreenivasan, R. R. Prasad, C. Meneveau and R. Ramshankar*
- 61 Fractal velocity models in refraction seismology, *D. J. Crossley and O. G. Jensen*
- 77 Variations in geometric measures of topographic surfaces underlain by fractured granitic plutons, *D. Norton and S. Sorenson*
- 99 Fractal sinuosity of stream channels, *R. S. Snow*
- 111 The fractal geometry of flow paths in natural fractures in rock and the approach to percolation, *D. D. Nolte, L. J. Pyrak-Nolte and N. G. W. Cook*
- 139 Testing linear models of sea-floor topography, *A. Malinverno*
- 157 Fractal dimension of fault systems in Japan: Fractal structure in rock fracture geometry at various scales, *T. Hirata*
- 171 Fractals in geology and geophysics, *D. L. Turcotte*
- 197 Fractal reconstruction of sea-floor topography, *J.-C. Mareschal*
- 211 Empirically derived relationships between fractal dimension and power law form frequency spectra, *C. G. Fox*
- 241 Are topographic data sets fractal?, *L. E. Gilbert*
- 255 Fractals, fault-gouge and friction, *C. G. Sammis and R. L. Biegel*
- 273 Scaling geology and seismic deconvolution, *J. P. Todoeschuk and O. G. Jensen*
- 289 Fractured but not fractal: Fragmentation of the Gulf of Suez basement, *G. Korvin*
- 307 Fractal analysis of long-range paleoclimatic data: Oxygen isotope record of Pacific core V28-239, *R. H. Fluegeman, Jr., and R. S. Snow*

Introduction

Many geological objects are jagged over a wide range of scales, so that their boundaries look the same when viewed at different magnifications or reductions. So too the objects themselves may exist at many different scales at which their appearance is very similar. Geologists have been intuitively aware of this basic fact for a very long time—hence the unspoken but inflexible rule that a photograph of any geological object must contain a scale reference, which in the case of photographs taken in the field is usually a common object of known size such as a lens cap, rock hammer or human figure.

This fundamental and pervasive scale invariance of geological objects may now be studied within the framework of fractal geometry. This truly represents a new paradigm for the practicing geologist or geophysicist, because it teaches one to view objects as part of fractal sets that obey certain scaling rules over a range of scales that is much broader than one is used to dealing with at any one time, and to provide sound reasons to expect such rules to hold. As examples of these rules, the size distribution of every characteristic of a fractal object is a power law, and a record of a fractal boundary or time series has a power spectrum that is also a power law, the exponents being in each case related to a fractal dimension.

This volume brings together several concrete problems of geophysics and several mathematical techniques of fractal geometry. Let us argue that this broad convergence is *good*, that it had been *inevitable*, and that it ought to act as a spur to *theory*. We shall not, however, make any attempt to forecast the future.

The convergence of fractals and of geophysics is something good because it answers an ancient need. To quote from BBM's 1982 book *The Fractal Geometry of Nature*, "mountains are not cones, coastlines are not circles, nor does lightning travel in a straight line." More generally, the shapes of Euclid are so far from the shapes of geophysics that they fail to provide even a first approximation. As to the context of the distribution of mass, the counterparts of Euclid are the uniform distribution or other distributions with a density that varies gradually in space. They are very far from the extremely fragmented distributions of such geophysical quantities as minerals, hence they also fail to provide even a first approximation. If not Euclid, then what? The only geometric alternative used to be "true chaos," namely the absence of quantitative natural law, and the reign of qualitative natural history. This is an extremity whose continuation a scientist finds horrible to contemplate! Moreover, the absence of realistic quantitative models used to create a real danger. The mathematically oriented scientist has an understandable but unfortunate tendency to try and forge ahead using unrealistic models that are the

"best one has." This is perhaps why some examples of applied mathematics have acquired the reputation of being as removed from nature as the "pure" kind.

In this environment, the major asset of fractal geometry consists in its having introduced an intermediate possibility between excessive order and true chaos. Fractal sets were first used in truly geometric fashion in BBM's 1967 study of coastlines, which certainly should be viewed as related to geophysics. And as models of distribution multifractal measures were first fully implemented in geometric fashion in BBM's 1974 study of the intermittency of turbulence, which is taken up again in his contribution to the present book.

The most striking feature of fractals is that all look very complex, but in fact many involve very simple formulas with very few parameters. Compared with true chaos, fractals and multifractals would have been viewed as pure good, were it not that they require fresh technique and fresh thought. The goal of the present *Special Issue* is to bring out those thoughts and techniques in the context of geophysics, so as to encourage new recruits to jump in.

Now to a second basic question. Good or bad, *is it true* that fractals and multifractals actually apply in geophysics? Even if the well-known "forgeries" of mountains or clouds had not referred to objects of geophysics, a casual inspection of these illustrations suffices to bring to mind many shapes familiar to every earth scientist. As proof, let us translate from an elegant and still timely review (unfortunately unsigned) that the *Annales de Géophysique* for 1976 gave of the first book on this topic, *Les objets fractals*.

"The author pursues two main goals in this work, and is successful in both. The first is to demolish the myth of the abstract mathematical concept with no concrete counterpart. . . The reader is astonished to find that, after all, he encounters every day objects whose dimension is not an integer! Having succeeded in this goal, the author seeks to stimulate. And he succeeds very well. The geophysicist who jumps into this Essay experiences the utmost difficulty to avoid daydreaming. Indeed, the crust of the earth having faults and fractures at all scales (from the scale of the continents down to the scale of crystals), can we not imagine that it is fractal. . . And a seismic fault, so simple from afar but so complex from close by, and practically beyond comprehension if one examines at the scale of the granular material that fills it, would it also be [a fractal]. . .

"We would very much like to do physics starting with these concepts. The author rouses the curiosity and fires the imagination of the reader, but abandons him helpless when the task is to go beyond elegant description. . . One can only hope that the use of a tool so nicely offered will not pose innumerable problems to the "classical" physicist or the naturalist."

These last words bring us to the third basic question to be tackled. Good or bad, obvious or not to the eye, *can we explain* the relevance of fractals to geophysics? Since fractals' "coming out" in 1975, the study of physics on fractals and of the physics of fractals have made very broad progress. the number of already explained

fractal phenomena has kept increasing, but so has the number of unexplained new ones. As a result, it is still true that a detailed explanation continues to be lacking in most cases. Let us briefly mention the two overall approaches that are characteristic of some fractals in every field, and are illustrated by the backs and the fronts of the jackets of BBM's books: the approaches via randomness and via chaotic dynamics.

One possible broad explanation of the role of fractals in geophysics may be found in probabilistic limit theorems, and in the existence of classical "universality classes" related to them. The reason is illustrated by the following fact. Wiener's scalar Brownian motion process $W(t)$ is the limit of the linearly rescaled random processes that belong to its very wide domain of attraction. Therefore, it is itself the fixed point of the rescaling process. That is, its graph is a self-affine fractal set, a curve. The argument suggests that geometric shapes relative to probabilistic limit theorems can be expected to be fractal sets.

A second possible broad explanation of the role of fractals in geophysics may come from deterministic chaotic dynamics. The field of mechanics used to be faced with the hard choice between excessive order, which characterized the problems that had been solved in the 18th century and have continued to fill the textbooks, and the remaining problems, about which little was known, and which could be called "nearly true chaos." They included such topics as the study of turbulence. But today, we know of a remarkable third intermediate possibility, deterministic chaos, which H. Poincaré had glimpsed long ago, and which has been developed in beautiful detail in recent years. Deterministic chaos and fractals arose independently, but it is no wonder that they have joined focus and now develop in parallel, with innumerable interactions, so that some observers view them as parts of the same broader endeavor. Chaotic dynamics has laid stress on chaos in the temporal behavior of systems and the resulting "strange attractors" and "strange repellers," but the study of these objects unavoidably brings in fractal sets and multifractal measures.

Of the interactions between fractals and chaos, one aspect is particularly relevant to geophysics, because it mixes the statistical and the dynamical faces of fractals. Let us accept for a moment, as an article of faith, the notion that all the complications of the time and space patterns of geophysics result from deterministic chaos. The ultimate goal would then be to unscramble these deterministic rules, on the basis of their strange attractors. The strange attractors are fractal sets, on which a multifractal measure is observed. Do we know that such unscrambling is possible, even on principle? We do not really know. How much information can conceivably be extracted from a given fractal set or multifractal measure? The present answer is, only a few numbers. Well, we surely do not lack problems to work on. Let us therefore stop philosophizing and set out to move on the field of fractals in geophysics.

During the past five years or so there have been an increasing number of studies

of the fractal nature of geological phenomena in many diverse fields of the earth sciences. In recognition of this, a symposium on "Fractals in Geophysics" was held at the fall meeting of the American Geophysical Union in San Francisco in December, 1987. This volume contains papers read at that meeting, as well as a few papers contributed later. The book does not pretend to provide a comprehensive view, only a fair sampling of the work being done in this rapidly expanding field.

The importance of multifractals is great and increasing. Mandelbrot describes them, using the approach he has been developing since the early 1970's. Sreenivasan *et al.* discuss multifractals and turbulence.

Geologists often work with maps, and the fractal nature of topography is the subject of a number of papers in this volume. Malinverno and Mareschal both discuss the fractal nature of sea-floor topography, and Gilbert expands that to include continental topography as well. Both Gilbert and Fox point out difficulties with straightforward analysis of self-affine fractals like topographic profiles.

Earthquakes have been long known to have the power law size distribution characteristic of fractals. In the paper by Hirata, the size distribution is extended to faults, which are the geological objects on which earthquakes are generated. In contrast, Korvin provides an example of a fracture system that he claims is not fractal. Norton and Sorenson make a connection between the fractal nature of topography and that of the fractures that underly it, and Nolte *et al.* describe a model for fluid flow through a medium containing a fractal network of fractures.

Turcotte provides a potpourri of fractal geophysical objects and phenomena. Among these are comminuted materials, a subject that is taken up in more detail by Sammis and Biegel, who show that communication during faulting produces gouge with a fractal distribution. A collection of the examples given so far provides an interesting hierarchy of fractal phenomena: tectonic topography—earthquakes—faults—fault surface topography—fault gouge.

Somewhat different phenomena were studied by Snow, who applies fractal analysis to meandering stream geometry, and Fluegeman and Snow, who use it to study that rather peculiar geological time series—data from a deep-sea core. Finally, we are beginning to see methodology developed for analysing geophysical data for an inhomogeneous medium that obeys fractal rules; namely in the papers by Todoeschuck & Jensen and Crossley & Jensen, who study the analysis of seismic wave propagation through such a medium. As the other studies show, this is a far more realistic way of representing the inhomogeneity of the earth.

February, 1989.

C. H. Scholz

Lamont-Doherty Geological Observatory
Columbia University
Palisades, NY 10964, USA

B. B. Mandelbrot

IBM, T. J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598, USA

Multifractal Measures, Especially for the Geophysicist*

BENOÎT B. MANDELBROT^{1,2}

Abstract—This text is addressed to both the beginner and the seasoned professional, geology being used as the main but not the sole illustration. The goal is to present an alternative approach to multifractals, extending and streamlining the original approach in MANDELBROT (1974). The generalization from fractal sets to multifractal measures involves the passage from geometric objects that are characterized primarily by one number, namely a fractal dimension, to geometric objects that are characterized primarily by a function. The best is to choose the function $\rho(\alpha)$, which is a limit probability distribution that has been plotted suitably, on double logarithmic scales. The quantity α is called Hölder exponent. In terms of the alternative function $f(\alpha)$ used in the approach of Frisch-Parisi and of Halsey *et al.*, one has $\rho(\alpha) = f(\alpha) - E$ for measures supported by the Euclidean space of dimension E . When $f(\alpha) \geq 0$, $f(\alpha)$ is a fractal dimension. However, one may have $f(\alpha) < 0$, in which case α is called “latent.” One may even have $\alpha < 0$, in which case α is called “virtual.” These anomalies’ implications are explored, and experiments are suggested. Of central concern in this paper is the study of low-dimensional cuts through high-dimensional multifractals. This introduces a quantity D_q , which is shown for $q > 1$ to be a critical dimension for the cuts. An “enhanced multifractal diagram” is drawn, including $f(\alpha)$, a function called $\tau(q)$ and D_q .

Key words: Fractal, multifractal, measure, Hölder, limit theorem.

1. Introduction and Motivation. Reasons Why Multifractals are Indispensable in Geophysics and in Other Sciences

The topic of multifractals is bound to become of increasing importance to geophysics, in particular if the present volume becomes influential.

In one phrase, the generalization from fractal sets to multifractal measures involves the passage from geometric objects characterized primarily by one number, to geometric objects characterized primarily by a function. This function can be a probability distribution that has been renormalized and plotted suitably.

In a different single phrase, the generalization for fractal sets to multifractal measures involves the passage from a finite number of fractal dimensions to an

* Note: This text incorporates and supersedes MANDELBROT (1988). A more detailed treatment, in preparation, will incorporate MANDELBROT (1989).

¹ Physics Department, IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, U.S.A.

² Mathematics Department, Yale University, New Haven, CT 06520, U.S.A.

infinite number of “dimensions.” Moreover (and this is a special point of this paper), these “dimensions” can be negative. We shall have gone far from the integer-valued dimensions of Euclid.

1.1. The Example of Copper

Correspondent to the simplest fractals, the basic idea is self-similarity, either exact or approximate. The closely related notions of self-similar fractal or self-similar multifractal can be phrased in many ways, but the geophysicist might best understand them in the context of the distribution of a rare mineral, such as copper. We first consider high-grade copper, then consider gradually lower grades.

High-grade copper is of course distributed nonuniformly: it concentrates in very few regions of the world. If one examines one such region in detail, however, copper continues to be found to be nonuniform: it concentrates in few sub-regions. And so on. It is reasonable, therefore, to suppose that the relative distribution of high-grade copper is the same (in the statistical sense) within each copper-bearing region, whether it is small or large. This being granted, take a large region, and cover it by a grid of equal “cells.” As the cell sizes are made smaller, the total area of the cells that contain high-grade copper is found to shrink.

Mathematics has long known a construction that follows this process, and fractal geometry has “tamed” this construction to make it a model of nature. For example see MANDELBROT (1982), *The Fractal Geometry of Nature (FGN)*. In the language of fractal geometry, high-grade copper is usefully viewed as “concentrated on,” or “supported by,” a self-similar fractal set of low fractal dimension.

Next, examine lower grade copper. The fact that it is more widespread in nature is expressed by its being supported by a fractal set of higher fractal dimension.

Overall, in order to give a full representation of the distribution of copper, it is seen that fractals are necessary and that no single fractal set is sufficient. A simple description consists of giving the fractal dimensions corresponding to each of a sequence of grades, as defined by thresholds varying from 0 to a very high value that is rarely exceeded.

The overall idea of the preceding paragraph has been combined with the generalization of the notion of self-similarity from sets to measures, and has thereby led to the notion of self-similar multifractal *measure*. To say that a multifractal is a *measure* and not a *set* is a very important distinction. It will be explained in Section 3.1. Our work on multifractals was initially concerned with the intermittency of turbulence, and was mostly carried out in the period 1968 to 1976, but it had started about 1962. My book *FGN* surveys multifractals on pp. 375–376, but this survey is overly sketchy and is now obsolete.

1.2. An Interesting Old Quote

The simplest of all multifractals, which is nonrandom, is called binomial and is

discussed in Section 5. This construction that has long been known to mathematicians, and has been tamed by fractal geometry, to make it a model of Nature. It happens that the basic circumstances that call for the binomial multifractal measure are very intuitive, and have nearly been rediscovered in the earth sciences context described in Section 1.1. Indeed, the geologist DE WIJS (1951) (quoted in *FGN*, p. 376) has described them as follows:

“Consider a [body of ore] with a tonnage W and an average grade M . With an imaginary cut we slash this body into two halves of equal tonnage $\frac{1}{2}W$, differing in average grade. Accepting for the grade of the richer half $(1+d)M$, the grade of the poorer half has to be $(1-d)M$ to satisfy the condition that the two halves together average again M . . . A second imaginary cut divides the body into four parts of equal tonnage $\frac{1}{4}W$, averaging $(1+d)^2M$, $(1+d)(1-d)M$, $(1+d)(1-d)M$, and $(1-d)^2M$. A third cut produces $2^3 = 8$ blocks, namely 1 block with an average grade of $(1+d)^3M$, 3 blocks of $(1+d)^2(1-d)M$, 3 blocks of $(1+d)(1-d)^2M$, and one block of $(1-d)^3M$. One can visualize the continued division into progressively smaller blocks . . . The coefficient d as a measure of variability adequately replaces the collective intangibles [dear to those who feel that ore estimation is an art rather than a science], and statistical deductions based upon this measure can abolish the maze of empirical and intuitive techniques.”

Of course, de Wijs did not even begin to explore the geometric aspects of his first sketch of a model, and neither he nor notable followers (including G. Matheron) had an inkling of fractals or of multifractals, e.g., of the basic notion of fractal dimension. However, assume that the ore density is independent of grade, making tonnage equivalent to volume, and allow the (reinterpreted) scheme of de Wijs to continue *ad infinitum*. We shall see that this leads to the conclusion that the ore “curdles” into a binomial multifractal.

1.3. Relative Intermittency in a Context Broader than that of Metals

To broaden the scope of multifractals, let us quote from the subsection on *Relative Intermittency* of my book *FGN*, p. 375 ss.

“The phenomena to which [multi] fractals are addressed are scattered throughout this Essay, in the sense that many of my case studies of natural fractals negate some unquestionable knowledge about Nature.

“We forget in Chapter 8 that the noise that causes fractal errors weakens between errors but does not desist.

“We neglect in Chapter 9 our knowledge of the existence of interstellar matter. Its distribution is doubtless *at least* as irregular as that of the stars. In fact, the notion that it is impossible to define a density is stronger and more widely accepted for interstellar than stellar matter. To quote deVaucouleurs, ‘it seems difficult to believe that, whereas visible matter is conspicuously clumpy and clustered on all

scales, the invisible intergalactic gas is uniform and homogeneous . . . [its] distribution must be closely related to . . . the distribution of galaxies . . . ?

“And in Chapter 10 the pastry-like sheets of turbulent dissipation are an obviously oversimplified view of reality.

“The end of Chapter 9 mentions very briefly the fractal view of the distribution of minerals. Here, the use of closed fractals implies that, between the regions where copper can be mined, the concentration of copper vanishes. In fact, it is very small in most places, but cannot be assumed to vanish everywhere.

“In each case, [portions of space] of less immediate interest were artificially emptied to make it possible to use *closed* fractal sets, but eventually these areas must be filled. This can be done using a fresh hybrid [namely, a] mass distribution in the cosmos such that no portion of space is empty, but, [given two] small thresholds θ and λ , a proportion of mass at least $1 - \lambda$ is concentrated on a portion of space of relative volume at most θ . ”

1.4. A Feature of Most Direct Importance in Many Sciences: Many Measures are not Observable Directly, only Through “Cuts”

The exploration of the earth cannot be carried out fully in three dimensions. Very often it must follow a straight bore-hole to obtain a straight 1-dimensional cross-cut through a real system that is intrinsically 3-dimensional. Often flat cuts are all that is available for inspection. The same constraint is encountered when turbulence in 3-dimensional space is explored via 1- or perhaps 2-dimensional cuts. Typically, the positions of these cuts bear no relation to the overall turbulence, and can therefore be thought of as having been chosen at random.

Consider also the context of strange attractors. Their full natural space has a very high dimensionality. But they are typically examined via a “Poincaré section” by a plane. The position of the plane, again, often bears no relation to the full attractor, and can be viewed as having been chosen at random. As we shall see by examining typical cases, the measure observed along a random cut has properties that are without counterpart in the measures studied in their natural space, and vice versa. This raises the issue of what can and what cannot be inferred from a cut to the whole measure. This issue is extraordinarily important and has motivated our early work of 1968–1976, especially MANDELBROT (1974).

2. Two Alternative Summaries

The present text begins with introductory material, continues with the binomial measure (Section 5) and then proceeds to step by step generalizations. One must wait until Section 8 to initiate a discussion of the cuts. The result is longer than we would have preferred.

We hope to interest both the reader who is not yet fully familiar with multifractals, and the reader who is already familiar with them via the variant approach of FRISCH and PARISI (1985), which was adopted by HALSEY *et al.* (1986).

As to the reader familiar with our earlier work, MANDELBROT (1988), he will find the present text to be much more detailed; for example, Sections 3.4 and 3.5 of the earlier work have been replaced by Sections 5 to 8.

Our goal is not to teach manipulations, but to present the reader who is new to multifractals with what we believe is the most understandable and simplest form of their theory, and to provide the skilled reader with the surprisingly simple but subtle explanation of the formal manipulations with which he is familiar.

2.1. Summary for the Reader New to Multifractals

The work is best summarized as follows:

The notion of self-similarity extends readily from fractal sets to measures. As we shall see in Section 3 a measure is simply a way of specifying a method of spreading mass, or probability, or other "stuff," over a supporting set. The distribution may be spread on a Euclidean "support", like an interval or a square, or it may be restricted to a fractal support, like a Cantor set. The function is the same, except for detail.

In order to describe quantitatively a self-similar method of spreading stuff around, one standard first characteristic is a sequence of moments, or an exponent in the cumulant generating function, which is denoted in this context by $\tau(q)$. The earliest graph of $-\tau(q)$ is given as Fig. 2 in MANDELBROT (1974).

Another characteristic, equivalent to the first one, is a limit distribution function. This limit is akin to an ordinary probability distribution function. Also, the renormalization must follow a very unusual and ill-known path: one needs a *multiplicative* renormalization that is unfamiliar. The *additive* renormalization, which is \bar{I} and familiar to every student of probability, and which leads to the central limit theorem, is worthless. Because of this renormalization, the limit distribution function is best considered after logarithmic transformations of both axes, that is, as plotted with doubly logarithmic coordinates. We shall denote the resulting function by $\rho(\alpha)$. An alternative quantity $f(\alpha)$, given by $f(\alpha) = \rho(\alpha) + 1$ for measures on the line, has become entrenched in the literature to specify the limit distribution.

One can obtain $f(\alpha)$ from $\tau(q)$, and $\tau(q)$ from $f(\alpha)$, by a direct or inverse Legendre transform. This property follows immediately from the Lagrange multipliers approach to the Gibbs distribution in thermodynamics, an approach long familiar to every physicist. Later, we give a full mathematical justification of the formalism, valid in a broader context in which $f(\alpha)$ can very well be negative. This justification is provided by reference to existing (but little-known) limit theorems of probability due to Harald Cramér, and concerned with "large deviations". The Legendre transform expresses the intuitive fact that $f(\alpha)$ can be obtained as the cap-convex envelope of an easily drawn family of straight lines.

On a plot of $f(\alpha)$, the quantity— $\tau(q)$ is the ordinate of the intercept of that tangent to $f(\alpha)$, whose slope is q , with the vertical axis. And the intercept of the same tangent with the main bisector of the coordinate axes is the quantity $D_q = \tau(q)/(q - 1)$, which is interpreted later as a critical dimension.

MANDELBROT (1974) has introduced two distinct kinds of random self-similar multifractals, respectively called *conservative* (or *microcanonical*) (Section 7) and *canonical* (Section 8). The latter are the more sophisticated, and needed for the study of low-dimensional cuts of multifractals embedded in a high-dimensional space (see Section 1.4). In particular, the canonical 1974 multifractals are necessary to understand the very peculiar standing of the lognormal distribution, which is widely mentioned when applying multifractals to turbulence. Observe that FRISCH and PARISI (1985) explicitly note that, while their approach had originally stemmed from MANDELBROT (1974), it was of lesser generality, because it did not accommodate the lognormal.

2.2. Summary for the Reader Familiar with the Multifractal Formalism, as Originated by Frisch and Parisi

To start with a question: Since “multifractal” is not a physical notion, but a probabilistic (and measure-theoretical) concept and tool, where is the probability theory in the many recent papers devoted to this topic? The answer is simple and is given in this work. It is best summarized as follows:

The function $\tau(q)$ with which you are familiar is a standard probabilistic tool to represent measures, called cumulant generating function.

The quantity α with which you are familiar is a standard notion, called Hölder exponent. There is no gain in calling it “strength of singularity.” HENTSCHEL and PROCACCIA (1983) call it dimension, but experience shows that many confusions are avoided if this term (in all its multiple and still multiplying forms!) is reserved to *sets* and never used to apply to *measures*.

There is a wide belief among those familiar with the multifractal formalism that the function $f(\alpha)$ with which you are familiar is a *new* concept that is *defined* via a Legendre transform from the function $\tau(q)$, therefore can be labeled at will (for example as the “spectrum of singularity”). In reality, if the measure is supported by a Euclidean interval of dimension 1, $f(\alpha)$ is simply equal to $\rho(\alpha) + 1$, where $\rho(\alpha)$ is a limit probability distribution function plotted on doubly logarithmic coordinates. But the set that supports the measure can be a fractal. This leads to a slightly more complex rule described in Section 6.6.

There are several methods for estimating $\rho(\alpha)$ and $f(\alpha)$. The Legendre transform starting with $\tau(q)$ yields one such method. But the estimation of the $\tau(q)$ involves many diverse smoothing operations. They are not innocuous, yet they have not been subjected to the careful analysis they deserve and demand.

The quantities $D_q = \tau(q)/(q - 1)$ are “critical dimensions” for $q > 1$. They were introduced in MANDELBROT (1974) through a theorem asserting that along a cut of dimension D through a measure, one has $\langle \mu^q(dx) \rangle = \infty$ when $q > q_{\text{crit}}(D)$. The function inverse of the function $q_{\text{crit}}(D)$ is D_q .

The attractiveness of multifractals may to some extent be due to an aura of mystery. All mystery is eliminated, however, when one understands the nature of the formal manipulations. But does it *really* matter that $\tau(q)$, α and $f(\alpha)$ are properly labeled? To respond, let us note that our approach obtains the Legendre transform between $\tau(q)$ and $f(\alpha)$ via Lagrange multipliers, as is usual in first courses of thermodynamics. In thermodynamics, there are several standard ways of making the use of these multipliers rigorous after the fact. One way is to follow the Darwin-Fowler method of steepest descents. This used to be a standard topic in textbooks, but is unfortunately no longer familiar to every physicist. Thus, in effect, the Frisch-Parisi method skips Lagrange multipliers, and proceeds to the Darwin-Fowler method immediately. No teacher of thermodynamics would proceed in this way in a first course! Nor should a teacher of multifractals.

The preceding remarks help explain that, in our experience, the Frisch-Parisi approach to multifractals has generated substantial confusion. No one has an intuition of what a “spectrum of singularities” can or cannot be like. Proper foundations flush out confusion. As a foremost example, the Frisch-Parisi interpretation of $f(\alpha)$ as a fractal dimension has led to paradoxes linked to negative $f(\alpha)$. These paradoxes had begun as surprising anecdotal evidence, and have of course ended by being explained away by suitable special developments of the theory. They do not even arise in our properly probabilistic approach.

The formalism originated by Frisch and Parisi has now given rise to a very extensive literature. But this is not the proper place to survey it, even in part. Our goal is to set out an alternative approach on its own terms, and the relations between the two approaches have better be developed elsewhere.

3. Spatial Variability Beyond Fractal Homogeneity

3.1. The Essential Distinction Between a Set and a Measure, and the Notion of “Singular Measure”

To understand multifractals beyond the crude formalism, a full appreciation of the distinction between the concepts of *set* and of *measure* is essential.

Fortunately, a knowledge of either set theory or measure theory is completely unnecessary. More precisely, the complications of measure theory enter into the theory at a stage beyond the topics discussed in this paper, for example when one wishes to follow the detailed proofs of some of the theorems to which we shall refer.

Basic ideas are always explained best when stated in the simplest context, which in the case of this paper is usually the interval $[0, 1]$, the square $[0, 1]^2$ or the cube $[0, 1]^3$. In a Euclidean space of dimension E , the interval, square or cube are examples of basic sets of the form $[0, 1]^E$.

The notion of a set. In order to define a set S in $[0, 1]^E$, one needs a rule that says for each point $P \in [0, 1]^E$ whether it satisfies $P \in S$ or $P \notin S$. Such a rule is provided by the “indicator function” $I(P)$ defined as $I(P) = 1$ if $P \in S$ and $I(P) = 0$ if $P \notin S$.

Uniform or homogeneous measure. It is well-known that one can associate with the whole $[0, 1]^E$ a uniform measure. The set and its uniform measure are mathematically equivalent. Next, consider a fractal set constructed recursively, e.g., a Sierpinski gasket, which is described in Figure 1. The most natural measure, again called “uniform,” gives equal weight to each of its thirds. Again, the set and its uniform measure are equivalent. However, the uniform measure is very special, and in many cases very unrealistic, as has already been stressed in Section 1.

The notion of the limit of a set. Multifractals play two main roles, each of which can begin with a complicated set of points. The first role is exemplified by the distribution of copper on earth. Since the mining engineer can see no harm in assuming

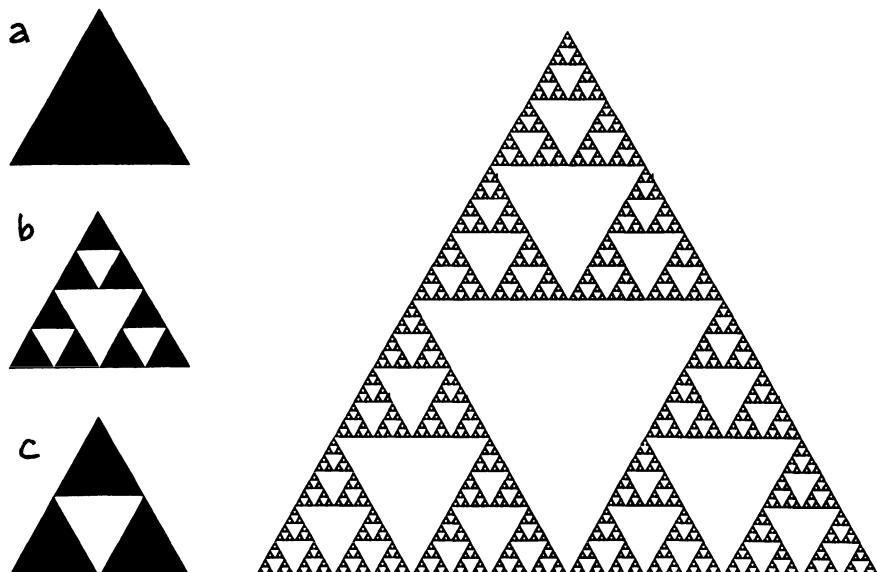


Figure 1

The construction of Sierpinski's gasket proceeds as follows. The initiator, shown in Figure 1a is a triangle. In the stage between Figures 1a and 1b, this triangle is divided into four and the middle fourth is erased, yielding 1b. Between 1b and 1c, each of the three remaining fourths is, divided into four and the middle fourth of each remaining triangle is erased. The same process is repeated *ad infinitum*.

that each copper atom reduces to a point, all the copper atoms on earth form a set of points. The second role is exemplified by the successive “sample values” of a random process examined at discrete instants of time, or the successive “orbit points” of a deterministic process. Again, one deals with a sequence of points or ordered set.

In all the cases of interest, however, these sets are very large, and to specify them completely would be an unmanageable task. It would also be pointless. This leads to a natural impulse, especially in the second example: let time run forever and take the limit of the set of successive positions of our process. Question: is the limit a set? In order to answer usefully, consider a collection of points that have been chosen at random on $[0, 1]$.

First suppose that their probability distribution is continuous, with a density $\mu'(x)$ such that $\mu'(x) > 0$ for all x . If so, the limit of an increasingly large set of sample points is indeed a set, namely the whole interval $[0, 1]$. The fact that this limit is independent of the function $\mu'(x)$ means of course that the limit contains less information than the finite set that tends to this limit. The answer is obvious: the least that we must do in either case is to specify, not only a limit set, but also a density on this set.

A second familiar case is when our random point can take only certain admissible values x_g , each of which has a known probability μ_g . In this case, the limit is a (denumerable) set $\{x_g\}$, but we must also specify the μ_g which are said to be “supported” by this set.

How to describe a random process that generates x with a density $\mu'(x)$? A first characteristic is $\mu'(x)$ itself, or its integral $\mu(x) = \int_0^x \mu'(s) ds$, which is a cumulative probability distribution. An alternative though less complete characteristic is the function

$$x(\mu) = \text{total length of the intervals of } x \text{ within which } \mu'(x) < \mu.$$

The densities $\mu'(x)$ and the probabilities μ_g happen to define the two classical examples of “measures.” Multifractals are also measures, but they belong to a third class, which has no density and is not discrete, and is called “singular.” The best is to first discuss them in an example.

3.2. An Old but Good Illustration of the Concept of Singular Measure

Since “to see is to believe”, Figure 2 reproduces the earliest completed illustration of a random multifractal, and appeared first in our earliest full paper on this topic (MANDELBROT, 1972). Many geologists and geophysicists may find it difficult *not* to find never-ending new interpretations for it.

The horizontal axis shows the “time” t , as “discretized” into small boxes of uniform width Δt , and the vertical axis shows the sequence of the “measures” of

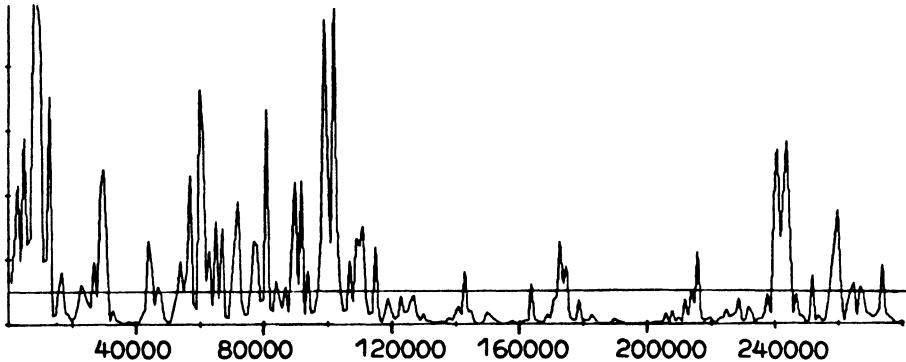


Figure 2

The earliest simulation of a sample from a multifractal measure, namely the limit lognormal measure of MANDELBROT (1972).

these boxes. Let us say again that, if the total integral measure over the time span $[0, 1]$ is set to 1, one can think of *the measure in a box as the probability of hitting this box*. But the measure in a box is, more generally, the amount of any kind of “stuff” contained within it.

The measure of the interval $[0, t]$, call it $\mu([0, t])$, is of course an increasing function of t , and what is plotted here is the sequence of its finite increments $\mu([0, t + dt]) - \mu([0, t])$. The increments’ values are joined to form a curve, but this curve is *not* what it seems to be. It is not the graph of an intrinsically interesting function, and it cannot readily be transformed into one.

To provide contrast and familiar background, let us draw the analogous diagram for a measure having a continuous density $\mu'(t)$. In that case, $\mu(\Delta t)/\Delta t$ is an approximation to the graph of the function representing $\mu'(t)$, and a first characterization of our measure would be provided by the distribution of this approximate density $\mu(\Delta t)/\Delta t$ in time. As $\Delta t \rightarrow 0$, the approximate density would tend to the true $\mu'(t)$.

In the present instance, however, the situation is extremely different. By design, the measure is approximately self-similar. This notion will be discussed in Section 3.4, but it may already be said that when the measure in Δt is examined on much finer scales $\Delta_2 t \ll \Delta t$, it proves to be about as irregular as the whole Figure 2. For example, if Δt is halved, the sharing of the $\mu(\Delta t)$ between the two halves is usually very unequal.

It follows that this $\mu([\Delta t])$ grossly fails to have a local density. Nor is it discrete. Therefore, taking the limit $\Delta t \rightarrow \infty$ transforms the wiggly curve on Figure 2 into a curve that is even more wiggly, hence does not become an increasingly close approximation to a function.

Suppose that the measure $\mu([0, t])$, considered as a function of t , is continuous but is not differentiable. Then the measure $\mu(dt)$ has no local density, and is called “singular.” (This is a technical term, due to Lebesgue.) For a singular measure,

the notion of asymptotic "distribution" for the values of $\mu([\Delta t])$ has no meaning. Fortunately, a suitable substitute is available.

3.3. The Limit Probability Distribution $\rho(\alpha)$, and the Corresponding Function $f(\alpha)$

Begin with the equivalent of Figure 2 for a sequence $\Delta_k t$ of decreasing values of Δt . Then, for each $\Delta_k t$, replace $\mu([\Delta_k t])$ by the quantity $\alpha_k = \log \mu([\Delta_k t]) / \log \Delta_k t$, which is called Hölder exponent. Next, form the probability density of α_k , and replace it by the quantity $\rho_k(\alpha) = -\log(\text{probability density of } \alpha_k) / \log \Delta_k t$. Each of these two steps consists of first taking logarithms, and then renormalizing them by dividing by the factor $-\log \Delta_k t$. This renormalizing factor is not familiar, and may seem peculiar and artificial. However, it will be justified in Section 7.2. There, we examine successive examples for which the measure is known to be multifractal, and we find that $\rho_k(\alpha)$ converges as $\Delta t \rightarrow 0$ to a limit that does not reduce to either 0 or ∞ . Any other renormalizing factor, to the contrary, would yield a degenerate limit.

The limit of $\rho_k(\alpha)$ will be denoted by $\rho(\alpha)$. It will be seen that $\rho(\alpha)$ is negative for all α , except for the value where $\rho(\alpha)$ reaches its maximum, which is the expectation $\langle \alpha \rangle$. But one must realize that the convergence to the limit may be slow. The property that $\rho_k(\alpha) \rightarrow \rho(\alpha)$ can be turned around, and used to *define* the notion of multifractal.

As has been first pointed out by FRISCH and PARISI (1985), it is also worth considering the quantity $f(\alpha) = \rho(\alpha) + 1$. In effect, Frisch and Parisi have discussed $f(\alpha)$ in the special case $\rho(\alpha) \geq -1$, so that $f(\alpha) > 0$, for all α . They have pointed out that, in this special case, one can interpret $f(\alpha)$ as being the fractal dimension of a suitable set. The replacement of $\rho(\alpha)$ by $f(\alpha)$ has virtues in some cases, but our feeling is that, fundamentally, it hides the nature of the multifractals.

Between my early papers and the preprint of FRISCH and PARISI (1985) (which was distributed in 1983), multifractals continued to develop only in the sense that the mathematical background was cleaned up and extended (e.g., in KAHANE and PEYRIÈRE, 1976). But they did not receive new applications (nor were they mentioned in *Physical Review Letters*). Their spread is a recent phenomenon, and most readers who have heard of them are likely to know presentations that follow the approach common to Frisch and Parisi and to HALSEY *et al.* (1986). Unfortunately, the algebra of these presentations is needlessly complicated, artificial, and of limited applicability and the terminology of HALSEY *et al.* hides the extremely simple nature of the underlying structure. We shall, therefore, adopt the notation of HALSEY *et al.*, but follow our original approach in the form into which it has lately developed.

3.4. Self-similar Measures and Beyond

We need a few more formal definitions concerning positive measures. The

measure of a set S will be written as $\mu(S)$. Again, those not familiar with measure can think of μ as being the probability of hitting the set S . The multifractal measures obtained as a result of multiplicative cascades are closest in spirit, to the exact self-similar fractal sets.

Recall that a fractal set is exactly self-similar, if it can be decomposed into parts, each of which is obtained from the whole by a transformation called isotropic contraction, or contracting similitude, to be denoted by \mathcal{K} . A self-similar set is fully determined by a collection of contractions. For example, each third of a basic fractal called a Sierpinski gasket is obtained from the whole by a contracting similitude of ratio $r = 1/2$. Starting with any triangle in a “prefractal collection of triangles”, the interpolation of the shape itself continues without regard to the “past” construction steps.

Now suppose that a (positive) measure $\mu(P)$ is defined for each third of the gasket, then is interpolated for each third of a third etc. More generally, when the part P' of a fractal is obtained from the part P by the contracting transformation \mathcal{K} , so that $P' = \mathcal{K}(P)$, the conditional measure of P' in P is defined precisely like a conditional probability. It is the ratio $\mu(P')/\mu(P)$ of the measure $\mu(P')$ to the conditioning measure $\mu(P)$.

Now, the idea of self-similarity for a measure expresses that the process of successive interpolations continues without regard to the “past” steps. The rough idea is that, as the parts contract, the measures they carry contract proportionately. To express this idea formally, take a second contracting transformation \mathcal{L} , and compare $\mu(P')/\mu(P) = \mu\{\mathcal{K}(P)\}/\mu(P)$ with $\mu\{\mathcal{L}(P')\}/\mu\{\mathcal{L}(P)\}$. If these conditional measures are identical for every choice of \mathcal{K} and \mathcal{L} , the measure μ will be called a *strictly self-similar multifractal*.

A random measure is called *statistically self-similar* if, given one or a finite collection of nonoverlapping parts $P_\gamma = \mathcal{K}_\gamma(P)$, the distribution or the joint distribution of the quantities $\mu(P_\gamma)/\mu(P)$ depends only on the contractions \mathcal{K}_γ .

Side remark. In a more general mathematical fractal set, the parts are obtained from the whole by transformations that are *nonlinear*. Examples where the contractions are in some sense *almost linear* include the Julia sets of polynomial maps. The corresponding multifractals include the harmonic measures on these sets. Other examples of multifractal measures concern the limit sets of groups based upon inversions in circles (*FGN*, Chapters 18 and 20). The limit set itself may be a straight line, as in the example examined by GUTZWILLER and MANDELBROT (1988). Finally, the “fat fractals” (a new term for the fractals investigated in *FGN*, Chapter 15) and the *Mandelbrot* set involve essentially nonlinear transformations. As yet, there is no general agreement about which transformations are acceptable in defining the terms “fractal” and “multifractal.”

3.5. Five Basic Concepts Concerning Multifractals

Aside from the distinction between set and measure, multifractals involve five basic notions, each of which will be introduced in the simplest possible context. A)

The Hölder α is introduced in Section 4, in the context of the homogeneous Cantor measure. B) The functions $\rho(\alpha)$ and $f(\alpha)$ are introduced in Section 3, in the context of the binomial measures. C) The Legendre transform “multifractal formalism” is introduced in Section 6, in the context of the multinomial measures. D) The notion of latent α 's is introduced in Section 7, together with a first form of randomness, in the context of “conservative 1974 multifractal measures.” E) The notions of critical exponent q_{crit} , and of critical dimension function D_q (D_q is the inverse function of q_{crit}), and the notion of virtual α are introduced in Section 8, together with a second form of randomness, in the context of “canonical 1974 multifractal measures.”

4. The Basic “Unifractal” is the Cantor Measure. It is Nonrandom and Self-similar.

Introduction of the Hölder α

4.1. Definitions

It is probably not necessary here, to repeat the definitions of the triadic Cantor dust \mathcal{C} , and of its devil staircase function $C(x)$ but it takes little space. To define \mathcal{C} , one starts with the interval $[0, 1]$ and one cuts out the middle third open interval $]1/3, 2/3[$. A closed interval is denoted by $[]$, and an open one by $] [$. Then one cuts out the middle third open interval of $[0, 1/3]$ and $[2/3, 1]$, etc. The staircase function $C(x)$ is a special case of the function $\mu([0, t])$ of Section 2.3. The left and right halves of \mathcal{C} are assigned the equal measures $1/2$, so that $C(1/3) = C(2/3) = 1/2$. Then the four quarters of \mathcal{C} are assigned the equal measures $1/4$, so that $C(1/9) = C(2/9) = 1/4$ and $C(7/9) = C(8/9) = 3/4$. And one continues to interpolate *ad infinitum*. Thus, the measure of every interval $[x', x'']$ in $[0, 1]$, is obtained from the limit function $C(x)$ as $\mu([x', x'']) = C(x'') - C(x')$.

This measure is called Cantor measure, or *fractally homogeneous measure* on the Cantor dust, for the following reason. If two pieces of Cantor dust contained in $[x'_1, x''_1]$ and $[x'_2, x''_2]$ can be superposed by translation, they carry identical measures: $\mu([x'_1, x''_1]) = \mu([x'_2, x''_2])$.

4.2. The Devil Staircase $C(x)$ is a Diagonally Self-affine Fractal Curve

Indeed, it subdivides into the following three parts. The “middle part” is deduced from the whole by a degenerate (noninvertible) affinity of matrix $\begin{pmatrix} 0 & 0 \\ 0 & 3^{-1} \end{pmatrix}$ and fixed point $(2^{-1}, 2^{-1})$. The two “side parts” are deduced from the whole by affinities that share the same matrix $\begin{pmatrix} 2^{-1} & 0 \\ 0 & 3^{-1} \end{pmatrix}$ and have fixed points $(0, 0)$ and $(1, 1)$, respectively. In the present context, degenerate matrices “do not count.” Therefore, the unique simplicity of the homogeneous multifractal on \mathcal{C} is due to the fact that the matrix has only one possible nondegenerate form.

More generally, consider a fractal obtained by a recursion such that there are N contractions, all having ratios equal to the same r , and that each recursive stage divides the measure equally between the N parts. Then the measure is self-affine, the matrix being either degenerate of the form $\begin{pmatrix} 0 & 0 \\ 0 & r \end{pmatrix}$ or nondegenerate of the form $\begin{pmatrix} N^{-1} & 0 \\ 0 & r \end{pmatrix} = \begin{pmatrix} r^D & 0 \\ 0 & r \end{pmatrix}$.

4.3. Application of the Notion of Hölder Exponent to Intervals Instead of Points. α and $f(\alpha)$ for the Unifractal Measures

Adapting a classical mathematical notion due to Hölder to apply to the dyadic cell $[dt]$, we write

$$\alpha(t) = \frac{\log[\mu(dt)]}{\log(dt)}.$$

For the Cantor dust of dimension D , the interval $[0, 1]$ contains $N^k = 2^k$ nonempty cells of length $dt = b^{-k}$, each containing the same measure $N^{-k} = b^{-kD}$. For each of these cells, one has $\mu(dx) = (dx)^D$. Hence $\alpha = D$.

The remaining cells are empty, and they form a set of fractal dimension equal to 1. In these cells, one can say that $\alpha = \infty$.

Infinitesimals of the form $(dy) = (dx)^\alpha$ with $\alpha \neq 1$ are called *nonstandard*, and Abraham Robinson has constructed around them a field called *nonstandard analysis*. It is not needed in this discussion.

The function $f(\alpha)$. For future reference, it is useful to introduce a function $f(\alpha)$ that is defined for two points only:

$$\text{for } \alpha = D, f(\alpha) = D, \text{ and for } \alpha = \infty, f(\alpha) = 1.$$

4.4. Finite Sample Prefactors

In general, the Cantor measure on a Cantor set of base b is studied using cells whose base b' is such that $b' \neq b$ and $b' \neq b^h$ for all integer h . In that case, the number of nonempty cells continues to be “of the order” of N^k , but it is no longer equal to N^k . One can write it in the form $F_1 N^k$, where the prefactor F_1 depends on k and remains positive and finite. The measure in a nonempty cell, similarly, can be written as $F_2 N^{-k}$, where the prefactor F_2 depends on k , varies between cells, is finite but can be arbitrarily small. Therefore, the sample values of α are of the form $D + \log F_2/k$. These values may overlap several of the bins in which the values of α are sorted for the purposes of statistics.

5. The Basic Multifractal is the Binomial Measure. It is Nonrandom and Self-similar. Introduction of the Function $\rho(\alpha)$ and $f(\alpha)$, and of the Multifractal Representation

The binomial measure on the interval $[0, 1]$ is the simplest example of a multifractal measure. By studying this example more carefully and in greater detail than is usual, this section hopes to make many of the important features of multifractals stand out clearly. In particular, the function $f(\alpha)$ is best understood when developed in a context in which no formalism is needed.

Under various names, the binomial and multinomial measures of Sections 5 and 6 have been long known in pure mathematics. But the first concrete application (to the distribution of turbulent diffusion, MANDELBROT, 1974), as well as various extensions in a “fractal” context have come with our work. *FGN*, calls them *Besicovitch measures*, to honor the principal early contributor of their study, but this terminology has never taken hold. It has now become obsolete, because it lends itself poorly to the necessity of distinguishing between the binomial and the multinomial cases.

5.1. Construction of the Binomial Measure, Derivation of the Equivalent Functions $f(\alpha)$ and $\rho(\alpha)$

The binomial measures are a family of multifractals that depend upon a single parameter m_0 , variously called a *multiplier* or a *mass*. One assumes that m_0 satisfies $0 < m_0 < 1$ and $m_0 \neq 1/2$, and one defines $m_1 = 1 - m_0$. Often it is nice to choose m_0 larger than m_1 , so that $1/2 < m_0 < 1$. (Note that, in terms of the de Wijs quote in Section 1.2, $m_0 = (1 + d)/2$.)

The mass in the “initiator” $[0, 1]$ is taken as unity, then the basic “generating step” is to spread mass over the halves of every dyadic interval, with the relative proportions m_0 and m_1 . The first stage yields the mass m_0 in $[0, 1/2]$ and the mass m_1 in $[1/2, 1]$, etc. After k stages, suppose that $t = 0, \eta_2\eta_2 \dots \eta_k$ is the development of t in the counting base $b = 2$ and let φ_0 and φ_1 denote the relative frequencies of 0’s and 1’s in the binary development of t . Then the binomial measure is defined as the measure that assigns to the dyadic interval $[dt] = [t, t + 2^{-k}]$ of length $dt = 2^{-k}$ the mass

$$\mu(dt) = m_0^{k\varphi_0} m_1^{k\varphi_1}.$$

The binomial measure is self-similar. By construction, $\mu(dt)$ satisfies the property of self-similarity described in Section 3.4.

The notion of “pre-multifractal”. One can define a sequence of measures $\mu_k(dt)$ such that $\mu_k(dt) = \mu(dt)$ if $[dt]$ is a dyadic cell of length 2^{-k} , and the mass is distributed uniformly over each cell. The measures $\mu_k(dt)$ can be called “pre-multifractal”, because they play the same role relative to the binomial measure, as the broken line “prefractal” approximations play relative to the fractal sets such as the Koch curve.

The function $\mu(t)$, an analog of the Cantor staircase function $C(t)$, is diagonally self-affine. As we know, this means that its variation within a cell of side b^{-k} is a reduced-size replica of its variation over the original cell of side 1. In every cell, the reduction is an affinity whose matrix is diagonal. The horizontal linear reduction is b^{-k} in every cell. The vertical reduction is the measure within the cell, and varies from cell to cell.

The Hölder α . In the present case,

$$\alpha = \alpha(\varphi_0, \varphi_1) = -\varphi_0 \log_2 m_0 - \varphi_1 \log_2 m_1,$$

and $0 < \alpha_{\min} = -\log_2 m_0 \leq \alpha \leq \alpha_{\max} = -\log_2 m_1 < \infty$.

The box dimension. The number of intervals leading to φ_0 and φ_1 is $N(k, \varphi_0, \varphi_1) = k!/(k\varphi_0)!(k\varphi_1)!$. One can use N to form the following expression:

$$\delta(k, \varphi_0, \varphi_1) = -\frac{\log N(k, \varphi_0, \varphi_1)}{\log(dt)} = -\frac{\log[k!/(k\varphi_0)!(k\varphi_1)!]}{\log(dt)}.$$

This quantity δ is of the form $-\log N/\log r$ that characterizes the similarity dimension of a set. Hence one can call δ a *box fractal dimension*. More precisely, since the boxes belong to a grid, it is a grid fractal dimension.

For large k , the replacement of the factorial by the leading term in the Stirling approximation shows that

$$\lim_{k \rightarrow \infty} \delta(k, \varphi_0, \varphi_1) = \delta(\varphi_0, \varphi_1),$$

with

$$\delta(\varphi_0, \varphi_1) = -\varphi_0 \log_2 \varphi_0 - \varphi_1 \log_2 \varphi_1.$$

Derivation of a function $f(\alpha)$. Eliminating φ_0 and φ_1 between α and δ , we obtain a function written in parametric form. It will be denoted as $f(\alpha)$, instead of $\delta(\alpha)$, the reason being explained in Section 6.2.

Note that $0 \leq f(\alpha) \leq \min\{\alpha, 1\}$, with equality being achieved in four points: 1) $f(\alpha) = 0$ for $\alpha = \alpha_{\min} = -\log m_0$, 2) $f(\alpha) = 0$ for $\alpha = \alpha_{\max} = -\log m_1$, 3) $f(\alpha) = 1$ for $\alpha = (1/2)(\alpha_{\min} + \alpha_{\max})$ and 4) $f(\alpha) = \alpha$ for $\alpha = -m_0 \log m_0 - m_1 \log m_1$.

A first reading of this paper may proceed directly to Section 6.

5.2. Entropy, Information, Entropy Dimension and Information Dimension

The notion of entropy, as every physicist knows, is extremely important but far from easy to handle. One of the basis forms of entropy is written as $-\sum p_\beta \log_b p_\beta$, where the p_β are probabilities, and entropy written in this form has since been reinterpreted by Claude Shannon as an “information”. Thus Section 5.1 proves the binomial case of the important result that $f(\alpha)$ is often formally an entropy-information. This role has led many authors to refer to it as “information dimension.”

Though I do not like to argue about words, this term has always seemed inappropriate. Even the simplest similarity dimension $\log_b N$ is already formally identical to one of the basic expressions for entropy-information.

5.3. Restatement of Section 5.1 in Terms of Rescaled Doubly Logarithmic Plots of Probability Densities

Having obtained $f(\alpha)$, we shall spend several sections in a discussion of what it really does and does not mean. First, let us reinterpret the combinatorics in Section 5.1 in probabilistic terms. This will be easy, and very important because it will show that $f(\alpha)$ is *not* a new notion requiring an entirely new intuition, but a new form of a familiar notion, only requiring the further development of old intuitions.

Let a dyadic cell of length 2^{-k} be picked at random among the 2^k such cells. For the probability of hitting a cell that corresponds to a prescribed α or the equivalent φ_0 and φ_1 , we shall use the notation $\Pr\{A = \alpha\}$. Here, A is to be read as “capital α ”, and we follow the probabilists’ custom of denoting a random variable by a capital letter, and a possible value of this variable by the corresponding small letters. We have of course $\Pr\{A = \alpha\} = 2^{-k}N(k, \varphi_0, \varphi_1)$. As $k \rightarrow \infty$, $\Pr\{A = \alpha\} \rightarrow 0$ for all $\alpha \neq \langle \alpha \rangle$. The next step is to raise the basic question of probability theory. It is possible to “renormalize” $\Pr\{A = \alpha\}$ in such a way that it has for $k \rightarrow \infty$ a limit that is neither 0 nor ∞ . Clearly, this is possible and even easy: it suffices to define

$$\rho(k, \varphi_0, \varphi_1) = -\frac{\log \Pr\{A = \alpha\}}{\log dt} = -\frac{\log 2^{-k}N(k, \varphi_0, \varphi_1)}{\log dt}.$$

Obviously, $\rho(k, \varphi_0, \varphi_1) = \delta(k, \varphi_0, \varphi_1) - 1$, with $\delta(k, \alpha)$ as in Section 5.1. Hence,

$$\lim_{k \rightarrow \infty} \rho(k, \varphi_0, \varphi_1) = \rho(\alpha) = f(\alpha) - 1.$$

5.4. Restatement of $f(\alpha)$ and $\rho(\alpha)$ in Terms of “Tail” Dimensions and Probabilities

What we have done is best stated in terms of “tail” quantities. After $k < \infty$ has been specified, the quantities φ_0 and φ_1 are both multiples of $1/k$. The “tail” quantities are defined by

$$N^+(k, \varphi_0, \varphi_1) = \sum_{\psi_0 \geq \varphi_0} N(k, \psi_0, -\psi_0)$$

$$N^-(k, \varphi_0, \varphi_1) = \sum_{\psi_0 \leq \varphi_0} N(k, \psi_0, -\psi_0).$$

By analogy with $-\log N/\log dt$, one can use N^+ and N^- to define the tail box dimensions $\delta^+(k, \varphi_0, \varphi_1)$ and $\delta^-(k, \varphi_0, \varphi_1)$, and then to seek and define their limits $\delta^+(\varphi_0, \varphi_1)$ and $\delta^-(\varphi_0, \varphi_1)$. A bit of combinatorics and Stirling approximations easily yields the following facts:

For $\varphi_0 > 1/2$, $\delta^+(\varphi_0, \varphi_1) = \delta(\varphi_0, \varphi_1)$ and $\delta^-(\varphi_0, \varphi_1) = 1$.

For $\varphi_0 = 1/2$, $\delta^+(\varphi_0, \varphi_1) = \delta^-(\varphi_0, \varphi_1) = 1$.

For $\varphi_0 < 1/2$, $\delta^+(\varphi_0, \varphi_1) = 1$ and $\delta^-(\varphi_0, \varphi_1) = \delta(\varphi_0, \varphi_1)$.

Next, the function δ^* is defined by stringing together the information-carrying portions of δ^+ and δ^- . Thus,

$$\delta^*(k, \varphi_0, \varphi_1) = \min\{\delta^+(k, \varphi_0, \varphi_1), \delta^-(k, \varphi_0, \varphi_1)\}.$$

The next obvious step is to define the positive (resp. negative) “tail probabilities” of α as being $\Pr\{A \geq \alpha\}$ (resp., $\Pr\{A \leq \alpha\}$), and to define “the” tail probability as being

$$\Pr\{A \geq \alpha\} \text{ when } \alpha > \langle \alpha \rangle$$

$$\Pr\{A \leq \alpha\} \text{ when } \alpha < \langle \alpha \rangle.$$

By a further analogy, we define the quantities $\rho^+(k, \varphi_0, \varphi_1) = \delta^+(k, \varphi_0, \varphi_1) - 1$, and similarly $\rho^+(k, \varphi_0, \varphi_1)$, $\delta^+(k, \varphi_0, \varphi_1)$ and $\rho^-(\varphi_0, \varphi_1)$. It is easy to see that, as $k \rightarrow \infty$,

For $\varphi_0 > 1/2$, $\rho^+(\varphi_0, \varphi_1) = \rho(\varphi_0, \varphi_1)$ and $\rho^-(\varphi_0, \varphi_1) = 0$.

For $\varphi_0 = 1/2$, $\rho^+(\varphi_0, \varphi_1) = \rho^-(\varphi_0, \varphi_1) = \rho(\varphi_0, \varphi_1) = 0$.

For $\varphi_0 < 1/2$, $\rho^+(\varphi_0, \varphi_1) = 0$ and $\rho^-(\varphi_0, \varphi_1) = \rho(\varphi_0, \varphi_1)$.

The quantities $\rho(k, \varphi_0, \varphi_1)$, $\rho^+(k, \varphi_0, \varphi_1)$, $\rho^-(k, \varphi_0, \varphi_1)$ and $\rho(\varphi_0, \varphi_1)$ will now be interpreted in probabilistic terms, and show to be nothing but probability distributions in doubly logarithmic coordinates.

5.5. The Notion of Multiplicative Measure and a Very Surprising Limit Theorem, a Special Case of a Theorem of Cramer

It is useful to restate the fact that the measure $\mu(dt)$ in a cell of length 2^{-k} is the product of k multipliers equal to either m_0 or m_1 . We consider the interval $[0, 1]$ globally and construct a cascade as follows. It begins with a uniform density $\mu'_0(t) = 1$. After the first stage, let the density become $\mu'_1(t) = m(2t)$, where $m(t)$ is the following periodic function of period 2:

$$m(t) = m_0 \text{ if } 0 < t < 1 \text{ and } m(t) = m_1 = 1 - m_0 \text{ if } 1 < t < 2.$$

After stage 2, $\mu'_2(t) = m(2t)m(2^2t)$.

After stage k , $\mu'_k(t) = m(2t)m(2^2t) \dots (2^kt)$.

(The value of $\mu'_k(t)$ at the dyadic points does not matter.) This $\mu'_k(t)$ is the density of the premultifractal $m_k(dt)$.

Hence, the binomial measure is called a *multiplicative* multifractal. It is the simplest example of multiplicative fractal constructed in a grid. This property makes it exactly renormalizable.

The process of multiplication of random quantities is unfamiliar, but it is easy to transform it into a process of addition, by taking logarithms. Thus,

$$\alpha(k) = -(1/k)[\log_2 m(\eta_1) + \log_2 m(\eta_2) + \cdots + \log_2 m(\eta_k)],$$

where $m(\eta_j) = m_0$ if $\eta_j = 0$ and $m(\eta_j) = m_j$ if $\eta_j = 1$.

Now, to select at random the cell $[dt]$, means to select at random the k first digits of t . Hence we see that $\alpha(k)$ is the average of k values of a random variable equal to

$-\log_2 m_0$ with the probability 1/2, and to

$-\log_2 m_0$ with the probability 1/2.

Comments concerning lognormality. In terms of the representation of $\alpha(k)$ as the average of k random variables. The results of Section 5.1 express a limit theorem of probability theory. The form of this theorem ought to surprise those readers who believe that the classical central limit theorem shows that $\log m(\eta_1) + \log m(\eta_2) + \cdots$ is asymptotically Gaussian in "every case." This would imply that the measure $\Pi M(\eta_k)$ is asymptotically lognormal. These familiar assertions are perfectly true, and they result from the application of a different renormalization. From the central limit theorem one does obtain information about the multiplicative multifractals, but this information only implies that $\rho(\alpha)$ and $f(\alpha)$ are parabolic in the central bell near their maximum. Concerning their form away from the maximum, the central limit theorem does not pretend to say anything. This is a very important issue, to which we shall return in a more general context in Sections 7 and 8.

5.6. The Two Forms of the Multifractal Representation

The term "multifractal," which is due to FRISCH and PARISI (1985), is attractive and has contributed to drawing attention to the multifractals. It is motivated by the following representation that, in turn, is often inverted and made into a definition of what a multifractal is.

"A multifractal measure can be represented as the union of a continuous infinity of addends. Each addend is an infinitesimal 'unifractal measure.' It is characterized by a single value of α , and it is supported by a fractal set having the fractal dimension $f(\alpha)$. The sets corresponding to different α 's are intertwined."

The situation after a finite number $k < \infty$ of cascade stages. After $k < \infty$ stages "fractal dimension" is defined for the premultifractal multiplicative measure as in Section 4.2. It is an informal box dimension that involves a single reading of $-\log N/\log r$, and does not involve a limit. The multifractal representation follows; it is mostly useful when k is large enough to take the bite off the critique to be described in Section 5.7.

The limit for $k \rightarrow \infty$. Role of the Hausdorff-Besicovitch dimension. For $k \rightarrow \infty$, the multifractal representation involves far subtler issues, but in the present case it

was justified in advance by theorems in EGGLESTON (1949) and VOLKMAN (1958) that were available “off-the-shelf,” a pleasant surprise. These theorems involve the Hausdorff-Besicovitch dimension, D_{HB} . The quantity D_{HB} had played an important early role in fractal geometry, but this role has been decreasing sharply in recent years.

The theorem in EGGLESTON (1949) covers the t 's that are “normal” for the probabilities φ_0 and φ_1 . These are the real values of t whose binary development $t = 0.\eta_1\eta_2\dots$ is such that the frequencies of 0 and 1 in the sequence η_1, η_2, \dots converge to well-defined limits φ_0 and φ_1 . Concerning the “Eggleston set” formed by these t 's, the Eggleston theorem tells us that its Hausdorff-Besicovitch dimension is the familiar $\delta = \varphi_0 \log_2 \varphi_0 - \varphi_1 \log_2 \varphi_1$. Furthermore, the Hölder exponent α can simply be defined by taking a sequence of increasingly small nested intervals dt that include t , for example the sequence of dyadic intervals, and one can write

$$\alpha(t) = \lim_{dt \rightarrow 0} \frac{\log \mu([dt])}{\log dt}.$$

When t is itself dyadic, the sequence of η_j ends by an infinity of either 0's or 1's. Hence, one must define left and right α exponents, using $\log \mu([t - dt, t])$ and $\log \mu([t, t + dt])$. But this hardly matters, because dyadic t 's are of fractal dimension equal to $f(\alpha) = 0$. As a matter of fact, the binomial measure's α_{\min} and α_{\max} are attained to the left and the right sides of the same dyadic values of t (except that the endpoint $t = 0$ yields only α_{\min} and the endpoint $t = 1$ yields only α_{\max}).

Now, let us go beyond the Eggleston theorem, and consider the t 's for which φ_0 and φ_1 fail to be defined. To account for these t 's, the definition of the Hölder exponent takes its original mathematical form: one must replace $\lim_{dt \rightarrow 0}$ by

$$\alpha^{\sup}(t) = \lim \sup_{dt \rightarrow 0} \frac{\log \mu([dt])}{\log dt} = \lim \sup_{dt \rightarrow 0} [-\varphi_0(k) \log_2 m_0 - \varphi_1(k) \log_2 m_1],$$

where $\varphi_0(k)$ and $\varphi_1(k)$ are the frequencies of 0 and of 1 in the k first digits of t . To continue, we must replace $\varphi_1 = 1 - \varphi_0$ by two independent quantities φ_0^{\sup} and φ_0^{\inf} , with

$$\varphi_0^{\sup} = \lim \sup_{k \rightarrow \infty} \varphi_0(k) = 1 - \varphi_1^{\inf} = 1 - \lim \inf_{k \rightarrow \infty} \varphi_1(k)$$

$$\varphi_0^{\inf} = \lim \inf_{k \rightarrow \infty} \varphi_0(k) = 1 - \varphi_1^{\sup} = 1 - \lim \sup_{k \rightarrow \infty} \varphi_1(k).$$

Volkmann (1958) proves that $-\Phi_0 \log_2 \Phi_0 - \Phi_1 \log_2 \Phi_1$ is the Hausdorff-Besicovitch dimension of the set of t where $\varphi_0^{\sup} = \Phi_0$, and also of the set of t where $\varphi_0^{\inf} = \Phi_0$. It follows that it is also the Hausdorff-Besicovitch dimension of the set of t where $\alpha^{\sup}(t) = \alpha = -\Phi_0 \log_2 m_0 - \Phi_1 \log_2 m_1$. Thus, Volkmann's theorem extends to all values of t the interpretation of $f(\alpha)$ as a dimension.

Interpretation of the two multifractal representations in terms of two different probability limit theorems. In terms of the probabilistic foundation of $f(\alpha)$ via $\rho(\alpha)$,

the above two forms of the multifractal representations are totally different in spirit. The box dimensional representation is the analog of the *weak law of large numbers*, which says that, as the size k of a sample increases,

$$\Pr\{|\text{sample average} - \text{expectation}| > \varepsilon\} \rightarrow 0.$$

The Hausdorff dimensional representation is the analog of the *strong law of large numbers* which is indeed far stronger (and harder to prove) and asserts that

$$\Pr\{\lim_{k \rightarrow \infty} \text{sample average} = \text{expectation}\} = 1.$$

Space lacks to elaborate upon this analogy.

5.7. Convergence Towards the Limit $f(\alpha)$ is Slow, and “Sampling Biases” are Ordinarily Very Large. Therefore, $f(\alpha)$ Represents the Generating Mechanism of the Multifractal Very Indirectly

Recall that the derivation of the function $f(\alpha)$ in Section 5.2 involves a limit argument based on the leading term in the Stirling formula for the factorial. The next question is, how fast do $\delta(k, \varphi_0, \varphi_1)$, δ^+ and δ^- converge to their limits, which we know is either $\delta(\varphi_0, \varphi_1)$ or 1?

Figure 3 compares the limit “Stirling” expression with the quantities $\delta(\alpha, \varphi_0, \varphi_1)$, δ^+ and δ^- based upon the full factorials. The value $k = 56$ corresponds to $dt = 2^{-56} \sim 10^{-17}$, which is minuscule. To study a probability distribution on such fine grids would require sample sizes well beyond any ordinary experiment. Yet, even at this stage, we see that the limit $f(\alpha)$ is far from having been reached.

Particularly visible is the value of α where the graphs of the functions δ^+ and δ^- cross each other. This is where the two tail probabilities are equal to each other, hence to 1/2. Probabilists call this the “median” value of α . Here, $\delta^+ = \delta^- = \log 2^{k-1}/\log 2^k = 1 - 1/k$. This is a special case of the extremely simple result, $\delta^+ = \delta^- = 1 + 1/\log_2(dt)$, which holds for *all* multifractals. Anyhow, whenever the distribution of α is symmetric, as it is in the binomial case, it is obvious that the median α is also the mode α (the most probable α) and the expectation $\langle \alpha \rangle$. For other multiplicative multifractals, the median, the mode and the expectation need not converge for $k \rightarrow \infty$, but a sufficient condition for convergence is that the bell of the limit distribution be asymptotically symmetric. This is the case when $\langle \log M \rangle < \infty$.

This bias elicits a subtle conceptual question. There is no question that $f(\alpha)$ does describe the process that generates the binomial measure. But what does $f(\alpha)$ really tell us about actually observable quantities, namely, those relative to the premultifractal measures that correspond to values of $k < \infty$ that are of “sensible size?”

To pinpoint the issue by contrast, recall what happens with self-similar fractals. For them, the concept of fractal dimension is also best defined by a limit. Yet, we argued (convincingly) in our books that it is concretely and practically representative and useful, because it also has a direct relevance to ordinary-sized fractals.

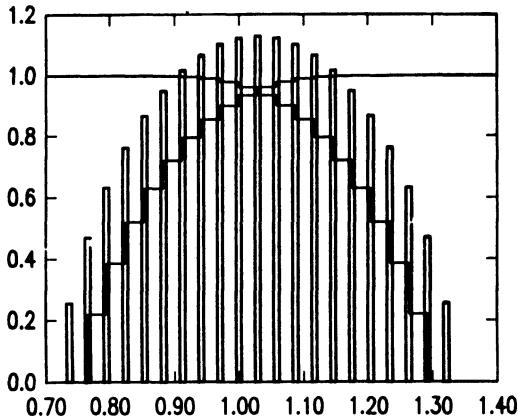


Figure 3

Three forms of the theoretical distribution of the binomial measure $\mu(dt)$. Calculations were carried out for $k = 20$, resulting in $dt = 2^{-20}$, namely, about a millionth. The three alternative distributions are plotted in “rescaled doubly logarithmic” form. That is, the horizontal is not μ itself, but $\alpha = \log \mu(dt)/\log(dt)$. The vertical bars simulate the “density” of α , i.e. (α being discrete) a probability divided by $\Delta\alpha$. The monotone increasing and decreasing step curves represent the tail probabilities $\Pr\{A < \alpha\}$ and $\Pr\{A > \alpha\}$. All three functions are plotted as $(\log \text{of a function})/\log(dt)$. For $k \rightarrow \infty$, all three curves have the same limit $\rho(\alpha)$. One has $\rho(\alpha) = 0$ and $f(\alpha)\rho(\alpha) + 1 = 1$ for the center of the diagram. Thus, the densities “overshoot” the limit, and the tails “undershoot” it. See Sections 5.7 and 6.6 for a discussion.

In a binomial multifractal (and *a fortiori* in more general ones), the situation is far harder. The $f(\alpha)$ is exclusively an asymptotic notion that is quite distant from representing the distribution of the measures over intervals of ordinary sizes dt .

The issue of the lack of direct significance of $f(\alpha)$ is especially acute in the case of turbulence, or of all other phenomena in which the fractal range has a lower cut-off that is not infinitely far, but reached by ordinary experiments. In those cases, indeed, the passage to the limit that seems to define $f(\alpha)$ has no concrete meaning. Thus, to say that the measure $\mu(dt)$ is *directly* represented by the limit would be a fiction. However, again, $f(\alpha)$ does indeed provide a useful *indirect* representation of the generating mechanism. Section 6.6 will tackle the problem of how $f(\alpha)$ can be inferred from a finite sample of data.

6. The Second Most Basic Multifractal is the Multinomial Measure. Introduction of the Legendre Multifractal formalism

6.1. Basic Background. The Domain of all Possible Points (α, δ)

To construct a multinomial of base $b > 2$ requires b masses m_β ($0 \leq \beta \leq b - 1$) adding to 1. Denote by Φ the point whose coordinates are the frequencies φ_β

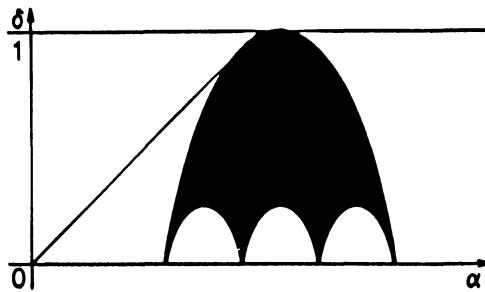


Figure 4

Rough idea of the domain of (α, δ) for a multinomial multifractal with $b = 4$. The domain's upper boundary defines the function $f(\alpha)$. Here, all the m_β are different. Clearly, $\alpha_{\min} = \min(-\log_b m_\beta) > 0$, and $\alpha_{\max} = \max(-\log_b m_\beta) < \infty$.

of the digits β in the base- b development $0, \eta_1 \eta_2 \dots \eta_k$. Every b -adic interval characterized by Φ yields

$$\mu(dt) = \prod m_\beta^{k\varphi_\beta}, \text{ hence } \alpha = -\sum \varphi_\beta \log_b m_\beta \quad \text{and} \quad \delta = -\sum \varphi_\beta \log_b \varphi_\beta.$$

In the binomial case, a single $\delta = f(\alpha)$ could be deduced from the value of α , but this possibility is not available here. A given α , indeed, allows a host of possible sets of values of φ_β constrained by $\sum \varphi_\beta = 1$. Each of these sets has its own δ , and the multinomial representation would not be a sum over a sample index α , but a sum over two indexes α and δ . More precisely, the Φ 's yielding a given α define a portion of hyperplane, on which the function δ varies continuously. Therefore, the possible values of δ for given α cover an interval. After the points (α, δ) corresponding to all the values of α have been combined, the result is a domain of the plane, as shown in black in Figure 4.

This domain is bounded *to the left* by $\alpha_{\min} = \min\{-\log_b m_\beta\}$, and *to the right* by $\alpha_{\max} = \max\{-\log_b m_\beta\}$. It is bounded from *below* by $\delta \geq 0$. In fact, it has as its more precise lower boundary a certain scalloped curve. For example, when no two m_β are identical, (including $\alpha = \alpha_{\min}$ and $\alpha = \alpha_{\max}$), $\delta_{\min} = 0$ when $\alpha = \alpha_\beta$, and between two successive α_β 's, δ_{\min} follows an entropy arch of height $\log_b 2$. When several values of β yield the same α_β , the arches on both sides of α_β are rescaled by an affinity. For example, if α_{\min} is attained for N_β distinct β 's, $\delta_{\min}(\alpha_{\min}) = \log_b N_\beta$.

6.2. Thermodynamical Manipulations. The Lagrange Multipliers Argument, and the Legendre Relation of the Gibbs Theory

For a given value of α , the δ 's are dominated by the term whose dimension is the highest. This is the term that maximizes $-\sum \varphi_\beta \log_b \varphi_\beta$, given $-\sum \varphi_\beta \log_b m_\beta = \alpha$, and $\sum \varphi_\beta = 1$. To solve this problem, we use the classical method of Lagrange

multipliers (HUANG, 1966, Chapter 8). It introduces a multiplier q , with $-\infty < q < \infty$, and yields

$$\varphi_\beta = \frac{b^{q \log_b m_\beta}}{\sum b^{q \log_b m_\beta}} = \frac{m_\beta^q}{\sum m_\beta^q}.$$

The customary roles of the “partition function” and the “free energy” are played here by the quantities $\sum m_\beta^q$ and $\tau(q) = -\log_b \sum m_\beta^q$.

In terms of $\tau(q)$, the Lagrange multipliers determine q and $f(\alpha)$ from α by

$$\begin{aligned} \alpha &= -\sum \varphi_\beta \log_b m_\beta = -\frac{\partial}{\partial q} \log_b \sum m_\beta^q \frac{\delta \tau(q)}{\delta q} \\ \max \delta = f(\alpha) &= -\frac{\sum \left(q \log_b m_\beta - \log_b \sum m_\beta^q \right) m_\beta^q}{\sum m_\beta^q}. \end{aligned}$$

That is,

$$\alpha = \frac{\partial \tau(q)}{\partial q} \quad \text{and} \quad f(\alpha) = q \frac{\partial \tau}{\partial q} - \tau = q\alpha - \tau.$$

The black domain in Figure 4 is now replaced by its upper boundary, which is the graph of a function $f(\alpha)$. Since in the multinomial case the black domain of α and δ satisfies $\alpha > 0$ and $\delta \geq 0$, we see that $f(\alpha) \geq 0$, $\alpha_{\min} > 0$, $f(\alpha_{\min}) \geq 0$ and $f'(\alpha_{\min}) = \infty$, and $\alpha_{\max} < \infty$, $f(\alpha_{\max}) \geq 0$ and $f'(\alpha_{\max}) = -\infty$. Multifractals that are not multinomial, yet possess these properties we call “pseudo-multinomial.”

Note, incidentally, that

$$\varphi_\beta = -\frac{1}{q} \frac{\partial \tau}{\partial (\log_b m_\beta)} = \log_b \frac{m_\beta}{q} \frac{\partial \tau}{\partial m_\beta}.$$

The moments of μ . When a cell is selected at random, the multiplier is a random variable M that takes the value m_β with the probability $1/b$. Its q -th moment is $\langle M^q \rangle = \sum (1/b)m_\beta^q$. The quantity $\log_b \langle M^q \rangle = -1 - \tau(q)$ in a common way of specifying a random variable in mathematical statistics, and it is called “cumulant generating function.” (Note: The sign attributed to τ in the usual multifractal notation is unfortunate, and a number of authors, e.g., FEDER (1988), spontaneously select the opposite sign.) Similarly,

$$\langle \mu^q(dt) \rangle = [\langle M^q \rangle]^k = \left[b^{-1} \sum m_\beta^q \right]^k = [b^{-1 - \tau(q)}]^k = (dt)^{1 + \tau(q)}.$$

Adding the contributions μ_i of the $(dt)^{-1}$ intervals of length dt , we find

$$\left\langle \sum \mu_i^q \right\rangle = (dt)^{-1} \langle \mu^q(dt) \rangle = (dt)^{\tau(q)}.$$

In the approach of HALSEY *et al.* (1986), this is the *definition* of $\tau(q)$.

6.3. Inverse Legendre Relations, Analytic and Graphical

Knowing $\tau(q)$ for all values of q , one can trace all the straight lines of equation $\delta_q(\alpha) = q\alpha - \tau$. The analytic solution is that these straight lines define $f(\alpha)$ as their envelope, namely

$$f(\alpha) = \min_q (q\alpha - \tau).$$

A graphical approach is often more direct. If the lines represented by $\delta_q(\alpha)$ are traced in black, they merge into a second black domain in the (α, δ) plane, which “surrounds” the black domain of $\{\alpha, \delta\}$ that we have considered previously.

6.4. Thermodynamical Analogies

If the present approach is used, the observations that, $q \sim$ inverse temperature, $\tau \sim$ Gibbs free energy, and $f \sim$ entropy are not after-the-fact “thermodynamic reinterpretations” of multifractals. They are embedded in the multifractals’ origin in the Lagrange multipliers. Further use of the thermodynamic interpretation is beyond the scope of this paper.

6.5. The “Darwin Fowler” Derivation of the “Multifractal Formalism” and Actual Computation

FRISCH and PARISI (1985) and HALSEY *et al.* (1986) obtain the same equations $\alpha = \tau'$ and $f = q\alpha - \tau$ via a steepest-descent argument. Experts who have learned statistical mechanics in older books recall that Darwin and Fowler have used steepest descents to justify rigorously the Lagrange multipliers procedure. (See also HUANG, 1966, Section 10.1.) However, *no one* teaching thermodynamics will start with the Darwin–Fowler method, without having first presented the Lagrange multipliers. Therefore, the path towards the formalism taken in Section 6.2. involves the least effort and the fullest understanding. Section 7 describes the next simplest generalization.

6.6. Estimation of $f(\alpha)$. Statistical Corrections for the Small Sample Bias

To deduce $f(\alpha)$ from a finite sample of data can be done in many different ways. We shall give one “quick and dirty” method, then several examples of inference from an observable fractal range to an inaccessible limit.

A quick and dirty correction from a single histogram. Section 5.7 shows that the convergence to the limit $\delta(\varphi_0, \varphi_1)$ is much faster for the tail functions $\delta^+(k, \varphi_0, \varphi_1)$ and $\delta^-(k, \varphi_0, \varphi_1)$ than it is for the density functions $\delta(k, \varphi_0, \varphi_1)$. Moreover, we already know that the value of the bias at the median where $\delta^+ = \delta^-$ is “universal”, in the sense that $\text{bias} = 1/\log_2(1/dt)$, independently of $f(\alpha)$. Assuming that the bias is the same for all α , which is the binomial case, suggests the following

“Recipe”: Suppose that $\mu(dt)$ is known down to cells of length $2^{-k_{\max}}$. Evaluate the sample tail function $\delta^*(k_{\max})$. An estimate of $\delta(\alpha)$ is provided by $*(k_{\max}) + 1/k_{\max}$.

The moments and Legendre transform route to the estimation of $f(\alpha)$ and $\zeta(\alpha)$. The preceding procedure is extremely crude, but there are several alternatives. Sections 6.2 and 6.3 suggest one alternative, which has become widely used. First, estimate $\tau(q)$ from the data; next, compute $f(\alpha)$ as the Legendre transform of the estimated $\tau(q)$. This path is deeply imbedded in the approach of Frisch and Parisi and of Halsey *et al.*, since their approach defines $f(\alpha)$ in this fashion.

However, the path via $\tau(q)$ is not unique. Let us now describe the principle underlying this and many other estimation procedures, then list alternative implementations.

How to infer $f(\alpha)$ from data within an observable fractal range. Suppose that $[0, 1]$ has been divided into $2^{k_{\max}}$ minimal intervals of length $2^{-k_{\max}}$. Then the measure $\mu(dt)$ is of course known for each of these intervals. But for every $k < k_{\max}$, $\mu(dt)$ is also known for 2^k intervals of length 2^{-k} . The values of $\mu(dt)$ carry considerable amounts of information. This information can be processed by evaluating $\tau(q)$, but have been a vocal advocate of alternative methods.

Two such methods are found in MENEVEAU and SREENIVASAN (1989) and in CHHABRA and JENSEN (1989). There is room for many more.

6.7. Generalization of the Scope of the Binomial and Multinomial Multifractals Defined on $[0, 1]$

The preceding discussion is of wider generality than has been suggested thus far. Many physicists seem to be under the impression that multifractal measures are always supported by fractal sets. The discussion thus far has shown, to the contrary, that the simplest multifractals are those supported by $[0, 1]$. Now, we proceed to the easy task of demonstrating that, when the bases are matched in a way to be described, a multinomial measure on $[0, 1]$ maps upon one that is supported by a cube or by a fractal set constructed recursively.

Mapping a fractal on $[0, 1]$. Consider a fractal set that is constructed recursively using a generator made of N “sticks”, which are intervals of equal lengths. The best known examples are the generalized Koch curves (FGN, p. 142 shows that this family also includes the Sierpinski gasket). Recursion defines a point P on such a fractal as being on stick number η_z in the first recursion stage, . . . stick number η_k in the k -th stage. Together, the integers η_k define a real number in the base N , namely $t = 0.\eta_1, \dots, \eta_k, \dots$, with $0 \leq t \leq 1$. In this way, every interval $[dt]$ of $[0, 1]$ is mapped on a piece of our fractal, which FGN calls a “fractal interval.”

Mapping a multifractal. Now distribute on $[0, 1]$ an arbitrary multinomial measure of base N . The above map transforms it automatically into a multinomial measure on our fractal.

Double points. It is known, of course, that some P ’s are characterized by several

values of t . Fortunately, however, the multifractal measures are continuous (though not differentiable). When the P 's for which t is ambiguous are denumerable, these ambiguous t 's do not matter. (There are several sources of ambiguity. The first is that if t is a multiple of N^{-k} , it can be represented as ending in an infinity of 0's or of $(N-1)$'s. A second source of ambiguity follows from double points in the Koch fractal; for example, the Sierpinski gasket has asymptotic double points.)

Cubes. A Peano motion (our term to replace the improper term, Peano curve) can, similarly, be used to map a given multinomial measure on $[0, 1]$ upon a cube $[0, 1]^E$. The only condition is that the multinomial's base N be such that $N^{1/E}$ is an integer.

Effect of mapping upon α , $f(\alpha)$ and $\rho(\alpha)$. This effect is, most fortunately, very simple and even amusing. To fit a widely used notation, denote our fractal's dimension by D_0 . Denote by $|dP|$ the Euclidean distance between the endpoints of the map dP of the linear interval dt . It is easy to see that $|dP| = |dt|^{1/D_0}$.

Therefore for the measure μ' on the map of $[0, 1]$, α is replaced by

$$\alpha' = \log \mu'(dP) / \log |dP| = D_0 \log \mu(dt) / \log(dt) = D_0 \alpha.$$

Moreover, our map multiplies the fractal dimension of a set in $[0, 1]$ by D_0 . Hence, the f -function is replaced by $f' = fD_0$. Thus,

$$f'(\alpha')/D_0 = f(\alpha'/D_0).$$

When $f'(\alpha')$ is plotted in the "reduced" coordinate states, α'/D_0 and f'/D_0 , the shape of its graph is independent of D_0 .

7. The Random 1974 Multifractal Measures, Part I. The Conservative Case and the Introduction of "Latent" α 's, with Negative $f(\alpha)$'s

Now we proceed to the first of two fully studied cases of the exactly renormalized multifractals. They are the "1974 multiplicative multifractals" introduced in MANDELBROT (1974). The reader who has access to MANDELBROT (1989) will find it fruitful to study that paper at this point.

7.1. The Conservative 1974 Measures on $[0, 1]$

First, let us modify the multinomial measure as follows: The indexes of the masses m_β are shuffled at random before each stage of the cascade that distributes the mass within each of the k -th level intervals into masses within the $(k+1)$ st level intervals. (Note that one must *not* shuffle mass between *all* the k -th level intervals!) The resulting "shuffled multinomials," are the next exactly renormalizable stage beyond the multinomials. Observe that shuffling does not modify $\rho(\alpha)$, hence does not modify $f(\alpha)$.

The 1974 multifractals enter logically as the next stage after the shuffled multinomials. They require a base $b \geq 2$ and a multiplier that is not limited to the multinomial's prescribed finite collection of m_β 's, but may be a more general random variable. Since it may depend on β , it will be denoted by $M_\beta \geq 0$.

Since mass is conserved while it is spread around within a cell, strong "conservation" rules constrain the multiplier. For example, the b random multipliers $M_\beta (0 \leq \beta \leq b - 1)$ must add to 1. Consider the representative point whose coordinates are these b multipliers M_β . This is a random point restricted to an "admissible set," namely, to the portion of b -dimensional space defined by $M_\beta \geq 0$ and $\sum M_\beta = 1$. When a cell is chosen at random, the mass is multiplied by one of the M_β , chosen with equal probabilities $1/b$. Therefore, the average multiplier M is the mixture of the b coordinates of the representative point. Clearly, $0 \leq M \leq 1$ and $\langle M \rangle = b^{-1}$. The conservative 1974 measures *can, but need not*, add the further assumption that the coordinates M_β are identically distributed. If they are, M is the same random variable as any of the coordinates M_β .

By a repetition of this scheme, the b -adic cell of length b^{-k} , starting at $t = 0, \eta_1 \eta_2 \dots \eta_k$ is given the mass

$$M(\eta_1)M(\eta_1, \eta_2) \dots M(\eta_1, \dots, \eta_k) = \prod M.$$

Here, the successive random multipliers M for given t are identically distributed and independent. Our b -adic cell, therefore, yields

$$\alpha = -(1/k)[\log_b M(\eta_1) + \log_b M(\eta_1, \eta_2) \dots].$$

Thus, α is the average of k independent random variables. Let us now show that, for the desired α 's, its limit distribution for $k \rightarrow \infty$ is far from "intuitive."

7.2. Five Examples of Sums of Independent Random Variables, for which the Probability Distribution is Known in Analytically Closed Form for Both the Individual Addends, and for their Sum or Average

To obtain the limit distribution of α for $k \rightarrow \infty$, the limit theorems known to every physicist (law of large numbers and central limit theorem) give results without interest. One needs, instead, the general results to be described in Section 7.3. In order to motivate these theorems, let us begin by special random multipliers M that allow an analytically explicit evaluation of the distribution of α .

Originally, these examples were singled out as being among the few known examples with the desired property. After the fact, however, the first three examples have proven to be of intrinsic interest. Example A) illustrates Section 5. Example B) illustrates the new feature that enters in this Section 7, and Example C) illustrates the new feature that enters in Section 8. Our random variables, whether discrete or

continuous, will stand for $-\log_b M$ or for $-\sum \log_b M_i$. To avoid unwieldy formulas, however, they will be denoted by X_i and their values by x_i . In each case, we denote the density by $p(x)$ for each addend, by $p_k(x)$ for the sum of k addends, by $kp_k(k\alpha)$ for the average α of k addends. We also consider the quantity $(1/k) \log_b [kp_k(k\alpha)]$, which (in effect) was already used in Section 5.3.

A) *The binomial distribution.* Section 5.3 implies that $(1/k) \log_2 [kp_k(k\alpha)] \rightarrow \rho(\alpha) = f(\alpha) - 1$, where $f(\alpha)$ is the “entropy” function described in Section 4.2.

The striking features of this example are that $\alpha_{\min} > 0$ with $f'(\alpha_{\min}) = \infty$ and $\alpha_{\max} < \infty$ with $f'(\alpha_{\max}) = -\infty$.

B) *The Gamma distribution.* When the parameter is γ , its density is $p(x) = x^{\gamma-1} e^{-x}/\Gamma(\gamma)$. The sum of two Gamma random variables of respective parameters γ' and γ'' is a Gamma random variable of parameter $\gamma' + \gamma''$. Therefore, $p_k(x) = x^{k\gamma-1} e^{-x}/\Gamma(k\gamma)$. By the Stirling approximation, $(1/k) \log_e [kp_k(k\alpha)] \rightarrow \gamma \log(\alpha/\gamma) - \alpha + \gamma = \rho_e(\alpha)$.

Let us show that this example is encountered for the conservative 1974 multifractals when $b = 2$, so that the admissible set for the multipliers M_0 and M_1 is the interval $[0, 1]$. If both multipliers are distributed over this admissible set with uniform density, one has $\Pr\{M > m\} = 1 - m$. Writing $Y = -\log M$ yields $\Pr\{Y > y\} = e^{-y}$. This exponential is simply the Gamma distribution corresponding to $\gamma = 1$. It leads to $\rho(\alpha) = \log_2 (\alpha \log_2 2) - \alpha + \log_2 e$.

The striking features of this example are that $\alpha_{\min} = 0$ and $\alpha_{\max} = \infty$, and that $\rho(\alpha_{\min}) = \rho(\alpha_{\max}) = -\infty$. These properties could not happen in the multinomial case, and serve in Section 7.4 to introduce the notion of *latent* α .

The Gamma distribution and related examples are discussed in detail (with several illustrations) in MANDELBROT (1989), which—again—the reader would benefit from at this point.

C) *The Gaussian distribution* $p(x) = (2\pi\sigma^2)^{-1/2} \exp(-x^2/2\sigma^2)$. Here, $[kp_k(k\alpha)]^{1/k} = (2\pi\sigma^2/k)^{-1/(2k)} \exp(-\alpha^2/2\sigma^2)$. Thus, $(1/k) \log [kp_k(k\alpha)] \rightarrow -\alpha^2/2\sigma^2 = \rho(\alpha)$.

Computationally, this example is the simplest of the present five. But it raises novel issues, which is why it has not been listed first. The Gaussian has two unbounded tails, while the conservative model requires $M < 1$. From $M_{\max} > 1$ follows $\alpha_{\min} = -\infty$, which will serve in Section 8 to introduce the notion of *virtual* α .

D) *The Poisson distribution.* When $\langle X \rangle = \gamma$, one deals with the discrete probability $p(x) = e^{-\gamma} \gamma^x / x!$. Here $p_k(x) = e^{-k\gamma} (k\gamma)^x / x!$ By the Stirling approximation,

$$(1/k) \log_e [kp_k(k\alpha)] \rightarrow -\gamma + \alpha \log_e (\gamma e / \alpha) = \rho_e(\alpha).$$

E) *The Cauchy distribution.* Its density is, of course, $p(x) = [\pi(1+x^2)]^{-1}$. By a property called “Lévy stability,” one has $p_k(x) = [\pi k(1+x^2/k^2)]^{-1}$, hence $\rho_e(\alpha) = 0$. This result means the Cauchy case falls quite outside the pattern found in cases A) to D). The renormalization that it requires is altogether different. But the random variable $M = \exp(\text{Cauchy variable})$ satisfies $\langle M^q \rangle = \infty$ for all q . Therefore, it is impossible to use it as a multiplier satisfying conservation on the average.

Features common to examples A) to D). 1) To form the quantity $(1/k) \log p_k(k\alpha)$ is to renormalize the average in a way that greatly accentuates the low probability “tails,” and deemphasizes the central “bell.” In examples, we have found that the newly renormalized average still converges to a limit, for which $\rho(\alpha) \leq 0$, with $\rho(\alpha) = 0$ for one value of α . Namely, for the expected value $\langle \alpha \rangle$ of α .

In examples B), C) and D), the analytic form of $\exp[\rho(\alpha)]$ is nearly identical to that of $p(\alpha)$, except that it is “wrongly normalized.” It does not integrate to 1, and would not be proper probability density. A consequence of nonintegrability is that, before the asymptotic range is reached, each of the above quantities $(1/k) \log p_k(k\alpha)$ is >0 over a range of α ’s near the maximum of $\rho(\alpha)$. That is, as in the binomial case, direct estimates from data collected for a finite k overestimate $\rho(\alpha)$ near its maximum.

Differences between the five examples. Gross nonuniversality. All the above limits differ from one another. For example, the limits for the non-Gaussian cases differ from the limit relative to the Gaussian case. In Section 7.3, this surprising result will be shown to hold very generally. It is the second most important probabilistic fact underlying the multifractals (the most important fact being the behavior of moments along low dimensional cuts; Section 8).

7.3. Generalization of the Multifractal Formalism by Application of the Cramèr Limit Theorem. Profound Nonuniversality

The above examples are very special cases of the “large deviations theorems” of H. Cramèr; see BOOK (1984), CHERNOFF (1952), DANIELS (1954, 1987). That the proper mathematical result is available “off-the-shelf” is a pleasant surprise. The “local form” of the Cramèr theorem asserts the following:

As $k \rightarrow \infty$, the quantity $(1/k)\log_B$ (probability density of α) converges to a limit, to be denoted as $\rho(\alpha)$.

The tail quantities $(1/k) \log_B$ (probability of $\alpha > \langle \alpha \rangle$) and $(1/k) \log$ (probability of $\alpha < \langle \alpha \rangle$) converges to the same limit. It is easily shown that

$$f(\alpha) = \rho(\alpha) + 1 = \rho(\alpha) + \text{dimension of the measure's support.}$$

The Legendre formalism. In this multinomial case, $\tau(q)$ could be written either as $-\log_b \sum m_i^q$ or as $-1 - \log_b \langle M^q \rangle$. This second form remains meaningful for all random variables M . With this definition of $\tau(q)$, the generalized Gibbs formalism resulting from the Cramèr theorem remains valid. That is, the relation between $\tau(q)$ and $f(\alpha)$ follows the Legendre rules we have encountered in Section 5 for the multinomial measures.

Failure of universality. Moreover, there is no universality, in the specific narrow sense exemplified in Section 7.2: different multipliers yield different $f(\alpha)$'s, and conversely.

Obviously, Cramér-type theorems extend to the case when the factors M are weakly dependent or weakly nonidentical. See AZENCOTT *et al.* (1980) and DAWNHA-CASTELLE (1979) for various generalizations.

7.4. Negative $f(\alpha)$'s, and the Distinction Between Manifest and Latent Values of α

Negative $f(\alpha)$. The 1974 multifractals introduce a very important theme. First, suppose that M is a discrete random multiplier and that its largest possible value M_{\max} satisfies $\Pr\{M = M_{\max}\} < b^{-1}$, which is perfectly allowable. It follows that $f(\alpha_{\min}) = \log_b [\Pr\{M_{\max}\}] + 1 < 0$. As a matter of fact, $f(\alpha_{\min})$ can be arbitrarily large in absolute value, and negative. Similarly, $\Pr\{M = M_{\min}\} < b^{-1}$ results in $f(\alpha_{\max}) < 0$.

Secondly, the example of the Gamma distribution suffices to show that the 1974 multifractals allow $\alpha_{\min} = 0$ and $\alpha_{\max} = \infty$, with $f(\alpha_{\min}) = -\infty$ and/or $f(\alpha_{\max}) = -\infty$. To show that this result is intuitive, take a more general M that is continuously distributed, but can be approximated arbitrarily close by a discrete multiplier with $\log M_{\max}$ arbitrarily close to 0 and/or $\log M_{\min}$ arbitrarily close to $-\infty$, with both extremes having correspondingly arbitrary tiny probabilities. Then $\alpha_{\min} = 0$ and/or $\alpha_{\max} = \infty$.

A new and important distinction. The α 's such that $f(\alpha) > 0$ will be called *manifest*. The remaining α 's will be called *latent*, which means "hidden but present." (For negative latent dimension in a different context, see MANDELBROT (1984).)

In the special case $\Pr\{M_{\max}\} \geq b^{-1}$ and $\Pr\{M_{\min}\} \geq b^{-1}$, all α 's are manifest. Such a multifractal can be called "pseudo-multinomial."

Crossover from the manifest to the latent α 's. The existence of latent α 's is characterized by crossover values α_{\min}^* and α_{\max}^* , such that $f(\alpha_{\min}^*) = 0$ and $f'(\alpha_{\min}^*) = q_{\max}^* < \infty$ and/or $f(\alpha_{\max}^*) = 0$ and $f'(\alpha_{\max}^*) = q_{\min}^* > -\infty$.

The effect of latent α 's on the moments. Let us call "manifest $f(\alpha)$ " the function $f(\alpha)$ restricted to manifest values of α . Its slopes, q_{\max}^* for α_{\min} and q_{\min}^* for α_{\max} fail to be infinite. Therefore, the manifest $f(\alpha)$ differs from the full $f(\alpha)$ function when there are latent α 's. How much do the latent α 's contribute to $\tau(q)$ and D_q ? For $q_{\min}^* < q < q_{\max}^*$, they contribute nothing. But for large positive and/or negative q , their contribution is predominant. Therefore, the estimates of $\tau(q)$ or D_q made on the basis of manifest α 's is truncated and biased.

Note that the "full" functions $\tau(q)$ and $D_q = \tau(q)/(q-1)$, evaluated on the basis of all α 's are analytic functions. But the "manifest" $\tau(q)$ or D_q evaluated on the basis of manifest α 's is incorrect (truncated and biased) for large $|q|$'s.

There is a seeming paradox here. On the one hand, the probability outside of the central bell tends to 0 as $k \rightarrow \infty$, meaning that the tails become thoroughly

insignificant. In the limit $k \rightarrow \infty$, the most probable value, the expectation and the other usual parameter of location all converge with each other. On the other hand, those “negligibly” few values in the tails are so huge that their contributions to all the moments of order $q \neq 0$, and to $\tau(q)$, are predominant. Moreover, the moments and $\tau(q)$ depend on the exact form of $f(\alpha)$, that is, are nonuniversal.

The meaning of negative values for the “dimension” $f(\alpha)$. When $f(\alpha)$ is viewed as a fractal dimension, the notion that $f(\alpha) < 0$ would be self-contradictory, that is, would be impossible. This self-contradiction does not deny that $\rho(\alpha)$ is useful. It only expresses the fact that $f(\alpha) = \rho(\alpha) + 1$ cannot always be interpreted as a fractal dimension. As a matter of fact, several authors have approached this issue from different angles, and have shown that restricting multifractals to $f(\alpha) \geq 0$ lead to self-contradiction (CATES and DEUTSCH, 1987), or is otherwise not acceptable (FOURCADE *et al.*, FOURCADE and TREMBLAY, 1987).

From the b^k data given by a single sample of measure, one can estimate $f(\alpha)$ if the probability of α is $> b^{-k}$, which means that $f(\alpha) \geq 0$. The range of the α 's is *not* expanded by increasing k , because the population from which the sample is drawn also depends upon k . This invalidates the intuitive belief that the sample value necessarily converges with the population expectation as the sample increases.

In order to measure the negative $f(\alpha)$ for latent α 's, the only generally valid method is to combine data from very many samples.

A remark linking Sections 7 and 8. Very often, negative $f(\alpha)$'s have a different cause, to be described in Section 8. They appear because one considers low-dimensional cuts through higher-dimensional measures. In that case, one extends the ranges of observable α 's by investigating the original measure through higher-dimensional cuts. The reason is that “generically” one should expect the overall original measure to include cells of higher or lower density, “hotter” or “colder”, than the hottest or coldest cells along the one-dimensional cut.

8. The Random 1974 Multifractal Measures, Part II. The Canonical Case and the Properties of Random Cuts. Introduction of the Critical Exponent q_{crit} and of its Inverse Function D_q

8.1. The Canonical 1974 Measures on $[0, 1]$ and Random Cuts

In this next random generalization of the binomial multifractal, the multipliers $M \geq 0$ are allowed to be statistically independent. This assumption implies, of course, that (almost surely) the cascade *does not conserve mass*. However, it is easy to insure that *mass is conserved on the average*. It suffices that the condition $\langle M \rangle = 1/b$, which used to be a corollary of conservation, be preserved as an autonomous requirement. The assumption of independence is motivated by an issue already discussed in Section 1.4. Many multifactor measures can only be observed

through cuts. To illustrate, we imagine a conservative cascade of cubic cells of dimension E . Each cubic cell contains b^E cubic subcells, but a straight edge of the initial cube intersects only b cells. If $b^E \gg b$, the conservation laws $\Sigma M(\eta_1) = 1$ hardly constrain the multipliers M within these subcells. These M can be assumed independent.

Independence of the multipliers brings essential new features.

8.2. The Random Variable Ω

In order to understand the structure of the cuts and the role of the “critical dimension” D_q , it is necessary to introduce a random quantity to be denoted by Ω (MANDELBROT, 1974). Consider a canonical cascade that has started with mass 1 in $[0, 1]$, and has continued over infinitely many stages. Because of the lack of conservation the ultimate mass is not identical to 1, but is a random variable. It will be denoted by Ω .

An alternative way to write Ω is to add the masses in the b subcells of length $1/b$. In the first cascade stage, the β -th interval is given the mass $M(\beta)$. Ultimately, it contains the mass $M(\beta)\Omega(\beta)$, the various quantities $M(\beta)$ and $\Omega(\beta)$ being statistically independent. The sum of the ultimate masses can be written as either $\Sigma M(\beta)\Omega(\beta)$ or Ω . This identity can be written as

$$\sum M_\beta \Omega_\beta \equiv \Omega \quad (\equiv \text{means “identical in distribution”}).$$

Thus Ω is the fixed point of the operation of randomly weighted averaging using as weights the random quantities $M(\beta)$. This Ω , first introduced in MANDELBROT (1974), has been the object of several papers of pure probabilists.

8.3. The Mass $\mu(dt)$ in a Cell of Length b^{-k} : Low- and High-Frequency Factors

Let the cell start at $t = 0$, $\eta_1 \eta_2 \dots \eta_k$. The first k -stages of the cascade can be called *low frequency* stages. They generate precisely the same effects as in the conservative cascade of Section 7.1. That is, they yield the multiplicative term

$$M(\eta_1)M(\eta_1, \eta_2) \dots M(\eta_1, \dots, \eta_k) = \prod M.$$

This $\prod M$ is to be called a *low frequency* (or *long wavelength*) *factor*, because it is generated by multipliers that are constant over b -adic intervals of length $dt = b^{-k}$ or shorter.

However, in contrast to the conservative case, that is not all! Each cell is also subjected to precisely the same process as has been described for $[0, 1]$. Therefore, the measure does *not* reduce to $\prod M$, but instead takes the form

$$\mu(dt) = \Omega(\eta_1, \dots, \eta_k)M(\eta_1)M(\eta_1, \eta_2) \dots M(\eta_1, \dots, \eta_k) = \Omega \prod M.$$

That is, each cell involves a sample value of Ω , to be called a *high frequency* (or *short wavelength*) *factor*, which results from multipliers that vary over intervals that are shorter than $dt = b^{-k}$. Successive M and Ω are independent, and all the M (or Ω) are identically distributed.

8.4. The Hölder α and the Function $f(\alpha)$. Virtual Values of α

Write $\alpha = \alpha_L + \alpha_H$, with

$$\alpha_L = -(1/k)[\log_b M(\eta_1) + \log_b M(\eta_1, \eta_2) \dots]; \quad \alpha_H = -(1/k) \log_b \Omega.$$

When $k \rightarrow \infty$, $\alpha_H = -(1/k) \log_b \Omega \rightarrow 0$, hence α_H does not affect $f(\alpha)$. That is, on the logarithmic scale of the Hölder α , the high frequency factor Ω is asymptotically negligible. Moreover, the term α_L continues to be covered by the theorem of Cramér, and the “Gibbs” quantities $\tau(q)$ and $f(\alpha)$ and D_q are precisely as for the conservative 1974 measures. However, a canonical 1974 cascade allows two possibilities that have not been seen in the preceding Sections, namely virtual α 's and misbehaving $\tau(\alpha)$.

