

Level sets and extrema of random processes and fields.

Jean-Marc Azaïs

Mario Wschebor

UNIVERSITÉ DE TOULOUSE UPS CNRS : INSTITUT DE MATHÉMATIQUES LABORATOIRE
DE STATISTIQUES ET PROBABILITÉS 118 ROUTE DE NARBONNE, 31062 TOULOUSE CEDEX 09,
FRANCE

CENTRO DE MATEMÁTICA. FACULTAD DE CIENCIAS., UNIVERSIDAD DE LA REPÚBLICA.,
CALLE IGUÁ 4225. 11400 MONTEVIDEO. URUGUAY.

E-mail address: `azais@cict.fr`

E-mail address: `wschebor@cmat.edu.uy`

Contents

Introduction	5
Reading diagram	12
Chapter 1. Classical results on the regularity of the paths.	15
1. Kolmogorov's Extension Theorem.	15
2. Reminder on the Normal Distribution.	17
3. 0-1 law for Gaussian processes.	20
4. Regularity of the paths.	22
Exercises	34
Chapter 2. Basic Inequalities for Gaussian Processes	39
1. Slepian type inequalities.	39
2. Ehrhard's inequality.	42
3. Gaussian isoperimetric inequality.	45
4. Inequalities for the tails of the distribution of the supremum.	45
5. Dudley's inequality.	51
Exercises	54
Chapter 3. Crossings and Rice formulas for 1-dimensional parameter processes	55
1. Rice Formulas	55
2. Variants and Examples.	62
Exercises	67
Chapter 4. Some Statistical Applications	73
1. Elementary bounds for $P\{M > u\}$	73
2. More detailed computation of the first two moments	78
3. Maximum of the absolute value	81
4. Application to quantitative gene detection	82
5. Mixtures of Gaussian distributions	92
Exercises	99
Chapter 5. The Rice Series	101
1. The Rice Series	102
2. Computation of Moments.	106
3. Numerical aspects of Rice Series.	111
4. Processes with Continuous Paths	116
Chapter 6. Rice formulas for random fields	119
1. Random fields from \mathbb{R}^d to \mathbb{R}^d .	119
2. Random fields from \mathbb{R}^d to $\mathbb{R}^{d'}$, $d > d'$	130
Exercises	134
Chapter 7. Regularity of the Distribution of the Maximum	137
1. The implicit formula for the density of the maximum	138
2. One parameter processes	140
3. Continuity of the density of the maximum of random fields	148
Exercises	150

Chapter 8. The tail of the distribution of the maximum	153
1. One-dimensional parameter: asymptotic behavior of the derivatives of F_M	154
2. An Application to Unbounded Processes.	157
3. A general bound for p_M	160
4. Computing $\bar{p}(x)$ for stationary isotropic Gaussian fields	161
5. Asymptotics as $x \rightarrow +\infty$	166
6. Examples	174
Exercises	178
Chapter 9. The record method	181
1. Smooth processes with one dimensional parameter	181
2. Non-smooth Gaussian processes	184
3. Two-parameter Gaussian processes	185
Exercises	188
Chapter 10. Asymptotic methods for infinite time horizon	189
1. Poisson character of “high” up-crossings	189
2. Central limit theorem for non-linear functionals	195
Exercises	206
Chapter 11. Geometric characteristics of random sea-waves	209
1. Gaussian model for infinitely deep sea	209
2. Some geometric characteristics of waves	211
3. Level curves, crests and velocities for space waves	212
4. Real Data	216
5. Generalizations of the Gaussian model	218
Exercises	219
Chapter 12. Systems of random equations	221
1. The Shub-Smale model	222
2. More general models	229
3. Non-centered systems (smoothed analysis)	235
4. Systems having a law invariant under orthogonal transformations and translations	243
Chapter 13. Random fields and condition numbers of random matrices	249
1. Condition numbers of non-Gaussian matrices	251
2. Condition numbers of centered Gaussian matrices	255
3. Non-centered Gaussian matrices	266
Notations	273
References	275
Index	291

Introduction

The theory of stochastic processes is a powerful tool to study a vast diversity of problems in the natural sciences, in which randomness is required as a component to describe phenomena. In this book we consider two general classes of problems arising when studying random models.

Let $\mathcal{X} = \{X(t) : t \in T\}$ be a stochastic process with parameter set T defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We will be mainly interested in the following subjects:

1) For each value u in the range space of $X(\cdot)$, understanding the properties of the level sets of the paths of \mathcal{X} , that is, the random sets $\{t \in T : X(t) = u\}$;

2) whenever the process has real values, studying and computing the distribution function of the random variable $M_T = \sup_{t \in T} X(t)$, that is, the function $F_{M_T}(u) = \mathbb{P}(M_T \leq u)$, $u \in \mathbb{R}$.

These are classical subjects in Probability Theory and have been considered for a long time in a variety of contexts. Generally speaking, our framework will be continuous parameter processes, which means here that T is a subset of the real line (such as a finite or infinite interval) or of some higher dimensional Euclidean space. When the parameter set T is multidimensional, we will call \mathcal{X} a “random field”.

In most of the theory and the applications that we will consider, the parameter set T will have some geometric regularity, such as being a manifold embedded in a finite dimensional Euclidean space or having some more general structure. As for the paths $t \rightsquigarrow X(t)$ of the stochastic function, we will require that they satisfy regularity properties, such as differentiability of a certain order. We will also need results on the supremum of random sequences, in which the geometry of the domain or the regularity of the paths does not play any role. This will provide us of basic and useful ingredients (such as in Chapter 2) but the emphasis will be on random functions possessing certain regularities.

For random level sets, our main tools are Rice Formulas. Assume that T is a Borel subset of the Euclidean space \mathbb{R}^d and \mathcal{X} a stochastic process or field having regular paths, defined on some open set containing T and taking values in \mathbb{R}^d . For given $u \in \mathbb{R}^d$ denote by $N_u(X; T)$ the number of roots of $X(t) = u$ lying in the set T . Rice Formulas allow one to express the k -th factorial moment of the random variable $N_u(X; T)$ as an integral over T^k of a some function, which depends upon the joint distribution of the process and its derivative and is evaluated at the k -tuples $(t_1, \dots, t_k) \in T^k$.

In fact, the main interest lies in the probability distribution of the random variable $N_u(X, T)$, which remains unknown: the authors are not aware of any non-trivial situation in which one can compute this distribution by means of some reasonable formula. Rice formulas appear to be a contribution to understand the distribution of $N_u(X, T)$, giving a general expression for its moments. One can measure to what extent this remains an open subject by the fact that no useful formula exists - for the time being - to compute the expectation of $f[N_u(X, T)]$ where $f : \mathbb{R} \rightarrow \mathbb{R}$ is some simple function, as for example $f(x) = x^\alpha$ when α is not an integer.

When the dimension d of the domain is strictly larger than that of the range space, say d' , generically the random level set has dimension $d - d'$ so that the interesting questions do not refer to the number of points lying in a set, but should aim to understand its geometry, which is of course richer than in the $d = d'$ case. The natural questions concern the probabilistic properties of the geometric measure of the level set, its Euler-Poincaré characteristic and so on. We give expressions and the corresponding proofs for the moments of the geometric measure of the level set under quite general conditions. Even though this is generally considered to be known, to our knowledge rigorous proofs are only available in special cases, such as real-valued random fields or particular probability laws. Chapter 11 presents some applications to sea-waves. We should say that the results we present are only a minor part of an important subject which, however, remains quite unexplored.

The name of Rice's formula honors the pioneering work of S.O. Rice, who computed the expectation of $N_u(X; T)$ for one-parameter Gaussian stationary processes in the 1940's and used it in telecommunications in an early stage of this discipline (this is formula 3.2 of Chapter 3 below). In the 1950's and 60's one can find applications of Rice's formula to physical oceanography (for example, in the series of papers by M.S. Longuet-Higgins and collaborators on the statistical properties of ocean waves, including some useful formulas in the multiparameter case). Applications to other areas of random mechanics were also developed somewhat later (see the book by Krée and Soize, 1983).

H. Cramér and M.R. Leadbetter's excellent book contains a systematic presentation of what was known by the time it was published (1967), together with connected subjects in Probability Theory and various applications. The book is on one-parameter processes and most of the material concerns Gaussian stationary processes. We still use it as a reference on various subjects.

Around 1970, for one-parameter Gaussian processes rigorous proofs of Rice formulas had already been given for certain classes of processes, with the contribution among others of K. Ito (1964), D. Ylvisaker (1965) and Y. Belayev (1966, 1972b). Some of this work included non-Gaussian processes and the multiparameter case but, to our knowledge, the first treatment of the multiparameter case in book form is Adler's "Geometry of Random Fields" (1981).

Our aim is to make a contribution to update the subject of Rice Formulas, including the improvements that have taken place during the last decades both in the basic theory and its applications.

There is a part of Probability Theory which refers to level sets of random functions which are not differentiable. This may have started with Paul Lévy's definition of local time (see for example Ikeda and Watanabe (1981) or Karatzas and Shreeve (1998) for a modern presentation of the subject) and has led to the study of the geometry of level sets of semimartingales or some other classes for one-parameter processes and similar problems for random fields. A short list of references on this subject can also consider the books by Kahane (1985) and Revuz-Yor (1998), and the paper by Ehm (1981). We are not considering this kind of problems in this book, our processes and fields have regular paths and for fixed height, almost surely, the level sets are nice sets. One can build up a bridge between these two worlds: take a process with non-differentiable paths, smooth it by means of some device, and try to reconstruct relevant properties of the geometry of the level sets of the original (irregular) paths, from the regularized ones. This leads to asymptotic expansions which are interesting by themselves and have important applications (See for example Jacod (1998, 2000) for polygonal and related approximations and Wschebor (2006) for a review without proofs).

With respect to the second main subject of this book, the general situation is that the computation of the distribution function of the supremum by means of a closed formula is known only for a very restricted number of stochastic processes (and trivial functions of them). The following

is a list of one-dimensional parameter processes for which - as far as the authors know - an actual formula exists for the distribution of $M = M_{[0,T]}$:

- The Brownian Motion or Wiener process $\{W(t) : t \geq 0\}$, for which the distribution of M has in fact been known since the 19 th century (Kelvin, D. André).
- The Brownian Bridge, $B(t) := W(t) - tW(1)$ ($0 \leq t \leq 1$).
- $B(t) - \int_0^1 B(s)ds$ (Darling, 1983).
- The Brownian Motion with a linear drift (Malmquist, 1954, Shepp, 1979).
- $\int_0^t W(s)ds + yt$ (McKean, 1963, Goldman, 1971, Lachal, 1991).
- The restriction to the boundary of the unit square of the Wiener sheet (Paranjape, Park, 1973)
- Each one of the stationary Gaussian processes with covariance equal to:
 - $\Gamma(t) = e^{-|t|}$ (Ornstein-Uhlenbeck process, DeLong, 1981),
 - $\Gamma(t) = (1 - |t|)^+$, T a positive integer (Slepian process, Slepian 1961, Shepp, 1971),
 - $\Gamma(t)$ even, periodic with with period 2, $\Gamma(t) = 1 - \alpha|t|$ for $0 \leq |t| \leq 1$, $0 < \alpha \leq 2$, (Shepp and Slepian 1976),
 - $\Gamma(t) = (1 - |t|/(1 - \beta)) \vee (-\beta/(1 - \beta))$, $|t| < (1 - \beta)/\beta$, $0 < \beta \leq 1/2$, $T = (1 - \beta)/\beta$ (Cressie 1980),
 - $\Gamma(t) = \cos t$ (Berman 1971b, Delmas 2003),
 - $\Gamma(t) = [2|t| - 1]^2$, $T = 1$ (Cabaña, 1991).

The methods to find formulas for the distribution of the supremum over an interval of this list of processes, are ad hoc, hence non transposable to more general random functions, even in the Gaussian context. Given the interest in the distribution of the random variable M_T , arising in a diversity of theoretical and technical questions, a large body of mathematics has been developed beyond these particular formulas.

A first way has been to obtain general inequalities for the distribution function $F_{M_T}(u)$. Of course, if one can not compute, then one tries to get upper or lower bounds. This is the subject of Chapter 2, which concerns Gaussian processes. The inequalities therein are essential starting points for the remainder of the book: a good part of the theory in the Gaussian case depends on these results.

However, generally speaking the situation is that these inequalities, when applied to the computation of the distribution of the random variable M_T , have two drawbacks. First, the bounds depend on certain parameters for which it is hard or impossible to obtain sharp estimates, implying that the actual computation of probabilities can become inaccurate or plainly useless, for statistical or other purposes. Second, these bounds hold for general classes of random functions, but may become rough when applied to a particular stochastic process or field. Hence, a crucial question is to improve the estimations derived from the general theory contained in Chapter 2. This is one of the purposes of this book, which is attained to a certain extent but at the same time, leaves a good deal of open problems. This question is considered in Chapters 4, 5, 8 and 9.

A second way has been to describe the behavior of $F_{M_T}(u)$ under various asymptotics. This subject is considered in several chapters. Especially, a part of Chapter 8 is devoted to the asymptotic behavior of the tail $1 - F_{M_T}(u)$ for a random field defined on a fixed domain as $u \rightarrow +\infty$. For extensions and a diversity of results concerning asymptotic behavior as $u \rightarrow +\infty$ that are not mentioned here, we refer the reader to the books (and the references therein) by Berman (1992 a), Lifshits (1995) and Piterbarg (1996).

The third way consists of studying the regularity of the function $u \rightsquigarrow F_{M_T}(u)$. This is considered in Chapter 7.

We have attempted to make the book as self-contained as possible. The reader is expected to have attained at least the level of a postgraduate student with basic training in Probability Theory and Analysis, including some elementary Fourier Analysis. However, given the intention to limit its total size, in certain points we have not respected the rule of being self-contained. Here is a list of what we consider the most relevant cases in which we use ideas and results and do not give proofs: Kolmogorov's extension theorem (Chapter 1); the definition and main properties of the Itô integrals and its basic properties, which are used in the first two chapters, including Itô's formula, the quadratic variation and the exponential martingales; the convergence of empirical measures and asymptotic methods in Statistics, of which we give a quick account without proofs of the results we need in Chapter 4; the co-area formula (Chapter 6) from integral geometry; ergodicity, which underlies certain number of results of Chapter 10; and finally, the computation of the density of the eigenvalues of random matrices, used in Chapter 8. From another point of view, in the applied examples the discussion of the underlying non-mathematical background has not been performed in detail, and other references should be consulted to go more deeply into the subjects. In all these cases we refer the reader to other books or scientific papers, expecting that this will suffice for a complete understanding .

At the same time, we would like this book to be useful for people working in research. In fact, we think that the mathematical reader looking for active research problems will find here a variety of open and interesting questions. These problems have a wide spectrum, from those which are tractable by the methods contained here to others which appear to be harder and seem to require new ideas. We would be happy if this book could be useful researchers in various areas of Mathematics (of course, Probability and Mathematical Statistics, but also Numerical Analysis and Algorithmic Complexity). At the same time, we believe the results should be useful for people using statistical modelling in Engineering, Econometrics and Biology, and hope to have contributed at least a little to build bridges in these directions.