Virtual α 's. First possibility: a canonical cascade allows $M_{\max} > 1$, hence $(\alpha_L)_{\min} < 0$, with $f[(\alpha_L)_{\min}] < 0$. If we forget Ω , our α_L would be a Hölder exponent, and $\alpha_L < 0$ would be absurd, because it would imply that a subinterval b^{-k-1} includes more mass than its embedding interval b^{-k} . But α_L is *not quite* a Hölder exponent. The paradox of the $\alpha_L < 0$, to be called *virtual*, can be shown to vanish if one considers Ω .

We shall see that, either both factors ΠM and Ω are “well behaved,” or both “misbehave” simultaneously, each in its own way.

8.5. The Prefactors $\langle \Omega^q \rangle$ Are Infinite when q is above the Critical Exponent q_{crit}

We have seen that Ω is negligible in the scale of the α . But in the scale of μ , taking account of Ω changes the familiar $\langle \mu^2(dt) \rangle = (dx)^{\tau(q)+1}$ into $\langle \mu^q(dt) \rangle = \langle \Omega^q \rangle (dt)^{\tau(q)+1}$. The new moment prefactors $\langle \Omega^q \rangle$ reflect high frequency effects due to scales $< b^{-k}$. They have far reaching effects, because *they need not be finite*. Indeed, MANDELBROT (1974) shows (and KAHANE and PEYRIÈRE, 1976 then proves rigorously) the following.

The inequality $\langle \Omega^q \rangle < \infty$ holds if, and only if, $q < q_{\text{crit}}$, where q_{crit} is the solution of the equation $\tau(q) = D_q = 0$. When $q > q_{\text{crit}}$, then $\tau(q) < 0$ and $D_q < 0$.

Without the presence of divergent prefactor $\langle \Omega^q \rangle$, the inequality $\tau(q) < 0$ for $q > q_{\text{crit}}$ would have meant that $\langle \mu^q(dt) \rangle$ actually *decreases as dt increases*. This would be an extreme anomaly, but it does *not* occur, because of $\langle \Omega^q \rangle = \infty$.

Furthermore, the paradox of *virtual α 's* can be shown to vanish if one considers the moments of Ω .

8.6. The Tail Behavior of Ω

MANDELBROT (1974) has conjectured, and GUVARCI'H (1987) has proven, that $\Pr(\Omega > \omega) \sim \omega^{-q_{\text{crit}}}$. This is the simplest tail behavior compatible with the role q_{crit} plays for Ω . Since q_{crit} plays the same role for $\mu(dt)$, the same tail behavior also holds for $\mu(dt)$. More precisely, such is the case if dt is well above the inner cutoff (if there is a cutoff > 0).

This prediction of the model is extremely important to the analysis of experimental data. But it should be stressed that it only holds asymptotically after the cascade has been allowed to proceed down to scales well below dt . In the case of turbulence, this requires an extremely high Reynolds number R , and indeed there are reports of ill-behaved sample moments of high q . But for the lower R reached in laboratory turbulence, Ω is effectively truncated at a level increasing with R , which avoids divergent moments. Experimentalists are urged to compare the dissipation's $\tau(q)$ for different R 's.

8.7. Sequences of Cuts with Different Bases, but Identical W Multipliers

To renormalize once for all bases, it suffices to consider a cascade that multiplies the densities (instead of the masses) by a fixed weight $W = bM$, satisfying $\langle W \rangle = 1$. When 1974 measure in E -dimensions, with weight W and base b , is intersected by a fractal of dimension $D = \log B / \log b < E$, it yields a measure with the same W and the new base B . In fact, by using random cuts, one may view B as continuously variable. One has $-\tau(q) = 1 + \log_b \langle M^q \rangle = \log_b \langle W^q \rangle - (q - 1)$. This is precisely the function that MANDELBROT (1974) plots (his Fig. 2) and discusses for different values of b . (The same paper also shows that the critical q_{crit} , which it denotes by α , depends on D .)

Now consider for different B 's or D 's the enhanced graphs showing $f(\alpha)$, $\tau(q)$ and D_q (as in Figure 5), and translate the origin of each graph to the point of coordinates $\alpha = \langle \alpha \rangle$ and $f(\alpha) = D_0$. This is the proper "renormalization" to use here, because one can show that the translated graphs coincide if one includes the manifest, the latent and the virtual α 's. As a corollary, increasing B eventually changes latent α 's into manifest α 's. That is, it changes negative $f(\alpha)$'s into positive $f(\alpha)$'s, which one can estimate through the B^k data from a single sample. Also, increasing B eventually changes virtual α 's into latent, then into manifest ones. Decreasing B eventually reaches $D = D_0$, beyond which the measure becomes degenerate (MANDELBROT, 1974; KAHANE and PEYRIÈRE, 1976).

As corollary, the manifest portion of $f(\alpha)$ fails to represent a 1974 multifractal in an intrinsic fashion. But a full description of the effects of the same W in different D_0 's is provided in the context of Cramèr's limit theorem by the intrinsic quantities $\alpha - \langle \alpha \rangle$, $\rho(\alpha)$, and $D_q - D_0$.

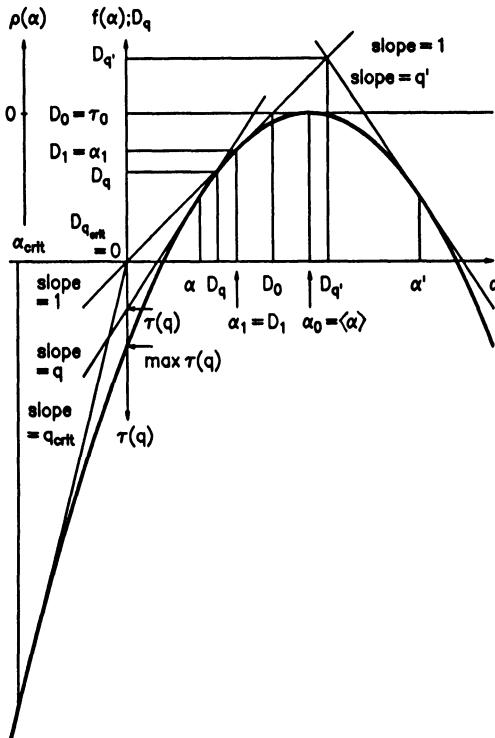


Figure 5

Enhanced version of the multifractal diagram. One vertical scale shows $\rho(\alpha) = \lim_{k \rightarrow \infty} (1/k) \log_b$ (probability) versus $\alpha = -(1/k) \log_b$ (measure). A second vertical scale shows $f(\alpha)$. In this particular graph, the weight W is lognormal, but μ itself is not lognormal at all. This example combines all three multifractal regimes described in the text: manifest α 's ($\alpha < 0$ and $0 \leq f(\alpha) \leq \min(\alpha, D_0)$), latent α 's ($\alpha > 0$ but $f(\alpha) < 0$) and virtual α 's ($\alpha < 0$ and $f(\alpha) < 0 < \alpha$). As is known, $q = f'(\alpha)$ and $-\tau(q)$ is the intercept of the tangent of slope q by the vertical axis. Let us also observe, in addition, that D_q is either coordinate of this tangent's intercept by the line $f(\alpha) = \alpha$.

8.8. Role of the D_q as Critical Dimensions

The last task is to explain why, for 1974 measures, D_q is a *critical* dimension for the moments of order q , as shown in MANDELBROT (1974). Later, the D_q were written down, in HENTSCHEL and PROCACCIA (1983), without any motivation, and called “generalized dimensions” (to call D_q a dimension can make no geometric sense when $D_q > E$ for low enough q 's).

The key to our motivation for D_q , again, resides in the study of low dimensional cuts of higher dimensional multifractals with the same $f(\alpha)$.

The discussion is simplest when $\max D_q < \infty$. The main step is to embed the cascade in a space of dimension $E > \max D_q$. We know that the equation $D_q = 0$ defines the q_{crit} relative to the property that $\langle \mu^q(dx) \rangle < \infty$ and that the variability of the Ω 's (which are independent in a sample) overwhelms the variability of the low frequency component of $\mu(dx)$. Therefore, the law of large numbers implies

that along the cut with $D > D_q$, the q -th sample moment is the sum of many contributing terms, each of which is negligible in relative value. Along a cut with $D < D_q$, to the contrary, $\tau(q) < 0$, which can be shown to imply that, while $\langle \mu^q \rangle = \infty$, all the percentiles of the q -th moment tend to 0 with dt . Hence, the sample moment is very small with a very high probability, and the exceptional instances when it is *not* very small occur because of the exceptional contribution of a single huge sample addend.

Thus, the set of points that contribute to the sample q -th moment of μ is hit by generic cut spaces of dimension $> D_q$, and missed by generic cut spaces of dimension $< D_q$.

9. Final Remarks

We realize that the later part of this paper is overly hasty. MANDELBROT (1974) is (unfortunately but perhaps unavoidably) a difficult paper, and not a good reference for the student. We are preparing a more complete presentation on this topic, which will incorporate MANDELBROT (1989).

Acknowledgements

Many thanks to Jens Feder, Tamas Vicsek, Jacques Peyrière, Carl Evertsz, and David di Vincenzo for pointing out rough and/or obscure spots in earlier versions of this text.

In Memoriam Abraham Robinson (1918–1974). From the mathematical viewpoint, expressions of the form $dy \sim (dx)^x$ are examples of “*nonstandard infinitesimals*”. This is a concept from *nonstandard analysis*, a creation of Abraham Robinson. He and I had once hoped to explore this joint field of interest together, but he fell ill and passed away. The memory of this project, together with my recent accession to the Abraham Robinson Professorship of Mathematical Sciences at Yale, bring my thoughts back to this remarkable man.

REFERENCES

- AZENCOTT, R., GIVAR'C'H, Y., and GUNDY, R. F., *Ecole de Saint-Flour 1978, Lecture Notes in Mathematics* (Saint Flour, 1978) Vol. 774 (Springer, New York 1980).
- BILLINGSLEY, P., *Ergodic Theory and Information* (J. Wiley, New York 1967) p. 139.
- BOOK, S. A. (1984), *Large deviations and applications*, In *Encyclopedia of Statistical Sciences* 4, 476.
- CATES, M. E., and DEUTSCH, J. M. (1987), *Spatial Correlations in Multifractals*, Phys. Rev. A35, 4907.
- CHERNOFF, H. (1952), *A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations*, Ann. Math. Stat. 23, 493.
- CHHABRA, A., and JENSEN, R. V. (1989), *Direct Determination of the $f(\alpha)$ Singularity Spectrum*, Phys. Rev. Lett. 62, 1327.
- DACUNHA-CASTELLE, D., *Grandes Déviations et Applications Statistiques* (Astérisque 68) (Société Mathématique de France, Paris 1979).

- DANIELS, H. E. (1954), *Saddlepoint Approximations in Statistics*, Ann. Math. Stat. 25, 631.
- DANIELS, H. E. (1987), *Tail Probability Approximations*, International Statistical Review 55, 37.
- ELLIS, R. S. (1984), *Large Deviations for a General Class of Random Vectors*, The Annals of Probability 12, 1.
- FEDER, J., *Fractals* (Plenum, New York 1988).
- FOURCADE, B., BRETON, P., and TREMBLAY, A.-M. S. (1987), *Multifractals and Critical Phenomena in Percolating Networks: Fixed Point, Gap Scaling and Universality*, Phys. Rev. B36, 8925.
- FOURCADE, B., and TREMBLAY, A.-M. S. (1987), *Anomalies in the Multifractal Analysis of Self-similar Resistor Networks*, Phys. Rev. A36, 2352.
- FRISCH, U., and PARISI, G., *Fully developed turbulence and intermittency*, In *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics* (International School of Physics "Enrico Fermi," Course 88) (eds. Ghil, M.) (North-Holland, Amsterdam 1985) p. 84.
- GRASSBERGER, P. (1983), *Generalized Dimensions of Strange Attractors*, Phys. Lett. 97A, 227.
- GUTZWILLER, M. C., and MANDELBROT, B. B. (1988), *Invariant Multifractal Measures in Chaotic Hamiltonian Systems, and Related Structures*, Phys. Rev. Lett. 60, 673.
- GUIVARC'H, Y. (1987), *Remarques sur les Solutions d'une Equation Fonctionnelle Non Linéaire de Benoît Mandelbrot*, Comptes Rendus (Paris) 305I, 139.
- HALSEY, T. C., JENSEN, M. H., KADANOFF, L. P., PROCACCIA, I., and SHRAIMAN, B. I. (1986), *Fractal Measure and their Singularities: The Characterization of Strange Sets*, Phys. Rev. A33, 1141.
- HENTSCHEL, H. G. E., and PROCACCIA, I. (1983), *The Infinite Number of Generalized Dimensions of Fractals and Strange Attractors*, Physica (Utrecht) 8D, 435.
- HUANG, K., *Statistical Mechanics* (J. Wiley, New York 1966).
- KAHANE, J. P., and PEYRIÈRE, J. (1976), *Sur Certaines Martingales de B. Mandelbrot*, Adv. in Math. 22, 131.
- MANDELBROT, B. B., *Possible refinement of the lognormal hypothesis concerning the distribution of energy dissipation in intermittent turbulence*, In *Statistical Models and Turbulence* (Lecture Note in Physics, Vol. 12), Proc. Symp. La Jolla, Calif. (eds. Rosenblatt, M., and Van Atta, C.) (Springer-Verlag, New York 1972) p. 333.
- MANDELBROT, B. B. (1974), *Intermittent Turbulence in Self Similar Cascades; Divergence of High Moments and Dimension of the Carrier*, J. Fluid Mech. 62, 331; also Comptes Rendus 278A, 289, 355.
- MANDELBROT, B. B., *The Fractal Geometry of Nature* (W. H. Freeman, New York 1982) pp. 373–381 discuss α and the "Lipshitz-Hölder heuristics," then the "nonlacunar fractals" = multifractals.
- MANDELBROT, B. B. (1984), *Fractals in Physics: Squig Clusters, Diffusions, Fractal Measures and the Unicity of Fractal Dimension* J. Stat. Phys. 34, 895.
- MANDELBROT, B. B. (1986), *Letter to the Editor: Multifractals and Fractals*, Physics Today 11.
- MANDELBROT, B. B., (1988), *An introduction to multifractal distribution functions*, In *Fluctuations and Pattern Formation* (Cargèse, 1988) (eds. Stanley, H. E., and Ostrowsky, N.) (Kluwer, Dordrecht-Boston 1988) pp. 345–360.
- MANDELBROT, B. B., *Examples of multinomial multifractal measures that have negative latent values for the dimension $f(\alpha)$* , In *Fractals ("Etteor Majorama" Centre for Scientific Culture, Special Seminar)* (ed. Pietronero, L.) (Plenum, New York 1989).
- MANDELBROT, B. B. (forthcoming), *Fractals and Multifractals: Noise, Turbulence and Galaxies* (Selecta, Vol. 1) (Springer, New York).
- MEAKIN, P., *The growth of fractal aggregates and their fractal measures*, In *Phase Transitions and Critical Phenomena* (eds. Domb, C., and Lebowitz, J. L.) (Academic Press, London 1988) 12, 335.
- MENEVEAU, C., and SREENIVASAN, K. R. (1987), *Simple Multifractal Cascade Model for Fully Developed Turbulence*, Phys. Rev. Lett. 59, 1424.
- MENEVEAU, C., and SREENIVASAN, K. R. (1989), *Measurement of $f(\alpha)$ from Scaling of Histograms, and Applications to Dynamic Systems and Fully Developed Turbulence*, Phys. Lett. A137, 103–112.
- PRASAD, R. P., MENEVEAU, C., and SREENIVASAN, K. R. (1988), *Multifractal Nature of the Dissipation Field of Passive Scalars in Fully Turbulent Flows*, Phys. Rev. Lett. 61, 74.
- VOLKMANN, (1958), *Ober Hausdorffsche Dimensionen Von Mengen, Die Durch Zifferneigenschaften Sind*, Math. Zeitschrift 68, 439.

(Received/accepted March 1, 1989)

The Fractal Geometry of Interfaces and the Multifractal Distribution of Dissipation in Fully Turbulent Flows

K. R. SREENIVASAN,¹ R. R. PRASAD,¹ C. MENEVEAU,¹ and R. RAMSHANKAR^{1,2}

Abstract—We describe scalar interfaces in turbulent flows *via* elementary notions from fractal geometry. It is shown by measurement that these interfaces possess a fractal dimension of 2.35 ± 0.05 in a variety of flows, and it is demonstrated that the uniqueness of this number is a consequence of the physical principle of Reynolds number similarity. Also, the spatial distribution of scalar and energy dissipation in physical space is shown to be multifractal. We compare the $f(\alpha)$ curves obtained from one- and two-dimensional cuts in several flows, and examine their value in describing features of turbulence in the three-dimensional physical space.

Key words: Fractals, multifractals, turbulent flows, interfaces, energy and scalar dissipation.

1. Introduction

One can define a variety of surfaces in turbulent flows. Some examples are the vorticity interface (that is, the conceptual surface separating domains of intense and zero vorticity fluctuations), iso-concentration surfaces (in reacting or nonreacting flows), iso-velocity surfaces, and iso-dissipation surfaces. A common property of these surfaces is that they are highly convoluted at many scales, and possess many shapes. Their complexity defeats attempts to describe them by means of classical geometry. MANDELBROT (in many papers and his 1982 book) has advanced the necessary framework for describing the geometry of such complex shapes; he also recognized that the self-similarity expected to hold in turbulence (according to the conventional wisdom succinctly described by RICHARDSON's (1922) rhyme) could permit fractal description of such surfaces. In Section 2 we demonstrate, using as an example the particular case of scalar surfaces (that is, surfaces marking the boundary of the scalar-marked regions in a turbulent flow), that the expectation is indeed valid. We also show in Section 3 that the experimentally measured value

¹ Mason Laboratory, Yale University, New Haven, CT 06520, U.S.A.

² Present Address: Department of Physics, Haverford College, Haverford, PA 19041, U.S.A.

of the fractal dimension can be deduced from the principle of Reynolds number similarity (that is, negligible dependence on viscosity of global properties such as the overall growth rates of turbulent flows). In Section 4, we summarize our earlier results regarding the multifractal distribution of dissipation, and interpret these results in the light of the more general multifractal formalism recently proposed by MANDELBROT (1988).

This paper overlaps with our earlier work (SREENIVASAN and MENEVEAU, 1986; MENEVEAU and SREENIVASAN, 1987a; SREENIVASAN *et al.*, 1989; PRASAD *et al.*, 1988—referred to respectively as I, II, III and IV below), but differs from those publications in two respects. First, some aspects related to different orientations of the intersecting planes used to measure fractal dimensions, dependence on thresholds, scaling ranges, etc., are described more fully here; some additional data on turbulent wakes are also provided. A second feature is that we revisit the interpretation of the observed multifractal features of turbulence obtained by low-dimensional intersections.

2. Experiments and Results on the Geometry of Interfaces

Figure 1 shows a thin longitudinal slice along the axis of a turbulent jet of water emerging from a well-contoured nozzle of circular cross-section into a tank of still water. The jet was made visible by mixing a small amount (of the order of 10 parts per million) of a fluorescing dye (sodium fluorescein) into the nozzle fluid, and exciting fluorescence by illuminating a thin section of the flow by a sheet of light. Care was taken to ensure that the fluorescence was not saturated. The light source was a pulsed Nd:YAG laser with a pulse width of about 8 ns (small enough to freeze the motion), and power density of up to $2 \times 10^7 \text{ Js}^{-1}$ per pulse; the light sheet had a thickness of the order of 200–250 μm , which is on the order of the estimated Kolmogorov scale (that is, the smallest *dynamical* scale in the flow). Not resolved here are even smaller scales in the dye concentration fluctuations, expected to occur because of the large Schmidt number, Sc , of the dye. (Schmidt number is the ratio of the kinematic viscosity ν of the fluid to the mass diffusivity of the scalar; in the present experiments its value is of the order of 10^3 .) It is legitimate to consider the plane intersection “mathematically thin” with respect to scales of the order of the Kolmogorov scale and larger. The visualized region, extending from 8 to 24 nozzle diameters downstream of the nozzle, was imaged on to a CCD camera with a 1300 (vertical) \times 1000 (horizontal) pixel array, yielding a pixel resolution on the order of $150 \mu\text{m}^2$. Further processing was done on a VAX station II/GPX.

The image shows a number of geometrically interesting features, one of which relates to the boundary that separates the nozzle fluid from the ambient tank fluid. The boundary is convoluted on a variety of scales, and appears to be disconnected at many places. It is possible that some out-of-plane connections may exist and that



Figure 1

A thin axial section of the nozzle fluid in an axisymmetric turbulent jet marked by a fluorescing dye. The nozzle Reynolds number is 4000.

the boundary is indeed connected; to establish this aspect properly, at least several simultaneous sections would be needed. Such measurements have now been completed and will be reported separately. In any case, one can imagine in three-dimensional space a surface that separates the nozzle fluid from the ambient tank fluid, a surface whose section by a plane is seen in Figure 1. This surface is of interest to us for many reasons, the primary one being that its geometry (which itself is a consequence of some dynamical constraints) will bear some relation to the amount and nature of mixing that occurs between the nozzle and tank fluids. For example, if the tank fluid were slightly acidic and the jet fluid slightly alkaline, the surface geometry will govern the amount and distribution of the product formed as a result of the reaction between the acid and the base.

As mentioned earlier, our objective is to characterize this surface (and in general all surfaces of interest in turbulent flows) by fractals; a primary property of a fractal surface being its fractal dimension, we want to measure it. We shall obtain the fractal dimension of the boundary seen in Figure 1, and later examine the sense

in which it relates to the fractal dimension of the surface embedded in three dimensions.

The first step is to specify how the boundary can be defined for further processing. Complex algorithms can be developed for the purpose, but we have shown elsewhere (see I and, PRASAD and SREENIVASAN, 1989) that it is adequate to use simple criteria based on the brightness threshold in the image (which is directly proportional to the concentration threshold on the nozzle fluid). Figure 2 shows the computer-drawn boundary obtained by setting the threshold at a brightness level that seems more or less satisfactory. One can now apply one of several techniques (described, for example, in Mandelbrot's book) to determine the fractal dimension of the boundary so marked. We have used both the box-counting and codimension methods. The codimension method was described in detail in I. In the box-counting methods, also briefly described in I, we cover the whole plane of



Figure 2

The boundary of the jet cross-section given in Figure 1, determined by prescribing a threshold on brightness (proportional to the concentration of the nozzle fluid). Also shown are typical (horizontal $a - a$ and vertical $b - b$) line intersections through the boundary; these will be discussed later in the text.

Figure 2 with square area elements of varying sizes, and count only the fraction $N(r)$ of elements containing the boundary, and plot $\log N(r)$ as a function of logarithm of the 'box' size r ; if the boundary is a fractal, we should expect an extensive straight part in this log-log plot, whose negative slope is the fractal dimension. A typical result (Figure 3a) shows that this is indeed the case, the straight part extending from the smallest scale resolved here to approximately a scale of the order of the nozzle diameter, giving a fractal dimension of 1.35 for the boundary. (It is worth mentioning that the computer programs for obtaining fractal dimensions have been checked extensively on several mathematically generated fractal sets of known dimension.)

Some minor ambiguities exist in defining the interface merely by means of a threshold, and so measurements have been repeated for several thresholds on a number of realizations of the jet. Figure 3b shows a plot of these results. It is clear that there exists a wide range of threshold values over which the fractal dimension of the boundary is essentially independent of the threshold, and that the mean value is 1.35. The spread of the data around this mean value is roughly in the range ± 0.05 . Figure 3c shows that the range of scale of similarity (that is, the range of scales over which the log-log plot has a straight part) varies somewhat with threshold, but is generally about 1.5 to 2 decades. This is approximately the ratio of the large scale L to the Kolmogorov scale η in the flow.

Similar experiments in various other flows have been conducted, and Figure 4 shows an axial section of the turbulent wake of a circular cylinder. The scalar boundary in each case was marked in a similar way, and fractal dimension was computed as before. The principal results are summarized in Table 1. All measurements were made by slicing flows with the plane of light sheet along the preferred direction of each flow. Without attaching significance to minor variations from one flow to another, we may conclude that a mean value is about 1.35. It would be useful, for later discussion, to take flow slices in different orientations. A typical result from such measurements is given in Figure 5 for the mixing layer between two streams of equal and opposite directions (RAMSHANKAR, 1988). It is seen that the boundaries of the flow in two orthogonal intersections possess the same fractal dimension. This conclusion also holds for all other flows investigated here.

We may now ask how the fractal dimension from planar intersections is related to the fractal dimension D of the surface itself—this being our major concern. This general problem has been discussed in the literature, and some specific results are available for special cases (see MARSTRAND, 1954, whose results have been generalized by MATTILA, 1975). The equivalent result in the present context, as stated by MANDELBROT (1982, p. 366), relates to the additive properties of codimensions of intersections. Specifically, if S_1 and S_2 are two *independent* sets embedded in a space of dimension d , and if codimension $(S_1) + \text{codimension } (S_2) < d$, the codimension of $S_1 \cap S_2$ is equal to the sum of the codimensions of S_1 and S_2 . For a fractal set F embedded in three-dimensional space and intersected by

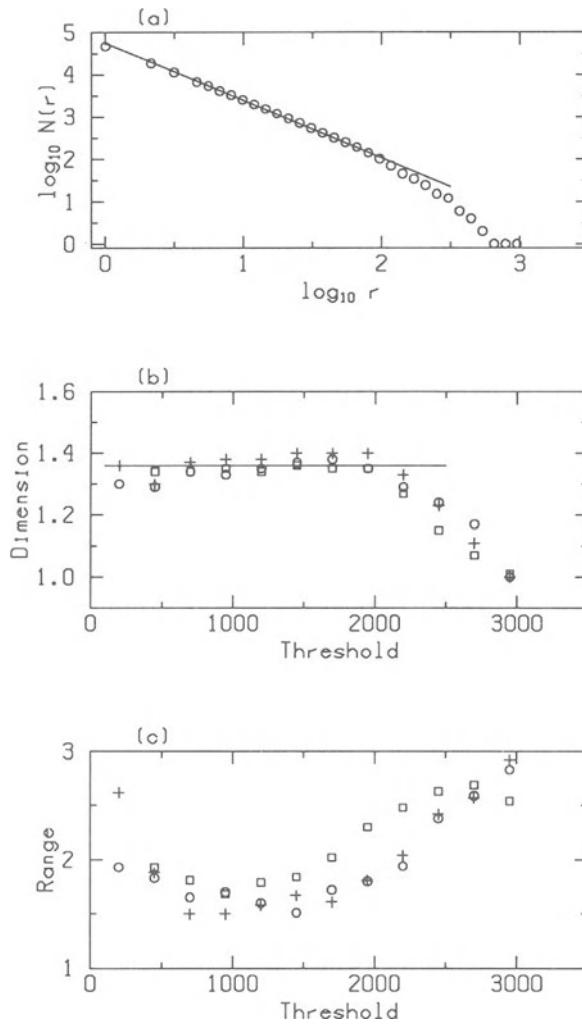


Figure 3

(a) The log-log plot of the number $N(r)$ of the area elements ('boxes') of size r containing the interface. The negative slope of the straight part gives the fractal dimension of the boundary ($= 1.35$). The abscissa ranging from 0 to 4096 in (b) is the dynamic range of the digital camera. Over a good fraction of this dynamic range, the measured fractal dimension is essentially independent of the threshold. Figure 3c shows the scaling range in decades for each threshold. For the Reynolds numbers typical of most experiments in present series (about 5000 based on the nozzle diameter and velocity), the typical large/small scale ratio is about 100. Different symbols in (b) and (c) correspond to different realizations.

a plane, the above statement implies that the dimension of the intersected set is one less than the dimension of F .

The results so far have shown that the fractal dimension of the boundary in longitudinal (that is, streamwise) as well as orthogonal sections of prototypical fluid

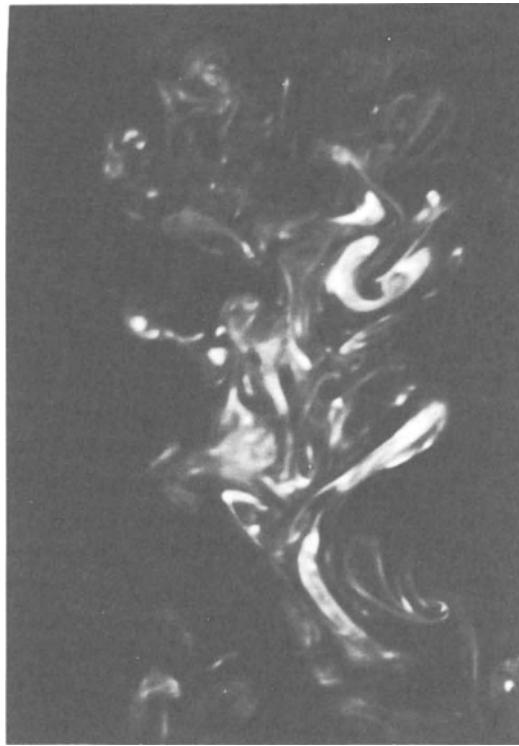


Figure 4

A thin axial section of the scalar-marked region in the wake of a circular cylinder. The cylinder Reynolds number is about 1500.

flows is 1.35. It then follows that the fractal dimension of the surface is one greater than 1.35, or $D = 2.35$.

It may be useful to expand briefly on the result that the fractal dimension of intersections is independent of the orientation of the intersecting plane. While this result can be expected intuitively to be valid for fractally isotropic objects, flows considered here do have a preferential direction. We should, however, emphasize two points: First, the possible anisotropic properties of the interface will be confined essentially to the largest scales in the flow, these being on the order of the flow width (and larger). Secondly, the smaller scales for which fractal-like behaviour has been found are expected to be more or less isotropic, thus explaining our observation. Although we have been unable to take simultaneous orthogonal sections, experiments with independent sections have shown that the anisotropy effect may determine the precise range of scale similarity in two orthogonal planes but not the fractal dimension itself.

The principle of additive codimensions also implies that the fractal dimension of line intersections is two less than that of the surface; again, we have an opportunity

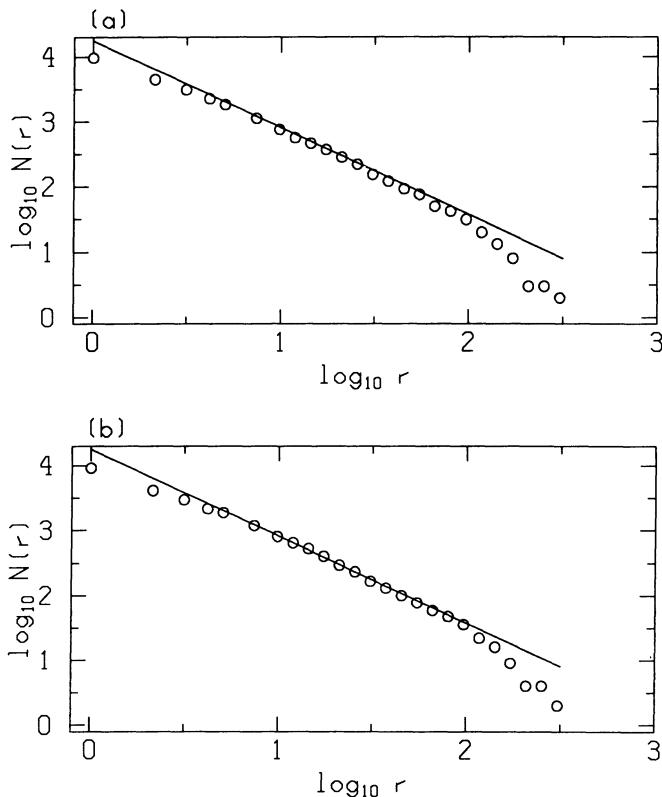


Figure 5

Results from the application of box counting methods to two orthogonal sections of a fully developed counter-current mixing layer at a moderate Reynolds number; (a) plan view; (b) side view. For both of these sections, the slope of the straight lines is nominally the same, giving the fractal dimension of the boundaries to be 1.34.

Table 1

Table summarizing the fractal dimension measurements of scalar interfaces in four classical turbulent flows. There are marginal differences among the results in these flows, but the variability from one realization to another does not warrant any significance to be attached to these differences. We may conclude that the mean value is about 1.35

Flow	Fractal dimension of plane intersection of interfaces
axisymmetric jet	1.36*
boundary layer	1.36**
mixing layer	1.34#
plane wake	1.35*

*Average over a number of realizations (of the order 20) covering a streamwise extent of 5 to 30 diameters from the nozzle.

**Single value in the outer region.

#Average over a number of realizations (of the order 10) covering approximately a streamwise extent of two large structures.

here for testing the independence of the fractal dimension on the orientation of the intersecting line. By using box counting methods as before, one can compute the fractal dimension of the set of discrete points corresponding to the intersection of the interface by a given line (see Figure 2 for examples). Figure 6 shows the measured fractal dimension of some representative intersections (horizontal as well as vertical) as a function of the location of the section. Figure 7 shows that the results from line-cuts of arbitrary orientation are also the same. The mean values of D are not far from 2.35. (We have already made use of the additive principle.)

Equivalently, if one assumes that the flow is frozen (TAYLOR, 1938), one can take line intersection by fixing a point probe in the Eulerian frame and letting the flow convect past it. There is no reason to think that this will be accurate at all scales, but the hypothesis has been known to work roughly in a variety of

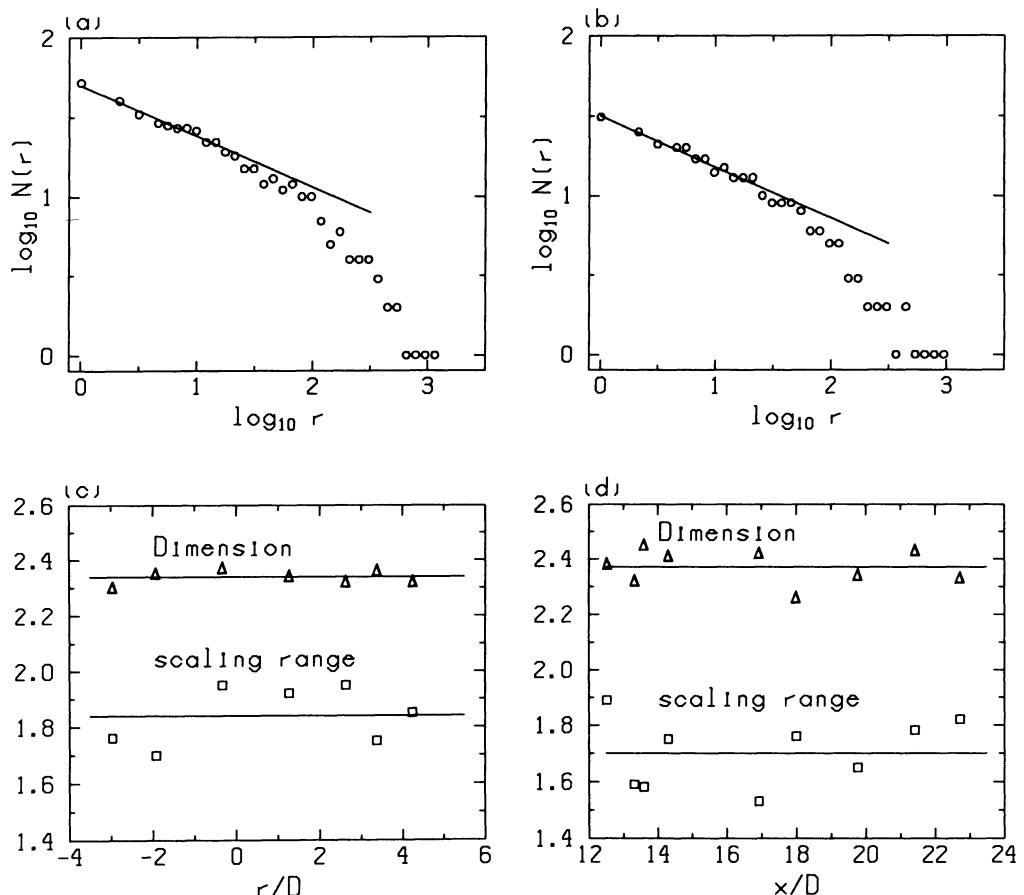


Figure 6

Typical results from one-dimensional cuts; (a) vertical and (b) horizontal, both referred to Figure 2. (c) and (d) indicate typical variability of the inferred D and the scaling range (in figures such as (a) and (b)) as a function of the position of the one-dimensional cut. The vertical cuts in (c) range from three diameters to the left of the axis to three diameters to the right. The horizontal cuts in (d) vary from 12 to 23 diameters downstream of the nozzle.

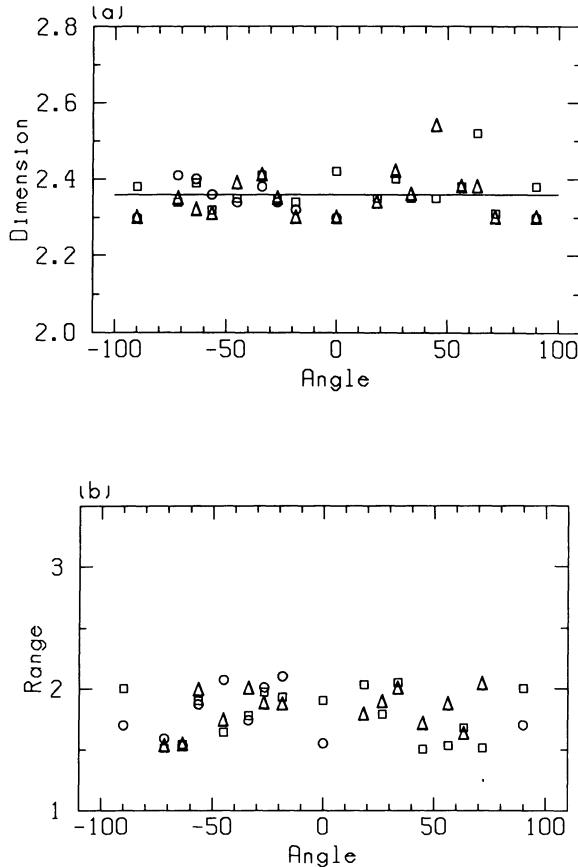


Figure 7

(a) Typical results from one-dimensional cuts passing through a fixed point as a function of the orientation of the intersecting line. Different symbols correspond to line cuts passing through different points. Figure (b) indicates the observed scaling range.

circumstances. We have used this approach also, and obtained results (Figure 8) for the nozzle fluid interface in a heated jet. Again, the inferred interface dimension has a mean value of 2.35, with the standard deviation on either side of about 0.05.

The last set of results obtained by the use of Taylor's frozen flow hypothesis are important because they enable us to conclude that the fractal dimension of scalar interfaces is independent of the mass diffusivity of the scalar: The dye and heat diffuse at rates which differ by about three orders of magnitude. Furthermore, the technique allows simple measurement of the fractal dimension of the vorticity interface as well. This has been discussed in I, where it was shown that the fractal dimension of this surface is also about 2.35.

We now have a very general result. To within the accuracy of measurement, the fractal dimension of vorticity and scalar interfaces (considering only scales

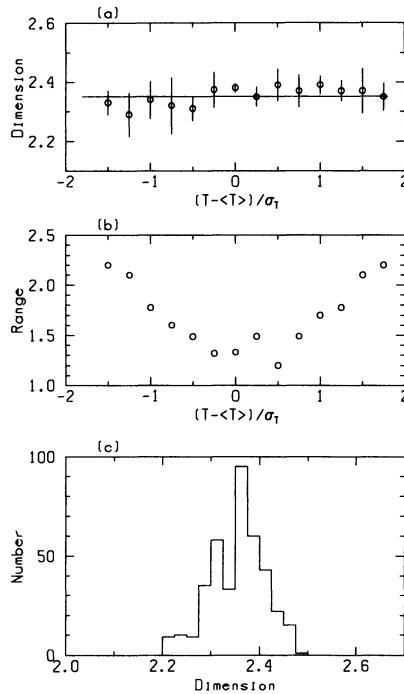


Figure 8

Fractal dimensions obtained by one-dimensional cuts using Taylor's frozen flow hypothesis in heated jets. (a) illustrates the constancy of the dimension with respect to the threshold, while (b) indicates the scaling range. In both (a) and (b), $\langle T \rangle$ is the mean temperature at the location of the measuring probe (a cold wire operated at 0.1 mA in the constant current mode), and σ is the root-mean-square fluctuation. Figure (c) is the histogram of D constructed using 390 realizations; the mean value is 2.35.

above η) in a number of fully developed turbulent flows, obtained by plane and line intersection techniques (whose validity we have established by a variety of measurements), is about 2.35. Clearly, the generality of the result demands an interpretation based on broad considerations. This is attempted in the next section.

3. Mixing and Reynolds Number Similarity

Let us consider transport by diffusion across interfaces of the type discussed so far. The flux is given by the product of the surface area, the concentration gradient normal to the surface and the molecular diffusivity. For fractals, the surface area S increases with the resolution of measurement r according to the relation (MANDELBROT, 1982)

$$S \sim r^{2-D}. \quad (1)$$

In all practical circumstances, the scale range over which (1) holds is bounded by cut-offs on both ends. For surfaces in turbulent flows, the outer cut-off is expected to occur at scales comparable to the integral scale, L , of turbulence, while the inner cut-off occurs at the smallest dynamical scale. For vorticity interfaces, the appropriate inner scale is the Kolmogorov scale $\eta = (v^3 \langle \epsilon \rangle)^{1/4}$, where $\langle \epsilon \rangle$ is the average rate of the turbulent energy dissipation. For with Schmidt number greater than unity, the relevant inner cut-off occurs at the Batchelor scale, $\eta_b = \eta Sc^{-1/2}$. The existence of a finite inner cut-off means that, as the surface area gets measured by covering it with increasingly finer area elements, a point is reached at which convolutions of even finer scales no longer exist, so that, thereafter, the area does not increase with increasing fineness of resolution; instead, it will saturate (abruptly in an ideal situation) at the maximum value corresponding essentially to the inner cut-off. The true area S_T of a fractal surface with finite inner cut-off is thus given (to within a constant) by the knowledge of the fractal dimension, and the inner cut-off r_i which theoretically truncates the power-law behaviour. Thus,

$$S_T = S_0(r_i/L)^{2-D} \quad (2)$$

where S_0 is some normalizing area. If the area levels off at L and beyond, S_0 becomes the surface area measured with the resolution equal to L . It was also shown in III that for the case of high Schmidt numbers, it is natural to assume (in analogy with the inertial and viscous-convective ranges in the power-spectral density of passive scalar fluctuations) that there are two scaling regimes for the area S . In the range $\eta < r < L$, the relevant exponent is D , whereas in the range $\eta_b < r < \eta$, the relevant exponent will be designated D' and equation (2) is modified accordingly (see III).

It was shown in III that the characteristic velocity and concentration gradient across interfaces are of the order U_c/η and $\Delta C/\eta_b$, respectively. U_c is a characteristic velocity, for example the centerline defect velocity for the wakes, the centerline excess velocity for the jets, the velocity difference between the two streams for the mixing layer, and the friction velocity (equal in kinematic units to the square root of the wall shear stress) for the boundary layer. ΔC is a characteristic concentration difference.

From these considerations, an expression for the flux of momentum across the interface can be written as

$$vS_T(U_c/\eta). \quad (3)$$

Defining the characteristic Reynolds number $Re = u'L/v$ (u' being the root-mean-square of the velocity fluctuations), we may note that $\eta/L \sim Re^{-3/4}$, and use equation (2) for the interface area S_T to write (after a little algebra) that the

$$\text{flux of momentum} \sim S_0 U_c^2 (u'/U_c) Re^{3(D - 7/3)/4}. \quad (4)$$

Note that S_0 , U_c and (u'/U_c) are all independent of Reynolds number. For nonunity Schmidt numbers, the corresponding result for the flux of a species with concentration difference ΔC is given by

$$\text{flux of contaminant} \sim S_0(U_c \cdot \Delta C)(u'/U_c) \text{Re}^{3(D-7/3)/4} Sc^\gamma, \quad (5)$$

where $\gamma = 0.5 (D' - 3)$ for $Sc > 1$ and $\gamma = 3 (D - 7/3)/4$ for $Sc < 1$.

Now, it is well-known that all fluxes (mass, momentum, energy) must be independent of Reynolds number in fully turbulent flows—the so-called Reynolds number similarity. (This is merely a statement of the observed fact that the growth rates of turbulent flows of a given configuration are independent of fluid viscosity.) According to (4) and (5) the Reynolds number similarity requires that

$$D = 7/3, \quad (6)$$

for both the vorticity and scalar interfaces, in rough agreement with experiments (see Table 1). Since the fluxes are expected to be also independent of Sc (the so-called Schmidt number similarity), equation (5) implies that $\gamma = 0$, or $D' = 3$. This means that the convolutions of the interface on scales between η_b and η are essentially space-filling.

In the above arguments we have assumed that it is appropriate to use a common characteristic velocity or concentration gradient everywhere along the interface. This is not strictly true, at least because the interface thickness varies from place to place because of the spatial intermittency of the dissipation rate ε (see Section 4). Furthermore, it is implied above that the globally averaged dissipation rate is the same as that averaged in the neighborhood of the interface alone. These two limitations were addressed in detail in III, where it was shown that the intermittent nature of the dissipation near the interface is statistically the same as that elsewhere, and that the inclusion of the intermittency will alter D from $7/3$ to 2.36 . This latter estimate is in excellent agreement with our experimentally determined mean value of 2.35 . It is worth mentioning that the reason for the relatively small correction is that the interface thickness depends on the quarter power of the dissipation, and so the strong variabilities in ε do not translate to comparable variations in the interface thickness.

4. The Multifractal Distributions of Dissipation Rates of Turbulent Kinetic Energy and Passive Scalar Fluctuations

There has been an explosive interest in recent years in the characterization of measures created by multiplicative processes. Starting with MANDELBROT (1974), several formalisms have been introduced (HENTSCHEL and PROCACCIA, 1983; FRISCH and PARISI, 1985; HALSEY *et al.*, 1986; MANDELBROT, 1988). In turbulence,

the transfer of kinetic energy from the large scales of motion to the smaller scales can be thought of as arising from a multiplicative cascade process. Therefore, the manifestation of this flux at the smallest scales, which is the dissipation ε , is expected to be a multifractal. The same can be said for the flux of the variance of a passive scalar and its dissipation χ . In such a description, the total dissipation of kinetic energy E_r that is contained in a box of size r obeys a local power law

$$E_r \sim r^\alpha, \quad (7)$$

where α varies from point to point. A similar power-law is expected to hold for X_r , the amount of χ in a box of size r . Since the values of α for the scalars need not be the same as those for the energy, the notation in (7) is to be regarded as generic. The number of boxes with α within in a band $d\alpha$ is assumed to scale according to

$$N(\alpha) d\alpha \sim r^{-f(\alpha)} d\alpha. \quad (8)$$

As one covers the measure with boxes of decreasing size, the number of occurrences of a certain α generally increases (that is, $f(\alpha) > 0$), or remains constant (that is, $f(\alpha) = 0$). In such cases, $f(\alpha)$ is interpreted as the fractal dimension of iso- α sets. (For a more detailed discussion of this point, see CHhabra *et al.*, 1989.) MANDELBROT (1988) has shown that cases exist for which $f(\alpha)$ is negative; we will return to this point below.

A dual description of multifractals is given in terms of the q -moments of E_r and X_r ; in that case one measures the ‘generalized box dimensions’ (HENTSCHEL and PROCACCIA, 1983), $D_q(\varepsilon)$ and $D_q(\chi)$ of ε and χ respectively, defined as

$$\Sigma E_r \sim r^{(q-1)D_q(\varepsilon)}, \quad \text{and} \quad \Sigma X_r \sim r^{(q-1)D_q(\chi)}. \quad (9)$$

The quantities $f(\alpha)$ and α are then obtained by the well-known Legendre transforms (HALSEY *et al.*, 1986)

$$\alpha = d/dq[(q-1)D_q] \quad \text{and} \quad f(\alpha) = q\alpha - (q-1)D_q. \quad (10)$$

Although $f(\alpha)$ can be determined directly, MENEVEAU and SREENIVASAN (1989) have shown that it is more practical to obtain them by first measuring the D_q exponents and using (10). This is due to second-order finite-size corrections to $N(\alpha)$ that appear in (8), which are important in experimental situations at moderate Reynolds numbers.

Measurements of the D_q curve from one-dimensional cuts through ε were reported in II, and the corresponding curves for the dissipation of passive scalar fluctuations from one- and two-dimensional intersections were reported in IV. The $f(\alpha)$ curves obtained by the use of (10) are shown in Figures 9a,b. (For readers

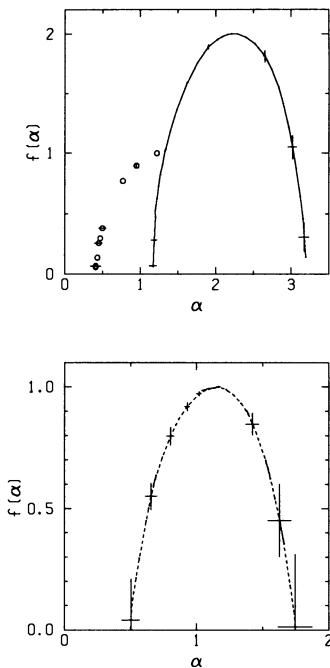


Figure 9

$f(\alpha)$ curves of lower-dimensional sections through the dissipation fields. (a) Dissipation of passive scalar fluctuations: The continuous line represents the mean results from two-dimensional measurements of concentration (a passive scalar) in water jets (see IV). Circles correspond to one-dimensional measurements of temperature (also a passive scalar) in heated air jets. Since there are some experimental difficulties associated with the determination of the D_q 's for $q < 0$ in the one-dimensional sections (see IV), only the left half of the curve is shown. (b) Dissipation of kinetic energy: The dashed line corresponds to the average $f(\alpha)$ curve for one-dimensional sections in several fully developed turbulent flows (see II). In both figures, bars show the variability in measurements.

familiar with our earlier work, it is useful to note a change of notation. In II and IV, α was defined in terms of the average dissipation on a domain r according to $\varepsilon_r \sim r^{\alpha-1}$ for any dimensionality of the box. Here, we use equation (7) as the basic definition and therefore $\varepsilon_r \sim r^{\alpha-d}$, where d is the dimensionality of the domain in which the dissipation is embedded: d is 1 for one-dimensional cuts, 2 for planar intersections and 3 for the case in three dimensions.)

Several properties of the lower-dimensional intersections can now be discussed. We shall assume here, and defer questions of exceptions to a later stage, that the additive properties of fractal dimensions discussed in Section 2 hold also for $f(\alpha)$; that is, $f(\alpha)$ in three-dimensional space can be obtained by adding 2 to that obtained by linear cuts and by adding 1 to the results from planar cuts. Similar additive properties hold for α , since lower-dimensional cuts measure densities that can be

integrated over boxes of different dimensionalities (for details see MANDELBROT, 1988). Figure 10 shows the $f(\alpha)$ curves of ε and χ obtained by translating according to the above additive law the values of α and $f(\alpha)$ of Figure 9. It is apparent that the $f(\alpha)$ curve for χ is wider than that for ε ; as discussed in IV, this is consistent with the well-known result that the dissipation of passive scalar is more intermittent than the energy dissipation.

The $f(\alpha)$ curve of ε obtained for one-dimensional cuts was modeled in MENEVEAU and SREENIVASAN (1987b) by a two-scale Cantor measure (the p -model). In order to reproduce the experimental observations, the following three-dimensional cascade process was proposed: After every step of the cascade, each eddy splits into 2^3 smaller eddies. Half of them each receive a fraction $p_1/4 = 0.7/4$ of the energy flux, while each of the others receive the remaining fraction $p_2/4 = 0.3/4$. As pointed out in SREENIVASAN and MENEVEAU (1988), this would imply that there are no singularities with $f(\alpha) < 2$, but this can be corrected by perturbing the p -model slightly and assuming that each eddy receives a fraction $p_1/4 \pm \delta$ or $p_2/4 \pm \delta$, where

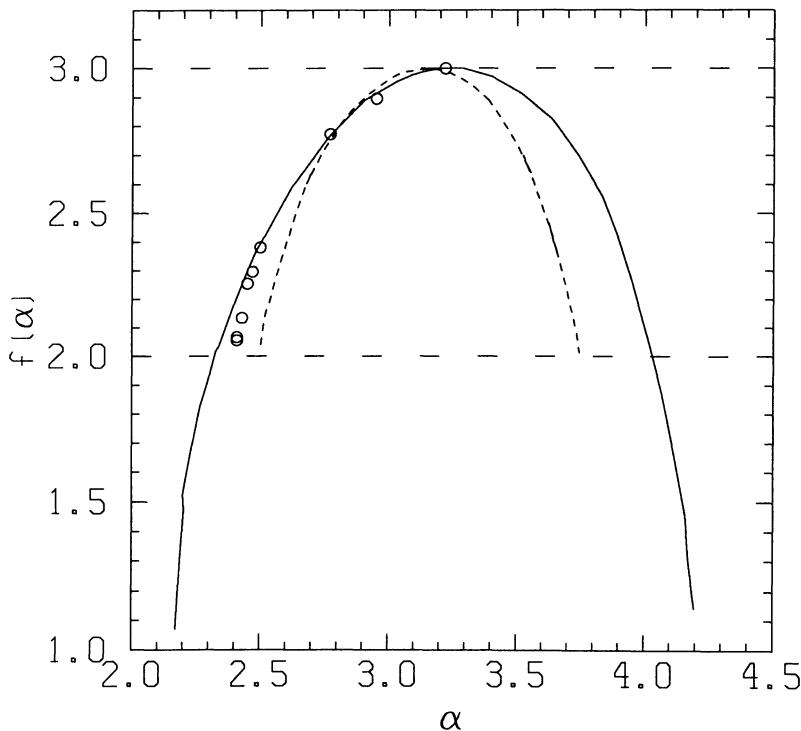


Figure 10

$f(\alpha)$ curves of lower-dimensional sections through the dissipation fields, translated in such a way as to correspond to the situation in three-dimensional space. Symbols as in Figure 9. Where they overlap, the passive scalar data obtained from one and two-dimensional sections agree well. The figure also shows that the $f(\alpha)$ curve for the scalar dissipation curve is wider than that for the energy dissipation.

δ is different for each eddy (MENEVEAU, 1989). Obviously, these perturbations are subject to the overall conservation of the energy flux.

We now discuss an issue raised by MANDELBROT (1988). By necessity, the discussion here is very brief, and will be expanded elsewhere using appropriate experimental evidence. The issue is the use of lower-dimensional cuts to obtain information on *very sparse* sets of singularities, whose $f(\alpha)$ in three dimensions is less than the co-dimension of the intersecting subspace. MANDELBROT (1988) has shown that this is in principle a valid procedure, provided that averages are performed over regions that are much larger than the integral scale of the flow. He has further shown (see also MANDELBROT, 1974) that if the energy cascade proceeds in such a way that a sub-eddy receives an amount of energy that is higher than a certain fraction, the implication is that lower-dimensional cuts would show negative values of α and D_q 's for $q > q_{cr}$. It follows that a possible criticism of the measurements reported above is that, since only the last step of the cascade is accessible in measurements, negative values of α and D_q (if they exist) are impossible to detect, and therefore D_q 's for $q > q_{cr}$ may have been biased.

If one extrapolates the curves of Figure 9a down to $f(\alpha) = 0$, it appears that α is always larger than 2 (especially for the energy dissipation), suggesting that no negative α 's are present even in one-dimensional cuts. The question now concerns the reliability of the $f(\alpha)$ curves in Figure 9 especially near the tails (which are determined by high order D_q 's—both positive and negative), as well as the extrapolation procedure. We point out that the positive part of the $f(\alpha)$ curve from lower-dimensional cuts can be obtained more or less completely using values of q up to about 6. An indication that D_q 's for q values of that order are quite accurate comes from the agreement of our results with those inferred from other experiments (ANSELMET *et al.*, 1984) in that range of q . From this and further evidence to be reported in the near future, we conclude that D_q 's at least up to order 6 are not biased, and, therefore, that there are no negative α 's. If this is so, then the cascade proceeds without ever surpassing the limit on the fluxes mentioned above, this being a very strong statement about the physics of the cascade. One can get an idea of this limit by using the three-dimensional p -model, for which this limit is 1/4.

At any rate, the quantification of a multiplicative measure by $f(\alpha)$ is degenerate at many levels (FEIGENBAUM, 1987; MANDELBROT, 1988; CHHABRA *et al.*, 1989), and it is not clear how much of the underlying physics can be extracted unambiguously from such a description.

Acknowledgements

We are thankful for financial support to the Air Force Office of Scientific Research (which supported the initial phases of this work) and the Defence Advanced Research Projects Agency (which has sustained the later phases).

REFERENCES

- ANSELMET, F., GAGNE, Y., HOPFINGER, E. J., and ANTONIA, R. A. (1984), *High-order Velocity Structure Functions in Turbulent Shear Flows*, J. Fluid Mech. 140, 63.
- CHHABRA, A., JENSEN, R., and SREENIVASAN, K. R. (1989), *Multifractals, Multiplicative Processes and the Thermodynamic Formalism*, Phys. Rev. A (in print).
- FEIGENBAUM, M. (1987), *Some Characterizations of Strange Sets*, J. Stat. Phys. 46, 919.
- FRISCH, U., and PARISI, G., *On the singularity structure of fully developed turbulence*, In *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics* (eds. Ghil, M., Benzi, R., and Parisi, G.) (North-Holland, New York 1985).
- HALSEY, T. C., JENSEN, M. H., KADANOFF, L. P., PROCACCIA, I., and SHRAIMAN, B. I. (1986), *Fractal Measures and their Singularities: The Characterization of Strange Sets*, Phys. Rev. A33, 1141.
- HENTSCHEL, H. G. E., and PROCACCIA, I. (1983), *The Infinite Number of Generalized Dimensions of Fractals and Strange Attractors*, Physica 8D, 435.
- MANDELBROT, B. B. (1974), *Intermittent Turbulence in Self-similar Cascades: Divergence of High Moments and Dimension of the Carrier*, J. Fluid Mech. 62, 331.
- MANDELBROT, B. B., *The Fractal Geometry of Nature* (Freeman, San Francisco 1982).
- MANDELBROT, B. B., *An introduction to multifractal distribution functions*, In *Fluctuations and Pattern Formation* (eds. Stanley, H. E., and Ostrowsky, N.) (Kluwer, Dordrecht-Boston 1988); see also this volume.
- MARSTRAND, J. M. (1954), *Some Fundamental Geometrical Properties of Plane Sets of Fractal Dimensions*, London Math. Soc. 3, 257.
- MATTILA, P. (1975), *Hausdorff Dimension, Orthogonal Projections and Intersections with Planes*, Ann. Acad. Sci. Fenn. Ser. A I Math. 1, 227.
- MENEVEAU, C. (1989), *The Multifractal Nature of Turbulence*, Ph.D. Thesis, Yale University.
- MENEVEAU, C., and SREENIVASAN, K. R. (1987a), *The Multifractal Dissipation Field in Turbulent Flows*, Nuclear Physics B (Proc. Suppl.) 2, 49.
- MENEVEAU, C., and SREENIVASAN, K. R. (1987b), *Simple Multifractal Cascade Model for Fully Developed Turbulence*, Phys. Rev. Lett. 59, 1424.
- MENEVEAU, C., and SREENIVASAN, K. R. (1989), *Measurement of $f(\alpha)$ from Scaling of Histograms, and Application to Dynamical Systems and Fully Developed Turbulence*, Phys. Lett. A (in print).
- PRASAD, R. R., MENEVEAU, C., and SREENIVASAN, K. R. (1988), *Multifractal Nature of the Dissipation Field of Passive Scalars in Fully Turbulent Flows*, Phys. Rev. Lett. 61, 74.
- PRASAD, R. R., and SREENIVASAN, K. R. (1989), *Scalar Interfaces in Digital Images of Turbulent Flows*, Experiments in Fluids 7, 259.
- RAMSHANKAR, R. (1988), *The Dynamics of Countercurrent Mixing Layers*, Ph.D. Thesis, Yale University.
- RICHARDSON, L. F., *Weather Prediction by Numerical Process* (Cambridge University Press, Cambridge, U.K. 1922).
- SREENIVASAN, K. R., and MENEVEAU, C. (1986), *The Fractal Facets of Turbulence*, J. Fluid Mech. 173, 357.
- SREENIVASAN, K. R., and MENEVEAU, C. (1988), *Singularities of the Equations of Fluid Motion*, Phys. Rev. A38, 6287.
- SREENIVASAN, K. R., RAMSHANKAR, R., and MENEVEAU, C. (1989), *Mixing, Entrainment, and Fractal Dimension of Interfaces in Turbulent Flows*, Proc. Roy. Soc. Lond. A421, 79.
- TAYLOR, G. I. (1938), *The Spectrum of Turbulence*, Proc. Roy. Soc. Lond. A164, 476.

(Received September 14, 1987, revised/accepted June 1, 1988)

Fractal Velocity Models in Refraction Seismology

DAVID J. CROSSLEY¹ and OLIVER G. JENSEN¹

Abstract—The introduction of noise components to a simple crustal velocity model is shown to markedly affect the appearance of synthetic seismograms calculated according to ray theory applied to refraction experiments. Here we simulate noise by a self-similar process with a power spectral density which falls off as inverse wavenumber to a simple power (0–2). The major effect is to destroy the coherency of the arrival branches normally expected from deterministic velocity models; the arrival amplitudes also show large trace-to-trace variations and considerable sensitivity to shot position. Some of these differences can be ascribed to the variety of noise model chosen (*i.e.*, white noise, flicker noise and brown noise). It is argued that there is no clear distinction between coherent noise and geological structure.

Key words: Refraction, fractal, raytracing, scattering, seismology.

1. Introduction

Small fluctuations in crustal velocities have been recognised as influencing seismic travel times and amplitudes for many decades. Until recently however, the effect of such noise has not been modelled explicitly but regarded (together with seismometer noise) as an interpretational nuisance. From a signal analysis perspective, velocity fluctuations can be regarded as convolutional, geological (and therefore potentially useful) noise. By contrast, seismometer noise (wind motion, etc.) is additive and of no interpretive use.

The detailed interaction of wave motion with stochastic media is a complicated subject, as demonstrated by HUDSON (1982), WENZEL (1982), SATO (1982), LERCHE (1985) and WU and AKI (1985). At the smallest distance scale, it has been proposed that a fractal-like roughness of the medium leads to a modification of the traditional smooth wave propagation mechanism (BERRY, 1979). A comprehensive treatment of scattering due to media irregularities has been recently given by FRANKEL and CLAYTON (1986) who used a finite difference approximation in

¹ Geophysics Laboratory, Department of Geological Sciences, McGill University, 3450 University Street, Montreal, Quebec, Canada H3A 2A7.

calculating synthetic seismograms for reflection surveys. For the analysis of large-scale refraction surveys, with which we are concerned here, some of the major effects of stochastic velocity functions can be studied using the traditional, asymptotic wave theory approach in two-dimensional media.

Using ray theory, MEREU and OJO (1981) showed that small random velocity fluctuations affect the travel times of simple reflected and refracted waves. They used a velocity model in which a white noise component is smoothed by a spatial filter to produce a medium 'tuned' to a specific correlation length. The result was an overall increase in the arrival times so that a least-squares fit overestimates the depth of discontinuities. Moreover, the smooth traveltimes branches are broken into several segments such that spurious structure can appear that is not present in the deterministic model. In a second paper OJO and MEREU (1986) argued that amplitudes would be affected by rays coalescing at some surface locations and avoiding others, thus producing seismograms similar to those observed in actual surveys.

In this paper we extend Mereu and Ojo's approach in two ways. Firstly we consider media with approximately fractal properties, *i.e.* correlations at all length scales (within the numerical simulation). Secondly we construct 'true' synthetic seismograms using the program RAYAMP by SPENCE *et al.* (1984) which enables us to see more clearly the multiplicity of arrivals produced by the media scattering. It should be noted that we are concerned only with first-order effects in the deviations in ray paths produced by the inhomogeneities. We ignore the backscattering (or multiple reflections) and P to S conversions produced by irregularities, though the inclusion of these would undoubtedly make the seismograms appear more realistic (as would the addition of simulated additive noise).

Our purpose in using RAYAMP was to profit from its relative ease of use, availability and speed and because this program, or similar programs using asymptotic ray theory, are widely used to interpret refraction surveys. We have not modified the program to take care of caustics and critical rays. We believe (based on what we know of the performance of this program for more conventional velocity structures) that the overall appearance of our seismograms would not be altered by the minor amplitude changes resulting from the use of a more accurate algorithm, *e.g.* the reflectivity method, though we cannot of course prove this *a priori*. On the other hand RAYAMP has been extensively redesigned for interactive microcomputer use. Many modifications have been made to speed up the raytracing and synthetics necessary in dealing with the large number of boundaries required by a two-dimensional grid of velocities.

2. Noise Models

We consider three flavours of self-similar processes, all derived from a one-dimensional white random pseudo-Gaussian process. The procedure for constructing

the noise is straightforward and begins with a finite two-dimensional array of uncorrelated Gaussian random numbers $w(i,j)$. This array is transformed to $W(k_1, k_2)$ in the spectral domain using an FFT algorithm where it is filtered using the process

$$F(k_1, k_2) = W(k_1, k_2)/k^{(n/2)}, \quad k > 0 \quad (1)$$

where k is the radial wavenumber $(k_1^2 + k_2^2)^{(1/2)}$. This is equivalent to dividing the Fourier spectrum of $w(i,j)$ by the factor $1/k^n$. Then the amplitudes of the Fourier coefficients F are scaled to maintain the same total power as for the original data $W(k_1, k_2)$. Finally we inverse-transform F to provide a new data set $f_n(i,j)$ whose properties depends on the index n in (1). For $n = 0$ we have the original data returned (*i.e.*, our original white noise sample), for $n = 1$ we have approximate flicker noise and for $n = 2$ we have approximate brown noise. The condition of these noises as ‘approximate’ arises from the fact that they are essentially band-limited between the period of the array (*i.e.*, the domain of the data) and the Nyquist period determined by the sample interval.

By the above procedure the constructed noise $f_n(i,j)$ will be stationary with a finite, predefined mean and variance. Properly a band-limited white noise ($n = 0$) process has a finite variance, upper band-limited flicker noise ($n = 1$), a log-infinite variance and upper band-limited brown noise ($n = 2$) an infinite variance and indeterminate mean. In such cases an extended notion of stationarity of these processes, related to the finiteness of their delta variances (MANDELBROT, 1982), may be assumed. Because our procedure (1) begins with the $k = 1$ spectral sample, thus leaving the $k = 0$ value (data mean) untouched, it automatically yields lower band-limited noise samples and we do not require this extended assumption of stationarity. We claim that our procedure constructs approximate noise samples which could (most) reasonably arise as samples of the more properly defined (*i.e.*, no lower band-limited) flicker and brown noise process (for one-dimensional examples of the latter, see *e.g.*, CROSSLEY *et al.*, 1986).

The flicker noise $f_1(i,j)$ and brown noise $f_2(i,j)$ functions represent a class of self-similar processes that have fractal character and autocorrelation functions in the space domain that fall smoothly with distance from the origin. We therefore expect the noise models to show increasing spatial correlation with $n > 0$ for length scales from the cell size up to the array size.

Examples of the data sets produced are shown in Figure 1 for grid sizes of (256×64) . As noted above the data shows stronger correlation as n increases. Also shown in Figure 2 are the power spectral maps corresponding to the flicker and brown noise data sets. These show the concentration of low-wavenumber energy as n increases. It is worth acknowledging that the maps in Figure 1 for $n > 1$ also show the folding effect of the FFT (*i.e.*, top and bottom edges wrap around). Though undesirable (and removable) in normal spectral processing, this feature was not considered a serious drawback for the present investigation.

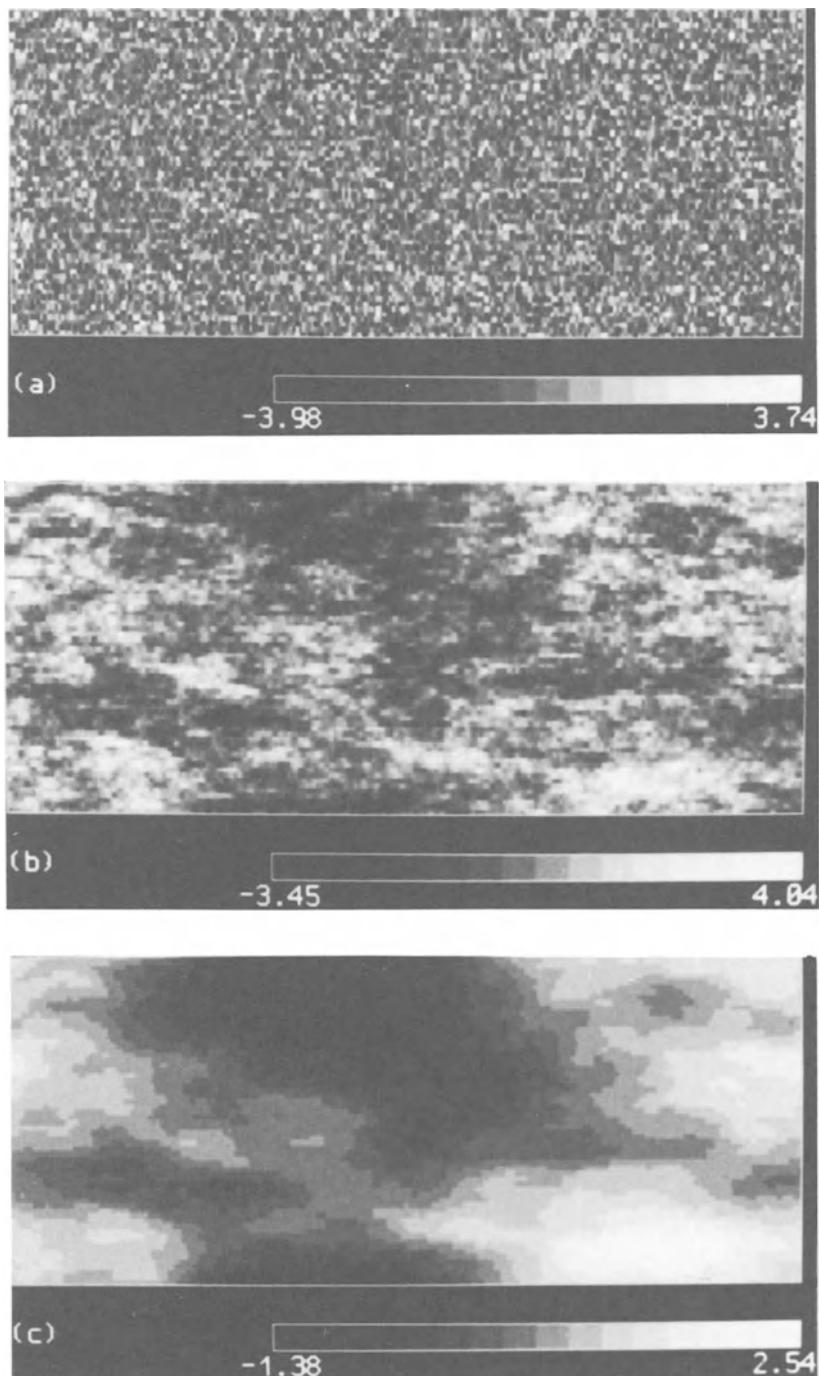


Figure 1

Three varieties of self-similar noise on a 256×64 grid—(a) white noise, (b) flicker noise and (c) brown noise.

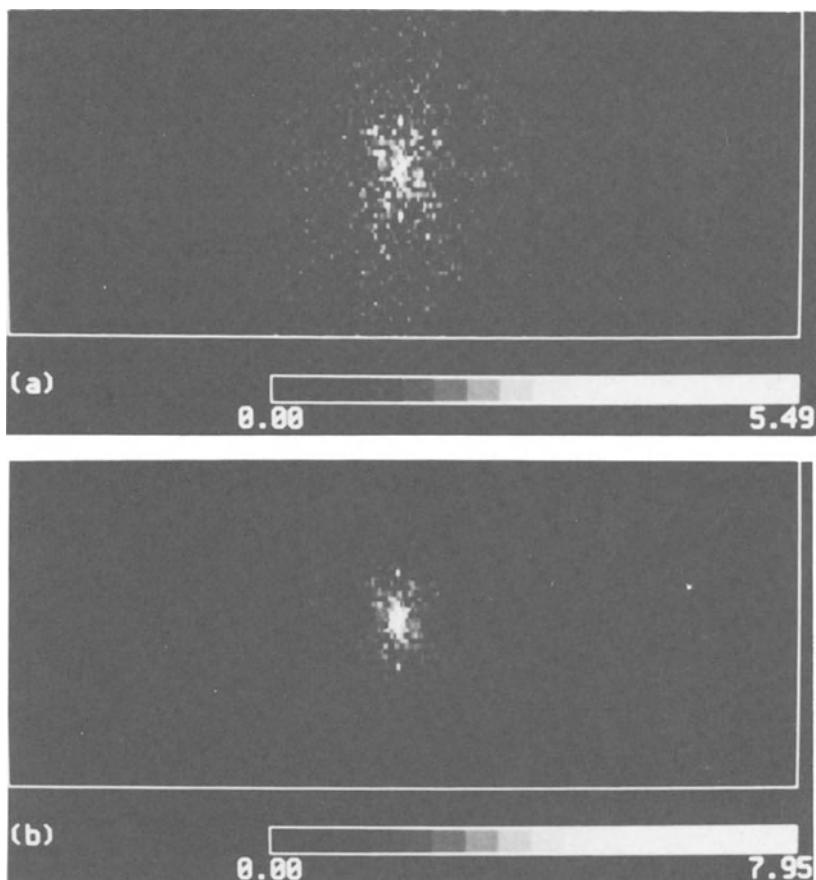


Figure 2

Power spectral estimates of two of the noise processes in Figure 1 for (a) flicker noise and (b) brown noise. Zero frequency is at the centre of each diagram and the map limits are \pm the Nyquist wavenumbers. (The white noise spectrum is not shown as it is virtually identical to Figure 1(a).)

3. Velocity Models

Since the pioneering work by MANDELBROT (1982) describing the many instances where nature can be represented by self-similar mathematical constructs, it has seemed virtually self-evident that fractals are ubiquitous in the physical world. Over the last decade numerous papers have covered the instances where, in the Earth sciences, fractal (*i.e.*, self-similar) data have been recognised. It seems entirely reasonable to extend geophysical models to the crust and upper mantle where seismic reflection surveys indicate a great range of acoustic reflectivity. One of the first studies of actual well-log data with fractals as a model was given by HOSKEN (1980).

If one were to consider casually which of the three noise data sets in Figure 1 best represents a notion of crustal velocity variability, then either of the latter two ($n = 1$ or $n = 2$) could be argued. As a representation of, for example, surface magnetisation, a geophysicist would probably pick $n = 2$ as closer to measured data (i.e., geology is more like brown noise than flicker noise).

None of the Figure 1 models are appropriate for the crust, however, as they ignore the overall increase in velocity which happens due to compaction of lithologies and fractures. Therefore a further processing was necessary to produce the velocity structures used for raytracing. We construct a base velocity model, qualitatively similar to that used by OJO and MEREU (1986), in which the velocity increases sharply from the surface to a depth of 3 km, followed by a more gradual increase to the 'Moho' at 33 km, then a sharp increase to the top of the mantle at 36 km, followed finally by a gradual increase in the upper mantle. The velocity depth function of this base model is clear in Figures 5 and 9.

To form a realistic 2-D velocity model we add appropriate noise to this base model. Consider the noise to represent a 2-D grid of velocity cells perpendicular to the surface. The velocity at any location P (horizontal distance x along, and depth z below, the surface) is calculated according to the relation

$$v(x, z) = v_0(j) + (z - z_j)g_j + af_n(i, j) \quad (2)$$

where

$v_0(j)$ is the velocity of the j -th boundary in the base model above P

z_j is the depth of the j -th boundary above P

g_j is the gradient at the top of the j -th boundary above P

a is a scalar amplitude constant

$f_n(i, j)$ is the value of noise in cell (i, j) of the model.

This model ensures that each cell has the gradient appropriate to the base model and so, as $a \rightarrow 0$, the model $v(x, z)$ approaches the base model smoothly. In all the models considered, a noise standard deviation (scalar a in (2)) of 0.1 km/sec was chosen.

Due to limitations in the computer memory (all calculations were done on an IBM-AT equipped with a DSI-32 coprocessor board), the number of cells that could be accommodated were limited to (64, 16) in (x, z) . The final velocity models are shown in Figures 3(a) and (b) for white and brown noise additions, respectively.

As is evident, the white noise models show more scattering in the velocity inhomogeneities compared to the brown model. It is worth noting that of course any simulated seismic experiment over such models must of necessity consider the models to be deterministic. Only by a large number of repeated trials over many realisations of each model can true differences between the noise models be ascertained. This limitation is exaggerated by the limited sampling of each model in any simulated refraction experiment (this is not true of reflection surveys which

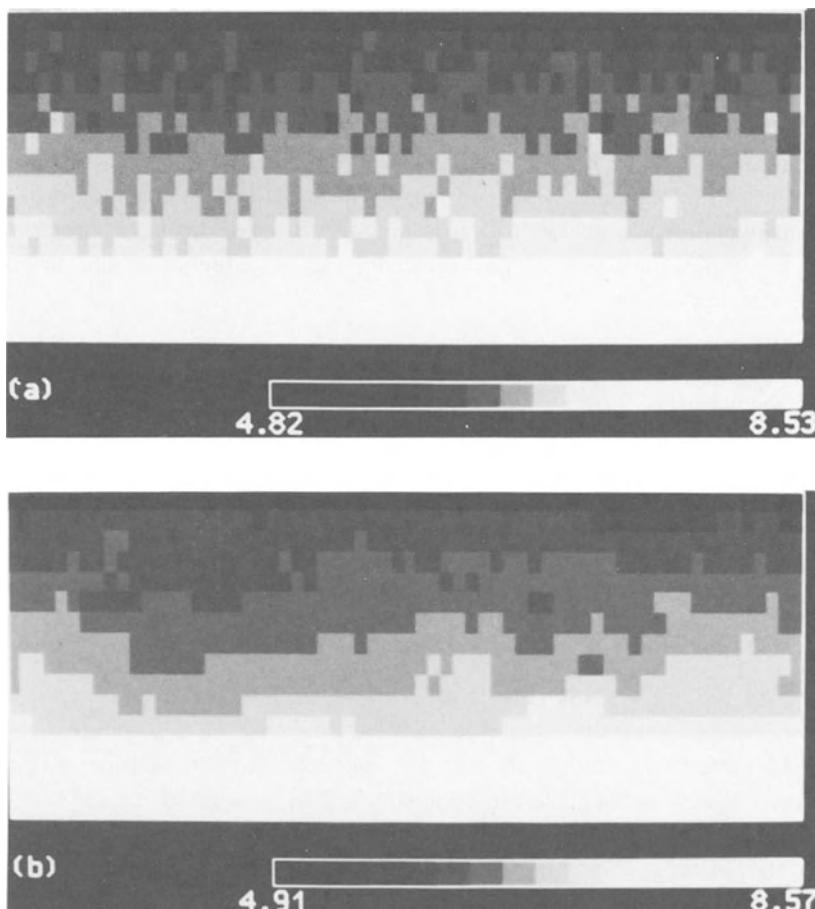


Figure 3

Crustal velocity models on a 64×16 grid in which the spacing is $4 \text{ km} \times 3 \text{ km}$. The base model is supplemented by a white noise model to produce (a) and by a brown noise model to produce (b).

sample the ground much more thoroughly). The very limited sample of cells in each of the two models will obviously limit the statistical differences in synthetics that might be produced.

4. Raytracing and Synthetics

The limitations in ray theory in dealing with scattering phenomena have been outlined by MEREU and OJO (1981). Briefly stated, the correlation length of inhomogeneities, here limited at the small scale to the cell size, should be larger than the acoustic wavelengths generated by the seismic source function. The minimum

cell size is (4 km, 3 km) in (x, z) respectively; for a velocity of 6 km/sec the seismic wavelengths vary from about 3 km to 0.5 km for typical refraction distances. Thus one could argue that the conditions for raytracing are barely met, but nonetheless boundaries only a few km apart are regularly used in refraction modelling, so in this respect we are no worse than common practise. Should it be possible to approach a 1×1 km grid numerically, in which velocity based on the noise models of Figure 1 (and hence representing a full scale refraction survey penetrating the upper mantle) could be constructed, the applicability of ray theory would have to be seriously questioned.

Base Model

The raytracing and synthetics for the base model are shown in Figure 4, and demonstrate three simple branches, the forward refraction or turning branch AB , the retrograde branch BC and the mantle refraction branch CD . The latter two branches are caused by the rapid increase in gradient at 33 km depth. In terms of sharp boundary models, BC is the branch containing the wide angle PmP reflections and CD is the mantle refraction Pn . The amplitudes decrease rapidly beyond the upper crustal velocity increase at 3 km and the mantle refractions are substantially weaker than AB , BC . These three arrival branches are the major ones sought in refraction surveys and they give the gross crustal parameters of any particular region. To construct geologically relevant models however, considerably greater information about crustal structure, in particular the location of major surface fault zones, is often sought. In a number of documented instances (one of them quoted by OJO and MEREU, 1986) investigators have attempted to model much shorter arrival branches as crustal layering. As we shall see it is easy to produce many smaller branches by adding small velocity inhomogeneities.

Base Model + Noise

The effect of the velocity construction (2) is summarised in the velocity depth profiles in Figure 5. In each case (white noise, brown noise), the solid lines give the maximum deviation of the model at any depth from the base model and the dotted lines give the respective mean deviations ($-$, $+$) for a value of a of 0.1. The synthetics and raytracing for the case of white and brown noise are shown in Figures 6 and 7.

We briefly give some numerical details of the constructions. Each model contains 2032 boundaries of which all vertical boundaries between cells have zero velocity. Rays are shot into the models at angular increments of 0.01 degrees. This leads to about 2,100 rays per model, of which about 1% are shown in the figures. These rays are automatically sorted into branches in which the surface distances

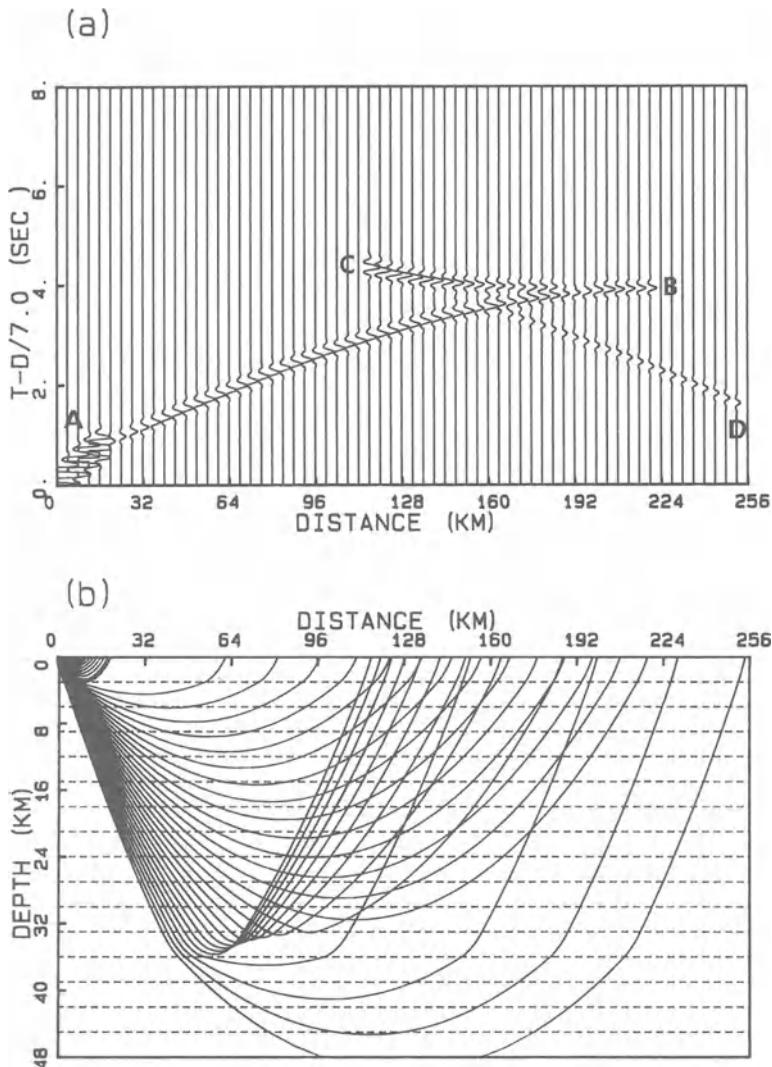


Figure 4

Synthetics (a) and raytracing (b) of the base model at surface stations spaced 4 km apart. The high amplitudes at *A* are produced by refractions within the strong gradient layer at the surface.

increase (or decrease) monotonically and for each branch an id number is assigned (see SPENCE *et al.*, 1984). We were careful to make sure that these id numbers were correctly identified for the amplitude interpolation. This was done by specifying that an arrival at any surface location could be interpolated only if the arrivals either side of that location (a) were on the same id branch and (b) that the arrivals were within twice the station spacing of that location.

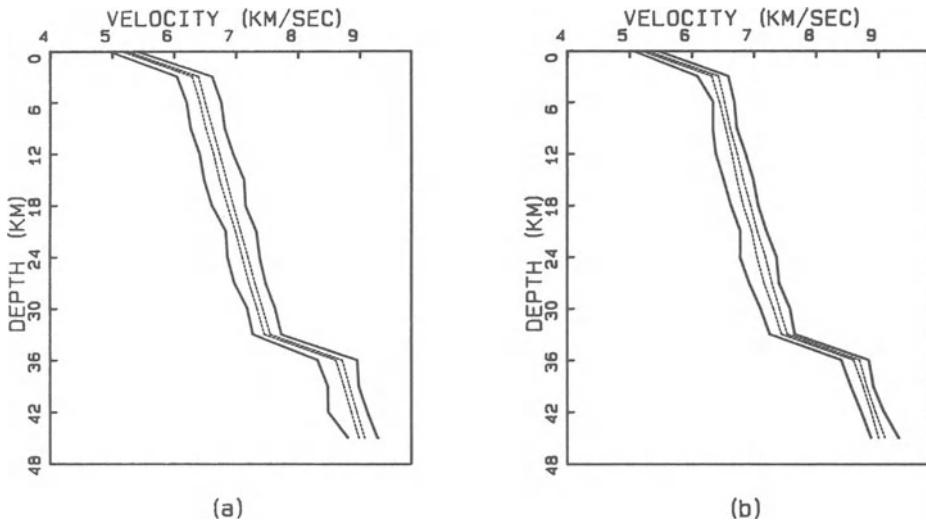


Figure 5

Velocity depth profiles for the base model plus (a) white noise, (b) brown noise. The wide limits (solid lines) are the maximum negative and positive deviations of the 64 horizontal cells from the base model at that depth. The inner (dotted) lines are the respective average deviations.

For the models shown, between 200 and 350 ray groups were found, many groups containing no more than 2 or 3 rays. While this indicates that the number of rays could have been increased by a large factor, one has also to realise that for such groups either the ray spreading is large (so the corresponding amplitudes are very small) or that the arrivals are extremely localised often within distances much less than the station spacing (4 km). Consequently, it was felt that no amplitude events of any significance were missed.

Despite the small number of cells actually sampled in the synthetics, some slight differences between the white and brown noise models can be seen (Figures 6 and 7). In particular, the white noise model shows a much weaker *BC* branch compared to the brown noise model indicating, perhaps, the smaller small-wavenumber correlation for white noise. Alternatively, one could instead note that the *BC* branch for the brown noise model is the only branch in both models which is relatively intact, perhaps indicating the presence of a more coherent structure for this model. Clearly a striking similarity between the figures is the intermittent nature of the primary *AB* branch and the virtual absence of the moho refraction *CD*. It is up to the reader's judgment to decide whether the synthetics on both figures could actually be interpreted as the same basic model with slightly different inhomogeneities.

We also decided to test two variations in shot position for the brown noise model. The first was to shift the shot point to +1 km (one quarter of a grid

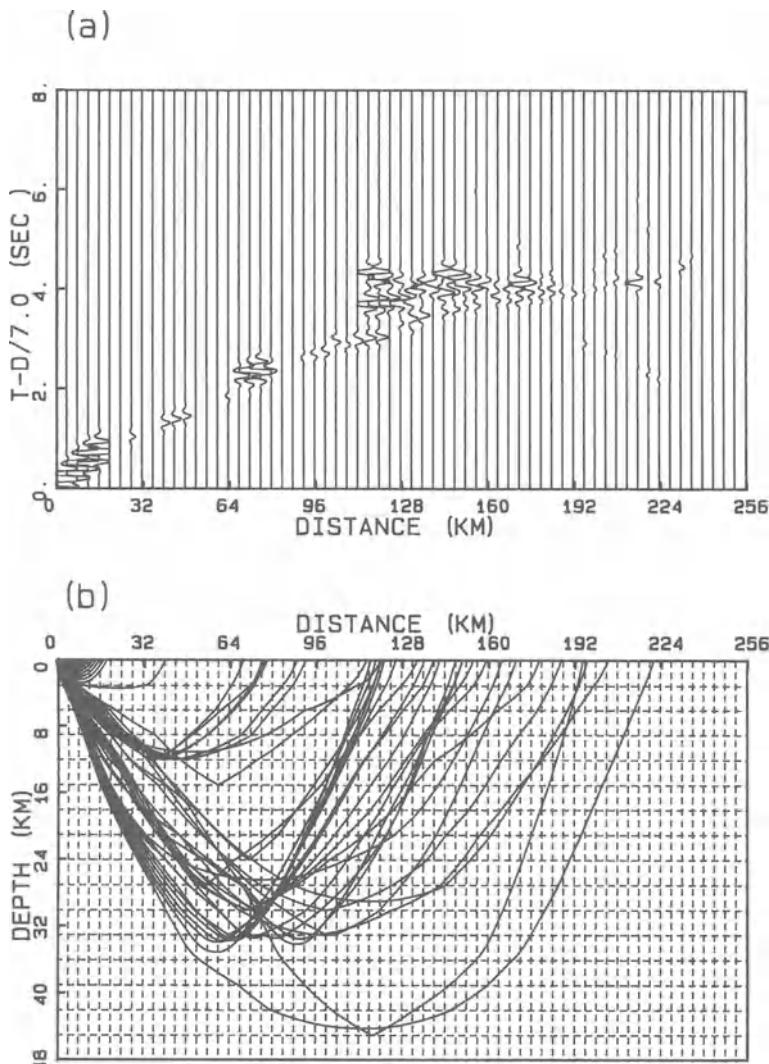


Figure 6

Synthetics (a) and raytracing (b) for the base model with additional white noise added at 64×16 cells are in Figure 3(a).

spacing) with the result shown in Figure 8(a). Some distinct differences between this figure and 7(a) are evident, particularly at distances less than 160 km, which indicates the known problem with shot point location. We note that the inhomogeneities introduced close to the shot point are in no way severe and one could argue convincingly for greater variation due to near-surface faulting and other inhomogeneities within the upper 3 km. As another example, we tried reverse

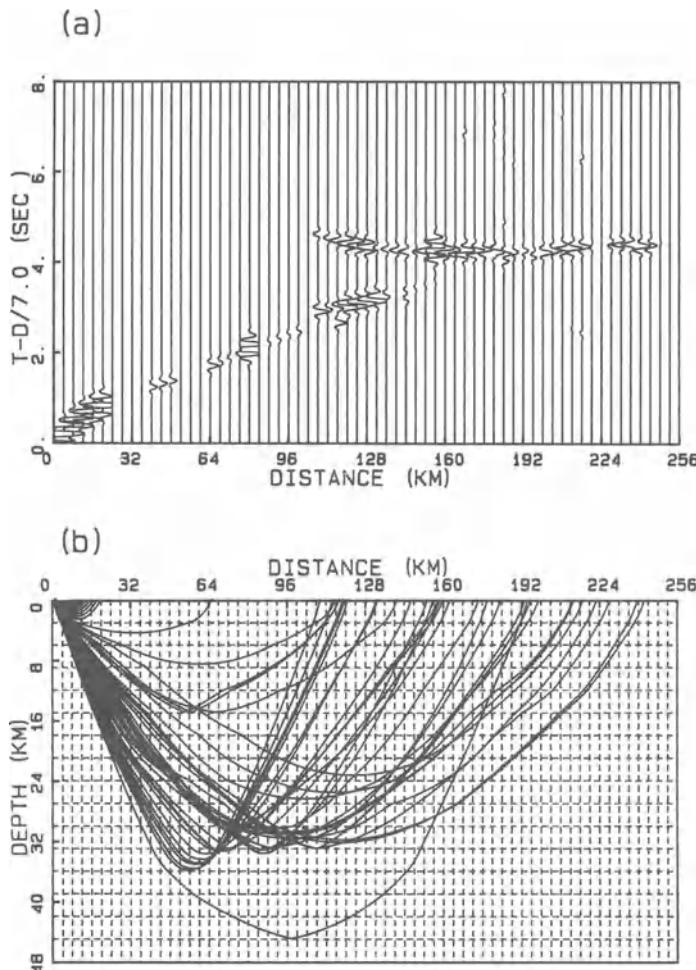


Figure 7
As Figure 6 but for the brown noise model in Figure 3(b).

shooting the brown noise model to produce the synthetics in Figure 8(b). Note in comparison to 7(a) the change in amplitudes between the late *BC* and *CD* branches and the different lengths over which coherent arrivals are present on the *AB* branches. Unsuspecting analysis could easily lead to a misinterpretation of such synthetics.

Base Model + Reflective Lower Crust

Many other variations of the distribution of noise models in the crust can be used in synthetic modelling. For example one could vary the amplitude a in

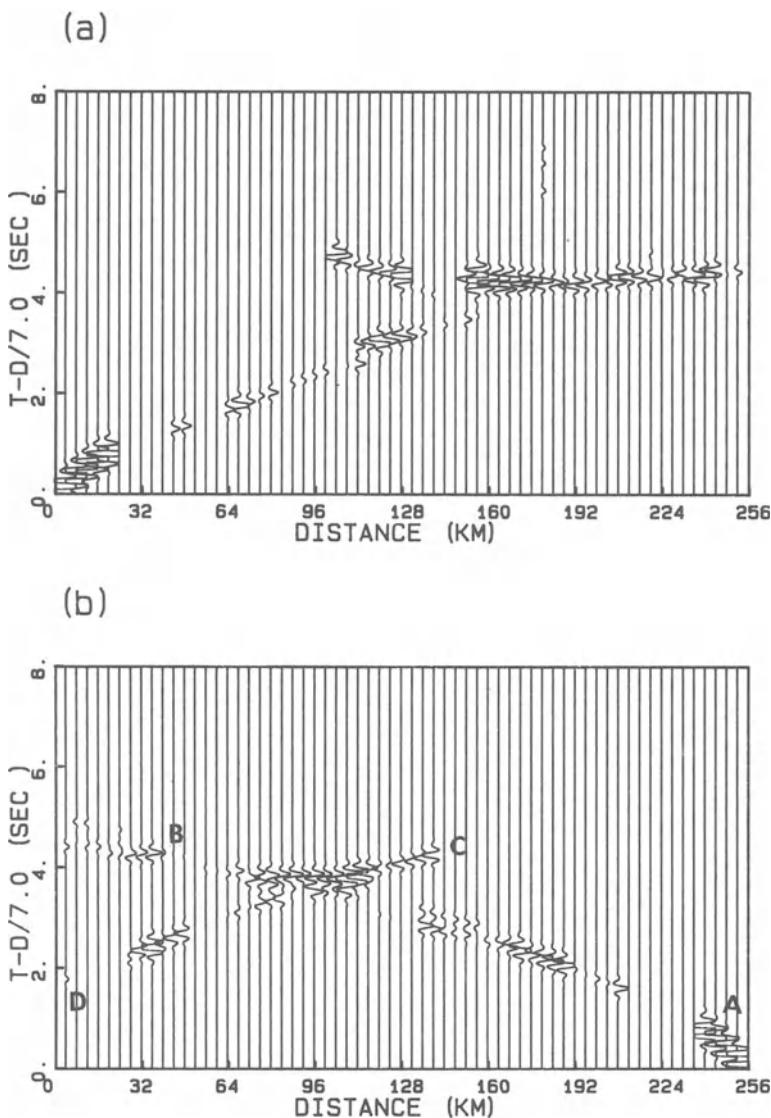


Figure 8
Synthetic seismograms for the brown noise model shot at surface locations of (a) 1 km and (b) 256 km.

equation (2), possibly together with the index n in equation (1), with lateral position or depth. One could thereby construct models in which the statistical measures (a, n) were indicative of various lithologies and geometry in the crust. As a simple example, we consider the effect of noise confined to the lower crust, to simulate the reflective lower crust found in many reflection surveys (MOONEY and BROCHER, 1987).

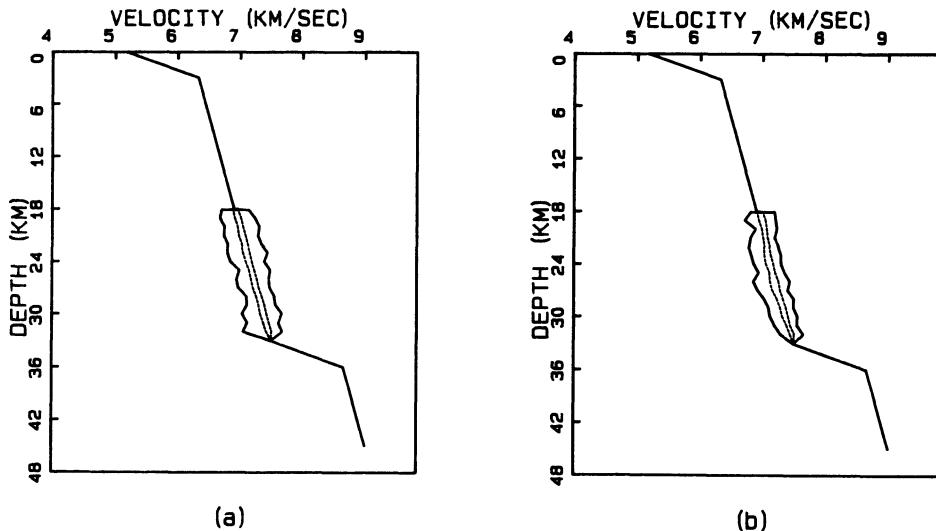


Figure 9

Velocity profiles for the base velocity model with a distribution of velocities in the lower crust between 18 km and 33 km—(a) white noise and (b) brown noise. The cell structure in the reflective zone is $4 \text{ km} \times 1 \text{ km}$ in x and z .

The appropriate velocity profiles are shown in Figure 9 where noise cells ($4 \text{ km} \times 1 \text{ km}$) were added to the base model between depths 18 and 33 km. Such a model roughly corresponds to the archetype of a transparent upper crust and reflective lower crust obtained from ECORS in France as quoted by MOONEY and BROCHER (1987). The interesting question in the present context is to what extent such a model allows the base model to be seen, especially the mantle transition below the reflective layer. Raytracing the brown noise model yields the results in Figure 10. In this case because no velocity inhomogeneity was present in the upper crust, branch *AB* is largely intact, at least up to the critical distance, though it would be expected that this branch would be disrupted by a near surface sedimentary sequence as in the ECORS profile.

As far as the later arrivals are concerned, the weak *Pn* branch is severely disrupted and so are the late *PmP* reflections. Even the early *CD* branch has erratic amplitudes. The synthetics for the white noise model had even less coherence. One would thus expect a reflective lower crust to pose severe interpretational problems for discerning the mantle transition in refraction surveys.

5. Conclusions

None of the results presented here run counter to previous work or to intuition. Though reassuring, this helps little in the analysis of refraction surveys. The inter-

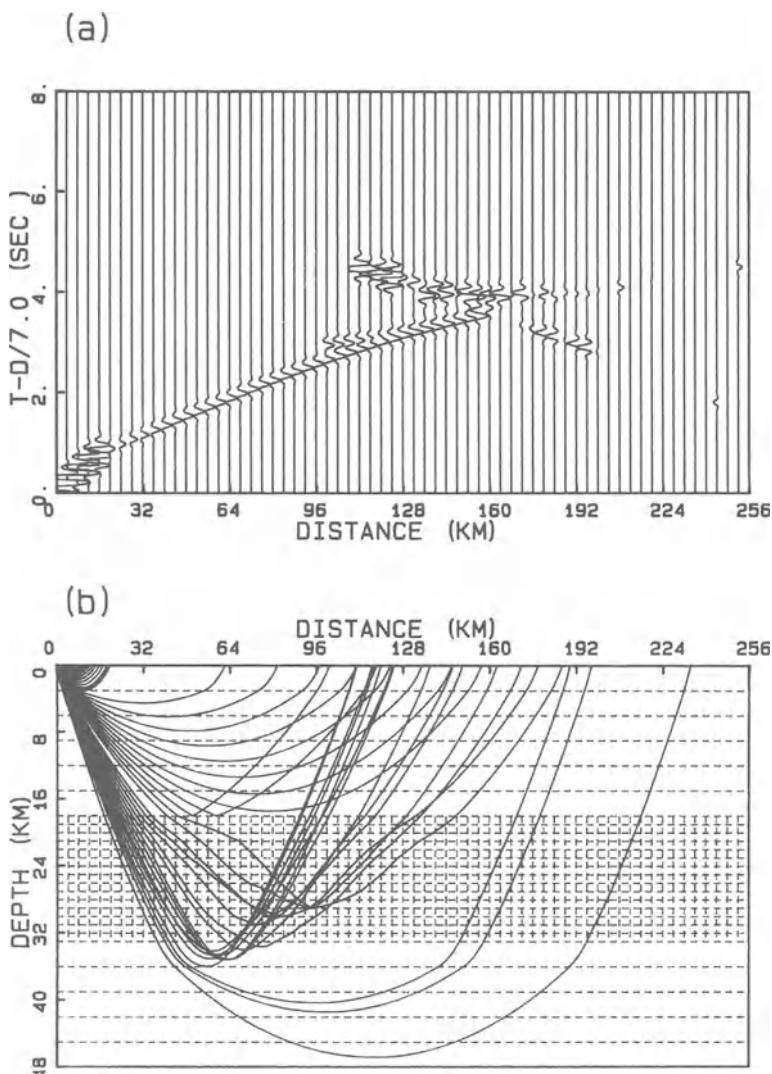


Figure 10
Synthetics (a) and raytracing (b) for a model with the velocity profile shown in Figure 9(b).

pretation of these results has to be made within a contextual framework. As the most obvious model, the brown noise velocity distribution on the one hand may be considered as a deterministic model with several structures within the crust (all of which have small velocity contrasts with the base model) with significant geological implications. On the other hand, this model is but one realisation of a random process which contains a large range of correlated small velocity variations and therefore is essentially devoid of geological meaning.

There is of course no answer as to the correctness of one or other of these

extremes. Presumably if one repeated a refraction experiment with many different shot locations, a certain component of the noise model may be smoothed or averaged, but for a truly self-similar process long-range correlations would remain embedded as essentially geological 'information'.

One fairly undeniable conclusion can be drawn however. The presence of velocity inhomogeneities, in addition to increasing observed traveltimes, also creates a clustering of amplitudes on the synthetics (through an interference effect) not unlike that observed in field data. In reality although little can be done to distinguish such features from coherent arrivals associated with 'genuine' structure, one should be extremely conservative about the interpretation of short travel-time branches in a geological region where no reflection data or other geophysical control is available.

Acknowledgements

We wish to thank Ron Clowes (UBC) for providing the program RAYAMP used in this study. The microcomputer based version, RAYAMP-PC, can be obtained by application to the authors. This research was supported by NSERC grant # A4240.

REFERENCES

- BERRY, M. V. (1979), *Difractals*, J. Phys. A: Math. Gen. 12, 781-797.
- CROSSLEY, D. J., JENSEN, O. J., and JACOBS, J. A. (1986), *The Stochastic Excitation of Reversals in Simple Dynamos*, Phys. Earth and Planet. Int. 42, 143-153.
- FRANKEL, A., and CLAYTON, R. W. (1986), *Finite Difference Simulations of Seismic Scattering: Implications for the Propagation of Short-period Seismic Waves in the Crust and Models of Crustal Heterogeneity*, J. Geophys. Res. 91, 6465-6489.
- HOSKEN, J. W. J. (1980), *A Stochastic Model of Seismic Reflections*, Paper G-69 presented at 50th Annual Meeting SEG, Houston, Texas.
- HUDSON, J. A. (1982), *Use of Stochastic Models in Seismology*, Geophys. J. Roy. Astr. Soc. 69, 649-657.
- LERCHE, I. (1985), *Multiple Scattering of Seismic Waves in Fractured Media: Cross-correlation as a Probe of Fracture Intensity*, Pure Appl. Geophys. 123, 503-542.
- MANDELBROT, B. B., *The Fractal Geometry of Nature* (Freeman, San Francisco 1982).
- MEREU, R. F., and OJO, S. B. (1981), *The Scattering of Seismic Waves Through a Crust and Upper Mantle with Random Lateral and Vertical Inhomogeneities*, Phys. Earth and Planet. Int. 26, 233-240.
- MOONEY, W. D., and BROCHER, T. M. (1987), *Coincident Seismic Reflection/Refraction Studies of the Continental Lithosphere: A Global Review*, Rev. Geophysics 25, 723-742.
- OJO, S. B., and MEREU, R. F. (1986), *The Effect of Random Velocity Functions on the Travel Times and Amplitudes of Seismic Waves*, Geophys. J. Roy. Astr. Soc. 84, 607-618.
- SATO, H. (1982), *Amplitude Attenuation of Impulsive Waves in Random Media Based on Travel Time Corrected Mean Wave Formalism*, J. Acoust. Soc. Am. 71, 559-564.
- SPENCE, G. D., WHITTAL, K. P., and CLOWES, R. M. (1984), *Practical Synthetic Seismograms for Laterally Varying Media Calculated by Asymptotic Ray Theory*, Bull. Seis. Soc. Am. 74, 1209-1223.
- WENZEL, A. R. (1982), *Radiation and Attenuation of Waves in a Random Medium*, J. Acoust. Soc. Am. 71, 26-35.
- WU, R-S., and AKI, K. (1985), *The Fractal Nature of the Inhomogeneities in the Lithosphere Evidenced from Seismic Wave Scattering*, Pure Appl. Geophys. 123, 805-818

(Received July 9, 1987, revised/accepted September 20, 1987)

Variations in Geometric Measures of Topographic Surfaces Underlain by Fractured Granitic Plutons

DENIS NORTON¹ and STEVE SORENSEN¹

Abstract—Silicate melts impose systematic fracture patterns on their hosts and their own subsolidus portions as they crystallize and cool to ambient conditions. Weathering of these plutonic bodies accentuates the fractures and produces fragmented topographic surfaces whose geometric measures exceed those of a Euclidean surface. Because geometric measures of these surfaces are potentially indicative of the percolation properties of the fracture network, contours and vertical sections were digitized from 1:250000 and 1:240000 scale maps, and their statistical fractal dimensions, \mathcal{D} , were computed. Regions underlain by granitic plutons with similar fragmentation patterns were found to have \mathcal{D} values ranging from 1.15 to 1.28 within a single pluton. This range of values is caused by noise and several geologic factors. The values contain noise introduced by the map-making and digitizing procedures that amounts to $\mathcal{D} \approx 1.05$. However, all values lie well above this threshold, and they correlate closely with local differences in weathering. Smaller \mathcal{D} values derive from contours where accumulation of screen or glacial erosion has smoothed the surface, whereas larger \mathcal{D} values derive from contours along aretes and in unglaciated terrains. In each of these regions, larger \mathcal{D} values occur locally where fractures are most frequent and/or continuous. Large \mathcal{D} values were also derived from regions underlain by host rocks that lie in the near-field region of the stress field caused by the pluton. The fractal values derived are apparently related to the magma-hydrothermal event, they can potentially be used to derive properties of the percolation networks that were active during the hydrothermal events.

Key words: Fractures, fractal geometry, granite permeability, topography.

1. Introduction

Exposures of granitic composition plutons that crystallized from volatile-rich magmas display a striking correlation between their topographic surfaces and fracture patterns within the pluton and nearby host rocks. Mountain range profiles that are underlain by these plutons have a diagnostic form; streams, game trails, and vegetation within the range follow systematic pathways through massive granitic blocks. Field relations indicate that these paths were developed from narrow, slit-like fractures formed deep beneath the earth's surface, that were then widened by erosion.

¹ Department of Geosciences, University of Arizona, Tucson, Arizona 85721, U.S.A.

Fracture patterns in these plutons and their host rocks are known from theory to be caused by mechanical deformation associated with the emplacement, crystallization, and subsolidus cooling of volatile-rich magmas (KNAPP and NORTON, 1979). Field studies of 20 such plutons in the western Americas confirmed that deformation during their thermal history produces discontinuous slit-like fractures that constitute a percolation network in the pluton and nearby lithologies (HAYNES and TITLEY, 1980; TITLEY *et al.*, 1986; NORTON, unpublished data). The weathering and erosion processes that exhume these systems also produce the patterns noted in their topographic surfaces through differential weathering along fractures. However, with minor exceptions the fractures can be attributed to the product of a dynamic process that included the dissipation of thermal, mechanical and chemical energy away from the cooling magma (NORTON, 1984).

Geometric properties of fracture sets in magma-hydrothermal systems have been widely used to reconstruct structural, fluid flow and chemical history (TITLEY *et al.*, 1986; VILLAS and NORTON, 1977; NORTON and TAYLOR, 1979; KNAPP and NORTON, 1979). Even though fractures have long been recognized as pathways for hydrothermal fluid flow, and inverse methods that estimate average values of rock permeability for the thermal event have been established, the functional relation between geometric properties of fracture networks observed in outcrop and the estimated permeability values has yet to be established for other than idealized situations (NORTON *et al.*, 1984).

Rock permeability is a function of geometry of the pore space through which the fluid flows and of the connectivity, intersection and abundance of fractures. Each of these characteristics also controls the erosion process and the topography. Therefore, the relationship between fragmentation patterns expressed in the topography and paleo-percolation properties of the fracture network is being examined. As a preliminary requisite to examining this relationship we have analyzed geometric variations of topographic surfaces underlain by the fossil remnants of volatile-rich magmas.

The topographic surfaces underlain by this class of pluton tend to be highly irregular; cut deeply by steep canyons that turn sharply as they follow first one fracture set and then another, these surfaces defy quantitative description. Although small-scale studies have established that topographic data is fractal and can perhaps be simulated by a fractional Brownian process (GOODCHILD, 1982), and analysis of topographic data have demonstrated that the regional faulting processes are fractal (AVILES *et al.*, 1987), fractal dimensions of topographic patterns have not been correlated with the underlying lithologic properties. In this manuscript we present the results of measurements made on topography underlain by a pluton whose structural state is known.

Beginning with a description of the Sawtooth Range—its form, rocks, and fractures—this paper discusses a method of measure for describing the surfaces and then applies it to lithologic units in the Sawtooth Range.

2. The Sawtooth Range

The Sawtooth Range is as topographically anomalous as its name implies. Aretes and horns in the central part of the range display profiles reminiscent of the cutter and raker teeth on a woodman's saw. As viewed from distant observation points, these features correlate directly with details of the fracture patterns that can be seen in outcrops of the granitic plutons underlying the region. A direct correlation is apparent at all scales between the topographic form of the erosion surface on this pluton and contiguous portions of its host, and the distribution of ancient fracture networks within the underlying bedrock.

2.1 Topographic Forms

The backbone of the Sawtooth Range is a sequence of central aretes whose base elevation is \approx 9000 ft; they connect a dozen prominent horns that rise to elevations above 10000 ft. Looming above Stanley Valley on the east, the aretes are clearly displayed by the contours between 8800 and 9400 ft elevation and can be followed

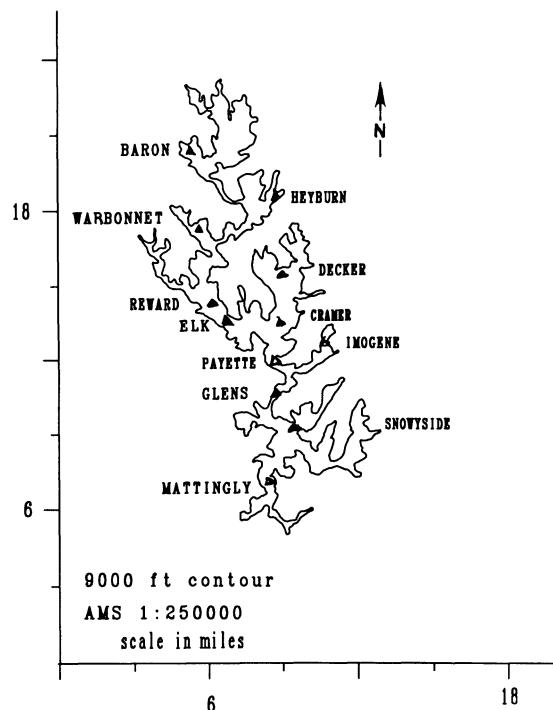


Figure 1

Plan map of the Sawtooth Range, Idaho, showing the 9000 ft contour and the principal peaks along the central arete. Data from the AMS 1:250000 and 1:24000 scale maps.

continuously from horn to horn starting with Mattingly Peak at the southern end of the range to Baron peak on the northern end, a distance of nearly 18 miles (Fig. 1). The form of these aretes derives from the northeastern and northwestern directed glacial valleys. At lower levels of detail each of these valleys also has many small cirques with similar orientations; a pattern that recurs at larger and larger scales of examination.

The scallop pattern in the contours is caused by cirques that occur approximately every 10000 ft and at amplitudes of \approx 10000 ft. This pattern recurs at wavelengths of \approx 2000 ft and amplitudes of \approx 2000 ft, where it is also caused by cirque development. At a level of even greater detail irregular, shorter-wavelength patterns occur that are also discernable at 1:250000 scale. These features are associated with small cols caused by preferential erosion locally along fractures spaced 50–400 ft apart.

The cirques and valleys are now free of glacial ice, but the transition from periglacial to subglacial features is still apparent. Blocky outcrops at high elevations give way to smooth outcrops below approximately 8400 ft, and even though they contain fracture patterns similar to those at the higher elevations they exert less control over the topography. However, the subglacial erosion process did accentuate some of the more continuous fractures, and the fracture system in many cirques caused roche moutonnée to form at steep angles to the ice transport direction. At elevations below \approx 7600 ft along the east side of the range, glacial detritus controls the topography. This transition zone varies in elevation by as much as \pm 400 ft from one portion of the Mt. Cramer region to another.

2.2 *Lithology*

The Sawtooth Range contains two major lithologic units. The core of the range is underlain by the Eocene (40–45 m.y.) Sawtooth Batholith (Figure 2) (REID, 1963; CRISS *et al.*, 1982; CRISS and TAYLOR, 1983). The lithologic properties of this granite pluton are the primary controls on the unusual topographic expression. The host for this pluton is the Mesozoic (\approx 100 m.y.) Idaho Batholith. The Idaho Batholith only occurs on the perimeter of the range and as small roof pendants along the western portions of the Sawtooth Pluton. Although the Idaho Batholith extends to the west, north, and south for tens of miles and contains a variety of rock types in this region, locally it is mostly quartz monzonite or leucocratic quartz diorite. Glacial deposits occur along the eastern margin of the range and along the deeply incised river canyons to the northwest. However, the contours over regions of moraine have not been included in the data in this study.

Because the magmatic event that generated the Sawtooth Pluton is the most recent activity in the region and it underlies most of the range, its history is of prime importance to this study. This history commenced with the accumulation of several thousand cubic kilometers of granitic magma into a chamber that deformed its own walls and deflected its overburden as it inflated. Within the Sawtooth Pluton

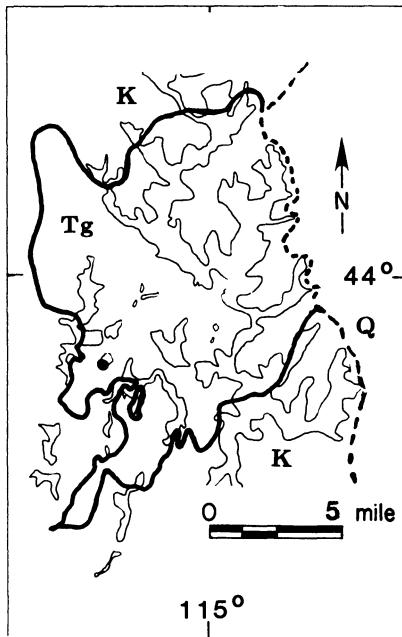


Figure 2

Geologic map of the Sawtooth Range, Idaho, showing the contact between the Eocene granite, Tgr, of the Sawtooth Batholith and the Cretaceous quartz monzonite, K, of the Atlanta Lobe of the Idaho Batholith and the glacial deposits, Q (from CRISS and TAYLOR, 1983). The 9000 ft contour from the AMS 1:250,000 scale map is shown for reference to subsequent figures.

miarolytic cavities are common, many of which occur in the form of bubble trains; these features are evidence that the magma was formed from a volatile-rich siliceous magma that crystallized into a granite. The emplacement process was accompanied by dissipation of thermal energy into the surroundings by combined processes of conduction and convection. During the post-emplacement cooling, a complex crystallization path coupled with differential heating of the hosts produced local stress conditions that were variable in intensity and orientation (KNAPP and NORTON, 1979). These conditions persisted over the thermal life of the pluton, $\approx 400,000$ years and produced the fracture patterns noted in outcrop today.

2.3 Fracture Orientation, Distribution, and Chronology

Thermal-mechanical events associated with the formation of the pluton were accelerated by fragmentation and consequent increase in rock permeability brought on by the formation of one fracture set with a northeasterly strike and one with a northwesterly strike (Figure 3). The interconnected fractures in these sets formed percolation networks for hydrothermal fluid flow, which increased the rate of heat

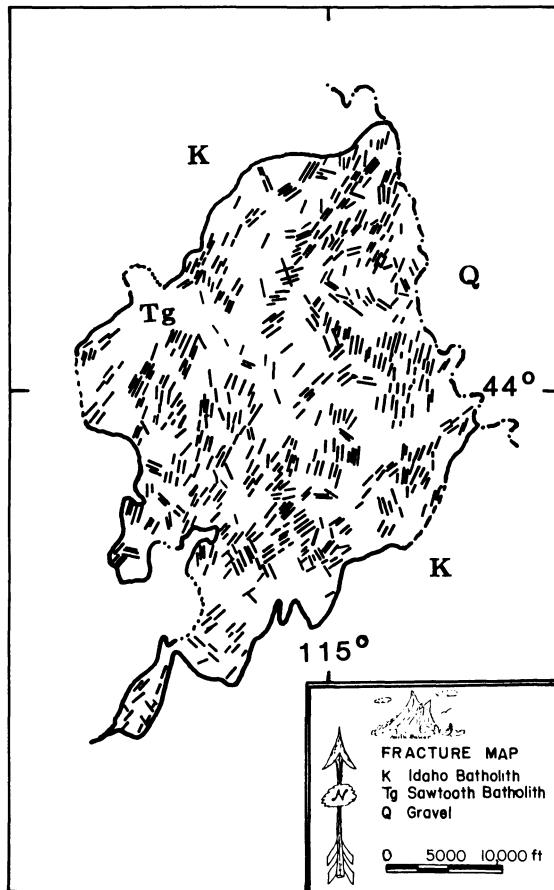


Figure 3

Fracture map of the Sawtooth Pluton, Idaho. Schematic of fractures discernible on 1:50000 aerial photography. Each line depicts the trend of fracture trace on the imagery.

and affinity transfer along the flow paths. These changes in the level of thermal and chemical energy in turn caused reactions of the fluids with the walls of the fractures and formed alteration effects that are diagnostic of these processes. The occurrence of hydrothermal mineral assemblages in the fractures provide evidence that the plutonic rocks were fragmented when the body was crystallizing and cooling, and that the fractures are not products of the current uplift and erosion cycle. The fracture patterns are most obvious in the towering aretes and horns that earned the range its name, but they also occur deep within the smoothly polished glaciated valleys. Seldom are the fractures farther than 100 ft apart or closer than a few inches, but commonly their spacing ranges from 10–50 ft. Although fractures within the valley floors exert only local control on the topographic surface, fracture spacing is much the same as on the high ridges. Locally the high valleys are filled

with prominent, angular roches moutonées carved by glacial action as it intersected the prominent fracture trends.

Near-vertical fractures oriented NW and NE are the most common, but small numbers of fractures that trend in all points of the compass occur locally (REID, 1963). These fracture trends cross boundaries between the Idaho and Sawtooth plutons, as well as subunits within them. This feature, together with the chemical and mineral alteration that is common on the dominant fracture sets, indicates that the fractures were percolation networks for hydrothermal fluids associated with the Sawtooth Pluton. These fractures are ≈ 35 m.y. old (CRISS *et al.*, 1982); products of the fragmentation process at that time.

The fracture sets and their associated alteration clearly reflect the time-integrated effects of thermal stresses and fluid flow. Therefore, the fracture sets that formed during the same thermal-mechanical episodes can be thought of as records of the processes that operated at the time. The fractures are estimated to have formed over a hundred-thousand-year interval of magma-hydrothermal activity (KNAPP and NORTON, 1979), that commenced ≈ 35 m.y. ago. Consequently they preserve information about the prevailing state of stress and local variations in stress trajectories during the activity. These fracture networks were also percolation paths for hydrothermal fluids as evidenced by the rock alteration associated with them.

The geologic observations described above verify that the fractures are the product of the magma-hydrothermal activity and lead to the following proposition: *Because most of the fractures were formed during the thermal event and now control the erosion process, the current topographic patterns are products of the dynamic processes during the thermal event and consequently are functionally related to conditions that prevailed during the pluton's thermal history.* To test this postulate, first we have measured the geometric properties of the topography.

3. Geometric Measures of Granitic Plutons

Fracture discontinuities subdivide lithologic units into blocks. Each subdivision by a fracture set produces N blocks that are self-similar. The similarity or scaling factor, r , among the blocks is defined by

$$r \equiv \frac{1}{N^{1/\mathcal{D}}} \quad (1)$$

where \mathcal{D} , the fractal dimension proposed by MANDELBROT (1983), is a real number that is a quantitative measure of the form of the fracture set and N is the number of self-similar blocks. Solving (1) for the geometric dimension, \mathcal{D} gives:

$$\mathcal{D} = \frac{\log N}{\log (\frac{1}{r})}. \quad (2)$$

Although fracture patterns in rocks may not display exact self-similarity, they are likely to display statistical similarity (MANDELBROT, 1983; Voss, 1985). Therefore, a statistical similarity ratio can be derived from measurements of the total length of any natural curve that is controlled by the fracture patterns. Where the total length is the product of the size of the ruler, ε , used to step off the length, \mathcal{L} , and the number of increments, N_ε , required to cover the curve:

$$\mathcal{L}(\varepsilon) = \varepsilon \times N(\varepsilon). \quad (3)$$

Because the number of steps is a function of ε and of \mathcal{D} from equation (1), the curve length is

$$\mathcal{L}(\varepsilon) = \frac{1}{\varepsilon^{\mathcal{D}-1}}. \quad (4)$$

Rearranging equation (4) and solving for the parameter \mathcal{D} gives

$$\log \mathcal{L}(\varepsilon) = (1 - \mathcal{D}) \log (\varepsilon) \quad (5)$$

where $(1 - \mathcal{D})$ is the slope of the locus of points of the log of curve length \mathcal{L} plotted against the log of ε used to obtain the curve length. Equation (5) forms the basis for testing the hypothesis that topographic surfaces underlain by fractured granitic plutons can be quantitatively represented by fractal dimension \mathcal{D} .

4. Digital and Analytical Procedures

Determining the fractal measure of an arbitrary space curve requires that it first be represented as a set of discrete points, the locus of which replicates the natural feature to within some desired degree of accuracy. Consequently, the study of topography in plutonic bodies reported here was based on digital data obtained from standard topographic maps.

The digital data were collected at several different scales using a standard high-precision digitizing tablet (Complot 7000). Tablet increments were selected to be consistent with the published map accuracies so that the data collected would have an accuracy at least as good as the maps. The map accuracy standards as legislated by the U.S. Mapping Agency (THOMPSON and MORRIS, 1981) are as follows:

1. *Horizontal Accuracy.* For maps at scales larger than 1:20000, not more than 10% of the positions are in error by more than $1/30^{\text{th}}$ of an inch; for maps at 1:20000 or smaller, the error is $1/50^{\text{th}}$ of an inch. These limits of accuracy apply to "positions of well-defined points only," *i.e.*, positions such as bench-marks, cultural features, monuments, and the like.
2. *Vertical Accuracy.* Not more than 10% of the positions are in error by more than one-half of a contour interval.

Each data set was collected using a digitizing increment approximately one-half the map accuracy. Because the legislated map accuracies pertain to the location of "well-defined" points, the digitized contours collected at a level of detail lower than the actual map accuracy are certain to contain noise introduced by the digital equipment and procedures.

All data sets were first analyzed with software that we developed to explore the empirical form of the function in equation (5). The tools are based on a ruler algorithm that uses the tolerance method similar to that reported by AVILES *et al.*, (1987). For each integral value of the contours a plot of $\log \mathcal{L}$ against $\log \varepsilon$ were made and the value of D determined by linear least-squares regression of the data. This part of the study revealed that the logarithm of contour length is actually a nonlinear function of the logarithm of ruler length for all the data sets. The functional relationship we found is typified by the 9000 ft contour from the Sawtooth Mountains (Figure 4).

The concave downward locus of the data points displayed in the example can be interpreted as three nearly linear trends:

1. For $0 < \varepsilon \leq 400$ ft, a linear trend with $D = 1.05$ is apparent in data from all data sets collected from 1:24000 scale maps (Figure 4). In some instances the upper bound of this domain is ≈ 200 ft. However, its exact location appears to depend on the degree of detail incorporated into the map at the stereo model stage of map production. The D value is attributed to a combination of digital, map

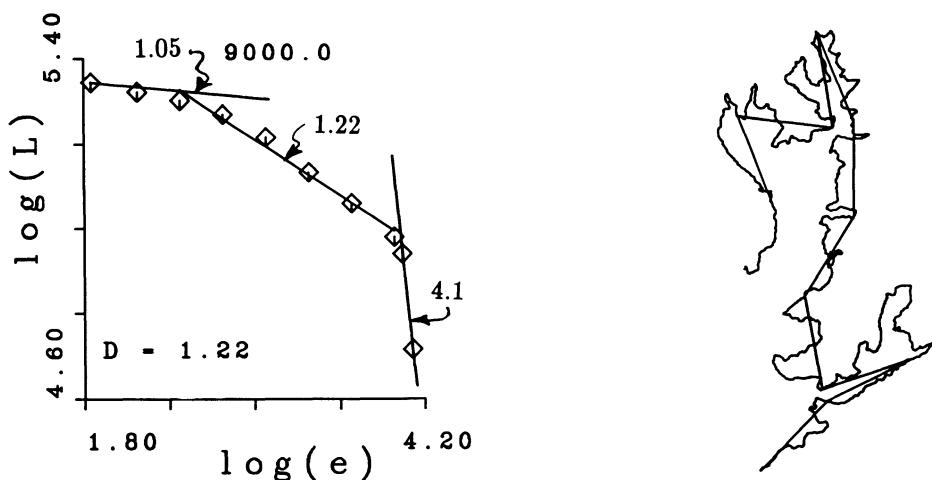


Figure 4

The logarithm of loop length, $\log \mathcal{L}$, as a function of the logarithm of ruler length, $\log \varepsilon$, for the 9000 ft contour loop in the Sawtooth Range, Idaho. Two approximately linear segments bound the domain of validity for calculation of the statistical D value. The spans of ε at length 9600 ft are displayed on the level zero loop. Digital data derived from: $7\frac{1}{2}$ minute 1:24000 topographic sheets for the Mt. Cramer quadrangle.

production, and natural effects. A lower limit of $\varepsilon = 400$ ft is used throughout this manuscript unless otherwise noted.

2. For $400 \text{ ft} < \varepsilon \leq E$, where E is the distance between extreme positions in the convex hull of the data set, there is a second linear trend in the function. This segment of the function appears to correlate closely with the variations in the contours caused by the topography, and therefore this portion of the function is used with equation (5) to compute the fractal measure \mathcal{D} . The small departures from linearity are attributed to the background effects introduced by the measurement technique as discussed above. Because only integral values of the ruler are included in the total length of the contour, some additional noise is introduced when the residual part of the ruler is large.
3. For $E < \varepsilon \leq \infty$, the number of applications of the ruler is less than a statistically significant population. In practice this domain of the function is easily recognized because the number of segments drops to ≈ 1 with a slight variation in ε . This segment is beyond the statistical validity of the technique.

These experiments lead to the conclusion that as the measuring ruler approaches zeroth order length—the average digitizing increment used to collect the data—the contour length is less dependent on ruler-length. At intermediate ruler lengths— $\varepsilon < 400$ ft in the 1:24000 scale data and $\varepsilon < 800$ ft in the 1:250,000 AMS data—the dependence increases. The upper ruler length is a function of the diameter of the contour's convex hull. *The statistical fractal dimension, \mathcal{D} , as determined from topographic maps is in part a function of the map-making process and the digitizing method.*

The computation of length using a wide range of ruler lengths is necessary to determine the optimum range in ε to be used in calculating the fractal measure \mathcal{D} . The upper limit, ε_{\max} , must generate a statistically valid number of segments. Consequently, its value cannot exceed the contour's maximum radius, and in practice the upper limit is a fraction of the radius. The effects discussed above were used as a basis for selecting that portion of the statistical function best suited for determining the natural geometric properties of the contours. The data in this range was fit to a straight line with the criteria that the regression coefficient in each case was better than 0.95.

5. Geometric Properties of the Topographic Surface

The regular but rough texture of the topographic surface that characterizes the Sawtooth Range defies description using traditional, straight-line geometry. A data structure that models this surface to within the accuracy of 7.5 min topographic maps requires nearly a megaword of computer memory. Condensation of this data into smaller, more general sets is necessary prior to analyzing the correlations among ancient and modern geologic processes and the form of the surface.

Therefore, we have reduced the topographic data set for a large portion of the range into scale-independent measures with the aid of concepts intrinsic to fractal geometry.

A range in fractal dimension was discovered in measurements made at both 1:24000 and 1:250000 map scales. The variations correlate with local geomorphic and geologic conditions in each case.

5.1 Measures at 1:24000 Scale

The individual contour patterns and their geometric properties as determined from the 7.5 min topographic maps of the Mt. Cramer and Warbonnet quadrangles are discussed below with respect to the two distinctly different geomorphic provinces. One province contains outcrops with a *sawtooth texture* and tends to be at the highest elevations; a second province contains outcrops with a *smooth texture* and occurs primarily below the 8400 ft level.

Sawtooth Texture: Aretes and horns above 8400 ft have been exposed to mountain periglacial conditions for several tens of thousands of years. These conditions have accentuated fracture zones within the granite to the extent that irregularities in the topography correlate closely with the occurrence of fractures.

The 9000, 9200, and 9400 contours are typical examples of this condition. They envelop the entire central arete region of the range and exhibit the strongest fractal character with \mathcal{D} measures that range from $\mathcal{D}_{9000} = 1.22$ to $\mathcal{D}_{9600} = 1.26$ (Figures 5, 6, 7 and 8, Table 1). These contours are unusual in that they are ≈ 100 miles in length when measured with an ε equal to the map resolution, a condition that permits statistical measures to be made over ≈ 1.5 orders of magnitude.

Table 1

Measures of the topographic contours in the Mt. Cramer Quadrangle, Sawtooth Mountains, Idaho, from the eastern portion of the central arete

Elevation (ft)	ε_0 (ft)	Length (ft) $\times 10^5$	Fractal \mathcal{D}^1	Residual
east	part	central	arete	
8000	56.6	1.85	1.15	0.95
8200	69.7	2.51	1.20	0.98
8400	64.1	3.06	1.24	0.93
8600	68.2	3.34	1.20	0.97
8800	73.4	2.13	1.23	0.99
9000	71.4	2.22	1.22	0.98
9200	71.9	2.37	1.25	0.96
9400	58.0	1.41	1.23	0.97
9600	70.0	1.29	1.26	0.98
9800	60.7	1.06	1.17	0.95

Table 1 (*Contd*)

Elevation (ft)	ε_0 (ft)	Length (ft) $\times 10^5$	Fractal D^1	Residual
west	part	central	arete	
7400	58.5	1.70	1.12	0.93
7600	57.4	1.88	1.13	0.92
7800	56.8	1.93	1.11	0.98
8000	56.8	2.12	1.15	0.98
8200	55.6	2.41	1.18	0.98
8400	56.4	2.73	1.17	0.99
8600	56.9	2.63	1.21	0.98
8800	57.3	2.75	1.21	0.98
closed	loops	on	cent	arete
9000	60.4	6.12	1.22	0.99
9200	64.3	4.25	1.26	0.97
9400	56.8	3.03	1.24	0.98

¹ Linear regression value from 14 values of ε .

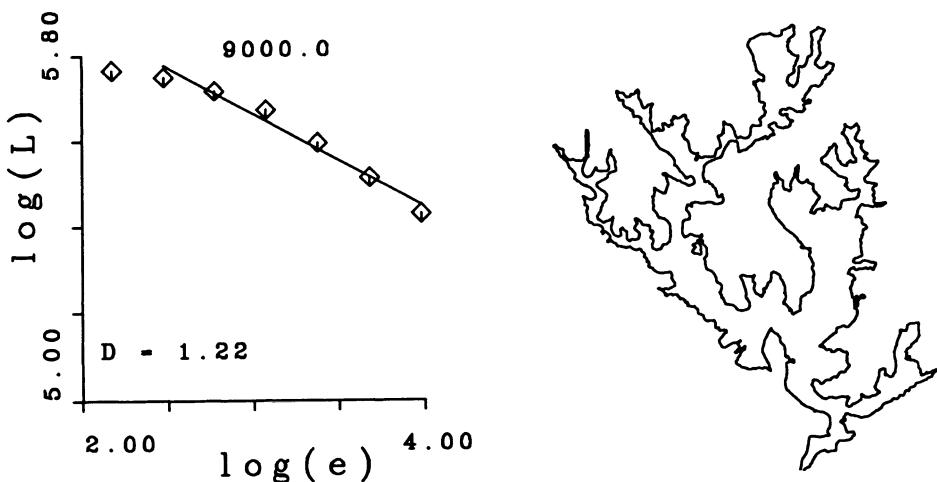


Figure 5

The logarithm of loop length, $\log \mathcal{L}$, as a function of the logarithm of ruler length, $\log \varepsilon$, for the 9000 ft contour loop in the Sawtooth Range, Idaho. Digital data derived from: $7\frac{1}{2}$ minute 1:24000 topographic sheets for the Mt. Cramer and Warbonnet quadrangles.

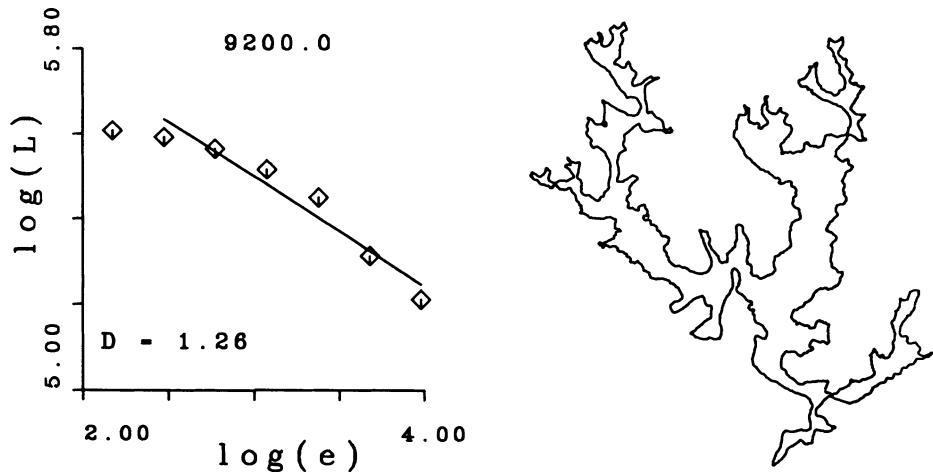


Figure 6

The logarithm of loop length, $\log \mathcal{L}$, as a function of the logarithm of ruler length, $\log \varepsilon$, for the 9200 ft contour loop in the Sawtooth Range, Idaho. Digital data derived from: $7\frac{1}{2}$ minute 1:24000 topographic sheets for the Mt. Cramer and Warbonnet quadrangles.

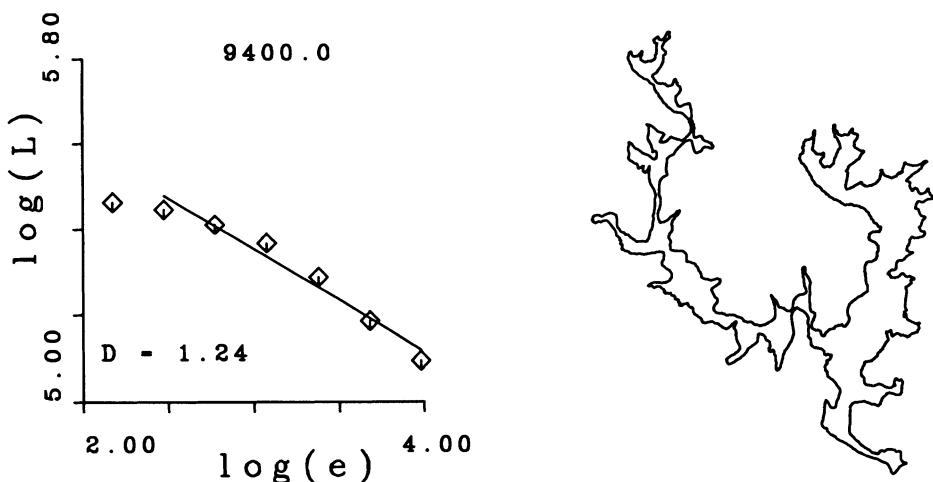


Figure 7

The logarithm of loop length, $\log \mathcal{L}$, as a function of the logarithm of ruler length, $\log \varepsilon$, for the 9400 ft contour loop in the Sawtooth Range, Idaho. Digital data derived from: $7\frac{1}{2}$ minute 1:24000 topographic sheets for the Mt. Cramer and Warbonnet quadrangles.

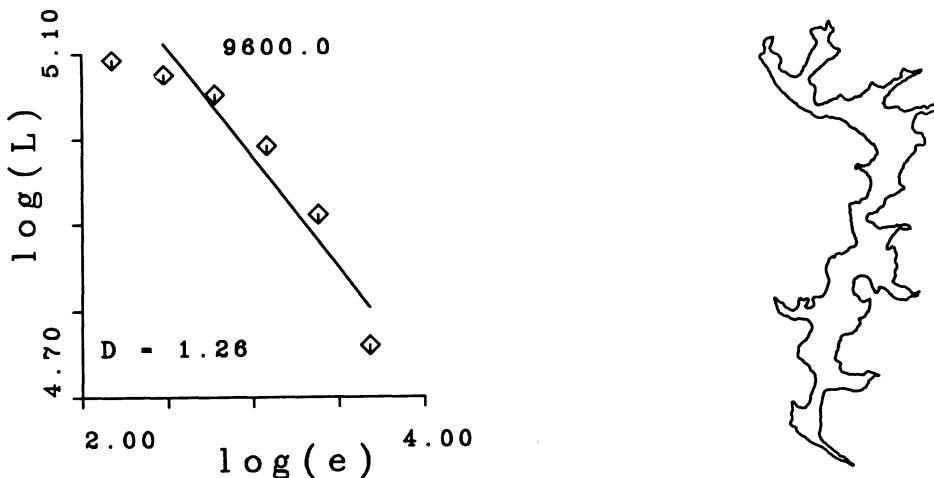


Figure 8

The logarithm of loop length, $\log \mathcal{L}$, as a function of the logarithm of ruler length, $\log \varepsilon$, for the 9600 ft contour loop along the eastern portion of the Mt. Cramer Quadrangle, Sawtooth Range, Idaho.
Digital data derived from: $7\frac{1}{2}$ minute 1:24000 topographic maps.

The contours are self-similar and have a pattern caused by the erosion bias of the glacial process for the northern exposures. This, combined with the dominant NE and NW fractures, produces a scallop pattern on the northerly side of the contours.

These contours display both the recursive pattern of cirques (scallop-like features also controlled by fractures) (Figures 5 and 6) at measurement lengths where $\varepsilon_n = 2400$ ft, and fracture-pattern control at $\varepsilon_1 = 150$ ft, in the south-easternmost promontory of the contour.

The 8800 and 8600 contours lie in the transition zone between the higher periglacial regions and the smooth glacial valleys. Their D values also lie between those for the contours of the two geomorphic environments (Figures 9 and 10).

Smooth Texture: Glacial valleys that have recently been freed of glacial action represent the second geomorphic province. Topographic contours at and below 8000 ft have fractal dimensions less than 1.20 (Figures 11–15). Although small D values suggest that a less fragmented bedrock controls the topography, field observations and interpretations of 1:10000 and 1:50000 photographs reveal a fracture distribution similar to the distributions at elevations above 8000 ft. The smoothing action of the valley glaciers has obliterated the topographic expression of the fragmented rock, as well as left a partial covering by debris.

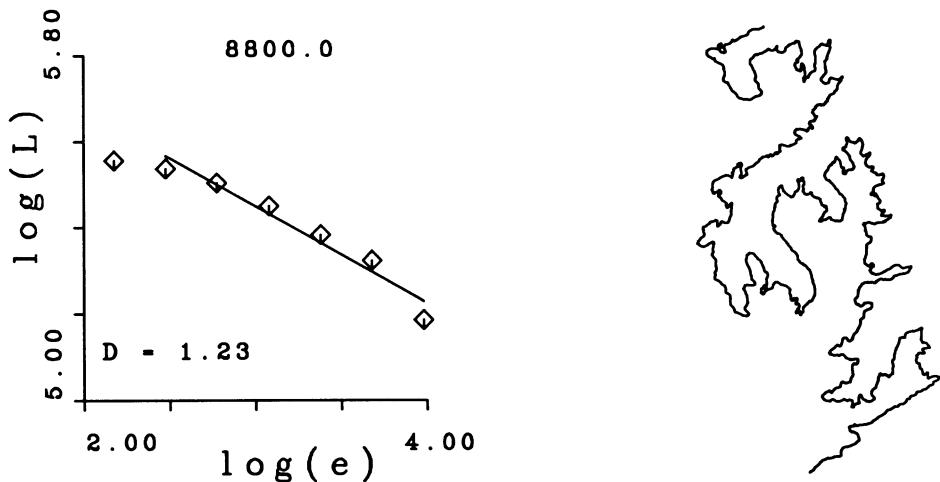


Figure 9

The logarithm of loop length, $\log \mathcal{L}$, as a function of the logarithm of ruler length, $\log \varepsilon$, for the 8800 ft contour loop along the eastern portion of the Mt. Cramer Quadrangle, Sawtooth Range, Idaho. Digital data derived from: $7\frac{1}{2}$ minute 1:24000 topographic maps.

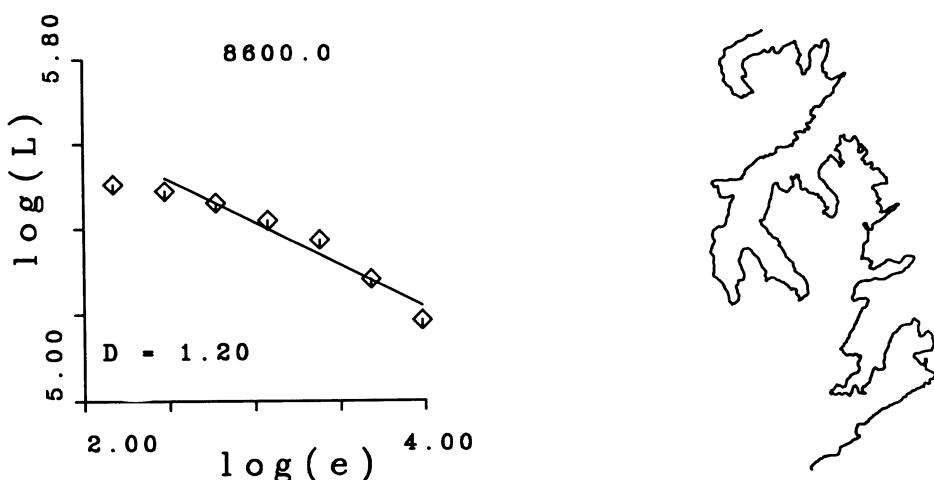


Figure 10

The logarithm of loop length, $\log \mathcal{L}$, as a function of the logarithm of ruler length, $\log \varepsilon$, for the 8600 ft contour loop along the eastern portion of the Mt. Cramer Quadrangle, Sawtooth Range, Idaho. Digital data derived from: $7\frac{1}{2}$ minute 1:24000 topographic maps.

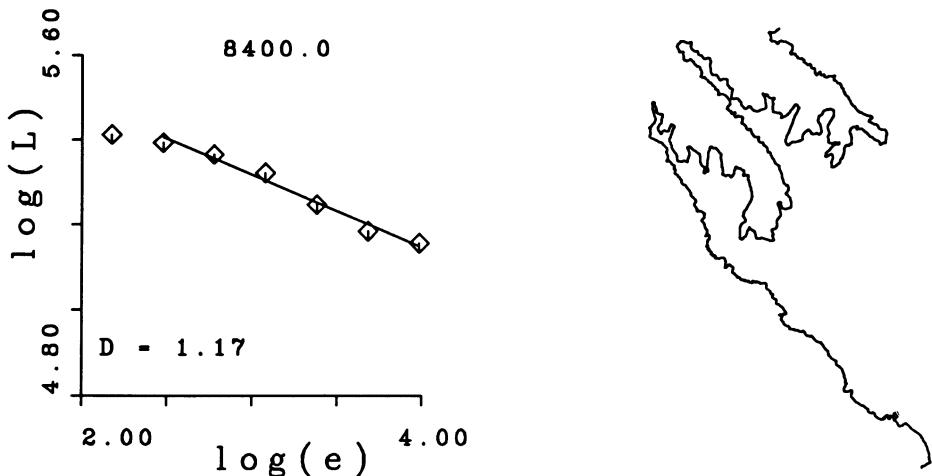


Figure 11

The logarithm of loop length, $\log \mathcal{L}$, as a function of the logarithm of ruler length, $\log \varepsilon$, for the 8400 ft contour loop along the central portion of the Warbonnet Quadrangle, Sawtooth Range, Idaho. Digital data derived from: $7\frac{1}{2}$ minute 1:24000 topographic maps.

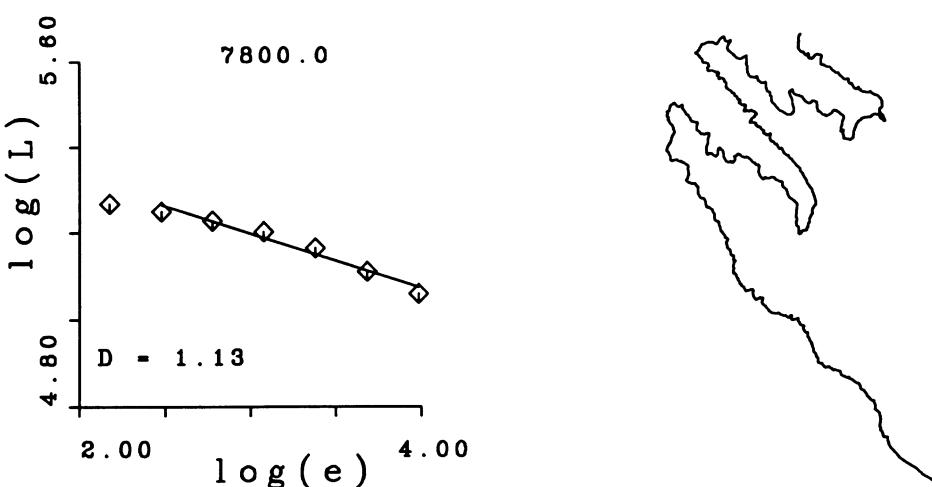


Figure 12

The logarithm of loop length, $\log \mathcal{L}$, as a function of the logarithm of ruler length, $\log \varepsilon$, for the 7800 ft contour loop along the central portion of the Warbonnet Quadrangle, Sawtooth Range, Idaho. Digital data derived from: $7\frac{1}{2}$ minute 1:24000 topographic maps.

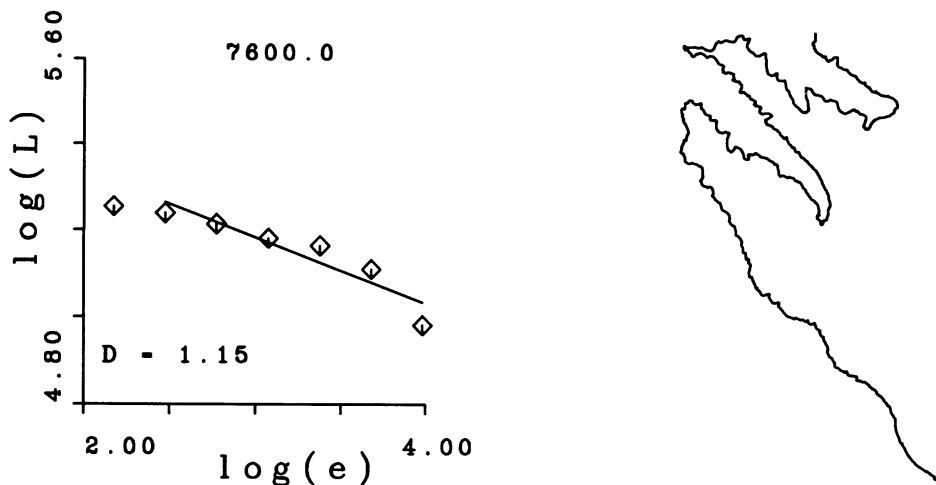


Figure 13

The logarithm of loop length, $\log \mathcal{L}$, as a function of the logarithm of ruler length, $\log \varepsilon$, for the 7600 ft contour loop along the central portion of the Warbonnet Quadrangle, Sawtooth Range, Idaho. Digital data derived from: $7\frac{1}{2}$ minute 1:24000 topographic maps.

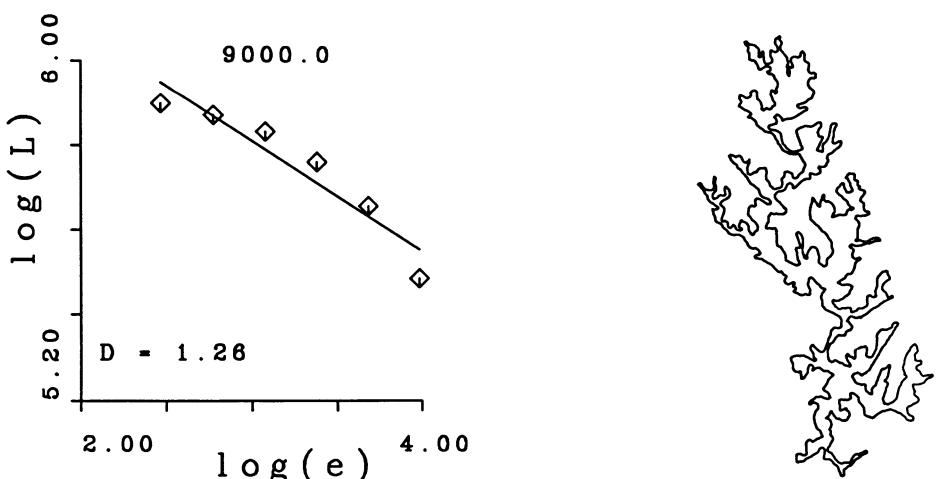


Figure 14

The logarithm of loop length, $\log \mathcal{L}$, as a function of the logarithm of ruler length, $\log \varepsilon$, for the 9000 ft contour loop along the central portion of the Sawtooth Range, Idaho. Digital data derived from 1:250000 AMS-series topographic maps.

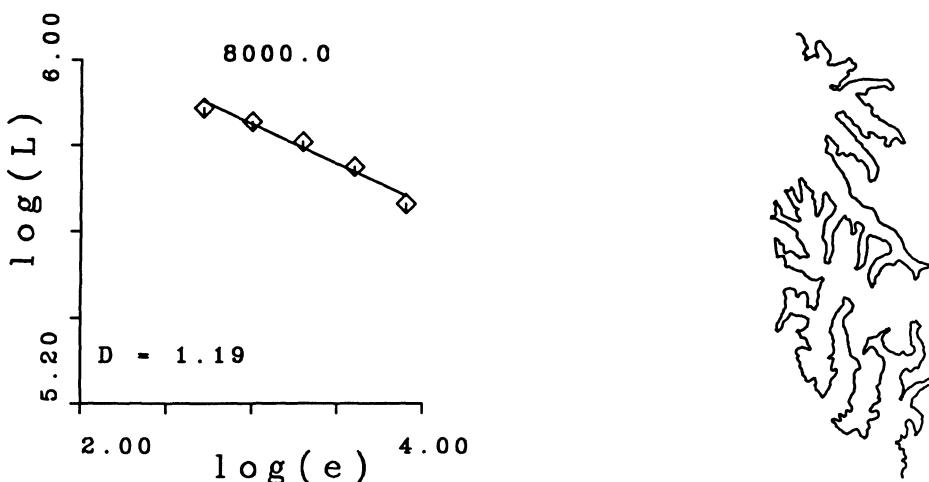


Figure 15

The logarithm of loop length, $\log \mathcal{L}$, as a function of the logarithm of ruler length, $\log \varepsilon$, for the 8000 ft contour loop along the central portion of the Sawtooth Range, Idaho. Digital data derived from 1:250000 AMS-series topographic maps.

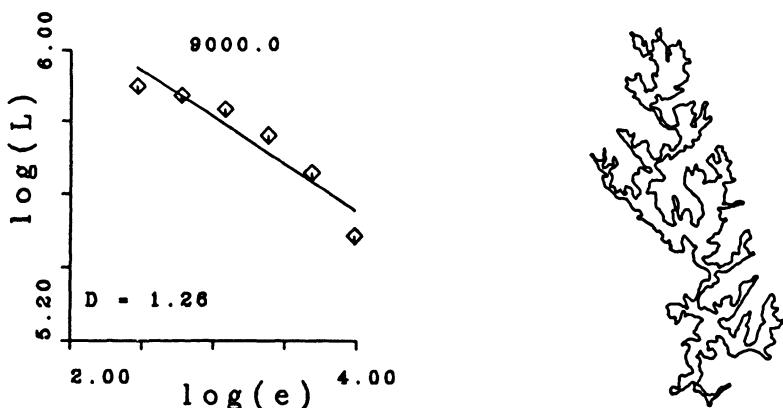


Figure 16.

The logarithm of loop length, $\log \mathcal{L}$, as a function of the logarithm of ruler length, $\log \varepsilon$, for the 9000 ft contour loop along the central portion of the Sawtooth Range, Idaho. Digital data derived from 1:250000 AMS-series topographic maps.

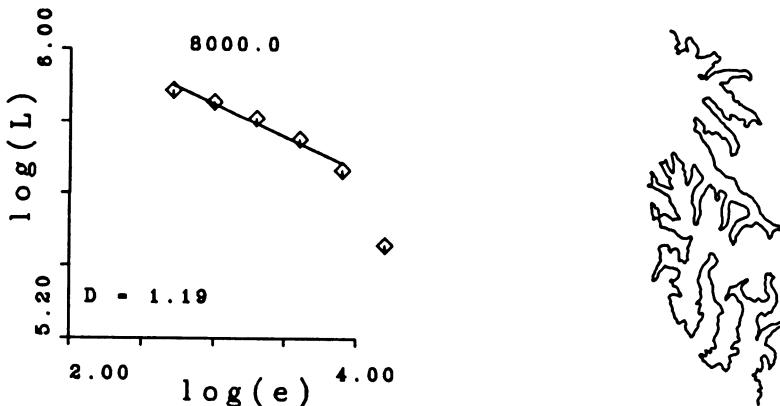


Figure 17.

The logarithm of loop length, $\log \mathcal{L}$, as a function of the logarithm of ruler length, $\log \varepsilon$, for the 8000 ft contour loop along the central portion of the Sawtooth Range, Idaho. Digital data derived from 1:250000 AMS-series topographic maps.

5.2 Measures at 1:250000 Scale

The contour patterns on the AMS 1:250000 Hailey and Challis topographic maps also display a prominent NW and NE grain in the region that coincides with the occurrence of the Sawtooth Pluton. Although the orientation of the ridges and valleys parallels the prominent fracture directions, strong variations in the depth of erosion are more evident at the smaller scale.

The fractal dimension of contours at the small scale is similar to those at the large scale but tends to be greater. The D value of 1.26 for the 9000 ft contour (Figure 16) is only slightly larger than the value determined from the large-scale map. This agreement, for two different map scales that were generated by two different methods of mapping, suggests that the texture of the contours is scale independent. The dimension for the 8000 ft contour, $D = 1.19$ (Figure 17) is also similar to the values from the large scale map values, $D = 1.15$.

6. Conclusions

The fractal dimension derived has potential use in several endeavors:

1. *Analysis of Strains.* The detection of regions in which the style of structural deformation is similar involves substantial judgment by the geologist. The large fractal measures in the near-field of the Sawtooth pluton appear to coincide with the unusual patterns of deformation caused by its formation. The dimensions of strain may also relate to the causative stress field and enable reconstructions

of stress conditions that have not been possible using traditional continuum mechanics.

2. *Correlation of Permeability Values with Fracture Networks.* The percolation property of fracture sets is a statistical measure of the rock's intrinsic permeability during a deformation epoch in which fractures formed. Such an epoch is defined as the interval during which the fracture set, \mathcal{F} , was active, where "activity" implies that the net heat flux caused by advective transport along the network provided by \mathcal{F} is greater than the conductive heat flux through the same region.

Although the diagnostic rugged pattern expressed in the topography of granitic pluton outcrops has long been recognized, the fundamental cause of these patterns has been attributed to uplift and weathering. Weathering of these bodies proceeds most rapidly along fractures and leaves the matrix blocks as towering pinnacles. This differential exhuming of the pluton accentuates fracture sets produced during the pluton's thermal history as well as those formed during the subsequent uplift by unloading of the vertical stress. The metasomatic effects associated with thermally related fractures permit recognition of those strains that are uniquely associated with the thermal event. Studies of several plutons reveal that only a small percentage of the fractures in outcrops were produced by the uplift and unloading process. Many of the thermal fractures have been enhanced by unloading but few new fractures of substantial extent were formed by this process. Consequently, the topographic patterns of plutons discussed in this text appear to be diagnostic of the fractures that formed in the body during its thermal history.

The fractal geometric measure, \mathcal{D} , of topographic contours on the 7.5 min maps of the Mt. Cramer and Warbonnet quadrangles ranges from 1.11–1.26. Although this variation correlates most directly with elevation, other contributors to the range of values include the rock type, fracture abundance and continuity, and extent of glacial smoothing. The largest values of \mathcal{D} are associated with the contours at high elevations along the central arete where weathering has accentuated fractures in the granite.

Acknowledgements

These studies are supported by NSF grant 8503114 to Norton. The manuscript has enjoyed the editorial attention of Emily Creigh and the critique of Christopher Barton and another anonymous reviewer. We would like to thank Jenny Miller, Bob Clark, Emily Creigh, Debbie Norton, Dan Davis and Stephanie Shakofsky who toiled over the digitizing tablet collecting data with great care and exercising patience with the quirks of the process. We wish to also thank Stephanie for her persistent and probing questions that helped improve the methodology and software.

REFERENCES

- AVILES, C. A., SCHOLZ, and BOATWRIGHT, J. (1987), *Fractal Analysis Applied to Characteristic Segments of the San Andreas Fault*, Journal of Geophysical Research 92, 331-344.
- CRISS, R. E., LANPHERE, M. A., and TAYLOR, H. P. Jr., (1982) *Effects of Regional Uplift, Deformation and Meteoric-hydrothermal Metamorphism on K-Ar Ages of Biotites in the Southern Half of the Idaho Batholith*, Journal of Geophysical Research 87, 7029-2046.
- CRISS, R. E., and TAYLOR, H. P. Jr., (1983), *An $^{18}\text{O}/^{16}\text{O}$ and D/H Study of the Tertiary Hydrothermal Systems in the Southern Half of the Idaho Batholith*, Geological Society of America 94, 640-663.
- GOODCHILD, M. F. (1982), *The fractional Brownian process as a terrain simulation model*, In *Modeling and Simulation*, Proceedings of the 13th Annual Pittsburg Conference on Modeling and Simulation, v. 13, pp. 1133-1137.
- HAYNES, F. M., and TITLEY, S. R. (1980), *The Evolution of Fracture-related Permeability within the Ruby Star Granodiorite, Sierrita Porphyry Copper Deposit, Pima County, Arizona*, Economic Geology 75, 673-683.
- KNAPP, R., and NORTON, D. (1979), *Preliminary Numerical Analysis of Processes Related to Magma Crystallization and Stress Evolution in Cooling Pluton Environments*, American Journal of Science 281, 35-68.
- MANDELBROT, B. B., *The Fractal Geometry of Nature* (New York, W. H. Freeman and Company 1983), 468 pp.
- NORTON, D. (1984) *A Theory of Hydrothermal Systems*, Annual Review of Earth and Planetary Sciences 12, 155-177.
- NORTON, D., and TAYLOR, H. P., Jr. (1979), *Quantitative Simulation of the Hydrothermal Systems of Crystallizing Magmas on the Basis of Transport Theory and Oxygen Isotope data: An analysis of the Skaergaard Intrusion*, Journal of Petrology, 20, 421-486.
- NORTON, D., TAYLOR, H. P., Jr., and BIRD, D. K. (1984), *The Geometry and High-temperature Brittle Deformation of the Skaergaard Intrusion*, Journal of Geophysical Research 89, 10178-10192.
- REID, R. R., (1963), *Reconnaissance Geology of the Sawtooth Range*, Idaho Bureau of Mines and Geology, Pamphlet 129, Moscow, Idaho.
- TITLEY, S. R., THOMPSON, R. C., HAYNES, F. M., MANSKE, S. L., ROBISON, L. C., and WHITE, J. L. (1986), *Evolution of Fractures and Alteration in the Sierrita-Esperanza Hydrothermal System, Pima County, Arizona*, Economic Geology 81, 343-370.
- THOMPSON, R. C., and MORRIS, M., *Maps for America* (U.S. Department of the Interior 1981) 265 pp.
- VILLAS, R. N., and NORTON, D. (1977), *Irreversible Mass Transfer between Circulating Hydrothermal Fluids and the Mayflower Stock*, Economic Geology, 72, 1471-1504.
- VOSS, R. F., *Random fractal forgeries*, In *Fundamental Algorithms for Computer Graphics* (Earnshaw, R. A. ed.) (NATO ASI Series, F17, 1985).

(Received May 11, 1987, revised/accepted September 4, 1987)

Fractal Sinuosity of Stream Channels

R. SCOTT SNOW¹

Abstract—Analysis of a diverse set of twelve stream channel planforms indicates that at scales relevant to river meandering, river traces are most reasonably treated as fractal curves. The atypically high degree of channel wandering apparent at such scales is a common characteristic of all single-channel streams, while identifiable meandering appears to be only one way in which this tendency is expressed. At smaller scales of view river paths have shapes of smooth curves appropriate to Euclidian geometry, and toward larger scales a distinct change in degree of wandering marks the transition to bends that are considered to be changes in general river course rather than parts of channel pattern. This analysis method provides a natural, objective calculation of river sinuosity as well as other parameters that more completely specify channel planform.

Key words: Rivers, channel pattern, fractal geometry, meandering, sinuosity, quantitative methods.

Introduction

River channels invariably wander from a straight line course, with bends and kinks in planform seen on all different scales of view, from the short channel reach to the full stream length. This wandering in planform is generally agreed to be the result of irregular, stochastic influences in all cases but one. Channel meandering, expressed at a scale one to two orders of magnitude larger than the channel width, is considered to be a regular, periodic form resulting deterministically from hydraulics and possibly other influences.

But even free meanders on broad floodplains fail to bear out the conceptual model of a series of regularly-spaced, smooth, equal-sized arcs. Recent reviews (RICHARDS, 1982; KNIGHTON, 1984) emphasize that meander shapes are intermediate between regular forms and purely random ones. Controls such as inhomogeneous valley fill can certainly be appealed to as factors distorting an otherwise regular, "ideal" planform, yet even the shape of this ideal planform is in dispute (CARSON and LAPOINTE, 1983).

¹ Department of Geology, Ball State University, Muncie, IN 47306, U.S.A.

This is a common predicament for the geologist. One is regularly confronted with phenomena that are not well represented by standard mathematical forms, but that nevertheless elicit a deeply felt sense of *structure*.

Analysis of meander planforms currently begins from an assumption of regular, periodic, "ideal" shape characteristics. The traditional description is in terms of radius of curvature, wavelength, and other parameters appropriate to simple geometric shapes. Evaluation of meander bend sequences by series analysis (SPEIGHT, 1965; FERGUSON, 1977) is more appropriate for treatment of irregular data, but still involves application of Fourier or other transforms to separate periodic signals from masking, presumably random noise. Workers applying current methods recognize the importance of the irregular component in meander planform, but the point of departure for the methods themselves remains an assumption of regularity, not irregularity.

The alternative mathematics of fractal geometry involve forms characterized by irregularity, structure, and expression of both of these over a large range of scales. There is some indication in the literature that river planforms have a fractal nature, at least when viewed at scales larger than those relevant to meandering. MANDELBROT (1983, p. 111) interprets HACK's (1957) exponent for ratio of stream length to basin area as an indicator that 1.2 is the general fractal dimension for river channels. Also, RICHARDSON'S (1961) analysis of national borders, interpreted by MANDELBROT (1967, 1983) as revealing fractal characteristics, includes the "land frontier" of Portugal which largely lies along rivers.

I have performed a detailed fractal analysis of planform for twelve single-channel stream segments in the western Ohio River drainage basin. The analyses treat planform characteristics on scales ranging from 0.1 km to the total segment length, in some cases exceeding 100 km. This allows exploration of fractal characteristics of stream channels at scales where meandering is expressed, and to some degree at larger and smaller scales.

Methods

The method of analysis for fractal characteristics used in this study is described in detail by RICHARDSON (1961). In brief, the length of a wiggly line on a map is measured by walking a map divider along it, with divider points set at a particular spacing. The spacing is changed and the measurement is repeated a number of times. In general, total length (measured line length) becomes greater as smaller unit lengths (point spacings) are used, taking smaller and smaller wiggles into account. Richardson determined that, for coastlines and other national boundaries, a log-log plot of total length against unit length approximates a straight line with finite slope. MANDELBROT (1967) indicates that such a result

expresses the statistical self-similarity that is characteristic of a fractal curve, with the slope of the plotted line equal to $1 - D$, where D is the fractal dimension of the wiggly line on the map.

Actual measurements by hand as described above were performed in a preliminary test of fractal analysis on river planforms (BERQUIST and SNOW, 1985). For the present study, river planform segments have been digitized and the Richardson procedure coded into a FORTRAN program (available from author) to allow more thorough analysis. The planforms used are mid-channel traces of streams as represented on 1:24000 scale topographic maps, digitized at data point spacing of approximately 25 m true scale, or 1 mm map distance. To assure that this limiting, practical data resolution does not influence the analysis results, the minimum unit step length for measurement has been set at 100 m true scale.

The set of planforms represents a wide range of natural channel conditions, including straight channels, free meanders on wide floodplains, confined meanders of first and second order (LEWIN and BRINDLE, 1977), incised meanders, and channel paths with high degrees of irregularity and angularity. All are located in plateau and lowland regions of Indiana and Kentucky, with various geologic controls including glacial deposits and karst development.

To better interpret the log-log graph of analysis results (here referred to as a "Richardson plot") for natural streams, we first consider such an analysis of idealized meander planforms (Figure 1). Regardless of whether the artificially generated planform is a sine-generated curve (LANGBEIN and LEOPOLD, 1966) or a series of upstream-skewed loops (CARSON and LAPOINTE, 1983), the same general characteristics appear on the plot. In the left portion of the graph, representing small unit lengths of measurement, the plot is a curve asymptotic to zero slope. This is an indicator of smoothness, lack of irregularities on a small scale of view. Throughout this paper "scale of view" is identified directly with the linear scale of some unit length used in the analysis. For example, the right portion of the plot has a slope of zero because at a scale of view of about 30 km the meander loops cannot be "resolved" and the planform appears as a straight line.

The second notable characteristic is the disrupted central portion of the plot, representing scales of view that approximate the uniform size and spacing of meander loops. Such disruption indicates a periodic signal in the analyzed data, with plot discontinuities and positive slopes resulting as progressively increasing unit lengths become in and out of "phase" with the signal. To the degree that regular signals are present in natural river planforms, the Richardson plot indicates their existence and characteristic scales by presence of disruptions. Alternatively, to the degree that the river planform exhibits similar degrees of irregularity across a range of different scales of view, the Richardson plot will be smooth, continuous, and straight with nonzero slope.

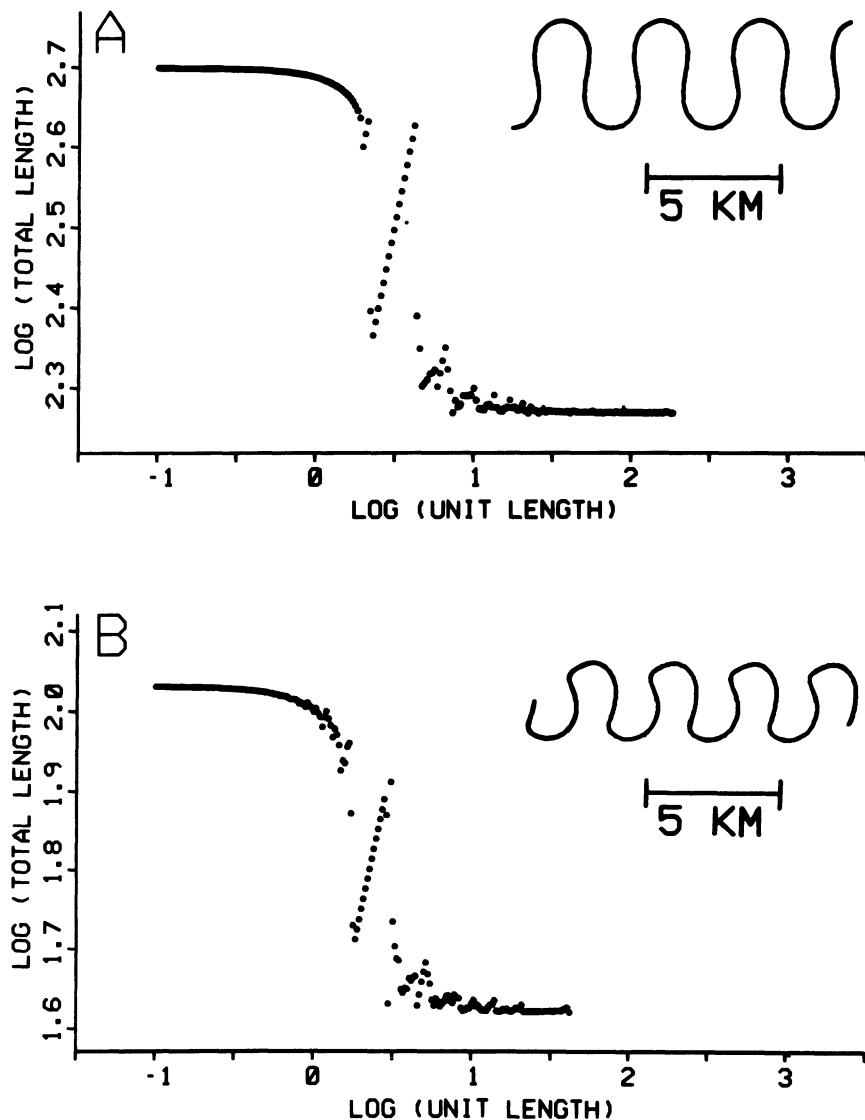


Figure 1

Richardson plots for ideal meander planforms, with representative subsegments of those planforms. The planform in Part A is a sine-generated curve with maximum deviation angle of 100 degrees, and that for Part B has been constructed from schematic, upstream-skewed loops in Figure 2 of CARSON and LAPOINTE (1982). The vertical scale is expanded. Plot slopes asymptotic to zero and disrupted plot sections including positive slopes indicate smoothness, periodicity, and lack of fractal character. With no large scale channel wandering involved, sinuosity in each case is simply the ratio of the stream lengths marked by the upper and lower asymptotes of the plot.

Results

Richardson plots and planform traces for six of the river segments studied appear in Figure 2. This group represents well the range of results obtained. Although the set of planforms includes a variety of types, the analysis results show a particular scheme applying universally: low degrees of wandering observed both at large and small scales, with an intermediate scale region showing a relatively high degree of wandering.

The smaller scale regions of these plots are generally curves asymptotic to zero slope, similar to those seen in Figure 1, confirming that river paths are generally smooth, Euclidian shapes when viewed at scales approaching those of channel width. While plot slopes in two of the examples (Figures 2D,F) are significantly different from zero at the smallest scales considered, these represent relatively narrow streams with values of channel width (Table 1) well below 100 meters.

Table 1
Channel planform parameters and related data for stream segments analyzed in the study

Stream segment	Channel length <i>L</i> (km)	Approx. channel width*	Sinuosity fractal dimension <i>D_p</i>	Limits of scale within which <i>D_p</i> applies (km)		Fractal sinuosity <i>P_f</i>
				lower	upper	
Big Indian Creek, from near Georgetown, Ind. to Ohio River	84.4	50	1.38	0.40	2.5	2.2
Little Blue River, from English, Ind. to near Ohio River	54.6	20	1.34	0.35	3.5	2.3
Patoka River, Ind. entering Wabash River	23.1	25	1.34	0.14	1.1	2.2
Green River, from Greensburg, Ky. to Asphalt, Ky.	177	60	1.32	0.7	6.3	2.2
Rockcastle River, Ky., between junctions with Horse Lick and Beech Creeks	67.5	40	1.29	0.6	2.5	1.7
Kentucky River, from Valley View, Ky. to near Frankfort, Ky.	159	100	1.28	0.9	7.1	1.9
E. Fork White River, from Tunnelton, Ind. to Portersville, Ind.	143	100	1.24	0.9	6.0	1.7
White River, from Petersburg, Ind. to Wabash River	74.9	140	1.23	0.7	5.0	1.7
N. Fork Salt Creek, from near Nashville, Ind. to Harrodsburg, Ind.	85.8	25	1.22	0.18	8.9	2.6
Wabash River, from Terre Haute, Ind. to Ohio River	332	220	1.17	0.9	11.2	1.6
Whitewater River, from Connersville, Ind. to Cedar Grove, Ind.	79.6	30	1.11	0.22	4.0	1.4
Wabash River, from Wabash, Ind. to Lafayette, Ind.	122	80	1.04	0.5	11.0	1.1

* As depicted on 1:24000 USGS topographic quadrangles.

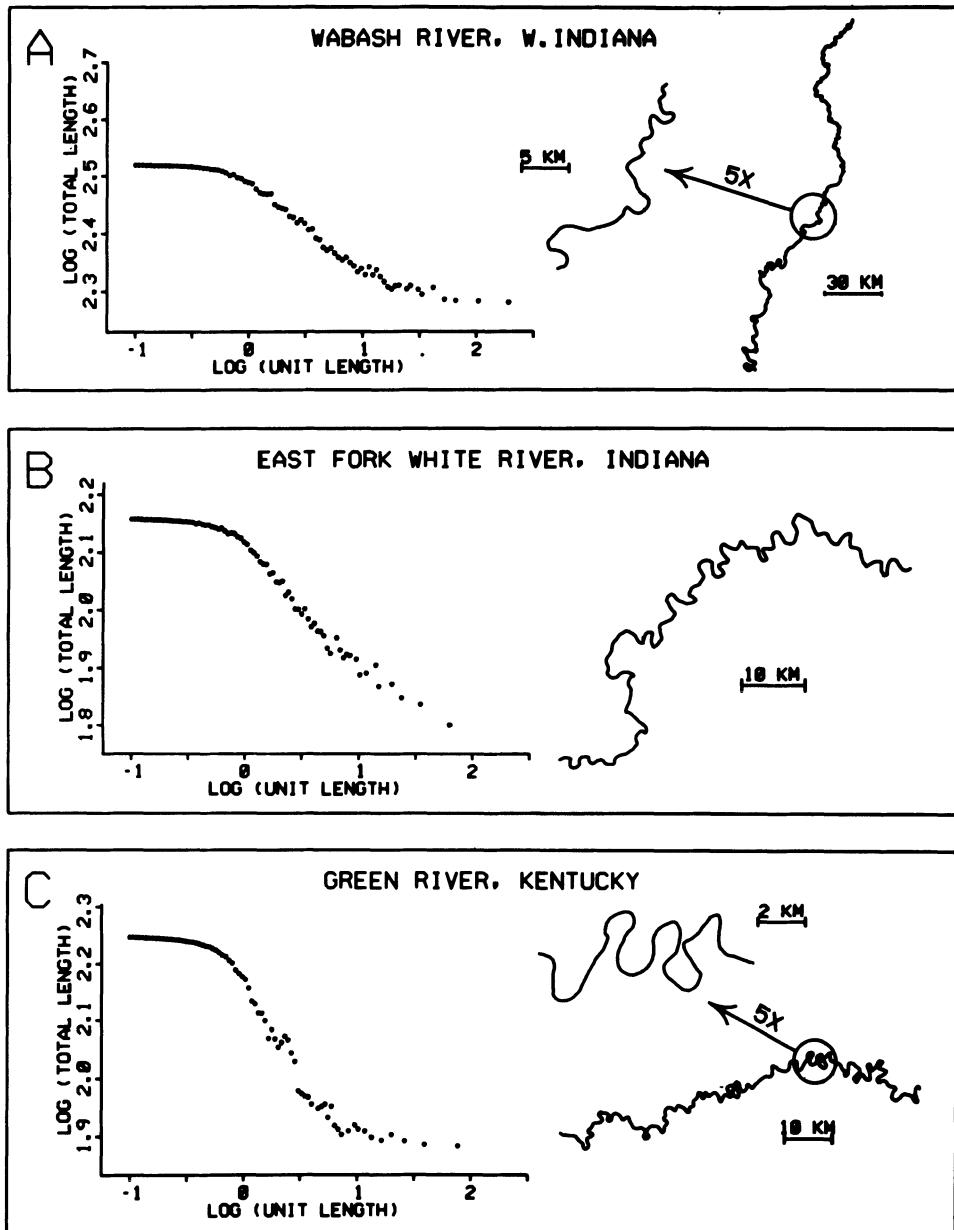


Fig. 2 (A, B, C)

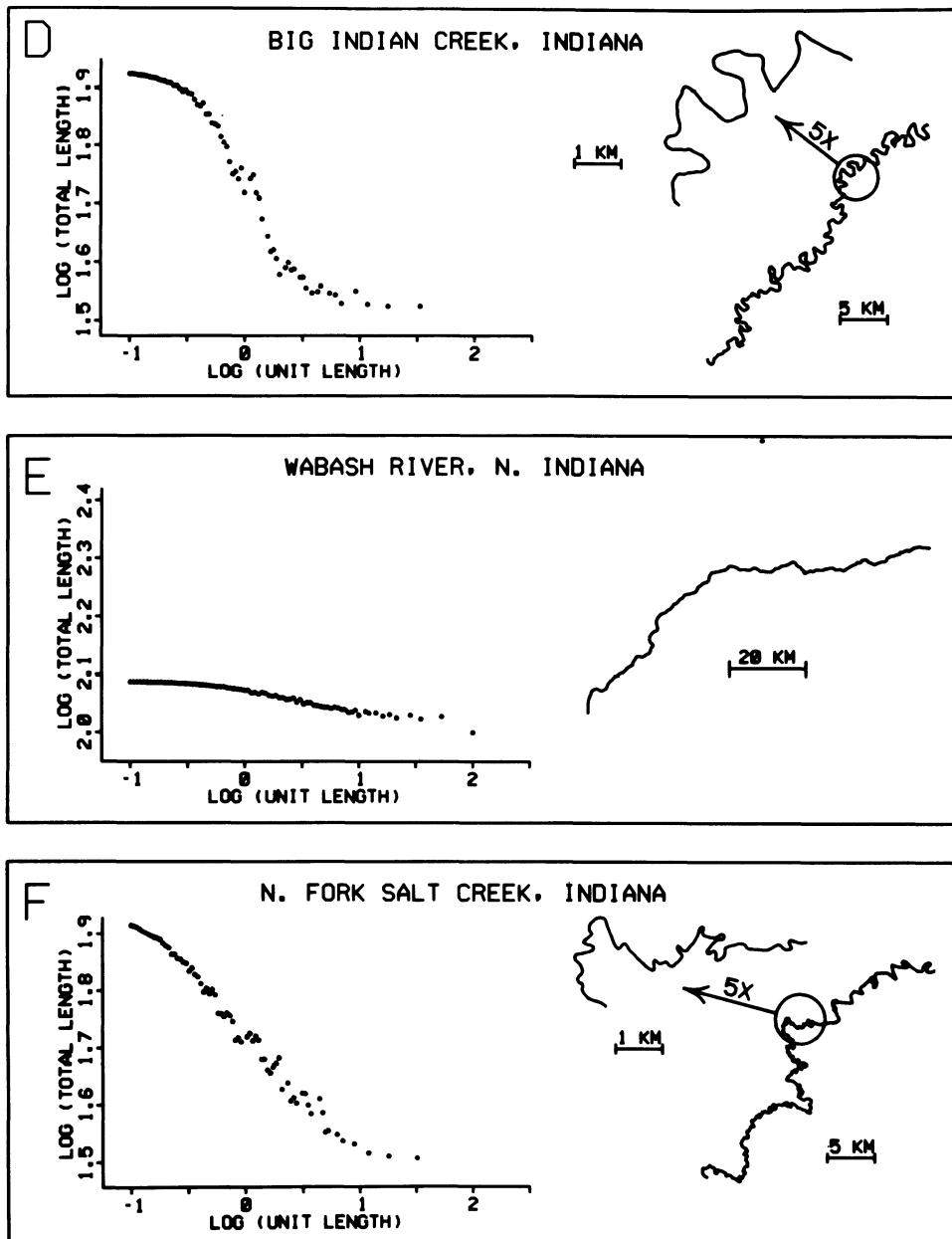


Fig. 2 (D, E, F)

Figure 2

Richardson plots and channel planforms for six stream segments. The channel conditions include free meanders (A), confined meanders of second order (B), incised meanders (C), "straight" channel (E), and angular, irregular planforms (D, F). Vertical scales are uniformly expanded, and all length values are expressed in kilometers.

Presumably if analysis were carried to a smaller scale, impractical with this data source, plots for these examples would continue to flatten progressively to the left.

The intermediate scale section of each plot approximates a straight line segment. In a few cases it might seem reasonable to treat this zone as merely the inflection point of a sigmoidal curve, but when applied to the full set of results the model is very unsatisfactory, while the straight line interpretation is reasonable in all cases. Geometry of channel planform is distinctively fractal when viewed at scales relevant to the most tortuous bends, whether those bends are classical meanders (Figure 2A) or highly irregular (Figure 2F). Significant disruptions indicating periodic signals are not evident in these scale ranges for most examples, but planforms of the Green River and Big Indian Creek (Figures 2C,D) do appear to have such regular components. It is interesting to note that neither is a case of free, active meandering. Instead both are incised, and the bends of Big Indian Creek are quite angular. At any rate, by visual comparison on the Richardson plots, the periodic signals in these examples remain secondary to the fractal component of channel wandering.

For scales of view approaching the scale of a whole segment length the analysis method involves only a few "steps" of "very large dividers," introducing a problem of small sample size. Chance hits or misses on large-scale channel bends can significantly affect the total lengths measured, resulting in a less reliable, more dispersed large scale end of the Richardson plot. Because of this and the limited scale ranges considered in this study, the results are insufficient to support or refute the indication given by others that river planforms are fractal shapes at scales much larger than those of meandering. It is sufficiently clear, however, that degree of channel wandering at these scales is-relatively low, as emphasized by a fairly distinct break in the Richardson plot slope. If river planforms at large scales are indeed fractal, it is reasonable to suspect additional inhomogeneities in fractal dimension with increasing scale, reflecting regional fracturing patterns and the like. A key advantage of the Richardson analysis in investigation of channel planforms is its potential to summarize the irregularity and tortuosity characteristics of channel traces over the full range of scales for which data is available.

Fractal Sinuosity

The commonly used measure of channel wandering is the sinuosity P , taken to be the ratio of real channel length to some general river course length; but actual definitions used for river course length vary widely. This is due partly to studies which employ P to atypical ends, such as work by SMART and SURKAN (1967) in which the use of a straight-line measure between stream segment end points is appropriate. However, in most applications P is intended to express channel wandering that is the result of stream action (one of the degrees of freedom of the stream system) rather than major changes in stream course due to other factors (external constraints on the stream system). The difficulty is that no objective

method of defining river course length is universally applicable. For example, use of a topographically defined valley length (SCHUMM, 1963) is awkward for work with incised meanders, and measures based on a series of line segments connecting channel bend inflection points (BRICE, 1964) lack utility for complex or very tortuous meanders.

The commonly understood, genetic meaning of sinuosity given above derives from a more basic observational conviction that the wandering of channel courses at particular scales, about one to two orders of magnitude larger than channel width, is qualitatively different from the wandering viewed at much larger scales. Results given here document that intermediate scale planform, with its enhanced wandering relative to large scale planform, is best understood as a fractal curve with consistent dimension that persists within distinct limits of scale (Table 1). In other words, the Richardson plots are showing clearly what we have meant by *sinuosity* all along. The scale region of enhanced wandering is a common characteristic of all stream segments studied, including ones with irregular patterns confined in valleys without significant floodplain and ones with "straight" patterns (Figure 2E), not just channels that can be recognized to meander.

Similarities in fractal analysis results, noted here for a variety of single-channel stream types, provide a natural basis for computation of the sinuosity P . The break in slope of the Richardson plot that is taken here as the boundary between intermediate and large scales of view also corresponds to a value of river course length L_c that accounts for large scale bends but none of the enhanced wandering discussed above (Figure 3). Actual length of the river channel L , taken as the

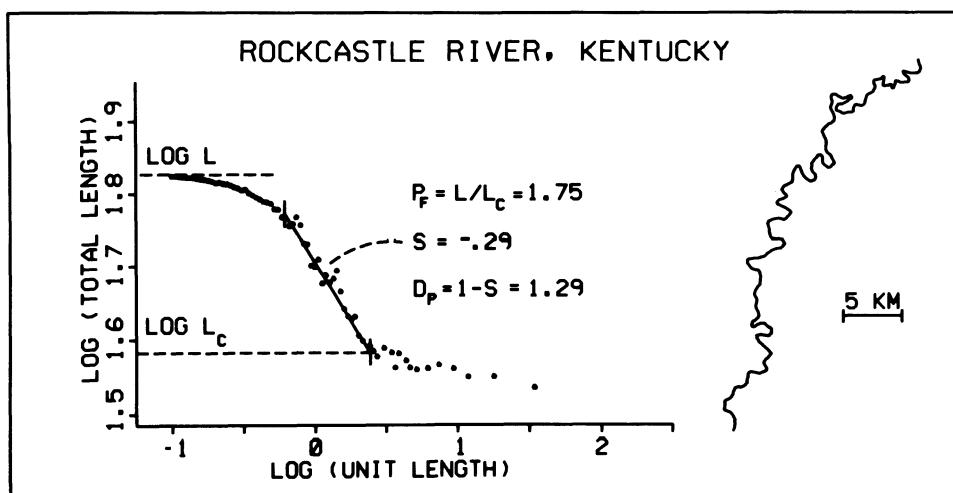


Figure 3

Definitions of values for quantitative analysis of channel wandering, shown on plot for planform of a segment of the Rockcastle River. Vertical scale is expanded, and all length values are expressed in kilometers.

small-scale asymptote on the Richardson plot or derived from other measurement, is divided by L_c to obtain a sinuosity value, for distinction termed P_f . This method is universally applicable for single-channel streams, because the only required information is a channel trace and no particular shape of channel bends is assumed. It is a fairly objective method and is robust in accommodating errors of interpretation. For example, there may be difficulty pinpointing the spot on Figure 2D where the lower section of the data plot first strongly diverges from the intermediate scale trend, but any of the reasonable choices of L_c will result only in a minor difference of sinuosity value.

The consistent fractal dimension value prevailing for intermediate scales of view, here termed "sinuosity fractal dimension" D_p (Figure 3), gives a further specification of river planform. It amounts to a measure of the additional channel wandering revealed by a unit increase in magnification of view, and provides a basis for direct comparison with the vast range of other natural phenomena with fractal characteristics. Upper and lower bounds of scale within which this single value of fractal dimension applies are typically separated by about an order of magnitude, but the extent of this range is quite variable from stream to stream (Table 1). These bounds provide appropriate keys to the characteristic scales of enhanced wandering, emphasizing as they do that concentration of the wandering upon some single wavelength is by far the exception, with the rule being a statistically self-similar distribution of channel bend sizes throughout the intermediate scale range.

Conclusion

Both the irregular and the structured aspects of single-channel river planforms can be accommodated by an approach based on fractal geometry. All planforms tested, regardless of place in traditional classification schemes, exhibit anomalously high degrees of channel wandering over distinct scale regions approximately one to two orders of magnitude greater than channel width. The planforms included in the study yielded fractal dimension values from 1.04 to 1.38 for these scale ranges.

Sinuosity and fractal dimension are related but not directly comparable measures of the degree of channel wandering. The Richardson plot shows clearly the bounds of scale within which the enhanced channel wandering exists and allows straightforward calculation of a sinuosity value based purely on planform shape.

Fractal analysis of this type leads to a more complete specification of planform at scales relevant to meandering than that provided by traditional sinuosity and wavelength values. In summarizing characteristics of channel wandering over a broad band of scales, this method should be useful in deriving information from river traces viewed at very large scales as well.

Acknowledgments

This research received funding from Ball State University, with computational support provided by the Intergraph Facility, College of Architecture. The author is deeply indebted to A. M. Snow for technical assistance in this project. The manuscript was improved following reviews by J. E. Adams and D. L. Turcotte.

REFERENCES

- BERQUIST, T. S., and SNOW, R. S. (1985), *Fractal Analysis of the Planforms of Rivers in Indiana and Kentucky* [abs], Geol. Soc. America Abstracts with Programs 17, p. 280.
- BRICE, J. C. (1964), *Channel Patterns and Terraces of the Loup River in Nebraska*, U.S. Geol. Survey Prof. Paper 422-D.
- CARSON, M. A., and LAPOINTE, M. F. (1983), *The Inherent Asymmetry of River Meander Planform*, J. Geology 91, 41-55.
- FERGUSON, R. I., *Meander migration: equilibrium and change*, In *River Channel Changes* (ed. Gregory, K. J.) (Wiley, New York 1977) pp. 235-248.
- HACK, J. T. (1957), *Studies of Longitudinal Stream Profiles in Virginia and Maryland*, U. S. Geol. Survey Prof. Paper 294-B.
- KNIGHTON, D., *Fluvial Forms and Processes* (Edward Arnold, London 1984).
- LANGBEIN, W. B., and LEOPOLD, L. B. (1966), *River Meanders—Theory of Minimum Variance*, U. S. Geol. Survey Prof. Paper 422-H.
- LEWIN, J., and BRINDLE, B. J., *Confined meanders*, In *River Channel Changes* (ed. Gregory, K. J.) (Wiley, New York 1977) pp. 221-233.
- MANDELBROT, B. B. (1967), *How Long Is the Coast of Britain?—Statistical Self-Similarity and Fractional Dimension*, Science 156, 636-638.
- MANDELBROT, B. B., *The Fractal Geometry of Nature* (Freeman and Co., New York 1983).
- RICHARDS, K., *Rivers: Form and Process in Alluvial Channels* (Methuen, New York 1982).
- RICHARDSON, L. F. (1961), *The Problem of Contiguity: An Appendix of Statistics of Deadly Quarrels*, General Systems Yearbook 6, 139-188.
- SCHUMM, S. A. (1963), *Sinuosity of Alluvial Rivers on the Great Plains*, Geol. Soc. America Bull. 74 1089-1100.
- SMART, J. S., and SURKAN, A. J. (1967), *The Relation between Mainstream Length and Area in Drainage Basins*, Water Resour. Research 3, 963-974.
- SPEIGHT, J. G. (1965), *Meander Spectra of the Angabunga River*, J. Hydrol. 3, 1-15.

(Received May 6, 1987, revised/accepted September 21, 1987)

The Fractal Geometry of Flow Paths in Natural Fractures in Rock and the Approach to Percolation

D. D. NOLTE¹, L. J. PYRAK-NOLTE² and N. G. W. COOK²

Abstract—The distributions of contact areas in single, natural fractures in quartz monzonite (Stripa granite) are found to have fractal dimensions which decrease from $D = 2.00$ to values near $D = 1.96$ as stress normal to the fractures is increased from 3 MPa up to 85 MPa. The effect of stress on fluid flow is studied in the same samples. Fluid transport through a fracture depends on two properties of the fracture void space geometry: the void aperture; and the tortuosity of the flow paths, determined through the distribution of contact area. Each of these quantities change under stress and contribute to changes observed in the flow rate. A general flow law is presented which separates these different effects. The effects of tortuosity on flow are largely governed by the proximity of the flow path distribution to a percolation threshold. A fractal model of correlated continuum percolation is presented which quantitatively reproduces the flow path geometries. The fractal dimension in this model is fit to the measured fractal dimensions of the flow systems to determine how far the flow systems are above the percolation threshold.

Key words: Fractals, fractures, fluid flow, percolation, rock mechanics, geohydrology.

1. Introduction

Whether for the recovery of oil, or for the isolation of nuclear or toxic wastes, it is essential to be able to predict the rates of flow and the flow patterns of fluids in bulk rock. Substantial research has been carried out on flow through porous media, for which permeability constants are well defined. However, the case for fractured rock is not nearly as well understood. In addition to the usually isotropic permeability of the bulk rock, directional permeability must also be assigned because of the flow through fracture networks (ROBINSON, 1983). Lately, there has been special interest in fracture networks caused by the need for nuclear waste isolation, especially for the case of isolation in impermeable, crystalline rock in

¹ Department of Physics, University of California, Berkeley, CA 94720, U.S.A. Now at AT&T Bell Laboratories, Crawfords Corner Rd., Holmdel, New Jersey 07733, U.S.A.

² Department of Material Science and Mineral Engineering, University of California, Berkeley, CA 94720, U.S.A.

which fluid flow occurs exclusively through fractures. Understanding the mechanisms of fluid flow through single fractures is the first step in constructing macroscopic models for fluid migration in the ground.

Flow in fractures has often been modeled as flow between parallel plates. The dependence of the flow on the separation between the two plates (this separation is called the aperture) can be expressed analytically for which the fluid rate varies as the cubic power of the aperture:

$$Q = \frac{Wb^3}{12\mu} \frac{dp}{dx} \quad (1)$$

where Q is the flow in units of volume per time, dp/dx is the gradient of the hydrostatic pressure, μ is the viscosity, W is the width of the fracture in the direction perpendicular to the flow direction, and the variable b is the aperture. This cubic law appears to be applicable for fluid flow through loosely mated, open fractures (IWAI, 1976) as well as for fractures with high correlation between the two surfaces (ENGELDER and SCHOLZ, 1981; WITHERSPOON *et al.*, 1980). Theoretical investigations, however, (TSANG and WITHERSPOON, 1981; BROWN, 1987) have predicted deviations from the cubic law. Experimental evidence (GALE and RAVEN, 1980) also has suggested that natural fractures (*i.e.*, fractures which have undergone mechanical or chemical alteration in the field) deviate from the parallel plate model. RAVEN and GALE (1985) related flow to the measured closure of fractures under uniaxial stress applied normal to the fracture and found a power-law-relationship between the flow and the fracture aperture (called mechanical aperture) derived from the deformation data, but with exponents much greater than cubic. We also relate flow to the measured closure of fractures under stress and find a similar power-law-relationship, again with exponents much larger than cubic. This dependence of flow on large exponents of mechanical aperture (even at low stresses) is one of the outstanding problems of fluid flow through fractures.

Clearly, the influence of changing aperture on fluid flow is of primary importance, but in addition to changes in fracture aperture, changes in fracture contact area occur when stress is applied normal to the fracture. Increasing contact area produces barriers to flow, increasing the "tortuosity" of the flow paths through the fracture and thereby increasing the resistance to flow. This changing contact area is implicitly related to changing aperture. To understand the large exponents, therefore, it is first necessary to separate the effects of changing aperture and changing contact area and explicitly relate flow to contact area as well as aperture. WALSH (1981) derived an expression for fluid flow through fractures as a function of applied stress which included the cubic aperture term as well as a tortuosity term related to the contact area. He argued that the tortuosity term had negligible effect compared to the aperture term, except at high stresses. On the other hand, computer simulations on two-dimensional resistor networks (TSANG, 1984) indicate that tortuosity can substantially reduce the flow rates below those predicted from the cubic law for tight fractures.

Questions about the effect of changing contact area and increasing tortuosity on fluid flow fall under the realm of percolation theory. This theory takes as its input the results of the topography of the flow paths, including both aperture and contact area distributions. The main goal of this paper is to directly image the changing contact area under applied stress and extract the effect of changing contact area alone on the fluid flow rate by relying on the results of two-dimensional percolation theory. In Section 2, we present direct measurements of flow path patterns in natural fractures. We also provide data of the fluid flow rate through the fracture as a function of the fracture closure in order to correlate flow rate with the flow path patterns. The flow path geometry is defined through the distribution of contact area, which is describable by fractal dimensions, discussed in Section 3. The flow area fractions and fractal dimensions are used to construct a two-dimensional correlated continuum-percolation model in Section 4.

2. Experimental Data

We studied the deformation and flow properties of fractures in three separate core sections of quartz monzonite (Stripa granite (OLKIEWICZ *et al.*, 1979)) which were 52 mm in diameter by 77 mm in height. These cores were obtained from extensometer holes from a waste isolation experiment in a drift in Stripa, Sweden, 340 meters below the surface. The fractures are orthogonal to the core axis. The core samples are labeled E30, E32, and E35, corresponding to the extensometer holes from which the cores were derived. The closure of the fractures under applied stress was measured by linear variable differential transformers (LVDT's). The mechanical deformation data for the samples are given in Figure 1. The stiffnesses of the fractures are defined from the inverse tangents of the deformation curves. At 20 MPa normal stress, the stiffnesses are: E35, 3×10^{12} Pa/m; E30, 7×10^{12} Pa/m; and E32, 15×10^{12} Pa/m. The stiff fracture in E32 deforms to a value of $d = 5.5$ microns, while the compliant fracture in E30 deforms considerably more to $d = 9.5$ microns at normal stresses of 85 MPa. The fluid flow data through the fractures as a function of the fracture closure are presented in Figure 2. There is a clear correlation between fracture stiffness and fluid flow in that the stiffer fractures support less flow. Under even the highest stresses, there remains an irreducible flow, denoted as Q_∞ . By analogy with the cubic law described by eq. (1), we fit the flow data to the equation

$$Q = Q_\infty + C(b_{\text{mech}})^n \quad (2)$$

where $b_{\text{mech}} = (d_{\text{max}} - d)$ is called the mechanical aperture. The quantity d is the measured closure for a given applied load, and n is the flow exponent. The maximum mechanical closure of the fracture is d_{max} , which can be understood as the average aperture which results when there are only three points of contact between the two fracture surfaces at zero normal load. The values for d_{max} are fit from the

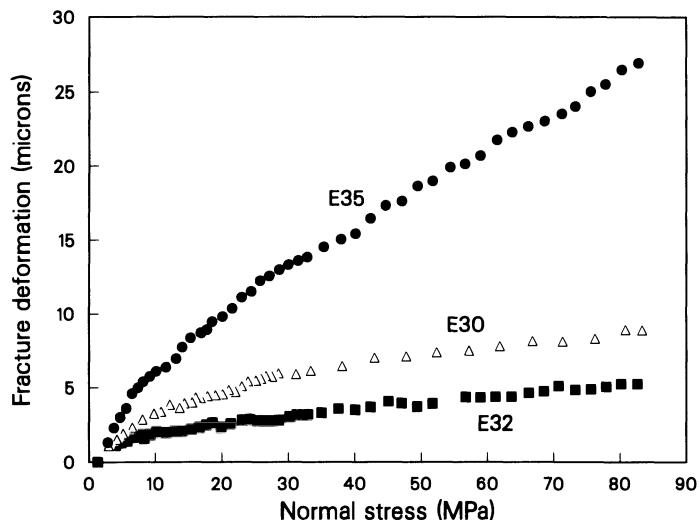


Figure 1

Mechanical closures of three fractures as functions of applied stress. The stiffnesses of the fractures are equal to the tangent slope of the deformation curves.

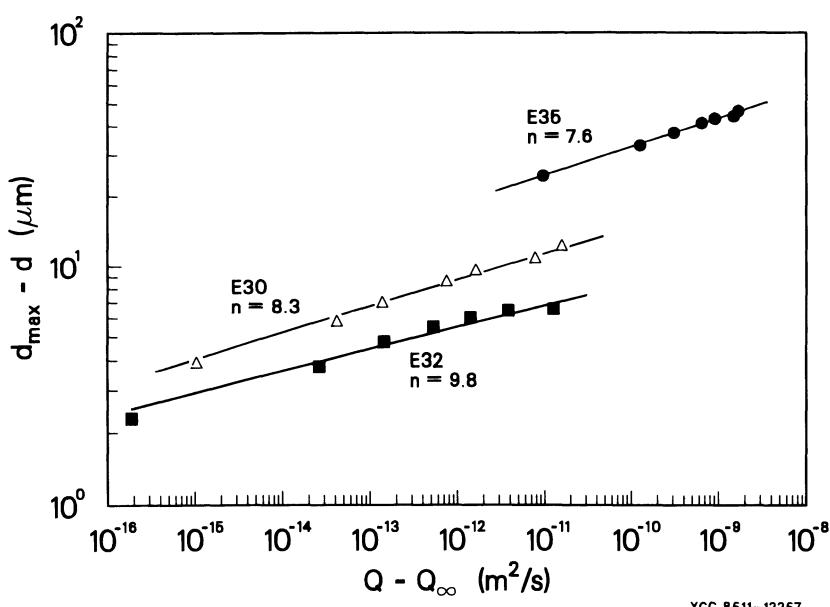


Figure 2

Hydraulic flow per unit head as a function of mechanical fracture closure for the three fractures. The compliant fractures support greater flow than the stiff fracture, as expected. The irreducible flow, Q_∞ , is that flow which remains when the fracture is hydraulically closed. The solid lines fit eq. (2) to the data for the given exponents, n . All exponents are much greater than cubic.

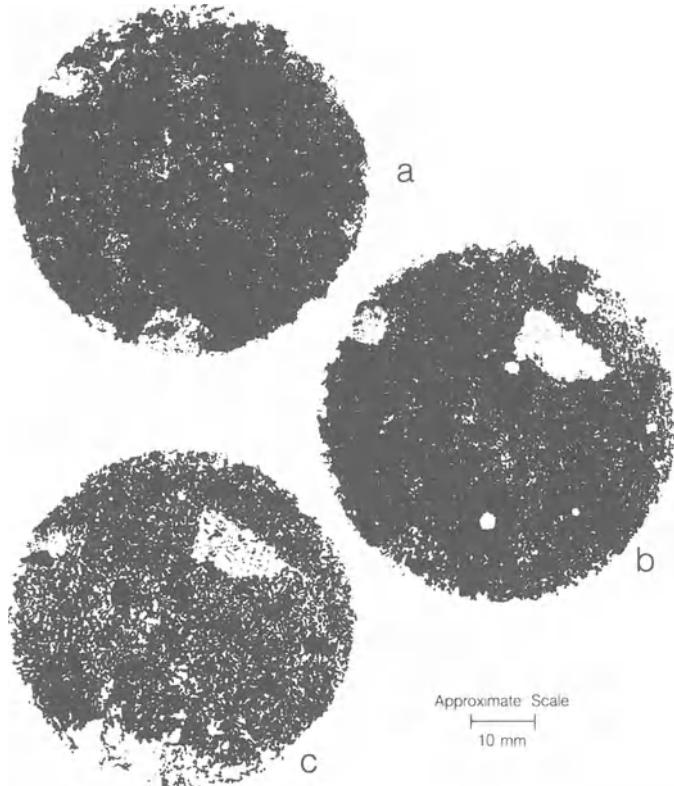
data: for E35 $d_{\max} = 46$ microns; for E30 $d_{\max} = 12.5$ microns; and for E32 $d_{\max} = 6.6$ microns. The quantity Q_∞ is the irreducible flow, defined as that flow which remains when the maximum closure d_{\max} has been reached. We consider this flow to result from permanent conduits carved in the fracture surface which are virtually unaffected by applied stress. This flow therefore occurs in parallel with the flow occurring over the fracture surfaces. Even parallel plates could have irreducible flow if there were small interconnected grooves in the plate surfaces. Experimental data for flow between ground quartz surfaces under high stress does, in fact, show this (ENGELDER and SCHOLZ, 1981).

Although eq. (2) is reminiscent of the cubic law, there is no reason, *a priori*, to believe that the exponent n should have a value near 3. Part of the difficulty in applying the cubic law to fracture lies in the problem of how to define an appropriate average aperture. The mechanical aperture defined above for the mechanical closure of a fracture under applied stress is likely to differ from an average hydraulic aperture defined by a weighted average over directly measured aperture profiles. In fact, mechanical aperture may not even be linearly proportional to hydraulic aperture (PYRAK-NOLTE *et al.*, 1987), and this could influence the exponent n measured for mechanical closure. Despite these difficulties, we emphasize and use mechanical aperture over aperture distributions obtained from profile or thin-section measurements because of the ease with which this data is obtained (requiring only a single deformation experiment compared to multiple profiles or thin-sections). Furthermore, the mechanical closure of a fracture is measured unambiguously, while aperture profiles require complicated statistical analyses in order to predict their effect on fluid flow. The use of the empirical eq. (1) should therefore simplify the process of fracture characterization, and speed up the prediction of flow properties in the field based on core samples measured in the laboratory. What remains is to explain the exponent n . The values of n in eq. (2), shown for different fractures in Figure 2, are found to lie near $n \sim 8$, much larger than cubic or even quartic (which holds for flow in a tube). Such strong deviations from the cubic law imply that the flow properties of these natural fractures must be heavily influenced by the effects of the flow path topology.

Despite the fact that changes in the contact area as a function of stress can have a pronounced effect on the flow properties of the fractures, fracture contact area has been an elusive and very difficult parameter to measure and image. Various indirect means have been tried previously to observe the contact area (DUNCAN and HANCOCK, 1966; IWAI, 1976; BANDIS *et al.*, 1983), with varying degrees of success. We have used a new technique (PYRAK-NOLTE *et al.*, 1987) developed to make a direct measurement of the contact area of fractures. This technique involves injecting molten Wood's metal into fractures which are initially evacuated. During the injection, stress is maintained on the fracture through a servo-controlled, stiff-testing machine. The pore pressure of the molten metal was maintained at 2 MPa through the injection and cooling of the cast in the fracture. The effective

stress on the fracture is calculated by subtracting this confining pressure from the applied load. After the metal cools, the fracture is taken apart, revealing a direct metal cast of the fracture void space. The metal adheres to either one or the other surface of the fracture, tearing along lines of contact between the rock surfaces. It was observed that the metal on one surface does not overlap with the metal on the other at the same location. The damage inflicted on the metal cast by tearing was observed under an electron scanning microscope to be limited to only several microns at most. In fact, SEM micrographs of the separate surfaces, probing sections 0.25 mm on a side, could be superposed with high accuracy, fitting together with minimal overlaps or gaps between the metal patterns on the two surfaces.

Images of the metal casts of the void spaces for E30 and E32 under 3 MPa, 33 MPa, and 85 MPa are shown in Figure 3. A computer enhanced composite of



XBL 8610-12615

Figure 3.1

Contact area composite images for sample E30 for (a) 3 MPa, (b) 33 MPa, and (c) 85 MPa. The black portion is the flow path geometry; the white areas are the contact area. The resolution of the patterns is about 3% of the diameter.