Applications deserve a good deal of our attention. In what concerns applications to problems outside Mathematics we must recognize that our choice is strongly dependent on the taste and experience of the authors: the reader will find a section on Genetics, one on Inference on mixtures of populations and another one on Statistical Modelling of ocean waves. However, we have not included applications to Mathematical Physics or Econometrics in which the fine properties of the distribution of the maximum of a stochastic process play a central role.

We have also included applications of random fields methods to other parts of Mathematics, especially to systems of equations and condition numbers. This is a new field, even though some of the problems it considers are quite old, and has become a very important theme that mixes up different branches of Mathematics. One of the aims is to help understanding algorithm complexity, via the randomization of the problems which algorithms are designed to solve. One can also apply random field methods to study the conditioning of systems of inequalities (as it has been done in Cucker and Wschebor, 2003) but this is a subject which is only starting and we preferred not to include it in this book.

Numerical methods appear in various chapters. They are by no means simple and are in fact crucial to be able to apply the mathematical results, so that we want to stress their importance. Some are solved, we can for example refer to the Matlab toolbox MAGP described in Chapter 9. Some of them appear to be difficult: even in the simplest cases of Gaussian one-dimensional parameter processes defined on an interval, the numerical computation of the moments of the number of crossings of the paths with a given level, presents hard obstacles and is a source of interesting open problems.

Let us now give a quick overview of the contents of each chapter. This may help the reader to choose. At the end, we also add our advice on the order in which the various chapters might be read, especially on account of the chain of pre-requisites they require.

Chapter 1 contains the basic definitions of stochastic processes and Kolmogorov-type conditions implying that, almost surely, the paths have a certain regularity property (continuity, Hölder condition, differentiability). These classical and well-known results are not optimal, but they are sufficient for most uses. We have also included a reminder on the Gaussian distribution and its elementary properties, since a large part of the book is devoted to Gaussian processes and the reader may appreciate to have them in situ. We have not included X. Fernique (1974) and M. Talagrand (1985) beautiful results giving necessary and sufficient conditions for the continuity of the paths of Gaussian processes in the stationary and non-stationary cases respectively. We are not using them in this book.

Chapter 2 is about inequalities for Gaussian processes, mainly related to the distribution of the supremum. For comparison inequalities (which we call of “Slepian-type”) the main result is the Li and Shao inequality (2002), which includes and improves a large set of similar results that have been in use during fifty years or so, apparently starting in a paper by Plackett in 1954, motivated by statistical linear models with Gaussian noise.

The remainder of this chapter is devoted to the classical upper-bounds for the tails of the distribution of the supremum. They roughly say that if \mathcal{X} is a centered Gaussian process with bounded paths, then $1 - F_{M_T}(u)$, is bounded above by some constant times a Gaussian density. This was already known around 1975, due to a series of key contributions: Dudley (1967), Landaу and Shepp (1970), Marcus and Shepp (1972), Fernique (1974), Sudakov and Tsirelson (1974) and Borell (1975). To present isoperimetric inequalities, the basis is a quite recent proof of C. Borell (2003) of the extension to a general case of Ehrhard’s inequality (1983), a Minkowski-type inequality for the Gaussian measure in \mathbb{R}^n .

Rice formulas are proved in Chapters 3 and 6. Chapter 3 starts with a proof for Gaussian one-parameter processes, which is very simple; then, we consider the general non-Gaussian case. For various uses, one only wants to know whether the moments of crossings are finite, or to give upper-bounds for them, but without direct use of Rice formulas, since these can lead to non-tractable calculations. This has been the motivation for a series of papers deducing bounds from hypothesis on the process, mainly in the Gaussian case, which are cited in Chapter 3. In the same chapter we give a general simple criterion to ensure finiteness and obtain some rough bounds for the moments. To illustrate this kind of results, a corollary is that if \mathcal{X} is a Gaussian process defined on a compact interval $[0, T]$ of the real line having C^∞ -paths and satisfying the non-degeneracy condition $\text{Var}(X(t)) > 0$ for every $t \in [0, T]$ then, all the moments of crossings of any level are finite.

Rice formulas for random fields are considered in Chapter 6. Proofs are new and self contained, except for the already mentioned co-area formula. In all cases, formulas for the moments of weighted (or “marked”) crossings are stated and proved. They are used in the sequel for various applications and moreover, are important by themselves.

Chapter 4 contains two parts: 1) in Sections 1,2,3, \mathcal{X} is a Gaussian process defined on a bounded interval of the real line, and some initial estimates for $P(M > u)$ are given, based on computations of the first two moments of the number of crossings; 2) in Sections 4, and 5 two statistical applications are considered, the first one to genomics and the second one to statistical inference on mixtures of populations. The common feature of these two applications is that the relevant statistic for hypothesis testing is the maximum of a certain Gaussian process, so that the calculation of its distribution appears to be naturally related to the methods in the previous sections.

Chapter 5 establishes a bridge between the distribution of the maximum on an interval of a one-parameter process and the factorial moments of up-crossings of the paths. The main result is the general formula 5.2, which expresses the tail of the distribution of the maximum as the sum of a series (the ‘‘Rice series’’) defined in terms of certain factorial moments of the up-crossings of the given process. Rice series have been used for a long time with the aim of computing the distribution of the maximum of some special one-parameter Gaussian processes, as for example in the work of Miroshin (1974). The main point in theorems 5.1, 5.6 and 5.7 is that they provide general sufficient conditions to compute or approximate the distribution of the maximum. Even though some of the theoretical results are valid for non-Gaussian processes, if one wishes to apply them in specific cases, it becomes hard to compute the factorial moments of up-crossings for non-Gaussian processes. An interesting feature of Rice series is its enveloping property: replacing the total sum of the series by partial sums gives upper and lower bounds for the distribution of the maximum, and a fortiori, the error when one replaces the total sum by a partial sum is bounded by the absolute value of the last computed term. This allows one to perform the calculation of the distribution of the maximum with some efficiency. We have included a comparison with the computation based upon Monte-Carlo simulation of the paths of the process. However, in various situations more efficient methods exist; they are considered in Chapter 9.

In the first section of Chapter 7 we prove a general formula for the density of the probability distribution of the maximum which is valid for a large class of random fields. This is used in the next section to give strong results on the regularity of the distribution of one-parameter Gaussian process; as an example, if the paths are of class C^∞ and the joint law of the process and its derivatives is non-degenerate (in the sense specified in the text), then the distribution of the maximum is also of class C^∞ . When it comes to random fields, the situation is more complicated and the known results are essentially weaker, as one can see in the last section of this chapter.

Chapters 4 and 5, as well as 8 and 9, point towards improving the computation of the distribution of the maximum on the basis of special properties of the process, such as the regularity of the paths and the domain. In chapter 8 one profits from the implicit formula for the density of the distribution of the maximum that has been proved in Chapter 7 to study second order approximation of the tails of the distribution as the level u tends to $+\infty$, as done in Adler and Taylor’s recent book (2007) by other methods. The ‘‘direct method’’ employed here is suited to obtain also non-asymptotic results.

Chapter 10 contains a short account of limit theorems when the time domain grows to infinity, including the Volkonskii-Rozanov theorem on the asymptotic Poisson character of the stream of up-crossings for one-parameter stationary Gaussian processes under an appropriate normalization. This implies that the distribution of the maximum, after re-scaling, converges to a Gumbel distribution. The next section establishes a central limit theorem for non-linear functionals of Gaussian process, which applies to the limit behavior of the number of crossings.

Chapter 11 describes the modelling of the surface of the sea using Gaussian random fields. Some geometric characteristics of waves, as length of crests and velocities of contours are introduced. The Rice formula is used to deduce, from the directional spectrum of the sea, some properties of the distribution of these characteristics. Some non-Gaussian generalizations are proposed in the last section.

In Chapter 12 we study random systems of equations over the reals, having more than one unknown. For polynomial systems, the first significant results on the number of roots have been published in the 1990’s, starting with the Shub-Smale theorem (1993) which gives a simple elegant formula for the expectation of the number of roots. This is a fascinating subject in which, the main questions remain unanswered. What can one say about the probability distribution of the number of solutions and how does it depend on the probability law of the coefficients? What is the probability of having no solution? What can one say about the distribution of the roots in

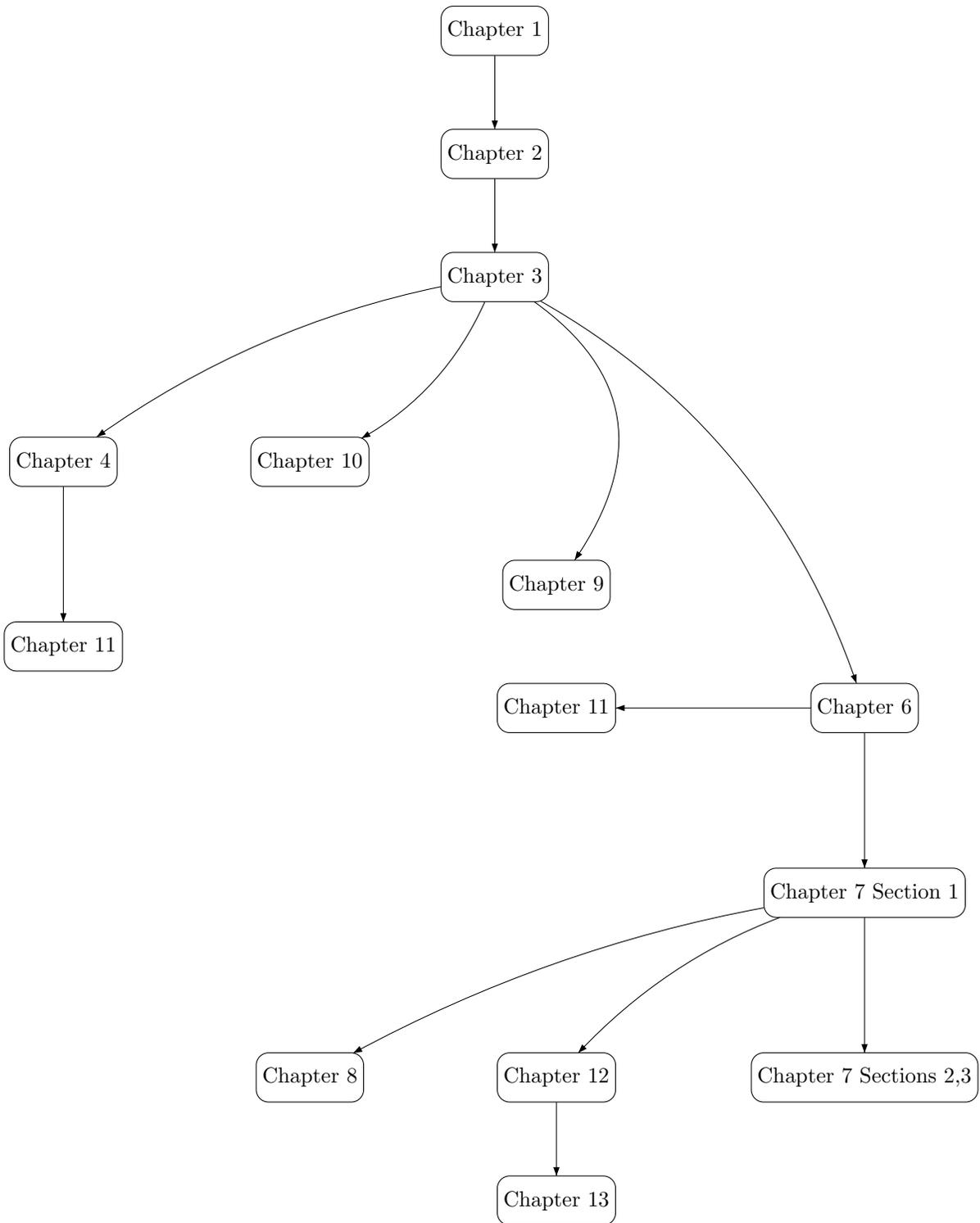
space and how can this help to solve the system numerically? How all these things behave as the number of unknowns grows indefinitely? What about the same questions for undetermined systems? And so on...

Answering some of these natural questions would imply at the same time making progress in key problems in Numerical Analysis and Algorithmic Complexity, as well as in other areas. The content of the chapter is extremely modest with respect to the above mentioned questions, and it is our hope that it may stimulate other people to work in the numerous problems arising in this field.

The last chapter is on condition numbers. Roughly speaking, the condition number of a problem measures the difficulty of the problem to be solved by any algorithm. Of course, there are many measures of this sort. The standard procedure is to give a metric on the space of problems and define the condition number of a problem as the inverse of its distance to the set of ill-posed problems, possibly with some additional normalization. In this chapter, we consider the simplest situation, in which a “problem” is a square system of linear equations and the set of ill-posed problems is the set of systems in which the matrix of coefficients is non-invertible. The condition number turns out to be related to the singular values of the matrix of coefficients. The role of condition numbers in numerical linear algebra is well-known since the 1940’s, when they were introduced by Turing (1948) and Von Neumann and Goldstine (1947). See also Smale (1997).

Condition numbers appear in the estimation of the complexity of algorithms in a natural way. When the choice of the problem is performed at random, the condition number becomes a random variable. Of course, its probability distribution will depend upon the underlying probability law on the set of problems. In the linear case, computing the probability distribution of the condition number becomes a problem on the spectrum of random matrices. The importance of studying the distribution of condition numbers of random matrices has been put forward by Smale (1985). The first precise result is due to A. Edelman (1988), who computed the equivalent of the expected value of the logarithm of the condition number, when the matrix size n tends to infinity.

The methods we present to study the probability distribution of the condition number of square random matrices rely again on random fields and Rice formula. They are not optimal in some standard cases but, in some other ones, produce at present the best known bounds. We do not consider the similar - and much harder - questions for non-linear systems. The reader can consult the book by L. Blum et al (1998) which includes a basic presentation of this subject or Cucker et al (2008) .

Reading diagram

Acknowledgements. We thank Christine Cierco-Ayrolles, Céline Delmas, Cécile Mercadier and Joaquin Ortega for allowing us to reproduce some tables and pictures of their published work (See Chapters 4, 9 and 11), as well as for valuable discussions.

We also debt gratitude to numerous colleagues and in particular we want to thank Alan Genz, Marie Kratz, Michel Ledoux and José León, for questions and comments that have contributed to improve the text.

Finally the ECOS cooperation Action U03E01 that enabled us to cross the Atlantic Ocean in both directions in order to achieve this book.

Classical results on the regularity of the paths.

This initial chapter contains a number of elements that will be repeatedly used along the book and constitute a necessary background to read it.

We will need to study the paths of random processes and fields, and the analytical properties of these functions play a relevant role. This arises a certain number of basic questions, such as whether the paths belong to a certain regularity class of functions, what can one say about their global or local extrema, about local inversion and so on. A typical situation is that the available knowledge on the random function is given by its probability law. So, one is willing to know what one can deduce from this probability law about these kind of properties of the paths. Generally speaking, the results one can expect is the existence of a “version” of the random function having good analytical properties. A version is a random function which, at each parameter value, coincides almost surely with the given one.

These are the contents of Section 4 of this chapter, which includes the classical theorems due to A.N. Kolmogorov and the results of Bulinskaya and Ylvisaker about the existence of critical points or local extrema having given values. The essence of all this has been well-known for a long time and in some cases proofs are only sketched. In other cases, we give full proofs and some refinements which will be necessary for further use.

As for the previous sections, the first one contains starting notational conventions and the statement of the Kolmogorov extension theorem of Measure Theory and the other two are a quick overview of the Gaussian distribution and some connected results. Even though this is completely elementary, we call the attention of the reader upon Proposition 1.2 which is the Gaussian regression formula, which will appear once and again along the book, and can be considered as the basis of the calculations with the Gaussian distribution.

1. Kolmogorov’s Extension Theorem.

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and (F, \mathcal{F}) a measurable space. For any measurable function

$$Y : (\Omega, \mathcal{A}) \rightarrow (F, \mathcal{F}),$$

that is, a random variable with values in F , the image measure

$$Q(A) = \mathbb{P}(Y^{-1}(A)) \quad (A \in \mathcal{F})$$

is called the “distribution” of Y .

Except for explicit statement of the contrary, we will assume that probability spaces are complete, that is, every subset of a set having zero probability, is measurable. Let us recall that if (F, \mathcal{F}, μ) is a measure space, one can always define its completion $(F, \mathcal{F}_1, \mu_1)$ by putting:

$$(1.1) \quad \mathcal{F}_1 = \{A : \exists B, C, A = B \triangle C, \text{ such that } B \in \mathcal{F}, C \subset D \in \mathcal{F}, \mu(D) = 0\}$$

and for $A \in \mathcal{F}_1$, $\mu_1(A) = \mu(B)$ whenever A admits the representation in (1.1). One can check that $(F, \mathcal{F}_1, \mu_1)$ is a complete measure space and μ_1 an extension of μ .

A “real-valued stochastic process indexed by the set T ” is a collection of random variables $\{X(t) : t \in T\}$ defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

In what follows, we will assume that the process is “bi-measurable”. This means that we have a σ -algebra \mathcal{T} of subsets of T and a Borel-measurable function of the pair (t, ω) to the reals:

$$X : (T \times \Omega, \mathcal{T} \times \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$$

($\mathcal{B}_{\mathbb{R}}$ denotes the Borel σ -algebra in \mathbb{R}) so that:

$$X(t)(\omega) = X(t, \omega).$$

Let T be a set and $\mathbb{R}^T = \{g : T \rightarrow \mathbb{R}\}$ the set of real-valued functions defined on T (in what follows in this section, one may replace \mathbb{R} by \mathbb{R}^d , $d > 1$).

For $n = 1, 2, \dots$, t_1, t_2, \dots, t_n n distinct elements of T and B_1, B_2, \dots, B_n Borel sets in \mathbb{R} , we denote:

$$C(t_1, t_2, \dots, t_n; B_1, B_2, \dots, B_n) = \{g \in \mathbb{R}^T : g(t_j) \in B_j, j = 1, 2, \dots, n\}$$

and \mathcal{C} the family of all sets of the form $C(t_1, t_2, \dots, t_n; B_1, B_2, \dots, B_n)$. These are usually called the “cylinder sets depending on t_1, t_2, \dots, t_n ”. The smallest σ -algebra of parts of \mathbb{R}^T containing \mathcal{C} will be called the Borel σ -algebra of \mathbb{R}^T and denoted by $\sigma(\mathcal{C})$.

Consider now a family of probability measures

$$(1.2) \quad \{\mathbb{P}_{t_1, t_2, \dots, t_n}\}_{t_1, t_2, \dots, t_n \in T; n=1, 2, \dots}$$

as follows: for each $n = 1, 2, \dots$ and each n -tuple t_1, t_2, \dots, t_n of distinct elements of T , $\mathbb{P}_{t_1, t_2, \dots, t_n}$ is a probability measure on the Borel sets of the product space $X_{t_1} \times X_{t_2} \times \dots \times X_{t_n}$ where $X_t = \mathbb{R}$ for each $t \in T$ (so that this product space is canonically identified to \mathbb{R}^n).

We say that the probability measures (1.2) satisfy the “consistency condition” if for any choice of $n = 1, 2, \dots$ and distinct $t_1, \dots, t_n, t_{n+1} \in T$ we have:

$$\mathbb{P}_{t_1, \dots, t_n, t_{n+1}}(B \times \mathbb{R}) = \mathbb{P}_{t_1, \dots, t_n}(B)$$

for any Borel set B in $X_{t_1} \times \dots \times X_{t_n}$. The following is the basic Kolmogorov’s Extension Theorem, that we state and do not prove here.

THEOREM 1.1 (Kolmogorov). $\{\mathbb{P}_{t_1, t_2, \dots, t_n}\}_{t_1, t_2, \dots, t_n \in T; n=1, 2, \dots}$ satisfy the consistency condition if and only if there exists one and only one probability measure \mathbb{P} on $\sigma(\mathcal{C})$ such that

$$(1.3) \quad \mathbb{P}(C(t_1, \dots, t_n; B_1, \dots, B_n)) = \mathbb{P}_{t_1, \dots, t_n}(B_1 \times \dots \times B_n)$$

for any choice of $n = 1, 2, \dots$, distinct $t_1, \dots, t_n \in T$ and B_j Borel sets in X_{t_j} , $j = 1, \dots, n$.

It is clear that if there exists a probability measure \mathbb{P} on $\sigma(\mathcal{C})$ satisfying (1.3) then the consistency conditions must hold since

$$C(t_1, \dots, t_n, t_{n+1}; B_1, \dots, B_n, X_{t_{n+1}}) = C(t_1, \dots, t_n; B_1, \dots, B_n)$$

So, the question is to prove the converse. This can be done in two steps: 1) define \mathbb{P} on the family of cylinders \mathcal{C} using (1.3) and show that the definition is unambiguous (note that each cylinder has more than one representation); 2) apply Caratheodory’s Theorem on extension of measures to prove that this \mathbb{P} can be extended in a unique form to $\sigma(\mathcal{C})$.

1.1. Remarks.

- (1) The above Theorem is interesting when T is an infinite set. The purpose is to be able to measure the probability of sets of functions from T to \mathbb{R} (i.e. subsets of \mathbb{R}^T) which can not be defined by means of a finite number of coordinates, which amounts to saying only looking at the values of the functions at a finite number of t -values.

Notice that in the case of cylinders, if one wants to know whether a given function $g : T \rightarrow \mathbb{R}$ belongs to $C(t_1, \dots, t_n; B_1, \dots, B_n)$ it suffices to look at the values of g at the finite set of points t_1, \dots, t_n and check if $g(t_j) \in B_j$ for $j = 1, \dots, n$. However, if one takes for example, $T = \mathbb{Z}$ (the integers) and considers the sets of functions:

$$A = \left\{ g : g : T \rightarrow \mathbb{R}, \lim_{t \rightarrow +\infty} g(t) \text{ exists and is finite} \right\}$$

or

$$B = \left\{ g : g : T \rightarrow \mathbb{R}, \sup_{t \in T} |g(t)| \leq 1 \right\}$$

then it is clear that these sets are in $\sigma(\mathcal{C})$ but are not cylinders (they “depend on an infinite number of coordinates”).

- (2) In general $\sigma(\mathcal{C})$ is strictly smaller than the family of all subsets of \mathbb{R}^T . To see this, one can check that

$$(1.4) \quad \sigma(\mathcal{C}) = \{A \subset \mathbb{R}^T : \exists T_A \subset T, T_A \text{ countable and } B_A \text{ a Borel set in } \mathbb{R}^{T_A}, \text{ such that } \\ g \in A \text{ if and only if } g|_{T_A} \in B_A\}.$$

The proof of (1.4) follows immediately from the fact that the right-hand side is a σ -algebra containing \mathcal{C} . (1.4) says that a subset of \mathbb{R}^T is a Borel set if and only if it “depends only on a countable set of parameter values”.

Hence, if T is uncountable, the sets

$$\{g \in \mathbb{R}^T : g \text{ is a bounded function}\}$$

or

$$\{g \in \mathbb{R}^T : g \text{ is a bounded function, } |g(t)| \leq 1 \text{ for all } t \in T\}$$

do not belong to $\sigma(\mathcal{C})$. Another simple example is the following: If $T = [0, 1]$, then

$$\{g \in \mathbb{R}^T : g \text{ is a continuous function}\}$$

is not a Borel set in \mathbb{R}^T , since it is obvious that there does not exist a countable subset of $[0, 1]$ having the determining property in (1.4). These examples lead to the notion of “separable process” that we will introduce later on.

- (3) In the special case when $\Omega = \mathbb{R}^T$, $\mathcal{A} = \sigma(\mathcal{C})$ and $X(t)(\omega) = \omega(t)$, $\{X(t) : t \in T\}$ is called a “canonical process”.
- (4) We will say that the stochastic process $\{Y(t) : t \in T\}$ is a “version” of the process $\{X(t) : t \in T\}$ if $P(X(t) = Y(t)) = 1$ for each $t \in T$.

2. Reminder on the Normal Distribution.

Let μ be a probability measure on the Borel subsets of \mathbb{R}^d . Its Fourier transform $\hat{\mu} : \mathbb{R}^d \rightarrow \mathbb{C}$ is defined as

$$\hat{\mu}(z) = \int_{\mathbb{R}^d} \exp(i \langle z, x \rangle) \mu(dx)$$

where $\langle \cdot, \cdot \rangle$ denotes the usual scalar product in \mathbb{R}^d .

We will use Bochner’s Theorem (see for example Feller, 1966): $\hat{\mu}$ is the Fourier transform of a Borel probability measure on \mathbb{R}^d if and only if the following three conditions hold true:

- (1) $\hat{\mu}(0) = 1$
- (2) $\hat{\mu}$ is continuous
- (3) $\hat{\mu}$ is positive semi-definite, that is for any $n = 1, 2, \dots$ and any choice of the complex numbers c_1, \dots, c_n and of the points z_1, \dots, z_n one has

$$\sum_{j,k=1}^n \hat{\mu}(z_j - z_k) c_j \bar{c}_k \geq 0.$$

The random vector ξ with values in \mathbb{R}^d is said to have the “normal distribution” - or the “Gaussian distribution” - with parameters (m, Σ) [$m \in \mathbb{R}^d$ and Σ a $d \times d$ positive semi-definite matrix] if the Fourier transform of the probability distribution μ_ξ of ξ is equal to:

$$\hat{\mu}_\xi(z) = \exp \left[i \langle m, z \rangle - \frac{1}{2} \langle z, \Sigma z \rangle \right], \quad z \in \mathbb{R}^d.$$

When $m = 0$ and $\Sigma = I =$ identity $d \times d$ matrix, the distribution of ξ is called "standard normal in \mathbb{R}^d ". For $d = 1$, we use the notation

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad \text{and} \quad \Phi(x) = \int_{-\infty}^x \varphi(y) dy$$

for the density and the cumulative distribution function of a standard normal random variable, respectively.

If Σ is non-singular, then μ_ξ is said to be "non-degenerate" and one can verify that it has a density with respect to Lebesgue measure given by

$$\mu_\xi(dx) = \frac{1}{(2\pi)^{\frac{d}{2}} (\det(\Sigma))^{\frac{1}{2}}} \exp \left[-\frac{1}{2} (x - m)^T \Sigma^{-1} (x - m) \right] dx$$

x^T denotes the transpose of x . One can check that:

$$m = \mathbb{E}(\xi), \quad \Sigma = \text{Var}(\xi) = \mathbb{E}((\xi - m)(\xi - m)^T),$$

so that m and Σ are respectively the mean and the variance of ξ .

From the above definition it follows that if the random vector ξ with values in \mathbb{R}^d has a normal distribution with parameters (m, Σ) , A is a real matrix with n rows and d columns and b is a non-random element of \mathbb{R}^n , then the random vector

$$A\xi + b$$

with values in \mathbb{R}^n has normal distribution with parameters $(Am + b, A\Sigma A^T)$. In particular, if Σ is non-singular, the coordinates of the random vector

$$\Sigma^{-1/2}(\xi - m)$$

are independent random variables with standard normal distribution on the real line.

Assume now that we have a pair ξ, η or random vectors in \mathbb{R}^d and $\mathbb{R}^{d'}$ respectively, having finite moments of order 2. We define the $d \times d'$ covariance matrix as

$$\text{Cov}(\xi, \eta) = \mathbb{E}((\xi - \mathbb{E}(\xi))(\eta - \mathbb{E}(\eta))^T).$$

It follows that if the distribution of the random vector (ξ, η) in $\mathbb{R}^{d+d'}$ is normal, and $\text{Cov}(\xi, \eta) = 0$, then the random vectors ξ, η are independent. A consequence of this is the following useful formula, which is standard in Statistics and gives a version for the conditional expectation of a function of ξ given the value of η .

PROPOSITION 1.2. *Let ξ, η be two random vectors with values in \mathbb{R}^d and $\mathbb{R}^{d'}$ respectively and assume that the distribution of (ξ, η) in $\mathbb{R}^{d+d'}$ is normal and $\text{Var}(\eta)$ is non-singular. Then, for any bounded function $f: \mathbb{R}^d \rightarrow \mathbb{R}$ we have:*

$$(1.5) \quad \mathbb{E}(f(\xi) | \eta = y) = \mathbb{E}(f(\zeta + Cy))$$

for almost every y , where

$$(1.6) \quad C = \text{Cov}(\xi, \eta) [\text{Var}(\eta)]^{-1}$$

and ζ is a random vector with values in \mathbb{R}^d , having normal distribution with parameters

$$(1.7) \quad (\mathbb{E}(\xi) - C\mathbb{E}(\eta), \text{Var}(\xi) - \text{Cov}(\xi, \eta) [\text{Var}(\eta)]^{-1} [\text{Cov}(\xi, \eta)]^T).$$

PROOF. The proof consists in choosing the matrix C so that the random vector

$$\zeta = \xi - C\eta$$

becomes independent of η . For this purpose, we need that

$$\text{Cov}(\xi - C\eta, \eta) = 0$$

and this leads to the value of C given by (1.6). (1.7) follows immediately. \square

In what follows, we will call “Gaussian regression” the version of the conditional expectation given by formula (1.5). To close this quick list of basic properties, we mention that a useful property of the Gaussian distribution is stability under passage to the limit (See Exercise 1.5 below).

Let $r : T \times T \rightarrow \mathbb{R}$ be a positive semi-definite function, and $m : T \rightarrow \mathbb{R}$ a function. In this more general context, that r is a positive semi-definite function means that for any $n = 1, 2, \dots$ and any choice of distinct $t_1, \dots, t_n \in T$ the matrix

$$((r(t_j, t_k)))_{j,k=1,\dots,n}$$

is positive semi-definite.

[This is consistent with the previous definition, which corresponds to saying that $r(s, t) = \widehat{\mu}(s - t)$, $s, t \in \mathbb{R}^d$ is positive semi definite].

Take now for P_{t_1, \dots, t_n} the Gaussian probability measure in \mathbb{R}^n with mean

$$m_{t_1, \dots, t_n} := (m(t_1), \dots, m(t_n))^T$$

and variance matrix

$$\Sigma_{t_1, \dots, t_n} := ((r(t_j, t_k)))_{j,k=1,\dots,n}.$$

It is easily verified that the set of probability measures $\{P_{t_1, \dots, t_n}\}$ verifies the consistency condition, so that Kolmogorov’s Theorem applies and there exists a unique probability measure P on the measurable space $(\mathbb{R}^T, \sigma(\mathcal{C}))$ which restricted to the cylinder sets depending on t_1, \dots, t_n is P_{t_1, \dots, t_n} for any choice of distinct parameter values t_1, \dots, t_n . P is called the “Gaussian measure generated by the pair (m, r) ”. If $\{X(t) : t \in T\}$ is a real-valued stochastic process with distribution P one verifies that:

- for any choice of distinct parameter values t_1, \dots, t_n , the joint distribution of the random variables $X(t_1), \dots, X(t_n)$ is Gaussian with mean m_{t_1, \dots, t_n} and variance Σ_{t_1, \dots, t_n}
- $E(X(t)) = m(t)$ for $t \in T$
- $\text{Cov}(X(s), X(t)) = E((X(s) - m(s))(X(t) - m(t))) = r(s, t)$ for $s, t \in T$.