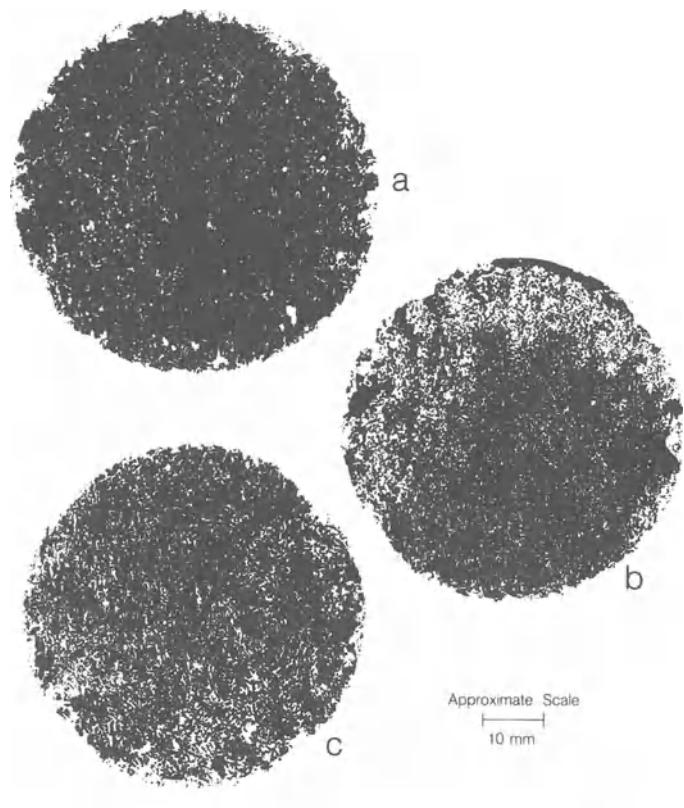
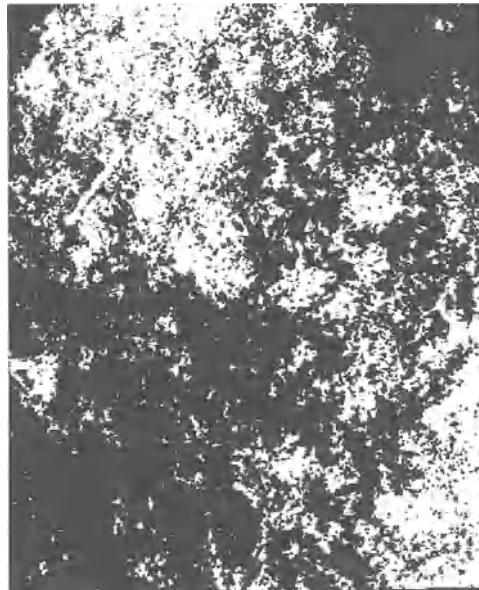


Figure 3.2
Contact area composite images for sample E32 for (a) 3 MPa, (b) 33 MPa, and (c) 85 MPa.

scanning electron micrographs of the two fracture surfaces of sample E30 under 85 MPa at 25 times magnification is shown in Figure 4. This figure represents an area 2 mm by 4 mm of the fracture. The error in this figure, caused by possible misalignment of the SEM micrographs or damage to the metal cast, is less than 1%. The hydraulic void space is black, while the contact area is white. As the stress is increased, the amount of contact area increases, producing barriers to flow which force the flow paths to become more tortuous. Large, single areas of contact are observable for the compliant fracture in E30, but are absent for the stiff fracture in E32. There is an important connection between the observed contact area and the surface tension of the injected metal. Specifically, a fluid with a given surface tension and at a given confining pressure can only enter into apertures larger than a certain size. This, in fact, is the principle on which mercury porosimetry is based: by injecting mercury into porous or fractured media under different confining pressure, one can measure pore and void space volume. Calculations for molten Wood's



XBL 8611-4760

Figure 4

A computer enhanced composite image of SEM micrographs of E30 at 85 MPa under 25 times magnification. The area of the figure is 2 mm by 4 mm. The fractal dimension of this section is $D = 1.93$.

metal predict that the metal, under 2 MPa confining pressure, can penetrate apertures on the size of tenths of a micron. In the SEM micrographs, filaments of the metal as small as 0.2 microns wide were observed. We consider this to be the smallest aperture which the metal can penetrate. The ability of the metal to penetrate small apertures is further borne out by the observation of areas in Figures 3 and 4 which are covered by metal, but which are surrounded entirely by contact area (at least to the resolution of the figures). This is a good indication that we are sampling nearly the entire void space of the fracture. The contact area fractions were measured using a Zeiss image analyzer. The resulting area fractions of void space are presented in Figure 5 as a function of stress. The error bars on the figure represent the uncertainties inherent in the image analysis, related to the scaling properties of the flow paths. In fact, fractal structures (we show later in this paper that the flow paths are fractal) are not required to have well-defined areas, rather the areas may depend on the resolution (or lower cut-off) of the pattern. This will be discussed in more detail later. For our resolution (3% of the diameter for the large composites), the void area fraction of the stiff fracture approaches a constant value at fairly low stress, while the void area fraction of the more compliant fracture continues to decrease up to 85 MPa. The void area fraction of flow of the stiff fracture is substantially smaller than the area fraction of the compliant fracture

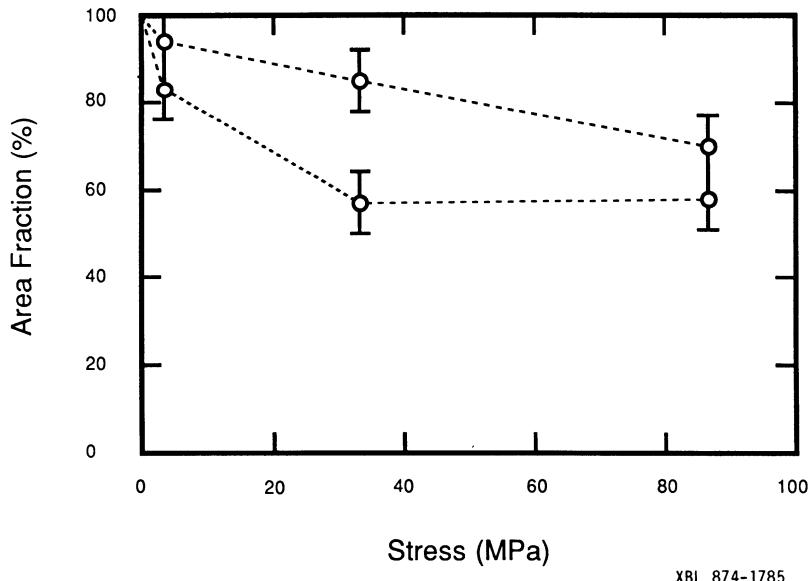


Figure 5

XBL 874-1785

The flow area fraction as a function of stress for the two samples at the resolution given in Figure 3. The area fraction is usually one of the few directly measurable parameters which can enter into percolation theory analysis. For fractal patterns, however, the area fraction may not be uniquely defined.

despite the absence of the large, bulky areas of contact which are present in the pattern of the compliant fracture. This indicates that the areas of contact of the stiff fracture are considerably more stretched out and filamentary than the areas of contact of the compliant fracture. These differences between the two fractures are likely to be a consequence of the low aspect angles of the void spaces in the stiff fracture compared to the high aspect angles in the compliant fracture. It is furthermore important to point out that the stiff fracture has greater contact area than the compliant fracture, and has correspondingly less flow.

Perhaps the most interesting features of the flow path geometries, and the most relevant to the flow properties of the fracture, are the distributions and shapes of the contact areas and flow paths. While the patterns certainly appear random, they have an unmistakable scaling structure which implies they might be fractal. This possibility is explored in the next section, and the patterns are, in fact, found to have fractal dimensions which vary with the applied stress.

3. Fractal Geometry of the Contact Area

There has been strong interest in the ability of fractal geometry to describe many of the characteristics of seemingly structureless patterns (MANDELBROT, 1983).

Much of this interest has come after the realization that an impressive number of random systems exhibit scale invariance, also known as self-similarity; that is, certain parameters describing the system (or pattern) remain the same regardless of the scale of magnification. Scale invariance lies at the heart of the notion of fractal dimension.

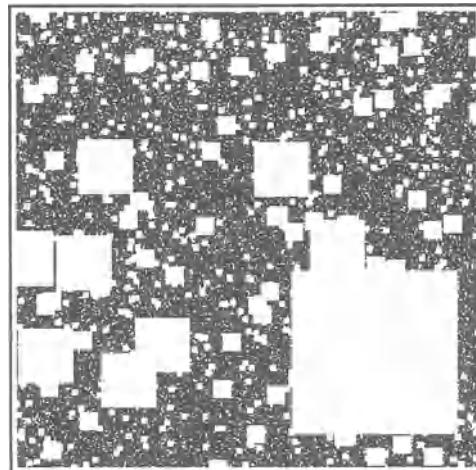
For a random pattern to have scale invariance, the distribution of the sizes of the features which define the pattern must vary as a power law with some scale, b : $P(b) \propto b^{-(D+1)}$. The exponent D is the fractal dimension of the object. As such, the concept of fractal dimension is simply the consequence of power-law statistics governing size distributions. Power-law statistics ensure scale invariance. Invariance under transformation is a powerful and recurring concept in physics describing widely different phenomena. In many of these phenomena, invariance under scale transformation offers valuable insight into the physical origins of the phenomena and often provides for the use of powerful analytic tools for describing their structure. For instance, the scale invariance of fractal objects allows the analytic results of renormalization group theory (DOMB and GREEN, 1976; MA, 1973) to be used.

In application, real systems rarely possess the same scaling properties for all scales. Namely, there are scales above or below which the scaling properties change. These scales are called cut-offs. Often the cut-offs can carry as much information about the physical processes creating the pattern as the scaling properties of the patterns. Typically, when a fractal dimension is assigned to a pattern, this dimension is only valid for scales above a lower cut-off and below an upper cut-off. In fact, measurement of certain fundamental properties, such as contact area, depends directly on the cut-offs. Also, the measurement of the fractal dimension can be influenced by cut-offs when the measurement scale approaches the cut-off scale. For these reasons, particular attention must be paid to the limits of the regimes of scaling when attempting to define the physical properties of a pattern.

With the power of fractal description comes considerable complexity. Fractal objects take on a tremendous variety of forms, and sometimes several fractal dimensions can be defined for the same object. For example, a random percolation network (in two dimensions) at its critical percolation threshold has a fractal dimension of $D = 1.89$ (AHARONY, 1984; STANLEY, 1977). At the same time, the fractal dimension of the backbone of the percolating cluster (the backbone is defined as that part of the cluster which carries flow, *i.e.*, dead-ends are excluded) is $D = 1.59$ (AHARONY, 1984; SCHLIFER *et al.*, 1979). Similarly, in the case of the fracture flow paths, many fractal models can be used to describe the void-space topology. Because the main goal of this paper is to understand the effects of the two-dimensional contact area on flow through fractures, we only consider two-dimensional fractal models, neglecting for the moment possible fractal models for aperture distributions. Although several different models could be considered, we find one fractal model particularly relevant and easy to relate to the observed flow

path patterns: the fracture contact areas can be viewed as random holes, or tremas (MANDELBROT, 1983), which puncture a conductive sheet. Therefore the flow paths can be modeled as lying on a random Sierpinski carpet.

A random Sierpinski carpet, shown in Figure 6, is constructed iteratively by removing successively smaller squares (or tremas) from the original black square. In the carpet shown, 8 out of 9 sub-squares, of scale $b = 1/3$, remain at each level. This gives the carpet the approximate fractal dimension $D = \ln 8/\ln 3 = 1.89$. The fractal dimension of a Sierpinski carpet can be measured, in principle, by counting the number of tremas (single areas of contact) with area, a , larger than some set value, A . In practice, the counting is carried out by superposing grids with successively smaller spacings and counting the number of grid squares at each level which are occluded more than 50% by a single trema. The fractal dimensions are derived from the slope of $\ln(b^2 - N)$ vs. $\ln b$, where b is the scale size of the grid and N is the number of squares occluded for that scale b . The quantity $(b^2 - N)$ is the number of grid squares which remain uncut by tremas at this scale size. The fractal dimensions for the large composites in Figure 3 of samples E30 and E32 are shown in Table 1. The largest grid size was 32 mm, corresponding to the scale $b = 1$. The resolution of this analysis for these large composites is 1 mm, corresponding to scale $b = 32$. Upper and lower cut-offs are defined as lengths above and below which the pattern ceases to scale, and therefore appears homogeneous. SEM micrographs of the flow patterns under 200 times magnification show a clear



XBL 8611-4762

Figure 6

A random Sierpinski Carpet with the approximate fractal dimension $D = \ln 8/\ln 3 = 1.89$. In the limit of true scaling, the black portion vanishes.

Table 1

*Measurement of fractal dimension of flow area. Area fraction ($b^2 - N$) as a function of scale size**

Scale l/b	E30			E32		
	3 MPa	33 MPa	85 MPa	3 MPa	33 MPa	85 MPa
1 sample size	1	1	1	1	1	1
1/2	4	4	4	4	4	4
1/4 cut-off	16	14	14	16	16	16
1/8	64	59	59	64	59	60
1/16	253	233	225	255	234	229
1/32 resolution	1016	922	894	1014	916	925
D =	1.996	1.975	1.964	1.997	1.955	1.955 ± 0.01

* Scale $b = 1$ corresponds to 32 mm.

change in the void-space morphology. This is not surprising because at these length scales the surface tension of the injected metal prevents the metal from penetrating void spaces smaller than $0.2 \mu\text{m}$. Also, the grain size of the crystalline quartz monzonite is comparable in size to, or larger than, these length scales. Therefore the lower cut-off of the flow pattern may be expected to lie somewhat above 5 microns. Similarly, an upper cut-off occurs at sizes comparable to the size of the sample. This change in scaling behavior can be seen in Table 1 to occur at a grid size of 4 mm. For grid sizes larger than this, the pattern appears homogeneous and has the corresponding fractal dimension $D = 2$. For grid sizes smaller than this, the pattern scales with a characteristic fractal dimension less than 2. The change in scaling between 4 mm and 8 mm is in agreement with work that has found correlation lengths on the order of millimeters between the two surfaces of a natural fracture (BROWN *et al.*, 1986). This upper cut-off may be an artifact of the experiment related to sample size. To test this, the scaling properties of metal injection in larger core samples should be investigated.

The fractal dimensions of the samples (in the regime in which they scale) are plotted in Figure 7 as a function of stress. The fractal dimensions clearly decrease with increasing stress, and appear to approach a value near $D = 1.95$. It is likely that the flowpath geometry in the fractures is only fractal to lowest order, and that deviations from true scaling behavior may be anticipated. As a check of the fractal dimensions calculated for the large composites (down to 1 mm length scale), the fractal dimension was calculated for the SEM micrograph shown in Figure 4 of Sample E30 at 85 MPa for 25 times magnification. We investigated a section of this micrograph 2 mm on a side with a smallest resolution of 60 microns. These sizes are substantially below the upper cut-off of 4 mm and therefore the fractal dimension

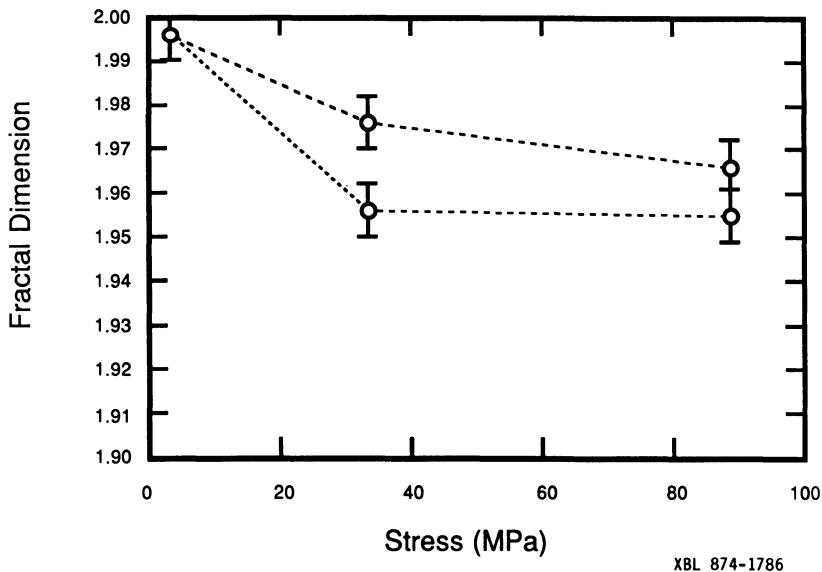


Figure 7

The fractal dimensions of the flow paths as functions of stress. The fractal dimensions of both fractures depend nonlinearly on stress and appear to asymptote to values near $D = 1.95$.

at this scale should not be influenced by the cut-off. The results of the measurement are shown in Table 2. The fractal dimension of this section was measured to be $D = 1.93$. This value is smaller than the value $D = 1.96$ measured for the same sample at lower resolution. This error may be due to simple fluctuations from site to site on the fracture, but may also be an indication that the fractal dimensions calculated for the large composites are artificially increased by their proximity to the

Table 2

*Measurement of fractal dimension of SEM micrograph.
Area fraction as a function of scale size**

Scale $1/b$	$(b^2 - N)$
1	1
1/2	3
1/4	13
1/8	50
1/16	193
1/32	741
$D =$	1.93 ± 0.01

* Scale $b = 1$ corresponds to 2 mm.

upper cut-off. However, the deviation is small, and to lowest order the flow pattern does appear to scale.

Fluid flow in a single fracture has a complicated topology imbedded in three dimensions. Any attempt to fully understand the flow properties of this topology must therefore be based on fully three-dimensional models. While three-dimensional information concerning the fracture surface may be obtainable through stereo projection or profilometer measurements, currently no theory (other than numerical computation) is available which can use such information. Therefore attempts to define this topology must resort to projections onto lower dimensions. For instance, our flow path data is the 2-d projection of this 3-d topology, while the effects of changing aperture are contained in measurements of 1-d displacements across the fracture. The simplest model which establishes the flow process in three dimensions multiplies the 2-d flow area by the 1-d average aperture. This approximation is equivalent to combining 2-d percolation theory with the cubic law. Such a model must be considered as incomplete. For example, the flow-path data is presented as a binary process in which regions either support flow or not. In fact, the flow constitutes a continuum, and the flow area fraction is a function of the threshold. In our experiment, that threshold is set ultimately by the surface tension of the injected molten metal. In addition, the assumption of an average aperture can grossly oversimplify the effects of changing aperture distributions. Changes in the fracture aperture, even at relatively low stresses, can dramatically alter the critical flows paths (defined as those paths which carry the bulk of the flow), changing the flow substantially more than would be expected for the case of flow between parallel plates. Despite these difficulties, the simple model provides a starting point for deriving a general flow equation whose parameters can be established empirically. To begin, it is necessary to separate the two-dimensional flow properties from the effects of changing aperture. This is achieved in the next section.

4. The Approach to Percolation

In fitting the flow data of Figure 3 to eq. (2), all effects of the changing flow path topology were implicitly included in the fracture closure. We now consider the closure and the flow area fraction separately. To do this, we discard for the moment the flow continuum, and consider only a binary process in which flow is either present or not. The effects of 2-d flow path topology on flow properties can be largely understood from percolation theory (KIRKPATRICK, 1973; STAUFFER, 1980; ESSAM, 1980). Standard percolation theory is performed on a lattice of sites connected by bonds. The number of bonds entering a site is called the coordination of the lattice. The lattices can be two-dimensional or three-dimensional, or have even larger integer dimensions. The two basic forms of percolation models are site-percolation and bond-percolation. In site percolation, the sites are occupied

with a probability, p_s . No flow can occur through an unoccupied site. In bond percolation, the bonds are occupied with a probability p_b , and no flow can occur through an unoccupied bond. The fundamental property of a random flow system is the existence of a critical probability, p_c , which defines the percolation threshold. For occupation probabilities below the critical probability, no connected path exists through the random network and the conductivity is zero. At the percolation threshold, only a single percolating cluster exists. This percolating cluster has a fractal dimension of $D = 1.89$. For values of the occupation probability increasing above the critical density more connected paths can be found, and the conductivity increases sharply. The critical probability is a function of the lattice dimension, and also of the specific model (site vs. bond). The qualitative trend of the conductivity as a function of density is shown in Figure 8. The power of percolation theory comes from its ability to define the critical threshold parameter (PIKE and SEAGER, 1974), as well as the functional form of the conductivity near the percolation

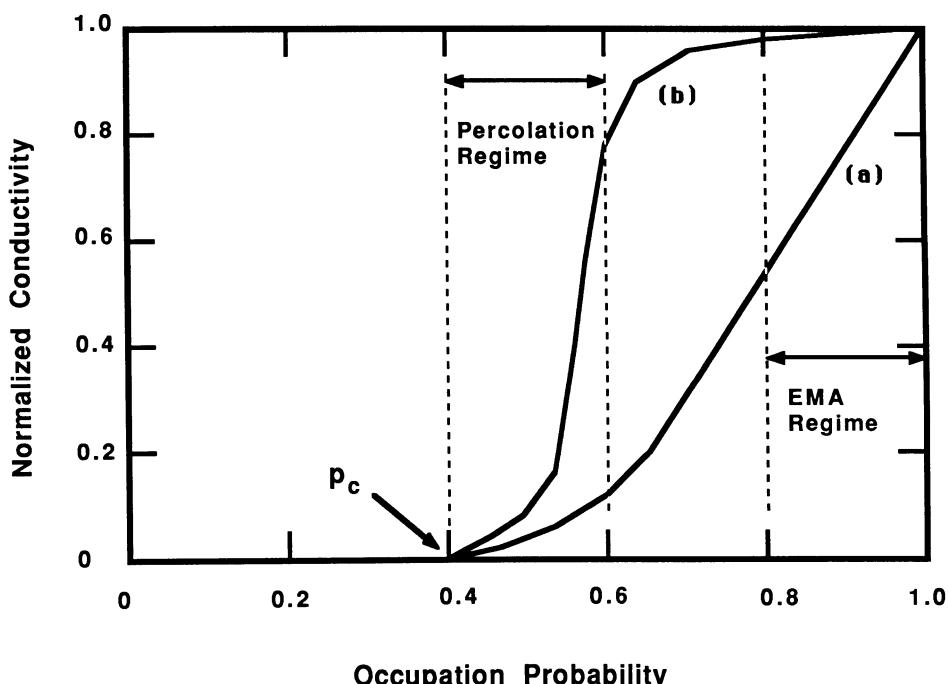


Figure 8

Conductivity as a function of the occupation probability for characteristic percolation systems. The curves represent the behavior for binary aperture distribution (a) and for continuous aperture distribution (b) (from SEAGER and PIKE, 1974). The region slightly above the percolation threshold, p_c , is the percolation regime in which the conductivity is described by a power law with an exponent, t . The region near unit occupation probability is the regime of the effective medium approximation (EMA) in which the conductivity depends linearly on the occupation probability p .

threshold. Because of scale invariance at the percolation threshold, the conductivity slightly above the critical threshold obeys a power-law-relation as a function of density: $\sigma \propto (p - p_c)^t$, where p is the flow path density, p_c is the critical density, and t is the conductivity exponent. As the system moves far above the percolation threshold, the random flow network can be characterized by applying the effective medium approximation (EMA) (KIRKPATRICK, 1973) from which a homogeneous conductivity can be defined for the system. The EMA region of Figure 8 occurs near unit density.

Under the application of stress, the fracture contact area increases, presenting barriers to flow and forcing the flow paths to take on increasingly tortuous routes. We first consider the effective medium approximation and its results for a general flow law as a function of stress. We then turn to percolation in a continuum in order to model the flow paths observed empirically. Developing a model which reproduces the qualitative and quantitative features of the experimentally observed contact areas should provide insight into the interplay between fracture morphology and flow path geometry. But most importantly, such a model can be used to quantitatively estimate the critical threshold parameter for flow through the fracture.

A. Effective Medium Approximation and Flow

The effective medium approximation (EMA) was developed to describe flow through a disordered system (KIRKPATRICK, 1973). In the EMA, the potential field in a random flow network is broken into two parts: a uniformly decreasing macroscopic potential field, and a random local field that averages to zero in a sufficiently large region. The conductance of the medium is set to a constant value such that the uniform macroscopic field is reproduced. The specific details of how the conductance varies as barriers to flow are introduced into a flow network depends on the details of the network. Different results are obtained depending on whether sites or bonds in a lattice are occupied (site- and bond-percolation), and depending on the coordination of the underlying lattice (number of connected neighbors), as well as the dimensionality (two dimensions or three), and on the details of the aperture distribution (SEAGER and PIKE, 1974). However, regardless of the network model, all the conductances derived from the EMA vary linearly with the occupation density (BÖTTGER and BRYSKIN, 1985). All that changes from model to model is the slope of the conductance.

The flow eq. (2) can now be expanded for the case when $AF \sim 1$ to include the effect of changing area fraction:

$$Q - Q_\infty = C \cdot (d_{\max} - d)^m \cdot (r \cdot AF + (1 - r)) \quad (3)$$

where the second term in parentheses is the two-dimensional normalized conductance which depends linearly on the occupation probability (total flow area fraction, AF) with the slope r . The quantity $(1 - r)$ is the y -intercept, C is an

arbitrary constant, and Q_∞ is the irreducible flow. The physical content of the above equation lies mostly in the exponent, m , and in the slope, r . We have insufficient data to define the functional form of the flow area fraction in order to fit either of these parameters accurately. However, the overall effect of including the area fraction in eq. (3) is to reduce the values of the exponents m from the values of n cited before. Rough estimates of m , however, indicate that the exponents are only reduced from $n \approx 8$ in eq. (2) to $m \approx 7$ in the above equation. In fact, the effective medium theory neglects the true three-dimensional character of the flow topology. The large exponents suggest that critical flow paths are changing more dramatically, even at low stresses, than for flow between parallel plates. In addition, it is clear from considering the tortuous flow paths in Figure 4 that under the highest stresses the flow paths must be approaching the percolation threshold. Therefore to characterize the flow properties of the fractures at all stresses it is necessary to consider the results of percolation theory.

B. Continuum Percolation

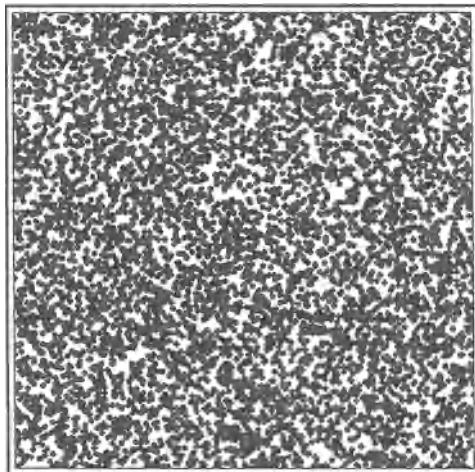
In order to estimate the effects of percolation, it is essential to know the critical area fraction for percolation. The most realistic percolation model should certainly be a continuum model (PIKE and SEAGER, 1974; ELAM *et al.*, 1984; PHANI and DHAR, 1984) because there is no physical basis for assuming an underlying lattice for the fracture. Continuum percolation is a fairly recent extension of the standard lattice percolation models, and therefore its properties have not been as extensively studied. However, continuum percolation is assumed to be in the same universality class as lattice percolation (VICSEK and KERTESZ, 1981; GAWLINSKI and STANLEY, 1981; BALBERG *et al.*, 1983), so many of the standard results are expected to hold for continuum percolation as well.

One possible model for the flow paths in the fractures involves randomly distributing circles in the plane. The occupied area fraction in this model is (PIKE and SEAGER, 1974)

$$AF(N) = 1 - \exp\left[-\frac{N\pi r^2}{\text{Area}}\right] \quad (4)$$

where r is the radius of the circles, and Area is the area in which N points are placed. The quantity in the exponential is simply the mean number of centers that fall within an area of radius r . The critical area fraction for the onset of percolation has been determined to be $AF_c = 0.688$ (PIKE and SEAGER, 1974), which corresponds to the mean value of 1.14 points per circle of radius r . A percolation plot near the percolation limit with $AF = 0.7$ is shown in Figure 9.

Because continuum percolation is assumed to be in the same universality class as lattice percolation, the mean size of the percolation clusters (correlation length) is expected to be the same if the occupation probability of lattice percolation is replaced by the area fraction (ELAM *et al.*, 1984; GAWLINSKI and STANLEY, 1981).



XBL 8611-4761

Figure 9

A random percolation plot near the percolation limit with area fraction $AF = 0.70$. The pattern cannot qualitatively reproduce the aggregated structures of the real flow path data in Figure 4. Specifically, the fractal dimension of the percolation plot is $D = 2.00$ regardless of the area fraction.

The conductivity exponent, t , defined by $\sigma \sim (p - p_c)^\beta$, is a different matter. The conductivity exponent in lattice percolation is the same for both electrical conductivity and for fluid conductivity because all the bonds have equal apertures. The accepted value is $t = 1.2$ for 2-d (BÖTTGER and BRYSKIN, 1985). However in continuum percolation the flow paths have different apertures, which affects the flow of electrical and fluid currents differently. These differences were investigated theoretically by HALPERIN *et al.* (1985) for a Swiss-cheese model in which circular holes were randomly punched in a two-dimensional conductive sheet. They found a substantial increase in the value of t for fluid flow over the value from lattice percolation. This was attributed to the cubic dependence of fluid flow on aperture which caused the narrow flow necks to dominate the fluid flow properties. The values of t were found to be model dependent, but the new estimates of t for fluid flow in a two-dimensional continuum ranged from $t \sim 1.7$ to 2.7.

This critical exponent for flow could be used to model the flow for values of high stress. However, several deficiencies exist in the random continuum percolation model for the flow paths through a natural fracture. At the highest stress, the area fractions of E30 and E32 are 0.69 and 0.59, respectively. From the random continuum percolation model the critical area fraction is 0.68 which would imply that E30 just barely percolates, while E32 is unable to support flow. Yet a cursory inspection of the flow paths shows that substantial flow is present at these values of area fraction. In addition, the random percolation plot of Figure 9 cannot qualitatively reproduce the large areas of contact in sample E30. Finally, for random

percolation, the fractal dimension of the entire pattern (percolating as well as nonpercolating clusters) is equal to 2.00 (only the single percolating cluster at threshold has a fractal dimension $D = 1.89$). However, we have already demonstrated that the complete flow path patterns have fractal dimensions at all values of stress.

C. Stratified Continuum Percolation

An improved continuum percolation model can be constructed by considering the auto-correlation function of the percolation plots. For random percolation, if a given site is occupied, the probability that a site which is a distance R away is also occupied is independent of R . From an inspection of the contact areas of the fractures, however, it is apparent that sites of contact have a high probability of being immediately surrounded by other sites of contact. In other words, the contact area is self- or auto-correlated. Correlation and fractal dimension are directly related. The auto-correlation function of a distribution $f(r)$ is defined by

$$F(R) = \int f(r)f(r + R)dr \quad (5)$$

where $F(R)$ is the probability that a feature of $f(r)$ present at site r is still present at site $r + R$. If $f(r)$ represents a power-law probability distribution, then the distribution is fractal, and its auto-correlation function varies as $F(R) \sim R^{-m}$, where $m = E - D$, and E is the Euclidean dimension. For the random continuum model, the auto-correlation function is a constant, which gives a fractal dimension $D = 2.00$. Therefore, a better model for the flow paths through the fractures should be a continuum percolation model that possesses correlations. Aspects of correlations in percolation models have been considered for nearly as long as percolation theory itself (PIKE and SEAGER, 1974). But the assumed interactions have been short-ranged, extending only to the first few nearest neighbors (KLEIN, 1982; TUTHILL and KLEIN, 1983). The correlation functions of these models do vary with R , but they do not scale (except at the critical threshold), which is required for the models to be fractal.

There are several ways to introduce correlations into percolation models. One approach transforms an originally random distribution of sites into a new plot according to defined rules that act to clump the occupied points. Different transformation rules produce different correlation functions, and the rules can be adjusted to fit the empirical data. However, this procedure is not easily implemented for a continuum model, and the properties of the resulting plots have no obvious connection to the properties of the random model. We choose a different approach which produces percolation plots that scale with adjustable fractal dimensions, but which preserves some of the properties of the standard continuum percolation model. This percolation model is a hybrid between a Sierpinski carpet and the

random percolation plot, and is constructed iteratively, with the standard continuum construction applied at each iteration. The procedure begins by randomly distributing N sites within a given square region, called a tier. Each site defines the center of a new tier which is smaller, by a scale factor b , than the parent tier. Within each of the new, smaller tiers, N points are randomly distributed which define the centers of yet new tiers that are again reduced in size by a factor of b from the immediately preceding tier. The procedure continues for as many iterations as is necessary, or possible, within the resolution of the graphics. At the last iteration, the squares of size $1/b^n$ are finally plotted, where n is the number of tiers. This model is a stratified continuum percolation model.

The fractal dimension of the pattern is estimated by considering the area fraction at each stage. Because the distributions are allowed to overlap, double counting must be avoided in the calculation of the area fraction. Starting with a given area fraction per tier, $AFpT = AF(N)$ (defined in eq. (4)), the fractal dimension is approximately

$$D \approx \frac{\ln[AFpT + (1 - AFpT)(AFpT)^2]}{\ln[b]} + 2. \quad (6)$$

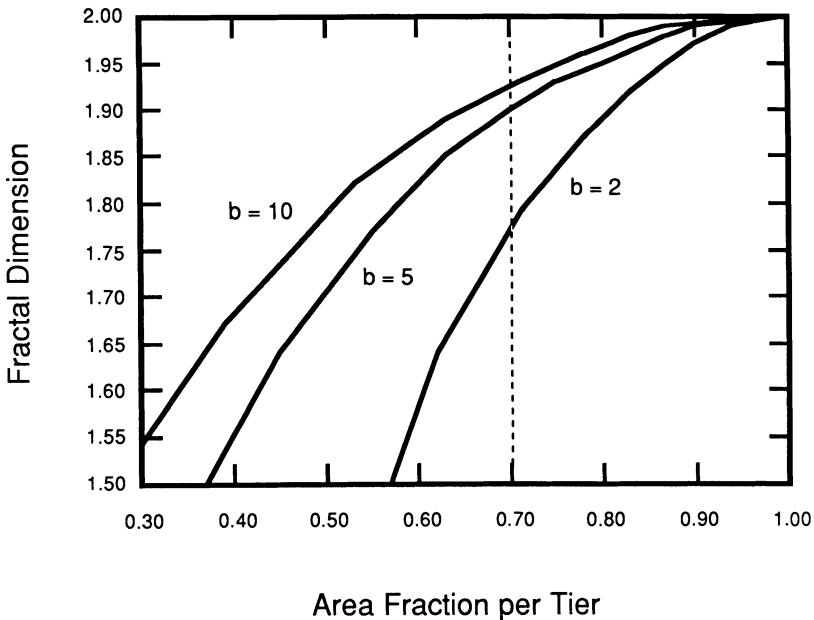
In the absence of overlap, the fractal dimension would be $D = 2 + \ln[AFpT]/\ln[b]$, but the presence of overlap increases the apparent area fraction per tier, which is reflected by the second term in the numerator of eq. (6). The overlap fraction per tier is $AFO = AF(N) * AF(N - 1) \approx (AFpT)^2$.

The total area fraction of the stratified percolation plots is not uniquely defined for a given scale b and $AFpT$. Instead, it depends on the resolution and the upper cut-off of the pattern: this is a fundamental property of fractals. For n tiers, the total area fraction is given approximately by

$$AF_n \approx AFpT[AFpT + (1 - AFpT)(AFpT)^2]^{n-1}. \quad (7)$$

The results from eqs. (6) and (7) are plotted in Figures 10 and 11. Monte Carlo computer simulations were run which measured the actual area fractions and correlation functions for varying $AFpT$ and b . The estimates of eqs. (6) and (7) are essentially correct up to around four tiers. Deviations of eq. (7) from the measured area fraction occur for more than four tiers due to errors in counting overlap of structures on different tiers. Likewise the fractal dimension given in eq. (6) is found to underestimate the measured values (see eq. (9)).

Fractal stratified percolation plots are shown in Figure 12 with increasing fractal dimensions. The fractal dimensions of the plots were verified by calculating the auto-correlation functions using two-dimensional fast Fourier transforms (FFT). Aliasing in the transforms was eliminated by zero-padding, which removes the effects of the artificial periodicity imposed by the FFT algorithm. It is interesting to compare the real data of Figure 4 to the simulation in Figure 12b, both of which have similar fractal dimensions: the qualitative match of the flow path patterns is



Area Fraction per Tier

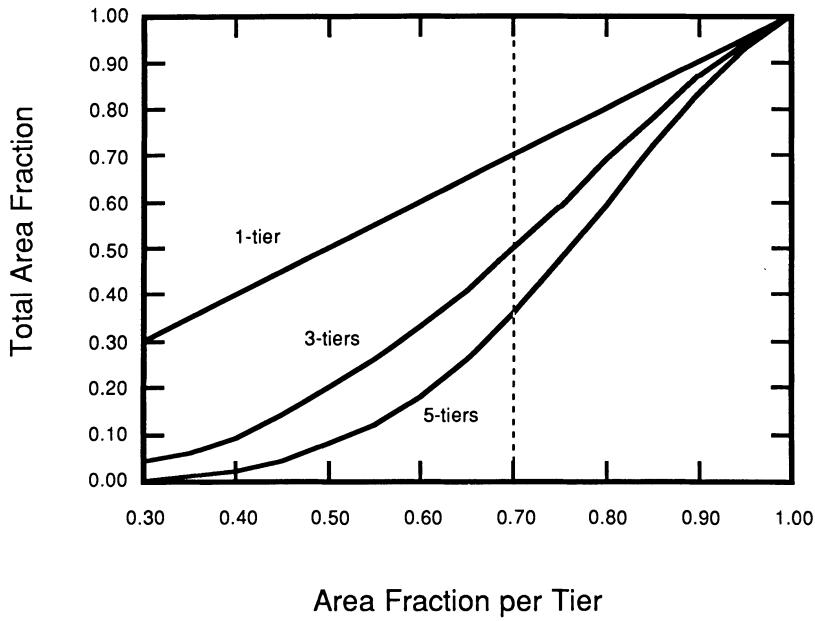
XBL 874-1787

Figure 10

Approximate fractal dimensions of stratified continuum percolation models as functions of area fraction per tier (AFpT) for three values of scale parameter: $b = 2$, $b = 5$, and $b = 10$. The percolation threshold occurs near $AFpT = 0.7$, independent of the resolution (number of tiers).

quite good. Overall, the results of the stratified percolation model match well with the data of the compliant fracture E30, but do not match as well with the data of the stiff fracture E32. The area fractions are substantially different, and the large, bulky areas of contact observed in E30 do not appear in E32. Large areas of contact are still present in E32, as they must be in order to produce the fractal dimension of the pattern; however they are drawn out into thin strings and filaments with coastlines of high fractal dimension. The stratified percolation model that we have presented tends to produce bulky structures rather than filaments. This feature is directly related to the lacunarity (MANDELBROT, 1983) of the patterns. Lacunarity is a qualitative measure of the connectedness of the tremas. Large lacunarity corresponds to large and bulky tremas as in sample E30, while small lacunarity corresponds to small and fragmented tremas as in sample E32. Some additional degree of freedom should therefore be introduced into our model that would influence the lacunarity of the tremas. We do not attempt to expand the model in the present work, but feel that the gross features of the flow paths through the fracture are reasonably well explained.

The percolation properties of the stratified continuum model require some discussion. The model is based on the standard continuum model, for which the



Area Fraction per Tier

XBL 874-1789

Figure 11

Approximate total area fraction as a function of $AFpT$ for a 1-tiered, 3-tiered, and 5-tiered stratified continuum percolation plot. The total area fraction is clearly a function of the number of tiers (resolution).

area fraction is the variable parameter which defines the percolation threshold and the percolation exponents. However, the total area fraction of the stratified plots is a function of the number of tiers (or the cut-off), and therefore the total area fraction is not uniquely defined. For this reason, the total area fraction of a fractal percolation model cannot be used as the critical parameter. The fractal dimension is well defined but the percolation threshold again does not occur at a uniquely defined fractal dimension: the fractal dimension can be used as an indicator of the nearness to threshold, but cannot be used quantitatively. On the other hand, the computer simulations do indicate that all the systems (for varying fractal dimension and scale factor b) have a threshold at $AFpT \approx 0.7$, which is the percolation threshold of the standard model. A well defined percolation threshold can therefore be identified for the stratified percolation simulations which use $AFpT$ as an input parameter.

Defining the percolation threshold is of great importance for helping understand the flow properties of a random system. Near the threshold, the flow equation should be

$$Q - Q_\infty = C \cdot (d_{\max} - d)^m \cdot (AFpT - 0.7)^t \quad (8)$$

where t is the flow percolation exponent. Extensive Monte Carlo renormalization

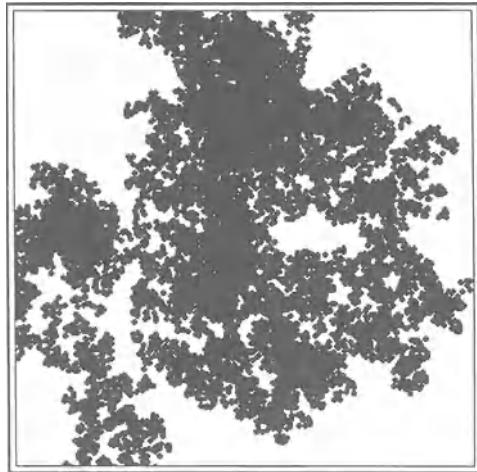


Fig. 12(a)

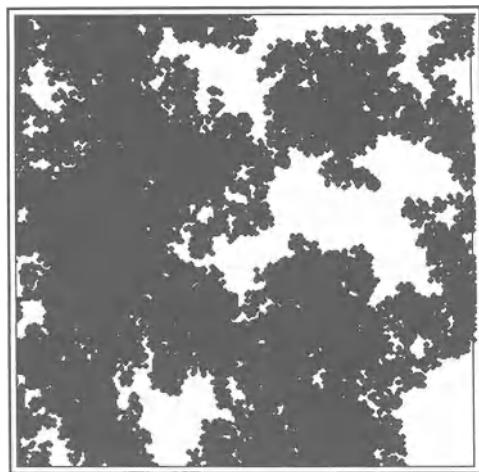


Fig. 12(b)

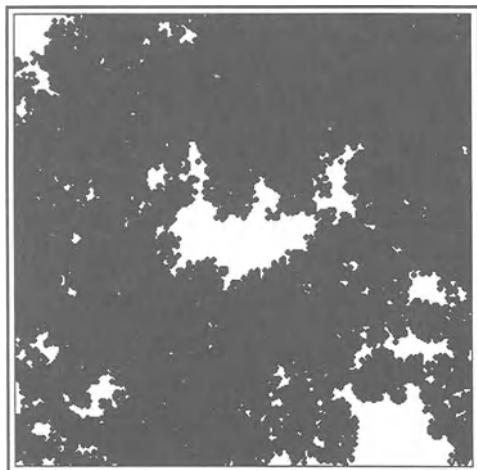


Fig. 12(c)

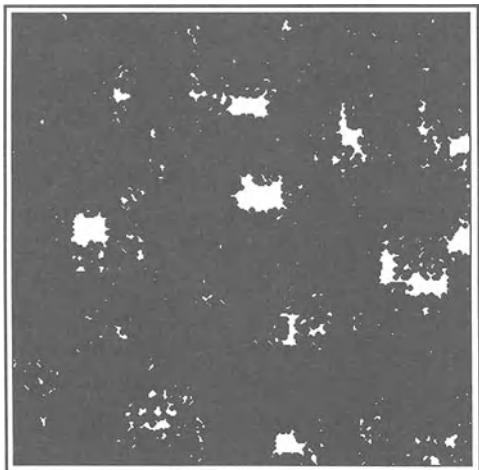


Fig. 12(d)

Figure 12

Stratified continuum percolation plots for a succession of fractal dimensions: (a) 5-tiers, $b = 2.4$, 7 points per tier, area fraction per tier at the percolation limit: $AFpT = 0.70$, $D = 1.80$; (b) 5-tiers, $b = 2.4$, 9 points per tier, $AFpT = 0.79$, $D = 1.94$. This pattern should be compared with the real flow path pattern in Figure 4; (c) 5-tiers, $b = 2.4$, 10 points per tier, $AFpT = 0.89$, $D = 1.97$; (d) 2-tiers, $b = 9$, 250 points per tier, $AFpT = 0.96$, $D = 1.99$.

group analysis (REYNOLDS *et al.*, 1980) would be required to accurately determine the value of the exponent, but arguments of universality between stratified percolation and standard percolation behavior may provide an estimate for the exponent. From Figure 11 it can be seen that the total area fractions of the stratified plots depend linearly on the $AFpT$ around the threshold. Therefore the conductivity exponent of the standard model may carry over unmodified into the stratified model. As stated earlier, estimates for t in the standard continuum model range from 1.7 to 2.7 (HALPERIN *et al.*, 1985).

The conductivity of the experimentally measured flow-paths can now be described semi-quantitatively by finding the values of $AFpT$ for the various fractures and loadings. We can calculate the area fraction per tier by defining a standard resolution, or cut-off, for both the simulation and the real pattern. For the given resolution, both the fractal dimension and the total area fraction can be measured and matched to the results of stratified continuum percolation in order to determine the $AFpT$ of the measured pattern. Comparing the calculated $AFpT$ of the real pattern to $AFpT_c \approx 0.7$ then gives a quantitative measure of the separation of the real system from the critical percolation threshold. Defining a set resolution for the patterns is paramount to setting the number of tiers used in the stratified percolation construction. Due to this imposed cut-off, the pattern no longer truly scales and measured fractal dimension is larger than the values predicted by eq. (6). To correct for the finite number of tiers, eq. (6) must be modified before the $AFpT$ can be derived from the fractal dimension and total area fraction of the pattern. For n tiers, the measured fractal dimension of the pattern is

$$D \approx \frac{(n-1)\ln[AFpT + (1 - AFpT)(AFpT)^2]}{(n)\ln[b]} + 2. \quad (9)$$

In the limit of many tiers, this expression reduces to eq. (6) which is valid for true scaling. By plotting the fractal dimension in eq. (9) against the total area fraction in eq. (7) for a given $AFpT$, we obtain a plot of isolars of constant $AFpT$. This is shown in Figure 13 with the data of samples E30 and E32 for various loads. The isolars are only plotted for $n = 1$ to $n = 5$ tiers because that is the range of validity of eqs. (7) and (9) (because of the neglect of higher orders of overlap). For $n = 1$, there is only one tier, and the fractal dimension $D = 2.00$ independent of the $AFpT$. This is the expected result of standard continuum percolation. Under the highest stress (85 MPa) the fracture in E30 has $AFpT \approx 0.8 \pm 0.1$ while the stiffer fracture in E32 has $AFpT \approx 0.7 \pm 0.1$. This places the fracture in E32 right at the percolation threshold, and the fracture in E30 within 15% of the threshold. Of course, the errors in the values of both D and AF for these fractures are reflected in the errors on $AFpT$.

For the fracture in E30 at the highest stress, the normalized conductivity from eq. (8) should be reduced roughly by a factor

$$\frac{(0.8 - 0.7)^{2.2}}{(1.0 - 0.7)^{2.2}} \approx 0.1 \quad (10)$$

from the maximum conductivity at unit area fraction. Here the exponent, 2.2, is taken as a probable value for the conductivity exponent, t . This reduction in conductivity is the result of only considering the changing flow path geometry. At the same time, the aperture is decreasing, further reducing the flow. We can now roughly estimate the exponent m in eq. (8) by comparing eqs. (8) and (10) with eq. (2) for which the exponents $n \approx 8$. For the fracture in E30, these are related by

$$0.1(b_{85}/b_0)^m = (b_{85}/b_0)^8 \quad (11)$$

where the mechanical apertures at 85 MPa and zero load are $b_{85} = 3$ microns and $b_0 = 12.5$ microns, respectively, and the factor 0.1 is the reduction in flow from increasing tortuosity from eq. (10). Solving for the exponent in eq. (11) yields $m = 6.4$. Therefore, we find that the empirical dependence of flow on large exponents of the mechanical aperture can be partially explained by explicitly

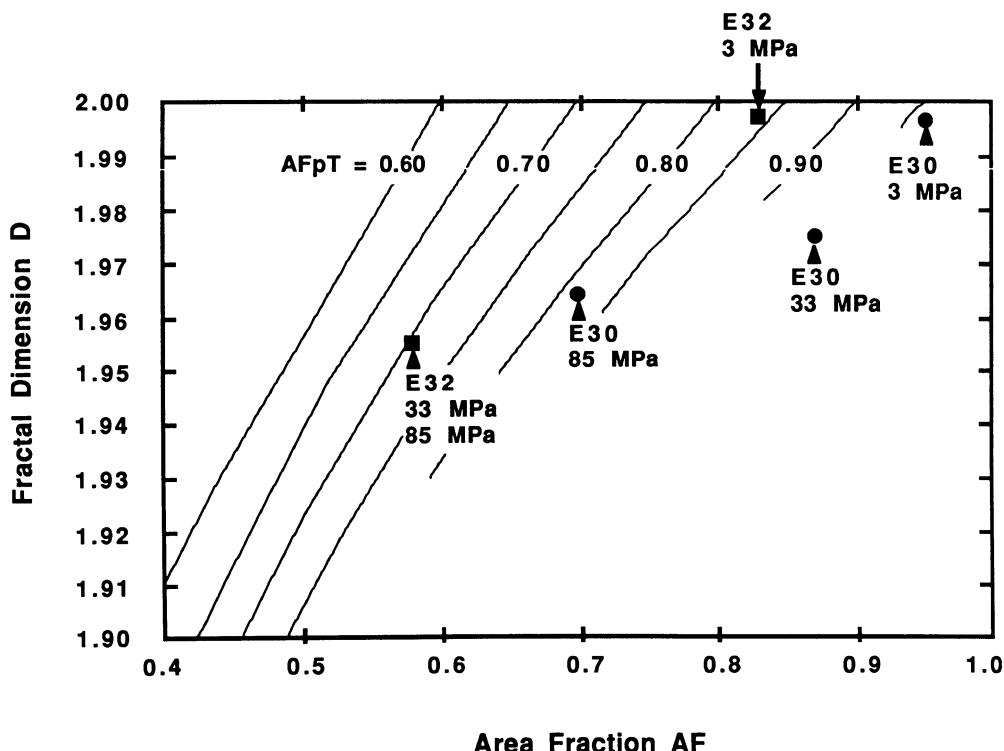


Figure 13
Fractal dimension D from eq. (9) correlated against the total area fraction AF from eq. (7). The lines are contours of constant $AFpT$ and are only plotted for one to five tiers. The fractal dimension for one tier is always $D = 2.00$. The data from the metal injection are also plotted.

considering the effects of increasing tortuosity on flow through fractures, derived from the principles of percolation theory. However, the dependence of flow on mechanical aperture alone continues to have exponents larger than cubic, which indicates that mechanical and hydraulic aperture are related in a nontrivial way.

An explanation of this remaining nontrivial relationship between mechanical and hydraulic aperture requires a detailed investigation of the specific details of the aperture distributions in the fracture and of the effect of mechanical deformation on that distribution (PYRAK-NOLTE *et al.*, 1988). In the presentation of the stratified continuum model, the $AFpT$ is not assumed to be a physical quantity in a real system. However, the concept of tiers may in fact be related to the aperture distribution. Also, the continuous dependence of flow on aperture has until now been discarded in favor of a binary process in which flow is either present or not. Even at low stress a percolating critical path must exist which is highly sensitive to changes in aperture. In this sense, the effects of percolation are present at all stresses, not only at the highest stress.

Finally, we must state that we do not claim that our fluid flow and contact area data are universal properties for all natural fractures. Our core samples are relatively small (~ 2 inches in diameter) and our fractures are well registered. These factors allow us to obtain considerable amounts of contact area and therefore approach the percolation limit. Larger core samples, or poorly registered fractures may not produce either the fractal contact area patterns or the extreme flow exponents that we obtained. More work needs to be undertaken to understand the detailed effects of core size and the stiffness of the fracture on flow properties.

5. Conclusions

The use of fractal geometry to describe physical patterns has often remained at the level of simple descriptive geometry, providing little insight into the physical processes which formed the pattern. The essence of fractal geometry lies in scale invariance, which seems to be a universal phenomenon in aggregation and fragmentation processes. A clue to understanding this ubiquitous property may be found in correlation. Microscopic correlations may well translate to scale invariant macroscopic geometry. In view of the pronounced interconnection between correlation and fractal geometry, it is to be expected that many real random flow processes (as through fractures in rock) should be highly correlated and in fact be fractal. While random percolation theory was developed primarily because of its simplicity, correlated percolation should probably be viewed as a more fundamental model of real percolation processes.

Fluid flow through a fracture in rock is clearly a three-dimensional process of high geometrical complexity. Significant simplification can be achieved by decoupling the two-dimensional effects from the effects of the third, perpendicular

direction. We have described in this paper how the two-dimensional projection of the void-space geometries can be obtained by making metal casts of the fracture under varying stresses. This two-dimensional projection defines a nominal flow path which can be analyzed using two-dimensional percolation theory. While standard continuum percolation theory can define the critical area fraction for percolation, the flow path areas of the natural fractures are found to be fractal, with no well defined area fraction. A stratified, or correlated, continuum percolation model was found to give semi-quantitative agreement with the observed flow path geometries. The critical threshold in this new model was found to be related to the area fraction covered by the flow path for a given level in the iterative procedure for modelling the flow path geometry. This area fraction per tier ($AFpT$) was determined for the measured flow paths by fitting the total area fraction (at a set resolution) as well as the fractal dimension of the pattern. For this reason, the fractal dimension becomes more than a mere descriptive value: it can be used directly to evaluate the physical behavior of the flow through the natural fracture.

Acknowledgment

We gratefully acknowledge critical comments from P. Reynolds. This work was supported by the Office of Civilian Radioactive Waste Management through the Office of Geologic Repositories through U.S. DOE contract AC03-76SF00098.

REFERENCES

- AHARONY, A. (1984), *Percolation, Fractals, and Anomalous Diffusion*, J. Stat. Phys. 34, 931–939.
- BALBERG, I., BINENBAUM, N., and ANDERSON, C. H. (1983), *Critical Behavior of the Two-dimensional Sticks System*, Phys. Rev. Lett. 51, 1605–1608.
- BANDIS, S., LUMSDEN, A. C., and BARTON, N. R. (1983), *Fundamentals of Rock Joint Deformation*, Int. J. Rock Mech. Min. Sci. and Geomech. Abstr. 20, 249–268.
- BÖTTGER, H., and BRYSKIN, V. V., *Hopping Conduction in Solids* (VCH, Altenburg 1985).
- BROWN, S. R., and SCHOLZ, C. H. (1985), *Closure of Random Elastic Surfaces in Contact*, J. Geophys. Res. 90(B7), 5531–5545.
- BROWN, S. R., KRANZ, R. L., and BONNER, B. P. (1986), *Correlation between the Surfaces of Natural Rock Joints*, Geophys. Res. Lett. 13, 1430.
- BROWN, S. R. (1987), *Fluid Flow through Rock Joints: The Effect of Surface Roughness*, J. Geophys. Res. 92(B2), 1337–1347.
- DOMB, C., and GREEN, M. S. ed., *Phase Transitions and Critical Phenomena* (Academic, London 1976).
- DUNCAN, N., and HANCOCK, K. E. (1966), *The Concept of Contact Stress in Assessment of the Behavior of Rock Masses as Structural Foundations*, Proc. First Cong. Int. Soc. Rock Mech., Lisbon, 2, 487–492.
- ELAM, W. T., KERSTEIN, A. R., and REHR, J. J. (1984), *Critical Properties of the Void Percolation Problems for Spheres*, Phys Rev. Lett. 17, 1516–1519.
- ENGELDER, T., and SCHOLZ, C. H. (1981), *Fluid Flow Along Very Smooth Faults at Effective Pressures up to 200 Megapascals*, Am. Geophys. Union Monogr. 24, 147–152.
- ESSAM, J. W. (1980), *Percolation Theory*, Rep. Prog. Phys. 43, 833–912.

- GALE, J. E., and RAVEN, K. G. (1980), *Effects of Sample Size on the Stress-permeability Relationship for Natural Fractures*, Lawrence Berkeley Laboratory Report, LBL-11865 (SAC-48), Berkeley, California.
- GAWLINSKI, E. T., and STANLEY, H. E. 1981, *Continuum Percolation in Two Dimensions: Monte Carlo Tests of Scaling and Universality for Noninteracting Discs*, *J. Phys. A14*, L291–L299.
- HALPERIN, B. I., FENG, S., and SEN, P. N. (1985), *Differences between Lattice and Continuum Percolation Transport Exponents*, *Phys. Rev. Lett.* *54*, 2391–2394.
- KLEIN, W. (1982), *Potts-model Formulation of Continuum Percolation*, *Phys. Rev. B26*, 2677–2678.
- IWAI, K. (1976), *Fundamentals of Fluid Flow through a Single Fracture*, Ph.D. Thesis Univ. of California, Berkeley.
- KIRKPATRICK, S. (1973), *Percolation and Conduction*, *Rev. Mod. Phys.* *45*, 574–588.
- MA, S-K. (1973), *Introduction to the Renormalization Group*, *Rev. Mod. Phys.* *45*, 589–615.
- MANDELBROT, B. B., *The Fractal Geometry of Nature* (W. H. Freeman, San Francisco 1983).
- MANDELBROT, B. B. (1984), *Fractals in Physics: Squig Clusters, Diffusions, Fractal Measures, and the Universality of Fractal Dimension*, *J. Stat. Phys.* *34*, 895–930.
- OLKIEWICZ, A., GALE, J. E., THORPE, R., and PAULSSON, B. (1979), *Geology and Fracture System at Stripa*, Lawrence Berkeley Laboratory Report, LBL-8907 (SAC-21), Berkeley, California.
- PHANI, M. K., and DHAR, D. J. (1984), *Continuum Percolation with Discs having a Distribution of Radii*, *Phys. A17*, L645–L649.
- PIKE, G. E., and SEAGER, C. H. (1974), *Percolation and Conductivity: A Computer Study*, *Phys. Rev. B10*, 1421–1434.
- PYRAK-NOLTE, L. J., MEYER, L. R., COOK, N. G. W., and WITHERSPOON, P. A. (1987), *Hydraulic and mechanical properties of natural fractures in low permeable rock*, Proc. 6th International Congress on Rock Mechanics (eds. Herget, G., and Vongpaisal, S.) Montreal, Canada (A. A. Balkema, Rotterdam 1987) pp. 225–231.
- PYRAK-NOLTE, L. J., COOK, N. G. W., and NOLTE, D. D. (1988), *Fluid Percolation through Single Fractures*, *Geophys. Res. Lett.* *15*, 1247–1250.
- RAVEN, K. G., and GALE, J. E. (1985), *Water Flow in a Natural Rock Fracture*, *Int. J. Rock Mech. Min. Sci. Geomech.* *22*, 251–261.
- REYNOLDS, P. J., STANLEY, H. E., and KLEIN, W. (1980), *Large-cell Monte Carlo Renormalization Group for Percolation*, *Phys. Rev. B21*, 1223–1245.
- ROBINSON, P. C. (1983), *Connectivity of Fracture Systems: A Percolation Theory Approach*, *J. Phys. A16*, 605–614.
- SCHLIFER, G. J., KLEIN, W., REYNOLDS, P. J., and STANLEY, H. E. (1979), *Large-cell Renormalization Group for the Backbone Problem in Percolation*, *Phys. A12*, L169–L174.
- SEAGER, C. H., and PIKE, G. E. (1974), *Percolation and Conductivity: A Computer Study. II*, *Phys. Rev. B10*, 1435–1446.
- STANLEY, H. E. (1977), *Cluster Shapes at the Percolation Threshold: An Effective Cluster Dimensionality and its Connection with Critical Point Exponents*, *J. Phys. A10*, L211–L220.
- STAUFFER, D. (1980), *Scaling Theory of Percolation Clusters*, *Phys. Rep.* *54*, 1–74.
- TSANG, Y. W., and WITHERSPOON, P. A. (1981), *Hydromechanical Behavior of a Deformable Rock Fracture Subject to Normal Stress*, *J. Geophys. Res.* *86*, 9287–9298.
- TSANG, Y. W. (1984), *The Effect of Tortuosity on Fluid Flow through a Single Fracture*, *Wat. Resources Res.* *20*, 1209–1215.
- TUTHILL, G. F., and KLEIN, W. (1986), *Renormalization Group for Percolation Using Correlation Parameters*, *J. Phys. A16*, 3561–3570.
- VICSEK, T., and KERTESZ, J. (1981), *Monte Carlo Renormalization Group Approach to Percolation on a Continuum: Test of Universality*, *J. Phys. A14*, L31–L37.
- VOSS, R. F. (1984), *The Fractal Dimension of Percolation Cluster Hulls*, *J. Phys. A17*, L373–L377.
- WALSH, J. B. (1981), *Effect of Pore Pressure and Confining Pressure on Fracture Permeability*, *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* *18*, 429–435.
- WITHERSPOON, P. A., WANG, Y. S., IWAI, K., and GALE, J. E. (1980), *Validity of Cubic Law for Fluid Flow in a Deformable Rock Fracture*, *Wat. Resources Res.* *16*, 1016–1024.

(Received April 29, 1987, revised/accepted August 8, 1987)

Testing Linear Models of Sea-Floor Topography

ALBERTO MALINVERNO¹

Abstract—Stochastic models can generate profiles that resemble topography by taking uncorrelated, zero-average noise as input, introducing some correlation in the time series of noise, and integrating the resulting correlated noise. The output profile will depict a nonstationary, randomly rough surface. Two models have been chosen for comparison: a fractal model, in which the noise is correlated even at large distances, and an autoregressive model of order 1, in which the correlation of the noise decays rapidly. Both models have as an end-member a random walk, which is the integration of uncorrelated noise. The models have been fitted to profiles of submarine topography, and the sample autocorrelation, power spectrum and variogram have been compared to the theoretical predictions. The results suggest that a linear system approach is a viable method to model and classify sea-floor topography. The comparison does not show substantial disagreement of the data with either the autoregressive or the fractal model, although a fractal model seems to give a better fit. However, the amplitudes predicted by a nonstationary fractal model for long wavelengths (of the order of 1000 km) are unreasonably large. When viewed through a large window, ocean floor topography is likely to have an expected value determined by isostasy, and to be stationary. Nonstationary models are best applied to wavelengths of the order of 100 km or less.

Key words: Autoregressive processes, fractals, roughness, stochastic modeling, time series analysis, topography.

Linear Models and Topography

A simple linear system can be used to generate a profile that resembles topography (Figure 1, left). The input is a time series a_t of uncorrelated noise, that has a spectral density constant at all frequencies (white noise). A filter is then applied to introduce some positive correlation between neighboring values of a_t . Such filtering will enhance the low frequencies of the input, so that the output Δz_t will have a “red” spectrum. Finally, the Δz_t series is integrated to produce an output z_t :

$$z_t = z_{t-1} + \Delta z_t.$$

The simplest case is that in which no filtering is applied to the uncorrelated input, so that $\Delta z_t = a_t$. Since the increments are uncorrelated, the resulting output z_t will

¹ Lamont-Doherty Geological Observatory, and Department of Geological Sciences of Columbia University, Palisades, New York 10964, U.S.A.

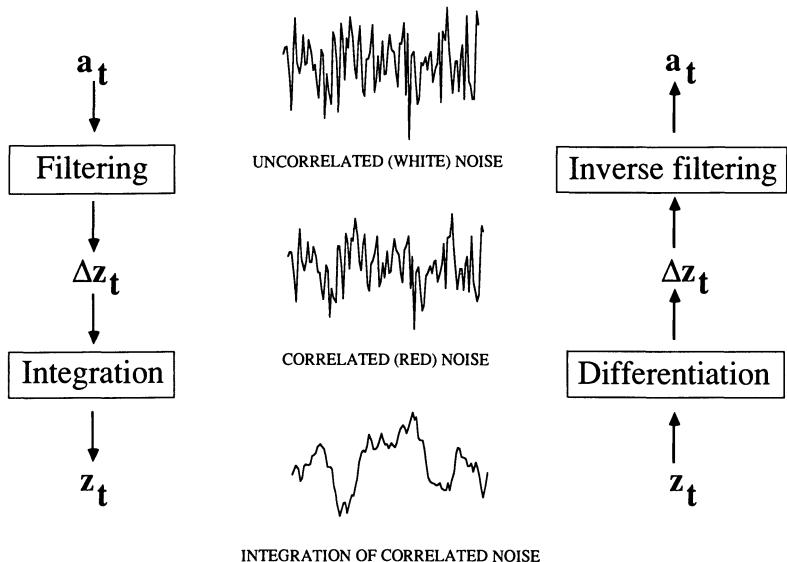


Figure 1

A profile of synthetic topography can be generated from white noise a_t (left). A filter introduces some correlation, and the resulting series Δz_t is integrated to produce the final output z_t . This scheme can be inverted to analyze a topographic profile (right). The data are first differentiated, and the correlation structure of the first differences is modeled. The final result should be uncorrelated noise.

be a random walk. A random walk, however, is usually a poor model for topography, because it is too rough (MANDELBROT, 1983). Therefore, the filtering step is necessary, and the characteristics of the filter will control the statistical structure of the synthetic topography z_t . An important difference between z_t and Δz_t is that Δz_t is stationary: its expectation is zero, and its variance is finite. On the other hand, z_t does not have a well-defined variance, and it is nonstationary.

In the analysis of a profile of real topography, the procedure described above is reversed (Figure 1, right). The input z_t is differentiated to produce a series of first differences Δz_t :

$$\Delta z_t = z_t - z_{t-1}.$$

The correlation structure of Δz_t is then the object of the analysis: the final goal is to find the inverse filter that will transform Δz_t in uncorrelated noise a_t . This can be accomplished by modeling the correlation structure of Δz_t . Two such models, a fractal and an autoregressive model, will be considered in the present work.

It has been suggested by a number of investigators that fractals provide an adequate representation of topography (MANDELBROT, 1967, 1973; BERRY and HANNAY, 1978; BURROUGH, 1981; BROWN and SCHOLZ, 1985; VOSS, 1985). Two basic attributes of fractals are self-similarity and nondifferentiability. A typical

example is that of an hypothetical coastline viewed from decreasing heights (BROWN and SCHOLZ, 1985; VOSS, 1985). As an observer gets closer, the overall jaggedness of the coastline will not change, since new bays and promontories will appear. If one were to draw the coastline from different heights, it will not be possible to distinguish pictures made at high altitudes from others made at low altitudes: the coastline is self-similar. At any given height, it will seem possible to draw tangents to points on the coastline. But as one gets nearer, more details will become visible and the concept of a tangent is meaningless: the coastline is not differentiable. Nevertheless, there will be a parameter that characterizes the roughness of such a curve, *i.e.*, its fractal dimension. The fractal dimension D will be a real number lying between the topological dimension D_T and the Euclidean dimension D_E . For a coastline, D will be between 1 and 2, and it will quantify the ability of the coastline to fill the Euclidean space. A measurement of the fractal dimension can be made by walking a pair of dividers of aperture G over a map of the coastline and by considering the local lengths measured as G varies, $L(G)$. If the coastline is "smooth", then $L(G)$ will tend to a constant as $G \rightarrow 0$; if the coastline is a fractal, we will have

$$L(G) = MG^{1-D}$$

where M is a constant and D is the fractal dimension (MANDELBROT, 1967).

A fractal coastline can be viewed as the intersection of a horizontal plane with a fractal surface. The intersection between any vertical plane and a fractal surface is also a fractal. Such a profile, or trace, is not strictly self-similar, but self-affine: if the trace is rescaled by multiplying the horizontal coordinate by a factor r , the vertical coordinate has to be multiplied by a factor r^{2-D} (MANDELBROT, 1983; VOSS, 1985).

A common model of a discrete fractal trace is the "fractional Brownian motion" (fBm) of MANDELBROT and VAN NESS (1968). In the scheme of Figure 1, z_i corresponds to fBm and its first differences Δz_i to "fractional Gaussian noise" (fGn). Realizations of fBm and fGn are characterized by a parameter H ($0 < H < 1$) that quantifies the correlation between successive values of fGn. A value of H between 1 and 0.5 corresponds to positive correlation; $H = 0.5$ corresponds to no correlation (fGn is white noise, fBm a random walk), and H between 0 and 0.5 to negative correlation (MANDELBROT and VAN NESS, 1968). The fractal dimension D of a trace of fBm is related to the parameter H (VOSS, 1985):

$$D = 2 - H \quad (1)$$

D then varies between 1 and 2, and for a random walk $D = 1.5$.

The self-affinity of fBm implies that its increments have a small but nonnegligible correlation even when they are separated by large distances (MANDELBROT and WALLIS, 1969). In fact, an approximation of fGn valid for $0.5 < H < 1$ can be written as follows (modified from MANDELBROT and WALLIS, 1969):

$$\Delta z_t = (H - 0.5) \sum_{j=1}^{\infty} j^{H-1.5} a_{t-j} \quad (2)$$

(note that this approximation is not adequate as H tends to 0.5). In other words, fGn can be viewed as the convolution of Gaussian white noise a , with a kernel that decays slowly with increasing lag j . The autocorrelation of fGn will also follow a power decay law (MANDELBROT and WALLIS, 1969):

$$\rho(s) = \frac{1}{2} [|s + 1|^{2H} - 2|s|^{2H} + |s - 1|^{2H}]. \quad (3)$$

This form of the autocorrelation function implies that the power spectra of fGn and fBm vary as $f^{-\beta}$, where f is frequency and β a positive real number (MANDELBROT and VAN NESS, 1968; MANDELBROT and WALLIS, 1969). In particular, if the power spectrum of fBm varies as $f^{-\beta}$, the power spectrum of the corresponding fGn varies as $f^{-\beta+2}$ (VOSS, 1985). In the case of a random walk, the power spectrum varies f^{-2} , and the power spectrum of its increments is constant. For fBm, it can be shown that β is related to H by the formula:

$$H = \frac{\beta - 1}{2}.$$

Using (1), we can write:

$$D = \frac{5 - \beta}{2}. \quad (4)$$

Equation (4) has been used to estimate D for profiles for natural surfaces from the slope of the logarithm of the power spectrum plotted versus the logarithm of frequency (BERRY and HANNAY, 1978; BURROUGH, 1981; BROWN and SCHOLZ, 1985).

An alternatively simple way to generate an output of correlated noise Δz_t from an input of white noise a_t is to introduce some dependence between the output at time t and the output at time $t - 1$:

$$\Delta z_t = \Phi_1 \Delta z_{t-1} + a_t \quad (5)$$

where Φ_1 is a correlation parameter ($0 \leq \Phi_1 < 1$ for positive correlation). This is an autoregressive process of order 1, or AR(1). The time series z_t is then defined as an autoregressive integrated process of order 1, or ARIMA (1, 1, 0). These linear models are treated comprehensively by Box and JENKINS (1970), and only a summary of their characteristics will be offered here.

By expanding Δz_{t-1} in terms of Δz_{t-2} , Δz_{t-3} , in terms of Δz_{t-3} and so on, (5) can be written as follows:

$$\Delta z_t = \sum_{j=0}^{\infty} \Phi_1^j a_{t-j} \quad (6)$$

i.e., an AR(1) process can be viewed as the convolution of white noise with a kernel that decreases rapidly with increasing lag j . The autocorrelation of an AR(1) process also follows an exponential decay law:

$$\rho(s) = \Phi_1^s. \quad (7)$$

The correlation decays much more rapidly with increasing lag in an AR(1) process than in fGn; compare (6) and (7) to (2) and (3). This is in fact what sets apart these two types of models (MANDELBROT and VAN NESS, 1968; MANDELBROT and WALLIS, 1969). On the other hand, they share common end members: white noise for fGn and AR(1), and a random walk for fBm and ARIMA(1, 1, 0). An AR(1) process with $\Phi_1 = 0$ is in fact white noise, and its integration a random walk. Finally, it can be shown that the power spectrum of an ARIMA(1, 1, 0) process takes the following form (BOX and JENKINS, 1970):

$$P(f) = \frac{\sigma_a^2}{2[1 - \cos(2\pi f)][1 + \Phi_1^2 - 2\Phi_1 \cos(2\pi f)]} \quad (8)$$

where f is frequency ($0 < f \leq 0.5$) and σ_a^2 the variance of the white noise a_t .

In the rest of this study, both a fractal and an autoregressive model will be applied to sea-floor topography profiles. The goals of this practical exercise will be to determine if the general approach described in Figure 1 indeed provides an adequate representation, and to decide if the data conclusively support one of the models and discount the other.

Sea-Floor Topography Data

The sea-floor topography measurements that will be used in the present study were collected by the deep-towed sonar vehicle Sea MARC I during a survey of the axial region of the Explorer Ridge, Northeast Pacific (Figure 2). During towing, a down-looking transducer on the sonar vehicle emitted periodically a short acoustic pulse, that was reflected by the sea-floor. The time elapsed between the emission and the arrival of the backscattered signal provided a measure of the altitude of the sonar vehicle on the sea-floor. A pressure gauge measured the depth of the sonar vehicle. For each acoustic cycle (i.e., every 4 sec for the data presented here), vehicle altitude and depth were recorded on tape (CHAYES, 1983). Vehicle altitude and depth were subsequently median filtered, to eliminate spot noise, and summed, to obtain sea-floor depth; finally, the sea-floor depth data were decimated. The decimation was necessary because the down-looking sonar has a beamwidth of about 30° , so that the area of sea-floor insonified during consecutive cycles overlap and some artificial correlation is introduced. For typical altitudes of 300 m, the diameter of the circle of sea-floor insonified at each cycle is about 160 m. For typical

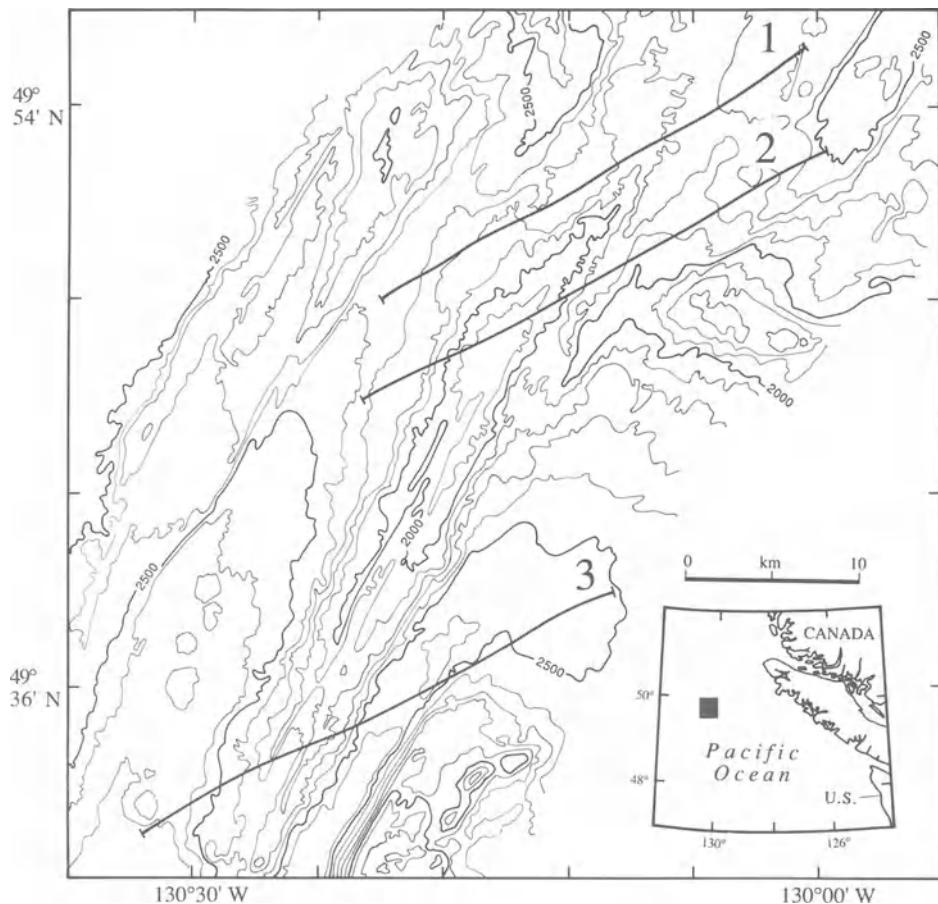


Figure 2

Bathymetry of the study area, after HAMMOND *et al.* (1984); contour interval 100 m. The three profiles studied are shown as continuous lines, and the overall location of the study area as a black square in the insert.

towing speeds of about 1 m/sec, the sonar vehicle moves about 4 m between consecutive pings, so that only one measurement in every forty is obtained from nonoverlapping areas of the sea-floor.

Three straight segments, each corresponding to 8 hours of continuous towing, have been used for the present study (Figure 3). The data were decimated by a factor of 45, obtaining 160 points in each profile. The length of the profiles is about 30 km, and the sampling interval about 185 m. The segments studied have all the same orientation with respect to the rift axis, so that the anisotropy of the sea-floor fabric should not have any effect.

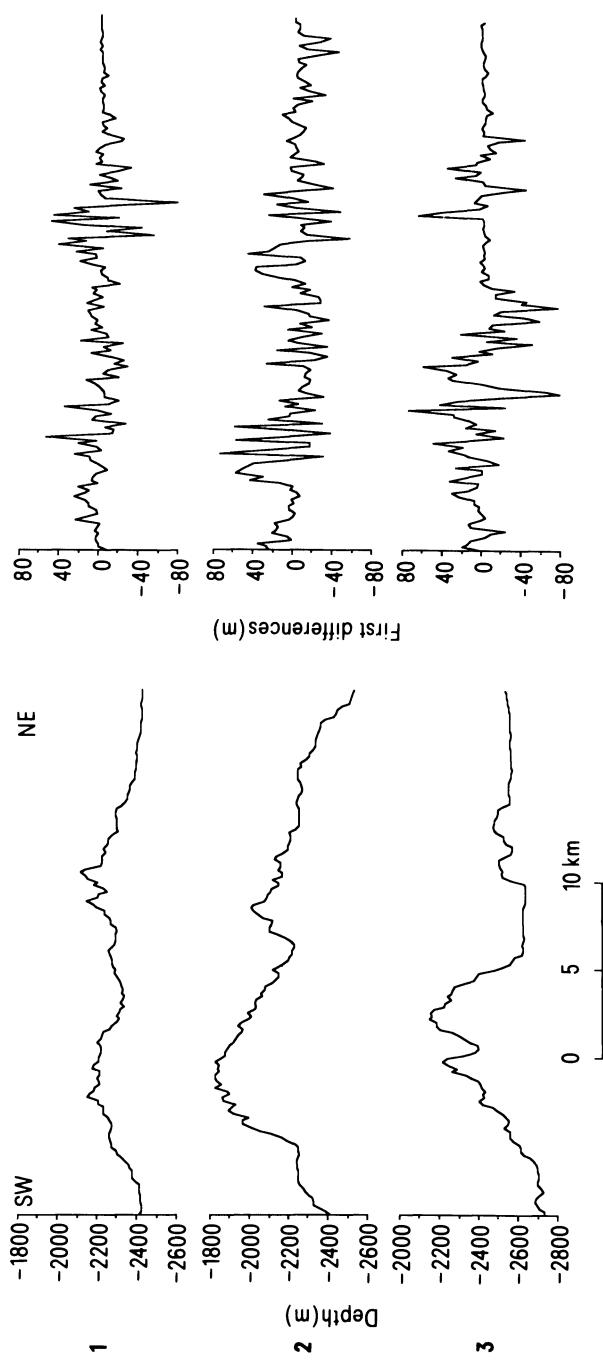


Figure 3
The three profiles studied (left) and their first difference (right).

Modeling Results

If topography is indeed comparable to fBm or to an ARIMA(1, 1, 0) process, then any profile of a statistically homogeneous sea-floor can be described by two parameters: a correlation parameter and an amplitude parameter. The correlation will be expressed by the fractal dimension D or by Φ_1 ; the amplitude by some measure of the variance of the process. For the purposes of the present work, the measures of correlation are the most important and will be treated below.

The three profiles have been considered as separate realizations of the same process, so that both the autocorrelations and the sample spectra of each profile have been averaged to smooth the resulting estimates. The average sample autocorrelation of the first differences is compared in Figure 4 to the theoretical autocorrelations of the AR(1) model (dashed line) and fGn (solid line). The fractal dimension used for the theoretical autocorrelation is $D = 1.32$, and has been determined by finding the value of D that gave the smallest squared error between the sample autocorrelation and the theoretical autocorrelation for lags ≤ 20 . The parameter Φ_1 of an AR(1) process Δz_t ($t = 1, 2, \dots, N$) can be estimated by a least-squares procedure. The sum of squares to be minimized is:

$$S.S. = \sum_{t=2}^N [\Delta z_t - \Phi_1 \Delta z_{t-1}]^2 = \sum_{t=2}^N [a_t]^2. \quad (9)$$

It can be shown (Box and JENKINS, 1970) that the value of Φ_1 that minimizes (9) is approximately the value of the autocorrelation at lag one, so that Φ_1 can be estimated from the sample autocorrelation:

$$\Phi_1 = \frac{\sum_{t=1}^{N-1} \Delta z_t \Delta z_{t-1}}{\sum_{t=1}^N [\Delta z_t]^2}$$

Φ_1 for the first differences of the profile studied is 0.298 ± 0.044 . The variance of Φ_1 has been computed from (Box and JENKINS, 1970):

$$\sigma_{\Phi_1}^2 = \frac{1 - \Phi_1^2}{N}.$$

For lags less than 10, the theoretical autocorrelation of fGn fits better the sample autocorrelation. For lags between 10 and about 20, the sample autocorrelation falls between the predictions of the two models. For lags larger than about 20, the sample autocorrelation becomes negative, which is against the predictions of both models.

The averaged sample power spectrum of the three profiles studied is shown in Figure 5, with theoretical power spectra computed for the fitted ARIMA(1, 1, 0) process and fBm. The spectrum has been computed using the differentiation as a

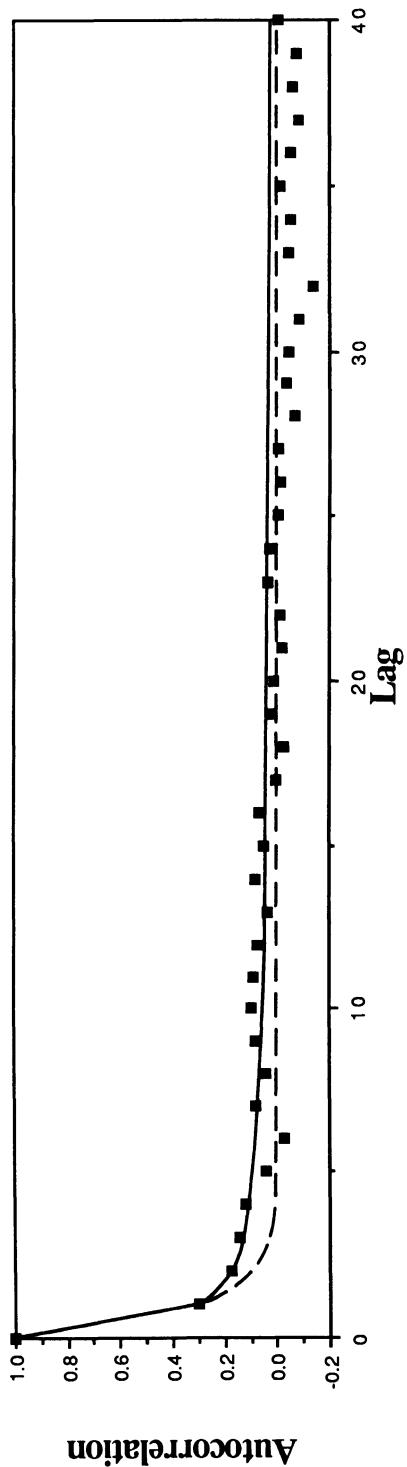


Figure 4
Average sample autocorrelation of the first differences of the profiles studied (black squares) compared to the theoretical autocorrelations of a fractal model with $D = 1.32$ (continuous line) and of an AR(1) model with $\Phi_1 = 0.298$ (dashed line).

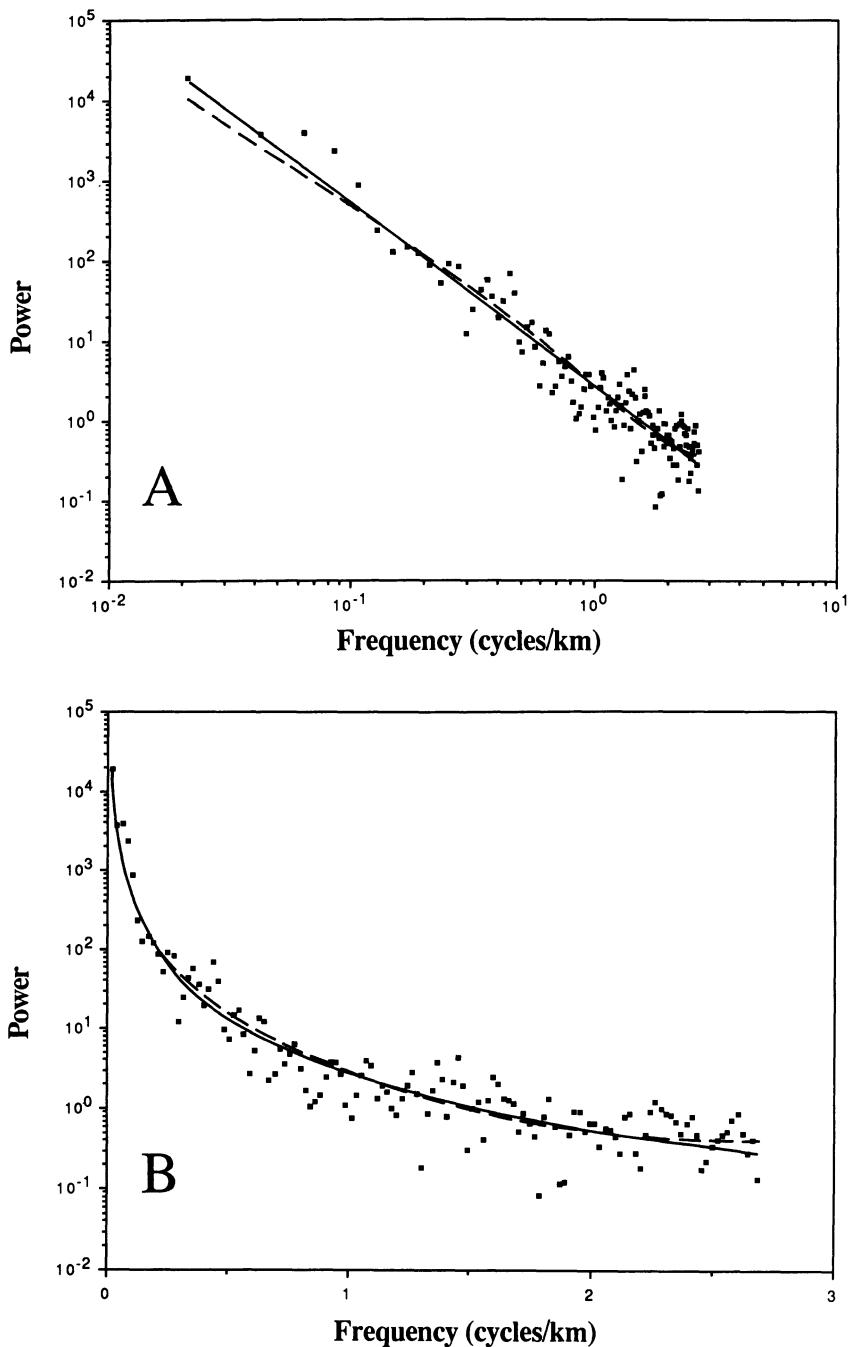


Figure 5

Average periodogram of the profiles studied (dots) compared to the theoretical spectra of a fractal model with $D = 1.36$ (continuous line) and an ARIMA (1, 1, 0) model with $\Phi_1 = 0.298$ (dashed line). *A* is a log power-log frequency plot; in *B* the frequency scale is linear.

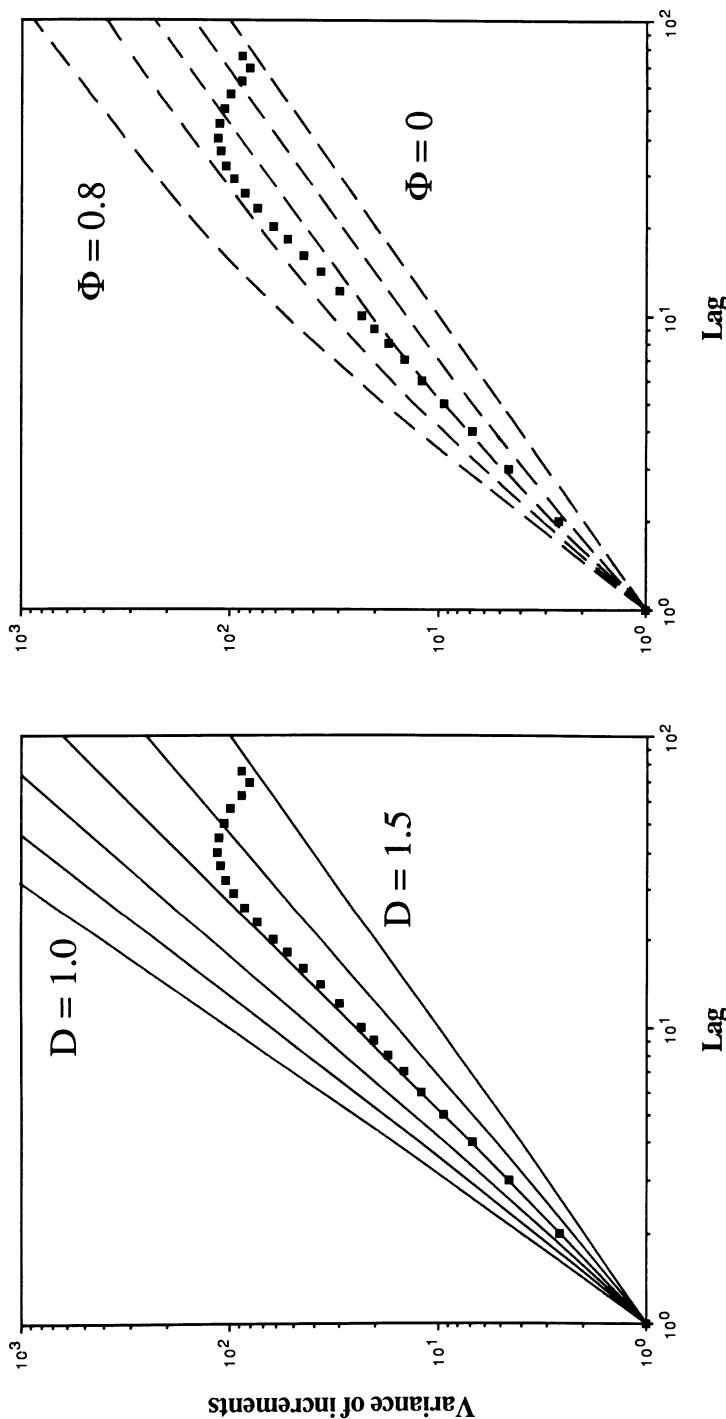


Figure 6
Average normalized variogram of the profiles studied (black squares) compared to the predictions of a fractal model (left; D varies between 1 and 1.5 with steps of 0.1) and for an ARIMA (1, 1, 0) model (right; Φ_1 varies between 0 and 0.8 with steps of 0.2). Note that the sample variance of the increments reaches a constant value for large lags.

prewhitening filter (see Appendix A). Prewhitening avoids leakage of power between frequencies, and is necessary if the original data have a spectrum that contains a substantial trend (FOX and HAYES, 1985). The fractal dimension in this case has been estimated by fitting a least squares line to the log-log power spectrum (Figure 5A) and computing D from the resulting slope β using (4). This gives $D = 1.36 \pm 0.033$ (the error is derived from the standard deviation of the slope estimate). Note that the theoretical spectra for the fractal and the autoregressive are practically indistinguishable in this case.

The variance of the increments can also be used to estimate the fractal dimension of a profile (BURROUGH, 1981). If z_t is fBm, it can be shown that (MANDELBROT and VAN NESS, 1968; VOSS, 1985).

$$V(s) = E\{[z_{t+s} - z_t]^2\} \propto s^{2H} = s^{4-2D} \quad (10)$$

where $V(s)$ is the variance of the increments, and $E\{\}$ the expectation operator. For a random walk, $V(s) \propto s$. If z_t is an ARIMA(1, 1, 0) process, it can be shown that (see Appendix B):

$$V(s) = E\{[z_{t+s} - z_t]^2\} \propto s + 2 \sum_{j=1}^{s-1} (s-j)\Phi_j^i.$$

The simplest estimator for $V(s)$ of a time series z_t ($t = 1, 2, \dots, N$) is:

$$V(s) = \frac{1}{N-s} \sum_{t=1}^{N-s} [a_{t+s} - z_t]^2.$$

A plot of the sample $V(s)$ versus s is commonly referred to as the variogram (JOURNEL and HUIJBREGHTS, 1978). The normalized variogram $V(s)/V(1)$ of the profile studied is compared to theoretical curves for fBm and an ARIMA(1, 1, 0) process in Figure 6. Note that the plot is in log-log space, so that a power law becomes a straight line. Up to lags of about 30, the sample $V(s)$ shows an excellent fit to a fractal model with a fractal dimension slightly larger than 1.3 (for lags ≤ 20 , a least-squares fit gives $D = 1.315 \pm 0.002$). On the other hand, the fit to the theoretical curves for a ARIMA(1, 1, 0) model is rather poor. For lags larger than 30, the sample $V(s)$ does not keep pace with the increase predicted by both models.

Discussion

The satisfactory fit to the theoretical predictions for both the autocorrelation and the spectrum demonstrates the validity of the general linear model illustrated in Figure 1 for the analysis of topographic profiles. On the other hand, it does not seem possible to conclude that the observations definitely disagree with either of the models proposed to account for the correlation in the Δz_t series. An obvious reason

for the failure of this test is that the trace studied is quite close to a random walk (D is between 1.315 and 1.36, Φ_1 about 0.3). Both fBm and an ARIMA (1,1,0) process converge to a random walk as D tends to 1.5 and Φ_1 tends to zero, and they become increasingly difficult to distinguish. For example, a test could be conducted on the values of the autocorrelation at large lags, that for fGn should be significantly larger than zero: but in our case, D is close enough to 1.5 for the theoretical autocorrelation of fGn to be close to zero even at small lags (Figure 4). A better test would be provided by data with a smaller D and a larger Φ_1 . However, although an ARIMA(1, 1, 0) model cannot be discounted, a fractal model seems a better approximation to the profile studied. The fractal model gives a superior fit to the autocorrelation for small lags (Figure 4) and to the general trend of the variance of the increments (Figure 6).

A log-log plot of the variance of the increments seems to provide a way to estimate the fractal dimension D of a given trace more robust than the use of the power spectrum: the standard deviation of the estimate of D from the variogram is substantially smaller than that obtained from the spectrum. However, the sample variogram departs from the predicted increase in variance at lags > 30 and seems to tend to a constant value (Figure 6). Such behavior can be explained if the samples separated by large lags are independent and the z_t series has a finite variance (JOURNEL and HUIJBREGHTS, 1978). A nonstationary model, such as fBm, predicts instead that the variance of the increments increases without bounds as the lag increases (see (10)) and overestimates the variance of the increments for large lags.

In the frequency domain, a fractal model predicts that the power spectrum varies as $f^{-\beta}$: in other words, as the spatial frequency of topography decreases, the power increases without bounds. In the case of ocean-floor topography, such unbounded increase seems unrealistic. For example, consider the fractal model of the data presented here, which span slightly more than two decades of spatial frequency. If ocean-floor topography is indeed self-affine, the power at long wavelengths must follow the law determined for shorter wavelengths. A straightforward extrapolation is shown in Figure 7, in which the amplitude of each Fourier component is plotted on the vertical axis. For a spatial frequency of 0.001 cycles/km, *i.e.*, a wavelength of 1000 km, the extrapolation predicts an amplitude of 4274 m, which is exceedingly large. It seems likely that amplitudes of long-wavelength ocean-floor topography will be substantially smaller than those predicted by straightforward extrapolation.

In the present paper and similar studies (*e.g.*, FOX and HAYES, 1985), the random nature of topography is the object of study. On the other hand, ocean-floor topography can be viewed as a deterministic process. As the newly created lithosphere moves away from the ridge axis, cools by losing heat through its upper boundary and increases its density, the ocean floor subsides isostatically (MCKENZIE, 1967; PARKER and OLDENBURG, 1973). Neglecting the effects of sedimentation

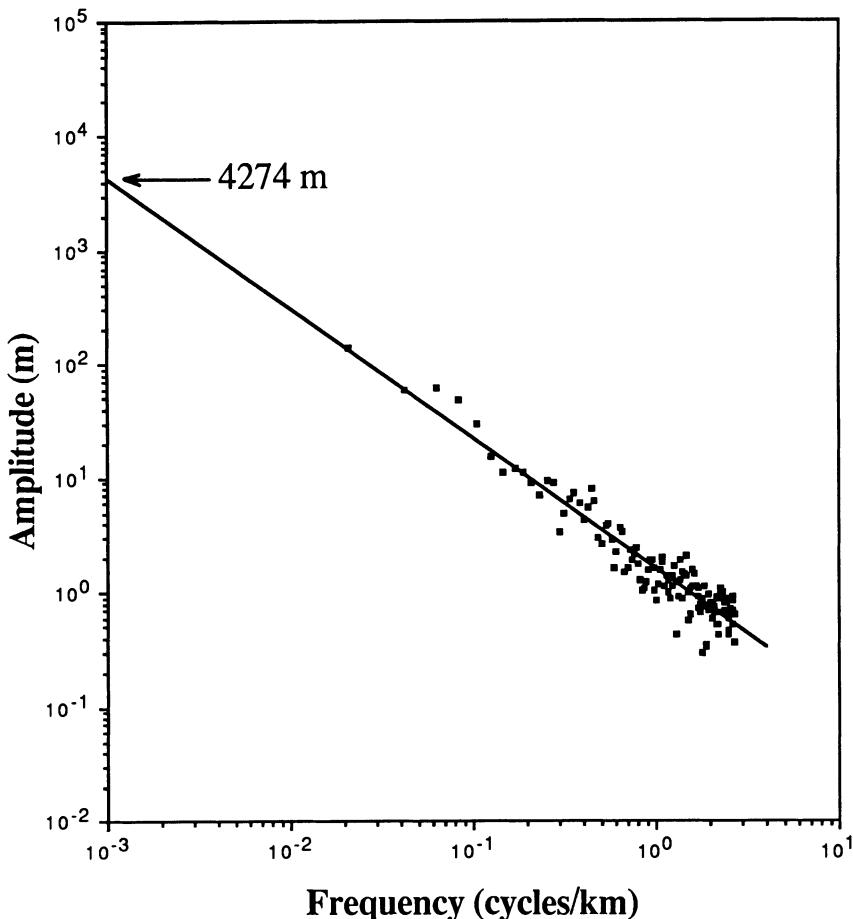


Figure 7

Extrapolation of the observed amplitude spectrum towards low spatial frequencies. If ocean-floor topography is self-similar, the power law determined for the profiles studied implies exceedingly large amplitudes for long wavelengths (4247 m for wavelengths of 1000 km).

and the local, short wavelength variation, the depth of the ocean floor is proportional to the square root of its age, for ages up to 70 million years (PARSONS and SCLATER, 1977). When viewed through a sufficiently large window, ocean-floor topography has an expected value, and should not be modeled as a nonstationary process. Therefore, the validity of a nonstationary, fractal or autoregressive model for the ocean-floor topography seems limited to the range of wavelengths in which volcanism, faulting and mass-wasting create a stochastic sea-floor, *i.e.*, to wavelengths of the order of 100 km or less.

Conclusions

(1) Ocean-floor topography can be modeled as a nonstationary random process, using a simple linear system approach. The original data are differenced, and subsequently an autoregressive or a fractal model is fitted to the correlation structure of the first differences.

(2) In a practical case, both the autoregressive and the fractal models fit reasonably well the observations. However, the fractal model seems to give a better fit to the sample autocorrelation and variogram.

(3) The fractal dimension D of the data examined can be estimated from a fit of the theoretical forms to the autocorrelation function, power spectrum and variogram. The error associated with the estimate of D from the variogram is one order of magnitude smaller than the error of the estimate of D from the power spectrum.

(4) A straightforward extrapolation of a fractal model predicts unreasonably large amplitudes for long wavelengths of ocean-floor topography (say 1000 km). When viewed through a sufficiently large window, ocean-floor topography is likely to be stationary, and to have an expected value determined by isostasy. Nonstationary models are best suited to account for the shorter-wavelength, random variability of ocean-floor topography.

Acknowledgements

Lengthy discussions with Lewis Gilbert had a large impact on the present work. The paper was revised by C. Fox, J. C. Mareschal, D. Martinson and W. B. F. Ryan, who offered valuable criticism. The work presented here was supported by the Office of Naval Research contract N00014-84-C-0132RK. This is Lamont-Doherty Geological Observatory contribution number 4415.

Appendix A: Prewhitening by Differentiation

As z_t is input and Δz_t , the output series, the differentiator is applied as follows:

$$\Delta z_t = z_t - z_{t-1}.$$

Therefore, the impulse response of the differentiator is:

$$b(t) = \begin{cases} 1 & t = 0 \\ -1 & t = 1 \\ 0 & \text{otherwise.} \end{cases}$$

The frequency response is:

$$B(f) = \sum_{t=-\infty}^{\infty} b(t)e^{-i2\pi ft} = 1 - e^{-i2\pi f} \quad (i = \sqrt{-1}; 0 < f \leq 0.5).$$

The frequency response squared is:

$$|B(f)|^2 = |1 - e^{-i2\pi f}|^2 = 2(1 - \cos(2\pi f)).$$

If differentiation is used for prewhitening, and the prewhitened power spectrum is $P_w(f)$, the power spectrum of the original process will be $P_w(f)$ divided by the squared frequency response of the differentiator:

$$P(f) = \frac{P_w(f)}{2(1 - \cos(2\pi f))}.$$

Appendix B: Variance of the Increments of an Arima (1, 1, 0) Process

Consider an ARIMA (1, 1, 0) process $z_t = z_{t-1} + \Delta z_t$, where

$$\Delta z_t = \Phi_1 \Delta z_{t-1} + a_t$$

and a_t is white noise. If $E\{\}$ is the expectation operator, we have (BOX and JENKINS, 1970):

$$E\{\Delta z_t^2\} = \sigma_c^2, \quad E\{\Delta z_t \Delta z_{t+s}\} = \sigma_c^2 \Phi_1^s. \quad (\text{B1})$$

We can also write:

$$z_{t+s} - z_t = z_t + \Delta z_{t+1} + \Delta z_{t+2} + \cdots + \Delta z_{t+s} - z_t = \Delta z_{t+1} + \Delta z_{t+2} + \cdots + \Delta z_{t+s}$$

so that the variance of the increments can be written as:

$$V(s) = E\{[z_{t+s} - z_t]^2\} = E\{[\Delta z_{t+1} + \Delta z_{t+2} + \cdots + \Delta z_{t+s}]^2\}. \quad (\text{B2})$$

From (B1) and (B2), it can be computed:

$$V(1) = E\{\Delta z_t^2\} = \sigma_c^2;$$

$$V(2) = E\{[\Delta z_{t+1} + \Delta z_{t+2}]^2\} = \sigma_c^2[2 + 2\Phi_1]$$

$$V(3) = E\{[\Delta z_{t+1} + \Delta z_{t+2} + \Delta z_{t+3}]^2\} = \sigma_c^2[3 + 2(2\Phi_1 + \Phi_1^2)]$$

$$V(4) = E\{[\Delta z_{t+1} + \Delta z_{t+2} + \Delta z_{t+3} + \Delta z_{t+4}]^2\} = \sigma_c^2[4 + 2(3\Phi_1 + 2\Phi_1^2 + \Phi_1^3)]; \dots$$

so that a general formula is:

$$V(s) = E\{[z_{t+s} - z_t]^2\} = \sigma_c^2 \left[s + 2 \sum_{j=1}^{s-1} (s-j)\Phi_1^j \right].$$

REFERENCES

- BERRY, M. V., and HANNAY, J. H. (1978), *Topography of Random Surfaces*, Nature 273, 573.
- BOX, G. E. P., and JENKINS, G. M., *Time Series Analysis, Forecasting and Control* (Holden-Day, San Francisco 1970).
- BROWN, S. R., and SCHOLZ, C. H. (1985), *Broad Bandwidth Study of the Topography of Natural Rock Surfaces*, J. Geophys. Res. 90, 12575–12582.
- BURROUGH, P. A. (1981), *Fractal Dimensions of Landscapes and Other Environmental Data*, Nature 294, 240–242.
- CHAYES, D. N. (1973), *Evolution of Sea MARC I*, IEEE Proc. of the third working Symposium on Oceanographic Data Systems, 103–108.
- FOX, C. G., and HAYES, D. E. (1985), *Quantitative Method for Analyzing the Roughness of the Sea-floor*, Rev. Geophys. Space Phys. 23, 1–48.
- HAMMOND, S. R., MALAHOFF, A., EMBLEY, R. W., CURRIE, R. G., DAVIS, E. E., RIDDIHOUGH, R. P., and SAWYER, B. S., *Preliminary Seabeam Bathymetry of the Juan de Fuca Ridge* (Earth Physics Branch, EMR, Open File Report 84-6, Canada 1984).
- JOURNEL, A. G., and HUIJBREGTS, Ch. J., *Mining Geostatistics* (Academic Press, New York 1978).
- MANDELBROT, B. B. (1967), *How Long is the Coast of Britain? Statistical Self-similarity and Fractional Dimension*, Science 156, 636–638.
- MANDELBROT, B. B., *The fractal Geometry of Nature* (W. H. Freeman and Company, New York 1983).
- MANDELBROT, B. B., and VAN NESS, J. W. (1968), *Fractional Brownian Motions, Fractional Noises and Applications*, SIAM Review 10, 422–437.
- MANDELBROT, B. B., and WALLIS, J. R. (1969), *Computer Experiments with Fractional Gaussian Noises*, Water Resources Res. 5, 228–267.
- McKENZIE, D. P. (1967), *Some Remarks on Heat Flow and Gravity Anomalies*, J. Geophys. Res. 72, 6261–6273.
- PARKER, R. L., and OLDENBURG, D. W. (1973), *Thermal Model of Ocean Ridges*, Nature Phys. Sci. 242, 137–139.
- PARSONS, B., and SCLATER, J. G. (1977), *An Analysis of the Variation of Ocean Floor Bathymetry and Heat Flow with age*, J. Geophys. Res. 82, 803–827.
- VOSS, R. F., *Random fractal forgeries*, In *Fundamental Algorithms for Computer Graphics*, NATO ASI Series, Vol. F17 (ed. Earnshaw, R. A.) (Springer-Verlag, Berlin Heidelberg 1985) pp. 805–835.