A class of examples which appears frequently in applications is the d -parameter real-valued Gaussian processes which are centered and stationary, which means that

$$T = \mathbb{R}^d, \quad m(t) = 0, \quad r(s, t) = \Gamma(t - s).$$

A general definition of strictly stationary processes will be given in Section 2 of Chapter 10.

If the function Γ is continuous, $\Gamma(0) \neq 0$, one can write

$$\Gamma(\tau) = \int_{\mathbb{R}^d} \exp(i \langle \tau, x \rangle) \mu(dx),$$

where μ is a Borel measure on \mathbb{R}^d with total mass equal to $\Gamma(0)$. μ is called the “spectral measure” of the process. We will usually assume that $\Gamma(0) = 1$, i.e. that μ is a probability measure which is simply obtained by replacing the original process $\{X(t) : t \in \mathbb{R}^d\}$ by the process

$$\left\{ \frac{X(t)}{(\Gamma(0))^{\frac{1}{2}}} : t \in \mathbb{R}^d \right\}.$$

EXAMPLE 1.1 (Trigonometric polynomials). *An important example of stationary Gaussian processes is the following.*

Suppose μ is a purely atomic probability symmetric measure on the real line, that is, there exists a sequence $\{x_n\}_{n=1,2,\dots}$ of positive real numbers such that

$$\mu(\{x_n\}) = \mu(\{-x_n\}) = \frac{1}{2}c_n \text{ for } n = 1, 2, \dots; \quad \mu(\{0\}) = c_0; \quad \sum_{n=0}^{\infty} c_n = 1.$$

Then a centered Gaussian process having μ as spectral measure is

$$(1.8) \quad X(t) = c_0^{\frac{1}{2}} \xi_0 + \sum_{n=1}^{\infty} c_n^{\frac{1}{2}} [\xi_n \cos(tx_n) + \xi_{-n} \sin(tx_n)], \quad t \in \mathbb{R},$$

where the $\{\xi_n\}_{n \in \mathbb{Z}}$ is a sequence of independent identically distributed random variables, each of them having a standard normal distribution. In fact, the series in (1.8) converges in $L^2(\Omega, \mathcal{F}, P)$ and

$$\mathbb{E}(X(t)) = 0 \quad \text{and} \quad \mathbb{E}(X(s)X(t)) = c_0 + \sum_{n=1}^{\infty} c_n \cos[(t-s)x_n] = \widehat{\mu}(t-s)$$

We will use the notation

$$(1.9) \quad \lambda_k := \int_{\mathbb{R}} x^k \mu(dx), \quad k = 0, 1, 2, \dots$$

whenever the integral exists. λ_k is the “ k -th spectral moment” of the process.

An extension of the preceding class of examples is the following.

Let (T, \mathcal{T}, ρ) be a measure space, $H = L^2_{\mathbb{R}}(T, \mathcal{T}, \rho)$ the Hilbert space of real-valued square integrable functions on it and $\{\varphi_n(t)\}_{n=1,2,\dots}$ an orthonormal sequence in H . We assume that each function $\varphi_n : T \rightarrow \mathbb{R}$ is bounded and denote $M_n = \sup_{t \in T} |\varphi_n(t)|$.

In addition, let $\{c_n\}_{n=1,2,\dots}$ be a sequence of positive numbers such that

$$\sum_{n=1}^{\infty} c_n < \infty, \quad \sum_{n=1}^{\infty} c_n M_n^2 < \infty$$

and $\{\xi_n\}_{n=1,2,\dots}$ a sequence of independent identically distributed random variables, each of them with standard normal distribution in \mathbb{R} .

Then, the stochastic process

$$(1.10) \quad X(t) = \sum_{n=1}^{\infty} c_n^{\frac{1}{2}} \xi_n \varphi_n(t)$$

is Gaussian, centered with covariance

$$r(s, t) = \mathbb{E}(X(s)X(t)) = \sum_{n=1}^{\infty} c_n \varphi_n(s) \varphi_n(t).$$

Formulas (1.8) and (1.10) are simple cases of spectral representations of Gaussian processes, which is an important subject, both for theoretical purposes and for applications. A compact presentation of this subject, including the Karhunen-Loève representation and the connection with Reproducing Kernel Hilbert Spaces, can be found in Fernique’s Lecture Notes (1974).

3. 0-1 law for Gaussian processes.

We will prove a 0-1 law for Gaussian processes in this section, without attempting full generality. This will be sufficient for our requirements in what follows. For a more general treatment, see Fernique (1974).

DEFINITION 1.3. Let $\mathcal{X} = \{X(t) : t \in T\}$, $\mathcal{Y} = \{Y(t) : t \in S\}$ be real-valued stochastic processes defined on some probability space (Ω, \mathcal{A}, P) . \mathcal{X} and \mathcal{Y} are said to be independent if for any choice of the parameter values $t_1, \dots, t_n \in T; s_1, \dots, s_m \in S$, $n, m \geq 1$, the random vectors

$$(X(t_1), \dots, X(t_n)), (Y(s_1), \dots, Y(s_m))$$

are independent.

PROPOSITION 1.4. Let the processes \mathcal{X} and \mathcal{Y} be independent and E (respectively F) belong to the σ -algebra generated by the cylinders in \mathbb{R}^T (respectively \mathbb{R}^S). Then:

$$(1.11) \quad \mathbb{P}(X(\cdot) \in E, Y(\cdot) \in F) = \mathbb{P}(X(\cdot) \in E) \mathbb{P}(Y(\cdot) \in F)$$

PROOF. (1.11) holds true for cylinders. Uniqueness in the extension Theorem provides the result. \square

THEOREM 1.5. [*0 or 1 Law for Gaussian Processes*]

Let $\mathcal{X} = \{X(t) : t \in T\}$ be a real-valued centered Gaussian process defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and (E, \mathcal{E}) a measurable space, where E is a linear subspace of \mathbb{R}^T and the σ -algebra \mathcal{E} has the property that for any choice of the scalars $a, b \in \mathbb{R}$, the function $(x, y) \rightsquigarrow ax + by$ defined on $E \times E$ is measurable with respect to the product σ -algebra. We assume that the function $X : \Omega \rightarrow E$ defined as $X(\omega) = X(\cdot, \omega)$ is measurable $(\Omega, \mathcal{A}) \rightarrow (E, \mathcal{E})$.

Then, if L is a measurable subspace of E , one has

$$\mathbb{P}(X(\cdot) \in L) = 0 \quad \text{or} \quad 1.$$

PROOF. Let $\{X^{(1)}(t) : t \in T\}, \{X^{(2)}(t) : t \in T\}$ be two independent processes each of them having the same distribution as the given process $\{X(t) : t \in T\}$.

For each $\lambda, 0 < \lambda < \frac{\pi}{2}$ consider a new pair of stochastic processes, defined for $t \in T$ by:

$$(1.12) \quad \begin{aligned} Z_\lambda^{(1)}(t) &= X^{(1)}(t) \cos \lambda + X^{(2)}(t) \sin \lambda \\ Z_\lambda^{(2)}(t) &= -X^{(1)}(t) \sin \lambda + X^{(2)}(t) \cos \lambda \end{aligned}$$

Each of the processes $Z_\lambda^{(i)}(t) (i = 1, 2)$ has the same distribution as \mathcal{X} .

In fact, $\mathbb{E}(Z_\lambda^{(1)}(t)) = 0$ and since $\mathbb{E}(X^{(1)}(s)X^{(2)}(t)) = 0$ we have $\mathbb{E}(Z_\lambda^{(1)}(s)Z_\lambda^{(1)}(t)) = \cos^2 \lambda \mathbb{E}(X^{(1)}(s)X^{(1)}(t)) + \sin^2 \lambda \mathbb{E}(X^{(2)}(s)X^{(2)}(t)) = \mathbb{E}(X(s)X(t))$.

A similar computation holds for $Z_\lambda^{(2)}$.

Also, the processes $Z_\lambda^{(1)}, Z_\lambda^{(2)}$ are independent. To prove this, note that for any choice of $t_1, \dots, t_n; s_1, \dots, s_m, n, m \geq 1$, the random vectors

$$(Z_\lambda^{(1)}(t_1), \dots, Z_\lambda^{(1)}(t_n)), (Z_\lambda^{(2)}(s_1), \dots, Z_\lambda^{(2)}(s_m))$$

have a joint Gaussian distribution so that it suffices to show that

$$\mathbb{E}(Z_\lambda^{(1)}(t)Z_\lambda^{(2)}(s)) = 0$$

for any choice of $s, t \in T$ to conclude that they are independent. This is easily checked.

Now, if we put $q = \mathbb{P}(X(\cdot) \in L)$, independence implies that for any λ :

$$q(1 - q) = \mathbb{P}(E_\lambda) \quad \text{where} \quad E_\lambda = \left\{ Z_\lambda^{(1)} \in L, Z_\lambda^{(2)} \notin L \right\}.$$

If $\lambda, \lambda' \in (0, \frac{\pi}{2}), \lambda \neq \lambda'$, the events E_λ and $E_{\lambda'}$ are disjoint. In fact, the matrix

$$\begin{pmatrix} \cos \lambda & \sin \lambda \\ \cos \lambda' & \sin \lambda' \end{pmatrix}$$

is non-singular and (1.12) implies that if at the same time $Z_\lambda^{(1)} \in L, Z_{\lambda'}^{(1)} \in L$ then, also $X^{(1)}(\cdot), X^{(2)}(\cdot) \in L$ since $X^{(1)}(\cdot), X^{(2)}(\cdot)$ are linear combinations of $Z_\lambda^{(1)}, Z_{\lambda'}^{(1)}$. Hence, $Z_\lambda^{(2)}, Z_{\lambda'}^{(2)} \in L$ and $E_\lambda, E_{\lambda'}$ can not occur simultaneously. To finish, the only way in which we can have an infinite family $\{E_\lambda\}_{0 < \lambda < \frac{\pi}{2}}$ of pairwise disjoint events with equal probability, is this probability to be equal to zero. That is, $q(1 - q) = 0$ so that $q = 0$ or 1 . \square

In case the parameter set T is countable, the above shows directly that any measurable linear subspace of \mathbb{R}^T has probability 0 or 1 under a centered Gaussian law. If T is a σ -compact topological space, E the set of real-valued continuous functions defined on T and \mathcal{E} the σ -algebra generated by the topology of uniform convergence on compact sets, one can conclude, for example, that the subspace of E of bounded functions has probability 0 or 1 under a centered Gaussian measure. The Theorem can be applied in a variety of similar situations to standard function spaces. For example, put a measure on the space (E, \mathcal{E}) and take for L an L^p of this measure space.

4. Regularity of the paths.

4.1. Conditions for continuity of the paths.

THEOREM 1.6 (Kolmogorov). *Let $\mathcal{Y} = \{Y(t) : t \in [0, 1]\}$ be a real-valued stochastic process that satisfies the condition*

(K) *For each pair $t, t+h \in [0, 1]$,*

$$\mathbb{P}\{|Y(t+h) - Y(t)| \geq \alpha(h)\} \leq \beta(h),$$

where α, β are even real-valued functions defined on $[-1, 1]$, increasing on $[0, 1]$ that verify

$$\sum_{n=1}^{\infty} \alpha(2^{-n}) < \infty, \quad \sum_{n=1}^{\infty} 2^n \beta(2^{-n}) < \infty$$

Then, there exists a version $\mathcal{X} = \{X(t) : t \in T\}$ of the process \mathcal{Y} such that the paths $t \rightsquigarrow X(t)$ are continuous on $[0, 1]$.

PROOF. For $n = 1, 2, \dots$; $k = 0, 1, \dots, 2^n - 1$, let:

$$E_{k,n} = \left\{ \left| Y\left(\frac{k+1}{2^n}\right) - Y\left(\frac{k}{2^n}\right) \right| \geq \alpha(2^{-n}) \right\}, \quad E_n = \bigcup_{k=0}^{2^n-1} E_{k,n}$$

From the hypothesis: $\mathbb{P}(E_n) \leq 2^n \beta(2^{-n})$, so that $\sum_{n=1}^{\infty} \mathbb{P}(E_n) < \infty$. The Borel-Cantelli Lemma implies that $\mathbb{P}(\limsup_{n \rightarrow \infty} E_n) = 0$, where

$$\limsup_{n \rightarrow \infty} E_n = \{\omega : \omega \text{ belongs to infinitely many } E_n \text{'s}\}$$

In other words, if $\omega \notin \limsup_{n \rightarrow \infty} E_n$ one can find $n_0(\omega)$ such that if $n \geq n_0(\omega)$ one has

$$\left| Y\left(\frac{k+1}{2^n}\right) - Y\left(\frac{k}{2^n}\right) \right| < \alpha(2^{-n}) \text{ for all } k = 0, 1, \dots, 2^n - 1.$$

Denote by $Y^{(n)}$ the function whose graph is the polygonal with vertices $(\frac{k}{2^n}, Y(\frac{k}{2^n}))$, $k = 0, 1, \dots, 2^n$, that is, if $\frac{k}{2^n} \leq t \leq \frac{k+1}{2^n}$ one has

$$Y^{(n)}(t) = (k+1 - 2^n t)Y\left(\frac{k}{2^n}\right) + (2^n t - k)Y\left(\frac{k+1}{2^n}\right).$$

The function $t \rightsquigarrow Y^{(n)}(t)$ is continuous. Now, if $\omega \notin \limsup_{n \rightarrow \infty} E_n$ one easily checks that there exists some integer $n_0(\omega)$ such that:

$$\|Y^{(n+1)} - Y^{(n)}\|_{\infty} \leq \alpha(2^{-(n+1)}) \text{ for } n+1 \geq n_0(\omega)$$

(here $\|\cdot\|_{\infty}$ denotes the sup norm on $[0, 1]$). Since $\sum_{n=1}^{\infty} \alpha(2^{-(n+1)}) < \infty$ by the hypothesis, the sequence of functions $\{Y^{(n)}\}$ converges uniformly on $[0, 1]$ to a continuous limit function that we denote $X(t), t \in [0, 1]$.

We put $X(t) \equiv 0$ when $\omega \in \limsup_{n \rightarrow \infty} E_n$.

To finish the proof, it suffices to show that for each $t \in [0, 1]$, $\mathbb{P}(X(t) = Y(t)) = 1$.