(Received April 29, 1987, revised/accepted September 11, 1987)

Fractal Dimension of Fault Systems in Japan: Fractal Structure in Rock Fracture Geometry at Various Scales

TAKAYUKI HIRATA^{1,2}

Abstract—Based on fault maps, whether or not the fracture geometry of rocks is self-similar, was examined by using a box-counting algorithm. The statistical self-similarity (fractal structure) of the fault fracture systems holds well at the scale of about 2 to 20 km. The fractal dimension in Japan varied from 1.05 to 1.60. The fractal dimension is about 1.5–1.6 at the central part of the Japan Arc, and decreases with distance from the center. At a smaller scale, the fractal structure also holds well in the rock fracture geometry. The fractal dimension of the North Izu Peninsula fault system (branching faults) is 1.49 at the scale of 0.625 to 10 km, the fractal dimension of rock fracture geometry at the scale order of 10^{-1} to 10^{-2} meters is about 1.49–1.61. The upper limit of the fractal dimension of rock fracture geometry is about 1.6, judging from the estimation of fractal dimension on actual fracture geometry of rocks. This value may impose a restraint on modeling of faulting and the fracture process of rocks.

Key words: Fracture geometry, fractal dimension, self-similarity.

Introduction

Stochastic self-similarity in rock fracture process has recently been recognized. The seismic process has a fractal structure at a broad scale in time, space and magnitude dimensions (in fracture size from magnitude $M = 1.5$ to largest earthquake, in space distribution from local regional catalog to worldwide catalog) (KAGAN and KNOPOFF, 1978, 1980, 1981; SADOVSKIY *et al.*, 1984). Gutenberg-Richter's frequency-magnitude relation indicates that earthquakes have a fractal structure in fracture size distribution (there is no characteristic fracture size) (AKI, 1981; KING, 1983). These self-similarities extend to the scale of micro-fracturing in a laboratory (MOGI, 1962; SCHOLZ, 1968; HIRATA *et al.*, 1987; HIRATA, 1987).

Based on the assumption that the geometry of the fault system is self-similar, KING (1983) discussed the geometrical origin of b-value, and interpreted the deformation of the earth's lithosphere by the movements of faults in a self-similar

¹ Department of Geophysics, Faculty of Science, Kyoto University, Sakyo, Kyoto 606, Japan.

² Now at Institute of Applied Physics, University of Tsukuba, Tsukuba, Ibaraki 305, Japan.

fault system. His study of the geometry of fault systems was based not on actual fault data, but on the seismicity (the magnitude-frequency relation). Although the fractal dimension has been calculated for the San Andreas fault (OKUBO and AKI, 1987; AVILES *et al.*, 1987), the emphasis was laid on the main fault line, and the fractal concept was only used as a tool to characterize the geometry of the fault line by the fractal dimension. Self-similarity (*i.e.*, an important profile of the fractal concept) in rock fracture geometry was not discussed adequately.

Does a fault system truly have a self-similar structure? We examined the actual fault maps by using a box-counting algorithm, and studied whether or not the fault system has a self-similar structure. After characterizing the geometry of the fault system by the fractal dimension, we will discuss the meaning of the fractal dimension.

Data

The data of the active faults used in this study were obtained from the sheet maps of the "Active Faults in Japan," published by the RESEARCH GROUP FOR ACTIVE FAULTS OF JAPAN (1980), based on the research on the active faults in Japan conducted cooperatively from 1976 through 1979 by forty-four researchers in geomorphology, geology and geophysics. This research group regarded any faults that had moved during the Quaternary (from 2 m.y. ago to the present) and those that could move in the future as active faults. The above publication gives the most complete map of Japanese faults currently available.

The active faults on land were selected primarily by careful interpretation of aerial photographs on a scale of 1:40,000, supplemented by geological maps and studies. Geological interpretation of aerial photographs has been most important for recognizing faults. Data were interpreted separately by two or more individuals to minimize any bias due to individual variation (MATSUDA *et al.*, 1977). In the "Active Faults in Japan," the certainty of the fault was classified into three ranks. The certainty of a fault refers to the reliability of both the reference plane (geomorphic surface such as a terrace surface, a slope, a low-relief erosional surface) and the reference line (geomorphic line such as ridge, drainage, a terrace edge) by which the amount of the displacement along the fault was estimated. Certainty I was a fault that was without doubt active during the Quaternary, certainty II was a fault whose activity was not as certain but which could be referred to have a displacement, and certainty III was a fault or mere lineament suspected of having been active during the Quaternary. In this study, the faults with a certainty of I, II and III were used. Sea faults were not used because the data uniformity was not as reliable as that of the land faults.

Fractal Dimension

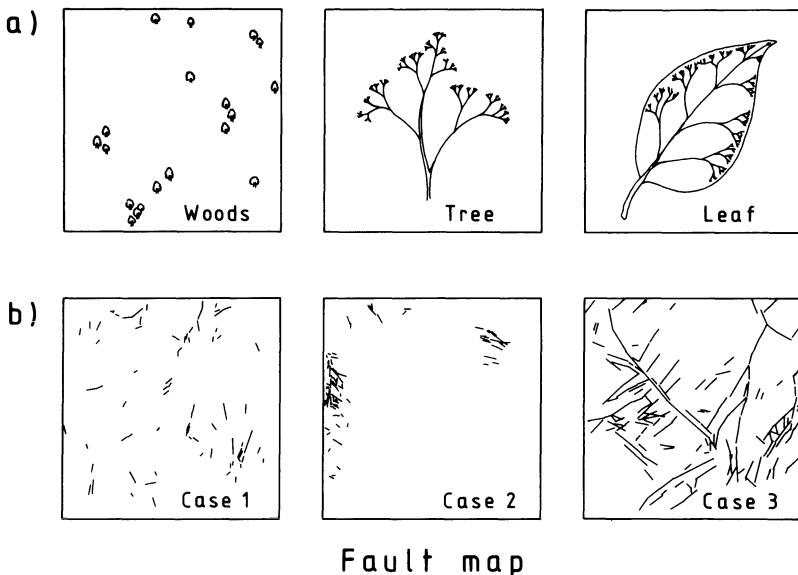
A fractal is the general concept for self-similarity introduced by MANDELBROT (1983), and is a powerful tool to characterize the geometry that has a self-similar structure. Many objects in nature, *e.g.*, clouds, mountains, coastlines, and trees, have complex shapes, which embarrass us when we want to describe their shapes quantitatively (for example, see Figure 1a). From the fractal point of view, however, we can find a self-similar structure in them and we can characterize their shapes by the fractal dimension (MANDELBROT, 1983). The geometry of the fault system also seems complex, and may be treated by a fractal approach.

Self-similarity manifests itself in a power law. Consider a fault system enclosed by a square region with a side length, R_0 (see Figure 2a). Divide the square region into (R_0/r^2) square boxes of side length, r . Let $N(r)$ be the number of boxes that the fault line enters. If a fault system has a self-similar structure, we get the following relation

$$N(r) \sim r^{-D} \quad (1)$$

where D is a fractal dimension (more strictly speaking, it is a capacity dimension (MANDELBROT, 1983)). Practically, $N(r)$ is plotted against r on double logarithmic scale, the graph is almost linear with slope $-D$. When relation (1) holds well, we can obtain the fractal dimensions from the slope of the graph. This is called the box-counting method (MANDELBROT, 1983). In this study, the box-counting method is used to obtain the fractal dimension of the fault system.

A fractal dimension is so convenient that we must be careful when we characterize the object by the fractal dimension. For example, consider a tree (see Figure 1a): a tree is a famous example of a fractal object in nature (MANDELBROT, 1983). Using the box-counting method, we can characterize the geometry of the branches of the tree by the fractal dimension. At the same time, we can also characterize the shape of the veins of a leaf by the fractal dimension. If we view the tree from an airplane, we can regard the tree as a point, and then we may also characterize the spatial distribution of the trees by the fractal dimension. These fractal dimensions at different scales can be obtained by using the same box-counting method, so we must note that structure in the object the fractal dimension reflects, especially when we treat the geometry over a broad range of scale, such as the fault system. Figure 1b shows the fault maps of the "Active Faults in Japan." Using the box-counting algorithm we can obtain the fractal dimension for the Cases 1, 2, and 3 fault maps in Figure 1b, but the meaning of each fractal dimension is different. For the Case 1 fault map, the fractal dimension is a measure of the spatial distribution of the faults. For the Case 3 fault map, the fractal dimension is a measure of the geometry of the fault system (branching geometry). For the Case 2 fault map, the meaning of the fractal dimension may be more complex (probably it



Fault map

Figure 1

Fractal objects in nature. Figure 1a; at different scales, tree shows different features. Figure 1b shows the fault maps at the same scale. However, fault maps may be classified into three groups: the geometry of Case 1 reflects the spatial distribution of the fault, the geometry of Case 3 reflects the fracture geometry, and then the geometry of Case 2 reflects both.

reflects both the spatial distribution and the geometry of the fault system). In this study, we classified the fault maps into two types: Type I fault map is the fault map that has large main faults (the fault length $> R_0/2 = 37.4$ km) and type II fault map is a fault map that has no large main fault. The type I fault map corresponds to Case 3 in Figure 1b and type II fault map corresponds to Case 1 or Case 2 in Figure 1b.

For type I fault maps, the fractal dimensions were calculated. Figure 2a shows the type I fault map (the square region of side length $R_0 = 74.8$ km). Changing the side length r of the box from 2.34 km ($= R_0/(2^5)$) to 37.4 km ($= R_0/2$), the number of boxes that the fault line enters $N(r)$ was counted. $\log N(r)$ was plotted against $\log r$, as shown in Figure 2b. The graph is almost linear between $r = 2.34$ km and $r = 18.7$ km, which indicates that the fault system has a self-similar structure over this range of scales. The fractal dimension D was calculated by the least-square method, and we obtained $D = 1.60 \pm 0.10$ (s.d.). Figure 3 shows the regional variation of the fractal dimensions for type I fault maps (only regions that have a standard deviation smaller than 0.12 in fractal dimension). The fractal dimension varied from 1.05 to 1.60. The fractal dimension at the central part of the Japan Arc was 1.5–1.6, and decreases with the distance from the center.

The fractal dimension of type I fault maps seem roughly corresponding to the density of faulting in Japan (highest in central Japan). Generally, D should be

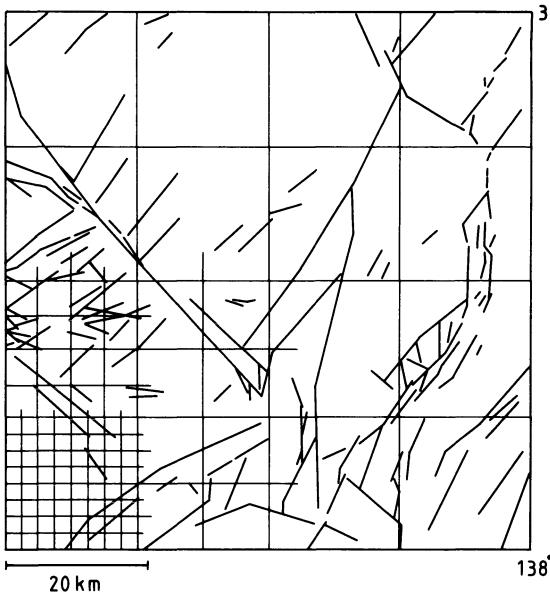


Fig. 2(a).

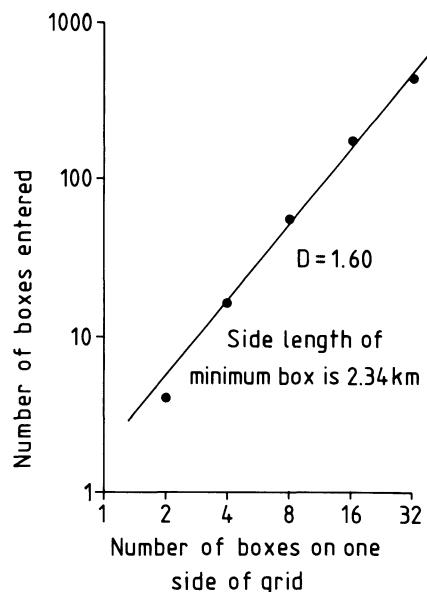


Fig. 2(b).

Figure 2

Fault systems of the Iida region shown in Figure 2a (the faults in the figure were drawn by tracing the active faults with a certainty of I to III in ref. "Active Faults in Japan"). The number of boxes entered by the lines of faults were counted, starting with a coarse grid of two large boxes on one side, then 2^n boxes, with n varying from 1 to 5. For clarity, the progressively finer divisions are only illustrated in one corner of the figure. The logarithm of the number of boxes entered by lines of faults was then plotted against the logarithm of the number of boxes along one side of the grid, as in Figure 2b. Fractal dimension is 1.60 ± 0.10 .

independent of density, so this seems to be a sampling problem. However, all type I fault maps used in this study are the high density regions of faults in Japan (see, MATSUDA, 1981). Moreover, the fractal dimension obtained at the scale range of 2.34–18.4 km was the fractal dimension of the branching geometry along the main fault. Therefore, the sampling does not produce an effect on the fractal dimension of type I fault maps.

For type II fault maps, the fractal dimensions were also calculated. Figure 4 shows the distribution of the fractal dimensions of regions with a standard deviation smaller than 0.12. Topological dimension of a fault D_T is 1, when the fault is considered as a straight line. Although the definition of fractal dimension requires $D > D_T$, there are the regions that have the fractal dimension $D < 1$. This may puzzle us. However, we must remember that the fractal dimensions for type II fault maps are concerned not with the geometry of the fracture system but the spatial distribution of the fault. For example, see Figure 5. Figure 5a is the fault map that has the smallest fractal dimension ($D = 0.72 \pm 0.01$ (s.d.)) in this study. The fault was considered as a broken line at the scale of the length of R_0 (in the extreme case,

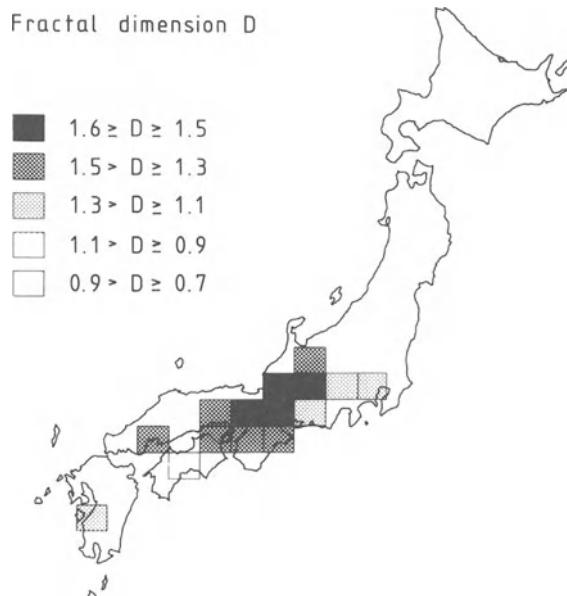


Figure 3

Regional variation of the fractal dimension in Japan Arc. The fractal dimension for type I fault maps were calculated, and only the regions with a standard deviation ≤ 0.12 are plotted.

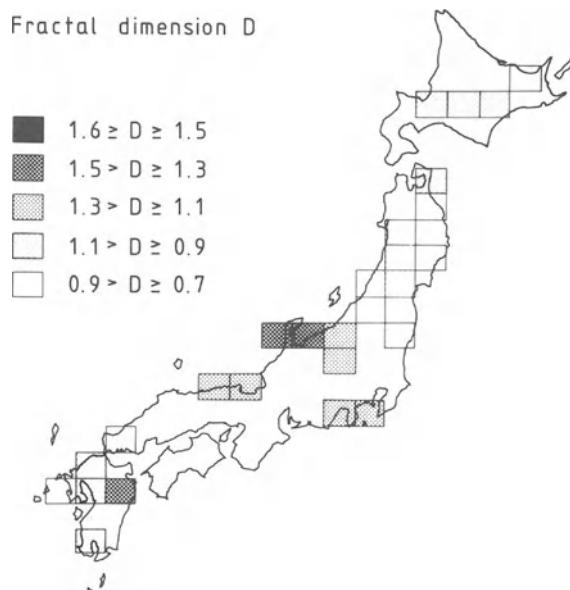


Figure 4

The fractal dimension for the type II fault maps. Only the regions with a standard deviation ≤ 0.12 are plotted.

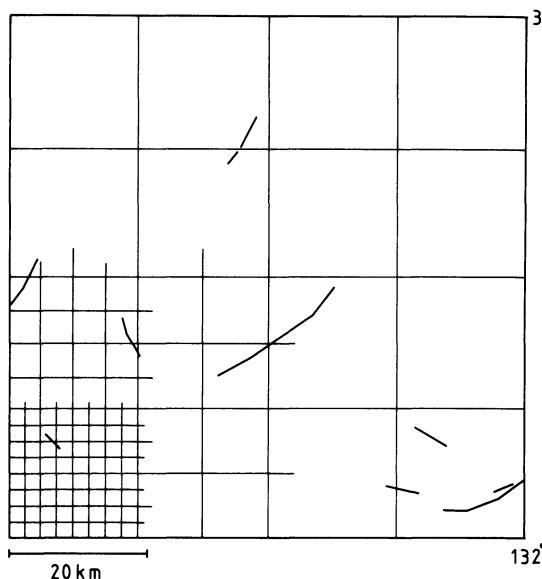


Fig. 5(a).

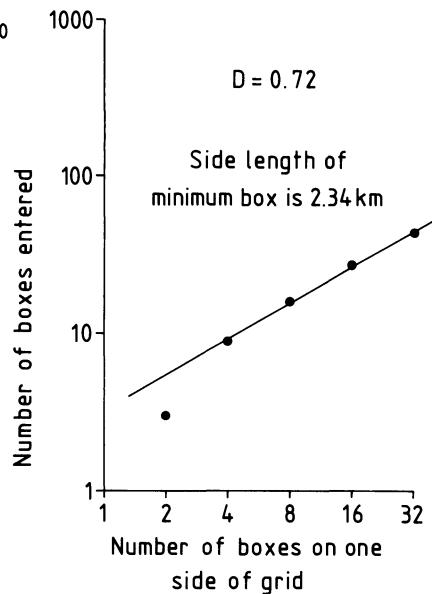


Fig. 5(b).

Figure 5

Fault system of Yamaguchi region, where the fractal dimension is the smallest value in Japan. The fractal dimension is 0.72 ± 0.01 .

the fault can be regarded as a point ($D_T = 0$). In such a case, the fractal dimension is controlled by the spatial distribution of faults, which probably coincides with the fractal dimension of the spatial distribution of the earthquake's epicenter in extreme case. However, the explanation for the fractal dimension of type II fault maps poses a problem because the fault actually has a finite length ($D_T = 1$) in spite of the treatment that the fault is a point ($D_T = 0$); this is a "thin fractal set" treatment. If we want to estimate the accurate fractal dimension of the spatial distribution of the faults based on the fault maps, we must treat the fault distribution as a "fat fractal set" (see, e.g., EYKHOLT and UMBERGER, 1986). This fractal dimension approach for the spatial distribution of the fault based on the fault map has an interesting feature, however, it will be difficult because of the data quality (that require the fault maps at the scale of 10^{-1} kilometers order). Moreover, the discussion for the fracture geometry in rocks (discussed so far) become dim. In this study, we used these fractal dimensions as a suggestion, and focus on the fractal dimension of the fracture geometry.

From the viewpoint of smaller scales, we can discuss the geometry of the fracture system based on the type II fault map. For example, we consider the faults in Izu peninsula, to be as shown in Figure 6a. This region has been investigated most intensively in Japan. The northern Izu earthquake ($M = 7.0$) occurred in 1930

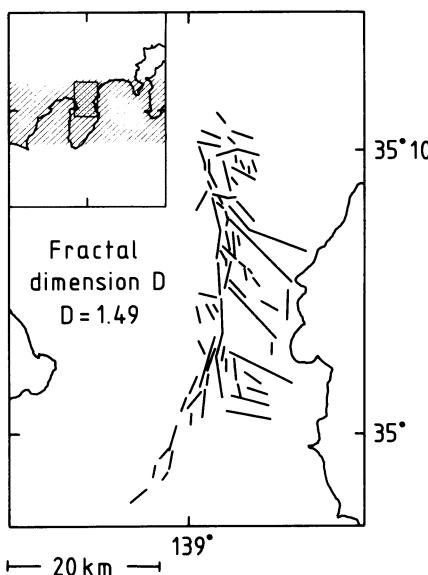


Fig. 6(a).

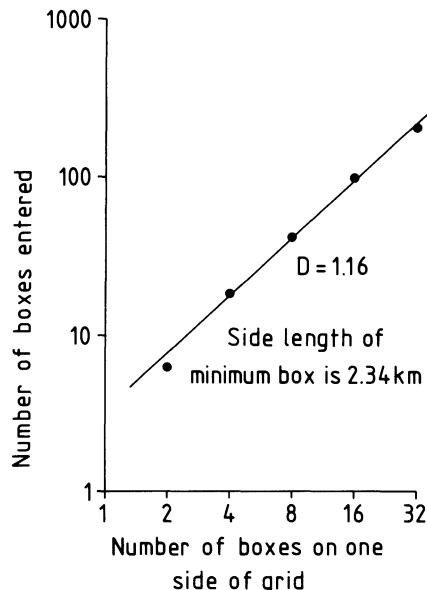


Fig. 6(b).

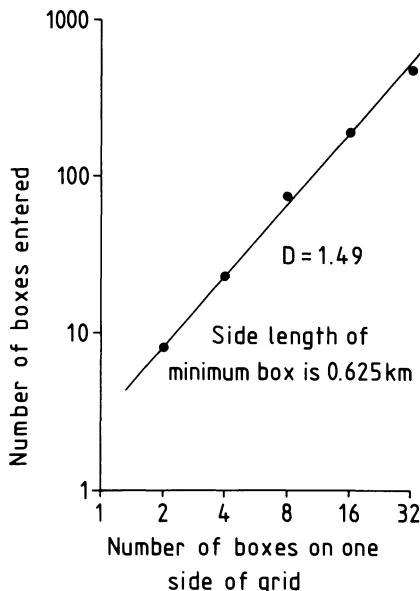


Fig. 6(c).

Figure 6

The North Izu fault group that is connected to the North Izu earthquake occurring in 1930. Figure 6b shows $\log N(r)$ vs. $\log r$ for the faults in the shadow region in Figure 6a. Fractal dimension is 1.16 ± 0.04 . Figure 6c shows $N(r)$ vs. $\log r$ for the North Izu fault group that is shown in Figure 6a. Fractal dimension is 1.49 ± 0.08 .

along Tanna, Ukihashi, Oono and Kadono faults (ABE, 1983). The motion on these faults is predominantly left-lateral. According to MATSUDA's field data (1972), the maximum offsets amount to 3.5 m along the Tanna fault (7 km long), 3.0 m along the central Ukihashi fault (4 km long), 1.5 m along the Oono fault (2.5 km long), and 2.0 m along the Kadono fault (1.5 km long). The average offset along the whole length is estimated as 2.9 m.

The fractal dimension of the geometry of the fault system (including Tanna, Ukihashi, Oono and Kadono faults) was estimated from the graph of $\log N(r)$ vs. $\log r$ (Figure 6c). Fractal dimension D is about 1.49 ± 0.08 (s.d.), which is higher than the fractal dimension of the fault system for the whole Izu region shown in Figure 6b ($D = 1.16 \pm 0.04$ (s.d.)). The fractal dimension at the scale 0.625 to 10 km is concerned with the geometry of branching faults. Unfortunately, it is impossible to estimate the fractal dimension in this scale range (about 500 meter order) for the whole of Japan because of incompleteness of current fault maps, and this remains a future problem.

Discussion

The fractal dimension of fault systems in Japan ranged from 1.05 to 1.60. The fractal dimension at the central part of the Japan Arc was about 1.5–1.6, and decreases with the distance from the center. This is explained by considering that the outer regions have few branch faults, and the geometry of the fault system approaches a line (main fault line). OKUBO and AKI (1987) showed that the fractal dimension for the entire San Andreas fault is 1.31, and the fractal dimension changed from 1.12 to 1.43 at the different segments. AVILES *et al.* (1987) found that the fractal dimension range from 1.0008 to 1.0191 along the main fault line of the San Andreas fault. The difference in the fractal dimension between OKUBO and AKI (1987) and AVILES *et al.* (1987) is due to the difference in the treatment of the branch faults. The fault data used by OKUBO and AKI (1987) include the branch faults, but that of AVILES *et al.* (1987) did not. The fractal dimension obtained in this study is the fractal dimension for the branching fault geometry, which reflects the branching rates of a mother-fault to daughter-faults, a daughter-fault to granddaughter-faults, a granddaughter-fault to great-granddaughter-faults, ... (the fractal dimension requires that they are the same rates).

The fractal dimension varied with the scale range concerned, and the meaning of the fractal dimension differed with the scale concerned: *e.g.*, in Figure 6, the fractal dimension ($D = 1.16$) obtained from the slope of the graph in Figure 6b reflects the spatial distribution of fault in the region, and the fractal dimension ($D = 1.49$) in Figure 6c is that of the geometry of the branching faults. This difference in the fractal dimension and the difference of the fractal dimension for the San Andreas fault between OKUBO and AKI (1987) and AVILES *et al.* (1987) suggest

that we must make clear what structure in the objects we want to characterize by fractal dimension.

Table 1 shows the values of the fractal dimension for the fracture geometry at various scale ranges. BROWN and SCHOLZ (1985) showed that the topography of natural rock surfaces (wavelength less than 20 microns to nearly 1 meter) can be characterized by the fractal dimension ($1.0 < D < 1.68$), and that the fractal dimension varies with the frequency band considered. SCHOLZ and AVILES (1986) showed that the fractal dimension ranged from 1 to 1.5 with the different scale bounds for the geometry of the San Andreas fault. To obtain the fractal dimension of rock fracture geometry at the scale order of 10^{-1} meter, we estimated the fractal dimension for the fracture system of Agematsu, and Ryu-o-Kyo areas in Japan mapped by WATANABE (1986). Figure 7a shows the fracture system of Agematsu area, and Figure 7b is the plot of $N(r)$ vs. r on a double logarithm scale. We obtained a geometry similar to that of the fault system shown in Figures 2a and 6a. Table 1 showed that there is an upper limit to the fractal dimension of the fracture geometry of rocks, of about 1.6.

What does this upper limit of the fractal dimension mean? Let us consider the case in which rock fracturing occurs obeying the Griffith energy balance concept: that is

$$U = (-W_L + U_E) + U_S \quad (2)$$

$$dU/dc = 0 \quad (3)$$

where c is the crack length, W_L is the amount of work done by external force, U_E is the strain potential energy reserved in the elastic body, and U_S is an expenditure energy of free surface energy to create a new crack surface. If U_E is almost constant

Table 1
Fractal dimension of the fracture geometry of rocks

Fault systems in Japan	$1.05 \leq D \leq 1.60$
San Andreas fault (SCHOLZ and AVILES, 1986)	$1.1 \leq D \leq 1.5$
North Izu fault group (branching geometry)	$D = 1.49$
*Fracture networks at Yucca Mountain (10^0 meter order, BARTON and LARSEN, 1985)	$1.12 \leq D \leq 1.16$
10^{-1} meter order fracture geometry (fracture map by WATANABE, 1985)	$1.4 \leq D \leq 1.6$
Topography of natural rock surface (BROWN and SCHOLZ, 1985)	$1 \leq D \leq 1.68$

*BARTON and CAMERON (1986) recently reported that the fractal dimension of fracture networks at Yucca Mountain ranged from 1.492 to 1.910 (at the same scale range of BARTON and LARSEN, 1985). Their joint and fault networks that were composed of many generations of fractures were composites. The very high value of the fractal dimension may be caused by this composites structure of fracture networks.

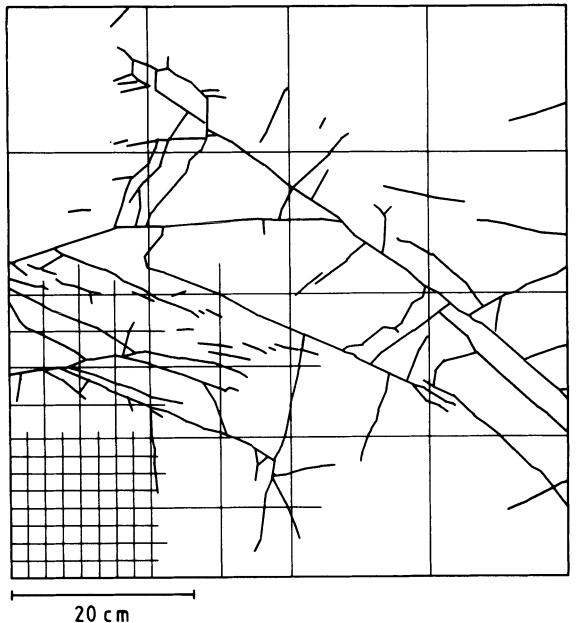


Fig. 7(a).

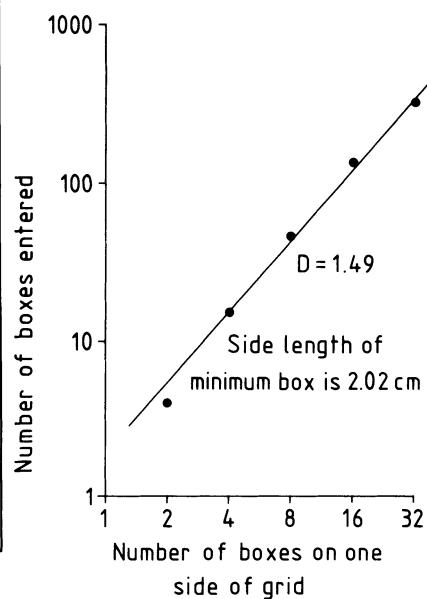


Fig. 7(b).

Figure 7

Fracture system of Agematsu region in Japan mapped by WATANABE (1986). Figure 7b is the plot of $\log N(r)$ vs. $\log r$. Fractal dimension is 1.49 ± 0.09 .

during the extension of crack and if the surface energy per unit area (U_s) is constant for the region, the Griffith energy balance concept tells us that W_L is proportional to the total area of the fracture surface that was produced by the crack extension. If we try to cover a square region (side length = R_0) with fractal dimension D by square planes (side length = $R_0/2^n$), we need 2^{nD} planes. The total area of fracture surface is $R_0^2 2^{(D-1)}$. (Fault maps can be regarded as the planner section of the fracture geometry of the crust in 3-D by the earth surface. The fractal dimension of the planner section D_{2D} is D_{3D} is the fractal dimension of 3-D geometry (MANDELBROT, 1983)). Table 2 shows the total area of fracture surface and the fractal dimension. For example, if self-similarity holds to the length of $R_0/32$, the fracture surface with $D = 3$ is 4 times larger than that with $D = 2.6$ (that corresponds to the upper value of the fractal dimension of fault fracture geometry $D = 1.6$). The production of a fracture surface with a high fractal dimension requires more external energy (NII *et al.*, 1985). If the fracture surface is made under minimum energy, the fractal dimension is 2. Crack paths in the materials including the obstacles depend on the relative strength, size and distribution of obstacles (see, for example, DAVIDGE and GREEN, 1968). Probably the asperities or barriers in the rock control the fractal dimension.

Table 2

Difference in the total area of fracture surface according to fractal dimension D

	$D = 3.0$	$D = 2.6$	$D = 2.4$	$D = 2.0$
R_0	1	1	1	1
$R_0/2$	2	1.5157	1.3195	1
$R_0/4$	4	2.2974	1.7411	1
$R_0/8$	8	3.4822	2.2974	1
$R_0/16$	16	5.2780	3.0314	1
$R_0/32$	32	8	4	1
$R_0/64$	64	12.1257	5.2780	1

* The amount of the areas are normalized by the unit area R_0^2 .

Fracture process of rock has been treated as a stochastic process by many investigators (e.g., OTSUKA, 1972; VERE-JONES, 1976; WATANABE, 1986). Goishi model of OTSUKA (1972) (that corresponds to a percolation model (SAITO *et al.*, 1973)) and the branching model of VERE-JONES (1976) and WATANABE (1986) give us the fracture geometry. Under a critical point, the fractal dimension of geometry of the percolation cluster is 1.56 in 2-dimension simulation, in spite of the fractal dimension $D = 1.89$ at the critical point (STAUFFER, 1985). We calculated the fractal dimension for the two-layer extension experiment carried out by KONDRA'TEV *et al.* (1985), and obtained the value of fractal dimension as about 1.6. In the experiment, wax or paraffin was used for the sample. The value of 1.6 appears not only in the fracture geometry of rocks but also in other material. Independence of the fractal dimension on the material suggests that we can treat the fracture process as a general theory.

The fractal dimension of the fracture geometry of rocks may be about 1.6 at all scales, and the decrease in the fractal dimension comes from both the lack of the branch faults in the observation and the contribution of the main fracture plane (the fractal dimension of the planner section of main fracture plane is almost 1). TAKAYASU (1985) proposed the deterministic model of fracture, and ascertained the fractal dimension to be about 1.65 for the fracture pattern. The fractal dimension of 1.6 suggests that Takayasu's model or a kind of percolation model under a critical point may be an important candidate for the general fracture theory.

Renormalization group approach for the fracture process is another powerful model for rock fracturing (ALLÈGRE *et al.*, 1982; MADDEN, 1983; SMALLEY *et al.*, 1985; TURCOTTE *et al.*, 1985). These models were supported by the rock fracture experiments (MOGI, 1962; SCHOLZ, 1968; HIRATA *et al.*, 1987; HIRATA 1987). TURCOTTE (1986) discussed the relationship between the fractal dimension and critical probability obtained from a renormalization-group approach model. The upper limit of the fractal dimension for the actual system obtained in the present

study imposes restraint on modeling of the faulting and fracture process of rocks. Barriers or asperities in the crust (or in the rocks) may control the fractal dimension of the fracture geometry of the rocks.

Acknowledgments

I would like to thank H. Takayasu, I. Tamada, N. Sumitomo, and S. Matsushima for their discussion and comments. I am especially grateful to K. Ito, T. Yanagidani, and T. Nakano for their discussion and advice. I am also grateful to C. H. Scholz and the reviewers for their constructive criticism of the original manuscript that significantly improved its contents.

REFERENCES

- ABE, K. (1978), *Dislocations, Source Dimensions and Stresses Associated with Earthquakes in the Izu Peninsula, Japan*, J. Phys. Earth 26, 253–274.
- AKI, K., *A probabilistic synthesis of precursor phenomena*, In *Earthquake Prediction* (eds. Simpson, D. W. and Richards, P. G.) (Maurice Ewing Volume 4, AGU, Washington, D. C. 1981) pp. 566–574.
- ALLÈGRE, C. J., LE MOUEL, J. L., and PROVOST, A. (1982), *Scaling Rules in Rock Fracture and Possible Implications for Earthquake Prediction*, Nature 297, 47–49.
- AVILES, C. A., SCHOLZ, C. H., and BOATWRIGHT, J. (1987), *Fractal Analysis Applied to Characteristic Segments of the San Andreas Fault*, J. Geophys. Res. 92, 331–344.
- BARTON, C. C., and LARSEN, E. (1985), *Fractal geometry of two-dimensional fracture networks at Yucca Mountain, Southwestern Nevada*, In *Fundamentals of Rock Joints*, Proc. Int. Sympo. on Fundamentals of Rock Joints, (ed. Stephansson, Ove), pp. 77–84.
- BARTON, C. C., and CAMERON, B. G. (1986), *Fractal Scaling of Fracture and Fault Maps at Yucca Mountain, Southern Nevada*, EOS (Transactions of the American Geophysical Union) 67 (44), 870.
- BROWN, S. R., and SCHOLZ, C. H. (1985), *Broad Bandwidth Study of the Topography of Natural Rock Surfaces*, J. Geophys. Res. 90, 12575–12582.
- DAVIDGE, R. W. and GREEN, T. J. (1968), *The Strength of Two-phase Ceramic/Glass Material*, J. Mater. Sci. 3, 629–634.
- EYKHOLT, R., and UMBERGER, D. K. (1986), *Characterization of Fat Fractals in Nonlinear Dynamical Systems*, Phys. Rev. Lett. 57, 2333–2336.
- HIRATA, T., SATOH, T., and ITO, K. (1987), *Fractal Structure of Spatial Distribution of Microfracturing in Rock*, Geophys. J. R. Astr. Soc. 90, 369–374.
- HIRATA, T. (1987), *Omori's Power Law Aftershock Sequence of Microfracturing in the Rock Fracture Experiment*, J. Geophys. Res. 92, 6215–6221.
- KONDRA'T'EV, V. N., KULUKIN, A. M., PONOMAREV, V. S., ROMASHOV, A. N., and CHUBAROV, V. M. (1985), *Investigation of a Two-layer Model of the Earth's Crust for Two-axis Extension of the Lower Layer*, Izvestiya, Earth Physics 21 (3), 176–183.
- KAGAN, Y. Y., and KNOPOFF, L. (1978), *Statistical Study of the Occurrence of Shallow Earthquakes*, Geophys. J. R. Astr. Soc. 55, 67–86.
- KAGAN, Y. Y., and KNOPOFF, L. (1980), *Spatial Distribution of Earthquakes: The Two-point Correlation Function*, Geophys. J. R. Astr. Soc. 62, 303–320.
- KAGAN, Y. Y., and KNOPOFF, L. (1981), *Stochastic Synthesis of Earthquake Catalogs*, J. Geophys. Res. 86, 2853–2862.
- KING, G. (1983), *The Accommodation of Large Strains in the Upper Lithosphere of the Earth and Other Solids by Self-similar Fault Systems: The Geometrical Origin of b-value*, Pure Appl. Geophys. 121, 761–815.

- MADDEN, T. R. (1983), *Microcrack Connectivity in Rocks: A Renormalization Group Approach to the Critical Phenomena of Conduction and Failure in Crystalline Rocks*, J. Geophys. Res. 88, 585–592.
- MANDELBROT, B. B., *The Fractal Geometry of Nature* (Freeman, New York 1983).
- MATSUDA, T., *Surface faults associated with Kita-Izu earthquake of 1930 in Izu Peninsula, Japan, in Izu Peninsula* (eds. Hoshino, M., and Aoki, H.) (Tokai Univ. Press, Tokyo 1972), pp. 73–93 (in Japanese).
- MATSUDA, T., OTA, Y., OKADA, A., SHIMIZU, F., and TOGO, M. (1977), *Aerial Photo-interpretation of Active Faults—The Individual Difference and Examples*, Bull. Earthq. Res. Inst. 52, 461–497.
- MATSUDA, T., *Active faults and damaging earthquakes in Japan—Macroseismic zoning and precaution fault zones*, In *Earthquake Prediction* (eds. Simpson, D. W. and Richards, P.G.) (Maurice Ewing Volume 4, AGU, Washington, D. C. 1981) pp. 279–289.
- MOGI, K. (1962), *Magnitude-frequency Relation for Elastic Shocks Accompanying Fractures of Various Materials and Some Related Problems in Earthquakes*, Bull. Earthq. Res. Inst. 40, 831–853.
- NII, Y., NAKAMURA, H., ITO, K., FUJII, N., and MATSUDA, J., *The fractal dimension of fractured fragments in relation with fracture energy*, In Proc. the 18th I.S.A.S. Lunar and Planet. Sympo. (eds. H. Mizutani, Oya, H., and Shimizu, M.) (Institute of Space and Astronautical Science, Tokyo 1985) pp. 58–59.
- OKUBO, P. G., and AKI, K. (1987), *Fractal Geometry in the San Andreas Fault System*, J. Geophys. Res. 92, 345–355.
- OTSUKA, M. (1972), *A Chain-reaction-type Source Model as a Tool to Interpret the Magnitude-frequency Relation of Earthquakes*, J. Phys. Earth 20, 35–45.
- RESEARCH GROUP FOR ACTIVE FAULTS OF JAPAN, *Active Faults in Japan, Sheet Maps and Inventories* (in Japanese) (Univ. of Tokyo Press, Tokyo 1980).
- SADOVSKIY, M. A., GOLUBEVA, T. V., PISARENKO, V. F., and SHNIRMAN, M.G. (1984), *Characteristic Dimensions of Rock and Hierarchical Properties of Seismicity*, Izvestiya, Earth Physics 20, 87–96.
- SAITO, M., KIKUCHI, M., and KUDO, K. (1973), *Analytical Solution of “Go-Game Model” of Earthquake*, Zisin 26, 19–25, (in Japanese).
- SCHOLZ, C. H., (1968), *Microfractures, Aftershocks, and Seismicity*, Bull. Seismol. Soc. Am. 58, 1117–1130.
- SCHOLZ, C. H., and AVILES, C. A. *The fractal geometry of faults and faulting*, In *Earthquake Source Mechanics* (eds. Das, S., Boatwright, J., and Scholz, C. H.) (Maurice Ewing Volume 6, AGU, Washington, D. C. 1986) pp. 147–155.
- SMALLEY, R. F., TURCOTTE, D. L., and SOLLA, S. A. (1985), *A Renormalization Group Approach to the Stick-slip Behavior of Faults*, J. Geophys. Res. 90, 1894–1900.
- STAUFFER, D., *Introduction to Percolation Theory* (Taylor and Francis, London 1985).
- TAKAYASU, H. (1985), *A Deterministic Model of Fracture*, Prog. Theor. Phys. 74, 1343–1345.
- TURCOTTE, D. L., SMALLEY, R. F., and SOLLA, S. A. (1985), *Collapse of Loaded Fractal Trees*, Nature 313, 671–672.
- TURCOTTE, D. L. (1986), *Fractals and Fragmentation*, J. Geophys. Res. 91, 1921–1926.
- VERE-JONES, D. (1976), *A Branching Model for Crack Propagation*, Pure Appl. Geophys. 114, 711–725.
- WATANABE, K. (1986), *Stochastic Evaluation of the Two Dimensional Continuity of Fractures in a Rock Mass*, Int. J. Rock Mech. Min. Sci. and Geomech. Abstr. 23, 431–437.

(Received April 17, 1987, revised September 30, 1987, accepted October 2, 1987)

Fractals in Geology and Geophysics

DONALD L. TURCOTTE¹

Abstract—The definition of a fractal distribution is that the number of objects N with a characteristic size greater than r scales with the relation $N \sim r^{-D}$. The frequency-size distributions for islands, earthquakes, fragments, ore deposits, and oil fields often satisfy this relation. Fractals were originally introduced by Mandelbrot to relate the length of a coastline to the length of the measuring stick. This application illustrates a fundamental aspect of fractal distributions, scale invariance. The requirement of an object to define a scale in photographs of many geological features is one indication of the wide applicability of scale invariance to geological problems; scale invariance can lead to fractal clustering. Geophysical spectra can also be related to fractals; these are self-affine fractals rather than self-similar fractals. Examples include the earth's topography and geoid.

Key words: Fractals, self-similarity, scale invariance.

Introduction

The concept of fractals was introduced by MANDELBROT (1967) in a geological context. Noting that the length of a rocky coastline increased as the length of the measuring rod decreased according to a power law, he associated the power with a fractal (fractional) dimension. The basis of this observation is that the rocky coastline is scale invariant. For example, photographs of a rocky coastline taken at altitudes of 10 m, 1 km, or 100 km are indistinguishable unless a scale is introduced.

The scale invariance of geological phenomena is one of the first concepts taught to a student of geology. It is pointed out that an object with a scale, *i.e.*, a coin, a rock hammer, a person, must be included whenever a photograph of a geological feature is taken. Without the scale it is often impossible to determine whether the photograph covers 10 cm or 10 km. For example, self-similar folds occur over this range of scales.

MANDELBROT (1982) has used fractal concepts to generate synthetic landscapes that look remarkably similar to actual landscapes. The fractal dimension is a measure of the roughness of the features. Because scale invariance is such a universal

¹ Department of Geological Sciences, Snee Hall, Cornell University, Ithaca, NY 14853, U.S.A.

phenomena in geology, it is not surprising that fractal distributions can be applied in a variety of ways to geological problems (BURROUGH, 1981, 1984).

Although fractal distributions would be useful simply as a means of quantifying scale-invariant distributions, it is now becoming evident that their applicability has a more fundamental basis. Statistical distributions have many fundamental applications in physics. One example is statistical mechanics. For example, the molecules in a gas have a statistical distribution of velocities. The temperature and entropy are measures of the average properties of the distribution. In quantum mechanics the location and momentum of an atomic particle have a statistical distribution. These statistical distributions are required to understand a variety of macroscopic phenomena. LORENZ (1963) derived a set of total, nonlinear differential equations that approximate thermal convection in a fluid. This set of equations was the first to exhibit chaotic behavior. Infinitesimal variations in initial conditions led to order one differences in the solutions obtained. Since the exact specification of initial conditions is precluded by quantum mechanics as well as the difficulty of making highly accurate physical measurements, it is concluded that the evolution of a wide variety of physical systems cannot be predicted in detail. Weather is an example. An important aspect of many of the dynamical systems that exhibit chaotic behavior is that they also exhibit fractal behavior (FARMER *et al.*, 1983). Fluid turbulence remains one of the major unsolved problems in physics. Major efforts are now being made to apply fractal concepts in order to better understand turbulence.

Another relatively new approach to the solution of a variety of physical problems is the renormalization group method. This method is particularly successful in predicting catastrophic changes. Examples include changes of phase, the onset of magnetic behavior, and predictions of the onset of permeability and electrical conductivity in media made up of networks. A characteristic of renormalization group solutions is that they often include fractal behavior.

Definition of a Fractal Distribution

If the number of objects N_i with a characteristic linear dimension r_i satisfies the relation

$$N_i = \frac{C}{r_i^D} \quad (1)$$

a fractal distribution is defined with D the fractal dimension and C a constant of proportionality. Fractal concepts can also be applied to a statistical distribution of objects. If the number of objects N with a characteristic linear dimension greater than r satisfies the relation

$$N = \frac{C}{r^D} \quad (2)$$

a fractal distribution is defined.

As mathematical representations (1) or (2) could be valid over an infinite range; however, for any physical application there will be upper and lower limits on the applicability of the fractal distribution. The essential feature of the fractal distribution is its scale invariance. No characteristic length scale enters into the definitions (1) or (2). If scale invariance extends over a sufficient range of length scales then the fractal distribution provides a useful description of the applicable statistical distribution. The fractal dimension D provides a measure of the relative importance of large versus small objects.

Under some circumstances different fractal scaling may be applicable at different scales. Also, other definitions of fractal distributions can be derived from (1) and (2). However, in a number of applications the basic definitions can be applied directly. For example, the Korčak empirical relation for the number of islands with area greater than a specified value fits (2) with $D = 1.30$ (MANDELBROT, 1975). We next turn to examples of direct applicability in fragmentation and seismology.

Fragmentation

A material can be fragmented in many ways. Rocks can be fragmented by weathering. The distribution of fragment sizes is likely to be related to the distribution of joints and other preexisting planes of weakness in the rock. Fragments can also be produced by explosives. Again, preexisting planes of weakness may determine the distribution of fragment sizes. Fragments can also be produced by impacts. Fragmentation is thought to play a dominant role in determining the size-frequency relation for asteroids and meteorites. A variety of statistical relationships have been used to correlate data on the size distribution of fragments; two of the most widely used distributions are the log normal and the power law. TURCOTTE (1986a) showed that the power-law distribution is equivalent in a variety of forms to the fractal distribution given in (2).

Several examples of power-law fragmentation are given in Figure 1. A classic study of the size-frequency distribution for broken coal was carried out by BENNETT (1936). The size-frequency for the chimney rubble above the PILEDRIVER nuclear explosion has been given by SCHOUTENS (1979). This was a 61-kT event at a depth of 457 m in granite. The size-frequency distribution for fragments resulting from the high velocity impact of a projectile on basalt has been given by FUJIWARA *et al.* (1977). In each example the fractal dimension for the distribution is near $D = 2.5$.

Further examples of power-law distributions of fragments are given in Table 1. It is seen that a great variety of fragmentation processes can be interpreted in terms of a fractal dimension. It has also been suggested that fault gouge has a fractal frequency-size distribution (SAMMIS *et al.*, 1986). It is seen that the values of the fractal dimension vary considerably but most lie in the range $2 < D < 3$. If the fractal dimension is less than 3 the volume integral diverges for large particles

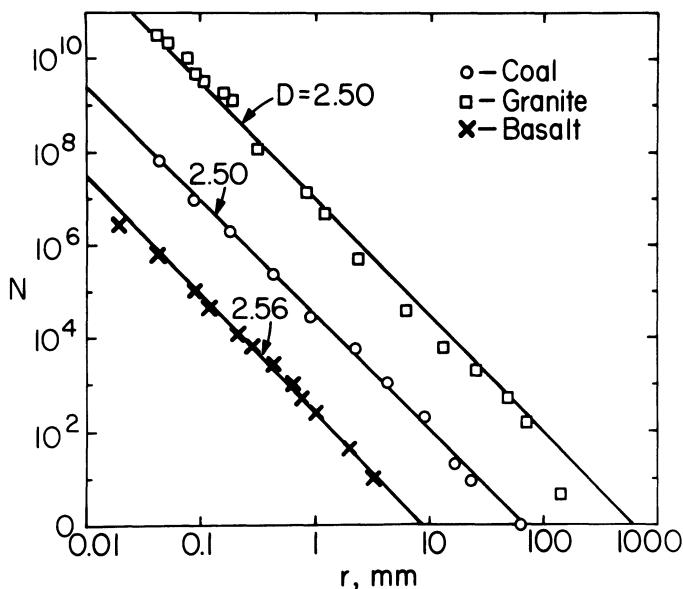


Figure 1

The number of fragments N with a characteristic linear dimension greater than $r = V^{1/3}$ as a function of r ; broken coal (BENNETT, 1936), broken granite resulting from a 61-kT underground nuclear detonation (SCHOUTENS, 1979), impact ejecta due to a 2.6 km/s polycarbonate projectile impacting on basalt (FUJIWARA *et al.*, 1977).

Table 1

Fractal dimensions for a variety of fragmented objects

Object	Reference	Fractal Dimension D
Projectile fragmentation of gabbro with lead	LANGE <i>et al.</i> (1984)	1.44
Projectile fragmentation of gabbro with steel	LANGE <i>et al.</i> (1984)	1.71
Meteorites (Prairie Network)	MCCROSKY (1968)	1.86
Artificially crushed quartz	HARTMANN (1969)	1.89
Disaggregated gneiss	HARTMANN (1969)	2.13
Disaggregated granite	HARTMANN (1969)	2.22
FLAT TOP I (chemical explosion, 0.2 kT)	SCHOUTENS (1979)	2.42
PILEDRIVER (nuclear explosion, 62 kT)	SCHOUTENS (1979)	2.50
Broken coal	BENNETT (1936)	2.50
Interstellar grains	MATHIS (1979)	2.50
Projectile fragmentation of quartizite	CURRAN <i>et al.</i> (1977)	2.55
Projectile fragmentation of basalt	FUJIWARA (1977)	2.56
Sandy clays	HARTMANN (1969)	2.61
Terrace sands and gravels	HARTMANN (1969)	2.82
Glacial till	HARTMANN (1969)	2.88
Stony meteorites	HAWKINS (1960)	3.00
Asteroids	DONNISON and SUGDEN (1984)	3.05
Ash and pumice	HARTMANN (1969)	3.54

$(V \rightarrow \infty)$; that is, the volume (mass) of small particles is negligible. If the fractal dimension is greater than 2 the area integral diverges for small particles $(V \rightarrow 0)$; that is, the area of small particles dominates. Volume (mass) is conserved upon fragmentation while area is not. However, the creation of area by fragmentation requires energy. Thus it is reasonable to hypothesize that the fractal dimension will increase with an increase in the energy density available for fragmentation.

In all cases it is expected that there will be an upper and lower limit on the validity of the fractal or power-law relation for fragmentation. The upper limit is generally controlled by the size of the object or region that is being fragmented. The lower limit is likely to be controlled by the scale of the heterogeneities responsible for fragmentation, *i.e.*, grain size.

A simple model illustrates how fragmentation can result in a fractal distribution. This model is illustrated in Figure 2. A cube with a linear dimension h is referred to as a cell; this cell is divided into eight cubic elements each with a dimension $h/2$. The probability that a cell will be fragmented into eight elements is f . Each of the fragmented elements with linear dimension $h/2$ is taken to be a first-order cell; each of these cells is divided into eight first-order cubic elements with a dimension $h/4$ as illustrated in Figure 2. The probability that a first-order cell will fragment into eight

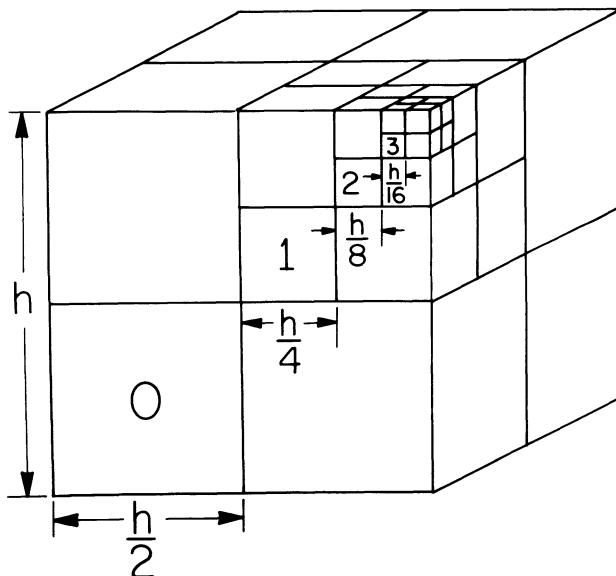


Figure 2

Idealized model for fractal fragmentation. A zero-order cubic cell with dimensions h is divided into eight cubic elements with dimensions $h/2$. The probability that a cell will be fragmented into eight elements is f . The fragments with dimensions $h/2$ become first-order cells; each of these have a probability f of being fragmented into first-order elements with dimensions $h/4$. The process is repeated to higher orders.

first-order elements is again taken to be f . The process is repeated at successively higher orders. The volume of the n -th order cell V_n is

$$V_n = \frac{1}{8^n} V_0 \quad (3)$$

where V_0 is the volume of the zero order cell. After fragmentation the number of zero order cells N_{0a} is

$$N_{0a} = (1-f)N_0 \quad (4)$$

where N_0 is the number of zero order before fragmentation. After fragmentation the number of n -th order cells N_n is

$$N_n = (8f)^n (1-f)N_0. \quad (5)$$

We can write (3), (4), and (5) as

$$\ln \left[\frac{V_n}{V_0} \right] = -n \ln 8 \quad (6)$$

$$\ln \left[\frac{N_n}{N_{0a}} \right] = n \ln(8f). \quad (7)$$

Elimination of n from (6) and (7) gives

$$\frac{N_n}{N_{0a}} = \left[\frac{V_n}{V_0} \right]^{-\ln(8f)/\ln 8} \quad (8)$$

Comparison with (1) shows that this is a fractal distribution with

$$D = 3 \frac{\ln(8f)}{\ln 8}. \quad (9)$$

Although the model is very idealized and nonunique, it illustrates the basic principles of scale-invariant fragmentation leading to a fractal distribution.

A number of authors (KATA and THOMPSON, 1985; KROHN and THOMPSON, 1986; DACCARD and LENORMAND, 1987) have suggested that sandstones have a fractal distribution of porosity. This can have an important influence on dispersion associated with two-phase flows.

Ore Grade and Tonnage

Statistical relations for the tonnage of ore with a grade (concentration) above a specified value have provided a basis for estimating ore reserves. The available data have been correlated using either log normal or power-law relations. This work has been reviewed by HARRIS (1984). TURCOTTE (1986b) showed that the power-law relations were equivalent to the fractal relation (2). If M is the tonnage of ore with

an average grade \bar{C} , the distribution is a fractal if

$$\bar{C} \sim M^{-D/3}. \quad (10)$$

It is of interest to compare (10) with available data.

The cumulative tonnage of mercury mined in the United States prior to a specified date was divided by the cumulative tonnage of ore by CARGILL *et al.* (1980) to give the cumulative average (weight) grade. Their results, given in Figure 3, are for five-year intervals. An excellent correlation with (10) is found for $D = 2.01$. A similar correlation for copper as given by CARGILL *et al.* (1981) is illustrated in Figure 4 and is compared with (10) taking $D = 1.16$. The data obtained prior to 1920 fall systematically low compared with the fractal correlation; however, this deviation can be attributed to a more efficient extraction process that was introduced at that time.

The simple model that was applied to fragmentation in Figure 2 can also be applied to the ore concentration problem. Instead of taking f to be the probability of fragmentation, it is taken to be the ratio of the concentration in an element to the concentration in a cell. As long as this concentration ratio is independent of scale the analysis given in (3) to (9) is applicable and a fractal grade-tonnage relation is found.

DREW *et al.* (1982) used the relation $N_{i-1} = 1.67 N_i$ in order to estimate the number of fields of order i , N_i , in the western Gulf of Mexico. Since the volume of oil in a field of order i is a factor of two greater than the volume of oil in a field of order $i - 1$, their relation is equivalent to a fractal distribution with $D = 2.22$.

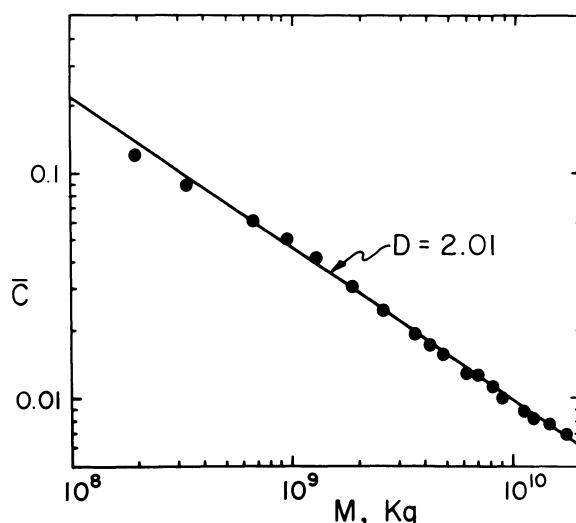


Figure 3

Dependence of ore grade on tonnage for mercury production in the United States (CARGILL *et al.*, 1980).

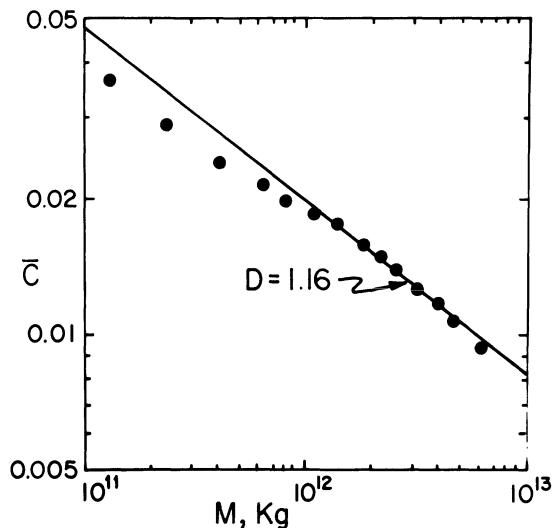


Figure 4

Dependence of ore grade on tonnage for copper production in the United States (CARGILL *et al.*, 1981).

Seismicity

Under many circumstances the number of earthquakes N with a magnitude greater than m satisfies the empirical relation (GUTENBERG and RICHTER, 1954)

$$\log N = -bm + a \quad (11)$$

where a and b are constants. Our discussion will be restricted to surface wave magnitudes. This relationship has been found to be applicable both regionally and on a world-wide basis. The b -value is widely used as a measure of regional seismicity. AKI (1981) showed that (11) is equivalent to the definition of a fractal distribution.

The moment of an earthquake is defined by:

$$M = \mu \delta A \quad (12)$$

where μ is the shear modulus, A the area of the fault break and δ is the mean displacement on the fault break. The moment of the earthquake can be related to its magnitude by

$$\log M = cm + d \quad (13)$$

where c and d are constants. KANAMORI and ANDERSON (1975) have established a theoretical basis for taking $c = 1.5$. These authors have also shown that it is a good approximation to take

$$M = \alpha r^3 \quad (14)$$

where $r = A^{1/2}$ is the linear dimension of the fault break. Combining (11), (13), and (14) gives

$$\log N = -2b \log r + \beta \quad (15)$$

with

$$\beta = \frac{bd}{1.5} + a - \frac{b}{1.5} \log \alpha \quad (16)$$

and (15) can be rewritten as

$$N = \beta r^{-2b}. \quad (17)$$

A comparison with the definition of a fractal given in (2) shows that

$$D = 2b. \quad (18)$$

Thus the fractal dimension of regional or world-wide seismic activity is simply twice the b value.

As an example of a regional frequency-magnitude distribution we consider seismicity in southern California as summarized by MAIN and BURTON (1986). Based on data from 1932 to 1972 the number of earthquakes per year N with magnitude greater than m is given as a function of m in Figure 5. In the magnitude range $4.25 < m < 6.5$ the data is in excellent agreement with (11) taking $b = 0.89$ or $D = 1.78$. With $c = 1.5$, $d = 16$, and $\alpha = 3.27 \times 10^7$ dyne/cm² the corresponding

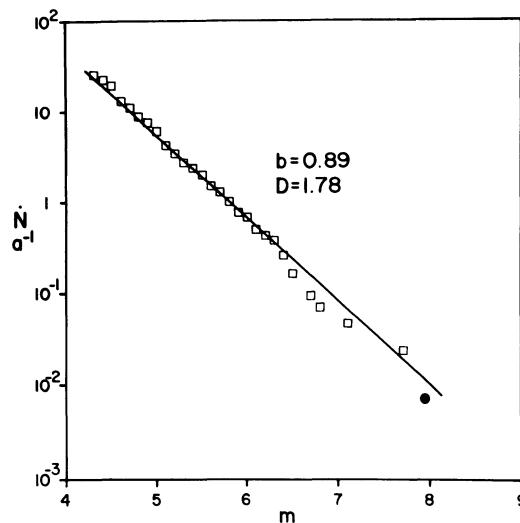


Figure 5

Number of earthquakes occurring per year N with a surface wave magnitude greater than m as a function of m . The open squares are data for southern California from 1932–1972 (MAIN and BURTON, 1986); the solid circle is the expected rate of occurrence of great earthquakes in southern California (SIEH, 1978).

range for r from (13) and (14) is $0.9 \text{ km} < r < 12 \text{ km}$. Data for smaller amplitude earthquakes; $m < 4.25$, would be expected to fit this relationship but in broad regions instrumented coverage is inadequate.

At magnitudes greater than 6.5 the data deviates from the fractal correlation. Two explanations for this deviation can be given:

1. The number of earthquakes with magnitude greater than 6.5 is so small (~ 6) that the statistics are not significant.
2. That geometrical self-similarity is not applicable to earthquakes greater than about 12 km since this is approximately the depth of the seismogenic zone in California and larger earthquakes are likely to have asymmetrical fault breaks.

Also included in Figure 5 is the N associated with great earthquakes on the southern section of the San Andreas fault as given by SIEH (1978); $m = 8.05$ and $N = 0.006a^{-1}$ (a repeat time of 163 years). An extrapolation of the fractal relation for regional seismicity appears to make a reasonable prediction of great earthquakes on this section of the San Andreas fault.

It should be emphasized that considerable caution must be exercised in using the fractal relation as a predictor for large earthquakes. Although the slope (b -value) in a particular region is likely to remain nearly constant, the level of seismicity may vary with time making any extrapolation subject to considerable error. For example, if a great earthquake reduces the stress level over a large region, the level of seismicity following the earthquake would be expected to decrease. Also, a largest characteristic earthquake would be expected in each region. The size of this earthquake would be related to the regional tectonics of the region. The seismic hazard would be strongly related to the magnitude of this largest earthquake, however, there is no reason to believe that the size would be related to other characteristics of the fractal distribution. Nevertheless, without the concept of fractals, a number of authors have used the fractal relation (11) to determine seismic hazards in a region (SMITH, 1976; MOLNAR, 1979; ANDERSON, 1979).

Tectonic models to explain the fractal distribution of seismicity have been proposed by KING (1983) and TURCOTTE (1986c). One limiting case is to attribute a characteristic earthquake to each fault. If this is the case, a fractal distribution of seismicity implies a fractal distribution of faults.

Perimeter Relations

Although we suggest that the basic definition of a fractal is the number-size relation given in (1), the original definition of a fractal given by MANDELBROT (1967) was the length of a trail or perimeter as a function of the step length. If the length of the step is r_i and if N_i is the number of steps required to obtain the length of the trail or perimeter P_i , we have

$$P_i = N_i r_i. \quad (19)$$

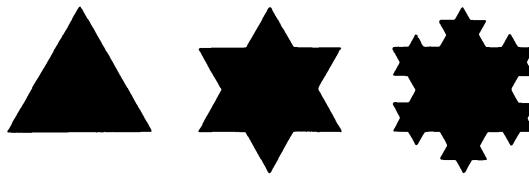


Figure 6

Triadic Koch island. A triangle with length $1/3$ is placed in the center of each side, $P_{i+1}/P_i = 4/3$, $r_{i+1}/r_i = 1/3$, $D = \ln 4/\ln 3$.

Substitution of (1) gives

$$P_i = Cr_i^{1-D}. \quad (20)$$

MANDELBROT (1967) showed that the west coast of Britain satisfied (20) with $D \approx 1.25$.

An example of the exact application of (20) is the triadic Koch island illustrated in Figure 6. The initial triangle with unit side length, $r_0 = 1$, has a perimeter $P_0 = 3$. Triangles one-third the size are placed on the centers of each side, thus $r_1 = 1/3$ and $P_1 = 12 \times 1/3 = 4$. The process is repeated with $r_2 = 1/9$ and $P_2 = 48 \times 1/9 = 16/3$. The fractal relation is exactly applicable with $D = \ln 4/\ln 3 = 1.2618$. Many other constructions with different fractal dimensions can be made.

Fractal dimensions for topography can be easily obtained from topographic maps. Several examples are given in Figure 7. The length along specified contours P is obtained using dividers of different lengths r . The results generally satisfy (20)

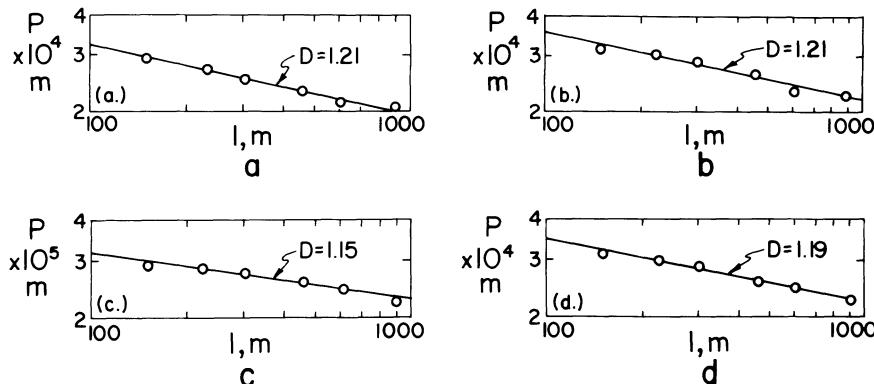


Figure 7

Fractal dimensions for specified topographic contours in several mountain belts, (a) 3,000 foot contour of the Cobblestone Mountain Transverse ranges, California ($D = 1.21$); (b) 5,400 foot contour of the Tatooh Buttes quadrangle, Cascade Mountains, Washington ($D = 1.21$); (c) 10,000 foot contour of the Byers Peak quadrangle, Rocky Mountains, Colorado ($D = 1.15$); (d) 1,000 foot contour of the Silver Bay quadrangle, Adirondack Mountains, New York ($D = 1.19$).

and a fractal dimension can be obtained. An important question is whether there are systematic variations in the fractal dimension. For example, does it depend systematically on the age of the geological province. It is seen in Figure 7 that quite diverse geological provinces have very similar fractal dimensions. Since most fractal dimensions are near $D = 1.2$ it does not appear possible to use fractal dimensions obtained by this method to characterize geological terrains.

AVILES *et al.* (1987) have used the length of trail technique to determine the fractal dimension of the San Andreas Fault. They found values ranging from $D = 1.0008$ to 1.0191. A similar analysis has been carried out by OKUBO and AKI (1987).

Another method of defining a fractal dimension is to relate perimeter to area with the relation

$$P \sim A^{D/2}. \quad (21)$$

LOVEJOY (1982) has shown that cross-sections of clouds obtained from radar or satellite photographs satisfy (21) with $D = 1.35$. RYS and WALDVOGEL (1986) have suggested that this relation is valid on scales greater than 3 km. Viscous fingering in the flow of immiscible fluids or two phases in porous media have been shown to produce fractal shapes (MALOY *et al.*, 1985; NITTMANN *et al.*; VAN DAMME *et al.*, 1986).

Many relationships in geomorphology have a fractal nature. An example is the relationship between the drainage area of a basin and the length of the principle river. For a number of drainage basins in the northeastern United States, HACK (1957) showed that (21) was satisfied with $D = 1.22$ as illustrated in Figure 8.

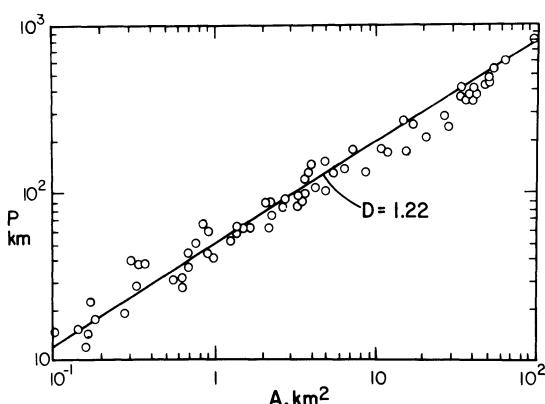


Figure 8

Dependence of the length of the principle river P on the area of the drainage basin A for several drainage basins in the northeastern United States (HACK, 1957).

Probability

We next relate fractal distributions to probability. A line has a Euclidian dimension of one while a point has a zero Euclidian dimension. The Cantor dust illustrated in Figure 9a has a fractal dimension between 0 and 1. The solid line of unit length is divided into three parts, $r = 1/3$, and the center third is removed, $N = 2$. From (1) $D = \ln 2/\ln 3 = 0.6309$. The process is repeated. The total line length decreases and this is referred to as curdling.

The probability that a step of length r includes a solid line is $Pr = 1$ when $r = 1$, $Pr = 2/3$ when $r = 1/3$, and $Pr = 4/9$ when $r = 1/9$. This is generalized to

$$Pr_i = N_i r_i \quad (22)$$

and substitution of (1) gives

$$Pr_i = r_i^{1-D}. \quad (23)$$

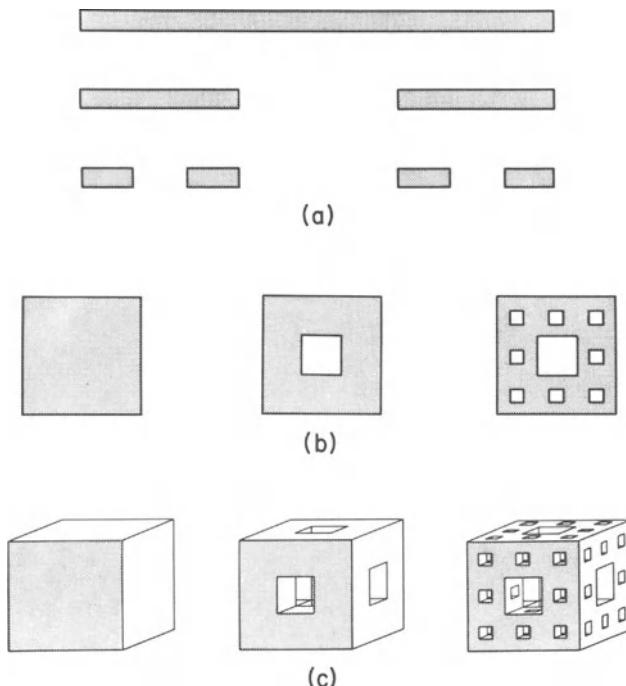


Figure 9

(a) Cantor dust. The middle one-third of each solid line is removed at each stage, $N_{i+1}/N_i = 2$, $r_{i+1}/r_i = 1/3$, $D = \ln 2/\ln 3$. (b) Sierpinski carpet. The middle one-third of each solid square is removed at each stage $N_{i+1}/N_i = 8$, $r_{i+1}/r_i = 1/3$, $D = \ln 8/\ln 3$. (c) Mengor sponge. Each solid cube has square passages cut through the center third of each face at each stage, $N_{i+1}/N_i = 20$, $r_{i+1}/r_i = 1/3$, $D = \ln 20/\ln 3$.

Although the Cantor dust is a determinant set, the probability relation (23) can be applied to statistical clustering.

The two-dimensional analog of the Cantor dust is the Sierpinski carpet illustrated in Figure 9b. The solid square of unit dimension has a square with dimensions $r = 1/3$ removed from its center. Eight solid squares with dimensions $r = 1/3$ remain so that $N = 8$. Thus from (1) $D = \ln 8/\ln 3 = 1.8928$. The process is repeated. The fractal dimension lies between 1 (the Euclidian dimension of a line) and 2 (the Euclidian dimension of an area). The probability that a square with dimensions r includes solids is $Pr = 1$ when $r = 1$, $Pr = 8/9$ when $r = 1/3$, and $Pr = 68/81$ when $r = 1/9$. This is generalized to

$$Pr_i = N_i r_i^2 \quad (24)$$

and substitution of (1) gives

$$Pr_i = r_i^{2-D}. \quad (25)$$

This result can be applied to clustering in an area.

The three-dimensional analog of the Cantor dust and Sierpinski carpet is the Mengor sponge illustrated in Figure 9c. A solid cube of unit dimensions has square passages with dimensions $r = 1/3$ cut through the centers of the six sides. Twenty solid cubes with dimensions $r = 1/3$ remain so that $N = 20$. Thus from (1) $D = \ln 20/\ln 3 = 2.7268$. The fractal dimension lies between 2 (the Euclidian dimension of an area) and 3 (the Euclidian dimension of a volume). The probability that a cube with dimensions r includes solid is $Pr = 1$ when $r = 1$, $Pr = 20/27$ when $r = 1/3$, and $Pr = 400/729$ when $r = 1/9$. This is generalized to

$$Pr_i = N_i r_i^3 \quad (26)$$

and substitution of (1) gives

$$Pr_i = r_i^{3-D}. \quad (27)$$

This result can be applied to clustering in a volume. Volumetric fractals such as the Mengor sponge can be used to develop models for porous flow through fractal distributions of permeability. This is particularly important in studies of dispersion (ALDER 1985a,b,c,d; ROSS, 1986). CURL (1986) has determined the fractal dimension of caves and has applied the Mengor sponge as a model.

The generalization of (23), (25), and (27) is

$$Pr_i = r_i^{d-D} \quad (28)$$

where d is the Euclidian dimension of the problem being considered. The probability methods introduced in this section will be used to obtain the fractal dimension of stationary, self-similar time series.

Cluster Analysis

The Cantor dust, Sierpinski carpet, and Mengor sponge form the basis for cluster analyses in one, two, and three dimensions. As a specific example of fractal cluster analysis in one dimension we consider the temporal distribution of N earthquakes occurring in a time interval τ_0 . The number of earthquakes is determined by the area of the region considered and the magnitude cut-off. For practical considerations the magnitude cut-off is made to ensure that all earthquakes in the region are detected by the regional network. The interval τ_0 is the interval over which data are collected. The natural period for the data set is τ_0/N .

In order to carry out a fractal cluster analysis we divide the interval τ_0 into a series of smaller intervals

$$\tau = \frac{\tau_0}{n}, \quad n = 2, 3, 4, \dots$$

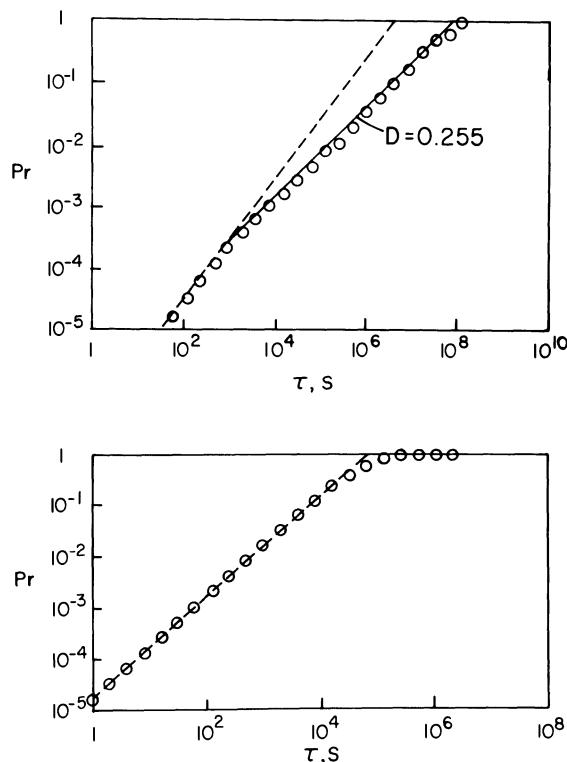


Figure 10

Fractal cluster analysis of 49 earthquakes that occurred near Efaté Island, New Hebrides in the period 1978 to 1984 (SMALLEY *et al.*, 1987). (a) The open circles fraction of intervals Pr of length τ with earthquakes as a function of τ . The solid line is the correlation with (23) taking $D = 0.255$. The dashed line is the result for uniformly spaced events. (b) The results for 49 randomly distributed events (Poisson process).

Our measure of clustering is the fraction Pr of the intervals in which an earthquake occurs as a function of τ . If Pr has a power law dependence on τ a fractal dimension for the clustering is defined.

This approach to fractal clustering of seismicity has been used by SADOVSKIY *et al.* (1985) and by SMALLEY *et al.* (1987). The latter authors considered the temporal variation of seismicity in several regions near Efate Island in the New Hebrides Island Arc for the period 1978 to 1984. One of their examples is given in Figure 10. During the period 49 events occurred in the region considered that satisfied the magnitude criteria. Time intervals between 8 minutes $\leq \tau \leq 524,288$ minutes were considered. The fraction of the intervals with earthquakes Pr as a function of the interval length τ is given in Figure 10a as the open circles. The solid line shows the correlation with the fractal relation (23) with $D = 0.255$. The dashed line is the result for uniformly spaced events. The results of a simulation for a random distribution of the 49 events in the time interval studied is given in Figure 10b. The random simulation (Poisson distribution) is significantly different than the earthquake data and is close to the uniform distribution.

Application to Time Series

Consider a single valued function of a variable, say $x(t)$, that is random but has a specified spectrum. For the case of a power law dependence of the spectral energy on frequency (wave number) it may be possible to define a fractal relation. We consider the specific case in which the increments $x(t + T) - x(t)$ satisfy the condition

$$Pr\left[\frac{x(t + T) - x(t)}{T^H} < x'\right] = F(x') \quad (29)$$

where H is a constant. In many examples $F(x')$ is a Gaussian. If this is the case and if $0 < H < 1$ then the random signal is known as fractional Brownian motion. For $H = 1/2$ we have Brownian motion.

Although the signal is continuous, it is not differentiable; thus it may be appropriate to consider the noise to be fractal in character. For the time series the fractal dimension would be between 1 and 2 so that the appropriate Euclidian dimension is 2. With $d = 2$, (28) can be written

$$\frac{Pr}{T^{2-D}} = \text{constant} \quad (30)$$

where the appropriate scale for the time series is T . Comparing (29) and (30) we define

$$H = 2 - D. \quad (31)$$

This is the basic definition of the fractal dimension for fractional Brownian motion.

The variance of the increment is defined by

$$V(T) = \langle [x(t + T) - x(t)]^2 \rangle \quad (32)$$

where the brackets $\langle \dots \rangle$ denote averages over many samples of $x(t)$. From (29), (31), and (32) we have

$$V(T) \sim T^{2H} \sim T^{(4 - 2D)}. \quad (33)$$

Random functions in time are often characterized by their spectral energy densities $S(f)$. If one defines $\bar{x}(f, t)$ as the Fourier transform of $x(t)$ for $0 < t < T$,

$$\bar{x}(f, t) = \frac{1}{T} \int_0^T x(t) e^{2\pi ift} dt \quad (34)$$

then

$$S(f) \sim T \bar{x}^2(f, t) \quad (35)$$

as $T \rightarrow \infty$. For Gaussian white noise $S(f) = \text{constant}$. In many cases of interest the spectral energy has a power law dependence on frequency

$$S(f) \sim f^{-\beta}. \quad (36)$$

The relationship between β and the fractal dimension is obtained from (32), (33), (35) and (36) with the result

$$S \sim f^{-\beta} \sim T^\beta \sim TV \sim T^{2H+1} \sim T^{5-2D} \sim f^{5-2D} \quad (37)$$

or

$$\beta = 5 - 2D. \quad (38)$$

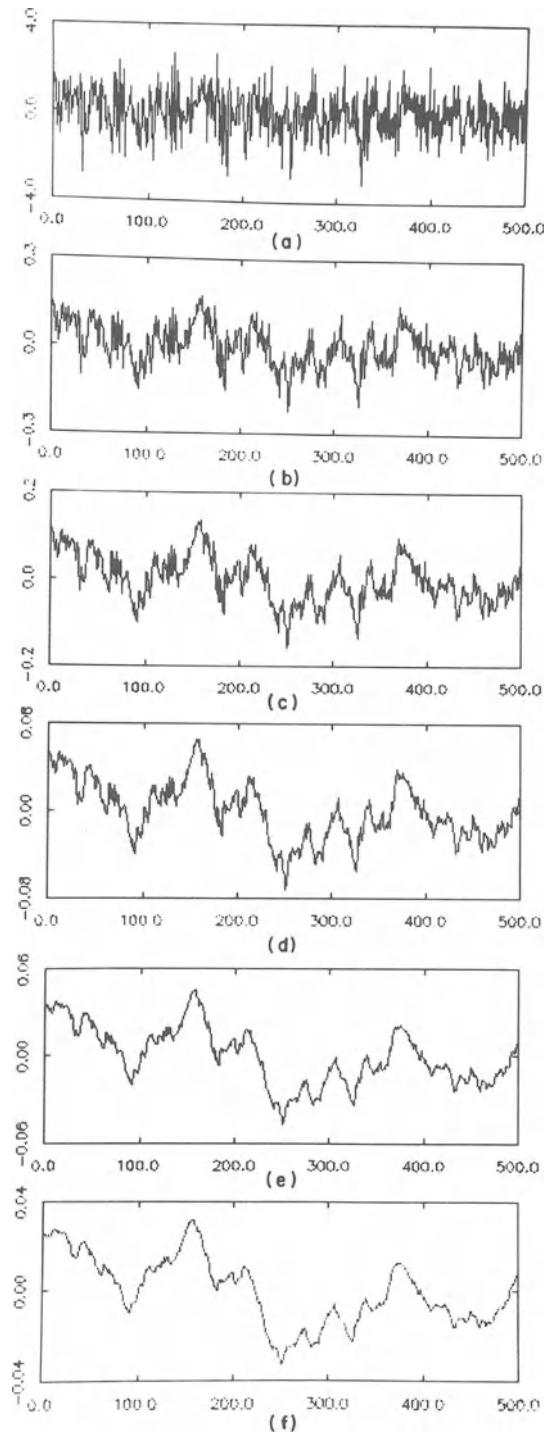
For a true fractal $1 < D < 2$ and $1 < \beta < 3$. For Brownian motion $\beta = 2$ and $D = 1.5$.

Examples of synthetically generated fractional Brownian motion are given in Figure 11. The method used to generate these random time series was as follows:

1. The time interval is divided into a large number of increments. Each increment is given a random number based on a Gaussian probability distribution. This is then a Gaussian white noise sequence.
2. The Fourier transform of the sequence is taken.
3. The resulting Fourier coefficients were filtered by multiplying by $f^{-\beta/2}$.
4. An inverse Fourier transform was taken using the filtered coefficients.
5. In order to remove edge effects (periodicities) only the central portion of the time series is retained.

Brownian motion is illustrated in Figure 11d.

An example of a continuum time series in geophysics is the earth's magnetic field. The power spectrum for the time variations of the north magnetic field at Victoria, British Columbia, Canada during July 1982 has been given by CHAVE *et al.* (1987). Their result is given in Figure 12 with robust averaging of the raw



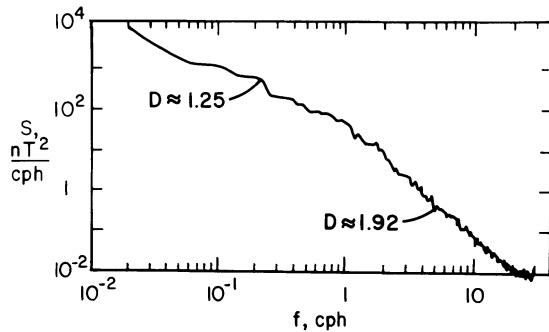


Figure 12

Power spectral density as a function of frequency for the north magnetic field at Victoria, British Columbia, Canada, during July 1982 (CHAVE *et al.*, 1987). A robust average of the raw spectra has been carried out to remove magnetic storms.

spectrum to remove high intensity magnetic storms. There appears to be a break in the slope at a frequency of about 1 cycle per hour. At lower frequencies the data appears to correlate reasonably well with (36) taking $\beta = 2.5$ ($D = 1.25$). At higher frequencies a correlation with (36) is reasonably good, taking $\beta = 1.15$ ($D = 1.92$).

In time series such as that illustrated in Figure 12 the dependent variable has a different dimension than the independent variable. In the example the amplitude of the magnetic field component is measured as a function of time. This illustrates the difference between self-similar and self-affine fractals (MANDELBROT, 1985). The perimeter relation defines a self-similar fractal. The spectral dependence of the magnetic field is an example of a self-affine fractal if it satisfies (36).

Synthetic fractal images can be generated using the same technique used to generate synthetic fractional Brownian motion. An example is given in Figure 13. The method used to generate the images was as follows:

1. Color coded random numbers ($0 - 1$) were generated on 256×256 pixels.
2. A two-dimensional Fourier transform generated a set of coefficients C_{nm} and S_{nm} .
3. A new set of coefficients was generated by taking $C'_{nm}k^{\beta/2}$ and $S'_{nm} = S_{nm}/k^{\beta/2}$ with $k = (k_x^2 + k_y^2)^{1/2}$.
4. Take an inverse Fourier transform to generate a new image.
5. Determine the fractal dimension from $D = 2 + (3 - \beta)/2$.

Examples with $D = 2.3$ and 2.5 , Figures 13d and e, resemble typical images of geological phenomena such as landsat or radar images of landforms.



Figure 11

Synthetically generated fractional Brownian motion. (a) Gaussian white noise sequence. (b) $\beta = 1.2$, $D = 1.9$. (c) $\beta = 1.6$, $D = 1.7$. (d) $\beta = 2.0$, $D = 1.5$, (Brownian motion). (e) $\beta = 2.4$, $D = 1.3$. (f) $\beta = 2.8$, $D = 1.1$.

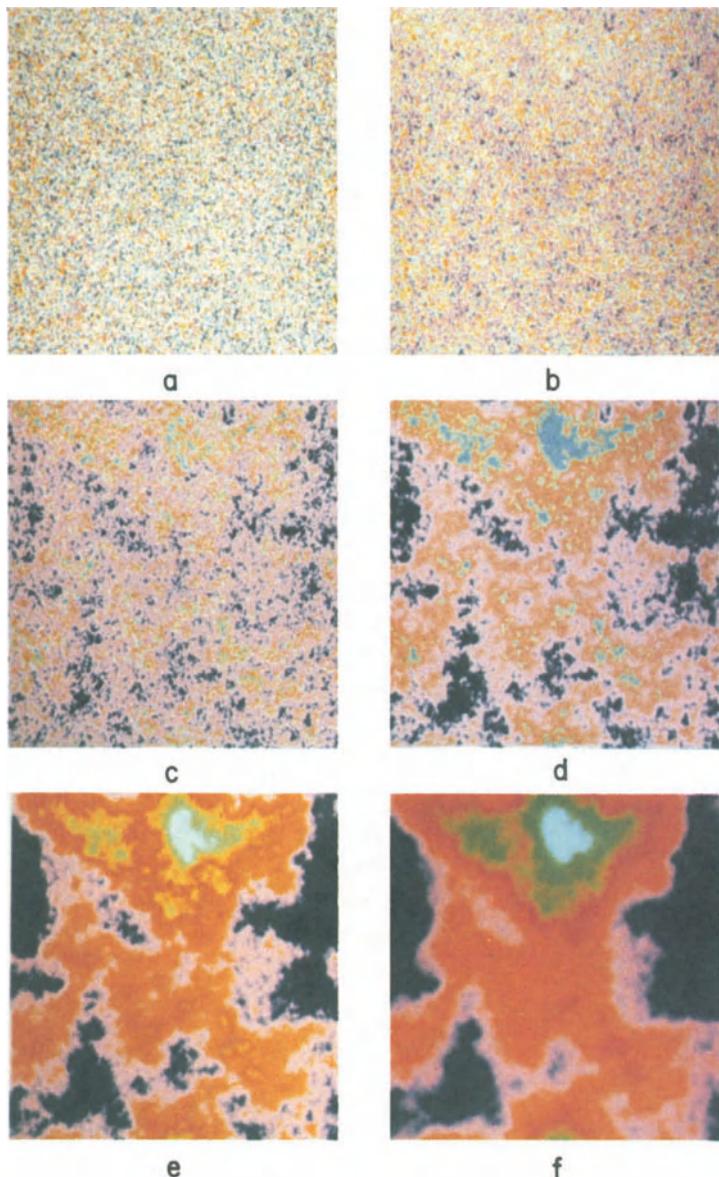


Figure 13

Synthetically generated images. (a) White noise image. (b) $\beta = 1.2$, $D = 2.9$. (c) $\beta = 1.6$, $D = 2.7$. (d) $\beta = 2.0$, $D = 2.5$. (e) $\beta = 2.4$, $D = 2.3$. (f) $\beta = 2.8$, $D = 2.1$.

Spherical Harmonics

It is common practice to expand data sets on the surface of the earth in terms of spherical harmonics; examples include topography and geoid. Using topography as an example the appropriate expansion for the radius r of the earth is

$$r(\theta, \phi) = a_0 \left[1 + \sum_{l=1}^{\infty} \sum_{m=0}^l (C_{lm} \cos m\phi + S_{lm} \sin m\phi) P_{lm}(\sin \theta) \right] \quad (39)$$

where a_0 is a reference radius, θ and ϕ are polar coordinates, C_{lm} and S_{lm} are coefficients, and P_{lm} are associated Legendre functions fully normalized so that

$$\frac{1}{4\pi} \int_0^{2\pi} \int_1^1 P_{lm}(\sin \theta) \begin{Bmatrix} \cos^2 m\phi \\ \sin^2 m\phi \end{Bmatrix} d(\sin \theta) d\phi = 1. \quad (40)$$

An identical expression can be used for the height of the geoid. The variance of the spectra for order l is defined by

$$V_l = a_0^2 \sum_{m=0}^l (C_{lm}^2 + S_{lm}^2) \quad (41)$$

and the power spectral density is defined by

$$S(k) = \frac{1}{k_0} V(k) = \lambda_0 V(k) \quad (42)$$

where k_0 is the wave number and $\lambda_0 = 1/k_0$ is the wave length over which data is included in the expansion. With $\lambda_0 = 2\pi a_0$ we have

$$S_l = 2\pi a_0^3 \sum_{m=0}^l (C_{lm}^2 + S_{lm}^2) \quad (43)$$

$$k_l = \frac{l}{2\pi a_0}. \quad (44)$$

A fractal dependence can be defined if S_l has a power law relation to k_l .

The power spectral density of the earth's topography is given in Figure 14. In addition to the spherical harmonic data of BALMINO *et al.* (1973) other spectral data for topography and bathymetry are included. An excellent correlation with (36) is obtained with $\beta = 2$ ($D = 1.5$). The spectral dependence of topography corresponds to Brown noise. This correlation has been previously noted by BELL (1975), BERKSON and MATTHEWS (1983), and FOX and HAYES (1985).

It is interesting to note that the fractal dimension associated with the topographic contours obtained by the perimeter method ($D = \approx 1.2$ as shown in Figure 7) differs from the fractal dimension obtained from the power spectral density of the earth's topography ($D \approx 1.5$ as shown in Figure 14). This difference is another example of self-similar and self-affine fractals. The vertical coordinate is not self-similar to the horizontal coordinates; gravity results in low angle slopes; perimeter relations cannot be derived for vertical cross-sections of topography.

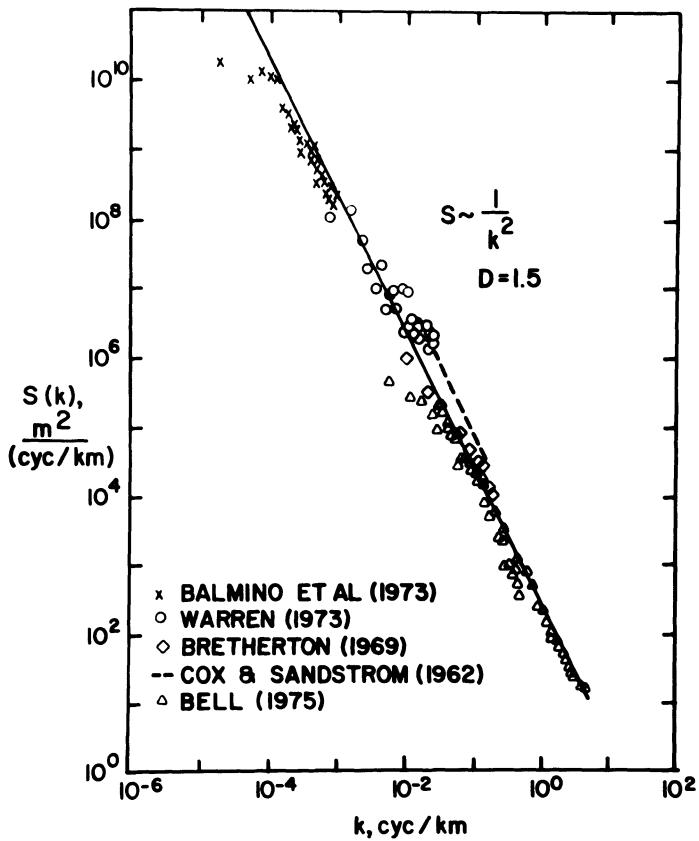


Figure 14

Power spectral density of the earth's topography (bathymetry) as a function of wave number. The correlation (solid line) is with (36) for $D = 1.5$ ($\beta = 2$).

Another example of a spherical harmonic application of fractals is illustrated in Figure 15. The power spectral density of the earth's geoid (gravitational potential) is given as a function of wave number (REIGBER *et al.*, 1985). An excellent correlation with (36) is obtained with $\beta = 3.5$ ($D = 0.75$); it should be noted that this lies outside the limits defined for a true fractal $1 < D < 2$. This correlation was previously noted by PHILLIPS and LAMBECK (1980) and is known as Kaula's law.

Conclusions

Scale invariance has long been recognized as an important feature of many geological problems. A fractal distribution is the only statistical distribution that is scale invariant. Thus it is not surprising that many geological and geophysical data

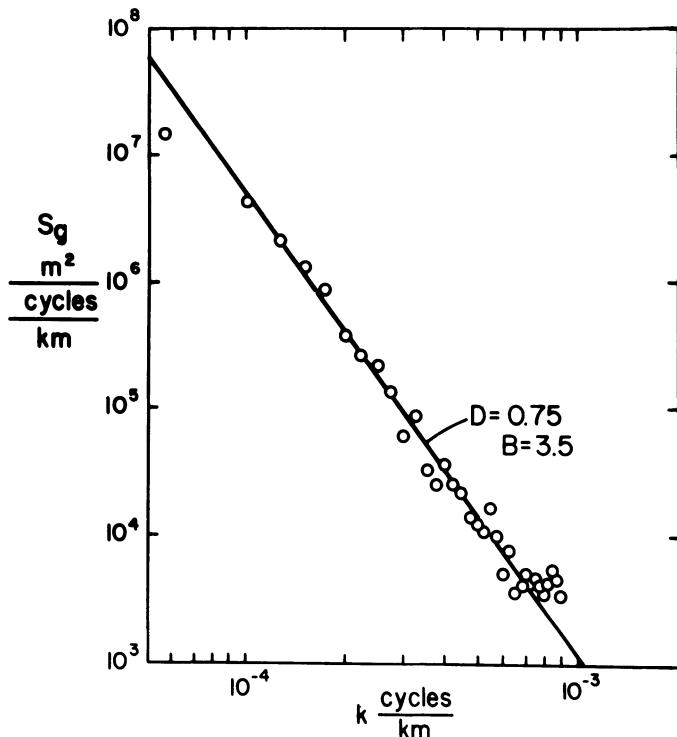


Figure 15

Power spectral density of the earth's geoid as a function of wave number. The correlation (solid line) is with (36) for $D = 0.75$ ($\beta = 3.5$).

sets are fractals. In this context fractals are a method of empirically correlating a variety of quantitative observations. Fractal concepts also suggest ways in which complex problems can be studied. Examples include:

1. Length of a perimeter as a function of the length of the measuring stick.
2. Perimeter length as a function of area.
3. Surface area as a function of volume.
4. Fraction of time increments that contain an event as a function of the length of the time increment.
5. Number of square boxes that contain material as a function of the box size.
6. Number of cubic boxes that contain material as a function of the box size.
7. Energy spectral density as a function of frequency or wave number.

In addition to being appropriate for empirical correlations, fractals also result from several classes of theoretical studies. These include:

1. Solutions to sets of differential equations and recursion relations that yield strange attractor behavior.
2. Solutions to problems obtained by using the renormalization group approach.

The use of fractal concepts in geology and geophysics is relatively new. Much of the work carried out so far represents a classification of existing data sets. However, fractal ideas should lead to new ways of collecting data. Also, a problem with fractal data indicates that a scale-invariant approach is necessary in order to obtain applicable solutions.

Acknowledgment

This research has been supported by the National Aeronautics and Space Administration under contract NAS 5-27340 and grant NAG 5-319.

REFERENCES

- ADLER, P. M. (1985a), *Transport Process in Fractals. I. Conductivity and Permeability of a Leibniz Packing in the Lubrication Limit*, Int. J. Multiphase Flow 11, 91-108.
- ADLER, P. M. (1985a), *Transport Processes in Fractals. II. Strokes Flow in Fractal Capillary Networks*, Int. J. Multiphase Flow 11, 213-239.
- ADLER, P. M. (1985c), *Transport Processes in Fractals. III. Taylor Dispersion in Two Examples of Fractal Capillary Networks*, Int. J. Multiphase Flow 11, 241-254.
- ADLER, P. M. (1985d), *Transport Processes in Fractals. IV. Nonlinear Flow Problems in Fractal Capillary Networks*, Int. J. Multiphase Flow 11, 853-871.
- AKI, K. (1981), *A probabilistic synthesis of precursory phenomena*, In *Earthquake Prediction* (Simpson, D. W., and Richards, P. G. eds.) (American Geophysical Union, Washington, D. C.) pp. 556-574.
- ANDERSON, J. G. (1979), *Estimating the Seismicity From Geological Structure for Seismic-risk Studies*, Seis. Soc. Amer. 69, 135-158.
- AVILES, C. A., SCHOLZ, C. H., and BOATWRIGHT, J. (1987), *Fractal Analysis Applied to Characteristic Segments of the San Andreas Fault*, J. Geophysical Res. 92, 331-334.
- BALMINO, C., LAMBECK, K., and KAULA, W. M. (1973), *A Spherical Harmonic Analysis of the Earth's Topography*, J. Geophys. Res. 78, 478-481.
- BELL, T. H. (1975), *Statistical Features of Sea-floor Topography*, Deep-Sea Res. 22, 883-892.
- BENNETT, J. G. (1936), *Broken Coal*, J. Inst. Fuel 10, 22-39.
- BERKSON, J. M., and MATTHEWS, J. E. (1983), *Statistical Properties of Sea-floor Roughness*, In *Acoustics and the Sea-Bed* (Pace, N. G. ed.) (Bath University Press, Bath) pp. 215-223.
- BRETHERTON, F. P. (1969), *Momentum Transport by Gravity Waves*, Quart. J. Roy. Meteor. Soc. 95, 213-243.
- BROWN, S. R., and SCHOLTZ, C. H. (1985), *Broad Bandwidth Study of the Topography of Natural Rock Surfaces*, J. Geophys. Res. 90, 12,575-12,582.
- BURROUGH, P. A., (1981), *Fractal Dimensions of Landscapes and Other Environmental Data*, Nature 294, 240-242.
- BURROUGH, P. A. (1984), *The Application of Fractal Ideas to Geophysical Phenomena*, Insti. Math. Applic. Bull. 20, 26-42.
- CARGILL, S. M., ROOT, D. H., and BAILEY, E. H. (1980), *Resources Estimation from Historical Data: Mercury, a Test Case*, J. Int. Assoc. Math. Geol. 12, 489-522.
- CARGILL, S. M., ROOT, D. H., and BAILEY, E. H. (1981), *Estimating Usable Resources from Historical Industry Data*, Econ. Geol. 76, 1081-1095.
- CHAVE, A. D., THOMSON, D. J., and ANDER, M. E. (1987), *On the Robust Estimation of Power Spectra, Coherences, and Transfer Functions*, J. Geophys. Res. 92, 633-648.
- COX, C., and SANDSTROM, H. (1962), *Coupling of Internal and Surface Waves in Water of Variable Depth*, J. Ocean. Soc. Japan 20, 499-513.

- CURL, R. L. (1986), *Fractal Dimensions and Geometrics of Caves*, J. Int. Assoc. Math. Geol. 18, 765–783.
- CURRAN, D. R., SHOCKEY, D. A., SEAMAN, L., and AUSTIN, M. (1977), *Mechanisms and models of cratering in earth media*, In *Impact and Explosion Carterway* (Eddy, D. J., Pepin, R. O., and Merrill, R. P., eds) (Pergamon Press, New York) pp. 1057–1087.
- DACCORD, G., and LENORMAND, R. (1987), *Fractal Patterns from Chemical Dissolution*, Nature 325, 41–43.
- DONNISON, J. R., and SUGDEN, R. A. (1984), *The Distribution of Asteroid Diameters*, Mon. Not. R. Astron. Soc. 210, 673–682.
- DREW, L. J., SCHUENEMEYER, J. H., and BAWIEC, W. J. (1982), *Estimation of the Future Rates of Oil and Gas Discoveries in the Western Gulf of Mexico*, U.S. Geol. Survey Prof. Paper 1252, 26 pp.
- FARMER, J. D., OTT, E., and YORKE, J. A. (1983), *The Dimension of Strange Attractors*, Physica 7D, 153–180.
- FOX, C. G., and HAYES, D. E. (1985), *Quantitative Methods for Analyzing the Roughness of the Seafloor*, Rev. Geophys. 23, 1–48.
- FUJIWARA, A., KAMIMOTO, G., and TSUKAMOTO, A. (1977), *Destruction of Basaltic Bodies by High-velocity Impact*, Icarus 31, 277–288.
- GUTENBERG, B., and RICHTER, C. F., *Seismicity of the Earth and Associated Phenomena*, 2nd ed. (Princeton University Press, Princeton 1954).
- HACK, J. T. (1957), *Studies of Longitudinal Stream Profiles in Virginia and Maryland*, U. S. Geol. Survey Prof. Paper 294B, 97 pp.
- HARRIS, D. P., *Mineral Resources Appraisal* (Oxford University Press 1984).
- HARTMANN, W. K. (1960a), *Terrestrial, Lunar, and Interplanetary Rock Fragmentation*, Icarus 10, 201–322.
- KANAMORI, H., and ANDERSON, D. L. (1975), *Theoretical Basis of Some Empirical Relations in Seismology*, Bull. Seis. Am. 65, 1073–1096.
- KATZ, A. J., and THOMPSON, A. H. (1985), *Fractal Sandstone Pores: Implications for Conductivity and Pore Formation*, Phys. Rev. Lett. 54, 1325–1328.
- KING, G. (1983), *The Accommodation of Large Strains in the Upper Lithosphere of the Earth and Other Solids by Self-similar Fault Systems: The Geometrical Origin of b-value*, Pure Appl. Geophys. 121, 761–815.
- KROHN, C. E., and THOMPSON, A. H. (1986), *Fractal Sandstone Pores: Automated Measurements Using Scanning-electron Microscope Images*, Phys. Rev. B33, 6366–6374.
- LANGE, M. A., AHRENS, T. J., and BOSLOUGH, M. B. (1984), *Impact Cratering and Spall Failure of Gabbro*, Icarus 58, 383–395.
- LORENZ, E. N. (1963), *Deterministic Nonperiodic Flow*, J. Atmos. Sci. 20, 130–141.
- LOVEJOY, S. (1982), *Area-perimeter Relation for Rain and Cloud Areas*, Science 216, 185–187.
- MAIN, I. G., and BURTON, P. W. (1986), *Long-term Earthquake Recurrence Constrained by Tectonic Seismic Moment Release Rates*, Seis. Soc. Am. Bull. 76, 297–304.
- MALOY, K. J., FEDER, J., and JOSSANG, T. (1985), *Viscous Fingering Fractals in Porous Media*, Phys. Rev. Lett. 55, 2688–2691.
- MANDELBROT, B. B. (1967), *How Long is the Coast of Britain? Statistical Self-similarity and Fractional Dimension*, Science 156, 636–638.
- MANDELBROT, B. B. (1975), *Stochastic Models for the Earth's Relief, the Shape and the Fractal Dimension of the Coastlines, and the Number-area Rule for Islands*, Proc. Natl. Acad. Sci. U.S.A. 72, 3825–3828.
- MANDELBROT, B. B., *The Fractal Geometry of Nature* (Freeman, San Francisco 1982).
- MANDELBROT, B. B. (1985), *Self-affine Fractals and Fractal Dimension*, Phys. Scripta 32, 257–260.
- MATHIS, J. S. (1979), *The Size Distribution of Interstellar Particles, II. Polarization*, Astrophys. J. 232, 747–753.
- MCCROSKEY, R. E. (1968), *Distributions of Large Meteoric Bodies*, Spec. Rep. 280, (Smithson. Astrophys. Observ., Washington, D. C.).
- MOLNAR, P. (1979), *Earthquake Recurrence Intervals and Plate Tectonics*, Seis. Soc. Am. Bull. 69, 115–133.
- NITTMANN, J., DACCORD, G., and STANLEY, H. E. (1985), *Fractal Growth of Viscous Fingers: Quantitive Characterization of a Fluid Instability Phenomenon*, Nature 314, 141–144.

- OKUBO, P. G., and AKI, A. (1987), *Fractal Geometry in the San Andreas Fault System*, J. Geophys. Res. 92, 345-355.
- PHILLIPS, R. J., and LAMBECK, K. (1980), *Gravity Fields of the Terrestrial Planets: Long-wavelength Anomalies and Tectonics*, Rev. Geophys. Space Phys. 18, 27-76.
- REIGBER, C., BALMINO, G., MULLER, H., BOSCH, W., and MOYNOT, B. (1985), *GRIM Gravity Model Improvement Using LAGEOS (GRIM 3-L1)*, J. Geophys. Res. 90, 9285-9299.
- ROSS, B., (1986), *Dispersion in Fractal Fracture Networks*, Water Resour. Res. 22, 823-827.
- RYS, F. S., and WALDVOGEL, A. (1986), *Fractal Shape of Hail Clouds*, Phys. Rev. Lett. 56, 784-787.
- SADOVSKIY, M. A., GOLUBEVA, T. V., PISAENKO, V. F., and SHNIRMAN, M. G. (1985), *Characteristic Dimensions of Rock and Hierarchical Properties of Seismicity*, Phys. Solid Earth 20, 87-95.
- SAMMIS, C. G., OSBORNE, R. H., ANDERSON, J. L., BANERDT, M., and WHITE, P. (1986), *Self-similar Cataclasis in the Formation of Fault Gauge*, Pure Appl. Geophys. 123, 53-78.
- SCHOUTENS, J. E. (1979), *Empirical Analysis of Nuclear and High-explosive Cratering and Ejecta*, Nuclear Geoplosics Sourcebook, v. 55, pt. 2, sec. 4, Rep. DNA OIH-4-2 (Def. Nucl. Agency, Bethesda, MD).
- SIEH, K. E. (1978), *Prehistoric Large Earthquakes Produced by Slip on the San Andreas Fault at Pallet Creek, CA*, J. Geophys. Res. 83, 3907-3939.
- SMALLEY, R. F., CHATELAIN, J. L., TURCOTTE, D. L., and PREVOT, R. (1987), *A Fractal Approach to the Clustering of Earthquakes: Applications to the Seismicity of the New Hebrides*, Seis. Soc. Am. Bull., 77, 1368-1381.
- SMITH, S. W. (1976), *Determination of Maximum Earthquake Magnitude*, Geophys. Res. Lett. 3, 351-354.
- TURCOTTE, D. L. (1986a), *Fractals and Fragmentation*, J. Geophys. Res. 91, 1921-1926.
- TURCOTTE, D. L. (1986b), *A Fractal Approach to the Relationship between Ore Grade and Tonnage*, Econ. Geol. 81, 1528-1532.
- TURCOTTE, D. L. (1986c), *A Fractal Model for Crustal Deformation*, Tectonophys. 132, 261-269.
- VAN DAMME, H., OBRECHT, F., LEVITZ, P., GATINEAU, L., and LAROCHE, C. (1986), *Fractal Viscous Fingering in Clay Slurries*, Nature 320, 731-733.
- WARREN, B. A. (1973), *Transpacific Hydrographic Sections at Lats. 43°S and 28°S: The SCORPIO Expeditions II. Deep Water*, Deep Sea Res. 20, 9-38.

(Received April 6, 1987, revised/accepted August 10, 1987)

Fractal Reconstruction of Sea-Floor Topography

JEAN-CLAUDE MARESCHAL¹

Abstract—Sea-floor bathymetric profiles exhibit features at many different scales of length; this suggests that they could be described as fractals. An algorithm interpolating a fractal line between points has been used to reconstruct bathymetric profiles from a few data points. In general, this fractal line has the same Fourier amplitude spectrum as real bathymetry, and, if the parameters of the interpolation are suitably chosen, it has a very similar appearance. The success of this fractal reconstruction algorithm for the sea-floor raises the possibility that it could be used to extrapolate, from data collected at one scale, the properties of the sea-floor at finer scales, and that similar techniques could be used to interpolate a surface between bathymetric profiles. The fractal character is a sign that the processes that shape the sea-floor are scale invariant and suggests that the renormalization group technique could be used to model these processes.

Key words: Fractal, bathymetry, interpolation, fracture zone, ridge, tectonics.

1. Introduction

At the largest scale, the bathymetry of the sea-floor is well described by deterministic models. The average bathymetry of the sea-floor can be explained by the cooling of the oceanic lithosphere as it moves away from the midoceanic ridges; the sea-floor depth increases in proportion to the square root of the age of the lithosphere or of the distance to the midoceanic ridge, if the spreading rate has remained constant (PARKER and OLDENBURG, 1973; OLDENBURG, 1975; DAVIS and LISTER, 1974). The average bathymetry data fit well the lithospheric cooling model, at least for ages less than 80 Ma (PARSONS and SCLATER, 1977). For ages greater than 80 Ma, the bathymetry flattens with respect to a square root of age reference curve. This flattening can be explained by the influx of heat at the base of the cooling plate, but there is still some debate on the implications of this flattening for mantle convection (RICHTER and PARSONS, 1975; JARVIS and PELTIER, 1980). Elongated swells often associated with volcanic islands are observed on the sea-floor and represent another type of departure from the depth age relationship. These

¹ Dept. des Sciences de la Terre and GEOTOP, Université du Québec à Montréal, P.O. 8888, Sta. A, Montréal, H3C 3P8, Canada.

swells are explainable by the thinning and rejuvenation of the lithosphere when it moves over a hot spot (CROUGH, 1978). On the ocean side of deep sea trenches, the long wavelength changes in depth to the sea bottom are explained by the flexure of the subducted lithosphere. At the largest scale, the bathymetry of the fracture zone is explained by a model of thermal and mechanical evolution (SANDWELL and SCHUBERT, 1982; SANDWELL, 1984).

Over the years, new techniques (Sea Beam, Sea Marc I and II, Seasat) have been used to map the sea-floor with an ever increasing resolving power (DETRICK, 1986). Figure 1 shows the bathymetry on different profiles across the Cocos spreading center in the East Pacific (KLITGORD and MAMMERICKX, 1982). Figure 2 shows the bathymetry on profiles across the Mendocino and Pioneer fracture zones of California (SANDWELL and SCHUBERT, 1982). These profiles extend over several hundred km, and they are long in comparison with the resolving power better than 20 cm achieved by the sea-floor mapping techniques. The figures, even at the largest scale, demonstrate clearly that the sea-floor bathymetry exhibits a structure far more complex than what the thermal and mechanical evolutionary models account for, and far more complicated than what such models could reasonably be expected to explain. The most obvious character of these profiles is that they contain features at many different scales, in addition to the largest tectonic scale governed by plate models. Increased resolving power would simply add more and more unaccountable details.

Such a hierarchy of scales appears in many different natural shapes where many finer features are superposed on larger ones. The concept of fractal was introduced by MANDELBROT (see 1983, and references therein), to provide a simple unifying description of those sets which contain structures with a complexity that increases as the scale becomes finer. Such sets have the property that, since new structures appear every time the scale changes, their measure depends on the scale considered (*i.e.*, on the length of the measuring rod). In order to quantify the fractal character of these sets, Mandelbrot has proposed to use the Hausdorff-Besicovitch dimension, H , sometimes referred to as the fractal dimension. This dimension is defined by the following relationship between $N(\varepsilon)$, the number of balls needed to cover the set and their radius ε

$$H = \lim_{\varepsilon \rightarrow 0} \left\{ \frac{\log N(\varepsilon)}{\log (1/\varepsilon)} \right\}.$$

Therefore, the length of a fractal line is undefined in the classical sense as it varies with the size of the measuring rod as follows:

$$L(\varepsilon) = N(\varepsilon) \cdot \varepsilon \propto \varepsilon^{1-H}.$$

In general, the Hausdorff-Besicovitch dimension is not, like the topological dimension, an integer; but, if it is an integer, it coincides with the topological dimension.

PACIFIC-RIVERA SPREADING CENTER BATHYMETRY

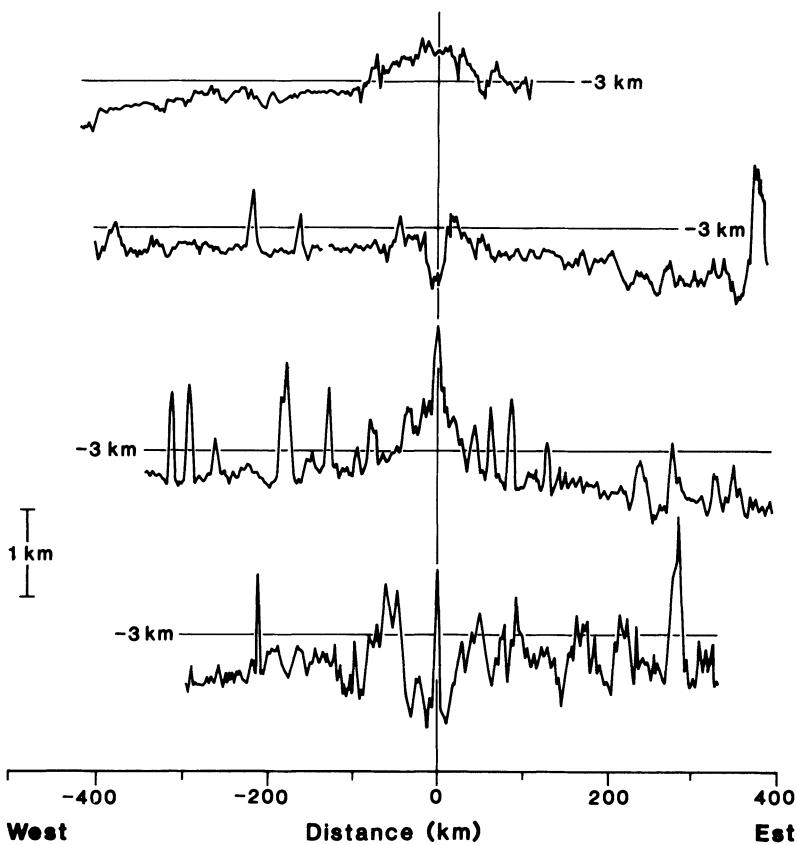


Figure 1

Bathymetry profiles near the Pacific Rivera spreading center (adapted from KLITGORD and MAMMERICKX, 1982).

Sea-floor bathymetric profiles from the North Atlantic ocean were analyzed by BARENBLATT *et al.* (1984) who determined their Hausdorff-Besicovitch dimension. They divided the profiles studied into segments, according to the distance to the ridge; they found that H varies from 1.46 near the ridge crest to 1.05 at distances greater than 2000 km away from the ridge.

Fourier analysis provides an alternate approach to relate the number of topographic features of a given scale to their wavelength. The spectral estimates of sea-floor topography obtained by BELL (1975) show that, in the range of wavenumbers between 10^{-5} km^{-1} and 10 km^{-1} , the spectral amplitude is inversely proportional to the wavenumber. In a more recent review, FOX and HAYES (1985) show that this relationship between amplitude spectrum and wavenumber is not perfectly

MENDOCINO-PIONEER FRACTURE ZONE

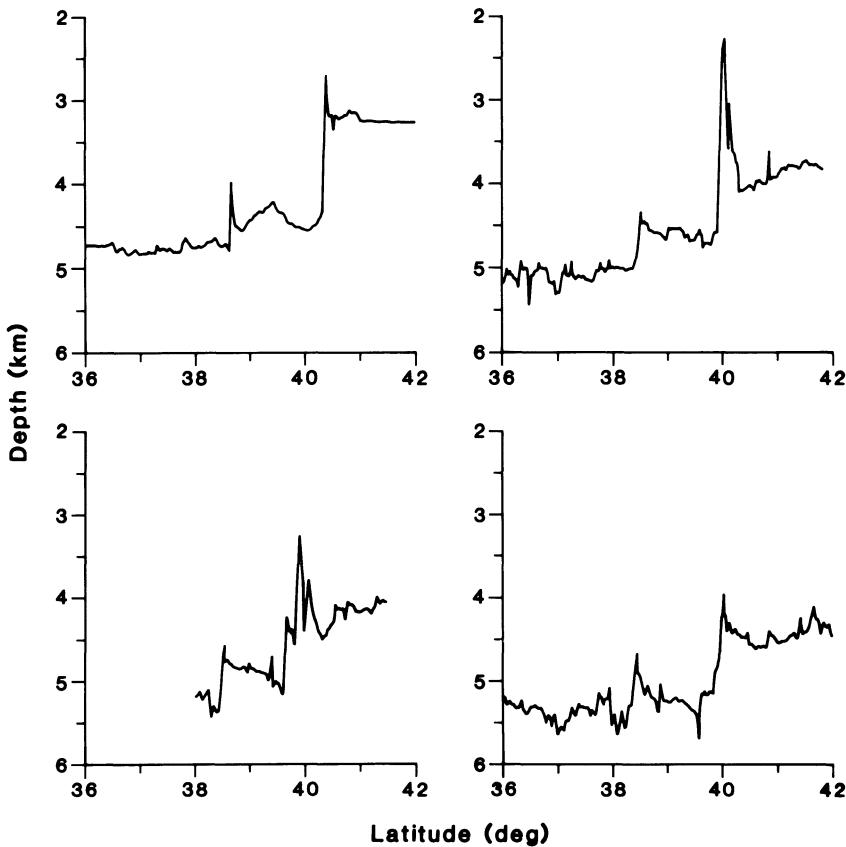


Figure 2

Bathymetry along several profiles across the fracture zones off the coast of California (adapted from SANDWELL and SCHUBERT, 1982).

isotropic and that it varies between morphological provinces. They suggest that the Fourier amplitude spectrum contains all the useful information about sea-floor morphology and its content could perhaps indicate the processes by which the sea-floor was shaped.

The purpose of this paper is to demonstrate the fractal character of the sea-floor by constructing synthetic sea-floor topography with a fractal interpolation algorithm introduced by BARNESLEY (1987). The algorithm will be summarized and examples of synthetic profiles compared to real profiles. The synthetic profiles have the same features and spectral content as real profiles. Some possible applications and the significance of this fractal character will be sketched.

2. Construction of Bathymetry Profiles

An algorithm interpolating a fractal line between $N + 1$ points (BARNESLEY, 1987) was used to construct, from a few data points, bathymetric profiles that look like sea-floor. The interpolation procedure is based on the application of iterated function systems (I.F.S.) (BARNESLEY and DEMKO, 1985). Given a compact metric space K with distance function $d(x,y)$, the I.F.S. is a set of continuous functions $\{w_n, n = 1, \dots, N\}$ mapping K on itself.

The image of any subset $A \subset K$ is defined as:

$$w(A) = \bigcup_{n=1}^N w_n(A) \quad w_n(A) = \{w_n(P), P \in A\}.$$

Any subset, $G \subset K$, such that:

$$G = w(G)$$

is called an attractor of the I.F.S. If w is a contractive mapping, i.e., if for some $s (0 < s < 1)$

$$d(w_n(x), w_n(y)) < sd(x, y) \quad \forall n; \forall x, y \in K$$

then the I.F.S. is called hyperbolic and the attractor is unique.

Given a set of data points $\{(x_i, y_i); i = 0, 1, \dots, N\}$ where $x_i \in [x_0, x_N]$ and $y_i \in \mathfrak{R}$, the interpolation functions are such that $y_i = f(x_i)$, $i = 0, 1, \dots, N$.

Interpolation functions for the set of data points can be constructed that are such that their graph $G = \{x, f(x)\}$ is the attractor of an I.F.S.; such functions can be referred to as fractal interpolation functions because they may be everywhere continuous and nowhere differentiable and their Hausdorff-Besicovitch dimension may be noninteger. The functional form of the maps w_n is arbitrary. One of the simplest I.F.S. is the linear fractal interpolation system which can be constructed with the following set of affine maps:

$$w_n(x, y) = \begin{cases} L_n(x) = x_{n-1} + \frac{(x_n - x_{n-1})}{(x_N - x_0)}(x - x_0) \\ F_n(x, y) = a_n x + b_n \end{cases}$$

where the scaling factors are such that $|s_n| < 1$ and the a_n and b_n are chosen such that:

$$\begin{cases} F_n(x_0, y_0) = y_{n-1} \\ F_N(x_N, y_N) = y_n \end{cases}$$

i.e.,

$$a_n = \frac{(y_{n-1} - y_n)}{(x_0 - x_N)} - s_n \frac{(y_0 - y_N)}{(x_0 - x_N)}$$

$$b_n = \frac{(y_{n-1}x_N - y_nx_0)}{(x_N - x_0)} - s_n \frac{(y_0x_N - y_Nx_0)}{(x_N - x_0)}.$$

These functions map the parallelogram P :

$$x_0 \leq x \leq x_N$$

$$y_0 + \frac{(y_N - y_0)}{(x_N - x_0)}(x - x_0) \leq y \leq y_0 + \frac{(y_N - y_0)}{(x_N - x_0)}(x - x_0) + 1$$

into the parallelogram P_n

$$x_{n-1} \leq x \leq x_n$$

$$y_{n-1} + \frac{(y_n - y_{n-1})}{(x_n - x_{n-1})}(x - x_{n-1}) \leq y \leq y_{n-1} + \frac{(y_n - y_{n-1})}{(x_n - x_{n-1})}(x - x_{n-1}) + s_n$$

if $s_n > 0$.

Given a set of $N + 1$ points $\{P_i(x_i, y_i), i = 0, 1, \dots, N\}$, the set of linear maps is completely defined by the arbitrary scaling factors. The values affected to the scaling factors s_n are related to the fractal dimension of the graph of the attractor for the I.F.S.

It is possible to associate a Fourier spectrum with this line. Let $F(k)$ by the Fourier spectrum of the function interpolating linearly between $N + 1$ points $\{(x_0, y_0), \dots, (x_N, y_N)\}$.

The Fourier spectrum of the function obtained by interpolating linearly between $N + 1$ points obtained by applying the map w_l to the original $N + 1$ points: $w_l(x_n, y_n)$ is given by:

$$M_l \odot F(k) = d_l \cdot s_l \cdot \exp\{ikx_l\} \cdot F(d_l \cdot k)$$

where

$$d_l = \frac{(x_l - x_{l-1})}{(x_N - x_0)}$$

is the horizontal scaling and s_l is the vertical scaling. The function obtained by successive application of the maps $w_{l_1}, w_{l_2}, \dots, w_{l_m}$ to the original points, will thus have the following spectrum:

$$M_{l_1} \odot M_{l_2} \odot \dots \odot M_{l_m} \odot F(k).$$

The spectrum of the line joining the points obtained by applying M times the iterated function system to the $N + 1$ points will thus contain the following terms:

$$\sum_{m=1}^N \left\{ \sum_{l_1=1}^N \dots \sum_{l_m=1}^N s_{l_1} \dots s_{l_m} d_{l_1} \dots d_{l_m} \cdot \exp\{ikx_{l_1 \dots l_m}\} \cdot F(d_{l_1} \dots d_{l_m} k) \right\}$$

where $x_{l_1 \dots l_m} = L_{l_m} \odot \dots \odot L_{l_2}(x_{l_1})$.

Fortran programs were written to implement such algorithms on an IBM PC. The programs were used to calculate and plot a synthetic bathymetric profile obtained by interpolating between data points from a line across the Mathematician Ridge, in the Eastern Pacific. The synthetic and the original profiles are compared on Figure 3. The data points were carefully selected and are not evenly distributed. Some experimenting with the scaling factors was necessary before the two profiles looked very roughly similar. It was found that the outcome of the interpolation is only moderately sensitive to the choice of the data points, provided that the gross features of the real profile are present. On the other hand, the outcome is extremely sensitive to the choice of scaling factors (*i.e.*, to the fractal dimension of the interpolant); larger values of the scaling factors would result in much larger features at the finer scales. The cross-correlation could be used to quantify the similarity

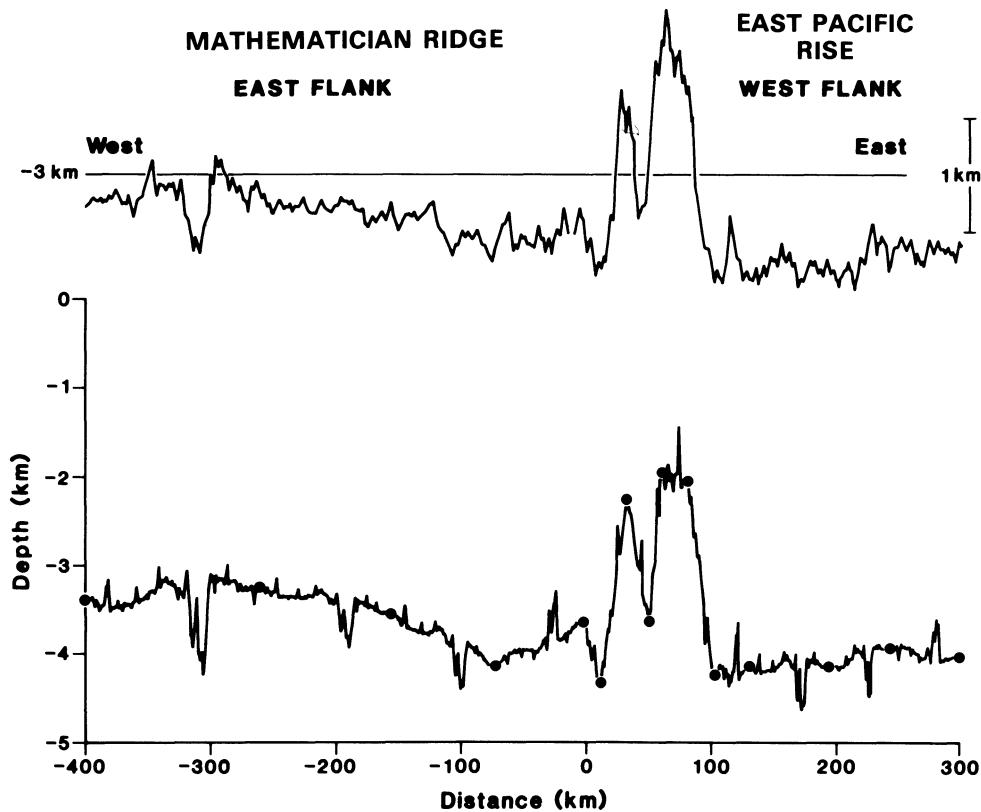


Figure 3

Comparison between a real bathymetry profile across the extinct Mathematician Ridge and a synthetic profile interpolated between 15 points marked as dots. The scaling factors for the interpolation algorithm were selected between -.5 and .5. Similar profiles can be obtained with different data points and with a different number of data points, provided that the selected points reproduce the gross feature of the data.

between real and synthetic profiles, but it is dominated by the choice of the data points which implies that the cross-correlation between the two profiles is *a priori* almost equal to their autocorrelation functions. The statistical similarity between the two profiles can be better demonstrated by the comparison of their Fourier amplitude spectra shown on Figure 4. On a bilogarithmic scale, both spectra show random fluctuations superposed to a linear trend with slope -1 (*i.e.*, the amplitude spectrum decreases as $1/k$ plus some random variation). Only the slopes of the spectra can be compared because the amplitude and wavenumbers have been normalized to the values corresponding to the lowest wavenumber in the spectrum. The data used for this study are not sampled closely enough to determine the behavior of the real data spectrum for higher wavenumbers but the studies of BELL (1975) have shown that, over many scales of wavenumbers, the spectrum of sea-floor topography decreases as $1/k$.

The second example examined is that of a sea-floor topographic profile near the Mendocino fracture zone. Figure 5 shows a real profile and a profile constructed by fractal interpolation between the data points. With the appropriate choice of the data points and of scaling factors, the two profiles can be made to look similar. The Fourier amplitude spectra are shown on Figure 6; they have the same behavior as the spectra examined on Figure 4.

BERRY and LEWIS (1980) have studied the Weierstrass-Mandelbrot function (whose spectrum spans the whole wavenumber range in geometric rather than arithmetic progression):

$$W(t) = \sum_{n=-\infty}^{n=\infty} \frac{1 - \exp\{i\gamma^n t\}}{\gamma^{(2-D)n}} \exp\{i\phi_n\}$$

where the phases ϕ_n are random, and the frequencies γ^n form a Weierstrass spectrum.

They have shown that the power spectrum of the Weierstrass-Mandelbrot function can be approximated by a continuous spectrum $s(k)$:

$$s(k) \propto \frac{1}{k^{5-2D}}$$

Also, they conjectured (see also MANDELBROT, 1983) that for the Weierstrass function, D coincides with the fractal dimension.

To the author's knowledge, it has not been demonstrated that these conjectures can be extended and applied to the spectra of functions other than the Weierstrass function, but such generalizations are often invoked to estimate the fractal dimension from a band averaged power spectrum. The following relationship yields the fractal dimension, D , from the logarithmic slope of the power spectrum, s :

$$D = (5 - s)/2.$$

If this relationship would be applied to the power spectra of the synthetic or real

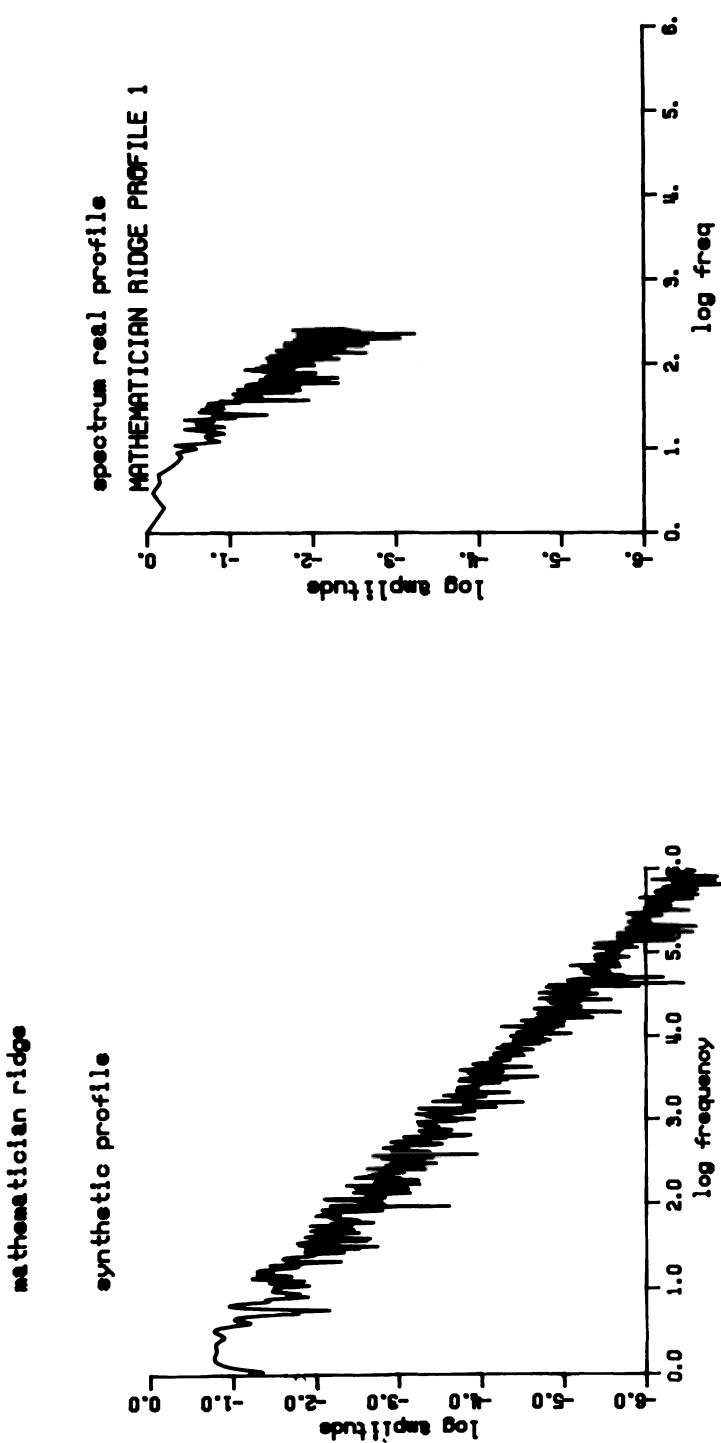


Figure 4
Fourier amplitude spectra of the real and of the synthetic profiles shown on Figure 3; the scale is bilogarithmic. The values are relative: the wavenumbers are normalized to the lowest wavenumber in the spectrum, the amplitude to the amplitude at the lowest wavenumber. Only the slopes can be compared. On log-log scale, both spectra have a slope close to -1 .

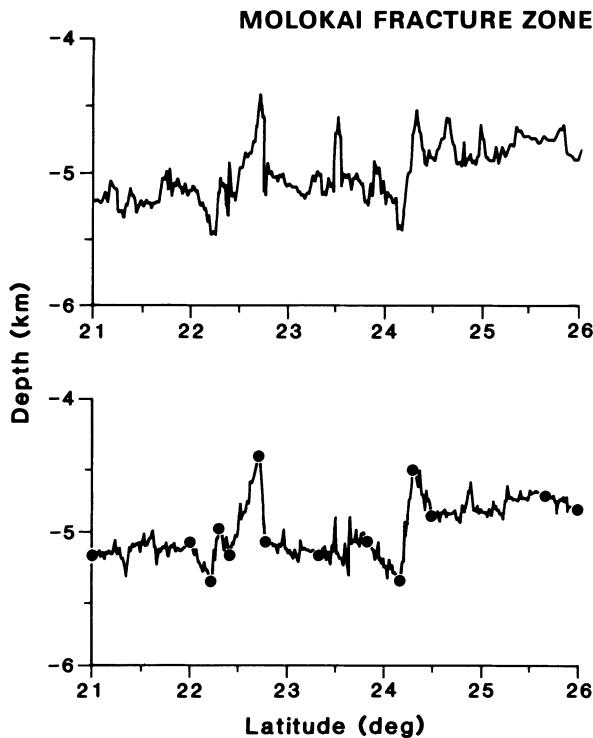


Figure 5

Comparison between a synthetic profile and a real profile across the Molokai fracture zone, off California. The synthetic profile was obtained by interpolating between the 14 points marked with as plain circles with scaling factors varying between -.4 and .4.

sea-floor bathymetric profiles (logarithmic slope 2), it would yield a fractal dimension of 1.5. The profiles examined above are close to the active ridges and in tectonically deformed regions; BARENBLATT *et al.* (1984) computed that, near the mid-Atlantic ridge, the fractal dimension of the sea-floor is 1.46.

3. Discussion and Conclusions

The examples presented in this paper, far from exhausting the question, simply suggest that it is possible to represent the sea-floor by a fractal. Although similar conclusions were reached by other studies, the fractal description is not the only one possible and other stochastic models could provide as adequate interpolants (see for instance the study of the ARIMA (1,1,0) by MALINVERNO (1989)). Also, the range of scales investigated in this study is restricted; one should not expect that the interpolation technique, which implies a high degree of self-similarity, will

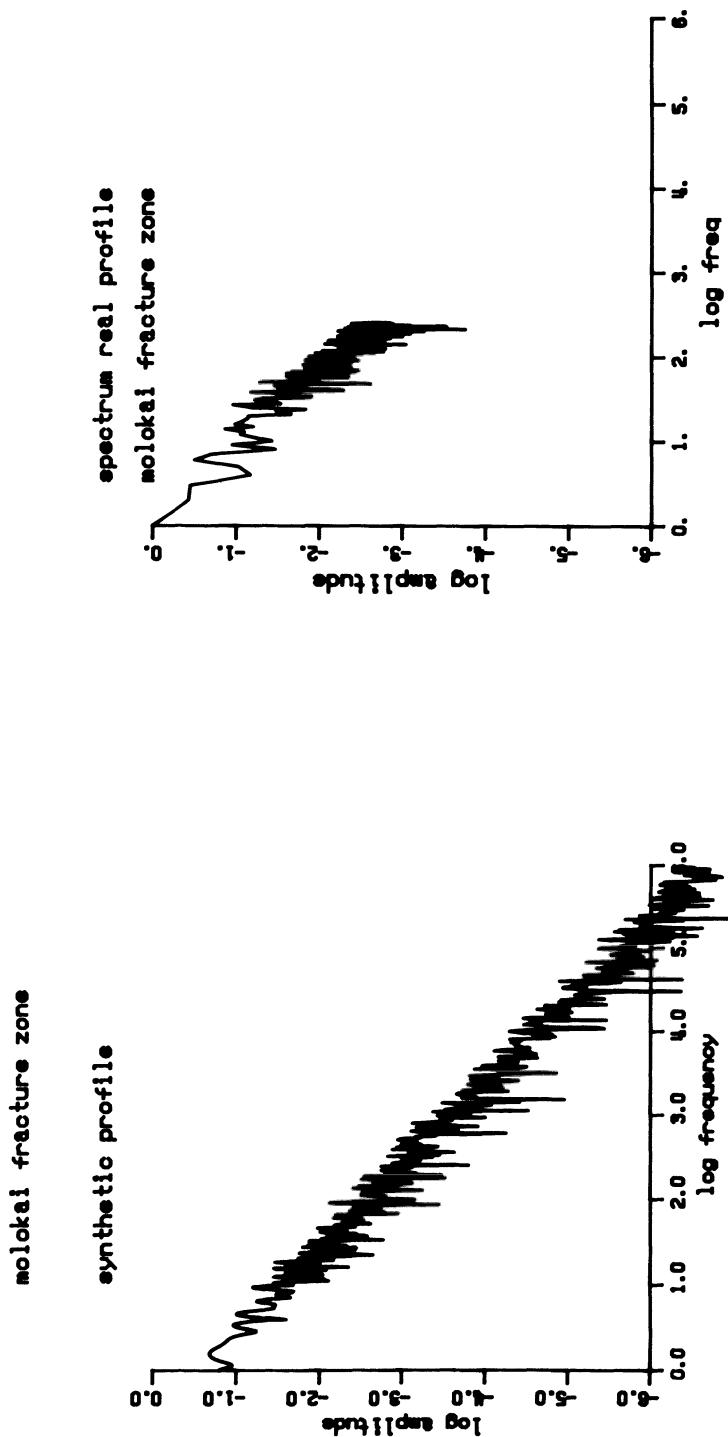


Figure 6
Fourier amplitude spectra of the real and synthetic profiles for the Molokai fracture zone shown in Figure 5. The conventions are the same as in Figure 3.

reproduce exactly the details of any profile at finer and finer scales, but only that it will carry them in a statistical sense (*i.e.*, only the band averaged power spectra of the real and synthetic profiles will be identical, both the amplitude and phase spectra will be different). Only profiles have been examined for this short study; the foremost important question of interpolating surfaces between profiles was left out because the generalization of the fractal interpolation algorithm to surfaces is not straightforward. A different procedure, following the approach of BARNESLEY and DEMKO (1985) was used and results were reported by MARESCHAL *et al.* (1985): the sea-floor topography was represented as the density function associated with a random walk experiment on the unit square; additional investigations are underway. The results presented above are thus still incomplete but, nevertheless, intriguing; they suggest that this type of reconstruction would permit extrapolation of the properties of the sea-floor from one scale to the next. If it is indeed the case, it would then be possible, from discrete data sets and with some additional information, to compute approximations and to determine the entire power spectrum of the sea-floor surface from its large scale properties. This may have important practical applications such as determining the reflexivity of the ocean floor (*e.g.*, Tolstoy, 1973).

Another question that must be asked is whether and how this fractal or scale invariant character of the sea-floor is related to the physical processes that have shaped the sea-floor surface. At this stage, the answer to this question can only be speculative and incomplete. The compilation of data on the size and frequency of transform faults (ABBOTT, 1986) shows that, in the range where they can be mapped and observed, transform faults follow a log-log distribution. Because the distribution of the transform faults determines the morphology of fracture zones, the fractal character of the sea-floor can be related to the distribution of faults. KING (1983) assumed that multiscale faulting is a natural consequence of geometrical self-similarity and that the distribution of faults is determined by spatial self-similarity. MANDELBROT (1983) proposed that fractal behavior is the consequence of underlying physical phenomena and that this could be demonstrated by the application of the renormalization group method. ALLEGRE *et al.* (1982) and TURCOTTE (1986) examined the processes of fragmentation and faulting and they suggested that fractal behavior can be derived by the application of the renormalization group method (WILSON and KOGUT, 1974). The renormalization group theory was introduced by SMALLEY *et al.* (1985) to explain the stick-slip behavior of faults. They showed that scale invariance becomes a natural consequence of the physical model. The morphology of the sea-floor in active regions is essentially determined by processes, such as faulting and volcanism; the introduction of a scale invariance requirement in the description of these phenomena could shed some light on the processes shaping the sea-floor and perhaps explain the fractal character of its topography.

Acknowledgements

The author is extremely grateful to the mathematicians who introduced him to the world of fractals, in particular to Prof. Daniel Bessis, to Prof. Michael Barnsley and his graduate student Douglas Hardin who patiently taught him the little he was able to learn. Michael Barnsley also provided the compiled Basic program which was used in the early stages of this research. Fortran programs for PC are available upon request from the author.

This work was partly supported by an NSERC operating grant.

REFERENCES

- ABBOTT, D. (1986), *A Statistical Correlation Between Ridge Crest Offsets and Spreading Rate*, Geophys. Res. Lett. 13, 157–160.
- ALLEGRE, C. J., LE MOUEL, J. L., and PROVOST, A. (1982), *Scaling Rules in Rock Fracture and Possible Implications for Earthquake Prediction*, Nature 297, 47–49.
- BARENBLATT, G. I., ZHIVAGO, A. V., NEPROCHNOV, Yu. P., and OSTROVSKYI, A. A. (1984), *The Fractal Dimension: A Quantitative Characteristic of Ocean Bottom Relief*, Oceanology 24, 695–697.
- BARNESLEY, M. F. (1987), *Fractal Functions and Interpolation*, Constr. Approx. 2, 303–329.
- BARNESLEY, M. F., and DEMKO, S. (1985), *Iterated Function Systems and the Global Construction of Fractals*, Proc. Roy. Soc. London Ser. A 399, 243–275.
- BELL, T. H. (1975), *Statistical Features of Sea-Floor Topography*, Deep Sea Res. 22, 883–892.
- BERRY, M. V., and LEWIS, Z. V. (1980), *On the Weierstrass-Mandelbrot Fractal Function*, Proc. Roy. Soc. London. Ser. A 370, 459–484.
- CROUGH, S. T. (1987), *Thermal Origin of Mid-Plate Hot Spot Swells*, Geophys. J. Roy. Astr. Soc. 55, 451–469.
- DAVIS, E. E., and LISTER, C. R. B. (1974), *Fundamentals of Ridge Crest Topography*, Earth Planet. Sci. Lett. 21, 405–413.
- DETTRICK, R. S. (1986), *Introduction to the Sea Floor Mapping Section*, J. Geophys. Res. 91, 3331–3333.
- FOX, C. G., and HAYES, D. E. (1985), *Quantitative Methods for Analyzing the Roughness of the Sea-Floor*, Rev. Geophys. Space Phys. 23, 1–48.
- JARVIS, G. T., and PELTIER, W. R. (1980), *Oceanic Bathymetry Profiles Flattened by Radiogenic Heating in a Conductive Mantle*, Nature 285, 649–651.
- KING, G. (1983), *The Accommodation of Large Strains in the Upper Lithosphere of the Earth and Other Solids by Self-Similar Fault Systems: The Geometrical Origin of b-Value*, Pure Appl. Geophys. 121, 761–815.
- KLITGORD, K. D., and MAMMERICKX, J. (1982), *Northern East Pacific Rise: Magnetic Anomaly and Bathymetric Framework*, J. Geophys. Res. 87, 6725–6750.
- MALINVERNO, A. (1989), *A Test of the Fractal Character of Sea-Floor Topography*, Pure Appl. Geophys. 131, 139–155.
- MANDELBROT, B. B., *The Fractal Geometry of Nature*, 3rd edition (W. H. Freeman and Co., New York 1983).
- MARESCHAL, J.-C., BARNESLEY, M., and HARDIN, D. (1985), *Fractal Reconstruction of Sea-Floor Topography*, EOS 66, 355.
- OLDENBURG, D. W. (1975), *A Physical Model for the Creation of the Lithosphere*, Geophys. J. Roy. Astr. Soc. 43, 425–451.
- PARKER, R., and OLDENBURG, D. D. (1973), *Thermal Model of Ocean Ridges*, Nature Phys. Sci. 242, 137–139.
- PARSONS, B., and SCLATER, J. G. (1977), *An Analysis of the Variation of Ocean-Floor Bathymetry and Heat Flow with Age*, J. Geophys. Res. 82, 803–827.

- RICHTER, F. M., and PARSONS, B. (1975), *On the Interaction of two Scales of Convection in the Mantle*, J. Geophys. Res. 80, 2529–2541.
- SANDWELL, D. T. (1984), *Thermomechanical Evolution of Oceanic Fracture Zones*, J. Geophys. Res. 89, 11401–11413.
- SANDWELL, D. T., and SCHUBERT, G. (1982), *Lithosphere Flexure at Fracture Zones*, J. Geophys. Res. 87, 4657–4667.
- SMALLEY, R. F. Jr., TURCOTTE, D. L., and SOLLA, S. A. (1985), *A Renormalization Group Approach to the Stick Slip Behavior of Faults*, J. Geophys. Res. 90, 1894–1900.
- TURCOTTE, D. L. (1986), *Fractals and Fragmentation*, J. Geophys. Res. 91, 1921–1926.
- TOLSTOY, I., *Wave Propagation* (McGraw Hill, New York 1973).
- WILSON, K. G., and KOGUT, J. (1974), *The Renormalization Group and the ε Expansion*, Phys. Rev. C. 12, 75–200.

(Received May 6, 1987, revised/accepted October 26, 1987)

Empirically Derived Relationships between Fractal Dimension and Power Law Form Frequency Spectra

CHRISTOPHER G. FOX¹

Abstract—Fractal analysis and Fourier analysis are independent techniques for quantitatively describing the variability of natural figures. Both methods have been applied to a variety of natural phenomena. Previous analytical work has formulated relationships between the fractal dimension and power law form frequency spectrum. MANDELBROT (1985) has shown that difficulties arise when the ruler method for measuring dimensionality is applied to other than self-similar figures. Since an investigator presumably does not know in advance the dimensionality of a natural profile, it is essential to quantify the nature of the discrepancy for self-affine cases. In this study, a series of experiments are conducted in which discrete random series of specified spectral forms are analyzed using the fractal ruler method. The various parameters of the fractal measurement are related to the parameters of the spectral model. In this way, empirical relationships between the techniques can be derived for discrete, finite series which simulate the results of applying the fractal method to observational data.

The results of the study indicate that there are considerable discrepancies between the results predicted by theory and those derived empirically. The fundamental power law form of length versus resolution pairs does not hold over the entire region of analysis. The predicted linear relationship between fractal dimension and exponent of the frequency spectrum does not hold, and the spectral signals can be extended beyond the limits of dimension inferred by theory. Root-mean-square variability is also shown to be linearly related to the fractal intercept term. An investigation of the effect of nonstationary sampling is conducted by generating signals composed of segments of differing spectral characteristics. Fractal analyses of these signals appear identical to those conducted on stationary series.

The discrepancies between theoretical prediction and empirical results described in this study reflect the difficulties of applying analytically derived techniques to measurement data. Both Fourier and fractal techniques are formulated through rigorous mathematics, assuming various conditions for the underlying signal. When these techniques are applied to discrete, finite length, nonstationary series, certain statistical transformations must be applied to the data. Methods such as windowing, prewhitening, and anti-aliasing filters have been developed over many years for use with Fourier analysis. At present, no such statistical theory exists for use with fractal analysis. It is apparent from the results of this study that such a statistical foundation is required before the fractal ruler method can be routinely applied to observational data.

Key words: Fourier analysis, fractal, frequency spectra, roughness, stationarity, numerical modeling.

¹ National Oceanic and Atmospheric Administration, Pacific Marine Environmental Laboratory, Hatfield Marine Science Center, Newport, Oregon 97365, U.S.A.

Introduction

Much interest has centered recently on methods for quantitatively describing the variability of surfaces found in nature. The concept of fractal dimension, formulated by MANDELBROT (1982), has been widely used to describe the roughness of surfaces, including deep-sea acoustic basement (BARENBLATT *et al.*, 1984), petrologic fracture planes (BROWN and SCHOLTZ, 1985), and the San Andreas Fault (AVILES *et al.*, 1987). An alternative method for describing surface roughness is through the use of frequency spectra derived by Fourier transform. BELL (1975), and later FOX and HAYES (1985), developed and used spectral techniques to quantify the roughness of the seafloor.

The underlying theories of fractal analysis and spectral analysis are rooted in different branches of mathematics. Earlier analytical work derived a functional relationship between the exponent of power law form spectra and fractal dimension (BERRY and LEWIS, 1980). The objective of this study is to investigate the correspondence between the two measurement techniques through empirically derived relationships gathered from discrete, finite time series. For simplicity, the discussion will be limited to one-dimensional spectra and therefore fractal dimensions between $D = 1.0$ and $D = 2.0$.

The approach taken in this study is straightforward. Signals of a specified spectral form are artificially generated via a computer algorithm described by FOX (1987). The algorithm allows the user to rigorously define the nature of the signal in terms of spectral form, statistical distribution, and statistical stationarity. This random signal is then passed through an additional algorithm which measures the fractal dimension D using a length versus resolution (ruler) method. The precision of the measurement is inferred using simple performance testing techniques. A series of numerical experiments are then performed using a range of input spectra, and the results interpreted.

Theoretical Background

The Weierstrass-Mandelbrot fractal function is both continuous and scaleless, but is nowhere differentiable (MANDELBROT, 1982). The Fourier transform assumes an infinite function within which the discontinuities are finite (BRACEWELL, 1978). Discontinuities are modeled by the addition of an infinite series of sinusoids of increasing frequency which approach the form of the discontinuity (Gibb's phenomenon). The functional relationship between the Weierstrass-Mandelbrot fractal function and the Fourier power spectrum was derived by BERRY and LEWIS (1980) as,

$$P \cong f^{-(5-2D)} \quad (1)$$

where P = power, f = frequency, and D = fractal (Hausdorff-Besicovitch) dimension. The relationship is approximate, becoming exact only under certain circumstances. The function is of power law form, which agrees with the results of many workers generating frequency spectra of natural surfaces (BELL, 1975; SAYLES and THOMAS, 1978; FOX and HAYES, 1985). Notice that the relationship in equation (1) has no proportionality term, reflecting the scaleless property of the fractal dimension.

For a linear figure, the range of values for D is bound by certain conditions (BERRY and LEWIS, 1980; MANDELBROT, 1982; SCHOLTZ and AVILES, 1986). For $D = 2.0$, the linear figure becomes plane filling, that is, two-dimensional. For $D = 1.0$, the figure becomes differentiable and describes the Euclidean notion of one-dimensional geometry. A fundamental concept of fractal mathematics is the property of self-similarity. A self-similar figure appears similar at all scales of observation. This property can be alternatively viewed in terms of Fourier spectra.

The power law form of the frequency spectrum can be represented as,

$$A = a \cdot f^b \quad (2)$$

where A = amplitude, f = frequency, and a and b are constants. In all further discussion, the amplitude spectrum, or square root of the power spectrum, will be used to enhance the visualization of the analysis. Since the amplitude spectrum can be normalized to simple units of length versus wavelength, it is compatible with the length versus length units of the fractal ruler method. Equation (1) can be rewritten in terms of amplitude as,

$$A \cong f^{-(5/2 - D)}. \quad (3)$$

The relationship between the exponent of the amplitude spectrum and the fractal dimension becomes,

$$b \cong -(5/2 - D) \quad (4)$$

or equivalently,

$$D \cong b + 5/2. \quad (5)$$

A special case of the power law form amplitude spectrum occurs when $b = -1$ or,

$$A = a \cdot f^{-1}. \quad (6)$$

As described by FOX and HAYES (1985, p. 19), this special case of the power law form amplitude spectrum implies that the ratio of amplitude to wavelength (aspect ratio) of all component sinusoids is constant over all frequencies, and the magnitude of that ratio depends upon the scaling factor a . Mixing fractal and spectral analysis terminology, the component sinusoids are self-similar (their shape is retained at all scales) in the case of $b = -1.0$. At all other values, the aspect ratio of the

component sinusoids changes as a function of frequency, analogous to the fractal concept of self-affinity. For values of $b > -1.0$, the aspect ratio increases with increasing frequency. That is, the profile appears rougher at finer scales. For $b < -1.0$, the aspect ratio decreases with increasing frequency, causing features to appear relatively smoother at smaller scales. For $b = -1.0$, equation (5) predicts that this "self-similar" case occurs at $D = 1.5$, which fractal mathematics describes as self-affine. BROWN (1987) addresses this apparent paradox between the two methods.

The limitation of linear fractal figures to $1.0 < D < 2.0$ implies a similar limitation in the exponent of the one-dimensional amplitude spectrum of $-1.5 < b < -0.5$. Obviously, no such limitation exists since spectra of any exponent can be used to generate a signal. For example, white noise with a constant amplitude spectrum ($b = 0.0$) implies a fractal dimension of $D = 2.5$, beyond the range of plane filling. The analytical relationship in equation (1) does not hold beyond the stated boundaries, however such signals can be subjected to fractal analysis, as well as those beyond the lower limit ($b < -1.5$). Since an investigator presumably does not know the dimensionality of a figure in advance, these nonfractal cases will be included in the following numerical experiments.

Both fractal and Fourier techniques have certain practical advantages in measuring natural surfaces. The Fourier transform requires a true mathematical function as input; that is, only one range value can correspond to a given value in the functional domain. No such restriction applies in fractal analysis, allowing the study of multivalued functions such as coastlines (MANDELBROT, 1967). In practice, the Fourier transform algorithms require data which are evenly spaced in the independent variable. In reducing the description of a figure to a single measure of dimension, as is generally done in fractal analysis, a great deal of information is lost. One of the advantages of Fourier spectral technique is in the ability to discriminate the power and phase of a signal with respect to frequency. FOX and HAYES (1985) found the scaling of the frequency spectra of seafloor topography to be of primary geological significance, while knowledge of the spectral exponent was required for extrapolation to other frequencies. In other applications, the phase content of a signal may be of primary importance (OPPENHEIM and LIN, 1981). The investigator must choose the technique which is most appropriate to the measurement problem.

In using either Fourier or fractal analysis as measurement techniques, one must take the step from well-behaved analytical boundary conditions to inferences concerning observational data. In the case of Fourier analysis, a continuous, infinite length, stationary time series is assumed in the underlying theory. In practice, the data series under analysis is often discrete, always finite, and may be highly nonstationary. Methods for minimizing the errors due to these effects have been developed through the use of various statistical techniques to reduce aliasing and spectral leakage (BLOOMFIELD, 1976), and additional techniques for minimizing

nonstationary effects (FOX and HAYES, 1985). Similar statistical theory is not yet available for workers measuring fractal dimensions of observational data. In this study, the relationships between the fractal and Fourier techniques are investigated through controlled experiments with synthetic signals which simulate profiles observed in nature.

Method of Analysis

The techniques used to empirically derive the relationships between fractal dimension D and the Fourier spectrum are quite direct. Initially, a random signal with a known spectral form is generated using an algorithm described by FOX (1987). A signal of known spectral form is calculated by first defining the desired amplitude spectrum over all discrete frequencies. For this study, a power law function in the form of equation (2) is used. A corresponding phase spectrum is modeled with uniformly distributed random noise, which agrees with the distribution of measured phase spectra of seafloor profiles as derived by FOX and HAYES (1985, Figure 20). The two series are rotated into the complex plane and inverse Fourier transformed to yield the desired real time series. This series is immediately passed to another routine which measures the fractal dimension D .

There are several methods for measuring the fractal dimension, many of which are described by MANDELBROT (1982). The algorithm used in this study is a form of the divider method, which was used by Richardson for analyzing coastline data. In this technique, the length (L) of a series is measured at a variety of different resolutions (r). The data are modeled as a power law function which plots linearly in log-log space as,

$$\log_{10} L(r) = (1 - D)\log_{10} r + \log_{10} C \quad (7)$$

where D is the fractal dimension and C is a constant. This relationship can be written for linear-linear space as,

$$L = C \cdot r^{(1-D)}. \quad (8)$$

A problem arises in how one handles the segments at the ends of the signal which are not even increments of the divider size (see AVILES *et al.*, 1987, for a complete discussion). Also, MANDELBROT (1985) has shown that serious problems can arise when the technique is applied to self-affine figures.

The particular algorithm for measuring fractal dimension used in this study is a simplification of the divider method. Throughout the various experiments described, a signal length of 2049 points (or 2048 intervals) is used. To simplify the analysis, the divider widths (r) selected are the powers of two between 1 and the data length (2048). Because the signal generation routine described by FOX (1987) uses the Fast Fourier Transform algorithm, the resulting signals always have a length

corresponding to a power of two. Therefore, random signals of 4096 points are generated, and the first 2049 points used. This approach has the advantage of not introducing biases through the estimation of remnant lengths. The disadvantage of the method is the reduced number of values used in the regression analysis for the fractal dimension calculation. As will be shown later, this deficiency can be minimized through ensemble averaging of multiple realizations.

Finally, the total length of the signal is calculated by summing the lengths of individual straightline segments for each resolution distance. Equation (7) is fitted using a linear least-squares regression in log-log space. Since the resolution distance is predefined, all estimation errors are assumed to occur in the measured lengths. If the errors of estimation for the lengths are constant in linear space, the errors in log-log space will be reduced with increasing length. The regression parameters are therefore weighted by the squared value of the corresponding length as described by SCARBOROUGH (1930). A further discussion of this technique and relevant FORTRAN software can be found in FOX (1985, Appendix A.2).

To quantify the random error in estimating the value of D , a series of performance tests were run on the underlying algorithm. The tests consist of estimating the fractal dimension D of random signals which are multiple realizations of identical amplitude spectra. More specifically, the signal generation algorithm was accessed sixteen separate times, always with an amplitude spectrum defined as

$$A = 1000 \cdot f^{-1}, \quad (9)$$

and randomly generated phase spectra with consistent uniform distributions. For each realization, a different, randomly selected seed value is used to generate the random series representing the phase spectrum. In this manner, multiple signals with the same underlying amplitude spectrum and phase spectrum distribution are generated. In theory, each of these signals should have the same fractal dimension D . Therefore by studying the variability of the measured values of D for these signals, the performance of the entire test algorithm can be evaluated.

Figure 1 illustrates a typical output from the algorithm. The figure shows the linear form of the data for log (resolution) versus log (length), indicating the power law form of equation (7). The measured fractal dimension of this analysis is 1.446, well below the value of 1.500 predicted by BERRY and LEWIS (1980). Notice that the data values correspond to powers of 2 from 1 to 2048. The measured D values from all sixteen analyses show an unbiased standard deviation value of $s = \pm .025$ and a mean value of $D = 1.4868$.

Although these results may be adequate for the study, the estimation error within an individual run can be reduced by ensemble averaging the length versus resolution data before performing the regression analysis to calculate D . Because the input signals are artificially generated, there is no restriction on the number of realizations that can be used in each analysis. For all analyses therefore, sixteen

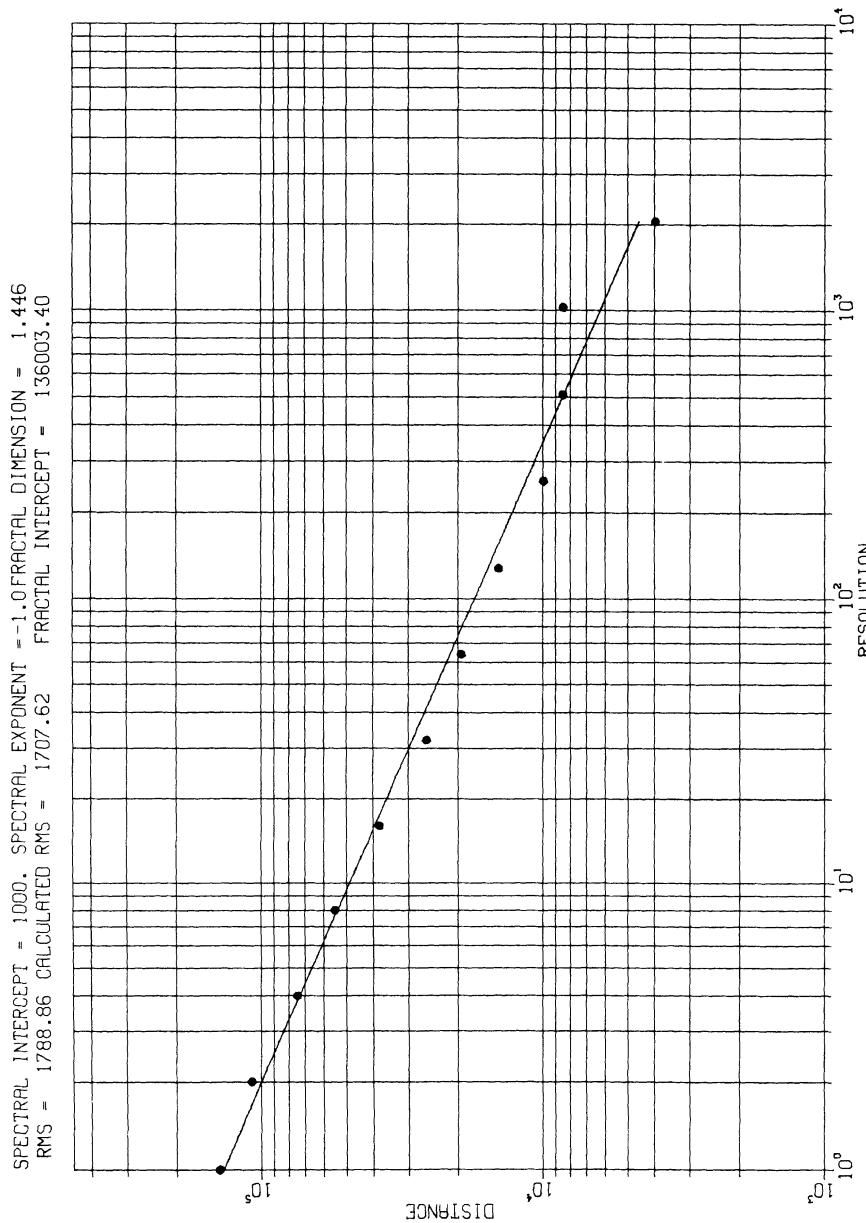


Figure 1
Typical fractal analysis of random data series. Linear form on log-log axes indicates a power law form for the relationship.
Ensemble average of sixteen analyses indicates a standard error of estimate for the fractal dimension of ± 0.025 units.

separate realizations of the artificial signals are generated and the average lengths for each resolution distance are averaged. This averaging should result in a theoretical reduction in the variance of the estimated lengths of $1/N$, or a reduction in the standard deviation of each length measurement of 75%. Figure 2 is a typical output from this improved estimating algorithm. The final estimates of D over a sixteen sample performance test yielded a standard deviation of $s = \pm .005$, and mean fractal dimension of $D = 1.490$.

Experimental Results

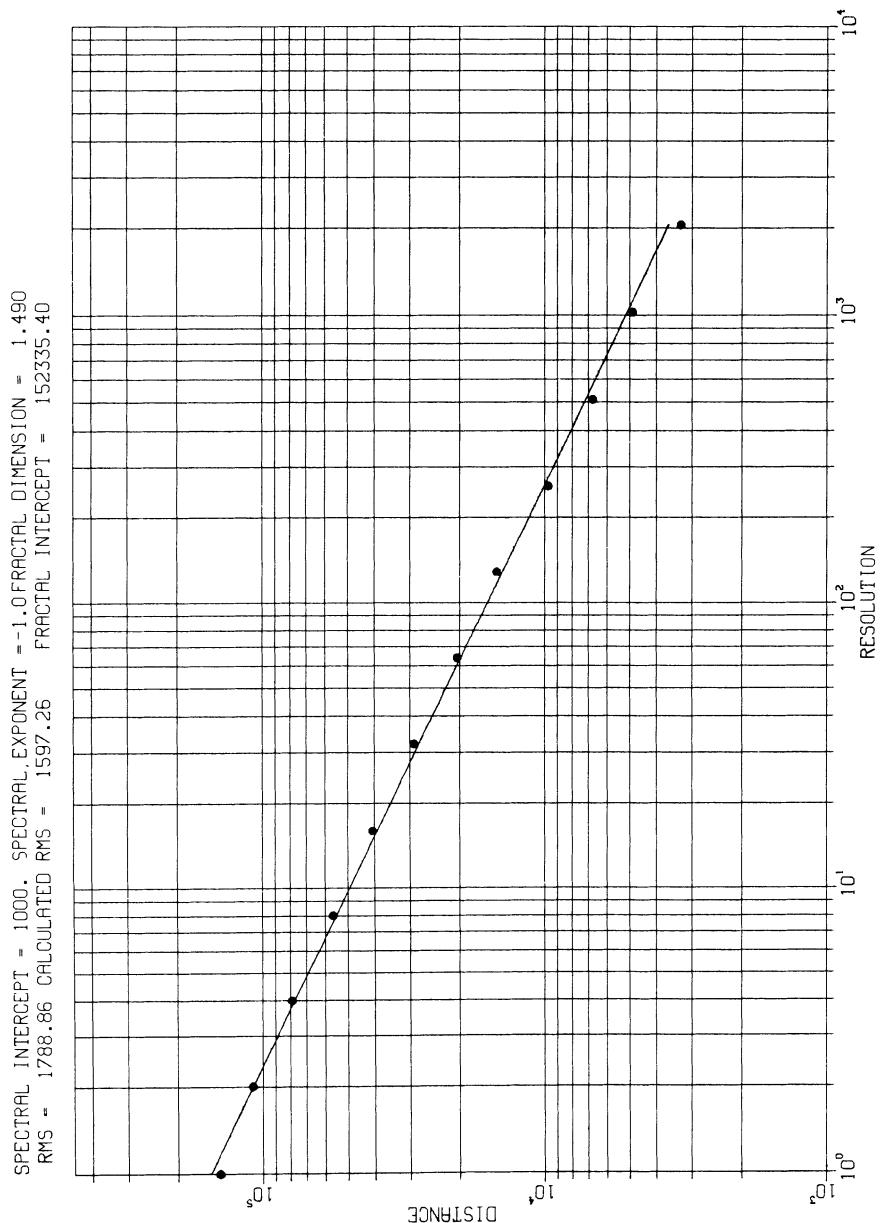
Using the method described, a battery of tests were performed to empirically investigate the relationships between the various parameters of the fractal model and the parameters of the power law spectral model. The model parameters will be referred to as spectral exponent (b), spectral intercept (a), fractal dimension (D), and fractal intercept (C). In addition, the root-mean-square (RMS) variability of the signal is compared in some analyses. It should be emphasized that all derived relationships apply only to the ruler method of fractal measurement.

Functional Form

Before conducting experiments to determine the relationship between the various functional parameters of the fractal and spectral models, it is necessary to ensure that the theoretical forms of each function are correct. The power law form for the amplitude spectrum (equation (2)) is assured since it is imposed by the signal generating algorithm. No windowing, tapering, or prewhitening was performed on the signal. The power law form for the fractal analysis (equation (8)) cannot be assumed, and a series of tests are performed to evaluate the validity of that model.

Figures 1 and 2 indicate that the power law form, which appears linear on the log-log axes, holds for fractal analyses of signals with the particular spectral form of equation (9). To evaluate the power law model over a range of spectral parameters, length versus resolution pairs were generated for a variety of spectral exponents (b) and intercepts (a). Spectral exponent values were extended beyond the range implied by equation (1) for a linear figure. The resulting family of curves were then plotted on log-log axes for visual examination.

Figure 3 illustrates the family of curves resulting from signals generated with amplitude spectra of the form of equation (2), with the spectral intercept held constant at $a = 1000.0$, and the spectral exponent parameter (b) varying from 0.0 to -2.0 in increments of -0.1 . The dots represent the various length versus resolution pairs, and the solid line connects these values with a rational spline, not a fitted function. It is obvious that the linear model of equation (7) does not hold for the full range of the analysis. As was previously mentioned, the measured lengths at



Similar analysis to that illustrated in Figure 1, however the length versus resolution pairs have been averaged over sixteen realizations. Prediction error for fractal dimension is reduced to $\pm .005$.

Figure 2

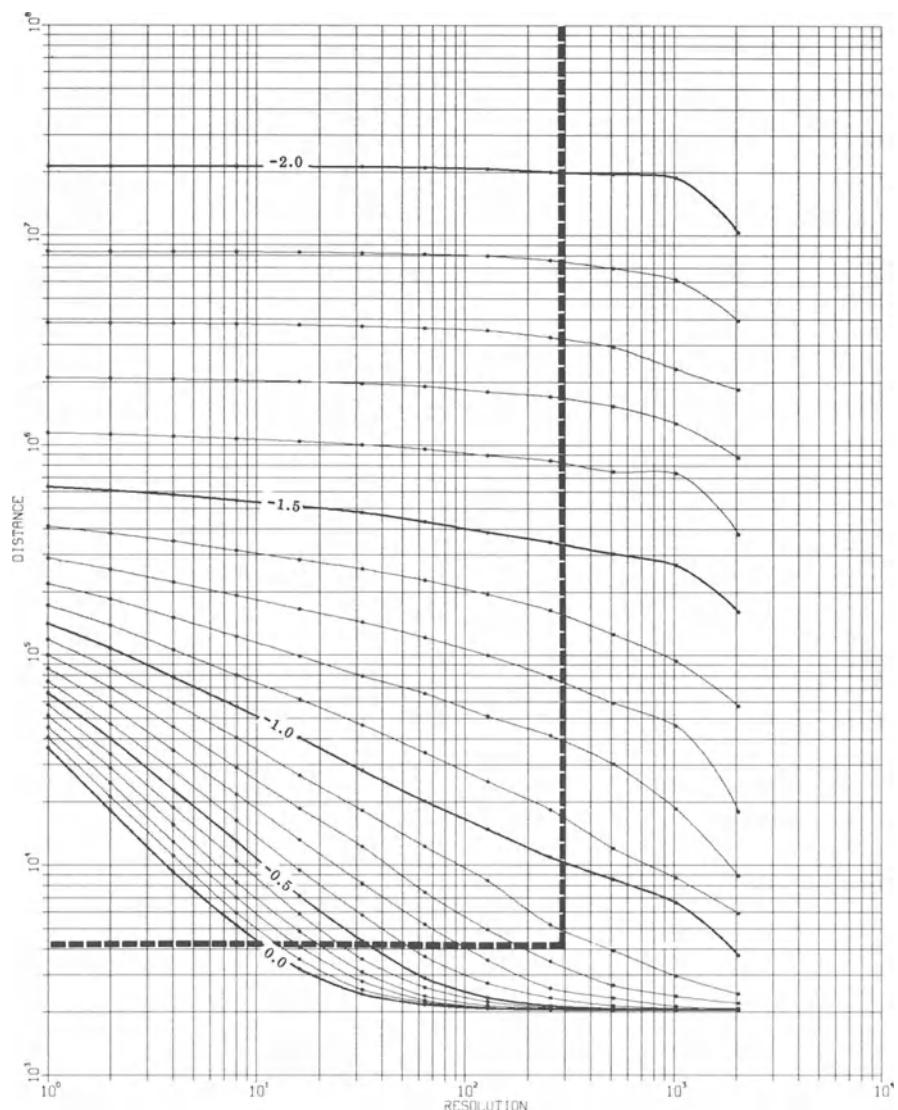


Figure 3

Family of curves for fractal analyses of signals of power law spectral form where $a = 1000.0$ and $0.0 > b > -2.0$. Signals consist of 2049 points (2048 intervals). Notice that data fall on resolutions which are powers of 2. The expected linear form for the curves does not hold over the full range of analysis, as predicted by theory. Future analyses include only those data within the region indicated by heavy dashed lines.

large resolution widths are subject to considerable random fluctuation and tend to curve down as the resolution length approaches the full length of the data series. For this reason, the lengths determined for large resolution spacings are not included in the regression analyses.

In addition, the linear form breaks down towards the base of Figure 3. The full length of the series provides a minimum distance value for the function, independent of the spectral form of the data. That is, the measured length can never be less than the divider width at its largest spacing. Figure 4 is an identical analysis with spectral intercept (a) reduced by an order of magnitude to $a = 100.0$. Inspection of these examples indicates that the breakdown of the linear model occurs as a function of measured length, independent of either spectral parameter a or b . Figure 5 illustrates a family of curves based on spectral models identical to those used to generate Figure 3, but with the data length reduced to 1024 points. The form of the results appears to be independent of data length, however the family of curves is displaced downwards, due to the shorter maximum resolution distance. The location of equivalent curvature on the log (distance) axis moves relative to the length of data and may be related to the "crossover length" described by MANDELBROT (1985) and WONG (1987).

It is apparent that an investigator using the fractal technique must examine the form of the analysis data, before routinely applying the power law form over the full range of the length versus resolution pairs. It is unlikely that a different functional form will be derived, and the power law model does appear to be adequate within certain boundaries. For future analyses, the power law functions are fitted to only the data within the arbitrary boundaries represented by heavy dashed lines in Figures 3, 4, and 5. These correspond to twice the data length for the lower boundary (length = 4096 for most analyses) and resolution widths less than 300.

Fractal Dimension versus Spectral Exponent

The analytical relationship derived by BERRY and LEWIS (1980) predicts an approximately linear relationship between fractal dimension D and spectral exponent b , as shown in equations (4) and (5). In order to test this relationship for the discrete case, the families of curves illustrated in Figures 3 and 4 were fitted with power law functions within the indicated boundaries, and the derived fractal dimensions (D) were plotted versus the known spectral exponents (b) for all curves. Figure 6 represents these results, with the theoretical linear relationship plotted as a dashed line. The form of the results is consistent in both cases ($a = 100.$ and $a = 1000.$), but the curve is somewhat shifted for the lower amplitude example. No estimates could be calculated for the spectral exponents where $b > -0.5$ and $a = 100.$, due to insufficient data for the regression.

The theoretical linear model does not hold over the full range of the analysis.

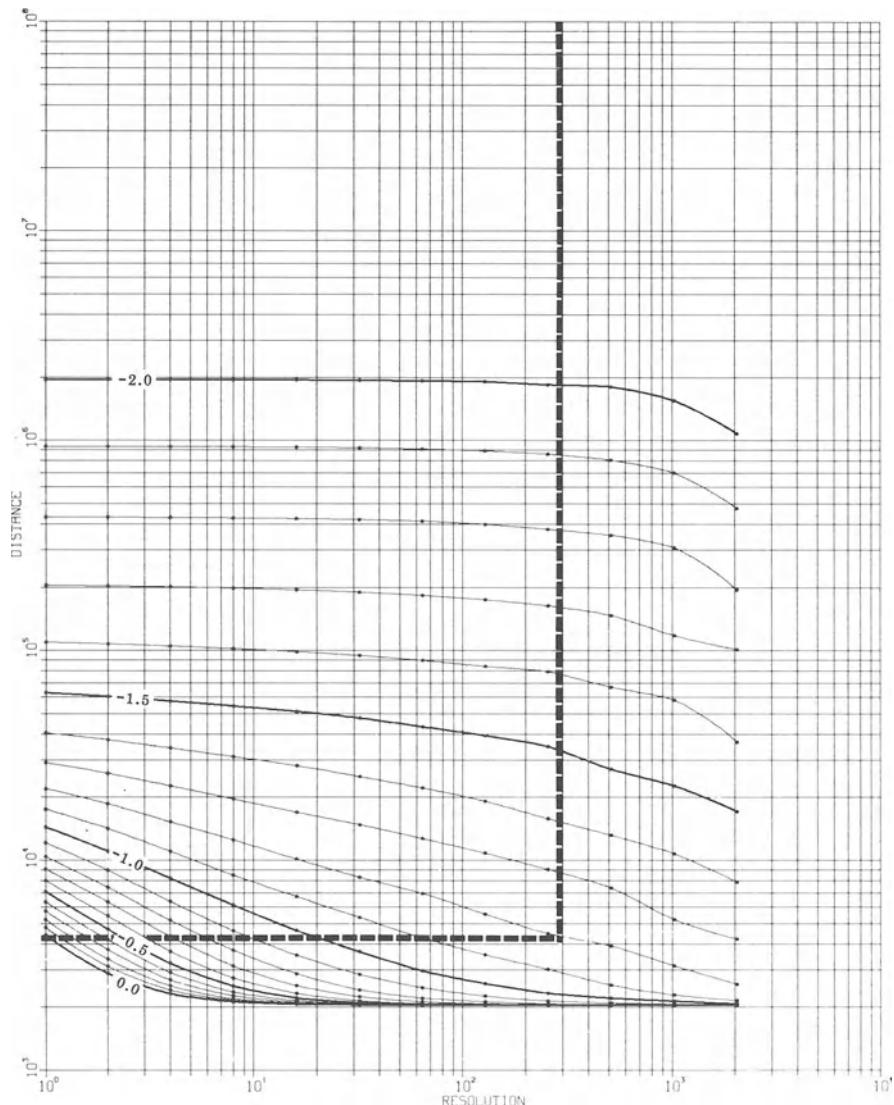


Figure 4

Similar analysis to that shown in Figure 3 with spectral intercept value reduced to $a = 100.0$. Notice that the overall lengths are reduced for all resolutions, but the boundaries of nonlinearity do not migrate.

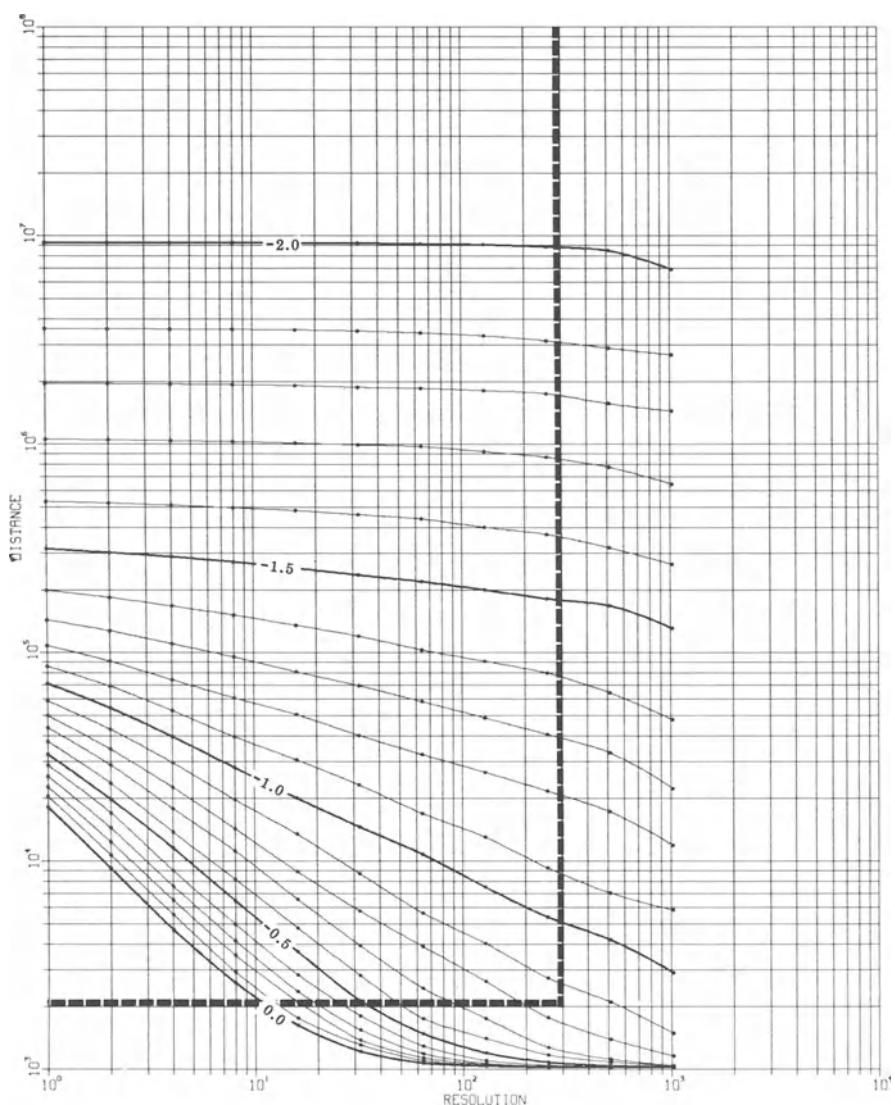


Figure 5

Similar analysis to that shown in Figure 3 with number of intervals in the input series reduced to 1024. Notice that the curves are identical but migrated downward in distance, and the lower boundary of nonlinearity is also shifted downward.

Spectral Exponent versus Fractal Dimension

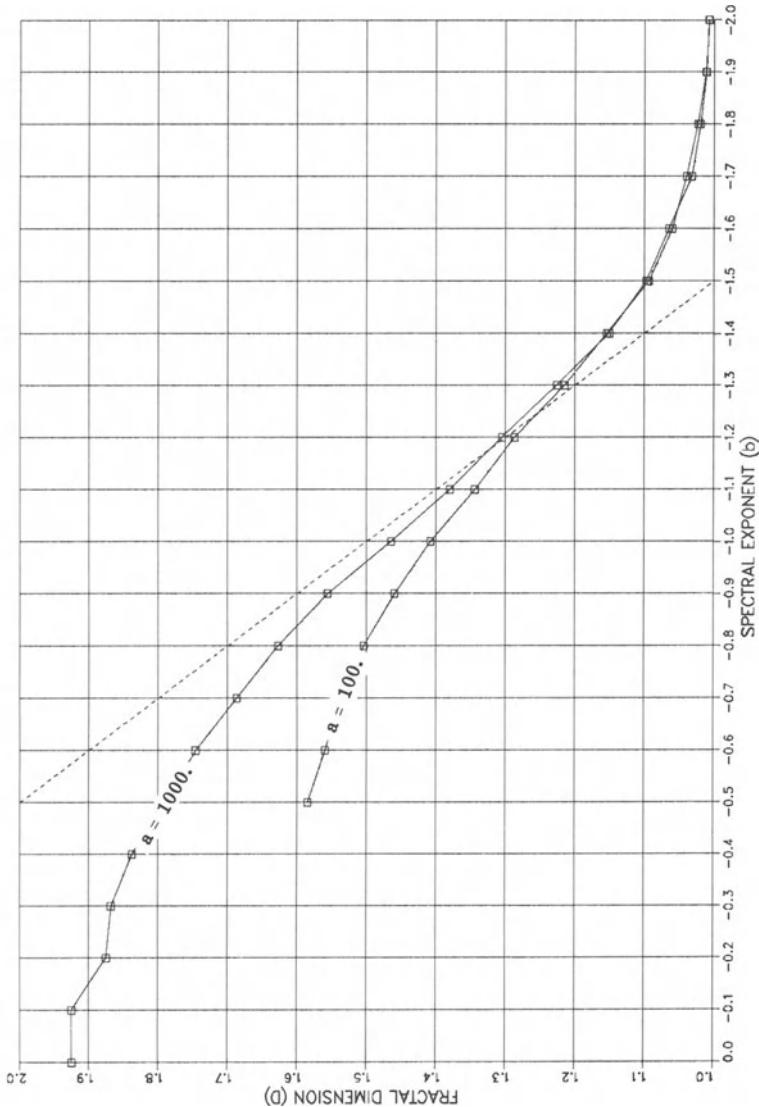


Figure 6

Summary of the relationship of fractal dimension to spectral exponent for data shown in Figures 3 and 4. The theoretical relationship of BERRY and LEWIS (1980) is shown as a dashed line. The measured results do not agree with the predicted results, the error increasing for lower amplitude input signals. The analysis is extended beyond the limits of the prediction equation.

The measured data appear nearly linear about an inflection point at $b = -1.0$. This is the particular spectrum for which the component sinusoids are self-similar. The corresponding value of fractal dimension (D) at that point falls somewhat below the predicted value of $D = 1.5$. The results away from this inflection point diverge substantially from the model of BERRY and LEWIS (1980), overestimating the fractal dimension at lower values (near $D = 1.0$) and underestimating higher values (near $D = 2.0$). This result confirms the observations of AVILES *et al.* (1987), who describe a significant overestimation of the spectral method over Richardson's method for fractal dimensions in the range $1.1 < D < 1.5$.

The behavior of signals with spectral exponent $b = -0.5$ is predicted by fractal theory and the BERRY and LEWIS (1980) model as corresponding to plane-filling curves ($D = 2.0$). The extent to which a time series of this form fills a plane, depends upon the absolute amplitudes of the component sinusoids. The higher the amplitude of each frequency component, the larger the statistical variability of the resultant signal, and the more nearly plane-filling the signal becomes. The magnitude of all amplitudes is determined by the spectral intercept term (a) in equation (2). Figure 7 plots measured fractal dimension (D) as a function of spectral intercept (a) for the two boundary limits of the model. The data for $b = -0.5$ appear to asymptotically approach $D = 2.0$ but are still well below that predicted limit, even for values of (a) which are larger than would be found on most naturally occurring surfaces. The data for $b = -1.5$ do not appear to converge to the predicted value of $D = 1.0$, which would imply a fully differentiable profile.

It is not possible for an ordered time series (that is, random only in the response variable) to completely fill a plane. There can be only one y -value for a given x -value. Neither is it possible for discrete series to be fully differentiable at all scales; they are by definition frequency band limited. Perhaps it is these fundamental differences between the analytical model and its application to discrete data which account for the divergence of empirically derived results from theoretical prediction.

The restriction of an ordered time series can be removed by generating figures which vary randomly in both x and y . Such doubly random figures are produced by combining two independently generated series using the inverse Fourier transform technique. These series provide a better model for certain applications, such as coastlines, and conform more closely to previous fractal studies. Figure 8 illustrates a variety of these doubly random series generated through Fourier techniques. In all cases, identical spectral parameters were used to produce the independent x and y series, although anisotropic figures are possible using the same technique.

The extremes of dimension are apparent in Figure 8. The $b = -0.5$ appears nearly plane filling as predicted by theory, yet the measured value of fractal dimension remains below $D = 2.0$. As in the singly random series, the measured fractal dimensions approach $D = 1.0$ for decreasing values of b , indicating the convergence to a differentiable series. Figure 9 compiles results for the full range of

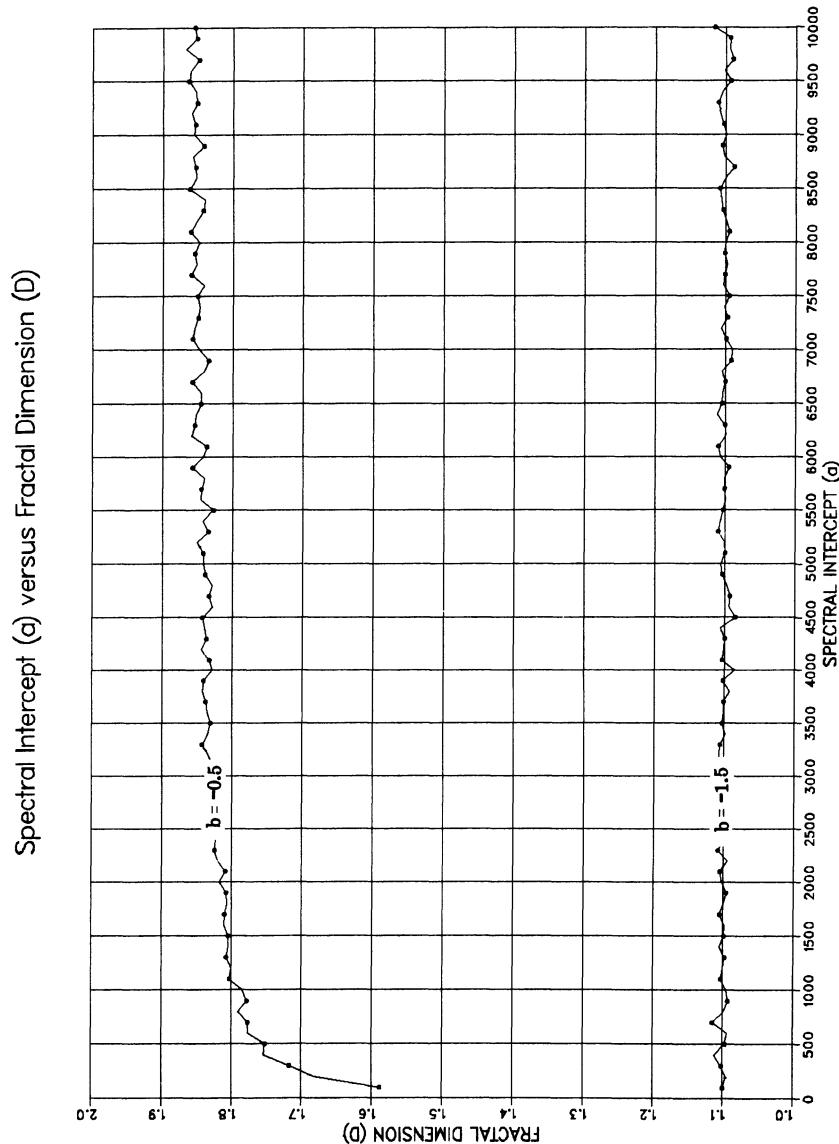


Figure 7

Fractal dimension versus spectral intercept for the two extremes of the BERRY and LEWIS (1980) model. Increasing values of (a) force the $b = -0.5$ results towards the predicted value of $D = 2.0$, but the result falls far short for higher amplitude curves than would be found in nature. The lower boundary ($b = -1.5$) appears unaffected by the increase in spectral energy and does not converge on $D = 1.0$.

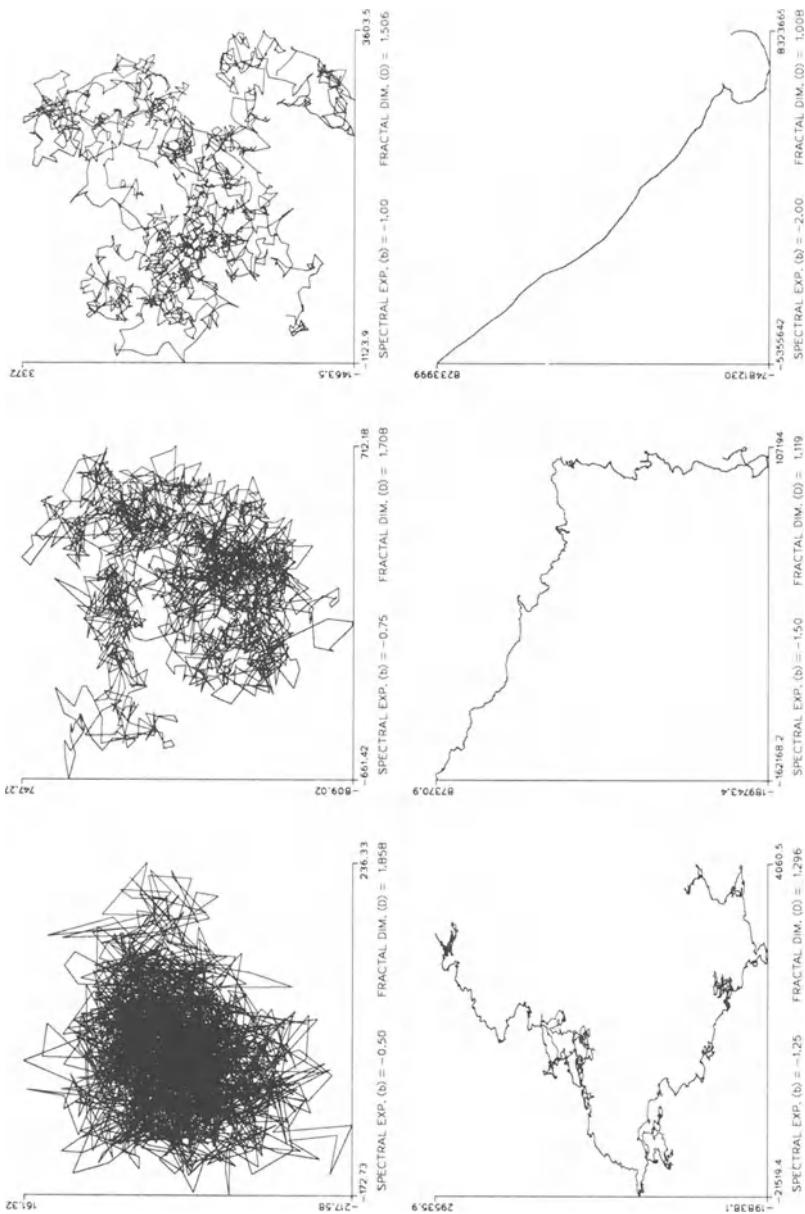


Figure 8
Doubly random series generated through Fourier techniques. In all cases, the spectral exponent (α) was held constant for the independent generation of the x and y series. The resulting curves range from nearly plane filling in the upper left to nearly differentiable in the lower right.

b values. As in the case of simple time series, these results do not conform to the theoretical model of BERRY and LEWIS (1980).

Fractal Intercept versus Spectral Intercept

One of the characteristic attributes to the fractal dimension (*D*) is its scale independence. The value of *D* reflects the rate of change in length of a signal as a function of measurement distance, without regard to the actual length at any resolution width. This is analogous to the spectral exponent (*b*), which determines the rate of change of amplitude with spatial frequency, without specifying an amplitude at a given wavelength. Given the fractal dimension (*D*) or spectral exponent (*b*) alone, one cannot infer any physical length of amplitude for the signal under investigation. Although this scaleless property is part of the elegance of fractal theory, a reference scale is often required to solve physical problems. For example, to study the interaction of sound with a rough surface, it is not enough to know the rate of change of roughness with scale; the actual variability at a specified scale (or wavelength) is required.

The proportionality constants *C* (in equation (8)) and *a* (in equation (2)) represent the measured length and amplitude respectively of a signal, at its *y*-intercept in log-log space. The fractal intercept (*C*) corresponds to the length of the signal at a resolution of one unit. Since this value is usually derived from a regression model, it is not necessarily the precise measured length, but rather the length as inferred by the full regression model. Similarly, the spectral intercept (*a*) represents the amplitude of a sinusoid of frequency $f = 1.0$ cycle/data interval, or a wavelength of one data interval (unit length) in the case of this study. With knowledge of both the intercept and exponent parameters in either model, the signal length (or amplitude) at any resolution (or spatial frequency) can be calculated.

To investigate the relationship of fractal intercept (*C*) to spectral intercept (*a*), a family of curves was generated for a range of *a*-values from 100.0 to 2000.0 in steps of 100.0, with spectral exponent (*b*) held constant. Separate runs were made for a full range of *b*-values. Figure 10 illustrates typical results for the analysis, in this case for *b* = -1.0. The linearity of the relationship is apparent. An analysis of variance for the regression model indicates a perfect linear association ($r = 1.000$). The slope (*u*) of the regression line varies as a function of spectral exponent (*b*). The *y*-intercept for the linear regression is assumed equal to zero, therefore the linear relationship can be simply written as,

$$C = u(b) \cdot a. \quad (10)$$

An additional measure of the roughness of a signal is the root-mean-square (RMS) variability. Band-limited RMS is used by FOX and HAYES (1985) as the basic parameter for delineating homogeneous provinces on the seafloor.

Spectral Exponent versus Fractal Dimension

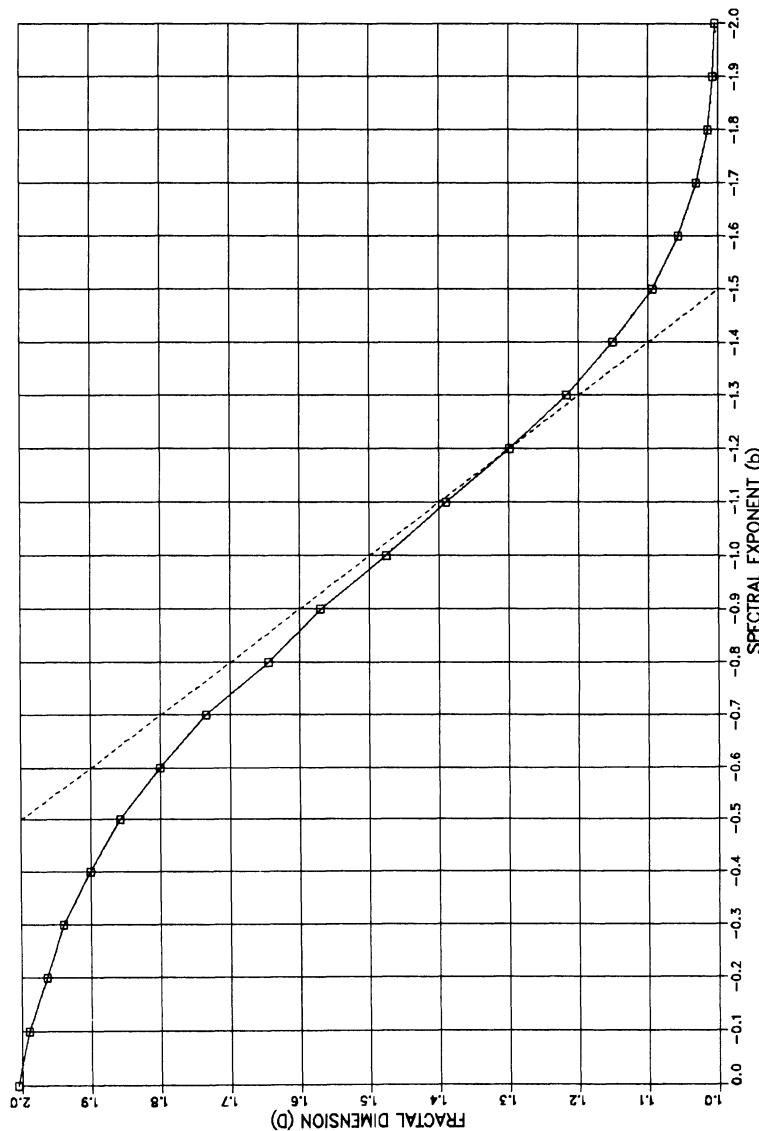


Figure 9

Summary of the relationship between fractal dimension and spectral exponent for doubly random series, such as those shown in Figure 8. The theoretical model of BERRY and LEWIS (1980) is shown as a dashed line. Like the ordered time series, the measured fractal dimension does not conform to the mathematically derived relationship.

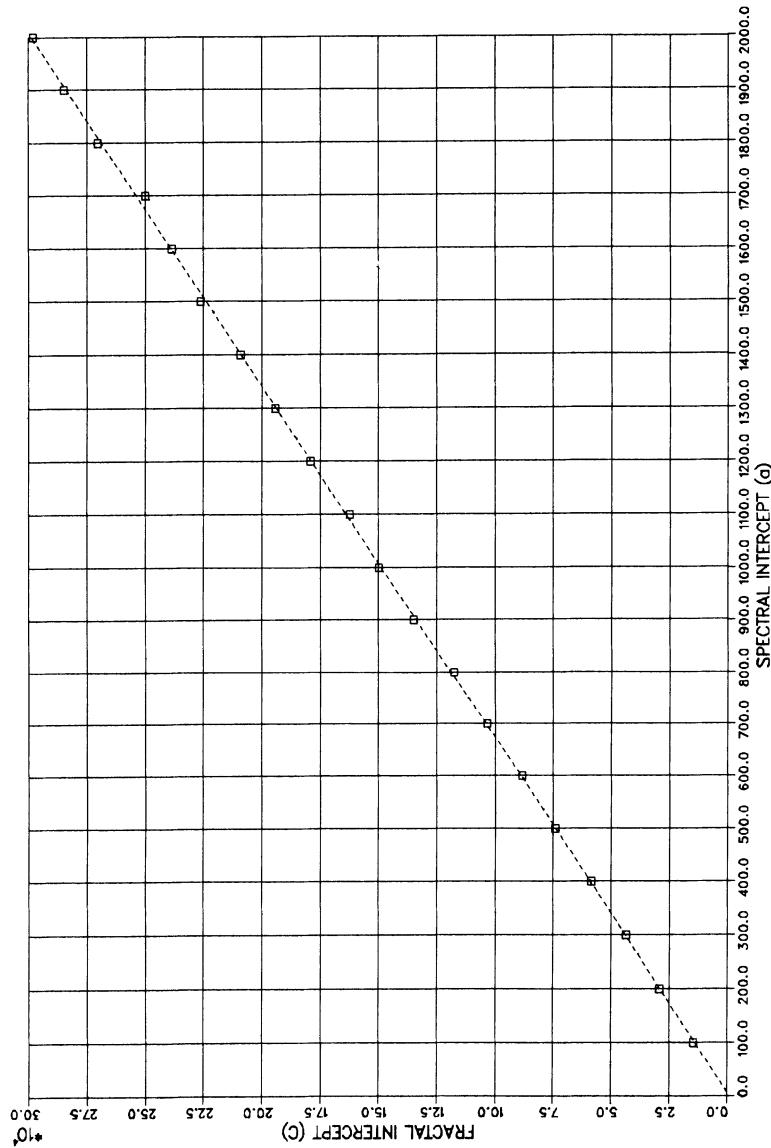
Spectral Intercept (α) versus Fractal Intercept (C)

Figure 10
Relationship of fractal intercept (C) to spectral intercept (α). The data show a perfect linear relationship with $r = 1.00$.

This fundamental statistic can be calculated using the standard formula from statistics,

$$\text{RMS} = \left(\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \right)^{1/2} \quad (11)$$

where \bar{x} is the mean. In this case, the RMS represents the sample standard deviation. Alternatively, the RMS can be estimated by integrating the frequency spectrum,

$$\text{RMS} = \left(\frac{\int_{-N}^N |X^2(f)| df}{2N} \right)^{1/2} \quad (12)$$

where N is the Nyquist frequency, and $X(f)$ is the amplitude spectrum. Figure 11 illustrates the relationship between fractal intercept (C) and root-mean-square (RMS) variability. Again, a simple linear model produces an excellent fit to the data. The estimates of RMS from integrating the signal generating spectrum (equation (12)) are much more stable than those derived by equation (11). Recalling the relationship between fractal intercept (C) and length, RMS variability is therefore linearly related to length for the power law form spectral model or fractal model.

Nonstationary Effects

The requirement of statistical stationarity is fundamental to Fourier techniques. Discrete Fourier analysis can be viewed as a multiple regression process in which the parameters of each component sinusoid are derived through least squares fitting (BLOOMFIELD, 1976). Interpretation of the resulting spectrum requires the assumption that the properties of the underlying signal remain essentially constant over the length of the signal. Statistical theory rigorously defines stationary series as signals in which mean, variance and autocovariance function do not vary with position (CHATFIELD, 1980). Fox (1985) argues that these definitions are overly restrictive to allow analyzing the earth's topography through spectral analysis and presents an algorithm for delineating provinces of "homogeneous" topography by discretely estimating the amplitude spectrum in the spatial domain.

In the preceding experiments, stationarity of the random signals was assured by the generating algorithm. Just as the Fourier transform requires a stationary input signal, the inverse Fourier technique developed by Fox (1987) ensures a stationary output signal. Similarly, the generation of fractal figures from a mathematical algorithm assures stationarity in some sense. In *measuring* natural signals by either technique, however, homogeneity of a profile over its entire length cannot be assumed. Fox and HAYES (1985) show that seafloor topography is quite nonstationary, and their provincing techniques can only attempt to minimize the errors

Root Mean Square versus Fractal Intercept

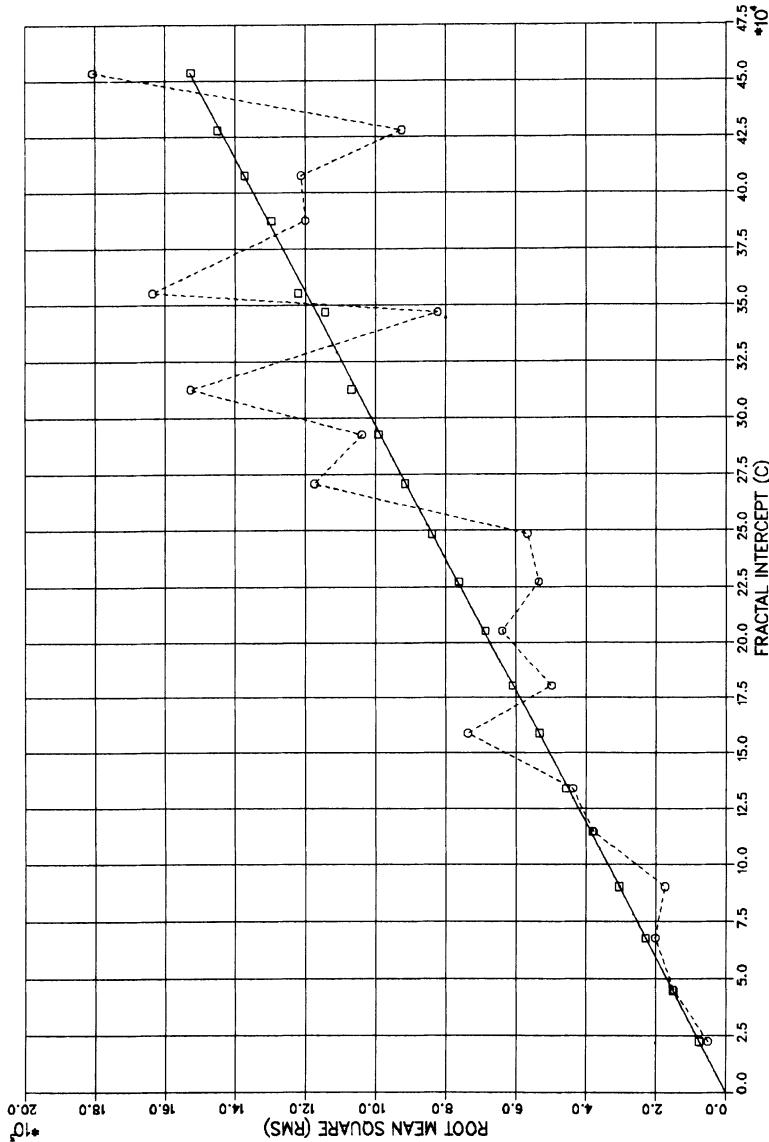


Figure 11
Root-mean-square (RMS) variability of a signal versus its fractal intercept term (C). RMS calculated from integration of the frequency spectrum is shown by small squares; RMS calculated by statistical formula is shown as circles. Solid line is regression curve.

due to nonstationary effects. To assess the effects of nonstationarity of a signal on fractal analysis, a series of experiments were conducted in which a signal composed of several segments of differing spectral characteristics was subjected to the standard analyses used in the preceding sections.

The first signal considered is composed of 8 segments of 256 points. The spectral intercept is held constant at $a = 1000.0$ in each segment, while the spectral exponent (b) is stepped from -0.7 to -1.4 , in steps of -0.1 . An example of this artificially generated nonstationary series is shown in the upper diagram in Figure 12. Note the increasing amplitude of low frequency components towards the right side of the profile, as b decreases below the self-similar value of $b = -1.0$. The distance versus resolution diagram illustrated below in Figure 12 maintains a consistent power law form and appears no different from an analysis of a stationary series, such as that illustrated in Figure 2. An investigator analyzing a similar profile in nature would have no indication of its inherent nonstationarity.