- If t is a dyadic point, say $t = \frac{k}{2^n}$, then given the definition of the sequence of functions $Y^{(n)}$, it is clear that $Y^{(m)}(t) = Y(t)$ for $m \geq n$. Hence for $\omega \notin \limsup_{n \rightarrow \infty} E_n$, one has $X(t) = \lim_{m \rightarrow \infty} Y^{(m)}(t) = Y(t)$.
The result follows from $\mathbb{P}((\limsup_{n \rightarrow \infty} E_n)^C) = 1$ (A^C is the complement of the set A).
- If t is not a dyadic point, for each $n, n = 1, 2, \dots$ let k_n be an integer such that $|t - \frac{k_n}{2^n}| \leq 2^{-n}$, $\frac{k_n}{2^n} \in [0, 1]$. Put

$$F_n = \left\{ \left| Y(t) - X\left(\frac{k_n}{2^n}\right) \right| \geq \alpha(2^{-n}) \right\}.$$

We have the inequalities:

$$\mathbb{P}(F_n) \leq \mathbb{P}\left(\left| Y(t) - X\left(\frac{k_n}{2^n}\right) \right| \geq \alpha\left(\left|t - \frac{k_n}{2^n}\right|\right)\right) \leq \beta\left(\left|t - \frac{k_n}{2^n}\right|\right) \leq \beta(2^{-n})$$

and a new application of the Borel-Cantelli Lemma gives $\mathbb{P}(\limsup_{n \rightarrow \infty} F_n) = 0$.

So, if $\omega \notin [\limsup_{n \rightarrow \infty} E_n] \cup [\limsup_{n \rightarrow \infty} F_n]$ we have at the same time $X(\frac{k_n}{2^n})(\omega) \rightarrow X(t)(\omega)$ as $n \rightarrow \infty$ - because $t \rightsquigarrow X(t)$ is continuous - and $X(\frac{k_n}{2^n})(\omega) \rightarrow Y(t)(\omega)$ - because $|Y(t) - X(\frac{k_n}{2^n})| < \alpha(2^{-n})$ for $n \geq n_1(\omega)$ for some integer $n_1(\omega)$.

This proves that $X(t)(\omega) = Y(t)(\omega)$ for almost every ω \square

COROLLARY 1.7. *Assume that the process $\mathcal{Y} = \{Y(t) : t \in [0, 1]\}$ satisfies one of the following conditions for $t, t+h \in [0, 1]$:*

a)

$$(1.13) \quad \mathbb{E}(|Y(t+h) - Y(t)|^p) \leq \frac{K|h|}{|\log|h||^{1+r}}$$

where p, r, K are positive constants, $p < r$.

b) \mathcal{Y} is Gaussian, $m(t) := \mathbb{E}(Y(t))$ is continuous and

$$(1.14) \quad \text{Var}(Y(t+h) - Y(t)) \leq \frac{C}{|\log|h||^a}$$

for all t , sufficiently small h , C some positive constant and $a > 3$.

Then, the conclusion of Theorem 1.6 holds.

PROOF. a) Put:

$$\alpha(h) = \frac{1}{|\log|h||^b}, \quad 1 < b < \frac{r}{p}$$

$$\beta(h) = \frac{|h|}{|\log|h||^{1+r-bp}}$$

and check condition (K) using a Markov inequality.

b) Since the expectation is continuous it can be subtracted from $Y(t)$, so that we may assume that \mathcal{Y} is centered. To apply Theorem 1.6, take

$$\alpha(h) = \frac{1}{|\log|h||^b} \text{ with } 1 < b < (a-1)/2 \text{ and } \beta(h) = \exp\left[-\frac{1}{4C}|\log|h||^{a-2b}\right]$$

Then,

$$\mathbb{P}(|Y(t+h) - Y(t)| \geq \alpha(h)) = \mathbb{P}\left(|\xi| \geq \frac{\alpha(h)}{\sqrt{\text{Var}(Y(t+h) - Y(t))}}\right)$$

where ξ stands for standard normal variable. We use the following usual bound for Gaussian tails, valid for $u > 0$:

$$\mathbb{P}(|\xi| \geq u) = 2\mathbb{P}(\xi \geq u) = \sqrt{\frac{2}{\pi}} \int_u^{+\infty} e^{-\frac{1}{2}x^2} dx \leq \sqrt{\frac{2}{\pi}} \frac{1}{u} e^{-\frac{1}{2}u^2}.$$

With the above choice of $\alpha(\cdot)$ and $\beta(\cdot)$, if $|h|$ is small enough one has

$$\frac{\alpha(h)}{\sqrt{\text{Var}(Y(t+h) - Y(t))}} > 1 \text{ and}$$

$$\mathbb{P}(|Y_{t+h} - Y(t)| \geq \alpha(h)) \leq (\text{const}) \beta(h).$$

where (const) denotes a generic constant which may vary from line to line. On the other hand, $\sum_1^\infty \alpha(2^{-n}) < \infty$ and $\sum_1^\infty 2^n \beta(2^{-n}) < \infty$ are easily verified. \square

Some examples.

- **Gaussian Stationary Processes:** Let $\{Y(t) : t \in \mathbb{R}\}$ be a real-valued Gaussian centered stationary process with covariance $\Gamma(\tau) = \mathbb{E}(Y(t)Y(t+\tau))$. Then Condition (1.14) is equivalent to :

$$\Gamma(0) - \Gamma(\tau) \leq \frac{C}{|\log|\tau||^a}$$

for sufficiently small $|\tau|$, with the same meaning for C and a .

- Wiener Process: Take $T = \mathbb{R}^+$. The function $r(s, t) = s \wedge t$ is positive semi-definite. In fact, if $0 \leq s_1 < \dots < s_n$ and $x_1, \dots, x_n \in \mathbb{R}$, one has:

$$(1.15) \quad \sum_{j,k=1}^n (s_j \wedge s_k) x_j x_k = \sum_{k=1}^n (s_k - s_{k-1})(x_k + \dots + x_n)^2 \geq 0,$$

where we have put $s_0 = 0$.

Then, according to Kolmogorov's extension Theorem, there exists a centered Gaussian process $\{Y(t) : t \in \mathbb{R}^+\}$ such that $E(Y(s)Y(t)) = s \wedge t$ for $s, t \geq 0$. One easily checks that this process satisfies the hypothesis in the Corollary 1.7 b), since the random variable $Y(t+h) - Y(t)$, $h \geq 0$ has the normal distribution $N(0, h)$ because of the simple computation

$$E\left([Y(t+h) - Y(t)]^2\right) = t + h - 2t + t = h.$$

It follows from Corollary 1.7 (b) that this process has a continuous version on every interval of the form $[n, n+1]$. The reader will verify that one can also find a version with continuous paths defined on all \mathbb{R}^+ . This version will be called "the Wiener process" and will be denoted $\{W(t) : t \in \mathbb{R}^+\}$.

- Ito integrals : Let $\{W(t) : t \geq 0\}$ be a Wiener process on a probability space (Ω, \mathcal{A}, P) . We define the "filtration" $\{\mathcal{F}_t : t \geq 0\}$ as $\mathcal{F}_t = \tilde{\sigma}\{W(s) : s \leq t\}$ where the notation means the σ -algebra generated by the set of random variables $\{W(s) : s \leq t\}$ (that is, the smallest σ -algebra with respect to which these random variables are all measurable) completed with respect to the probability measure P .

Let $\{a_t : t \geq 0\}$ be a stochastic process adapted to the filtration $\{\mathcal{F}_t : t \geq 0\}$. This means that a_t is \mathcal{F}_t -measurable for each $t \geq 0$. For simplicity we will assume that $\{a_t : t \geq 0\}$ is uniformly locally bounded in the sense that for each $T > 0$ there exists a constant C_T such that $|a_t(\omega)| \leq C_T$ for every ω and all $t \in [0, T]$.

For each $t > 0$ one can define the stochastic Ito integral

$$Y(t) = \int_0^t a_s dW(s)$$

as the limit in $L^2 = L^2(\Omega, \mathcal{A}, P)$ of the Riemann sums

$$S_Q = \sum_{j=0}^{m-1} \tilde{a}_{t_j} (W(t_{j+1}) - W(t_j))$$

when $N_Q = \sup\{(t_{j+1} - t_j) : 0 \leq j < m-1\}$ tends to 0. Here Q denotes the partition $0 = t_0 < t_1 < \dots < t_m = t$ of the interval $[0, t]$ and $\{\tilde{a}_t : t \geq 0\}$ an adapted stochastic process, bounded by the same constant as $\{a_t : t \geq 0\}$ and such that

$$\sum_{j=0}^{m-1} \tilde{a}_{t_j} \mathbf{1}_{\{t_j \leq s < t_{j+1}\}}$$

tends to $\{a_t : 0 \leq s \leq t\}$ in the space $L^2([0, t] \times \Omega, \lambda \times P)$ as $N_Q \rightarrow 0$. λ is Lebesgue measure on the line.

Of course the statements above should be proved to be able to define $Y(t)$ in this way (see for example Mc Kean, 1969). Our aim here is to prove that the process $\{Y(t) : t \geq 0\}$ thus defined has a version with continuous paths. With no loss of generality, we assume that t varies on the interval $[0, 1]$ and apply Corollary 1.7 (a) with $p = 4$.

We will prove that:

$$E\left((Y(t+h) - Y(t))^4\right) \leq (const)h^2$$

For this, it is sufficient to see that if Q is a partition of the interval $[t, t+h]$, $h > 0$,

$$(1.16) \quad E(S_Q^4) \leq (const)h^2$$

where the (*const*) does not depend on t, h, Q and then apply Fatou's Lemma when $N_Q \rightarrow 0$.

Let us compute the left-hand side of (1.16). Put $\Delta_j = W(t_{j+1}) - W(t_j)$. We have:

$$(1.17) \quad \mathbb{E}(S_Q^4) = \sum_{j_1, j_2, j_3, j_4=0}^{m-1} \mathbb{E}(\tilde{a}_{t_{j_1}} \tilde{a}_{t_{j_2}} \tilde{a}_{t_{j_3}} \tilde{a}_{t_{j_4}} \Delta_{j_1} \Delta_{j_2} \Delta_{j_3} \Delta_{j_4})$$

If one of the indices - say j_4 - satisfies $j_4 > j_1, j_2, j_3$ then the corresponding term becomes:

$$\begin{aligned} \mathbb{E}\left(\prod_{h=1}^4 (\tilde{a}_{t_{j_h}} \Delta_{j_h})\right) &= \mathbb{E}\left(\mathbb{E}\left(\prod_{h=1}^4 (\tilde{a}_{t_{j_h}} \Delta_{j_h}) \mid \mathcal{F}_{t_{j_4}}\right)\right) \\ &= \mathbb{E}\left(\prod_{h=1}^3 (\tilde{a}_{t_{j_h}} \Delta_{j_h}) \tilde{a}_{t_{j_4}} \mathbb{E}(\Delta_{j_4} \mid \mathcal{F}_{t_{j_4}})\right) = 0 \end{aligned}$$

since

$$\mathbb{E}(\Delta_j \mid \mathcal{F}_{t_j}) = \mathbb{E}(\Delta_j) = 0 \quad \text{and} \quad \prod_{h=1}^3 (\tilde{a}_{t_{j_h}} \Delta_{j_h}) \tilde{a}_{t_{j_4}} \text{ is } \mathcal{F}_{t_{j_4}} \text{ - measurable.}$$

In a similar way, if $j_4 < j_1 = j_2 = j_3$ (and similarly if any one of the indices is strictly smaller than the others and these are all equal), the corresponding term vanishes since in this case

$$\begin{aligned} \mathbb{E}\left(\prod_{h=1}^4 (\tilde{a}_{t_{j_h}} \Delta_{j_h})\right) &= \mathbb{E}\left(\mathbb{E}\left((\tilde{a}_{t_{j_1}} \Delta_{j_1})^3 \tilde{a}_{t_{j_4}} \Delta_{j_4} \mid \mathcal{F}_{t_{j_1}}\right)\right) \\ &= \mathbb{E}\left(\tilde{a}_{t_{j_1}}^3 \tilde{a}_{t_{j_4}} \Delta_{j_4} \mathbb{E}(\Delta_{j_1}^3 \mid \mathcal{F}_{t_{j_1}})\right) = 0 \end{aligned}$$

because

$$\mathbb{E}(\Delta_j^3 \mid \mathcal{F}_{t_j}) = \mathbb{E}(\Delta_j^3) = 0.$$

The terms with $j_1 = j_2 = j_3 = j_4$ give the sum:

$$\sum_{j=0}^{m-1} \mathbb{E}\left((\tilde{a}_{t_j} \Delta_j)^4\right) \leq C_1^4 \sum_{j=0}^{m-1} 3(t_{j+1} - t_j)^2 \leq 3C_1^4 h^2$$

Finally we have the sum of the terms corresponding to 4-tuples of indices j_1, j_2, j_3, j_4 such that for some permutation (i_1, i_2, i_3, i_4) of $(1, 2, 3, 4)$ one has $j_{i_1}, j_{i_2} < j_{i_3} = j_{i_4}$. This is:

$$6 \sum_{j_3=1}^{m-1} \sum_{0 \leq j_1, j_2 < j_3} \mathbb{E}\left(\tilde{a}_{t_{j_1}} \tilde{a}_{t_{j_2}} \tilde{a}_{t_{j_3}}^2 \Delta_{j_1} \Delta_{j_2} \Delta_{j_3}^2\right).$$

Conditioning on $\mathcal{F}_{t_{j_3}}$ in each term yields for this sum:

$$\begin{aligned} 6 \sum_{j_3=1}^{m-1} \sum_{0 \leq j_1, j_2 < j_3} (t_{j_3+1} - t_{j_3}) \mathbb{E}\left(\tilde{a}_{t_{j_1}} \tilde{a}_{t_{j_2}} \tilde{a}_{t_{j_3}}^2 \Delta_{j_1} \Delta_{j_2}\right) \\ = 6 \mathbb{E}\left(\sum_{j_3=1}^{m-1} (t_{j_3+1} - t_{j_3}) \tilde{a}_{t_{j_3}}^2 \left(\sum_{j=0}^{j_3-1} \tilde{a}_{t_j} \Delta_j\right)^2\right) \\ \leq 6C_1^2 \sum_{j_3=1}^{m-1} (t_{j_3+1} - t_{j_3}) \mathbb{E}\left(\left(\sum_{j=0}^{j_3-1} \tilde{a}_{t_j} \Delta_j\right)^2\right) \\ = 6C_1^2 \sum_{j_3=1}^{m-1} (t_{j_3+1} - t_{j_3}) \sum_{j=0}^{j_3-1} \mathbb{E}\left(\tilde{a}_{t_j}^2\right) (t_{j+1} - t_j) \leq 3C_1^4 h^2. \end{aligned}$$

Using (1.17) one obtains (1.16) and hence, the existence of a version of the Itô integral possessing continuous paths.

Separability. Next, we consider separability of stochastic processes. The separability condition is shaped to avoid the measurability problems that we have already mentioned and to use without further reference, versions of stochastic processes having good path properties. We begin with the definition.