The nonstationary series shown in Figure 12 has a measured fractal dimension of $D = 1.3661$. Applying equation (5) to the arithmetic mean of the spectral exponents comprising the signal ($b = -1.05$) yields an estimated fractal dimension of $D = 1.45$. A further test was performed in which a series of nonstationary signals were analyzed for fractal dimension. Each signal is composed of eight segments of 256 points, each with $a = 1000.$ and b values stepped by -0.05 , the initial value ranging from $b = 0.0$ to $b = -1.65$. The results summarized in Figure 13 are quite similar to the results shown in Figure 6, which are derived from stationary series. Apparently, the fractal dimension measured on a nonstationary series of a given mean spectral exponent, cannot be discerned from that measured on a stationary series of the same spectral exponent. Since the inferred dimension may apply to only part or indeed none of the profile under study, this effect could result in erroneous interpretations of natural processes.

An additional experiment was conducted in which the 8 segments comprising the input signal are maintained at a constant spectral exponent (b), while the proportionality constant (a) is varied. For the example shown in Figure 14, the spectral exponent is held at $b = -1.0$, while the spectral intercept is varied from $a = 1000.0$ to $a = 2750.0$ in steps of $+250.0$. The change in the profile appears as a general increase in variability from left to right. Again, the length versus distance pairs plot in power law form, and yield for this example a measured fractal dimension of $D = 1.5265$. No experiments were performed to examine the effects of nonstationarity on fractal intercept (C).

In light of these conclusions, an unexplained result obtained by FOX and HAYES (1985, p. 24) can be reexamined. In that study, profiles of bathymetry from the northeastern Pacific Ocean were found to vary widely with respect to their measured spectral parameters. Presumably, the measured fractal dimension of these surfaces also varies geographically, in accordance with the relationship derived in this study. These areal variations were related to differing materials and relief-forming

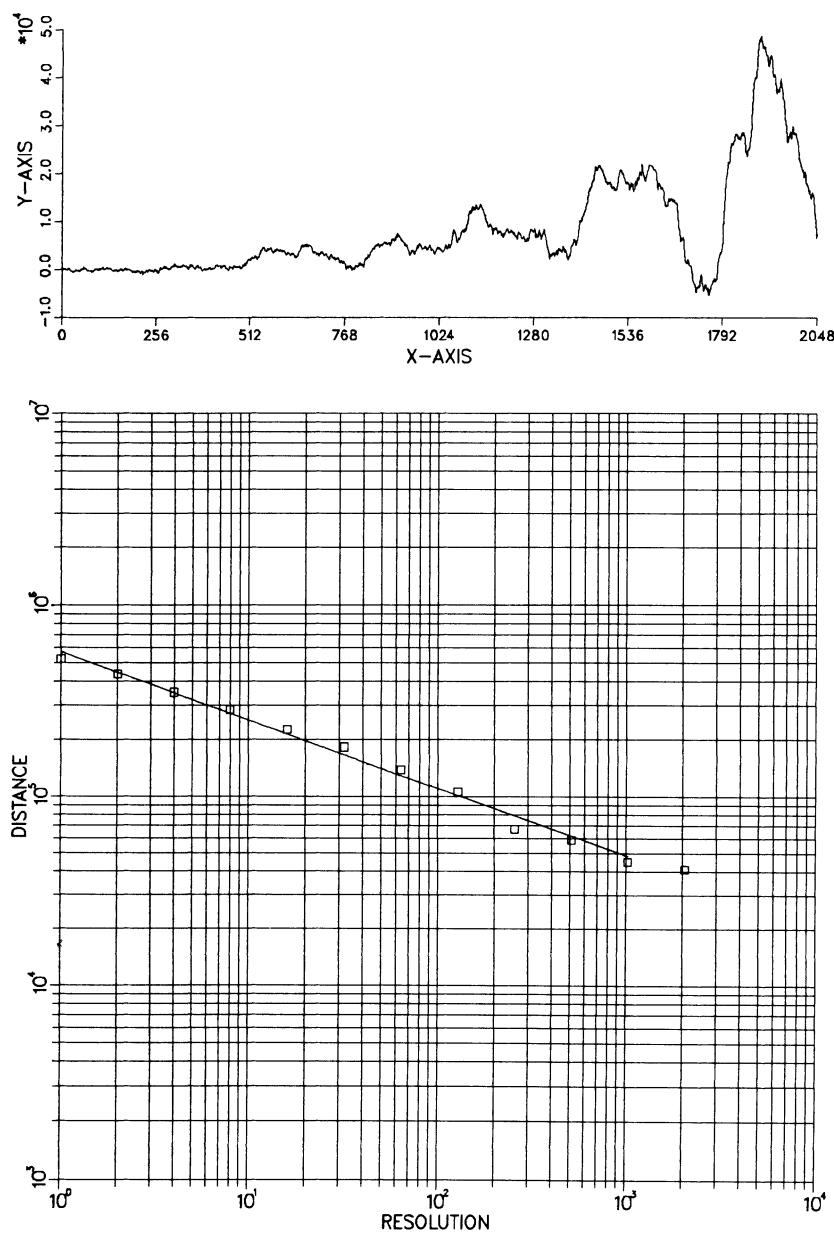


Figure 12

Fractal analysis of signal which is nonstationary with respect to its spectral characteristics. Series (shown above) is constructed of eight segments of 256 points which range from $b = -0.7$ to $b = -1.4$ in steps of -0.1 . Spectral intercept is held constant at $a = 1000.0$. Despite the nonstationarity of the input signal, the fractal analysis (below) appears no different than for stationary series.

Measured Fractal Dimension of Nonstationary Series

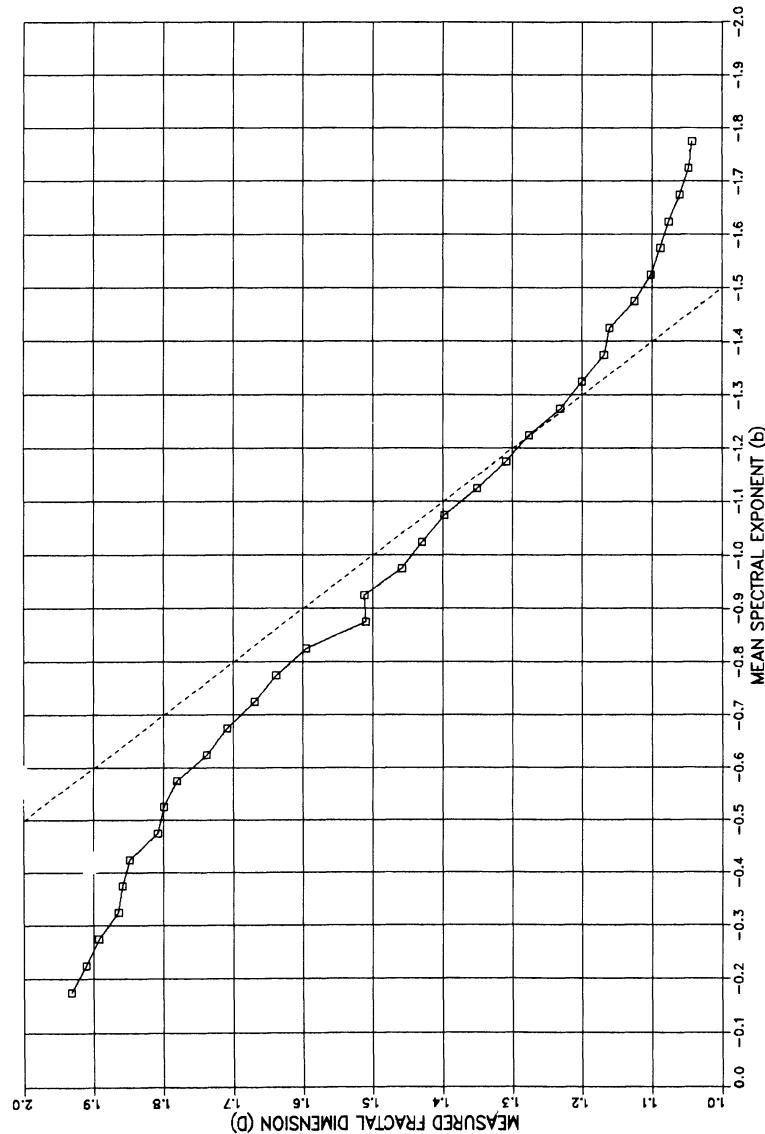


Figure 13

Summary of measured fractal dimensions for various nonstationary signals. The results look similar to the results from stationary series illustrated in Figure 6, when plotted against the mean spectral exponent (\bar{b}) for all segments.

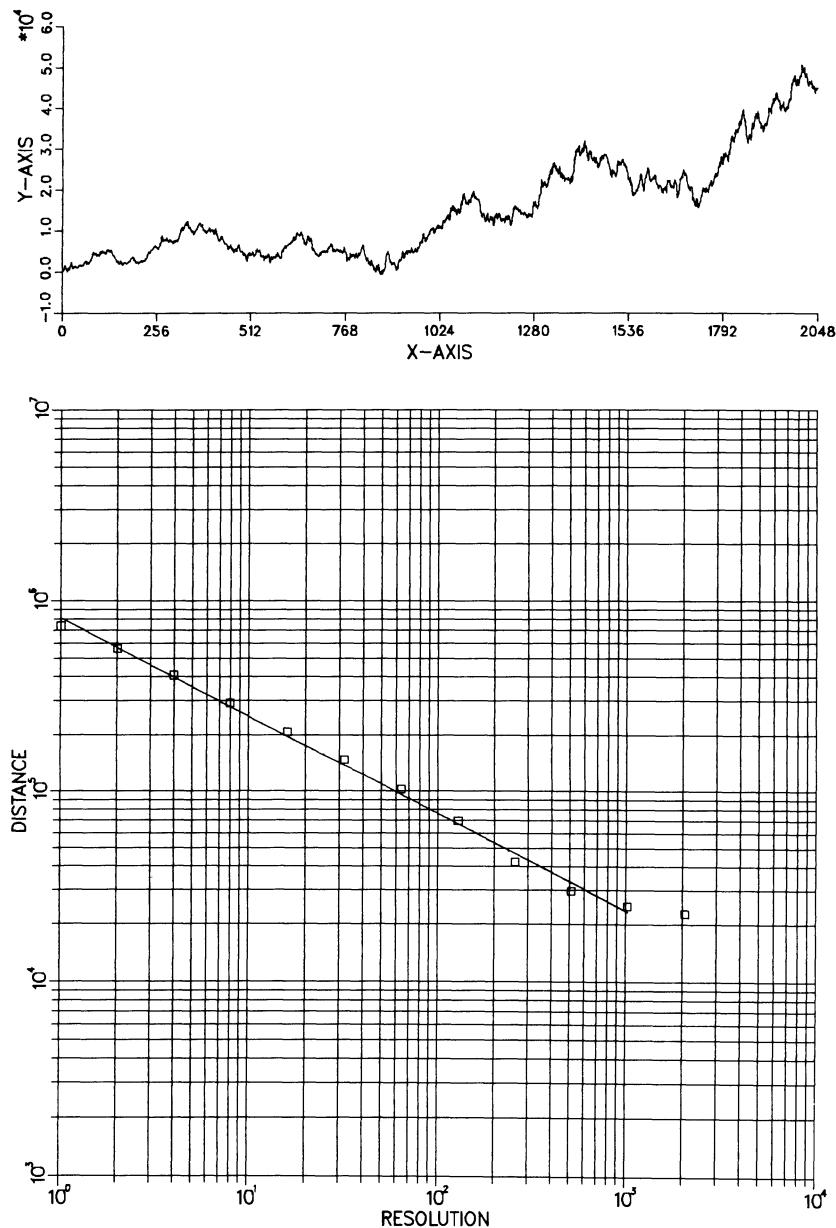


Figure 14

Additional analysis of nonstationary series where spectral exponent is held constant at $b = -1.0$, and spectral intercept is increased from $a = 1000$. to $a = 2750$. in steps of 250. Again, the analysis appears identical to that of a stationary series.

processes at work on the seafloor. The earlier study of BELL (1975) which analyzed long profiles without restricting nonstationary effects, measured spectra with consistent spectral exponents of $b = -1.0$. FOX and HAYES (1985) found spectral exponents in the range of $-1.2 > b > -1.7$ for profiles which had been previously divided into quasi-stationary segments. However, when these segments were recombined into long (and presumably nonstationary) profiles, the measured spectral exponents consistently approached the self-similar value of $b = -1.0$.

The convergence of long topographic profiles to the self-similar model seems to indicate a general rule for nonstationary series. However, the results obtained above indicate that the measured fractal dimension of a nonstationary series should correspond to the mean fractal dimension of its component segments. Lacking a mathematical explanation, the self-similar dimension of long profiles may reflect some physical law. There are magnitude limits for the aspect ratio of naturally occurring features. Gravitational and erosional forces will not allow an infinitely peaked mountain. Sedimentary features have a limiting angle of repose which restricts the gradients, and therefore the aspect ratios, in sedimentary provinces. It is also difficult to envision a natural surface being perfectly smooth. Since these physical limits apply globally and at all scales, the earth's surface may appear self-similar in the mean, although self-affine within particular quasi-stationary provinces and/or spatial frequency bands.

Conclusions

A method has been developed for investigating the relationships of fractal methods and Fourier techniques as they are applied to discrete random signals. A series of experiments compare the several parameters of each model over a range of input signals. These signals are constructed to simulate profiles found in nature, without the uncertainties associated with analyzing natural measured profiles. The empirically derived results are compared to theoretical relationships derived through analytical mathematical techniques. In many cases these empirically derived results differ significantly from the results predicted from theory. Some of the major results are the following:

- 1) The predicted power law form for the fractal analysis (measured length versus resolution) does not hold over the entire range of the analysis. The functional form diverges as the resolution approaches the data length and as the measured length approaches the minimum overall length. The boundaries within which the power law form does hold, varies with data length and the spectral parameters of the input signal. Investigators using the fractal ruler technique for quantifying the roughness of surfaces must examine the form of the length versus resolution pairs previous to curve fitting.
- 2) The relationship of fractal dimension to spectral exponent does not conform to the linear relationship derived by BERRY and LEWIS (1980) over all fractal

dimensions. The spectral technique tends to overestimate dimension at lower values of D and underestimate dimension at higher values of D . The magnitude of the discrepancy decreases in higher amplitude signals, but the effect can be substantial for signals typically found in nature. This nonlinear form is true for single-value time series, as well as multivalued, doubly random series.

- 3) In addition to the fractal dimension, the proportionality constant in the fractal analysis (or fractal intercept) can provide important information on the actual scale of the feature. This scaling parameter, which is a least squares estimation of the length of the signal at resolution $r = 1$, is functionally related to the two parameters of the power law form amplitude spectrum. This correspondence may allow a functional relationship between the two techniques to be developed based on length as the common parameter. Root-mean-square variability of the signal is shown to be linearly related to the fractal intercept.
- 4) Profiles which are nonstationary with respect to their spectral parameters (and therefore fractal dimension), appear no different than stationary profiles when subjected to fractal analysis. Since both fractal and spectral techniques assume homogeneous conditions over the length of the analysis, this effect could easily lead to false interpretations of the underlying profiles. Such difficulties do not arise in generating fractal figures, but only when the technique is used for measurement of natural figures. The self-similar geometry observed in long deep-sea profiles cannot be explained as an artifact of nonstationary sampling, and must therefore reflect some physical limits on the relief-forming processes.

The discrepancies between theoretical prediction and empirical results described in this study reflect the difficulties of applying analytically derived techniques to measurement data. Both Fourier and fractal techniques are formulated through rigorous mathematics, assuming various conditions for the underlying signal. When these techniques are applied to discrete, finite length, nonstationary series, certain statistical transformations must be applied to the data. Methods such as windowing, prewhitening, and anti-aliasing filters have been developed over many years for use with Fourier analysis. At present, no such statistical theory exists for use with fractal analysis. It is apparent from the results of this study that such a statistical foundation is required before the fractal ruler method can be routinely applied to observational data. It is quite possible that fractal measurement techniques other than the ruler method produce results which better compare to theory for self-affine cases. The methods described in this study could be applied equally to testing other algorithms.

Acknowledgements

Thanks to Andy Lau, Kurt Molley, and Marijke van Heeswijk for mathematical assistance. Special thanks to Alberto Malinverno for a thorough review. This study was supported under NOAA's VENTS research program. PMEL contribution number 928.

REFERENCES

- AVILES, C. A., SCHOLTZ, C. H., and BOATWRIGHT, J. (1987), *Fractal Analysis Applied to Characteristic Segments of the San Andreas Fault*, Journal of Geophysical Research 92 (B1), 331–344.
- BARENBLATT, G. I., ZHIVAGO, A. V., NEPROCHNOV, Yu. P., and OSTROVSKIY A. A. (1984), *The Fractal Dimension: A Quantitative Characteristic of Ocean-bottom Relief*, Oceanology 24 (6), 695–697.
- BELL, T. H. (1975), *Statistical Features of Sea-Floor Topography*, Deep-Sea Research 22, 883–892.
- BERRY, M. V. and LEWIS, Z. V. (1980), *On the Weierstrass-Mandelbrot Fractal Function*, Proceedings of the Royal Society of London A370, 459–484.
- BLOOMFIELD, P. R., *Fourier Analysis of Time Series: An Introduction* (John Wiley and Sons, New York 1976) pp. 258.
- BRACEWELL, R. N., *The Fourier Transform and its Applications* (McGraw-Hill, New York 1978) pp. 381.
- BROWN, S. R. (1987), *A Note on the Description of Surface Roughness Using Fractal Dimension*, Geophysical Research Letters, (submitted).
- BROWN, S., and SCHOLTZ, C. H. 1985, *Broad Bandwidth Study of the Topography of Natural Rock Surfaces*, Journal of Geophysical Research 90 (12) 12,575–12,582.
- CHATFIELD, C., *The Analysis of Time Series: An Introduction* (Chapman and Hill, New York 1980) pp. 268.
- FOX, C. G. (1987), *An Inverse Fourier Transform Algorithm for Generating Random Signals of a Specified Spectral Form*, Computers and Geoscience 13 (3), 1–6.
- FOX, C. G. (1985), *Description, Analysis and Prediction of Sea-floor Roughness Using Spectral Models*, Naval Oceanographic Office Technical Report 279, Bay St. Louis, MS., pp. 218.
- FOX, C. G., and HAYES, D. E. (1985), *Quantitative Methods for Analyzing the Roughness of the Seafloor*, Reviews of the Geophysics and Space Physics 23 (1), 1–48.
- MANDELBROT, B. B. (1985), *Self-affine Fractals and Fractal Dimension*, Physica Scripta 32, 257–260.
- MANDELBROT, B. B., *The Fractal Geometry of Nature* (W. H. Freeman, San Francisco, CA 1982), pp. 460.
- MANDELBROT, B. B. (1967), *How Long is the Coastline of Britain? Statistical Self-similarity and Fractal Dimension*, Science 155, 636–638.
- OPPENHEIM, A. V., and LIN, J. S. (1981), *The Importance of Phase in Signals*, Proceedings of the IEEE 69 (5) 529–541.
- SAYLES, R. S., and THOMAS, T. R. (1978), *Surface Topography as a Nonstationary Random Process*, Nature 271, 431–434.
- SCARBOROUGH, J. B., *Numerical Mathematical Analysis* (The Johns Hopkins Press, Baltimore, MD 1930) pp. 46.
- SCHOLTZ, C. H., and AVILES, C. A. (1986), *The fractal geometry of faults and faulting*, In *Earthquake Source Mechanics*, Geophysical Monograph 37, 147–155.
- WONG, P., *Fractal surfaces in porous media*, in *Physics and Chemistry of Porous Media II*, AIP Conference Proceedings, 154 (eds. Banavar, J., Koplik, J., and Winkler, K.) (Amer. Inst. Phys., New York 1987) pp. 304–318.

(Received May 1, 1987, revised November 4, 1987, accepted November 5, 1987)

Are Topographic Data Sets Fractal?

LEWIS E. GILBERT¹

Abstract—The scale invariant properties of fractal sets make them attractive models for topographic profiles because those profiles are the end product of a complex system of physical processes operating over many spatial scales. If topographic data sets are fractal, their power spectra will be well represented by lines in log-log space with slopes s such that $-3 \leq s < -1$. The power spectra from a Digital Elevation Model (30 meter sample spacing) of the Sierra Nevada Batholith and from Seabeam center beam depths (425 meter sample spacing) along a flowline in the South Atlantic are curved. Straight sections in the spectra can be identified but the slopes of those sections are strongly dependent upon the particulars of the data analysis. Fractal geometry must be used with caution in the discussion of topographic data sets.

Key words: Topography, fractal.

Introduction

Many phenomena in nature possess characteristics which appear to be independent of the length scale of observation; for example, the cross-section of a joint surface may be indistinguishable from that of an oceanic ridge if no relative length scales are provided (Figure 1). This requires one either to specify a single length scale or to generalize over all scale lengths when discussing such scale dependent features as topography. Topographic profiles represent the complex interaction of many physical processes which operate over a wide range of length scales; thus in the study of topography one would hope to be able to generalize over many spatial scales. Fractal geometry (MANDELBROT, 1977) provides a framework for such a generalization. Central to fractal geometry is the single parameter fractal dimension (D) which characterizes the scale invariance of a data set. The notion of fractal dimension was first introduced with the question, "How long is the coast of Britain?" (MANDELBROT, 1967). Consider the length as measured from a 1:5,000,000 map, from a 1:250,000 map and the length measured with a meter stick along the shore. The total length will increase as the unit length decreases and the relationship

¹ Lamont-Doherty Geological Observatory, and Department of Geological Sciences, Columbia University, Palisades, NY, U.S.A.

between the total length and the unit length can be quantified as

$$L(r) \propto r^{1-D} \quad (1)$$

where L is the total length measured with a ruler of unit length r and D is the fractal dimension of the data. The coastline is a fractal set whose fractal dimension describes the persistency of its detail over changes in spatial scale. For the west coast of Britain, $D = 1.2$ is commonly accepted.

There are two classes of algorithms which are commonly used to estimate the fractal dimension of data sets with Euclidean dimension (E) 2; they are the ruler and spectral methods. Ruler algorithms, in addition to being extremely sensitive to the details of the particular algorithm, yield meaningful results only when the data under consideration are strictly self-similar. Topographic profiles are self-affine; therefore ruler method estimates of their fractal dimension will not be easily interpreted (MANDELBROT, 1986). (Figure 1 demonstrates the self-affinity of topography. The similarity in the appearance of the two profiles is the result of different vertical exaggerations. In order to preserve the statistics of the profiles after a change in scale, it was necessary to scale the vertical and horizontal axes in different proportions. If the data were self-similar, the statistics would be preserved by scaling the axes in the same proportion.)

The spectral method is the second commonly used algorithm and can be applied to any single valued data set and has no restrictions regarding similarities. The spectral method relates the change in power with frequency of the power spectra of a data set to that set's fractal dimension by

$$D = \frac{s + 5}{2} \quad (2)$$

where s is the slope of a straight line which characterizes the spectra in log-log space (MANDELBROT and VAN NESS, 1968). In this study (and in general (BROWN and SCHOLZ, 1986; Fox and HAYES, 1985)), s is the slope of a line which satisfies a least squares (L2 norm) condition.

This study evaluates the appropriateness of straight lines as characterizations of the power spectra of topographic data sets and addresses the appropriateness of fractal geometry as a quantifier of topography. If lines are adequate descriptions of topographic power spectra then the slope and intercept of those lines will be useful in the quantification of topographic data sets. If the slopes of those lines fall within the range $-3 \leq s < -1$ then the vocabulary and concepts of fractal geometry will be meaningful when applied to topographic data sets.

History

BELL (1975) noted that power spectra of seafloor bathymetry tended to exhibit an "inverse square law dependence on wavenumber." He related this observation to

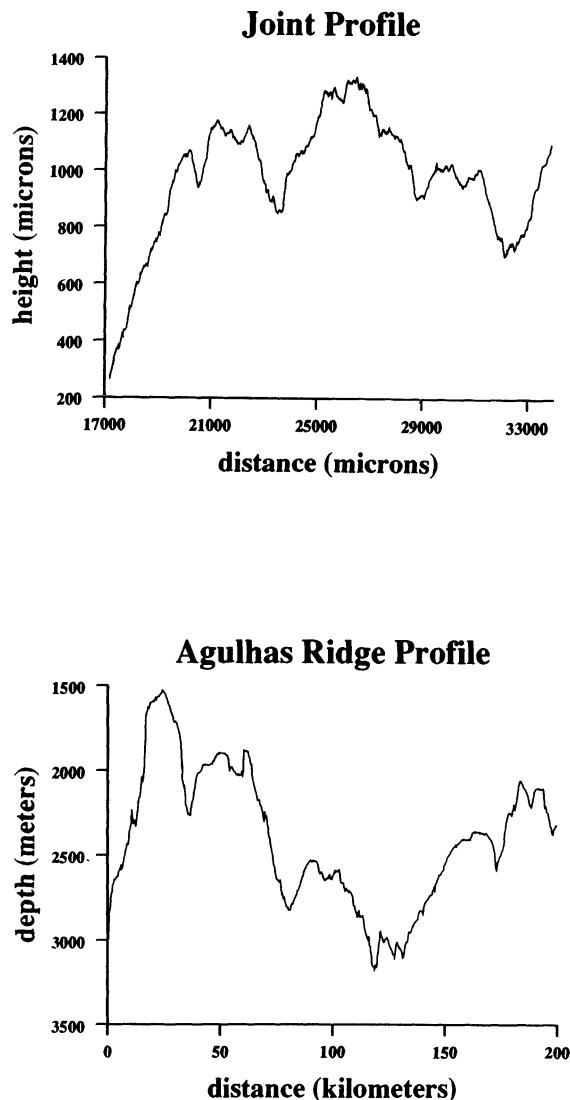


Figure 1

The similarity of these two profiles demonstrates the scale invariance of topography. The Joint Profile is from a joint in the basalt of the Palisades Sill. The Agulhas Ridge Profile is Seabeam bathymetry collected in the South Atlantic on *Robert. D. Conrad* cruise 2710. The Joint Profile is plotted with a vertical exaggeration of 10:1, while the Agulhas Ridge Profile is plotted with a vertical exaggeration of 100:1. The necessity of scaling the vertical and horizontal axes in different proportions to preserve the character of the profiles is a demonstration of the self-affinity of topography. If topography was self-similar, the profiles would "look the same" when plotted at the same vertical exaggeration.

a random distribution of features and an equipartition of the energy of formation over all spatial scales.

SAYLES and THOMAS (1978) compiled topographic data from a variety of sources which spanned nearly 8 decades of spatial scale. They assumed that the increments of the elevations (Δh_{12} and Δh_{23}) were uncorrelated and showed that their data were consistent with a law of the form

$$G(\omega) = 2\pi k \omega^{-2} \quad (3)$$

where $G(\omega)$ is power spectral density, ω is angular frequency, and k is a constant which Sayles and Thomas dubbed *topothesy*. They assert that the value of k uniquely defines the statistical properties of a topographic profile (they also imply that for all topography $D = 1.5$, equation (2)). A corollary to the relationship in equation (3) is that samples of finite length will never completely represent the statistical properties (e.g., variance) of a surface; natural topographic surfaces are nonstationary (SAYLES and THOMAS, 1978).

BERRY and HANNAY (1978) objected to the assumption by SAYLES and THOMAS (1978) that increments of topography are necessarily uncorrelated. They note that fractal profiles will have spectra of the form

$$G(\omega) = 2\pi k \omega^{-\alpha} \quad (4)$$

where $1 < \alpha \leq 3$ ($2 > D \geq 1$) and α is related to the degree of correlation of the increments (Voss, 1983). They note that the relationship derived by SAYLES and THOMAS is too strict; when individual surfaces from the set compiled by SAYLES and THOMAS are considered, the best fit values of α fall in a range from 1.07 to 3.03 (roughly consistent with theoretically acceptable values of D). The discussion between SAYLES and THOMAS (1978) and BERRY and HANNAY (1978) boils to an argument over the importance of the slope versus the importance of the intercept of a line which best describes the power spectra in log-log space.

FOX and HAYES (1985) avoided the language of fractal geometry and used both the slope and intercept of a line fit to the amplitude spectra of seafloor topography to characterize those sets; the slope (analogous to D) characterizes the relationship between spatial scales and the intercept (analogous to k , equation (3)) is a normalized amplitude for the profile. BROWN and SCHOLZ (1986) used power spectra and D to describe naturally occurring rock surfaces. They argue that their spectra exhibit several "straight line" sections and suggest that D may vary somewhat with spatial scale.

Data and Processing

Two digital topographic data sets are considered in this study; one is from the South Atlantic seafloor and the other is from the North American continent (Figure 2).

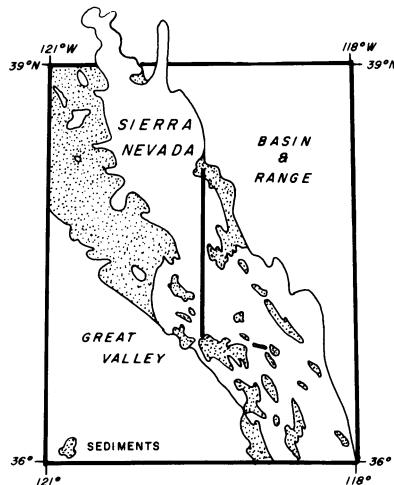
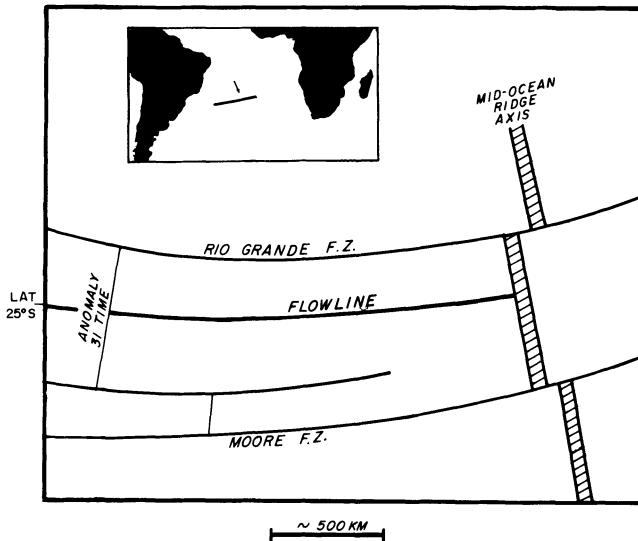


Figure 2

Location maps for the data presented in this study. The seafloor data follow a flowline from the mid-ocean ridge to anomaly 31 in the western basin of the South Atlantic. (A flowline is the hypothetical path which a piece of oceanic crust follows as it moves away from the ridge crest; thus the material which forms this profile was all formed at the same location on the mid-ocean ridge.) The continental data follow north-south Universal Transverse Mercator grid lines in the Sierra Nevada Batholith of the western United States.

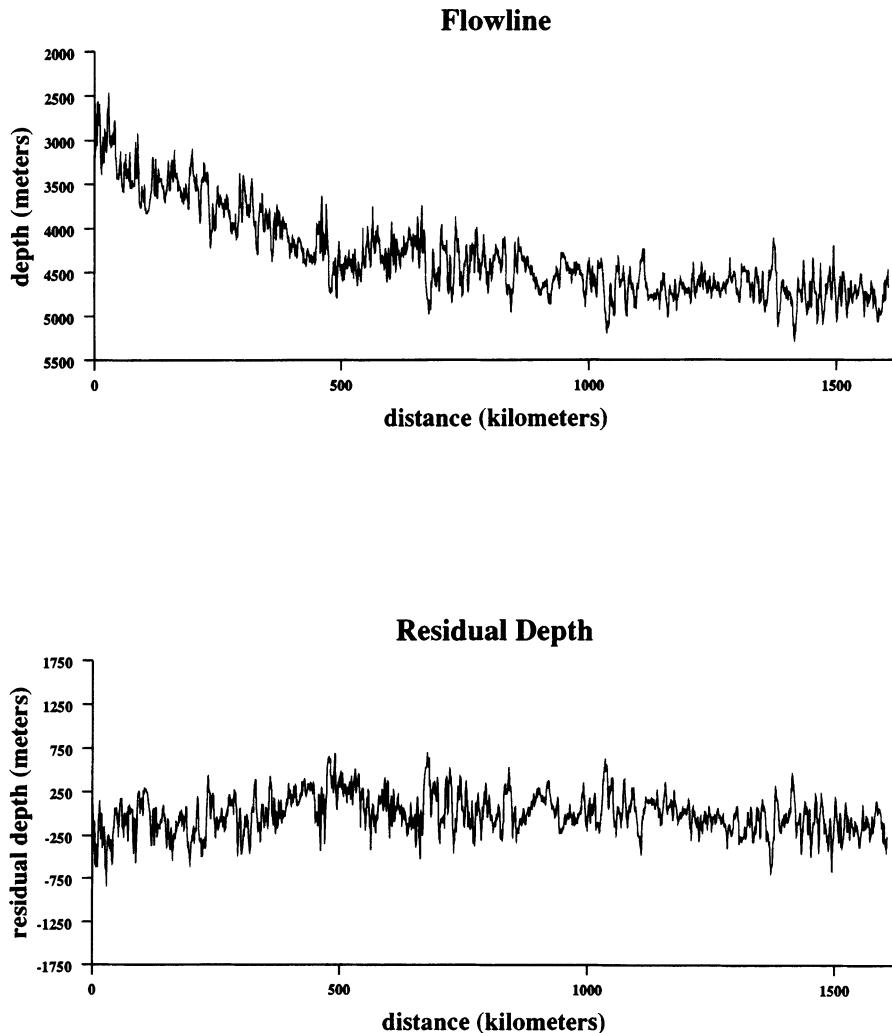


Figure 3

The seafloor data analyzed in this study. The location of these data is shown in Figure 2. The data are from the center beam of a Seabeam data set collected on *Robert D. Conrad* cruise 2711. The Residual Depth data is used in all of the calculations and is a result of removing the square root of age dependence of depth. This is done by assuming constant spreading rate and removing the best fitting square root of distance vs. depth line.

The seafloor data (Figure 3) are Seabeam center beam depths from a flowline which runs from the ridge crest (on the east) to anomaly 31 (on the west) between the Moore and Rio Grande fracture zones. (A flowline is the hypothetical path which a piece of oceanic crust follows as it moves away from the ridge crest; thus the material which forms the seafloor profile under consideration was all formed at

the same location on the mid-ocean ridge.) These data were collected on *Robert D. Conrad* cruise RC2711.

Adjacent Seabeam depth observations are not independent because the footprint of the beam (~ 200 m) is greater than the sampling interval (~ 40 m); furthermore, the data are not evenly spaced because they are collected with reference to a clock on a platform which moves with nonuniform velocity. The nonindependence of adjacent samples requires that the original data set be resampled. Three resampling schemes have been used and their results compared:

- 1) A 10 point *moving average* yields a data set with a sampling interval standard deviation of 66.65 m. This data set contains no original observations and is a smoothed version of the original data.
- 2) *Decimation* (carried out by choosing the next point which exceeds some minimum distance from the current point) yields a sub-set of the original observations with a sampling interval standard deviation of 44.14 m. The minimum distance used here is 425 m.
- 3) *Linear interpolation* yields a data set which is exactly evenly spaced, but as with the moving average, will contain no original observations. Linear interpolation (to a sample spacing of 425 m) was carried out on the original data, the averaged data and the decimated data.

The square root of age dependence of depth (PARSONS and SCLATER, 1977) has been removed from the seafloor data by assuming that spreading rate has been constant and removing the best fitting square root of distance versus depth line. The residual depth data have been used in all of the calculations which follow.

The continental data are from a Digital Elevation Model (DEM) for the Sierra Nevada Batholith (Figure 4). These data have been bilinearly interpolated from original air-photo interpretations to a 30 meter Universal Transverse Mercator grid. Four profiles 120 meters apart were extracted from the larger data set.

The spectra in this study are all computed using a FFT (this assumes that unevenness in the sampling interval of the decimated and averaged data sets is unimportant). In all instances the mean is removed and a Hanning taper is applied to the input data set. The seafloor profile is broken into 5 segments and a smoothed spectra is produced by ensemble averaging the pieces. The continental spectra are smoothed by ensemble averaging the spectra of adjacent profiles.

After computing the spectra, a line can be fit to the spectral estimates in log-log space. Errors associated with slope of that line are 95% confidence limits based on scatter in the spectral estimates. Errors associated with the spectral estimates themselves are constant in log-log space and will contribute only to the *error* associated with slope estimates; therefore errors in the spectral estimates are ignored. The computed line was subtracted from the original power spectra and the residuals were inspected. If a line is an adequate description of the spectra, the residuals will lie about zero and have no structure.

Sierra Nevada North

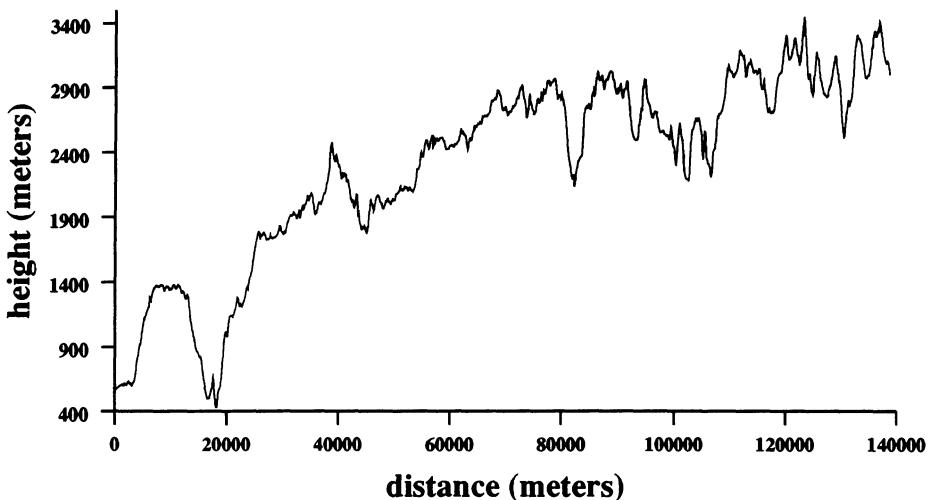


Figure 4

An example of the continental data analyzed in this study. The location of these data is shown in Figure 2. These data are from a Digital Elevation Model and are evenly spaced on a 30 meter Universal Transverse Mercator grid.

Observations and Discussion

Topographic spectra are poorly represented by straight lines. When considered over the entire frequency band sampled, all of the seafloor spectra show striking curvature (Figure 5). It can be argued that the seafloor spectra consist of a straight section in the wavelengths shorter than about 10 kilometers and a section whose slope tends to flatten as wavelength increases above several 10's of kilometers. The Sierra Nevada data also fail to be straight over their entire band width; however when the data are decimated straight lines become more appropriate (Figure 6). This is a result of the fact that decimation reduces the band width sampled. As the band width is reduced the curvature of the spectra becomes less apparent and straight lines become more appropriate.

The fractal dimensions (slopes) implied by the straight sections are a function of the resampling technique. This is clear from inspection of Table 1. (For the purposes of comparison, the range from 10 km to 1 km is deemed straight and is implied whenever a fractal dimension for the seafloor is stated.) The D 's for the interpolated and decimated sets are equivalent (approximately 1.3) at the 95% confidence level; however, for the averaged data D is significantly lower (approximately 1.1). This reflects the smoothing effect of a moving average. It is important to note that interpolation can also act as a strong smoothing operation (Figure 7). Interpolation of the decimated (or averaged) data sets yields a set which is significantly smoother

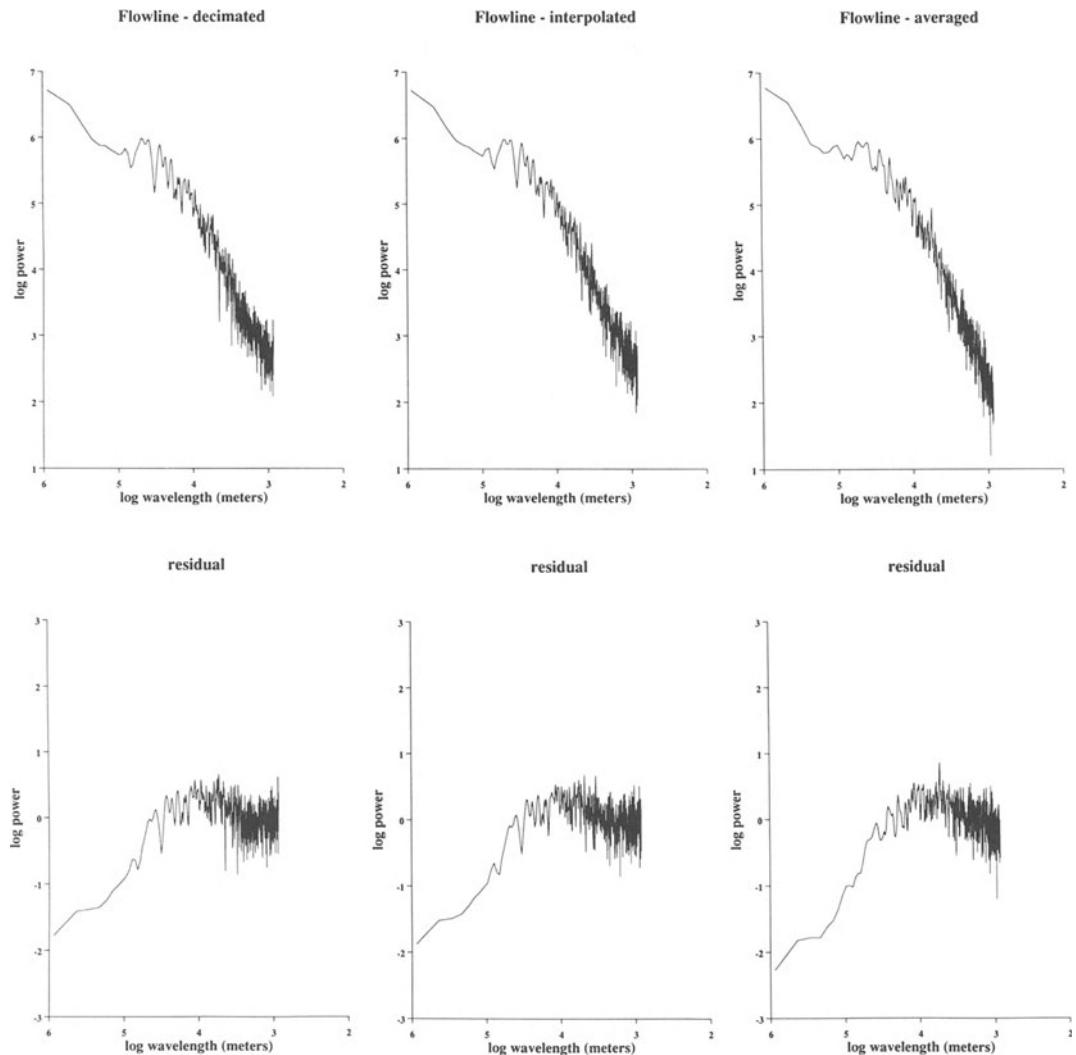


Figure 5

The seafloor spectra. The residuals are the result of subtracting the best fitting line to the entire spectra. Further details on the processing are presented in "Data and Processing" section of the text.

than the set produced by interpolating the original data. This is manifested by fractal dimensions which are close to 1 for evenly spaced resamplings of the decimated and averaged data sets and approximately 1.3 for an evenly spaced resampling of the original data (Table 1). Fractal dimensions for the Sierra Nevada data demonstrate a dependence on the decimation interval; D increases from less than 1 to nearly 1.5 as the data are further decimated (Table 2). This further

Table 1

FLOWLINE STRAIGHT SECTIONS (10 km to 1 km wavelengths)		
Method	D	95% confidence
decimate	1.325	0.032
interpolate	1.297	0.030
average	1.133	0.030
decimate then interpolate	1.040	0.030
average then interpolate	0.919	0.031

Table 2

SIERRA NEVADA (entire band width)		
Data	D	95% confidence
All data	0.835	0.013
decimated by 5	1.208	0.034
decimated by 10	1.363	0.052
decimated by 15	1.471	0.053

illustrates the curved nature of the Sierra Nevada spectra. As noted above, as the data are decimated the band width represented narrows; it also moves to longer wavelengths. It can be seen in the spectra of all of the data (Figure 6) that slope (D) shallows (increases) with increasing wavelength.

Conclusion

The appeal of fractal geometry as a statistical model for topography is its apparent simplicity. Fractal geometry offers D as a single parameter which describes roughness and the dependence of roughness on scale. Using fractal geometry one can extrapolate from properties observed at one scale to the properties of a scale which has not been observed. A topographic profile can be completely described by D and an amplitude parameter (FOX and HAYES, 1985).

Balancing this simplicity are necessary conditions which must be satisfied if the profile is to be adequately described by fractal geometry. (It should be noted that the conditions which follow are necessary but not necessarily sufficient.) The very least of these is that the power density spectrum be linear in log-log space. The power spectrum of a quintessential fractal will be linear over all frequencies from zero to infinity. It is unreasonable to be so rigorous with the spectra of natural

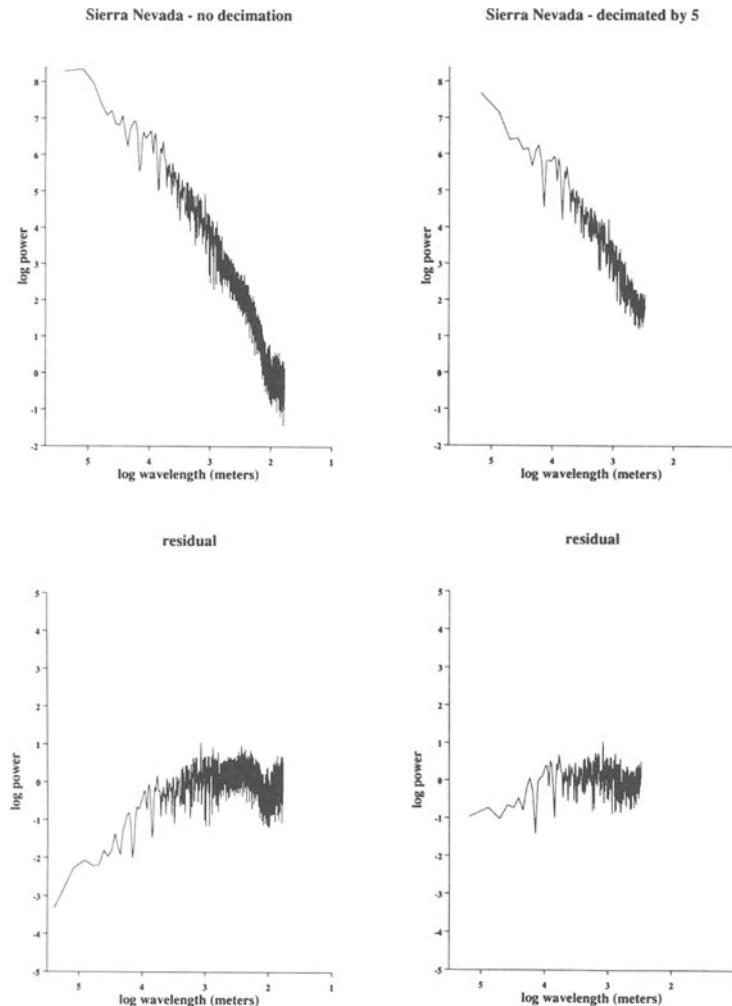


Fig. 6(a)

surfaces; however a fractal interpretation of a data set does imply fairly wide band scale invariance. Coupled with linearity of the spectrum as another necessary condition is a theoretically acceptable range of slopes (fractal dimensions); the slope of the best fitting line through the spectrum in log-log space must fall between -3 and -1 ($1 \leq D < 2$). Departures from these conditions on the power spectrum are inconsistent with a fractal interpretation of the spectrum.

Considered over their entire frequency range, the spectra of this study exhibit striking departures from the linearity required by fractal geometry. The spectra

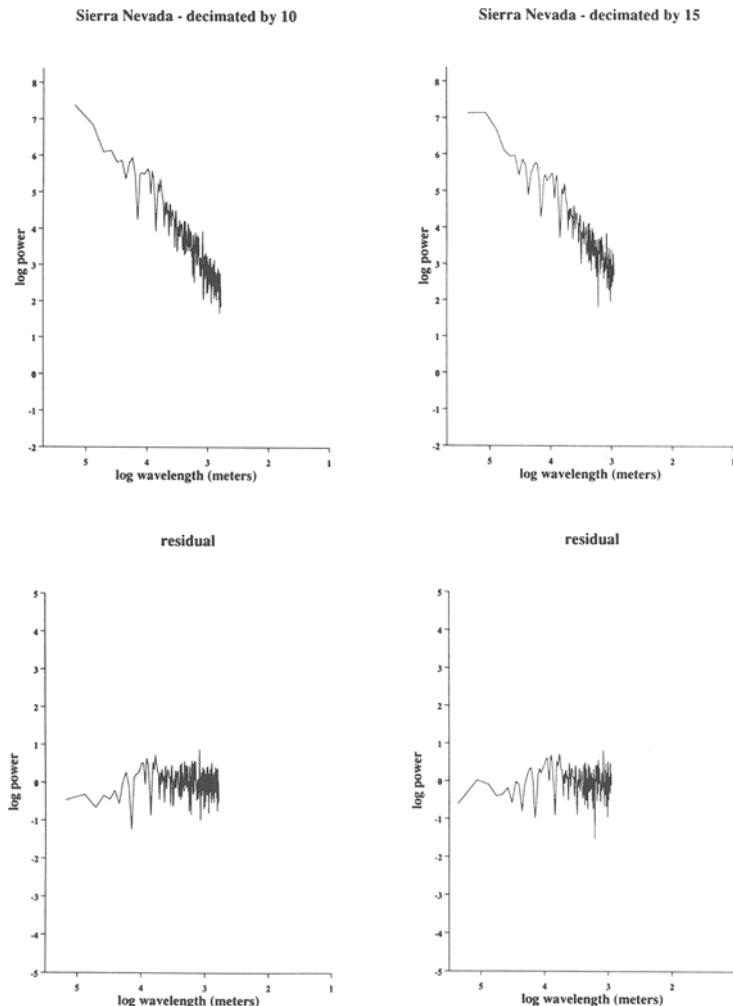


Fig. 6(b)

Figure 6

The continental spectra. The residuals are the result of subtracting the best fitting line to the entire spectra. Further details on the processing are presented in "Data and Processing" section of the text.

considered as a whole are curved which implies that the profiles are not fractal over all of the represented spatial scales. The strongest curvatures are in the neighborhood of 10 km wavelengths in the seafloor data. Sections of the Sierra Nevada spectra which are acceptably straight yield slopes which are a function of the band width considered; in a strict sense, this is not acceptable under the tenets of fractal geometry.

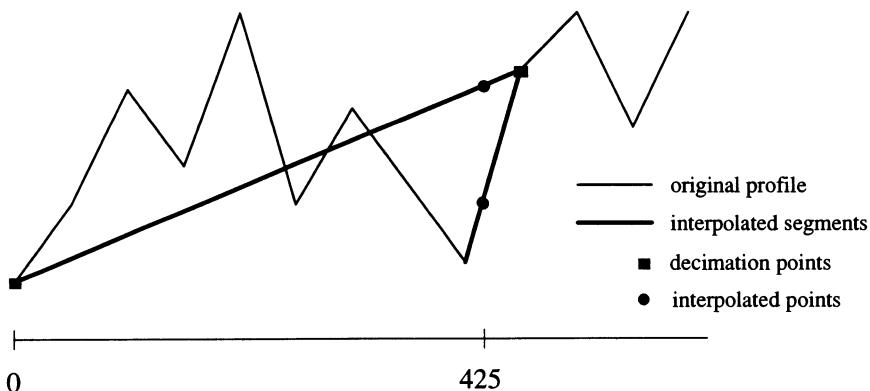


Figure 7

The interpolant of the decimated data (long bold segment) will yield a set which is more smooth than that produced by the interpolant of the original data (short bold segment) because it is less sensitive to small-scale variations in the observations.

It has been argued that the section of the seafloor spectra between 10 and 1 km wavelengths satisfies the requirements of fractal geometry; however the slope of this section is strongly dependent upon the particulars of the data analysis. If the observations are used and the power spectra is calculated directly, then values of D in the neighborhood of 1.4 are gained and one would conclude that the data are fractal over these scales. If the data are linearly interpolated (a smoothing operation with respect to a fractal), then the slope of that section yields a fractal dimension of 1, and one might conclude that the data set is differentiable rather than fractal. The conclusion drawn concerning the fractal nature of that band is entirely dependent on the data analysis techniques.

The use of fractal geometry as a quantifier of topography must be qualified by two considerations 1) Scale must be specified in the form of a band width under consideration; and 2) the particulars of the data analysis techniques must be explicitly stated.

Acknowledgements

Dennis Hayes has provided support throughout. Discussions with Alberto Malinverno have helped to smooth some of the rough spots and have contributed greatly to the results presented here. Steve Cande and Carol Raymond have been especially encouraging in my seafloor efforts. Thanks go out to the people of the *Robert D. Conrad* who shared their home with me for cruises 2710 and 2711. Chris Barton provided the Sierra Nevada data. This work has been supported by ONR contract number 14-84-C-9132.

Robert D. Conrad who shared their home with me for cruises 2710 and 2711. Chris Barton provided the Sierra Nevada data. This work has been supported by ONR contract number 14-84-C-9132.

REFERENCES

- AVILES, C. A., SCHOLZ, C. H., and BOATWRIGHT, J. (1987), *Fractal Analysis Applied to Characteristic Segments of the San Andreas Fault*, JGR 92, 331-334.
- BELL, T. H. (1975), *Statistical Features of Sea-floor Topography*, Deep-Sea Research 22, 883-892.
- BERRY, M. V., and HANNAY, T. R. (1978), *Topography of Random Surfaces*, Nature 273, 573.
- BROWN, S. R., and SCHOLZ, C. H. (1986), *Broad Bandwidth Study of the Topography of Natural Rock Surfaces*, JGR 90, 12,575-12,582.
- FOX, C. G., and HAYES, D. E. (1985), *Quantitive Methods for Analyzing the Roughness of the Sea-floor*, Rev. Geophy. and Space Phys. 23, 1-48.
- MANDELBROT, B. B. (1967), *How Long is the Coast of Britain? Statistical Self-similarity and Fractional Dimension*, Science 155, 636-638.
- MANDELBROT, B. B., *Fractals: Form, Chance and Dimension* (W. H. Freeman and Co., San Francisco, CA 1977).
- MANDELBROT, B. B., *The Fractal Geometry of Nature* (W. H. Freeman and Co., San Francisco CA 1982) pp. 460.
- MANDELBROT, B. B., *Self-affine fractal sets, I-III*, in (Pietronero, L. and Tossatti, E., eds.) *Fractals in Physics* (Elsevier Science Publishers B. V. 1986) pp. 3-28.
- MANDELBROT, B. B., and VAN NESS, J. W. (1968), *Fractional Brownian Motions, Fractional Noises and Applicatons*, SIAM Review 10, 422-437.
- PARSONS, B., and SCLATER, J. G., (1977), *An Analysis of the Variation of Ocean-floor Bathymetry and Heat Flow with Age*, JGR 82, 803-827.
- SAYLES, R. S., and THOMAS, T. R. (1978), *Surface Topography as a Nonstationary Random Process*, Nature 271, 431-434.
- VOSS, R. F., *Random fractals: characterization and measurement*, in (Pynn, R. and Skjeltorp, A., eds.) *Scaling Phenomena in Disordered Systems* (Plenum Press, New York 1983), 580 pp.

(Received May 18, 1987, revised/accepted November 16, 1987)

Fractals, Fault-Gouge, and Friction

CHARLES G. SAMMIS¹ and RONALD L. BIEGEL¹

Abstract—The particle-size distribution of a natural fault-gouge has been determined over the range from 5 μm to 40 cm. The gouge is self-similar over the range from 5 μm to 1 cm having a fractal dimension of 2.60 ± 0.11 . The lower and upper fractal limits were also determined. The lower fractal limit occurs at a dimension of about 1–10 μm where mineral cleavage and intergranular porosity dominate the cataclasis. The upper fractal limit occurs at particle sizes on the order of 1 cm where the scaled particle density decreases abruptly by a factor of about three. By analogy to soil-mechanics shear box tests and laboratory rock friction experiments, it is argued that the upper fractal limit of the gouge determines the characteristic displacement parameter in stick-slip friction models. A characteristic displacement on the order of 1 cm is consistent with estimates based on numerical simulations of faulting.

Key words: Fractals, fault-gouge, friction, faulting, stick-slip, velocity-weakening, self-similarity, fractal dimension.

Introduction

Since the pioneering work of BRACE and BYERLEE (1966), the earthquake mechanism has been closely associated with the stick-slip frictional instability observed in the laboratory. DIETERICH (1978) demonstrated that the stick-slip instability is due to velocity weakening, a phenomenon in which the coefficient of dynamic friction decreases as the sliding velocity increases. He further related velocity weakening to the time dependence of static friction through the concept of a characteristic displacement. The characteristic displacement, d_c , can be observed in laboratory experiments as the sliding distance necessary to achieve the new value of dynamic friction following an instantaneous velocity change. Evidence that the characteristic displacement is a property of the sliding surfaces, which is correlated with the surface roughness, led Dieterich to propose that d_c may be interpreted as the distance necessary to completely change the contacting asperity distribution thus establishing a new contact regime characteristic of the new sliding velocity. OKUBO (1986) observed that a characteristic displacement is also associated with slip

¹ Department of Geological Sciences, University of Southern California, Los Angeles, CA 90089-0741, U.S.A.

weakening instabilities in a large-scale (2 m) laboratory faulting experiment. As in the small-scale velocity change experiments described above, d_c was correlated with surface roughness.

In order to scale frictional constitutive relations from the laboratory to the field, it is necessary to establish the characteristic displacement appropriate for natural faults. Toward this goal, POWER *et al.* (1987) have measured the surface roughness of natural fault surfaces at scales ranging from 10^{-5} to 1 m. Rather than discovering a dominant wavelength which might be identified with the characteristic displacement, they found that the surface roughness is scale-independent over the range studied. Further, they point out that their distribution is consistent with that found by SCHOLZ and AVILES (1986) who established scale independence for the San Andreas Fault over the range 100 m to 100 km. The fractal nature of fault surfaces found in these studies implies that there is no one scale-length which might be associated with the characteristic displacement over the entire range of 10^{-5} to 10^5 m.

If the surface roughness of natural faults does not contain a characteristic scale length which might be associated with the characteristic displacement, what, if anything, does? Because most natural faults contain a layer of gouge between their sliding surfaces, the particle distribution within the gouge is a likely candidate. DIETERICH (1981) studied the frictional behavior of laboratory surfaces separated by a layer of artificial gouge. He found that d_c was correlated with both the scale of surface roughness and with the particle size of the gouge. MORROW and BYERLEE (1986) presented evidence that velocity effects on the coefficient of friction are related to the pore volume in the gouge. They proposed that a critical void ratio corresponds to each sliding velocity and that, following a step change in velocity, the gouge either dilates or compacts to the new value. From this soil mechanics viewpoint, d_c is the sliding distance necessary to establish the new state of dilatancy; it thus seems reasonable that d_c should be related to the particle-size distribution within the gouge layer.

Soil mechanics shear box experiments are a source of data on the transient stress response of granular materials to a change of strain rate. Such experiments differ from the standard friction experiments in that the normal stress is low and comminution of the particles is minimal. Figure 1 shows the results of a shear box experiment on dry sand from FEDA (1982). According to MORROW and BYERLEE'S (1986) hypothesis, the open circles are analogous to a velocity increase in a sliding friction experiment; the sand dilates to an equilibrium porosity appropriate to the new velocity while the shear stress first increases and then decreases to the new equilibrium value. From this viewpoint then, the solid data points are analogous to a velocity decrease; the sand contracts as the stress adjusts to the new velocity. Note that in both cases the characteristic displacement required to reach equilibrium is on the order of 3–4 mm, about twice the largest grain diameter of 2 mm.

It is possible that stick-slip on initially clean surfaces may also require a layer of

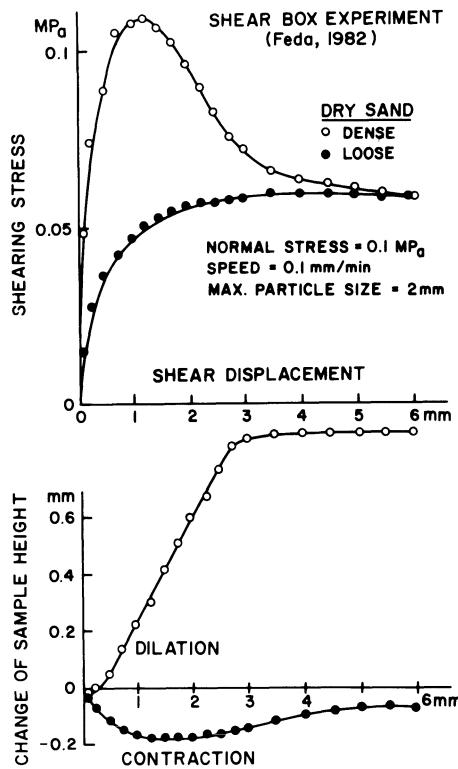


Figure 1

Shear box test of dense and loose samples of dry sand (from FEDA, 1982). The maximum grain diameter was 2 mm. The normal stress was 0.1 MPa and the speed was 0.1 mm/min. Note that the shear displacement required to establish the equilibrium stress and porosity is slightly larger than the maximum particle diameter. Other shear box data is summarized in Table 2 and Figure 3.

gouge (DIETERICH, 1978). TULLIS and WEEKS (1986) measured the friction between initially clean surfaces of Westerly granite in a rotary shear apparatus. The development of a layer of gouge 100 to 200 μm thick between the rock rings led them to conclude "...the experiments represent the behavior of a gouge layer produced during sliding rather than the behavior of bare rock."

SAMMIS *et al.* (1987) measured the particle-size distribution of a natural gouge. We found the distribution to be scale-independent over the range 5 μm to 0.4 mm. However, unlike surface roughness which can be (and apparently is) self-similar at all scale lengths up to the total fault length, gouge can only be self-similar at scales up to the fault width. In this paper we extend the measurements of the particle-size distribution in the Lopez Canyon fault gouge from an upper limit of 0.4 mm reported by SAMMIS *et al.* (1987) to 400 mm. We find a breakdown in self-similar behavior at a particle size of about 1 cm which we identify as the "upper fractal limit." We speculate that this upper fractal limit is a natural

candidate for the “asperity size” which determines the scale length d_c in stick-slip friction on natural faults.

Measuring the Fractal Limits of a Natural Fault Gouge

SAMMIS *et al.* (1987) [hereafter referred to as paper I] measured the particle distribution of a natural fault gouge over the range 5 μm to 4 mm. A gouge sample was retrieved intact from the Lopez Fault in the San Gabriel Mountains of Southern California, and circular thin sections (about 2.5 cm in diameter and 100 μm thick) were photographed at magnifications of 12.5, 25, 50, 100, 200, 400, 800, and 1600 using optical and electron microscopy. At each magnification, the particles were sorted by diameter into four classes differing in mean diameter by factors of 2, using a transparent overlay of concentric circles. It is important to point out that the area analyzed on each photomosaic was chosen such that it contained no particles larger than the largest circle on the overlay. If portions of larger particles intruded into a field of analysis, their area was subtracted for the total area before particle densities were calculated. This sampling strategy imposes constraints on the test for self-similarity and determination of the fractal dimension discussed below.

A self-similar distribution may be characterized by a fractal dimension, D , which may be calculated as:

$$D = \log[N(n)/A]/\log[1/L(n)] \quad (1)$$

where $N(n)/A$ is the number of particles per unit area, and $L(n)$ is the diameter of particles in the class n . The fractal dimension of gouge from the Lopez Fault was found in paper I to be $1.60 \pm .11$ in two-dimensional cross-section, and thus $2.60 \pm .11$ in an isotropic three-dimensional volume. In 3-D this corresponds to an increase in the number of particles by a factor of six each time the diameter is reduced by a factor of one half. Because our sampling avoided larger particles at each scale, D had to be determined separately for each photograph. The consistency of D at each magnification is evidence for self-similarity across the range.

In paper I, we related the fractal dimension to the comminution process in terms of the relative size of nearest neighbors. For an idealized gouge, $D = 2.58$ characterized a geometry in which no two particles of the same size are nearest neighbors at any scale. Thus, for a comminution process in which the fracture probability is a maximum for nearest neighbors of the same size, any initial particle distribution will evolve toward a fractal distribution having a dimension of 2.58. A simple model is presented in paper I showing that the probability of fracture should be a maximum for a particle with nearest neighbors the same size. We have subsequently found data on the fracture probabilities in binary mixtures of glass beads which support the model (HOFFMANN and SCHONERT, 1971).

It is apparent that self-similarity cannot extend over an infinite range of particle sizes; the scale length at which self-similarity breaks down is called the fractal limit. In paper I, it was observed that the nature of comminution appeared to change at particle diameters on the order of 10 μm . Below this dimension the fracture pattern was dominated by intracrystalline porosity and mineral cleavage. We identify this as the lower fractal limit and note that a similar shift in comminution mechanism has been identified in commercial crushing and grinding processes at about this scale which has been attributed to the absence of Griffith flaws in smaller particles (STEIER and SCHONERT, 1972). No upper fractal limit was observed in paper I.

A second test for self-similarity, which was developed in paper I, will now be used to search for the upper fractal limit of the Lopez Canyon fault gouge. This test quantifies the property of a self-similar distribution that it must "look the same" at each scale. Stated another way, it should be impossible to tell at which magnification a photograph of a self-similar system was taken. If a distribution is scale-independent, then the number of particles in an area scaled to the mean particle dimension, $N_A(n)$, should be a constant, where

$$N_A(n) = N(n) \langle L(n)^2 \rangle / A \quad (2)$$

and

$$\langle L(n)^2 \rangle = \{L^2(n) + L^2(n+1)\}/2 \quad (3)$$

is proportional to the mean particle size in each class n .

Our sampling strategy of avoiding large particles at each scale also imposes constraints on this analysis in that $N_A(n)$ is only constant for the same pair of rings on the overlay when it is applied across the range of magnifications. For the observed fractal dimension of $D = 1.6$, we expect $N_A(n) = 0.25$ for the largest particles in each photograph. (Actually, it should be slightly smaller because of the way we define average area in Equation (3).) For all smaller particles, we expect $N_A(n)$ to be smaller because of the larger particles in the area of analysis. This trend is apparent in Figure 12 of paper I. We can analyze any size class on a photograph as if it were the largest by subtracting the area of all larger particles from the total area. We have made this correction to the data in paper I by treating each class as if it were the largest in the photograph. The resulting $N_A(n)$ is given for each particle size in Table 1 and shown as the closed squares in Figure 2. The error bars represent the 95% confidence limits on $N(n)$ under the assumption that we are sampling a Poissonian spatial distribution (WILSON, 1952, pp. 191–195). The constancy of $N(n)$ for the closed squares is evidence for self-similarity over the range investigated.

To search for the upper fractal limit, we have extended this analysis to larger particle sizes by significantly increasing the area sampled beyond the original 2.5 cm thin sections. This was done in two ways; first by preparing flat circular mouldings (~ 15 cm in diameter) in the field which were brought to the laboratory for counting particles with diameters in the range 1.6 to 25.6 mm, and, second, by clearing a flat surface in the exhumed fault gouge and counting, *in situ*, particles with diameters

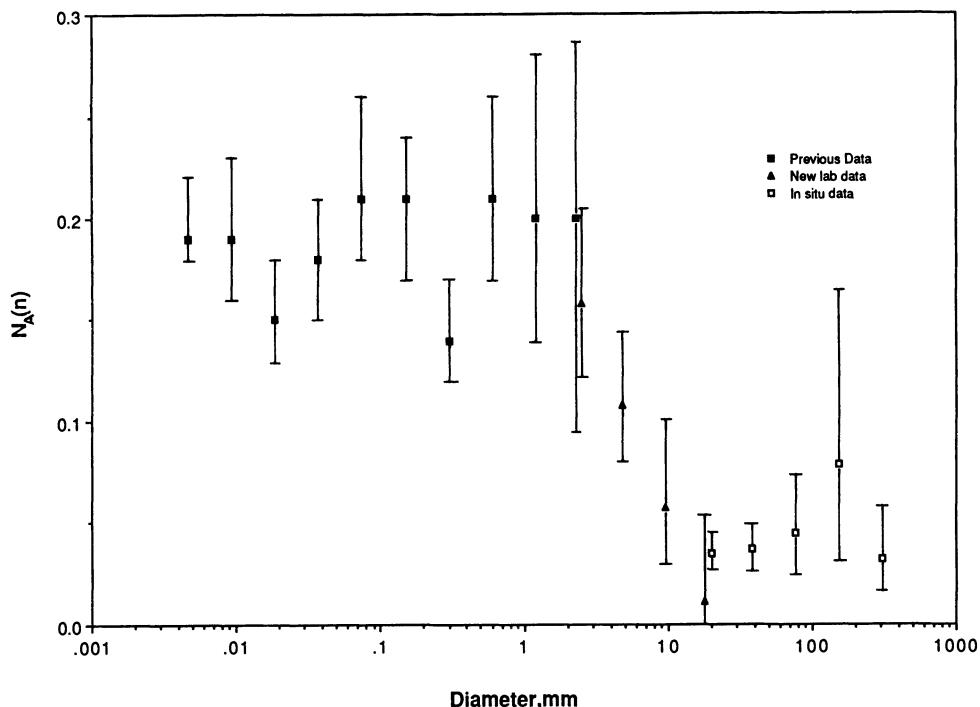


Figure 2

The number of particles in an area scaled to the particle diameter ($N_A(n)$ defined in the text) as a function of the particle diameter. Closed squares are microscopic determinations from SAMMIS *et al.* (1987). Open triangles are microscopic measurements from large mouldings and open squares are *in situ* field determinations. Note the decrease in scaled density at a particle diameter on the order of 1 cm which we identify as the upper fractal limit.

larger than 25.6 mm. By this second method, we extended the distribution to particles with diameters up to ~ 410 mm.

The circular mouldings were made by driving cylindrical moulds into a horizontal surface of the gouge, and then pouring in clear liquid polyurethane. After the polyurethane had soaked into the gouge and hardened, the mouldings were removed to the lab and vacuum impregnated with a low viscosity epoxy. Finally, a flat polished surface was prepared, and particle densities were determined as in Paper I. Results of these counting experiments are given in Table 1.

In order to extend the study to still larger particles, a flat surface with area $\sim 2.6 \text{ m}^2$ was prepared in the face of the gouge. Particles with diameters larger than 25 mm were identified and counted. These results are also given in Table 1. Finally, an area approximately 41 m^2 was marked off on the weathered face of the gouge deposit and inspected for particles with diameters in excess of 200 mm. Because the surface of the fault-zone was flat on the 200 mm scale, we were spared the task of clearing such a large surface. The particle densities for this large area are also

Table 1
Scaled Particle Densities for Lopez Fault Gouge

	Particle Diameter (mm)	Particles counted	A (mm ²)	N _A (n)		Particle Diameter (mm)	Particles counted	A (mm ²)	N _A (n)
F	410–205	8	4.15×10^7	.037(max) .020(mean) .010(min)	OL*	3.2–1.6	7	295	.37 .02 .095
F	205–102	5	2.67×10^6	.10 .049 .020	OL	1.6–.8	21	295	.28 .20 .14
F	102–51.2	11	2.67×10^6	.044 .027 .015	OL	.8–.4	66	295	.26 .21 .17
F	51.2–25.6	36	2.67×10^6	0.29 .022 .016	OL	.4–.2	100	167	.17 .14 .12
F*	25.6–12.8	42	861,250	.025 .020 .015	OL	.2–.1	137	39.3	.24 .21 .17
NL*	25.6–12.8	1	57,710	.033 .007 .0004	OL	.1–.05	152	10.5	.24 .21 .18
NL	12.8–6.4	10	27,623	.063 .037 .019	OL	.05–.025	118	2.5	.21 .18 .15
NL	6.4–3.2	35	13,684	.087 .065 .048	OL	.025–.0125	134	.838	.18 .15 .13
NL*	3.2–1.6	59	4,185	.12 .090 .070	OL	.0125–.00625	134	.163	.23 .19 .16
					OL	.00625–.0031	128	.038	.22 .19 .18

*: overlapping particle diameters

F: field data

NL: data from large samples taken into lab.

OL: data from SAMMIS *et al.* (1987)

included in Table 1. In all cases these new data were analyzed as if they were the largest particles in the section.

The new data from Table 1 are plotted as open symbols in Figure 2 where they may be compared with the microscopic data from paper I. Note that the scaled density decreases significantly for particles larger than about 1 cm. We identify this abrupt decrease in scaled particle density as the upper fractal limit of the gouge.

The Upper Fractal Limit and Stick-Slip Friction

Having established that the upper fractal limit for the Lopez Canyon fault gouge is on the order of 1 cm, we now argue that this upper fractal limit may determine the characteristic displacement in stick-slip friction. We begin with a brief review of the concept of a characteristic displacement in the slip-weakening model. We then compare characteristic displacements which have been determined in laboratory experiments with the size of the largest gouge particles. Finally we review determinations of the characteristic displacement from seismic data, and show that they yield dimensions which are comparable to those we determine from the gouge.

The Characteristic Displacement in Stick-slip Friction

To a first approximation, sliding friction can be described by

$$\tau = \mu_0 \sigma_n \quad (4)$$

where τ is the shear stress, σ_n is the normal stress, and μ_0 is the coefficient of friction which is usually about 0.6 for rock, and is, with the exception of a few clays, insensitive to rock type (BYERLEE, 1978). In the context of the mechanics of earthquake faulting, interest has focused on the observation that under some conditions the sliding motion is smooth, while under other conditions it occurs as a series of unstable jerks. The former has been termed "stable-sliding," and the latter "stick-slip." It was recognised early on that material properties alone do not completely determine stability during sliding, but that stability also depends upon the stiffness of the system applying the loads (BRACE, 1972). In the laboratory, stiffness is determined by the force-displacement characteristics of the loading apparatus; in the field, it is the rock adjacent to the fault. WALSH (1971) has shown that the stiffness of natural faults is probably much less than that of a typical laboratory apparatus. Other key observations which led to a theoretical model for stick-slip displacement were 1) static friction is time-dependent, 2) dynamic friction is velocity-dependent with the coefficient of dynamic steady-state friction decreasing with increasing velocity, and, 3) following a change in sliding velocity, a characteristic sliding distance is required before the new coefficient is attained. DIETERICH (1978) combined these observations into a theory for the stick-slip instability which we briefly discuss.

Time dependency of the static coefficient of rock was first noted by DIETERICH (1970, 1972). For granite, greywacke, quartzite, and sandstone he found a logarithmic relation between the coefficient of friction μ and the time of contact, t , for contact times as long as 10^5 seconds, which he expressed as

$$\mu = \mu_0 + A \log(Bt + 1) \quad (5)$$

where A , B , and μ_0 are constants. A and μ_0 were found to be insensitive to normal stress. Possible mechanisms causing this time dependency were discussed by DIETERICH (1972), SCHOLZ *et al.* (1972), and SCHOLZ and ENGELDER (1976). DIETERICH (1978) concluded that the evidence strongly favored an increase in the area of adhesive contacts due to creep of the point contacts between asperities. This conclusion was subsequently strengthened by the observation that time-dependent friction and velocity weakening are not observed in friction experiments run under ultra-dry conditions (DIETERICH and CONRAD, 1984).

DIETERICH (1978) pointed out that the logarithmic time dependence of static friction in Equation (5) also explains the inverse logarithmic dependence of μ on the slip velocity demonstrated by SCHOLZ and ENGELDER (1976). The key to this link between static and dynamic friction was DIETERICH'S (1978) observation that, following a step change in the sliding velocity, a characteristic sliding distance, d_c is required to establish the new value of friction. Because d_c was found to be independent of the sliding velocity and the normal stress, but to be proportional to the surface roughness, Dieterich concluded that it is a characteristic of the sliding surface. He proposed that d_c is the distance it takes to break the point contacts formed at the old sliding velocity and establish a new set of contacts characteristic of the new sliding velocity. An important implication of this interpretation is that the time of contact, T_c , of any given population of contacts is inversely proportional to the sliding velocity v ,

$$T_c = d_c/v. \quad (6)$$

Hence, Equation (5) may be rewritten as

$$\mu = \mu_0 + A \log(Bd_c/v + 1) \quad (7)$$

which has the observed dependence of μ on $\log(v^{-1})$.

DIETERICH (1978) used this model to calculate the motion of a one-dimensional frictional slider which is driven by a spring having stiffness k , which is itself driven at constant velocity v_1 . He showed that the minimum normal stress required to produce stick-slip motion can be expressed as

$$\sigma_n^* = kd_c/\Delta\mu \quad (8)$$

where

$$\Delta\mu = A[\log(Bt_1 + 1) - \log(Bd_c/v_1 + 1)] \quad (9)$$

and t_1 is the time that the slider remains stationary before the motion begins. He verified experimentally the linear dependence of σ_n^* on the system stiffness k , and further demonstrated that σ_n^* was larger for rougher surfaces—an observation which supports the identification of d_c with surface roughness.

Although subsequent work by DIETERICH (1979), RUINA (1983), and RICE and RUINA (1983) has extended these concepts for stability of spring-slider systems, and experiments by DIETERICH (1979), RUINA (1983), TULLIS and WEEKS (1986), and others have indicated the need for a direct velocity dependence in the state variable formulation, the basic physics of velocity weakening has remained unchanged.

Characteristic Displacements Observed in Laboratory Experiments

DIETERICH (1981) explored the effect of surface roughness and gouge particle size on the parameters of the slip-weakening model, and found that both the roughness and the particle distribution affected d_c . For fine surfaces, he observed slip localization at the gouge interface, while for the rough surface, localization occurred at low-angle "Riedel shears" within the gouge. We suggest that

Table 2
Characteristic Displacements in Laboratory Experiments

		Maximum Particle Diameter (mm)	Characteristic Displacement (mm)
<i>Shear box tests</i>			
FEDA (1982)	Dry Zbraslav sand	2	3 ± 1
HOUGH (1969)	Dry fine-to-coarse sand	2	10 ± 5
TAYLOR (1948)	Dry Ottawa standard sand	0.84	4 ± 2
<i>Double shear friction tests</i>			
DIETERICH (1979)	Granite	.060* .012* .0025*	.017 \pm .007 .010 \pm .005 .002 \pm .001
DIETERICH (1981)	Gouge	.250 .250 .085	.212 \pm .43 ⁺ .151 \pm .047 ⁺ .120 \pm .062 ⁺
<i>Rotary friction tests</i>			
TULLIS and WEEKS (1986)	Granite	.050*	.90 \pm .060
<i>Large block fault tests</i>			
OKUBO (1986)	Granite	.090* .012*	.027 \pm .013 .0055 \pm .0045

*Amplitude of surface polish damage (see Figure 4 and text)

⁺Mean and standard deviation of data

localization at the boundary of the fine surfaces may be a consequence of poor coupling due to the difference in scale between the surface ($1 \mu\text{m}$) and the gouge particles ($85\text{--}250 \mu\text{m}$). In the natural faults, material is continually added to the gouge by plucking bits from the wall-rock (HUNDLEY-GOFF and MOODY, 1980; MOODY and HUNDLEY-GOFF, 1980; ROBERTSON, 1982). Although the wear processes by which gouge is formed are not yet well understood (see e.g., SCHOLZ, 1987), it seems unlikely that the fault surface will evolve to be significantly smoother than the largest particles which are being added to the gouge. Table 2 and Figure 3 summarize the d_c values measured by DIETERICH for the case of rough surfaces. Note that d_c is of the same order as the largest particle size.

Table 2 and Figure 3 also contain data from experiments on initially bare surfaces. For these experiments, we hypothesize that a layer of gouge is formed by plucking from the wall rock during the preseismic slip in large-scale experiments, or during conditioning displacements in the smaller scale experiments (see e.g., DIETERICH, 1978; TULLIS and WEEKS, 1986). We estimate the maximum particle diameter in this case from the amplitude of surface damage, since it seems reasonable that the scale of plucked particles should be set by the scale of the surface roughness. The amplitude of surface damage is arbitrarily defined as $1/2$ the depth of the deepest grooves produced by the various grits of abrasive compound.

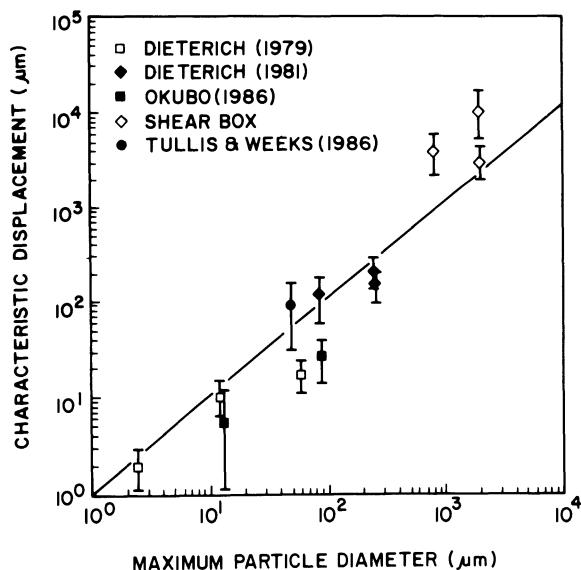


Figure 3

Characteristic displacement as a function of the largest particle size in the gouge. For initially gouge-free surfaces, the largest particle is assumed to be set by the amplitude of surface polishing damage (see Figure 4).

Surface profiles used in this analysis are from OKUBO (1986) and TULLIS (personal communication); the results are shown in Figure 4. TULLIS and WEEKS (1986) fit their data for Westerly granite to a two state variable friction law. We have used their longer decay constant L_2 in this analysis since it is more comparable with other graphical estimates of d_c . It is possible that their shorter decay constant $L_1 = 3-8 \mu\text{m}$ may reflect the size of particles at the lower fractal limit. As discussed above, this is about the size at which quartz deformation changes from fracture to crystalline slip (STEIER and SCHONERT, 1972). Given the above assumptions relating surface roughness to gouge particle size, d_c also appears to be comparable to the largest gouge particles for the case of initially barren surfaces.

Finally, the shear box experiment shown in Figure 1 and other shear box data from Table 2 are also included in Figure 3. For the shear box, d_c is taken to be the displacement required to establish equilibrium porosity. It must be pointed out that these shear box data represent the initial loading transient for a velocity jump from zero to a fixed value. In this sense they are not strictly comparable to d_c in friction experiments which are defined by the transient response to a step change in velocity

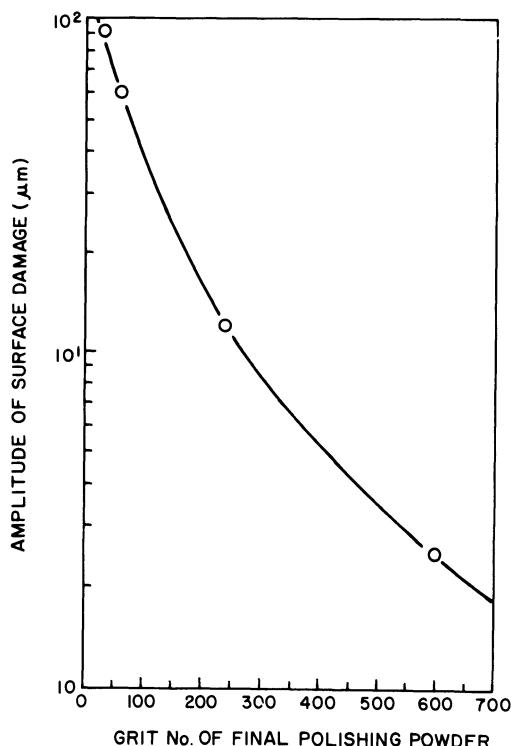


Figure 4

Amplitude of surface damage as a function of the grit used in the final polishing. The amplitude of surface damage was measured on surface profiles from OKUBO (1986) and TULLIS (personal communication).

from an approximate steady state. However, the observation that the length of such transients does not depend upon the initial packing (see Figure 1) argues that one might get the same result in a standard velocity change experiment. Such velocity change experiments on a 4 mm thick layer of sand (MARONE *et al.*, 1987) have found that d_c is comparable to the particle size (SCHOLZ, private communication). The fact that initial transients in some friction experiments are longer may reflect a more complex path to steady state involving comminution and wear.

A significant source of uncertainty in this analysis is our assumption that the particle size remains essentially unchanged throughout the experiment. While this is probably true for the shear box experiments at low normal stress, significant comminution is known to occur during frictional runs. The evolution of the particle distribution depends upon the relative rates of comminution within the gouge and resupply of large particles from the wall rock. BIEGEL *et al.* (1987) have shown that d_c decreases with increasing displacement in double-shear friction experiments. Our interpretation of their results is that comminution is more efficient than resupply in these experiments and the gouge becomes finer with increasing displacement. It is interesting that d_c approaches about 10 μm at large displacements, regardless of the starting particle distribution in the gouge or surface roughness.

Estimates of Characteristic Displacements for Natural Faults

There have been several attempts to determine a characteristic distance from seismic data. Many of these (IDA, 1973; AKI, 1979; PAPAGEORGIOU and AKI, 1983) are based on a model for the slip necessary to propagate a mode II crack, and are thus not directly comparable to d_c discussed above. More relevant is the fault simulation by CAO and AKI (1984/85) which was based on a fifty block, one-dimensional frictional slider model originally developed by BURRIDGE and KNOPOFF (1967). Using STUART's (1979) displacement softening model for friction, they simulated seismicity patterns. Only when the critical displacement associated with slip-weakening was about 10 cm did the pattern show seismic quiescence before major events. For smaller values, seismicity was well simulated but quiescence did not occur. In a later study, CAO and AKI (1987) noted that the seismic quiescence is only observed above a threshold magnitude between 2.5 and 3.5; smaller events actually increased in number before the main event. They proposed that the critical displacement changes abruptly from values in excess of 1 cm above magnitude 3 to smaller values below. The only fault simulation which used the rate and state dependent friction laws discussed above was made by TSE and RICE (1986). They modeled conditions on the San Andreas Fault using a temperature dependent state variable friction law of the type discussed above and found that earthquake instabilities were possible only for $d_c < 8$ cm. These estimates of d_c for natural faults are thus of the same order as the upper fractal limit of 1 cm which we observed for the Lopez Canyon fault gouge.

Discussion

Early work on friction was mostly in metallic systems (BOWDIN and TABOR, 1964). Therefore most of the micromechanical modeling focused on surface asperities. For the case of rock, however, the sliding surfaces become rapidly covered with a layer of powdered rock gouge. Recent laboratory studies of friction in rock systems which contain a layer of gouge suggest that sliding behavior may be intimately related to the soil-like behavior of the gouge. Indeed the formation of Riedel shears and the concepts of over and under consolidated gouge (MORROW and BYERLEE, 1986) are all from soil mechanics. Hence, although the formalism of friction theory for rock may parallel that for metals, the micromechanical interpretation may be quite different. The observation that stick-slip behavior on initially clean rock surfaces seems to require preslip conditioning of the surfaces (DIETERICH, 1978) or an episode of preseismic slip (OKUBO, 1986) suggests that a thin layer of gouge may be required. DIETERICH (private communication) has observed that if the wear products are cleaned from between sliding surfaces, then a new preconditioning displacement is required before stick-slip is observed. For the case of initially clean surfaces, correlations with surface roughness may actually reflect corresponding variations in the particle size of the newly formed gouge.

In the case of a system which contains gouge, the particle size distribution may play the same role as the surface asperity distribution of a clean metallic surface. The characteristic displacement in such a system was shown by DIETERICH (1981) to be independent of the gouge thickness; an observation which he interpreted as further evidence that flow is localized in shear bands within the gouge. If slip is occurring in shear-bands, it seems reasonable that the characteristic displacement should be set by the largest particle or, for a self-similar gouge, by the upper fractal limit. Figure 5 shows a slip surface through an idealized self-similar gouge of cubic particles. In paper I, we showed that this geometry has a fractal dimension of 2.58 (1.58 in planar section) which is consistent with the 2.60 ± 0.11 which was measured for the Lopez Canyon fault gouge. In this figure the upper fractal limit corresponds to a sharp cutoff at the largest squares; it is apparent that a displacement equal to the dimension of the largest particle to establish a new set of neighbors at all scales. This is equivalent to Dieterich's interpretation of d_c as "the distance required to establish a new set of asperity contacts" except, in the case of gouge, the asperities are identified as the particles.

The micromechanics of stick-slip friction on surfaces separated by a gouge layer are still poorly understood. The relative importance of dilatancy, shear localization, time dependent adhesion, and further comminution within the gouge layer are not known. However, the observation that natural gouge appears to have a fractal geometry, and that characteristic scale lengths associated with stick-slip friction may be determined by the upper fractal limit, suggests that it may be possible to estimate d_c from the gouge particle-size distribution. We have shown above that, for the

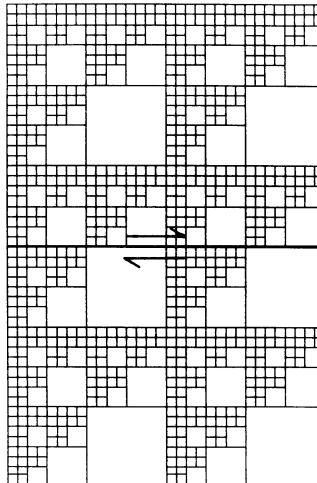


Figure 5

Cross-section through an idealized self-similar gouge of cubic particles which has a fractal dimension of 2.58. The upper and lower fractal limits are sharp cut-offs at the largest and smallest particle sizes. A displacement equal to the upper fractal limit is required to completely change nearest neighbors across the indicated shear localization.

Lopez Fault, this upper fractal limit occurs at a particle size on the order of 1 cm. This value is consistent with the TSE and RICE (1986) estimate that d_c must be less than 8 cm for the instabilities to occur in a San Andreas like fault environment.

Acknowledgements

We wish to thank T. Tullis for providing unpublished surface profiles for Sierra White granite, T. Tullis and J. Dieterich for helpful reviews of the manuscript, J. Byerlee for pointing out similarities between soil mechanics and frictional data, and T. Chao and K. Aki for helpful discussions regarding seismic estimates of the characteristic displacement. This work was supported by National Science Foundation Grant EAR8608246.

REFERENCES

- AKI, K. (1979), *Characterization of Barriers on an Earthquake Fault*, J. Geophys. Res. 84, 6140–6148.
 BIEGEL, R. L., SAMMIS, C. G., and DIETERICH, J. H. (1987), *Frictional Properties of Simulated Fractal Gouge in Westerly Granite*, EOS Trans. Amer. Geophys. Union 68, 1478.
 BURRIDGE, R., and KNOPOFF, L. (1967), *Model and Theoretical Seismology*, Bull. Seis. Soc. Am. 57, 341–371.
 BOWDIN, F. P., and D. TABOR, *The Friction and Lubrication of Solids*, Vol 2 (Clarendon, Oxford 1964).

- BRACE, W. F. (1972), *Laboratory Studies of Stick-slip and their Application to Earthquakes*, Tectonophysics 14, 189–200.
- BRACE, W. F., and BYERLEE, J. D. (1966), *Stick-slip as a Mechanism for earthquakes*, Science 153, 990–992.
- BYERLEE, J. (1978), *Friction of Rocks*, Pure Appl. Geophys. 116, 615–626.
- CAO, T., and AKI, K. (1984/85), *Seismicity Simulation using a Mass-spring Model Characterized by the Displacement Hardening-softening Friction Law*, Pure Appl. Geophys. 122, 10–24.
- CAO, T., and AKI, K. (1987), *Physical Basis for the Magnitude Cut-off Dependence of Seismicity Quiescence*, Pure Appl. Geophys., in press.
- DIETERICH, J. H. (1970), *Time Dependence in Stick-slip Sliding*, Trans. Am. Geophys. Union 51, 423.
- DIETERICH, J. H. (1972), *Time-dependent Friction and the Mechanism of Stick-slip*, J. Geophys. Res. 77, 3690–3697.
- DIETERICH, J. H. (1978), *Time-dependent Friction and the Mechanism of Stick-slip*, Pure Appl. Geophys. 116, 790–806.
- DIETERICH, J. H. (1979), *Modeling of Rock Friction. I. Experimental Results and Constitutive Equations*, J. Geophys. Res. 84, 2161–2168.
- DIETERICH, J. H., *Constitutive properties of faults with simulated gouge*, in *Mechanical Behavior of Crustal Rocks*, Geophys. Monogr. Ser. 24 (AGU, Washington D. C. 1981) (eds. N. L. Carter, M. Friedman, J. M. Logan, and D. W. Stearns) pp. 103–120.
- DIETERICH, J. H., and CONRAD, G. (1984), *Effect of Humidity on Time- and Velocity-dependent Friction in Rocks*, J. Geophys. Res. 89, 4196–4202.
- FEDA, J., *Mechanics of Particulate Materials* (Elsevier, New York, 1982).
- HOFFMANN, N., and SCHONERT, K. (1971), *Bruchanteil von Glaskugeln in Packungen von Franktionen und binaren Mischungen*, Aubereit. Tech. 12, 513–518.
- HOUGH, B. K., *Basic Soils Engineering* (J. Wiley and Sons, New York 1969).
- HUNDLEY-GOFF, E. M., and MOODY, J. B. (1980), *Microscopic Characteristics of Orthoquartzite from Sliding Friction Experiments, I: Sliding Surface*, Tectonophysics 62, 279–299.
- IDA, Y. (1973), *The Maximum Acceleration of Seismic Ground Motion*, Bull. Seismol. Soc. Am. 63, 959–968.
- MARONE, C., SCHOLZ, C. H., and RALEIGH, C. B. (1987), *Constitutive Modeling and Frictional Behavior of Thick Gouge Layers*, EOS Trans. Am. Geophys. Union 68, 1242.
- MOODY, J. B., and HUNDLEY-GOFF, E. M. (1980), *Microscopic Characteristics of Orthoquartzite from Sliding Friction Experiments, II: Gouge*, Tectonophysics 62, 301–319.
- MORROW, C. A., and BYERLEE, J. D. (1986), *The Effect of Velocity Dependent Consolidation on the Frictional Strength of Fault Gouges*, (abst.) EOS Trans. Am. Geophys. Union 67, 1187.
- OKUBO, P. (1986), *Experimental and Numerical Model Studies of Frictional Instability Seismic Sources*, Ph. D. Thesis, Massachusetts Institute of Technology.
- PAPAGEORGIOU, A., and AKI, K. (1983), *A Specific Barrier Model for the Quantitative Description of Inhomogeneous Faulting and the Prediction of Strong Ground Motion II-Applications of the Model*, Bull. Seismol. Soc. Am. 73, 953–978.
- POWER, W. L., TULLIS, T. E., BROWN, S. R., BOITNOTT, G. N., and SCHOLZ C. H. (1987), *Roughness of Natural Fault Surfaces*, submitted to Geophys. Res. Letters.
- RICE, J. R., *The mechanics of earthquake rupture*, in *Proc. Int. School Phys. "Enrico Fermi"*, Course 78, 1980 (eds. A. Dziewonski and E. Boschi) (Italian Physical Society, North Holland Publishing Co. 1980) pp. 555–644.
- RICE, J. R., and RUINA, A. L. (1983) *Stability of Steady Frictional Slipping*, J. Appl. Mech. 50, 343–349.
- ROBERTSON, E. C., *Continuous formation of fault gouge and breccia during fault displacement*, in *Issues in Rock Mechanics* (eds. R. E. Goodman and F. E. Henze) (A.I.M.M.P.E., New York 1982).
- RUINA, A. (1983), *Slip instability and State Variable Friction Laws*, J. Geophys. Res. 88, 10359–10370.
- SAMMIS, C. G., KING, G., and BIEGEL, R. (1987), *The Kinematics of Gouge Deformation*, Pure Appl. Geophys. 125, 777–812.
- SCHOLZ, C. H., (1987), *Wear and Gouge Formation in Brittle Faulting*, Geology 15, 493–495.
- SCHOLZ, C. H., and AVILES, C. A. (1986), *The fractal geometry of faults and faulting*, in 5th Ewing Symposium, *Earthquake Source Mechanisms*, Geophys. Monogr. Ser. 37, (AGU, Washington D.C. 1986) (eds. S. Das, J. Boatwright, and C. H. Scholz) pp. 147–156.

- SCHOLZ, C. H., and ENGELDER, J. T. (1976), *The Role of Asperity Indentation and Ploughing in Rock Friction. I. Asperity Creep and Stick-slip*, Int. J. Rock Mech. Min. Sci. 13, 149-154.
- SCHOLZ, C. H., MOLNAR, P., and JOHNSON, T. (1972), *Detailed Studies of Frictional Sliding of Granite and Implications for the Earthquake Mechanism*, J. Geophys. Res. 77, 6392-6406.
- STEIER, K., and SCHONERT, K. (1972), *Verformung und Bruchphänomene unter Druckbeanspruchung von sehr kleinen Körnern aus Kalkstein, Quarz, und Polystyrol*, Dechema Monograph 69, 167-192.
- STUART, W. D. (1979), *Strain-softening Instability Model for the San Fernando Earthquake*, Science 203, 907-910.
- TAYLOR, D. W., *Fundamentals of Solid Mechanics* (J. Wiley and Sons, New York 1948).
- TSE, S. T., and RICE, J. R. (1986), *Crustal Earthquake Instability in Relation to the Depth Variation of Frictional Slip Properties*, J. Geophys. Res. 91, 9452-9472.
- TULLIS, T. E., and WEEKS, J. D. (1986), *Constitutive Behavior and Stability of Frictional Sliding of Granite*, Pure Appl. Geophys. 124, 383-414.
- WALSH, J. B. (1971), *Stiffness in Faulting and in Friction Experiments*, J. Geophys. Res. 76, 8597-8598.
- WILSON, E. B. Jr., *An Introduction to Scientific Research* (McGraw Hill, New York 1952).

(Received May 6, 1987, revised December 1, 1987, accepted December 14, 1987).

Scaling Geology and Seismic Deconvolution

J. P. TODOESCHUCK^{1,2} and O. G. JENSEN³

Abstract—The reflection seismic signal observed at the surface is the convolution of a wavelet with a reflection sequence representing the geology. Deconvolution of the observations without prior knowledge of the wavelet can be done by making assumptions about the statistics of the reflection sequence. In particular, the widely used prediction error filter is obtained by assuming that the power spectra of reflection sequences are white. However, evidence from well logs suggests that the power spectra are in fact proportional to a power of the frequency f , that is, to f^α , with α equal approximately to 1.

We have found a simple modification to the prediction error filter that markedly improves deconvolution for reflection sequences with such scaling behaviour. We have calculated three reflection sequences from sonic logs of a well off Newfoundland and two wells in Quebec. The three values of α were 0.84, 0.95 and 1.20. We made artificial seismograms from the sequences and deconvolved them with the prediction error filter and our new filters. The errors between the known reflection sequences and the recovered ones for the prediction error filter were 20%, 26%, and 31%; for the new filters 0.5%, 2.0% and 0.5%.

Key words: Deconvolution, scaling geology, fractal geology.

1. Introduction

The reflection sequence in a sedimentary section arises from contrasts in the acoustic impedance of the rocks in each layer. HOSKEN (1980) and WALDEN and HOSKEN (1985) have argued from well-log data that in general the acoustic impedance is scaling, with a power spectrum inversely proportional to spatial frequency. As a consequence, reflection sequences have power spectra proportional to frequency. This hypothesis has considerable implications for the deconvolution of seismic signals (WALDEN and HOSKEN, 1986).

The reflection seismic signal recorded at the surface is often very complicated and hard to interpret. The signal must be deconvolved with an appropriate filter.

¹ Mineral Exploration Research Institute, P.O. Box 6079, Station A, Montreal, P.Q. H3C 2A7, Canada.

² Present address: Defence Research Establishment, Pacific, FMO, Victoria, B.C. V0S 1B0, Canada.

³ Geophysical Laboratory, Department of Geological Sciences, McGill University, 3450 University Street, Montreal, P.Q. H3A 2A7, Canada.

The oldest and most widely used deconvolution technique is predictive decomposition (ROBINSON, 1957). It assumes that the reflection sequence is uncorrelated; that is, it has a white power spectrum and a delta function autocorrelation.

In two previous papers (TODOESCHUCK and JENSEN, 1988, 1989), we have presented a simple modification to predictive decomposition to deal with reflection sequences with power spectra strictly proportional to frequency. When the resulting filter was used to deconvolve an artificial seismogram made from reflection coefficients obtained from well log, the new filter recovered the actual sequence with an RMS error of 9.5% compared with 20% for the unmodified filter.

Now, what if the power spectrum of the reflection sequence is not strictly proportional to frequency, but varies as f^α , say? In the limit of small reflections, this means that the acoustic impedance varies as $1/f^\beta$ with $\beta = 2 - \alpha$ (WALDEN and HOSKEN, 1985). Indeed, this is a more natural expectation than that the proportionality should always be exact. This would be the case if the acoustic impedance function could be modelled as a generalized Brownian process (MANDELBROT, 1983, p. 351). The extension from our previous filter is straightforward.

2. Filters for $1/f^\beta$ Geology

We may regard a set of seismic observations $\{x_k\}$ as the result of a convolution of wavelet $\{w_k\}$ with a reflection sequence $\{r_k\}$. We seek the filter $\{a_n\}$ which when applied to the observations produces an output $\{e_k\}$ which we desire to be the same as $\{r_k\}$.

The autocorrelation function of the observations is $\{R_m\}$ given by

$$R_m = E[x_k x_{k-m}] \quad (1)$$

where E denotes the expectation operator.

Assuming for the moment that we have an infinite data set, the output is given by

$$e_k = \sum_n a_n x_{k-n}. \quad (2)$$

The autocorrelation Y_j of the output is $E[e_k e_{k-j}]$ so

$$Y_j = \sum_m a_m \sum_n a_n R_{m+j-n}. \quad (3)$$

The operator for predictive decomposition, the prediction error filter (PEF), $\{a_n\}$, is obtained by minimizing Y_0 with respect to each parameter a_n , $n > 1$. (We normalize $a_0 = 1$.) This gives the set of normal equations

$$\sum_n a_n R_{p-n} = 0, \quad p \geq 1 \quad (4)$$

which are solved to yield the filter $\{a_n\}$.

We see immediately on substituting each of the normal equations back into the expression for the Y_j 's that the prediction error filter requires that all the Y_j 's are zero for $j \geq 1$, that is, that the power spectrum of the reflection sequence is white. Suppose the reflection sequence has a number of nonzero Y_j 's, say, for $0 \leq j \leq J$. (The reason for posing the problem in this manner will become clear below.) We may still minimize Y_0 with respect to the parameters a_n , $n \geq J + 1$, leading to a reduced set of normal equations

$$\sum_n a_n R_{p-n} = 0, \quad p \geq J + 1 \quad (5)$$

which only requires that the Y_j 's with $j > J$ are zero. These equations are identical with those of the PEF. The normal equations with $p \leq J$ can be obtained directly from the equations for the Y_j 's.

From (3)

$$Y_J = a_0 \sum_n a_n R_{J-n} \quad (6)$$

so, since by our normalization $a_0 = 1$

$$\sum_n a_n R_{J-n} = Y_J. \quad (7)$$

Likewise

$$Y_{J-1} = a_0 \sum_n a_n R_{J-1-n} + a_1 \sum_n a_n R_{J-n} \quad (8a)$$

so-

$$\sum_n a_n R_{J-1-n} = Y_{J-1} - a_1 Y_J \quad (8b)$$

while

$$\sum_n a_n R_{J-2-n} = Y_{J-2} - a_1 Y_{J-1} - (a_2 - a_1^2) Y_J \quad (9)$$

and so on. We can thus recover a full set of equations corresponding to the normal equations for the PEF.

When the power spectrum $P(f)$ of the reflection coefficients is strictly proportional to frequency up to the Nyquist frequency $f_0 = 1/2 \Delta t$, where Δt is the sampling interval of the reflection sequence, the discretized autocorrelation function Y_j is

$$Y_j = -\frac{4Y_0}{j^2 \pi^2}, \quad j \text{ odd}$$

$$= 0, \quad j \text{ even} \quad (10)$$

(TODOESCHUCK and JENSEN, 1988). The $1/j^2$ means that the autocorrelation dies off rapidly with lag. The autocorrelation functions corresponding to power spectra proportional to f^α have similar behaviour for a wide range of α . Figure 1 shows the first three terms of the normalized autocorrelation function as α varies from 0 to 2. Note that the lag = 1 term (Y_1/Y_0) is negative throughout the range and is of larger magnitude than the other terms. It has been known for some time that negative values at small lags are characteristic of the autocorrelation function of reflection sequences generated from well logs (O'DOHERTY and ANSTEY, 1971). The size of the lag = 1 term compared with the rest suggests that we should truncate the autocorrelation function after this term. (The standard PEF results from truncating after the lag = 0 term.) We will therefore consider the case where the only nonzero values of the autocorrelation function are Y_0 and Y_1 .