DEFINITION 1.8. *We say that a real-valued stochastic process $\{X(t) : t \in T\}$, T a topological space, is “separable” if there exists a fixed countable subset D of T such that with probability one:*

$$\sup_{t \in V \cap D} X(t) = \sup_{t \in V} X(t) \quad \text{and} \quad \inf_{t \in V \cap D} X(t) = \inf_{t \in V} X(t) \quad \text{for all open sets } V$$

A consequence of the previous Theorem is the following:

PROPOSITION 1.9. *Let $\{Y(t) : t \in I\}$, I an interval in the line, be a separable random process process that satisfies the hypotheses of Theorem 1.6. Then, almost surely, its paths are continuous.*

PROOF. Denote by D the countable set in the definition of separability. With no loss of generality, we may assume that D is dense in I . The theorem states that there exists a version $\{X(t) : t \in I\}$ that has continuous paths, so that

$$P(X(t) = Y(t) \quad \text{for all } t \in D) = 1$$

Let

$$E = \{X(t) = Y(t) \quad \text{for all } t \in D\}$$

and

$$F = \bigcap_{J \subset I, J = (r_1, r_2), r_1, r_2 \in Q} \left\{ \sup_{t \in J \cap D} Y(t) = \sup_{t \in J} Y(t) \quad \text{and} \quad \inf_{t \in J \cap D} Y(t) = \inf_{t \in J} Y(t) \right\}$$

Since $P(E \cap F) = 1$, it is sufficient to prove that if $\omega \in E \cap F$, then $X(s)(\omega) = Y(s)(\omega)$ for all $s \in I$.

So, let $\omega \in E \cap F$ and $s \in I$. For any $\varepsilon > 0$, choose $r_1, r_2 \in Q$ such that

$$s - \varepsilon < r_1 < s < r_2 < s + \varepsilon$$

Then, putting $J = (r_1, r_2)$

$$Y(s)(\omega) \leq \sup_{t \in J} Y(t)(\omega) = \sup_{t \in J \cap D} Y(t)(\omega) = \sup_{t \in J \cap D} X(t)(\omega) \leq \sup_{t \in J} X(t)(\omega)$$

Letting $\varepsilon \rightarrow 0$ it follows that

$$Y(s)(\omega) \leq \limsup_{t \rightarrow s} X(t)(\omega) = X(s)(\omega)$$

since $t \rightsquigarrow X(t)(\omega)$ is continuous:

In a similar way one proves that $Y(s)(\omega) \geq X(s)(\omega)$. □

The separability condition is usually met when the paths have some minimal regularity. See exercise 1.7 at the end of this chapter. For example, if $\{X(t) : t \in \mathbb{R}\}$ is a real-valued process having almost surely càd-làg paths (that is, paths that are right-continuous with left limits), then it is separable. All processes considered in the sequel will be separable.

Some additional remarks and references.

A reference for Kolmogorov’s Extension Theorem and the regularity of paths, at the level of generality we have considered here, is the book by Cramér & Leadbetter (1967), where the reader can find proofs that we have skipped as well as related results, examples and details. For d -parameter Gaussian processes, a subject that we will consider in more detail in Chapter 6, in the stationary case, necessary and sufficient conditions to have continuous paths are due to Fernique (see his St. Flour 1974 lecture notes) and to Talagrand (1987) in the general non-stationary case. In the Gaussian stationary case Belayev (1961) has shown that either with probability 1 the paths

are continuous or with probability 1 the supremum (resp. the infimum) on every interval is $+\infty$ (resp. $-\infty$).

General references on Gaussian processes are also Adler's (1990) and Lifshits'(1995) books.

4.2. Sample path differentiability and Hölder conditions.

In this section we will state some results, without detailed proofs. These follow the same lines of the previous section.

THEOREM 1.10. *Let $\mathcal{Y} = \{Y(t) : t \in [0, 1]\}$ be a real-valued stochastic process that satisfies the hypotheses of Theorem 1.6 and additionally, for any triplet $t - h, t, t + h \in [0, 1]$ one has:*

$$P(|Y(t+h) + Y(t-h) - 2Y(t)| \geq \alpha_1(h)) \leq \beta_1(h)$$

where α_1, β_1 are two even functions, increasing for $h > 0$ and such that

$$\sum_{n=1}^{\infty} 2^n \alpha_1(2^{-n}) < \infty, \quad \sum_{n=1}^{\infty} 2^n \beta_1(2^{-n}) < \infty.$$

Then, there exists a version $\mathcal{X} = \{X(t) : t \in T\}$ of the process \mathcal{Y} such that almost surely the paths of \mathcal{X} are of class C^1 .

SKETCH OF THE PROOF. Consider the sequence $\{Y^{(n)}(t) : t \in [0, 1]\}_{n=1,2,\dots}$ of polygonal processes introduced in the proof of Theorem 1.6. We know that a.s. this sequence converges uniformly to $\mathcal{X} = \{X(t) : t \in [0, 1]\}$, a continuous version of \mathcal{Y} .

Define:

$$\begin{aligned} \tilde{Y}^{(n)}(t) &:= Y^{(n)'}(t^-) \text{ for } 0 < t \leq 1 \text{ (left derivative)} \\ \tilde{Y}^{(n)}(0) &:= Y^{(n)'}(0^+) \text{ (right derivative)} \end{aligned}$$

One can show that the hypotheses imply

- 1) almost surely, as $n \rightarrow \infty$ $\tilde{Y}^{(n)}(\cdot)$ converges uniformly on $[0, 1]$ to a function $\tilde{X}(\cdot)$.
- 2) almost surely, as $n \rightarrow \infty$ $\sup_{t \in [0,1]} |\tilde{Y}^{(n)}(t^+) - \tilde{Y}^{(n)}(t)| \rightarrow 0$.

To complete the proof check that almost surely the function $t \rightsquigarrow \tilde{X}(t)$ is continuous and coincides with the derivative of $X(t)$ at every $t \in [0, 1]$. \square

EXAMPLE 1.2 (Stationary Gaussian Processes). *Let $\mathcal{Y} = \{Y(t) : t \in \mathbb{R}\}$ be a centered stationary Gaussian process with covariance of the form*

$$\Gamma(\tau) = E(Y(t)Y(t+\tau)) = \Gamma(0) - \frac{1}{2}\lambda_2\tau^2 + O\left(\frac{\tau^2}{|\log|\tau||^a}\right)$$

with $\lambda_2 > 0$, $a > 3$. Then there exists a version of \mathcal{Y} with paths of class C^1 .

For the proof apply the previous Theorem.

A related result is the following. The proof is also left to the reader.

PROPOSITION 1.11 (Hölder Conditions). *Assume that*

$$(1.18) \quad E(|Y(t+h) - Y(t)|^p) \leq K|h|^{1+r} \quad \text{for } t, t+h \in [0, 1].$$

where K, p, r are positive constants, $r \leq p$.

Then, there exists a version of the process $\mathcal{Y} = \{Y(t) : t \in [0, 1]\}$ with paths that satisfy a Hölder condition with exponent α for any α such that $0 < \alpha < \frac{r}{p}$.

Note that, for example, this proposition can be applied to the Wiener process (Brownian motion) with $r = (p-2)/2$ showing that it satisfies a Hölder condition for every $\alpha < 1/2$.

4.3. Higher derivatives. Let $\mathcal{X} = \{X(t) : t \in \mathbb{R}\}$ be a stochastic process and assume that for each $t \in \mathbb{R}$ one has $X(t) \in L^2(\Omega, \mathcal{A}, \mathbb{P})$.

DEFINITION 1.12. \mathcal{X} is differentiable in quadratic mean (q.m.) if for all $t \in \mathbb{R}$

$$\frac{X(t+h) - X(t)}{h}$$

converges in quadratic mean as $h \rightarrow 0$ to some limit that will be denoted $X'(t)$.

The stability of Gaussian random variables under passage to the limit, implies that the derivative in q.m. of a Gaussian process remains Gaussian.

PROPOSITION 1.13. Let $\mathcal{X} = \{X(t) : t \in \mathbb{R}\}$ be a stochastic process with mean $m(t)$ and covariance $r(s, t)$ and suppose that m is \mathcal{C}^1 and that r is \mathcal{C}^2 . Then, \mathcal{X} is differentiable in quadratic mean.

PROOF. We use the following result which is easy to prove: the sequence Z_1, \dots, Z_n of real random variables converges in q.m. if and only if there exists some constant C such that $E(Z_m \cdot Z_n) \rightarrow C$ as the pair (m, n) tends to infinity.

Since $m(t)$ is differentiable it can be subtracted from $X(t)$ without changing its differentiability. So we can assume that the process is centered. Then for all real h and k

$$E\left(\frac{X(t+h) - X(t)}{h} \cdot \frac{X(t+k) - X(t)}{k}\right) = \frac{1}{hk} [r(t+h, t+k) - r(t, t+k) - r(t, t+h) + r(t, t)] \\ \rightarrow r_{11}(t, t) \text{ as } (k, h) \rightarrow (0, 0),$$

where $r_{11}(s, t) := \frac{\partial^2 r}{\partial s \partial t}(s, t)$. This shows differentiability in q.m. \square

We assume, using the remark in the proof above, that \mathcal{X} is centered and satisfies the conditions of the proposition. It is easy to prove that

$$E(X(s)X'(t)) = r_{01}(s, t) := \frac{\partial r}{\partial t}(s, t).$$

and similarly, that the covariance of $\mathcal{X}' = \{X'(t) : t \in \mathbb{R}\}$ is $r_{11}(s, t)$. Let now \mathcal{X} be a Gaussian process and \mathcal{X}' its derivative in quadratic mean. If this one satisfies, for example, the criterion in Corollary 1.7 b), then it admits a continuous version $\mathcal{Y}' = \{Y'(t) : t \in \mathbb{R}\}$. Set

$$Y(t) := X(0) + \int_0^t Y'(s) ds$$

Clearly \mathcal{Y} has \mathcal{C}^1 paths and $E(X(s), Y(s)) = r(s, 0) + \int_0^s r_{01}(s, t) dt = r(s, s)$. In the same way, $E(Y(s)^2) = r(s, s)$, so that $E([X(s) - Y(s)]^2) = 0$. As a consequence, \mathcal{X} admits a version with \mathcal{C}^1 paths.

Using this construction inductively, one can prove the following:

- Let \mathcal{X} be a Gaussian process with mean \mathcal{C}^k and covariance \mathcal{C}^{2k} and such that its k -th derivative in quadratic mean satisfies the weak condition of Corollary 1.7 b). Then, \mathcal{X} admits a version with paths of class \mathcal{C}^k .
- If \mathcal{X} is a Gaussian process with mean of class \mathcal{C}^∞ and covariance of class \mathcal{C}^∞ , then \mathcal{X} admits a version with paths of class \mathcal{C}^∞ .

In the converse direction, regularity of the paths implies regularity of the expectation and of the covariance function. For example, if \mathcal{X} has continuous sample paths the mean and the variance are continuous. In fact, if $t_n, n = 1, 2, \dots$ converges to t , then $X(t_n)$ converges almost surely to $X(t)$, hence, also in distribution. Using the form of the Fourier transform of the Gaussian distribution, one easily proves that this implies convergence of the mean and the variance. Since for Gaussian variables, all the moments are polynomial functions of the mean and the variance, they are also continuous.

If the process has differentiable sample paths, in a similar way one shows the convergence

$$\frac{m(t+h) - m(t)}{h} \rightarrow E(X'(t))$$

as $h \rightarrow 0$, showing that the mean is differentiable.

For the covariance, restricting ourselves to stationary Gaussian processes defined on the real line, without loss of generality we may assume that the process is centered. Put $\Gamma(t) = r(s, s + t)$. The convergence in distribution of $(X(h) - X(0))/h$ to $X'(0)$ plus the Gaussianity, imply that $\text{Var}((X(h) - X(0))/h)$ has a finite limit as $h \rightarrow 0$. On the other hand,

$$\text{Var}\left(\frac{X(h) - X(0)}{h}\right) = 2 \int_{-\infty}^{+\infty} \frac{1 - \cos(hx)}{h^2} \mu(dx),$$

where μ is the spectral measure.

Letting $h \rightarrow 0$ and applying Fatou's Lemma, it follows that

$$\lambda_2 = \int_{-\infty}^{+\infty} x^2 \mu(dx) \leq \liminf_{h \rightarrow 0} \text{Var}\left(\frac{X(h) - X(0)}{h}\right) < \infty.$$

Using the result in Exercice 1.4, Γ is of class \mathcal{C}^2 .

This argument can be used in a similar form to show that if the process has paths of class \mathcal{C}^k , then the covariance is of class \mathcal{C}^{2k} . As a conclusion, roughly speaking, for Gaussian stationary processes, the order of differentiability of the sample paths is half of the order of differentiability of the covariance.

4.4. More general tools.

In this section we consider the case when the parameter of the process lies in \mathbb{R}^d or, more generally, in some general metric space. We begin with an extension of Theorem 1.6.

THEOREM 1.14. *Let $\mathcal{Y} = \{Y(t) : t \in [0, 1]^d\}$ be a real-valued random field that satisfies the condition*

(K_d) : *For each pair $t, t + h \in [0, 1]^d$,*

$$\mathbb{P}\{|Y(t + h) - Y(t)| \geq \alpha(\bar{h})\} \leq \beta(\bar{h})$$

where $h = (h_1, \dots, h_d)$, $\bar{h} = \sup_{1 \leq i \leq d} |h_i|$ and α, β are even real-valued functions defined on $[-1, 1]$, increasing on $[0, 1]$ which verify

$$\sum_{n=1}^{\infty} \alpha(2^{-n}) < \infty, \quad \sum_{n=1}^{\infty} 2^{dn} \beta(2^{-n}) < \infty$$

Then, there exists a version $\mathcal{X} = \{X(t) : t \in [0, 1]^d\}$ of the process \mathcal{Y} such that the paths $t \rightsquigarrow X(t)$ are continuous on $[0, 1]^d$.

PROOF. The main change with respect to the proof of Theorem 1.6 is that we replace the polygonal approximation, adapted to one-variable functions by another interpolating procedure. Denote by \mathcal{D}_n the set of dyadic points of order n in $[0, 1]^d$, that is

$$\mathcal{D}_n = \{t = (t_1, \dots, t_d) : t_i = \frac{k_i}{2^n}, k_i \text{ integers}, 0 \leq k_i \leq 2^n, i = 1, \dots, d\}.$$

Let $f : [0, 1]^d \rightarrow \mathbb{R}$ be a function. For each $n = 1, 2, \dots$ one can construct a function $f^{(n)} : [0, 1]^d \rightarrow \mathbb{R}$ with the following properties:

- $f^{(n)}$ is continuous,
- $f^{(n)}(t) = f(t)$ for all $t \in \mathcal{D}_n$,
- $\|f^{(n+1)} - f^{(n)}\|_{\infty} = \max_{t \in \mathcal{D}_{n+1} \setminus \mathcal{D}_n} |f(t) - f^{(n)}(t)|$

where $\|\cdot\|_{\infty}$ denotes sup-norm on $[0, 1]^d$. A way to define $f^{(n)}$ is the following: let us consider a cube $\mathcal{C}_{t,n}$ of the n -th order partition of $[0, 1]^d$, that is:

$$\mathcal{C}_{t,n} = t + [0, \frac{1}{2^n}]^d,$$

where $t \in \mathcal{D}_n$ and the obvious notation for the sum. For each vertex τ put

$$f^{(n)}(\tau) = f(\tau).$$

Now, for each permutation π of $\{1, 2, \dots, d\}$ let \mathcal{S}_π be the simplex

$$\mathcal{S}_\pi = \left\{ t + s : s = (s_{\pi(1)}, \dots, s_{\pi(d)}), 0 \leq s_{\pi(1)} \leq \dots \leq s_{\pi(d)} \leq \frac{1}{2^n} \right\}.$$

It is clear that $\mathcal{C}_{t,n}$ is the union of the \mathcal{S}_π 's over all permutations. Extend now -in a unique way - $f^{(n)}$ to \mathcal{S}_π as an affine function. It is then easy to verify the above mentioned properties and that

$$\|f^{(n+1)} - f^{(n)}\|_\infty \leq d \sup_{s,t \in \mathcal{D}_{n+1}, |t-s|=2^{-(n+1)}} |f(s) - f(t)|.$$

The rest of the proof is essentially similar to that of Theorem 1.6. \square

From this we deduce easily

COROLLARY 1.15. *If the process $\mathcal{Y} = \{Y(t) : t \in [0, 1]^d\}$ verifies one of these two conditions*

$$(1.19) \quad \mathbb{E}(|Y(t+h) - Y(t)|^p) \leq \frac{K_d |h|^d}{|\log |h||^{1+r}}$$

where p, r, K are positive constants, $p < r$,

b) If \mathcal{Y} is Gaussian, $m(t) = \mathbb{E}(Y(t))$ is continuous and

$$(1.20) \quad \text{Var}(Y(t+h) - Y(t)) \leq \frac{C}{|\log |h||^a}$$

for all t and sufficiently small h and $a > 3$,

then the process has a version with continuous paths.

Note that the case of processes with values in $\mathbb{R}^{d'}$ needs not to be considered separately, since continuity can be addressed coordinate by coordinate. For Hölder regularity we have

PROPOSITION 1.16. *Let $\mathcal{Y} = \{Y(t) : t \in [0, 1]^d\}$ be a real-valued stochastic process with continuous paths such that for some $q > 1$, $\alpha > 0$,*

$$\mathbb{E}(|Y(s) - Y(t)|^q) \leq (\text{const}) \|s - t\|^{d+\alpha}$$

Then almost surely \mathcal{Y} has Hölder paths with exponent $\alpha/(2q)$.

Until now, we have deliberately chosen elementary methods which apply to general random processes, not necessarily Gaussian. In the Gaussian case, even when the parameter varies in a set that does not have a restricted geometric structure, the question of continuity can be addressed using specific methods. As we have remarked several times already, we only need to consider centered processes.

Let $\{X(t) : t \in T\}$ be a centered Gaussian process taking values in \mathbb{R} . We assume that T is some metric space with distance denoted by τ . On T we define the canonical distance d

$$d(s, t) := \sqrt{E(X(t) - X(s))^2}.$$

In fact d is a pseudo distance because two distinct points can be at d distance zero. A first point is that when the covariance $r(s, t)$ function is τ -continuous, which is the only relevant case (otherwise there is no hope of having continuous paths), d -continuity and τ -continuity are equivalent. The reader is referred to Adler (1990) for complements and proofs.

DEFINITION 1.17. *Let (T, d) be a metric space. For $\varepsilon > 0$ denote by $N_\varepsilon = N(T, d, \varepsilon)$ the minimum number of closed balls of radius ε with which we can cover T (the value of N_ε can be $+\infty$).*

We have the following theorem

THEOREM 1.18 (Dudley (1973)). *A sufficient condition for $\{X(t) : t \in T\}$ to have continuous sample paths is*

$$\int_0^{+\infty} (\log(N(\varepsilon)))^{\frac{1}{2}} < \infty$$

$\log(N(\varepsilon))$ is called the “entropy” of the set T .

A very important fact is that this condition is necessary in some relevant cases:

THEOREM 1.19 (Fernique (1974)). *Let $\{X(t) : t \in T\}$, T compact a subset of \mathbb{R}^d be a **stationary Gaussian process**. Then the three following statements are equivalent*

- *almost surely $X(\cdot)$ is bounded.*
- *almost surely $X(\cdot)$ is continuous.*
- $\int_0^{+\infty} (\log(N(\varepsilon)))^{\frac{1}{2}} < \infty$.

This condition can be compared with Kolmogorov’s Theorem. The reader can check that Theorem 1.19 permits to weaken the condition of Corollary 1.7 b) to $a > 1$. On the other hand, one can construct counterexamples - that is, processes not having continuous paths, such that (1.14) holds true with $a = 1$. This shows that the condition of Corollary 1.7 b) is nearly optimal and sufficient for most applications. When the Gaussian process is no more stationary, M. Talagrand has given necessary and sufficient condition for sample path continuity in terms of the existence of majorizing measures (see Talagrand, 1987).

The problem of differentiability can be addressed in the same manner as for $d = 1$. A sufficient condition for a Gaussian process to have a version with \mathcal{C}^k sample paths, is its mean to be \mathcal{C}^k , its covariance \mathcal{C}^{2k} and its k -th derivative in quadratic mean to satisfy some of the criteria of continuity above.

4.5. Tangencies. Local extrema.

In this section we give two classical results that will be used several times in this book. The first one gives a simple sufficient condition for a one-parameter random process not to have, almost surely, critical points at a certain specified level. The second result states that under mild conditions, a Gaussian process defined on a quite general parameter set, with probability one does not have local extrema at a given level.

We will use systematically the following notation: If ξ is a random variable with values in \mathbb{R}^d and its distribution has a density with respect to Lebesgue measure, this density will be denoted as

$$p_\xi(x), \quad x \in \mathbb{R}^d.$$

PROPOSITION 1.20 (Bulinskaya (1961)). *Let $\{X(t) : t \in I\}$ be a stochastic process with paths of class C^1 defined on the interval I of the real line. Assume that for each $t \in I$, the random variable $X(t)$ has a density $p_{X(t)}(x)$ which is bounded as t varies in a compact subset of I and x in a neighborhood of $u \in \mathbb{R}$.*

Then

$$P(T_u^X \neq \emptyset) = 0$$

where $T_u^X = \{t : t \in I, X(t) = u, X'(t) = 0\}$ is the set of critical points with value u of the random path $X(\cdot)$.

PROOF. It suffices to prove that $P(T_u^X \cap J \neq \emptyset) = 0$ for any compact subinterval J of I . Let ℓ be the length of J and $t_0 < t_1 < \dots < t_m$ a uniform partition of J , i.e. $t_{j+1} - t_j = \frac{\ell}{m}$ for $j = 0, 1, \dots, m - 1$.

Denote by $\omega_{X'}(\delta, J)$ the modulus of continuity X' on the interval J and $E_{\delta, \varepsilon}$ the event:

$$E_{\delta, \varepsilon} = \{\omega_{X'}(\delta, J) \geq \varepsilon\}$$

Let $\varepsilon > 0$ be given; choose $\delta > 0$ so that $P(E_{\delta,\varepsilon}) < \varepsilon$ and m so that $\frac{\ell}{m} < \delta$. We have:

$$\begin{aligned} P(T_u^X \cap J \neq \emptyset) &\leq P(E_{\delta,\varepsilon}) + \sum_{j=0}^{m-1} P(\{T_u^X \cap [t_j, t_{j+1}] \neq \emptyset\} \cap E_{\delta,\varepsilon}^C) \\ &< \varepsilon + \sum_{j=0}^{m-1} P(|X(t_j) - u| \leq \varepsilon \frac{\ell}{m}) = \varepsilon + \sum_{j=0}^{m-1} \int_{|x-u| \leq \varepsilon \frac{\ell}{m}} p_{X(t_j)}(x) dx \end{aligned}$$

If C is an upper bound for $p_{X(t)}(x)$, $t \in J$, $|x - u| \leq \varepsilon \ell$, we obtain

$$P(T_u^X \cap J \neq \emptyset) \leq \varepsilon + C\varepsilon\ell$$

Since $\varepsilon > 0$ is arbitrary, the result follows. \square

The second result is an extension of Ylvisaker's Theorem, which has the following statement:

THEOREM 1.21 (Ylvisaker (1968)). *Let $\{Z(t) : t \in T\}$ be a real-valued Gaussian process indexed on a compact separable topological space T , having continuous paths and $\text{Var}(Z(t)) > 0$ for all $t \in T$.*

Then, for fixed $u \in \mathbb{R}$, one has $P(E_u^Z \neq \emptyset) = 0$, where E_u^Z is the set of local extrema of $Z(\cdot)$ having value equal to u .

The extension is the following:

THEOREM 1.22. *Let $\{Z(t) : t \in T\}$ be a real-valued Gaussian process on some parameter set T and denote by $M^Z = \sup_{t \in T} Z(t)$ its supremum (which takes values in $\mathbb{R} \cup \{+\infty\}$).*

We assume that there exists a non-random countable set \mathcal{D} , $\mathcal{D} \subset T$, such that a.s. $M^Z = \sup_{t \in \mathcal{D}} Z(t)$.

Assume further that there exist $\sigma_0^2 > 0$, $m_- > -\infty$ such that

$$m(t) = \mathbf{E}(Z(t)) \geq m_-$$

$$\sigma^2(t) = \text{Var}(Z(t)) \geq \sigma_0^2 \text{ for every } t \in T.$$

Then the distribution of the random variable M^Z is the sum of an atom at $+\infty$ and a - possibly defective - probability measure on \mathbb{R} which has a locally bounded density.

PROOF. Step 1. Suppose first that $\{X(t) : t \in T\}$ satisfies the hypotheses of the theorem, and moreover:

$$\text{Var}(X(t)) = 1 ; \mathbf{E}(X(t)) \geq 0,$$

for every $t \in T$.

We prove that the supremum M^X has a density p_{M^X} which satisfies the inequality:

$$(1.21) \quad p_{M^X}(u) \leq \psi(u) := \frac{\exp(-u^2/2)}{\int_u^\infty \exp(-v^2/2) dv} \text{ for every } u \in \mathbb{R}.$$

Let $\mathcal{D} = \{t_k\}_{k=1,2,\dots}$. Almost surely, $M^X = \sup\{X(t_1) \dots X(t_n) \dots\}$. We put:

$$M_n := \sup_{1 \leq k \leq n} X(t_k).$$

Since the joint distribution of $X(t_k)$, $k = 1, \dots, n$ is Gaussian, for any choice of $k, \ell = 1, \dots, n$; $k \neq \ell$, the probability $P\{X(t_k) = X(t_\ell)\}$ is equal to 0 or to 1. Hence, possibly excluding some of these random variables, we may assume that these probabilities are all equal to 0 without changing the value of M_n on a set of probability 1. Then, the distribution of the random variable M_n has a density $g_n(\cdot)$ that can be written as:

$$g_n(x) = \sum_{k=1}^n P(X(t_j) < x, j = 1, \dots, n; j \neq k | X(t_k) = x) \frac{e^{-\frac{1}{2}(x-m(t_k))^2}}{\sqrt{2\pi}} = \varphi(x)G_n(x)$$

where φ denotes the standard normal density and

$$(1.22) \quad G_n(x) = \sum_{k=1}^n P(Y_j < x - m(t_j), j = 1, \dots, n; j \neq k | Y_k = x - m(t_k)) e^{x \cdot m(t_k) - \frac{1}{2} m^2(t_k)}$$

with

$$Y_j = X(t_j) - m(t_j) \quad (j = 1, \dots, n).$$

Let us prove that $x \rightsquigarrow G_n(x)$ is an increasing function.

Since $m(t) \geq 0$ it is sufficient the conditional probability in each term of (1.22) to be increasing as a function of x . Write the Gaussian regression

$$Y_j = Y_j - c_{jk} Y_k + c_{jk} Y_k \quad \text{with} \quad c_{jk} = E(Y_j Y_k)$$

where the random variables $Y_j - c_{jk} Y_k$ and Y_k are independent. Then, the conditional probability becomes

$$P(Y_j - c_{jk} Y_k < x - m(t_j) - c_{jk}(x - m(t_k)), j = 1, \dots, n; j \neq k).$$

This probability increases with x because $(1 - c_{jk}) \geq 0$ due to the Cauchy-Schwarz inequality. Now, if $a, b \in \mathbb{R}, a < b$, since $M_n \uparrow M^X$:

$$P\{a < M^X \leq b\} = \lim_{n \rightarrow \infty} P\{a < M_n \leq b\}$$

Using the monotonicity of G_n ,

$$G_n(b) \int_b^{+\infty} \varphi(x) dx \leq \int_b^{+\infty} G_n(x) \varphi(x) dx = \int_b^{+\infty} g_n(x) dx \leq 1.$$

So that

$$\begin{aligned} P\{a < M_n \leq b\} &= \int_a^b g_n(x) dx \leq G_n(b) \int_a^b \varphi(x) dx \\ &\leq \int_a^b \varphi(x) dx \left(\int_b^{+\infty} \varphi(x) dx \right)^{-1}. \end{aligned}$$

This proves (1.21).

Step 2. Let now Z satisfy the hypotheses of the theorem, without assuming the added ones in Step 1. For given $a, b \in \mathbb{R}, a < b$, choose $A \in \mathbb{R}^+$ so that $|a| < A$ and consider the process:

$$X(t) = \frac{Z(t) - a}{\sigma(t)} + \frac{|m_-| + A}{\sigma_0}.$$

Clearly for every $t \in T$:

$$E(X(t)) = \frac{m(t) - a}{\sigma(t)} + \frac{|m_-| + A}{\sigma_0} \geq -\frac{|m_-| + |a|}{\sigma_0} + \frac{|m_-| + A}{\sigma_0} \geq 0,$$

and

$$\text{Var}(X(t)) = 1,$$

so that (1.21) holds for the process X .

On the other hand:

$$\{a < M^Z \leq b\} \subset \{\mu_1 < M^X \leq \mu_2\},$$

where

$$\mu_1 = \frac{|m_-| + A}{\sigma_0}, \quad \mu_2 = \frac{|m_-| + A}{\sigma_0} + \frac{b - a}{\sigma_0}$$

It follows that

$$P\{a < M^Z \leq b\} \leq \int_{\mu_1}^{\mu_2} \psi(u) du = \int_a^b \frac{1}{\sigma_0} \psi\left(\frac{v - a + |m_-| + A}{\sigma_0}\right) dv.$$

which proves the statement. \square

Theorem 1.21 follows directly from Theorem 1.22, since under the hypotheses of Theorem 1.21, we can write:

$$\{E_u^X \neq \phi\} \subset \bigcup_{U \in \mathcal{F}} (\{M_U = u\} \cup \{m_U = u\}).$$

where M_U (resp. m_u) is the maximum (resp. the minimum) of the process on the set U and \mathcal{F} denotes a countable family of open sets being a basis for the topology of T .

REMARK 1.1. *We will come back later to the subject of the regularity properties of the probability distribution of the supremum of a Gaussian process in the chapters dedicated to this subject.*

Exercises

EXERCISE 1.1. Let $T = \mathcal{N}$ be the set of natural numbers. Prove that the following sets belong to $\sigma(\mathcal{C})$.

- c_0 (the set of real-valued sequences $\{a_n\}$ such that $a_n \rightarrow 0$)
- ℓ^2 (the set of real-valued sequences $\{a_n\}$ such that $\sum_n |a_n|^2 < \infty$)
- the set of real-valued sequences $\{a_n\}$ such that $\overline{\lim}_{n \rightarrow \infty} a_n \leq 1$

Suggestion for the first: Note that

$$c_0 = \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n \geq m} \left\{ |a_n| < \frac{1}{k} \right\}$$

EXERCISE 1.2. Take $T = \mathbb{R}$, $\mathcal{T} = \mathcal{B}_{\mathbb{R}}$. Then if for each $\omega \in \Omega$ the function

$$(1.23) \quad t \rightsquigarrow X(t, \omega)$$

”the path corresponding to ω ” is a continuous function, then the process is bimeasurable.