The normal equations become

$$\sum_n a_n R_n = Y_0 - a_1 Y_1,$$

$$\sum_n a_n R_{1-n} = Y_1,$$

and

$$\sum_n a_n R_{p-n} = 0, \quad p \geq 2. \quad (11)$$

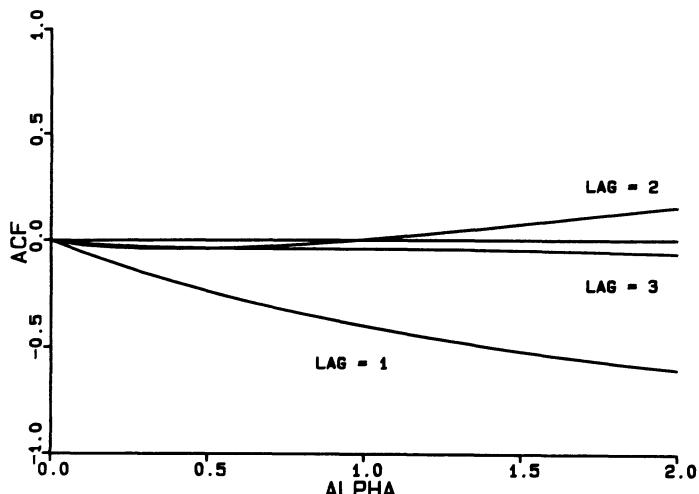


Figure 1

Variation with α of the lag = 1, 2, 3 terms of the normalized autocorrelation function of a f^α process.

The first of these equations provides the means of establishing the scale of the Y 's (since we know the ratio for a given value of α between Y_0 and Y_1 but not their size). The presence of a filter coefficient on the right-hand side of the equation can be dealt with by an iterative method. We will generally be interested in obtaining a filter of finite length, say of length $L + 1$, so we can write the equations in matrix form as

$$\begin{bmatrix} R_0 & R_1 & \dots & R_L \\ R_1 & R_0 & \dots & R_{L-1} \\ \vdots & \vdots & \ddots & \vdots \\ R_L & R_{L-1} & \dots & R_0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_L \end{bmatrix} = \begin{bmatrix} Y_0 - a_1 Y_1 \\ Y_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (12)$$

The matrix is of Toeplitz form which of course offers advantages in the solution of the system of equations. Our technique has been to assume initially that Y_1 is zero, calculate the filter coefficients using the FORTRAN subroutine EUREKA (ROBINSON, 1978, pp. 44–45), use these to evaluate Y_0 and therefore Y_1 and, finally, to substitute the new value of Y_1 into (12). The solution converges satisfactorily after a small number of iterations.

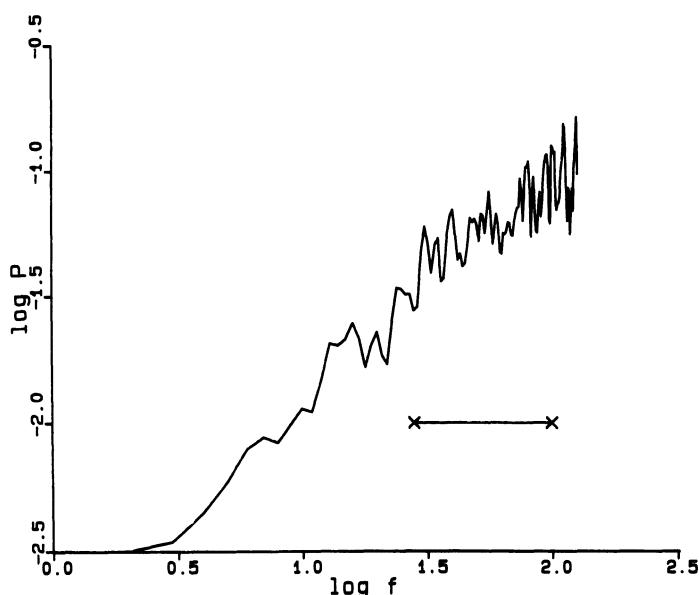


Figure 2

Power spectrum of reflection sequence calculated from a sonic log of the Puffin well. Least squares slope of the entire spectrum is 0.84; of the part marked by the horizontal line, 0.63. Arbitrary units.

3. Examples

Figure 2 shows the power spectrum of the reflection sequence calculated from a sonic log digitized at 1 millisecond intervals from the Amoco IOE A-1 Puffin B-90 well on the Grand Banks of Newfoundland ($44^{\circ}39'09''\text{N}$, $53^{\circ}42'28''\text{W}$). In our previous papers, we argued that the slope on log-log paper (that is, the parameter α) of the power spectrum was close to unity. In fact, the least squares slope is $\alpha = 0.84 \pm 0.02$.

We generated an artificial seismogram from the reflection sequence by direct recursion filtering, using a simple two pole wavelet and then deconvolved it with a set of filters, each assuming a different value of α . The optimum filter length was found, by trial, to be 3. We can compare the recovered sequence $\{e_i\}$ with the true sequence $\{r_i\}$. Figure 3 shows the RMS error, Λ , given by

$$\Lambda^2 = \sum_i (e_i - r_i)^2 / \sum_i r_i^2 \quad (13)$$

as a function of the value of α used in the filter. The error is smaller for $\alpha = 0.84$ than for either α equal to zero or one but is even smaller, falling to 0.5%, for $\alpha = 0.63$, corresponding to $\beta = 1.37$.

The reason for discrepancy is the restricted frequency content of our wavelet. The wavelet and its power spectrum are shown in Figures 4 and 5. The energy of the wavelet is concentrated between 110 and 390 Hz as shown on Figure 5. The least

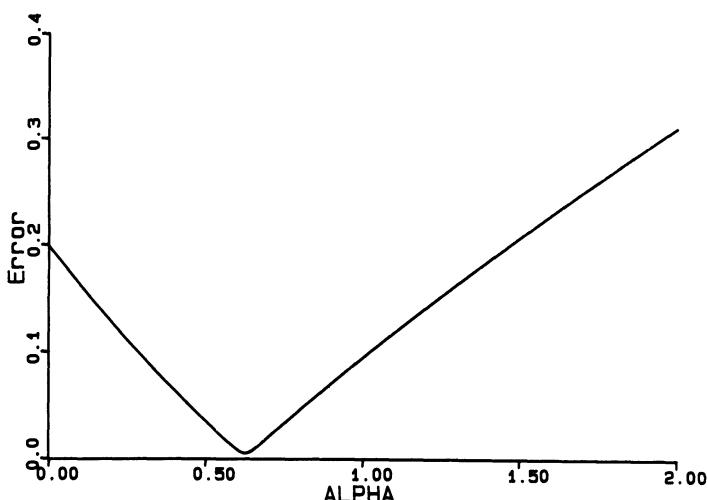


Figure 3

The RMS error between the true reflection coefficients and those recovered by deconvolution with filters for various values of α . The minimum error is for $\alpha = 0.63$.

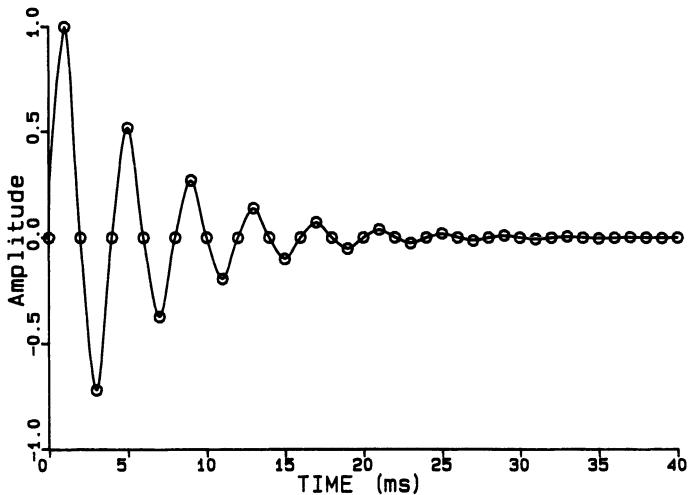


Figure 4
The wavelet used in producing the artificial seismograms.

squares slope of the power spectrum from the Puffin well between these limits is, as might be expected, 0.63 ± 0.01 .

Figure 6 shows the Puffin well reflection sequence and the artificial seismogram calculated from it. Figure 7 shows, to the same scale, the errors between the true and recovered sequences for the PEF with $\alpha = 0$, and the filter for $\alpha = 1$. The top part of Figure 8 shows the error for the filter with $\alpha = 0.63$ to the same scale as before; the bottom part at a magnification of ten times. We attribute much of the exactness with which the true result is recovered to the simplicity of our wavelet. Nonetheless, the improvement is undeniable.

We will repeat this analysis on two other wells, with much the same results. Figure 9 is the power spectrum of the reflection sequence calculated from a sonic log, again digitized at 1 millisecond intervals, from the Shell Wickham No. 1 well, in the St. Lawrence Lowlands of Quebec ($45^{\circ}48'04''N$, $72^{\circ}25'42''W$). The least squares slope is 1.02 ± 0.05 . We found that the optimum value of α in the deconvolution was 0.95. Figure 10 shows the original reflection sequence and the error when the seismogram was deconvolved with the PEF. The RMS error was 26%. Figure 11 shows the errors for deconvolution with $\alpha = 1.0$ and with $\alpha = 0.95$. The RMS errors are 2.4% and 2.0%, respectively. The error in the optimum case is somewhat higher than for the Puffin well. This is probably due to the length of the sequence. We found in tests with artificial sequences that the error with the filter for $\alpha = 1$ increased as the sequences grew shorter.

For this reason, the next well was digitized at 0.5 millisecond intervals. Figure 12 is the power spectrum for the Soquip Ste. Helene well, again in the St. Lawrence Lowlands of Quebec ($45^{\circ}44'12''N$, $72^{\circ}46'24''W$). The least squares slope is

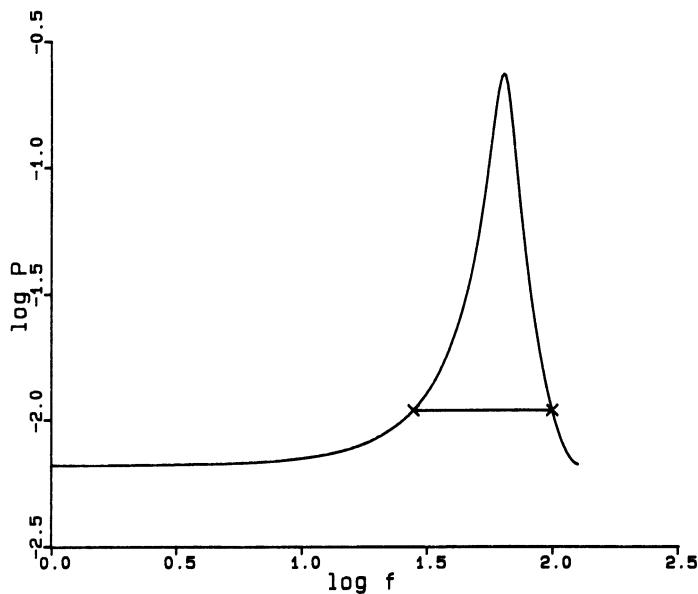


Figure 5
Power spectrum of the wavelet to the same scale as Figure 2.

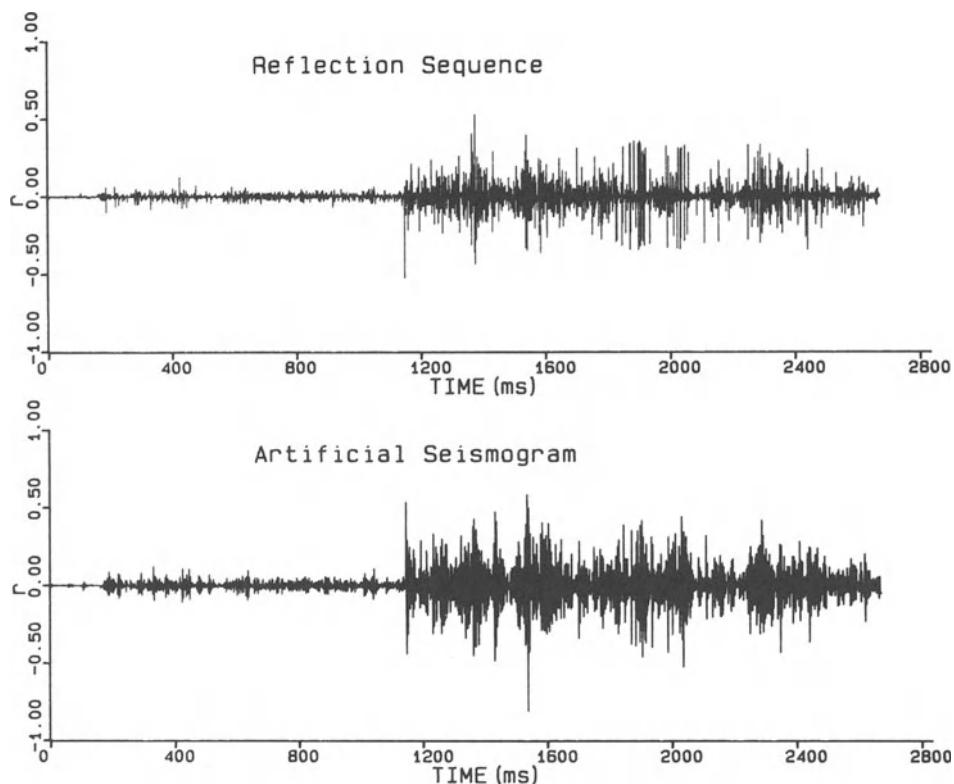


Figure 6
Top. The reflection sequence from the Puffin well. Bottom. The artificial seismogram for the well.

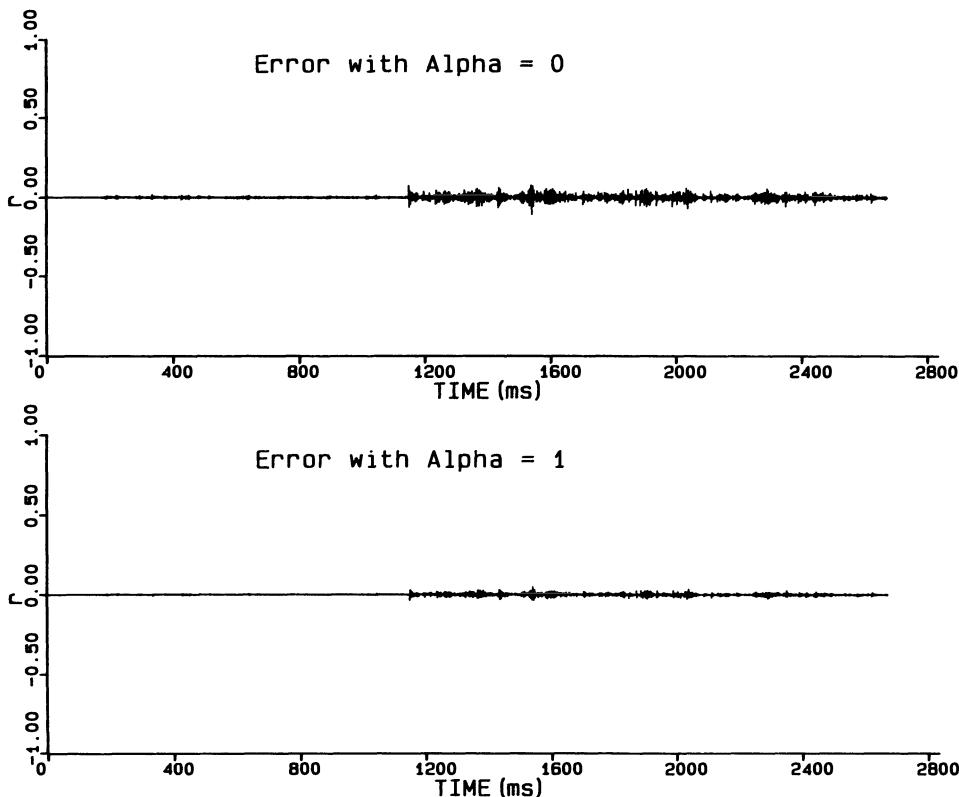


Figure 7

Top. The error between the real reflection sequence and that recovered by deconvolution with a filter with $\alpha = 0$, the prediction error filter. The RMS error is 20%. *Bottom.* The error between the real reflection sequence and that recovered by deconvolution with a filter with $\alpha = 1.0$. The RMS error is 9.5%.

1.19 ± 0.04 . We found the optimum value of α to be 1.20. Figure 13 shows the original reflection sequence and the error when the seismogram was deconvolved with the PEF. The RMS error was 31%. Figure 14 shows the errors for deconvolution with $\alpha = 1.0$ and with $\alpha = 1.20$. The RMS errors are 4% and 0.5%, respectively.

The slopes in the three real examples were all close to unity. We might expect from Figure 1 that the approximation involved in truncating after two terms would break down for some values of α , in particular around $\alpha = 2$. This is borne out by Figure 15, which shows the error in deconvolving seismograms made with artificial reflection sequences having power spectra proportional to f^α with filters for the appropriate value of α in the two term case. Evidently, two terms are enough up to $\alpha = 1.5$ or so, but the filters break down quite badly after that. Taking more terms may correct this.

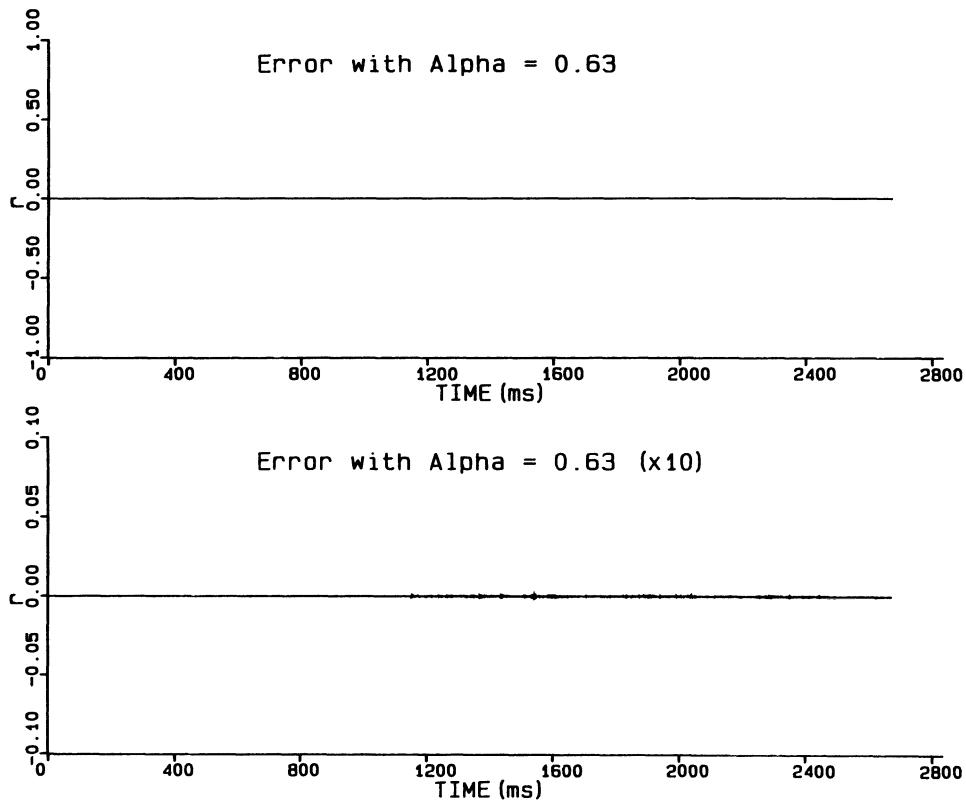


Figure 8

Top. The error between the real reflection sequence and that recovered by deconvolution with a filter with $\alpha = 0.63$. *Bottom.* The error between the real reflection sequence and that recovered by deconvolution with a filter with $\alpha = 0.63$ expanded by 10. The RMS error is 0.5%.

It is interesting to see how the prediction error filter, which of course is a one term truncation, performs on the same set of artificial sequences. This is shown in Figure 16. As might be expected, the error increases as α moves away from zero, where the truncation is exact. Note that the error does not blow up for $\alpha > 1.5$ as in the two term case.

4. Conclusion

Do reflection sequences obey scaling laws? The few examples which have been presented are obviously inadequate to settle the point. It appears from the discussion above that a knowledge of the scaling law obeyed by a reflection sequence is of great value in the deconvolution of a seismogram from that sequence. Given a

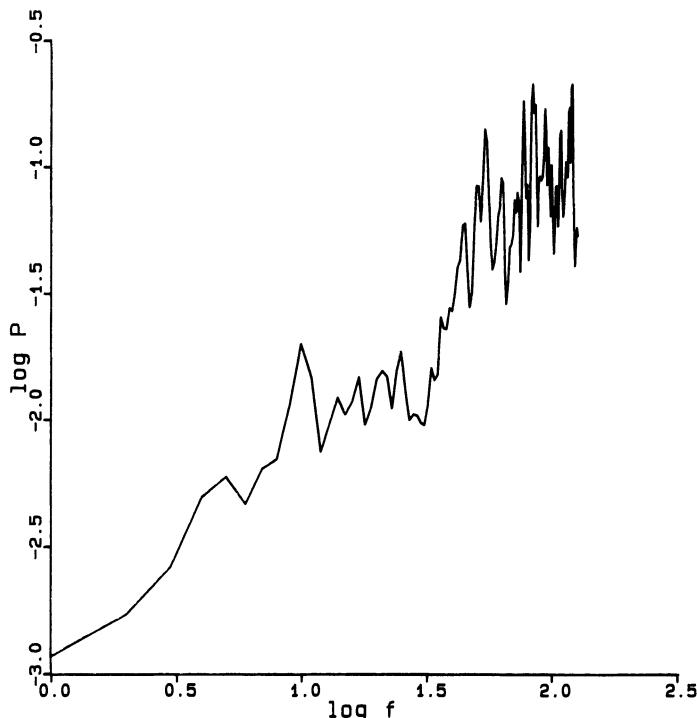


Figure 9

Power spectrum of reflection sequence calculated from a sonic log of the Wickham No. 1 well. Least squares slope of the spectrum is 1.02. Arbitrary units.

scaling behaviour, this is not surprising. What is surprising is the extent to which keeping one extra term can improve the deconvolution. Evidently much information resides in that term.

In the common seismic case, we will not know beforehand what the best value of α is for deconvolution. Several approaches suggest themselves. It may be possible to determine the optimum value from the seismogram itself. For example, the RMS differences between deconvolved sequences using different α 's goes through a minimum in the neighbourhood of the correct value. However, this minimum is far too broad and shallow to be of much use. There may be some other test which can be used.

It may be that α is sufficiently constant for the wells of a geologic province so that a given value can be used for that region. This is an obvious point for further investigation. On the other hand, α can be used as an adjustable parameter. Heuristically, one expects the error in the deconvolution to look like noise in a section. The value of α could be stepped over a range and the best value chosen on either the appearance of the section or by some coherence test. A coherency test might allow the best value of α to be chosen on a trace by trace basis.

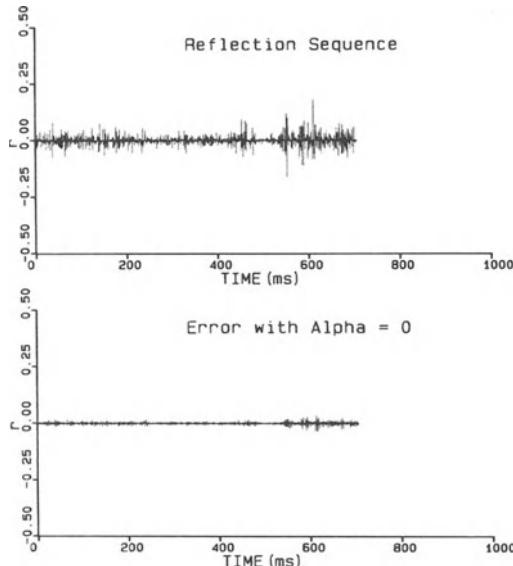


Figure 10

Top. The reflection sequence for the Wickham No. 1 well. *Bottom.* The error between the real reflection sequence and that recovered by deconvolution with a filter with $\alpha = 0$. The RMS error is 26%.

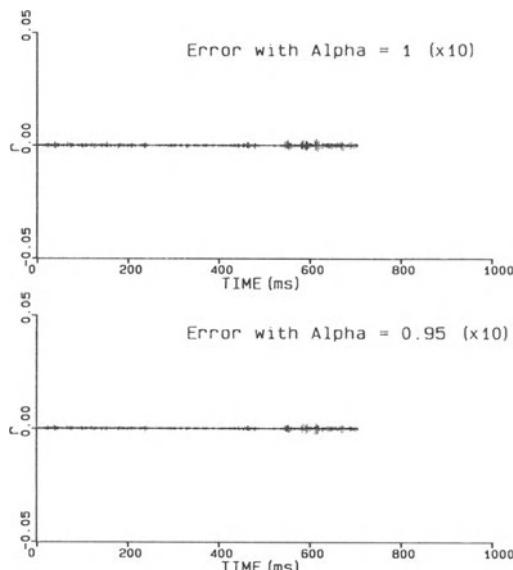


Figure 11

Top. The error between the real reflection sequence and that recovered by deconvolution with a filter with $\alpha = 1.0$ expanded by 10. The RMS error is 2.4%. *Bottom.* The error between the real reflection sequence and that recovered by deconvolution with a filter with $\alpha = 0.95$ expanded by 10. The RMS error is 2.0%.

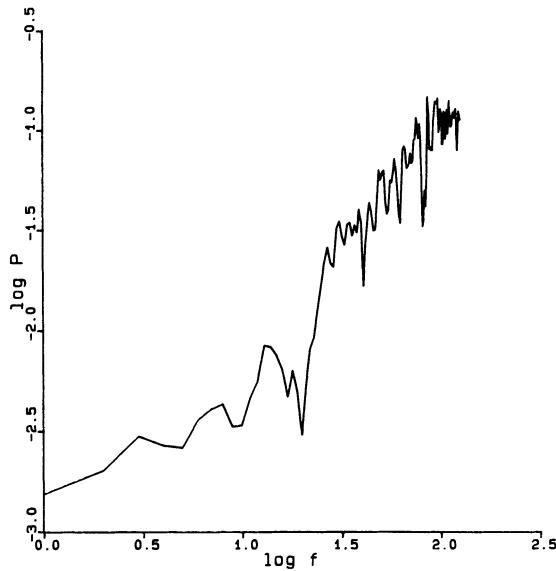


Figure 12

Power spectrum of reflection sequence calculated from a sonic log of the Ste. Helene well. Least squares slope of the spectrum is 1.19. Arbitrary units.

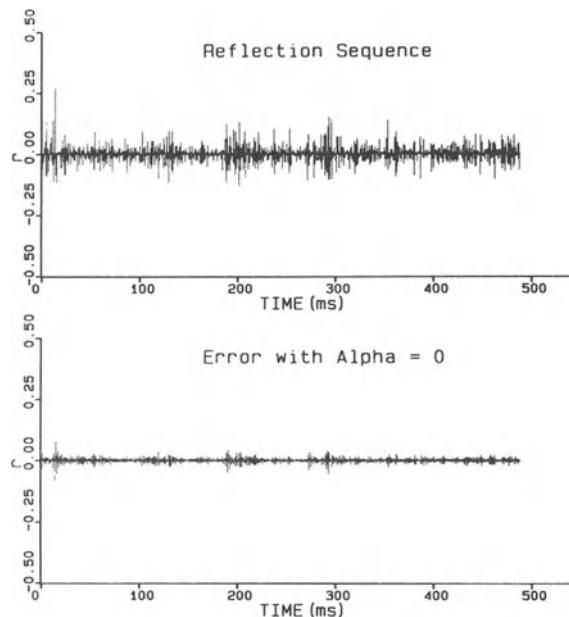


Figure 13

Top. The reflection sequence for the Ste. Helene well. *Bottom.* The error between the real reflection sequence and that recovered by deconvolution with a filter with $\alpha = 0$. The RMS error is 31%.

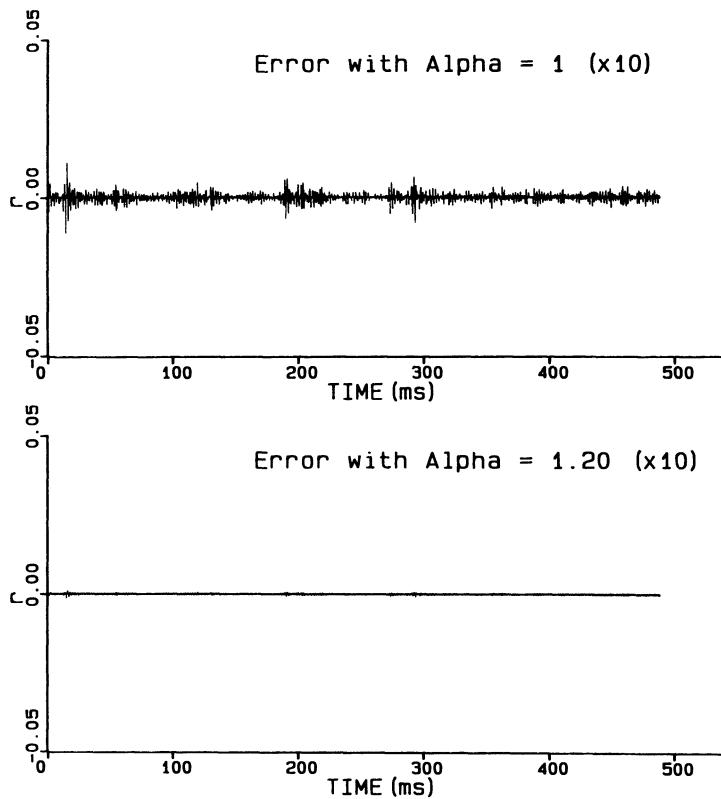


Figure 14

Top. The error between the real reflection sequence and that recovered by deconvolution with a filter with $\alpha = 1.0$ expanded by 10. The RMS error is 4%. *Bottom.* The error between the real reflection sequence and that recovered by deconvolution with a filter with $\alpha = 1.20$ expanded by 10. The RMS error is 0.5%.

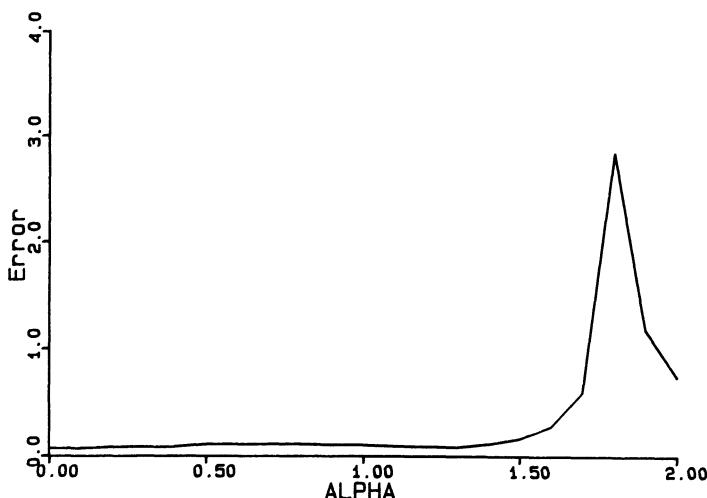


Figure 15

Test of the two term filter. Error in deconvolving seismograms from artificial reflection sequences with different values of α with filters for the same α . The two term filter works well for $\alpha < 1.5$.

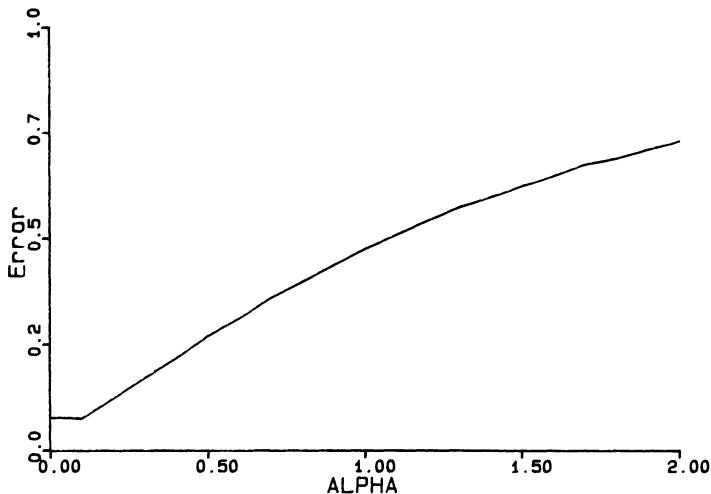


Figure 16

Error in deconvolving seismograms from the same reflection sequences with the prediction error filter.

REFERENCES

- HOSKEN, J. W. J. (1980), *A Stochastic Model of Seismic Reflections*, 50th Annual Meeting Soc. Explor. Geophys., Houston. (Abstract G-69, Geophysics 46, 419.)
- MANDELBROT, B. B., *The Fractal Geometry of Nature* (W. H. Freeman, San Francisco 1983).
- O'DOHERTY R. F., and ANSTEY, N. A. (1971), *Reflections on Amplitudes*, Geophysical Prospecting 19, 430-458.
- ROBINSON, E. A. (1957), *Predictive Decomposition of Seismic Traces*, Geophysics 22, 767-778.
- ROBINSON, E. A., *Multichannel Time Series Analysis with Digital Computer Programs*, revised (Holden-Day, San Francisco 1978).
- TODESCHUCK, J. P., and JENSEN, O. G. (1988), *Joseph Geology and Seismic Deconvolution*, Geophysics 53, 1410-1414.
- TODESCHUCK, J. P., and JENSEN, O. G., *1/f Geology and Seismic Deconvolution*, to appear in Nonlinear Variability and Geophysics (eds. Lovejoy, S. and Schertzer, D.) (Reidel, Dordrecht 1989).
- WALDEN, A. T., and HOSKEN, J. W. J. (1985), *An Investigation of the Spectral Properties of Primary Reflection Coefficients*, Geophysical Prospecting 33, 400-435.
- WALDEN, A. T., and HOSKEN, J. W. J. (1986), *The Nature of the Non-Gaussianity of Primary Reflection Coefficients and its Significance for Deconvolution*, Geophysical Prospecting 34, 1038-1066.

(Received March 30, 1987, revised/accepted December 23, 1987)

Fractured but Not Fractal: Fragmentation of the Gulf of Suez Basement

GABOR KORVIN¹

Abstract—Recent geophysical studies revealed that the Palaeozoic basement of the Gulf of Suez consists of an enormous number of fault blocks whose network qualitatively resembles the contraction-crack polygons which can be found in nature in a wide variety of materials and on all scales (mud cracks, hardening concrete, age cracking in paintings, etc.). The fault network of the Gulf of Suez basement forms a rather uniformly spaced polygonal pattern, most of the blocks are four-sided, the lengths of block sides parallel with the Gulf of Suez axis are exponentially distributed. The power-law size distribution associated with the fractal (scale-free) fragmentation can be possibly ruled out.

The paper calls attention to the necessity of classifying the physical processes leading to fragmentations with exponential-, lognormal-, and power-law size distributions, respectively.

Key words: Fractal, fragmentation, size distribution, Gulf of Suez.

1. Introduction

Fracturing and fragmentation occur on all scales: rocks fragmented by joints and weathering (ALLEGRE *et al.*, 1982), ejecta produced by impacts (FUJIWARA *et al.*, 1977), drift sea ice (ROTHROCK and THORNDIKE, 1984; MATSUSHITA, 1985) or asteroids and meteorites (HARTMANN, 1969; BROWN *et al.*, 1983) are well-known examples.

The cumulative frequency distribution of fragment size is generally approximated by one of the following functions: by an exponential distribution (BROWN *et al.*, 1983):

$$N(r) \sim \exp[-(r/r_0)^v] \quad (1)$$

where $N(r)$ is number of fragments greater than r ; r_0 and v are constants; by a logarithmic normal distribution (EPSTEIN, 1947):

$$N(r) \sim \int_r^{\infty} \frac{1}{\sigma(x - r_1)} \exp\left[-\frac{1}{2\sigma^2} \left\{\log \frac{x - r_1}{b}\right\}^2\right] dx \quad (2)$$

¹ Department of Geology and Geophysics, University of Adelaide, G.P.O. Box 498, Adelaide, South Australia, 5001.

(σ , r_1 , and b are constants); or by the power-law distribution (HARTMANN, 1969; BADER, 1970; FUJIWARA *et al.*, 1977; MANDELBROT, 1982; ROTHROCK and THORNDIKE, 1984; MATSUSHITA, 1985; TURCOTTE, 1986):

$$N(r) \sim (r/r_0)^\alpha \quad (\alpha < 0, r \neq 0). \quad (3)$$

In equations (1)–(3) the “size” r of the fragments is not uniquely defined except perhaps for the most simple geometrical objects as a sphere (diameter) or a cube (length of an edge). For irregular particles the size commonly depends on the method of measurement. ALLEN (1968, Table 2.1) lists eight differing definitions for the size of sedimentary particles. ROTHROCK and THORNDIKE (1984) in their study of sea ice floe distribution use several measures of size (area, diameter, perimeter) and prove that these quantities are highly correlated and lead to the same type of probability distribution. In experimental studies the most commonly used measure of size is the “sieve diameter”: for the particles passing through the sieve with opening diameter r their size is regarded as less than r (see *e.g.*, EPSTEIN, 1947; FUJIWARA *et al.*, 1977). In certain cases equations (1)–(3) are expressed in terms of total weight or mass of particles larger than a given mass m . For example, FUJIWARA *et al.* (1977) use instead of equation (3),

$$M(< s) \sim s^k \quad (4)$$

or

$$N(> m) \sim m^{-k} \quad (5)$$

where $M(< s)$ is the total mass of fragments of size less than s and $N(> m)$ is the number of fragments with mass larger than m . In equations (4)–(5) the connection between mass and size is provided by

$$m(s) = \frac{4}{3} \pi \beta \rho \left(\frac{s}{2}\right)^3$$

where β is an empirical shape factor and ρ is density of the fragmented material.

The power-law distribution of particle size has been observed in a wide variety of fragmentation processes of earth materials (HARTMANN, 1969; BADER, 1970; FUJIWARA *et al.*, 1977; MANDELBROT, 1982; ROTHROCK and THORNDIKE, 1984; MATSUSHITA, 1985; TURCOTTE, 1986). Formally, equation (3) is similar to the power-law size distribution which is exhibited by many geographical objects over a limited range of sizes. KORČAK (1940) first described this distribution for the areas of islands; recent examples include sizes of lakes (KENT and WONG, 1982) and caves (CURL, 1986). MANDELBROT (1975) proved that KORČAK's law is the consequence of the repetitive subdivision of geometric figures and the exponent α in KORČAK's law

$$N(r) \sim (r/r_0)^\alpha$$

is related to the fractal dimension of the objects which are usually self-similar, that is their arbitrarily small substructures look statistically similar to the whole object

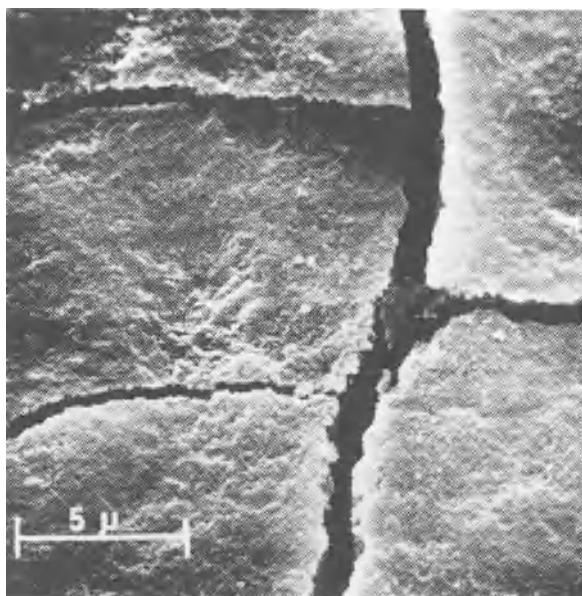


Fig. 1(a)

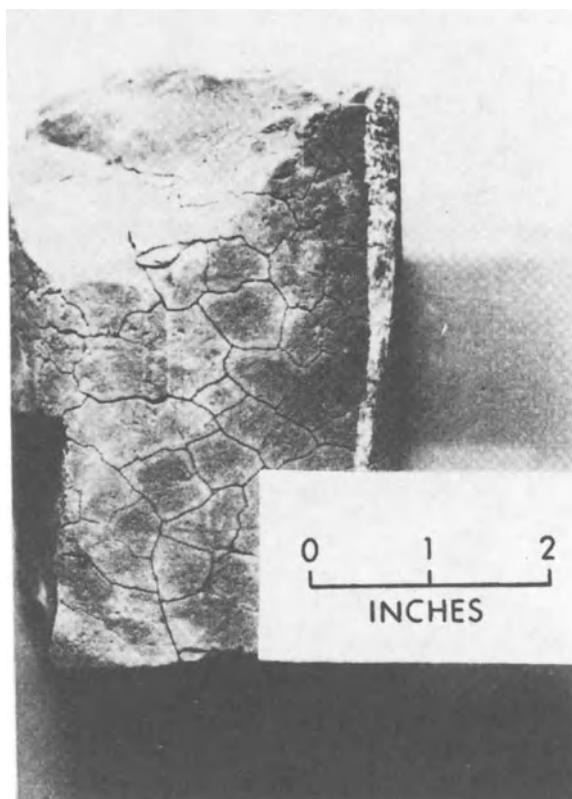


Fig. 1(b)



Fig. 1(c)

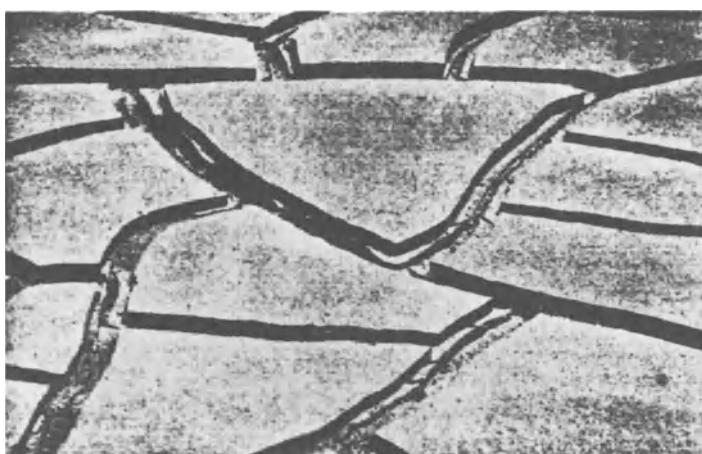


Fig. 1(d)

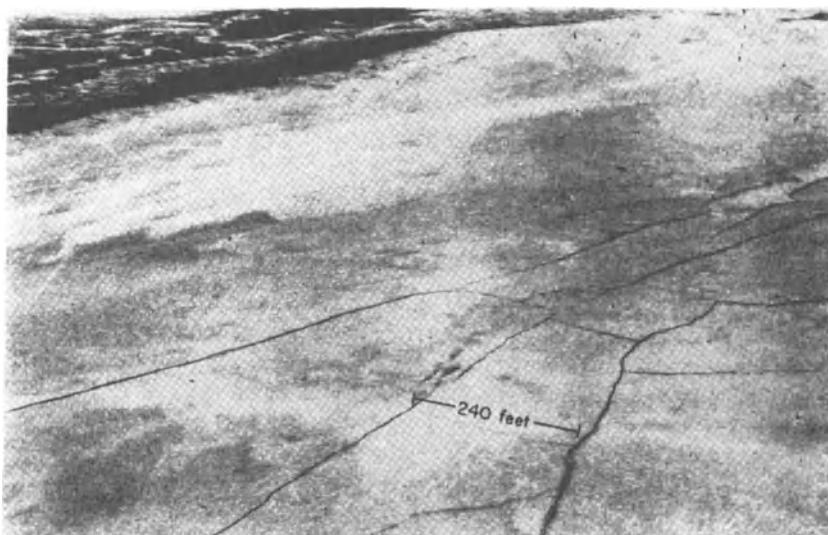


Fig. 1(e)

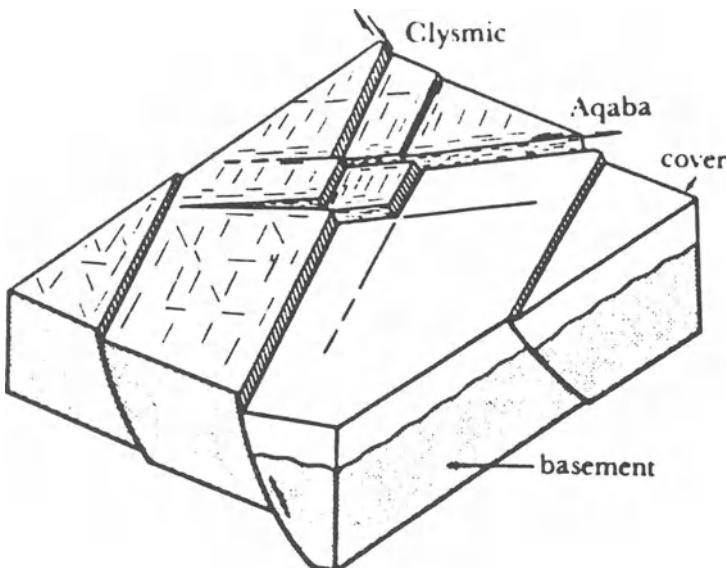


Fig. 1(f)

Figure 1

Examples of contraction-crack polygons at different scales. a) Cracks in wood opal, from SEGNET *et al.* (1970); b) contractional fractures in sedimentary rock, from NELSON (1979); c) age cracking in a desiccated canvas painting, from KECK (1965); d) desiccation mudcrack, France. Width of the foreground about 50 cms. From PLUMMER and GOSTIN (1981); e) desiccation fissures, Black Rock Desert, Nevada. From WILLDEN and MABEY (1961); f) schematic zig-zag faulting pattern of the Gulf of Suez Basement. From JARRIGE *et al.* (1986).

under the proper magnification. (A rigorous definition of self-similarity can be found in MANDELBROT, 1982). It has recently been recognised that the power-law size distribution observed in the fragmentation of earth materials (rocks and sea ice) is a consequence of the scale invariance of the fragmentation mechanism, that is the pre-existing zones or planes of weakness where breakage occurs, exist on all scales (MATSUSHITA, 1985; TURCOTTE, 1986). In these fractal fragmentation theories it is assumed that the flaws have a hierarchical structure where a fracture at the macroscopic scale is a consequence of the accumulation of fractures at lesser scales. The formalization of this principle has recently lead to the renormalization group methods of predicting rock failure (ALLEGRE *et al.*, 1982; MADDEN, 1983; TURCOTTE, 1986).

The scale invariance of the fragmentation process seems to be corroborated by the fractal nature of fracture surfaces on the microscopic scale (MANDELBROT *et al.*, 1984; LOUIS *et al.*, 1986; TERMONIA and MEAKIN, 1986) and by the fractal shape of faults and the power-law distribution of the trace length of faults and joints on the macroscopic scale (KING, 1983; SEGALL and POLLARD, 1983; SCHOLZ and AVILES, 1986; OKUBO and AKI, 1987).

A spectacular type of fragmentation is associated with the polygonal network of fractures that may be seen in ceramic glazes, concrete, varnish and paint, mud,

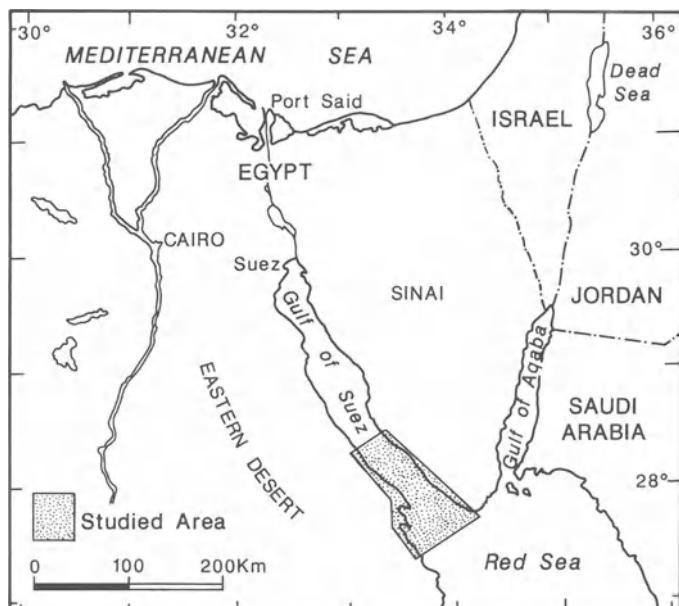


Figure 2
Location map of the Gulf of Suez region.

frozen ground, limestone and other media which undergo volume change. Such a pattern is generally the result of contraction of a layer of homogeneous material perpendicular to a cooling or shrinking surface (LACHENBRUCH, 1962; NEAL *et al.*, 1968; WALKER, 1986). Though these polygonal networks can be observed on all scales (Figure 1), the size distribution of the "islands" bounded by the cracks (or faults, or fractures) has apparently never been analysed.

Recently, HAMMOUDA (1986) applied direct and indirect aeromagnetic and gravity modelling to map the structural configuration of the Paleozoic basement in the Southern part of the Gulf of Suez (Figures 2, 3). The Gulf of Suez is one of the most intensively faulted areas on the face of the earth (SAID, 1962; JARRIGE *et al.*, 1986) and its basement consists of an enormous number of fault blocks (Figure 3). As the network of faults qualitatively resembles the contraction-crack polygons of Figure 1, a study of the size distribution of the basement blocks might give a clue to the general statistical properties of contraction-crack polygons.

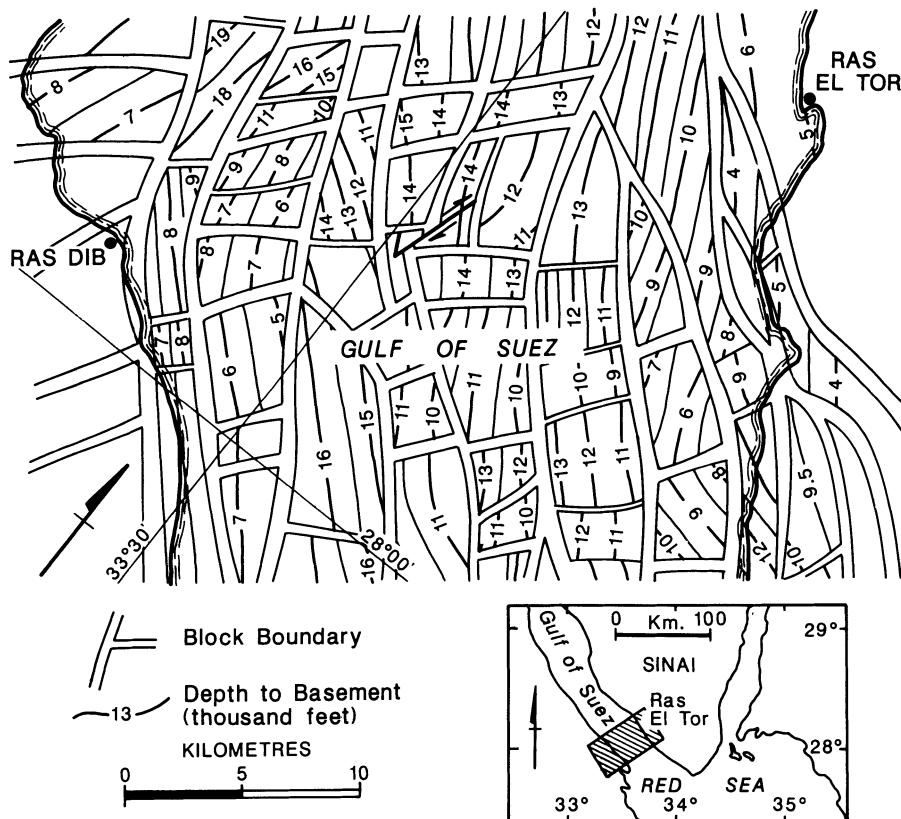


Figure 3

A detail of the structural map of the Palaeozoic basement of the Gulf of Suez. Contour lines show depth to basement in thousand feet. After HAMMOUDA (1986).

2. Statistical Analysis of the Block Size Distribution

The basement map (partly shown in Figure 3) contains 242 blocks. Almost all blocks are four-sided (202), eight are two-sided, seventeen are three-sided and fifteen are five-sided. (DIONNE, 1971, also reports that the contraction crack polygons in muddy tidal flats are three- to five-sided, but the greatest number have four sides.) As in the absence of asymmetrical tectonic forces the surface of a homogeneous medium should be divided by a hexagonal crack system (THOMPSON, 1942; BILLINGS, 1954), the predominance of four-sided blocks suggests anisotropic stress. This is corroborated by the rose diagram of the directional distribution of block sides (Figure 4). There are two distinct sets of faults: one parallel to the Gulf of Suez axis ($N20\text{--}40^\circ\text{W}$) and another almost perpendicular "cross trend" ($N40\text{--}50^\circ\text{E}$, cf., JARRIGE *et al.*, 1986). Thus, by LACHENBRUCH's (1962) classification the system is oriented orthogonal. Because of the possible directional dependence of the distribution of the length of block sides, the following size parameters were separately analysed:

- length of the "cross trend" block sides ($N20\text{--}70^\circ\text{E}$): curve 1 in Figures 5a-c;
- length of the "Gulf of Suez trend" sides ($N10\text{--}60^\circ\text{W}$): curve 2 in Figures 5a-c;
- perimeters of the blocks: curve 3 in Figures 5a-c, see also Figure 7;
- “sieve diameter” defined as the diameter of the smallest circumscribed circle around the given block: Figure 6;
- area of the blocks: Figure 7.

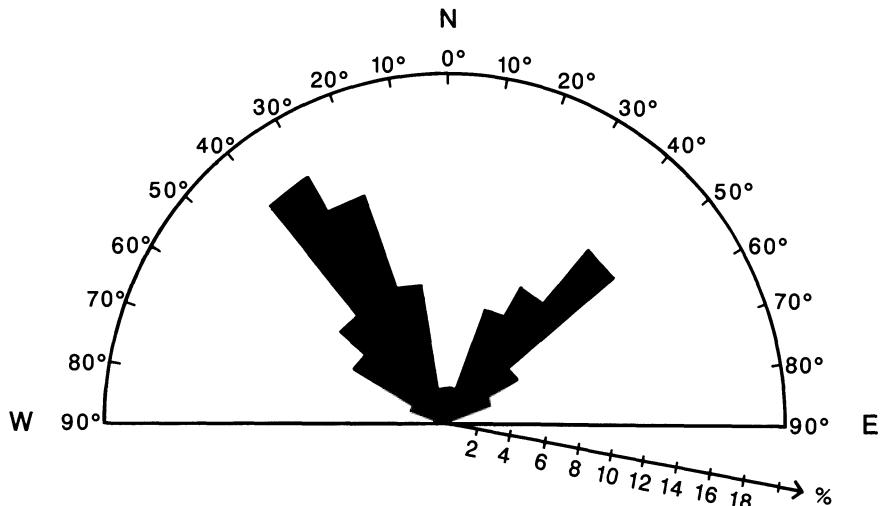


Figure 4

Rose diagram showing the directional distribution of the sides of the fault blocks of the Gulf of Suez basement. The two main peaks correspond to the Clysmic (or "Gulf of Suez" trend $N20\text{--}40^\circ\text{W}$) and the "cross" trend ($N40\text{--}50^\circ\text{E}$) structural directions.

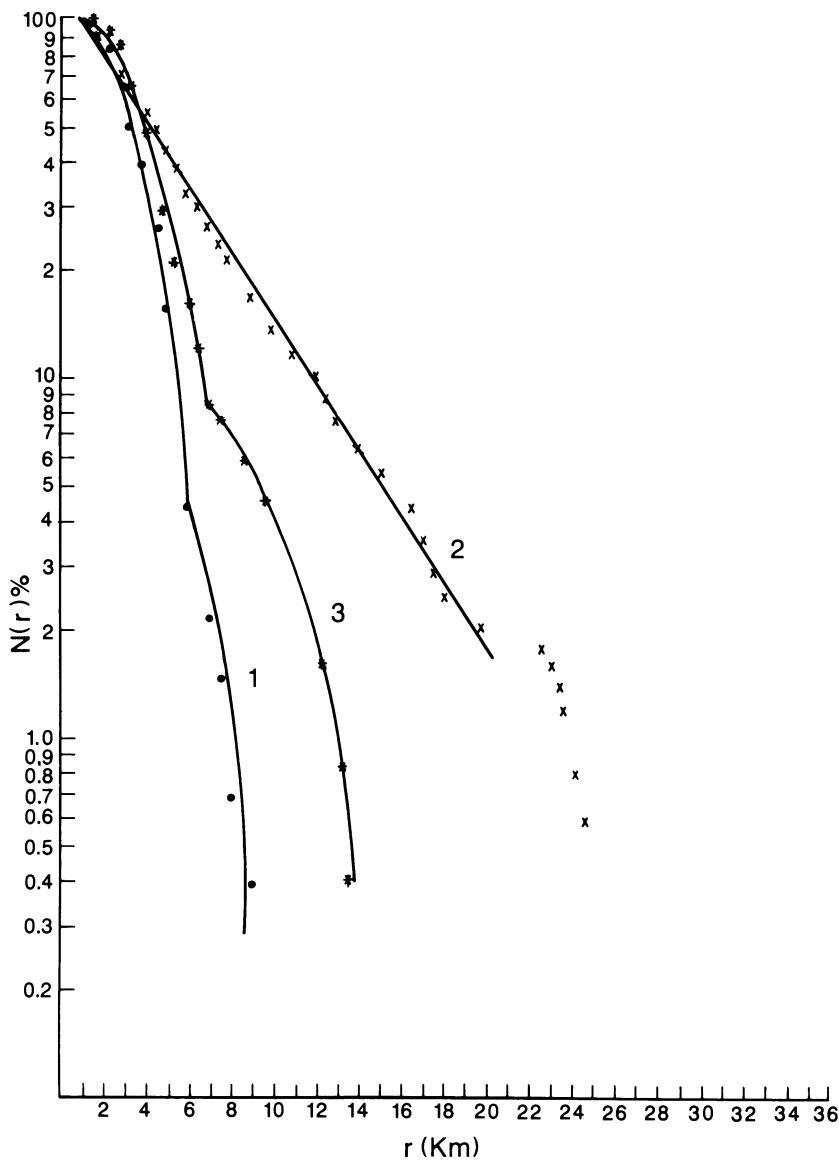


Fig. 5(a)

Figures 5a-c and 6 represent the empirical cumulative frequency curves $N(r)$ of the relative number of block sides greater than r , where r is the respective size parameter. Figure 5a is a semilogarithmic plot, Figure 5b a lognormal probability plot, Figures 5c and 6 are log-log plots. As well-known, exponential distributions

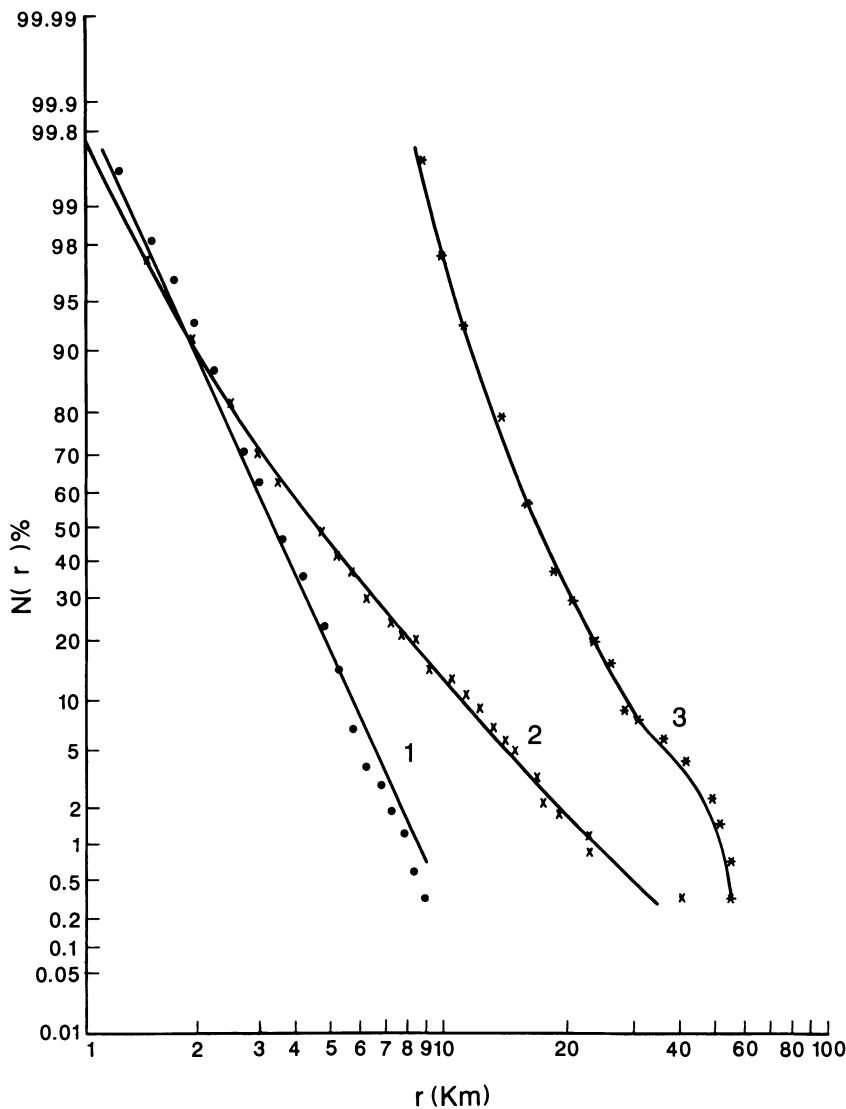


Fig. 5(b)

show up as straight lines on grid *a*, lognormal distributions on the grid *b*, and power-law distributions on the grid of Figures 5c and 6. Figure 7 presents the relationship between the perimeter *P* and the area *A* of the basement blocks, on a double logarithmic grid.

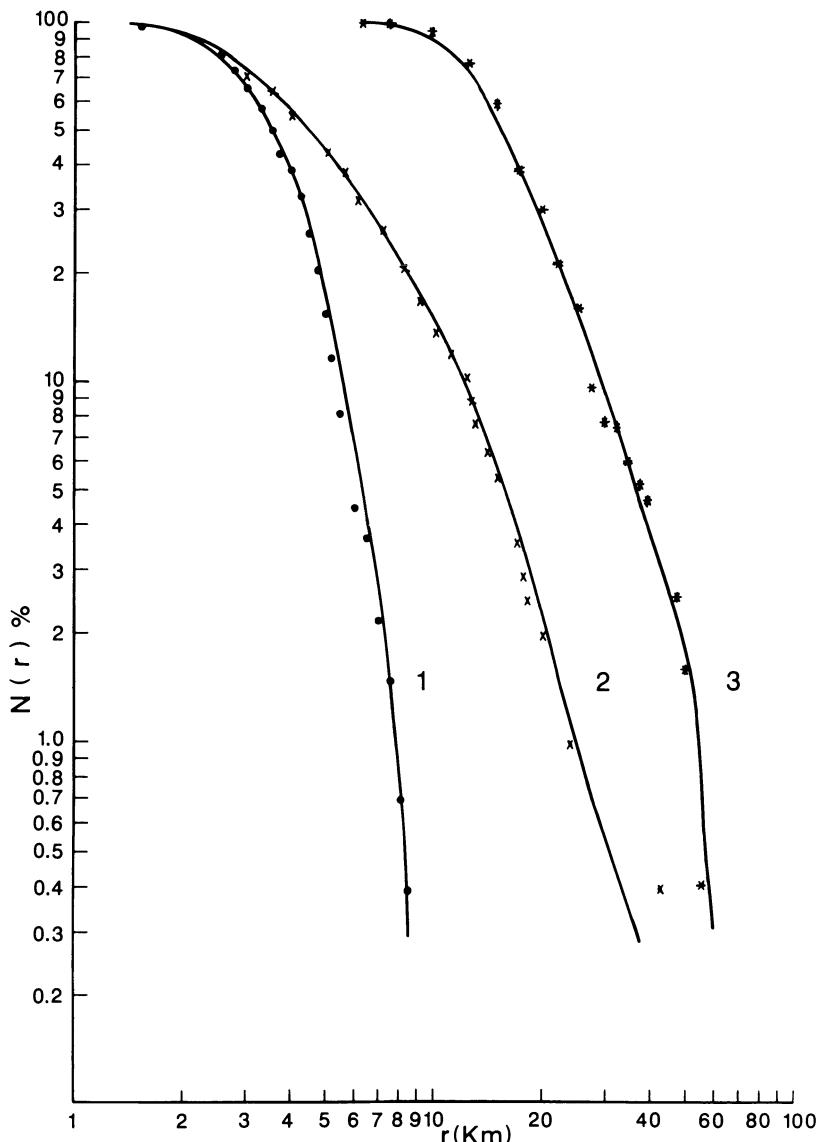


Fig. 5(c)

Figure 5

Empirical cumulative frequency curves $N(r)$ of the relative number of block sides greater than the size parameter r . Curve 1: lengths of the "cross trend" sides ($N20-70^{\circ}$ E); Curve 2: lengths of the "Gulf of Suez trend" sides ($N10-60^{\circ}$ W); Curve 3: perimeter of the basement blocks in case of Figures 5b and 5c, and quarter perimeters in Figure 5a. Figure 5a: semilogarithmic plot; Figure 5b: lognormal probability plot; Figure 5c: log-log plot. Exponential distributions show up as straight lines on grid a , lognormal distributions on grid b and power-law distributions on grid c .

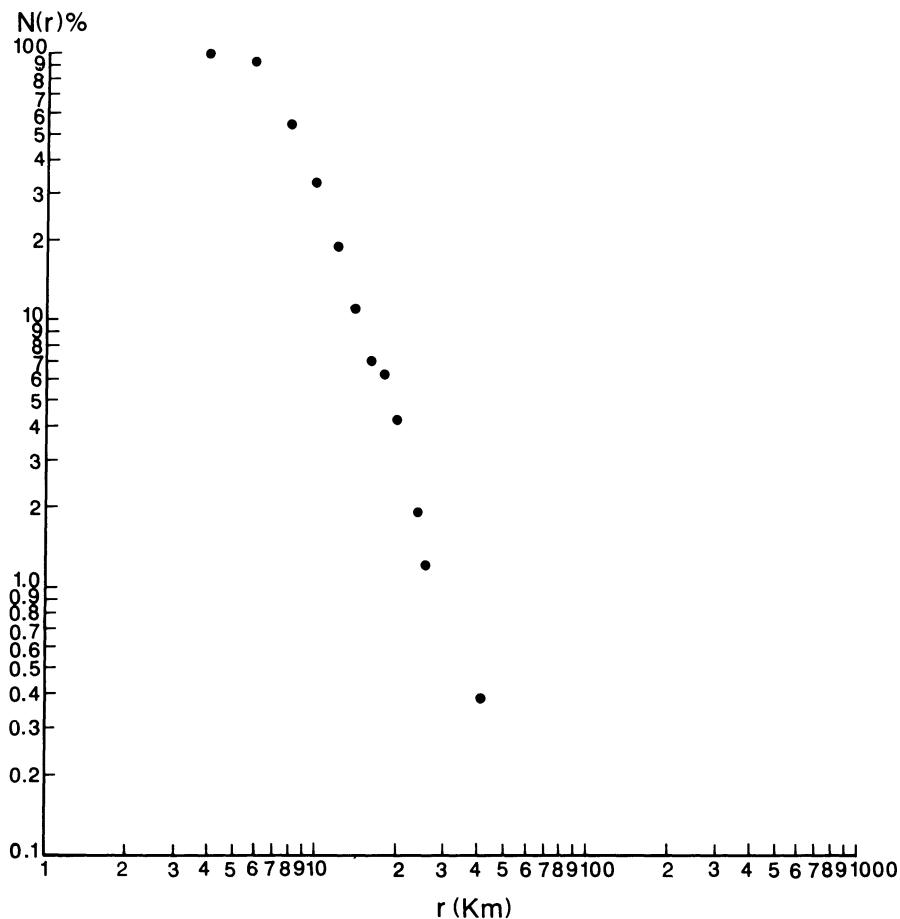


Figure 6

Empirical cumulative frequency curve of the relative number $N(r)$ of blocks whose sieve diameter (*i.e.*, diameter of the smallest circumscribed circle) is larger than r .

3. Discussion

No statistically significant conclusions can be made on the distribution of the size parameters because of the minimality of the data set. The size parameters do not span more than a decade in scale range, there are very few representatives of the largest features (as the few “outlying points” for $r > 20$ km in Figure 5a corresponding to the large elongated “islands” of the map) and the fine features are not resolved because of the discreteness of the grid both in geophysical measurements and in the computer modelling when constructing the map. As in crack networks usually only one or two generations of cracks occur and the very small-sized polygons are absent (DIONNE, 1971; PLUMMER and GOSTIN, 1981;

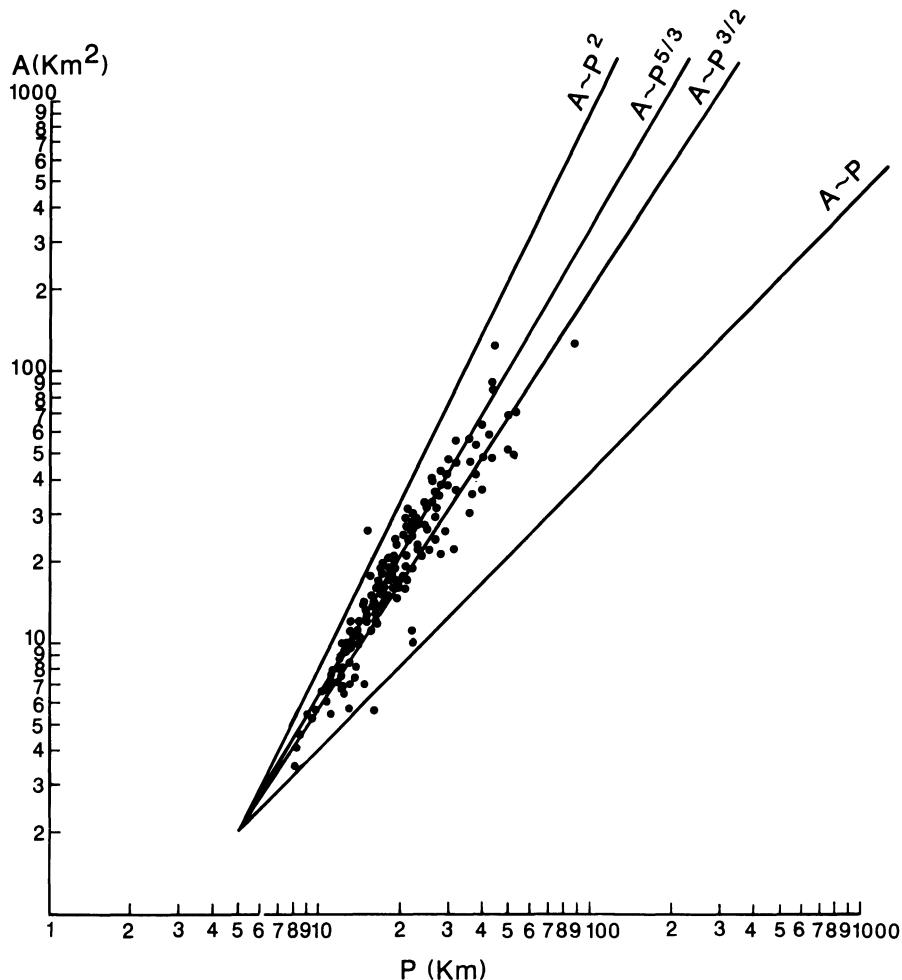


Figure 7

Area-perimeter relation for the Gulf of Suez basement blocks. The lines $A \sim P^2$, $P^{5/3}$, $P^{3/2}$ and P correspond, in turn, to the fractal dimensions $D = 1$, $6/5$, $4/3$ and 2 .

WALKER, 1986), the small sized blocks are possibly missing in the present case as well.

On the basis of Figures 5c and 6, and by visually inspecting Figure 3, which shows a uniform spacing between the fault lines (and does not resemble a scale-free Apollonian gasket associated with fractal fragmentation, cf. ROTHROCK and THORNDIKE, 1984; MATSUSHITA, 1985), the power-law distribution of the size parameters possibly can be ruled out. As a matter of fact, the power-law distribution corresponds to scale invariance, while contraction-crack polygons generally have a characteristic length related to the elastic properties and thickness of the

contracting medium. For example, in the giant desiccation polygons of the Great Basin playas, NEAL *et al.* (1968) established the empirical rule-of-thumb that fissure spacing (polygon width) is generally about 10 times the fissure depth.

Those lengths of sides of the blocks aligned in the direction of the Gulf of Suez axis ($N10\text{--}60^\circ\text{W}$) are exponentially distributed, $N(r) \sim \exp(-r/r_0)$, at least for $r < 20$ km (curve 2 in Figure 5a) with $r_0 = 4.66$ km. The distribution of the length of sides of direction $N20\text{--}70^\circ\text{E}$ is approximately lognormal (curve 1 in Figure 5b). The cumulative frequency distribution of the block perimeters (curve 3 in Figures 5a, b, c), of the "sieve diameter" (*i.e.*, diameter of the smallest circumscribed circle, Figure 6) and of the areas of the blocks (not shown) cannot be expressed analytically.

In connection with the possibly exponential distribution of the length of the ($N10\text{--}60^\circ\text{W}$)-oriented block sides let us recall a few fragmentation processes leading to exponential size distribution. GRIFFITH (1943) showed that if the energy consumed in breaking is proportional to the new surface formed, then Maxwell-Boltzmann energy partition leads to an exponential size distribution of the resulting particles. GILVARRY (1964) also derived the exponential distribution of fragment size in processes where the breakage proceeds along pre-existing Poisson-distributed volume-, facial-, and edge-flaws. Within the framework of GILVARRY's theory the value of $r_0 = 4.66$ km figuring in the size distribution $N(r) \sim \exp(-r/r_0)$ of curve 2 in Figure 5a would correspond to the mean spacing between the pre-existing Poisson-distributed flaws. The Poisson distribution of flaws prior to fragmentation is also in concord with the finding that most of the blocks are four-sided, as we know from statistical geometry that if a plane is dissected by Poisson-distributed random straight lines, the expected number of sides of the resulting polygons will be four (KENDALL and MORAN, 1963).

It is much more difficult to attach any physical significance to the possible lognormal distribution of the length of the ($N20\text{--}70^\circ\text{E}$)-directed block sides. Still, this finding calls for a reconsideration of the work of EPSTEIN (1947) who derived the lognormal size distribution of fragments under very general assumptions of the scale invariance and iterativeness of the breaking process; while today these very properties are solely associated with fractal (*i.e.*, power-law) fragmentation (ALLEGRE *et al.*, 1982; MANDELBROT, 1982; MATSUSHITA, 1985; CURL, 1986; TURCOTTE, 1986).

As Figure 7 shows, there is a fair correlation between the area of the blocks and a power of their perimeter with an exponent slightly less than two. It is well-known that for a set of random planar figures bounded by irregular curves of fractal dimension D , the area and perimeter of the figures are related as

$$P \sim \sqrt{A^D}. \quad (7)$$

(See *e.g.*, KENT and WONG, 1982 for lakes; LOVEJOY, 1982 and HENDERSON-SELLERS, 1986, for terrestrial and Martian clouds.)

Thus, the exponents 2, 5/3, 3/2, 1 indicated in Figure 7 correspond to perimeters of fractal dimension $D = 1$, $D = 6/5$, $D = 4/3$, $D = 2$, respectively. As most of the points cluster in the range $1.2 < D < 1.33$ corresponding to low fractal dimensions, this is a further indication that the fragmentation of the Gulf of Suez basement is *not* fractal. However, the very fact that a natural assemblage of convex polygons (*viz.*, the fault blocks of the Gulf of Suez basement), having a limited number of sides and very simple regular boundaries, could lead to an area-perimeter relation $A \sim P^n$ with $n < 2$ shows that the fractal nature of irregular boundaries cannot be established on the basis of the area-perimeter relation alone.

4. Conclusions

It is outside the scope of this paper to discuss the geological-tectonical implications of the size distributions found. I only wanted to present an example for a wide class of fragmentation processes occurring on all scales (the contraction-crack polygons) where exponential or lognormal fragment size distribution seems to prevail, rather than the power-law distribution associated with fractal fragmentation. It still seems premature to quote KORČAK (1940) (as CURL, 1986 does) that "only two fundamental types of statistical distributions exist" (*viz.*, the normal and the power-law). Even though the cases documented by HARTMANN (1969), ROTHROCK and THORNDIKE (1984), MATSUSHITA (1985) or TURCOTTE (1986) as being power-law are all fragmentations under compressive loading while the fracture pattern analysed in the present paper has clearly formed under biaxial tension (see Figure 4), in the axiomatic derivation of the power-law size distribution (MATSUSHITA, 1985; TURCOTTE, 1986) no explicit use is made of the compressive nature of the stress field. On the other hand, these derivations assume that "the pre-existing zones or planes of weakness where failure occurs exist on all scales" (TURCOTTE, 1986) while fragmentation theories leading to exponential size distribution start out from Poisson-distributed pre-existing sites of failure (GILVARRY, 1964). The axiomatic derivation of the lognormal size distribution in certain breakage mechanisms (EPSTEIN, 1947) is posing a most challenging problem as its basic assumptions (scale invariance and iterativeness of the breaking process) are strikingly similar to the conditions that are generally thought to characterise fractal fragmentation.

As there is an abundance of examples relevant to earth sciences for all three kinds of fragment-size distribution (see the papers cited in connection with Eqs. (1)-(3) and the recent review of KRANZ, 1983, who claims that the length of microcracks in rocks is exponentially or lognormally distributed), it would be of importance to clearly define those physical processes leading to the respective fragmentations.

Finally, it should be noted that certain rough seismic boundaries have been

shown to possess fractal topographies (BARENBLATT *et al.*, 1984). In an attempt to model the fractal relief of the earth, MANDELBROT (1975) started out from Poisson-distributed random straight lines dissecting a plane, in each case subjected the two sides to random vertical displacements in order to create "cliffs" and repeated this process *ad infinitum*. The fault network, and the vertical displacement of the blocks, observed in the Gulf of Suez Palaeozoic Basement, resemble an early stage of this random geomorphological process.

Acknowledgment

Thanks are due to Dr. Hammouda and to Professor A. Bayoumi of the Geophysical Department, Cairo University, for sending me a copy of Dr. Hammouda's Ph.D dissertation; and to Professor C. H. Scholz and to the two unknown referees who kindly pointed out the deficiencies in the original version of my paper.

REFERENCES

- ALLÈGRE C. J., LE MOUEL, J. L., and PROVOST, A. (1982), *Scaling Rules in Rock Fracture and Possible Implications for Earthquake Prediction*, Nature 297, 47–49.
- ALLEN, T., *Particle Size Measurement* (Chapman and Hall, London 1968).
- BADER, H. (1970), *The Hyperbolic Distribution of Particle Sizes*, Journal Geoph. Res. 75, 2822–2830.
- BARENBLATT, G. I.; ZHIVAGO, A. V., NEPROCHNOV, YU. P., and OSTROVSKIY, A. A. (1984), *Fractal Dimension: A Quantitive Characteristic of the Ocean Bottom Geomorphological Evolution Processes*, Okeonologiya 24, 924–928 (In Russian).
- BILLINGS, M. P., *Structural Geology* (Prentice Hall, New York 1954).
- BROWN, W. K., KARP, R. R., and GRADY, D. Z. (1983), *Fragmentation of the Universe*, Astrophys. and Space Science 94, 401–412.
- CURL, R. L. (1986), *Fractal Dimensions and Geometries of Caves*, Math. Geol. 18, 765–783.
- DIONNE, J. C. (1971), *Polygonal Patterns in Muddy Tidal Flats*, J. Sed. Petr. 41, 835–883.
- EPSTEIN, B. (1947), *The Mathematical Description of Certain Breakage Mechanisms Leading to the Logarithmico-normal Distribution*, J. Franklin Inst. 244, 471–477.
- FUJIWARA, A., KAMIMOTO, G., and TSUKAMOTO, A. (1977), *Destruction of Basaltic Bodies by High Velocity Impact*, ICARUS 31, 277–288.
- GILVARRY, J. J. (1964), *Fracture of Brittle Solids. I. Distribution Function for Fragment Size in Single Fracture. (Theoretical)*, J. Appl. Phys. 32, 391–399.
- GRIFFITH, L. (1943), *A Theory of the Size Distribution of Particles in a Comminuted System*, Can. J. Research 21A, 57–64.
- HAMMOUDA, H. M. (1986), *Study and Interpretation of Basement Structural Configuration in the Southern Part of Gulf of Suez using Aeromagnetic and Gravity data*, Ph.D. Thesis, Faculty of Science, Cairo University.
- HARTMANN, W. K. (1969), *Terrestrial, Lunar and Interplanetary Rock Fragmentation*, ICARUS 10, 201–213.
- HENDERSON-SELLERS, A. (1986), *Are Martian Clouds Fractals?* Quart. J. R. Astr. Soc. 27, 90–93.
- JARRIGE, J. J., D'ESTEVOU, P. O., BUROLLET, P. F., THIRIET, J. P., ICART, J. C., RICHERT, J. P., SEHANS, P., MONTENAT, C., and PRAT, P. (1986), *Inherited Discontinuities and Neogene Structure: The Gulf of Suez and the Northwestern Edge of the Red Sea*, Phil. Trans. R. Soc. Lond. A317, 129–139.
- KECK, C. K., *A Handbook on the Care of Paintings* (Am. Assoc. for State and Local History 1965).
- KENDALL, M. G., and MORAN, P. A. P., *Geometrical Probability* (Griffin and Co., London 1963).

- KENT, C., and WONG, J. (1982), *An Index of Littoral Zone Complexity and its Measurement*, Can. J. Fish. Aquat. Sci. 39, 847–853.
- KING, G. (1983), *The Accommodation of Large Strains in the Upper Lithosphere of the Earth and Other Solids by Self-similar Fault-systems: The Geometrical Origin of b-value*, Pure Appl. Geophys. 121, 761–815.
- KORČAK, J. (1940), *Deux types fondamentaux de distribution statistique*, Bull. Inst. Int. Stat. 30, 295–299.
- KRANZ, R. L. (1983), *Microcracks in rocks: A review*, Tectonophysics 100, 449–480.
- LACHENBRUCH, A. H. (1962), *Mechanics of Thermal Contraction Cracks and Ice-wedge Polygons in Permafrost*, Geol. Soc. Am. Spec. Paper 70, 69 pp.
- LOUIS, E., GUINEA, F., and FLORES, F., *The fractal nature of fracture*, In *Fractals in Physics* (ed. Pietronero, L. and Tosatti, E.) (Elsevier, Amsterdam-New York 1986) pp. 177–180.
- LOVEJOY, S. (1982), *Area-perimeter Relation for Rain and Cloud Areas*, Science 216, 185–187.
- LUNG, C. W., *Fractals and the fracture of cracked metals*, In *Fractals in Physics* (ed. Pietronero, L. and Tosatti, E.) (Elsevier, Amsterdam-New York 1986) pp. 189–192.
- MADDEN, T. R. (1983), *Microcrack Connectivity in Rocks: A Renormalisation Group Approach to the Critical Phenomena of Conduction and Failure in Crystalline Rocks*, J. Geophys. Res. 88, 585–592.
- MANDELBROT, B. B. (1975), *Stochastic models for the Earth's Relief, the Shape and the Fractal Dimension of the Coastlines, and the Number-area Rule for Islands*, Proc. Nat. Acad. Sci. USA 72, 3825–3828.
- MANDELBROT, B. B., *The Fractal Geometry of Nature* (Freeman, San Francisco 1982).
- MANDELBROT, B. B., PASSOJA, D. E., and PAULLAY, A. J. (1984), *Fractal Character of Fracture Surfaces of Metals*, Nature 308, 721–722.
- MATSUSHITA, M. (1985), *Fractal viewpoint of Fracture and Accretion*, J. Phys. Soc. Japan. 54, 857–860.
- NEAL, J. T., LANGER, A. M., and KERR, P. F. (1968), *Giant Desiccation Polygons of Great Basin Playas*, Geol. Soc. Amer. Bull. 79, 69–70.
- NELSON, R. A. (1979), *Natural Fracture Systems: Description and Classification*, Am. Assoc. Petr. Geologists Bull. 63, 2214–2232.
- OKUBO, P. G., and AKI, K. (1987), *Fractal Geometry in the San Andreas Fault System*, Journal Geophys. Res. 92B, 345–355.
- PLUMMER, P. S., and GOSTIN, V. A. (1981), *Shrinkage Cracks: Desiccation or Synaeresis?* J. Sed. Petr. 51, 1147–1156.
- ROTHROCK, D. A., and THORNDIKE, A. S. (1984), *Measuring the Sea Ice Floe Size Distribution*, J. Geophys. Res. 89C, 6477–6486.
- SAID, R., *The Geology of Egypt* (Elsevier, Amsterdam-New York 1962).
- SCHOLZ, Ch. H., and AVILES, C. A., *The fractal geometry of faults and faulting*, In *Earthquake Source Mechanics* (eds. Das, S., Boatwright, J. and Scholz, Ch. H.) (Geoph. Monograph 37, Maurice Ewing Volume 6, American Geoph. Union, Wash. D.C. 1986), pp. 147–155.
- SEGALL, P., and POLLARD, D. D. (1983), *Joint Formation in Granite Rock of the Sierra Nevada*, Geol. Soc. Am. Bull. 94, 563–575.
- SEGNIT, E. R., ANDERSON, C. A., and JONES, J. B. (1970), *A Scanning Microscope Study of the Morphology of Opal*, Search 1, 349–351.
- TERMONIA, Y., and MEAKIN, P. (1986), *Formation of Fractal Cracks in a Kinetic Fracture Model*, Nature 320, 429–431.
- THOMPSON, D'ARCY W., *On Growth and Form* (Cambridge University Press, Cambridge 1942).
- TURCOTTE, D. L. (1986), *Fractals and Fragmentation*, J. Geophys. Res. 91B, 1921–1926.
- WALKER, J. (1986), *Cracks in a Surface Look Intricately Random but Actually Develop Rather Systematically*, Sci. Am. 25 (4), 178–186.
- WILLDEN, R., and MABEY, D. R. (1961), *Giant Desiccation Fissures on the Black Rock and Smoke Creek Deserts, Nevada*, Science 133, 1359–1360.

(Received June 15, 1987, revised February 1, 1988, accepted August 31, 1988)

Fractal Analysis of Long-Range Paleoclimatic Data: Oxygen Isotope Record of Pacific Core V28-239

RICHARD H. FLUEGEMAN, JR.¹ and R. SCOTT SNOW¹

Abstract—*R/S* analysis of the oxygen isotope curve of Pacific core V28-239 yields a fractal dimension of 1.22. This value is considered to characterize global climatic change over the last 2 million years as expressed by changing δO^{18} ratios and confirms that climatic variations are characterized by long-term persistence. The fractal dimension of 1.22 compares favorably with the approximate fractal dimension of 1.26 for annual precipitation records for nine major cities in the United States. Although the precipitation and oxygen isotope data are measured in different physical units and recorded at different time scales, fractal analysis allows for a mathematical comparison of the two phenomena. Additionally, since the fractal dimensions of the oxygen isotope and precipitation records are similar, it is implied that such fractal dimensions are characteristic of climate change over the spectral range of 10 to 10^6 years. Given this, temperature curves based on fractal parameters of long-term δO^{18} data could be constructed which would allow examination of characteristics of temperature variation over tens and hundreds of years. Such studies may allow the establishment of limits on natural temperature variation and document the persistence of temperature trends through time. If these trends and limits can be resolved, long-range climatic prediction is feasible.

Key words: *R/S* analysis, paleoclimate, oxygen isotopes, Pacific, comparison of phenomena, long-term persistence, fractal.

Introduction

Accurate long-term climatic forecasting would be of great value to society. Such predictions, if reliable, would make economic and civil planning practical and eliminate costly preparations which may be unnecessary. Although many “long-term” forecasts based on observational and deterministic meteorological approaches are available, the accuracy limit is about 90 days (AHERNS, 1985, p. 63). Such methods provide no lead time to prepare entire populations for truly long-term effects of climatic change such as rising sea level, changes in annual mean temperature, and changes in precipitation amounts. These changes occur over many decades and if preparations to meet changes such as these ever become necessary, minimum advance warning of several decades would be needed.

Studies of man-induced climate change, whether observational or predictive,

¹ Department of Geology, Ball State University, Muncie, Indiana 47306, U.S.A.

must assume a baseline of natural fluctuations in precipitation and temperature extending over many decades. Historical records provide only short, inadequate models for such applications. Recently need has arisen for even longer term climatic prediction. Sites considered for terminal storage of high-level radioactive waste must be evaluated to determine potential climate changes over decay lives of radionuclides that span thousands of years. It is necessary to predict potential ranges of climatic fluctuations and determine the effects of those climate changes on hydrology and erosion of geological deposits. The methods of paleoclimatology must provide the data for investigations like these.

The work of EMILIANI (1955) demonstrated the use of oxygen isotope ratios on foraminiferal ooze from Atlantic and Caribbean cores to determine Pleistocene paleotemperatures. Such techniques have subsequently been employed by many workers studying the foraminiferal record of deep sea cores. That the oxygen isotope record from deep sea cores might give some insight to future climatic trends was shown by CHAPPELL (1973). In this paper, Chappell was able to relate various climatically controlled variables, including Emiliani's δO^{18} temperature curve, to the Milankovitch hypothesis. This states that changes in insolation brought about by the earth's orbital perturbations have controlled the incidence of glacial episodes during the Pleistocene. Since glacial episodes have recurred during the Pleistocene, the characteristics of the δO^{18} curve during glacials and interglacials could be used to make long-term predictions of climatic trends at the present based on the type of episode we are in, the amount of time elapsed since the beginning of the episode, and the likely climatic characteristics associated with the part of the episode remaining.

A problem with using δO^{18} or any other paleoclimatic data derived from deep sea cores to make statements about future climatic trends is resolution. The sampling interval in most paleoclimatic studies is 5 or 10 cm. This may represent an interval of time of between 1,000 to 10,000 years, depending on sedimentation rates. Additionally, vertical mixing of sediment in deep sea cores may place a real minimum limit of oxygen isotope event resolution at 1000 years (RUDDIMAN and GLOVER, 1972). If any climatic predictions are to be meaningful to the population in general, they must be expressed in terms of a human lifetime, that is decades or perhaps one century.

Fractal geometry may allow the characteristics of a δO^{18} curve at a given sample interval (in time) to be used to reconstruct the characteristics of a curve with much finer resolution. While natural fractal shapes and phenomena are typically interpreted as inhomogeneous fractals (LA BRECQUE, 1986), there are cases in nature where constant values of fractal dimension are found to extend over many orders of magnitude of relevant scale (*i.e.*, clouds in LOVEJOY, 1982). If climate fluctuations over time spans ranging from decades to hundreds of thousands of years can be characterized by a single, general fractal dimension, then the longer term fluctuations become models for the shorter term ones. Thus while an absolute curve

may not be developed for finer resolution scales the characteristics of this curve are known. In the case of the δO^{18} curve, exact data for a given point in time will not be obtained but the degree of irregularity and long-term persistence of the climatic trends can be derived for a given time span that is of interest.

The purpose of this study is to examine a specific δO^{18} curve derived from a deep sea core and to determine the fractal dimension of this curve. If the obtained dimension is significantly greater than 1.0 and less than 1.5, then a fractal model of long-term persistence is uniquely appropriate for this data set. If the dimension approximates ones obtained by others from climatic measurements at much shorter time scales, then there is reason to assume that climate fluctuations are characterized by one general fractal dimension across the intervening range of time spans as well.

The Data

For our study we have selected oxygen isotope data derived from Pacific core V28-239 from the Solomon Rise and published by SHACKLETON and OPDYKE (1976). The core is located at lat. $3^{\circ}15'N$, long. $150^{\circ}11'E$ and is from a depth of 3490 m. The sample interval was 5 cm and was calculated to represent about 5000 years. Data values were given from 408 of the 420 samples studied. Where values were not provided, we estimated the δO^{18} value from the published oxygen curve for core V28-239 (SHACKLETON and OPDYKE, 1976, pp. 450-451).

The oxygen isotope data from core V28-239 is an excellent subject for a representative fractal analysis of Pleistocene to recent global climatic trends for several reasons. First, a well preserved carbonate fraction made it possible for SHACKLETON and OPDYKE to obtain a fairly complete δO^{18} record. Second, excellent paleomagnetic control makes calibration with the geochronologic scale possible. Third, a relatively long record of geologic history (2.1 m.y.) is represented by the core. Fourth, a known rate of sedimentation throughout the core gives the constant sample interval a constant time value (5 cm = 5000 years). This is important because of the assumption of an approximately even sampling interval in the R/S analysis discussed below. Although the Cenozoic geochronologic scale has recently been revised by BERGGREN *et al.* (1985), the calculated sedimentation rate does not change significantly with the recalibration.

The oxygen isotope curve of core V28-239 was figured by HARLAND *et al.* (1982) as a standard for Pleistocene climatic stratigraphy. The fact that core V28-239 is a Pacific core is also important as the effect of local temperature change is reduced and the isotopic changes reflect glacially induced global climatic trends (SHACKLETON and OPDYKE, 1973). The oxygen isotope record of core V28-239 correlates closely with that from the nearby core V28-238. Core V28-238 has been proposed as a global stratotype for the Upper Pleistocene (SHACKLETON and OPDYKE, 1973, p. 48).

Methods

The set of oxygen isotope values were subjected to R/S analysis (rescaled range analysis), according to the method specified by MANDELBROT and WALLIS (1969a). This results in a logarithmic plot of R/S values versus time spans for which those values are derived. The plot is asymptotic to a constant slope H with increasing time span (increasing number of data values included). For any particular time span except the longest ones, many nonoverlapping subsets of the total data set are available for analysis, providing a range of R/S values. We have taken the mean of these various values as representative for each specific time span.

MANDELBROT (1983) indicates that the exponent H , if greater than 0.5 and less than 1.0, is related to the fractal dimension D by,

$$D = 2 - H$$

giving values of D greater than 1.0 and less than 1.5. Records with such values of D and H have some degree of long-term persistence: even those measured values separated by very long time intervals are not completely independent. MANDELBROT and WALLIS (1969a) have shown this to be a common characteristic of a variety of geophysical records.

Results

The results of fractal analysis of the core V28-239 oxygen isotope data are represented graphically in Figure 1. The calculated value for H is 0.78 with a fractal dimension of 1.22. A slight deviation from the best-fit line is seen for measurements

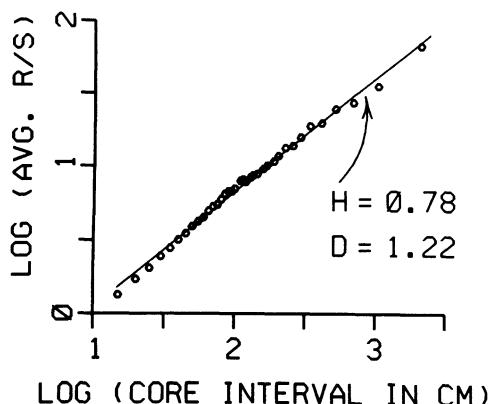


Figure 1

Plot of average rescaled range (R/S) values versus core depth, oxygen isotope record of Pacific core V28-239. Sedimentation rate is approximately 1 cm per 1000 years. Plotted line represents a linear regression on all data points.

taken over short time spans, up to a time span of approximately 40,000 years (equivalent to approximately 8 successive data values). Such an "initial transient" range is apparent in several analyses of geophysical records conducted by MANDELBROT and WALLIS (1969a). Some of the transients in that study extended to time spans encompassing 20 successive data values, and in all cases the transients were treated as artifacts of the analysis method rather than indicators of real change in fractal dimension at shorter time scales. Similar effects were obtained in analysis of records of artificially generated noise (MANDELBROT and WALLIS, 1969b). Several breaks in the trendline are visible, the most apparent one corresponding to a time interval of approximately 100,000 years. A similar break identified by MANDELBROT and WALLIS (1969a) in the *R/S* analysis of monthly sunspot activity records is associated with the well-known sunspot cyclicity of 11 years. It is inferred that the prominent break noted here is the result of a climatic cycle with a period of approximately 100,000 years. Overall, these breaks along with the initial transient do not obscure the strongly linear plot trend that indicates a consistent *H* value (*H* = 0.78).

The results of the analysis therefore indicate that, for time spans (time scales) 5000 to 2,000,000 years in duration a fractal dimension of about 1.22 characterizes global climate change over the last 2 m.y. as expressed in changing δO^{18} ratios. This confirms that climate variations are characterized by long-term persistence, and that the fractal model is specially applicable.

Discussion

The fractal nature of the oxygen isotope curve of deep sea core V28-239 provides two major opportunities for the study of paleoclimates. First, the fractal dimension provides an index by which oxygen isotope records can be compared to other phenomena related to climate.

The available records of climate change on different time scales come from a variety of data sources, with results expressed in quite different units. To the degree that each data source (tree rings, for example) directly expresses climate change, there should be a basis for comparison with other data sources. But even though there may be a relationship based on a common control, it is not easy to objectively compare phenomena with physical characteristics so different.

Fractal analysis provides a basis for comparison which is not only independent of particular units of measure, but also allows comparison of irregularity characteristics of records with vastly different time scales. For example, our results from δO^{18} ratios can be compared with those results obtained for local United States annual precipitation records from between 1820 and 1963 (MANDELBROT and WALLIS, 1969a, p. 339). For nine major cities, the majority being in the eastern U.S., precipitation data had various *H* values ranging from 0.64 to 0.89, with a mean

value of 0.74 corresponding to a fractal dimension of 1.26. Although we are looking at an indicator of temperatures (and ice volume) over relatively long time spans on one hand, and at precipitation over relatively short time spans on the other, the fractal dimension values indicate that the irregularity characteristics of these records are very similar. The fractal analysis gives a mathematical basis for a comparison and will help make resolution of obscure relationships between phenomena possible.

The second opportunity fractal analysis of oxygen isotope curves provides to the paleoclimatologist is the possibility of fine resolution. If a certain curve has a fractal dimension that can reasonably be extrapolated beyond the limits of the data set, then it is possible to reconstruct the oxygen isotope curve at smaller scales than the sample interval.

In the case of core V28-239, the sample interval is 5 cm, roughly equivalent to 5000 years. Because analyses of the core data and the precipitation data noted above yield similar fractal dimensions, we can come to a preliminary conclusion that such D values are characteristic of climate change over the spectral range 10 to 10^6 years. In the case of the oxygen isotope data, randomized fractal functions can be developed to mimic the general behavior of the empirical record and these same functions can be in effect superimposed at smaller time scales to model shorter term variations in temperature.

Rather than precipitation records, a more direct basis for extrapolation of our observed D to smaller time scales would be the detailed temperature records which have been kept during the last century. If the modern data can be found to have characteristics similar to those of the "extrapolated" oxygen isotope data, it is reasonable to associate the same general temperature variation characteristics to intervening time scales as well.

The utility of such studies may be to establish limits of natural temperature variation and persistence of temperature trends for long-range climatic prediction with a scale of decades rather than centuries or millenia. If such trends can be resolved, then informed long-range climatic prediction, based on the paleotemperature record of the Pleistocene and modern temperature records is feasible.

Summary

1. R/S analysis of the oxygen isotope curve of Pacific core V28-239 has resulted in an estimated fractal dimension of 1.22 for the recorded climate changes over time spans from several thousand years to two million years.

2. As shown by an example comparing the oxygen isotope results to analyses of historical precipitation records by others, the fractal dimension is a useful tool in making comparisons between phenomena with different time scales and different physical characteristics.

3. The possibility of creating expanded time-scale-range paleotemperature curves using a constant fractal dimension may make accurate prediction of some important characteristics of long-range climatic change possible.

REFERENCES

- AHRENS, C. D., *Meteorology Today* (West Publishing, St. Paul 1985).
- BERGGREN, W. A., KENT, D. V., FLYNN, J. J., and VAN COUVERING, J. A., (1985), *Cenozoic Geochronology*, Geol. Soc. America Bull. 96, 1407-1418.
- CHAPPELL, J. (1973), *Astronomical Theory of Climatic Change: Status and Problem*, Quaternary Research 3, 221-236.
- EMILIANI, C. (1955), *Pleistocene Temperatures*, J. Geology 63, 538-578.
- HARLAND, W. B., COX, A. V., LLEWELLYN, P. G., PICKTON, C. A. G., SMITH, A. G., and WALTERS, R., *A Geologic Time Scale* (Cambridge University Press 1982).
- LA BRECQUE, M. (1986), *Fractal Applications*, National Science Foundation Mosaic 17(4), 34-48.
- LOVEJOY, S. (1982), *Area-Perimeter Relation for Rain and Cloud Areas*, Science 216, 185-187.
- MANDELBROT, B. B., and WALLIS, J. R. (1969a), *Some Long-run Properties of Geophysical Records*, Water Resources Res. 5, 321-340.
- MANDELBROT, B. B., and WALLIS, J. R. (1969b), *Robustness of the Rescaled Range R/S in the Measurement of Noncyclic Long-run Statistical Dependence*, Water Resources Res. 5, 967-988.
- MANDELBROT, B. B., *The Fractal Geometry of Nature* (Freeman and Co., New York 1983).
- RUDDIMAN, W. F., and GLOVER, L. K. (1972), *Vertical Mixing of Ice-raftered Volcanic Ash in North Atlantic Sediments*, Geol. Soc. America Bull. 83, 2817-2836.
- SHACKLETON, N. J., and OPPDYKE, N. D. (1973), *Oxygen Isotope and Paleomagnetic Stratigraphy of Equatorial Pacific Core V28-238: Oxygen Isotope Temperatures and Ice Volumes on a 10^5 Year and 10^6 Year Scale*, Quaternary Research 3, 39-55.
- SHACKLETON, N. J., and OPPDYKE, N. D. (1976), *Oxygen isotope and paleomagnetic stratigraphy of Pacific Core V28-239 late Pliocene to latest Pleistocene*, In *Investigation of Late Quaternary Pale-oceanography and Paleoclimatology* (eds. Cline, R. M., and Hays, J. D.) Geol. Soc. America Memoir 145, 449-464.

(Received May 18, 1987, revised/accepted October 1, 1988)

PAGEOPH

Reprints from *Pure and Applied Geophysics*

New in the series:

Subduction Zones

Edited by

Larry J. Ruff

Department of Geological Sciences, University of Michigan, Ann Arbor, MI, USA

Hiroo Kanamori

Seismological Laboratory, California Institute of Technology, Pasadena, CA, USA

Part I:

1988. 352 pages. Paperback

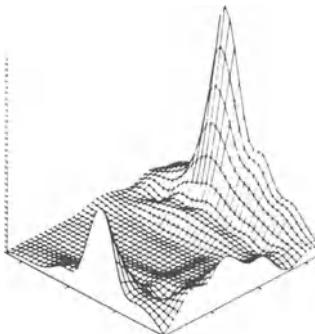
ISBN 3-7643-1928-3

Part II:

1989. 288 pages. Paperback

ISBN 3-7643-2272-1

Subduction Zones is a diverse collection of contributions that portray many different facets of the subduction process. Accretionary prism dynamics, subduction initiation, deep mantle penetration of subducted lithosphere, continental crust recycling, and seismicity are some of the topics discussed in the two volumes of *Subduction Zones*. Volume I consists of two sections, *Global-scale reviews* and *Accretionary prism processes*, while Volume II contains the two sections *Tectonics & subduction initiation* and *Earthquake occurrence*.



Seismicity in Mines

Edited by

S. J. Gibowicz

Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland

1989. 398 pages,

Paperback

ISBN 3-7643-2273-X

In this collection of 22 papers, new observations describing seismicity patterns and source mechanism of seismic events induced by mining are reported. The methods describing location of mine tremors with accuracy of the order of a few tens of meters are also included, as well as seismic tomography and its applications in mines. The research on rockbursts undertaken recently in Canada, and recent progress in research on seismicity in mines in South Africa are described in some detail.

Middle Atmosphere

Edited by

Alan R. Plumb

MIT, Cambridge, MA, USA

Robert A. Vincent

University of Adelaide, Australia

1989. 472 pages.

Paperback

ISBN 3-7643-2290-X

This book presents an up-to-date summary of current research into the dynamics of the middle atmosphere, addressing an ever-increasing need to understand the complex interplay of processes which control the distribution of atmospheric ozone and the impact of stratospheric pollutants on ozone. Special emphasis is given to studies of the southern hemisphere.

Most of the papers contained here were presented at workshops held in Adelaide, South Australia, in May 1987.

Please order through your bookseller or Birkhäuser Verlag, P.O. Box 133, CH-4010 Basel / Switzerland or, for orders originating from the USA and Canada, through Birkhäuser Boston, Inc., c/o Springer-Verlag New York, Inc., 44 Hartz Way, Secaucus, NJ 07096-2491 / USA

Springer
Basel AG