In fact, check that

$$X(t, \omega) = \lim_{n \rightarrow +\infty} X^{(n)}(t, \omega)$$

where for $n = 1, 2, \dots$, $X^{(n)}$ is defined by

$$X^{(n)}(t, \omega) = \sum_{k=-\infty}^{k=+\infty} X_{\frac{k}{2^n}}(\omega) \mathbf{1}_{\{\frac{k}{2^n} \leq t < \frac{k+1}{2^n}\}}$$

which is obviously measurable as a function of the pair (t, ω) . So, the limit function X has the required property.

If one replaces continuity of the path (1.23) by some other regularity properties such as right continuity, bi-measurability follows in a similar way.

EXERCISE 1.3. Let U be a random variable defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$, having uniform distribution on the interval $[0, 1]$.

Consider the two stochastic processes:

$$\begin{aligned} Y(t) &= \mathbf{1}_{t=U} \\ X(t) &\equiv 0 \end{aligned}$$

The process $Y(t)$ is sometimes called “the random parasite”.

- (1) Prove that for all $t \in [0, 1]$, a.s. $X(t) = Y(t)$.
- (2) Deduce that the processes $X(t)$ and $Y(t)$ have the same probability distribution \mathbb{P} on $\mathbb{R}^{[0,1]}$ equipped with its Borel σ -algebra.
- (3) Notice that for each ω in the probability space, $\sup_{t \in [0,1]} Y(t) = 1$ and $\sup_{t \in [0,1]} X(t) = 0$, so that the suprema of both processes are completely different. Is there a contradiction with the previous point?

EXERCISE 1.4. Let μ be a Borel probability measure on the real line and Γ its Fourier transform, that is:

$$\Gamma(\tau) = \int_{\mathbb{R}} \exp(i\tau x) \mu(dx)$$

a) Prove that if

$$\lambda_k = \int_{\mathbb{R}} |x|^k \mu(dx) < \infty$$

for some positive integer k , then the covariance $\Gamma(\cdot)$ is of class C^k and

$$\Gamma^{(k)}(\tau) = \int_{\mathbb{R}} (ix)^k \exp(i\tau x) \mu(dx).$$

b) Prove that if k is even $k = 2p$ the reciprocal is true: If Γ is of class C^{2p} , then λ_{2p} is finite and

$$\Gamma(t) = 1 - \lambda_2 \frac{t^2}{2!} + \lambda_4 \frac{t^4}{4!} + \cdots + (-1)^{2p} \lambda_{2p} \frac{t^{2p}}{(2p)!} + o(t^{2p})$$

Hint for b): use induction on p , suppose that λ_k is infinite, then for every $A > 0$, one can find some $M > 0$ such that

$$\int_M^M x^k \mu(dx) \geq A.$$

Show that it implies that

$$(-1)^k \frac{k!}{t^k} \left[\Gamma(t) - \left(1 - \lambda_2 \frac{t^2}{2!} + \cdots + (-1)^{k-2} \lambda_{k-2} \frac{t^{k-2}}{(k-2)!} \right) \right]$$

has a limit, when t tends to zero, greater than A which contradicts differentiability.

c) When k is odd the result is false, see Feller (1966, Chap. XVII example (c)).

EXERCISE 1.5. Let $\{\xi_n\}_{n=1,2,\dots}$ be a sequence of random vectors defined on some probability space taking values in \mathbb{R}^d , and assume that $\xi_n \rightarrow \xi$ in probability, for some random vector ξ . Prove that if each ξ_n is Gaussian, then ξ is also Gaussian.

EXERCISE 1.6. Prove the following statements on the process defined by (1.10)

a) For each $t \in T$ the series (1.10) converges almost surely.

b) Almost surely, the function $t \rightsquigarrow X(t)$ is in H and $\|X(\cdot)\|_H^2 = \sum_{n=1}^{\infty} c_n \xi_n^2$

c) $\{\varphi_n\}_{n=1,2,\dots}$ are eigenfunctions - with eigenvalues $\{c_n\}_{n=1,2,\dots}$ respectively - of the linear operator $A : H \rightarrow H$ defined by

$$(Af)(s) = \int_T r(s,t) f(t) \rho(dt).$$

EXERCISE 1.7. Let $\{X(t) : t \in T\}$ be a stochastic process defined on some separable topological space T .

a) Prove that if $X(t)$ has continuous paths, then it is separable. b) Let $T = \mathbb{R}$. Prove that if the paths of $X(t)$ are càd-làg, then $X(t)$ is separable.

EXERCISE 1.8. Let $\{X(t) : t \in \mathbb{R}^d\}$ be a separable stochastic process defined on some (complete) probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

a) Prove that the subset of Ω $\{X(\cdot) \text{ is continuous}\}$ is in \mathcal{A} .

b) Prove that the conclusion in a) remains valid if one replaces "continuous" by "upper continuous", or "lower continuous, or "continuous on the right" (a real-valued function f defined on \mathbb{R}^d is said to be continuous on the right if for each t , $f(t)$ is equal to the limit of $f(s)$ when each coordinate of s tends to the corresponding coordinate of t on its right).

EXERCISE 1.9. Show that in the case of the Wiener process, condition (1.18) holds for every $p \geq 2$, with $r = \frac{p}{2} - 1$. Hence, the proposition implies that almost surely, the paths of the Wiener process satisfy a Hölder condition with exponent α , for every $\alpha < \frac{1}{2}$.

EXERCISE 1.10. (Wiener integral) Let $\{W_1(t) : t \geq 0\}$, $\{W_2(t) : t \geq 0\}$ be two independent Wiener processes defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$, and denote $\{W(t) : t \in \mathbb{R}\}$ the process defined as:

$$W(t) = W_1(t) \text{ if } t \geq 0, \text{ and } W(t) = W_2(-t) \text{ if } t \leq 0.$$

$L^2(\mathbb{R}, \lambda)$ denotes the standard L^2 -space of real-valued measurable functions on the real line with respect to Lebesgue measure and $L^2(\Omega, \mathcal{A}, \mathbb{P})$ the L^2 of the probability space. $\mathcal{C}_K^1(\mathbb{R})$ denotes the subspace of $L^2(\mathbb{R}, \lambda)$ of \mathcal{C}^1 -functions with compact support.

Define the function $I : \mathcal{C}_K^1(\mathbb{R}) \rightarrow L^2(\Omega, \mathcal{A}, \mathbb{P})$ as:

$$(1.24) \quad I(f) = - \int_{\mathbb{R}} f'(t)W(t)dt$$

for each non-random $f \in \mathcal{C}_K^1(\mathbb{R})$. (1.24) is well defined for each $\omega \in \Omega$ since the integrand is a continuous function with compact support.

(a) Prove that I is an isometry, in the sense that

$$\int_{\mathbb{R}} f^2(t)dt = \mathbb{E}(I^2(f)).$$

(b) Show that for each f , $I(f)$ is a centered Gaussian random variable. Moreover, for any choice of $f_1, \dots, f_p \in \mathcal{C}_K^1(\mathbb{R})$, the joint distribution of $(I(f_1), \dots, I(f_p))$ is centered Gaussian. Compute its covariance matrix.

(c) Prove that I admits a unique isometric extension \tilde{I} to $L^2(\mathbb{R}, \lambda)$ such that:

(1) $\tilde{I}(f)$ is a centered Gaussian random variable with variance equal to $\int_{\mathbb{R}} f^2(t)dt$.

Similarly for joint distributions.

(2) $\int_{\mathbb{R}} f(t)g(t)dt = \mathbb{E}(\tilde{I}(f)\tilde{I}(g))$.

[Comment: $\tilde{I}(f)$ is called the “Wiener integral of f ”].

EXERCISE 1.11. (Fractional Brownian motion)

Let H be a real number, $0 < H < 1$.

We use the notation and definitions of the previous exercise.

(a) For $t \geq 0$ define the function $K_t : \mathbb{R} \rightarrow \mathbb{R}$:

$$K_t(u) = [(t-u)^{H-1/2} - (-u)^{H-1/2}] \mathbf{1}_{u < 0} + (t-u)^{H-1/2} \mathbf{1}_{0 < u < t}.$$

Prove that $K_t \in L^2(\mathbb{R}, \lambda)$.

(b) For $t \geq 0$, define the Wiener integral $\tilde{I}(K_t)$ and for $s, t \geq 0$, prove the formula:

$$\mathbb{E}(\tilde{I}(K_s)\tilde{I}(K_t)) = \frac{C_H}{2} [s^{2H} + t^{2H} - |t-s|^{2H}]$$

where C_H is a positive constant depending only on H . Compute C_H .

(c) Prove that the stochastic process $\{C_H^{-1/2} \tilde{I}(K_t) : t \geq 0\}$ has a version with continuous paths.

This normalized version with continuous paths is usually called the “fractional Brownian motion with Hurst exponent H ”, and is denoted $\{W_H(t) : t \geq 0\}$.

(d) Show that if $H = 1/2$, then $\{W_H(t) : t \geq 0\}$ is the standard Wiener process.

(e) Prove that for any $\delta > 0$, almost surely the paths of the fractional Brownian motion with Hurst exponent H satisfy a Hölder condition with exponent $H - \delta$.

EXERCISE 1.12. Local time

Let $\{W(t) : t \geq 0\}$ be a Wiener process defined in some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. For $u \in \mathbb{R}$, I an interval $I \subset [0, +\infty]$ and $\delta > 0$, define:

$$\mu_\delta(u, I) = \frac{1}{2\delta} \int_I \mathbf{1}_{|W(t)-u|<\delta} dt = \frac{1}{2\delta} \lambda(\{t \in I : |W(t) - u| < \delta\}).$$

(a) Prove that for fixed u and I , $\mu_\delta(u, I)$ converges in $L^2(\Omega, \mathcal{A}, \mathbb{P})$ as $\delta \rightarrow 0$.

Denote the limit by $\mu_0(u, I)$

(Hint: Use Cauchy's criterion).

(b) Denote $Z(t) = \mu_0(u, [0, t])$. Prove that the random process $Z(t) : t \geq 0$ has a version with continuous paths. We will call this version the “local time of the Wiener process at the level u ”, and denote by $L^W(u, t)$.

(c) For fixed u , $L^W(u, t)$ is a continuous increasing function of $t \geq 0$. Prove that almost surely, it induces a measure on \mathbb{R}^+ which is singular with respect to Lebesgue measure, that is, its support is contained in a set of Lebesgue measure zero.

(d) Study the Hölder continuity properties of $L^W(u, t)$.

For future reference, with a slight abuse of notation, we will write, for any interval $I = [t_1, t_2]$, $0 \leq t_1 \leq t_2$:

$$L^W(u, I) = L^W(u, t_2) - L^W(u, t_1).$$

Basic Inequalities for Gaussian Processes

This chapter is on inequalities for the probability distribution of the supremum of Gaussian processes. Among the numerous results giving upper and lower bounds, we have chosen the ones we consider to be more useful for the subjects considered in this book: comparison inequalities, isoperimetric inequalities and their applications to obtain bounds for the tails of the distribution of the supremum and its moments.

The results in this chapter are very general, in the sense that beyond Gaussianity and almost sure boundedness of the paths, we do not require the random function to satisfy other hypotheses. They are essential basic tools and, at the same time, provide bounds that may turn out to be rough when applied to special families of random functions. One of our purposes in the next chapters is to refine these inequalities under additional hypotheses, as explained in the Introduction.

A good part of the theory was already well-established more than 30 years ago. However, some results have been significantly improved more recently. Two relevant examples of this evolution are the Li-Shao comparison inequality and the C. Borell proof of the Ehrhard conjecture, which we consider in Sections 1 and 2 respectively.

1. Slepian type inequalities.

LEMMA 2.1 (Li and Shao (2002)). *Let $X := (X_1, \dots, X_n)^T$, and $Y = (Y_1, \dots, Y_n)^T$ be two centered Gaussian random vectors in \mathbb{R}^n , $n \geq 2$. Denote*

$$\Sigma^X = ((r_{jk}^X))_{j,k=1,\dots,n} \text{ with } r_{jk}^X = E(X_j X_k).$$

and use similar notation for Y .

We will assume that $r_{jj}^X = r_{jj}^Y$ for all $j = 1, \dots, n$, i.e. that the variances are respectively equal, and with no loss of generality for our purposes, that their common value is equal to 1.

Then, for any choice of the real numbers a_1, \dots, a_n one has:

$$(2.1) \quad \begin{aligned} & \mathbb{P}\{X_1 \leq a_1, \dots, X_n \leq a_n\} - \mathbb{P}\{Y_1 \leq a_1, \dots, Y_n \leq a_n\} \\ & \leq \frac{1}{2\pi} \sum_{1 \leq i < j \leq n} (\arcsin(r_{ij}^X) - \arcsin(r_{ij}^Y))^+ \exp\left(-\frac{a_i^2 + a_j^2}{2(1 + \rho_{ij})}\right), \end{aligned}$$

where $\rho_{ij} = \max(|r_{ij}^X|, |r_{ij}^Y|)$.

This lemma, which is known under the generic name of “Normal Comparison Lemma”, has a quite long history. As far as we know, its first version is due to Plackett (1954) who proved that if $r_{jk}^X \leq r_{jk}^Y$ for all $1 \leq j < k \leq n$, then

$$(2.2) \quad \mathbb{P}\{X_1 \leq a_1, \dots, X_n \leq a_n\} \leq \mathbb{P}\{Y_1 \leq a_1, \dots, Y_n \leq a_n\}.$$

We will call this original version the Plackett-Slepian Comparison Lemma. Further versions have been given by Slepian (1961), Berman (1964, 1971) and Leadbetter, Lindgren and Rootzén (1983). The present statement, due to Li and Shao (2002), contains and refines the previous ones.

PROOF. We introduce some additional notation. For $t \in [0, 1]$, let

$$\Sigma_t = (1-t)\Sigma^X + t\Sigma^Y$$

It is obvious that Σ_t is positive semi-definite. Let $Z = (Z_1, \dots, Z_n)^T$ be a centered Gaussian random vector in \mathbb{R}^n with covariance Σ_t , that is, $E(Z Z^T) = \Sigma_t$ and

$$F(t) = P(Z_1 \leq a_1, \dots, Z_n \leq a_n)$$

Our aim is to give an upper bound for $F(0) - F(1)$, and for this purpose, we will consider the derivative $F'(t)$.

Let us first notice that it is sufficient to prove the result when Σ_t is non-singular for all $t \in [0, 1]$. In fact, if this has been proved, in the general case we proceed as follows.

Take n i.i.d. random variables ξ_1, \dots, ξ_n , each one of them having a standard normal distribution. $\xi = (\xi_1, \dots, \xi_n)^T$ is also assumed to be independent of X and Y .

For any $\varepsilon > 0$, the variance of the centered Gaussian vector

$$((1-t)X_1 + tY_1 + \varepsilon\xi_1, \dots, (1-t)X_n + tY_n + \varepsilon\xi_n)$$

is $\Sigma_t + \varepsilon I_n$, which is non-singular for any $t \in [0, 1]$. Hence, we may apply the inequality (2.1) to the pair of random vectors $X_1 + \varepsilon\xi_1, \dots, X_n + \varepsilon\xi_n$ and $Y_1 + \varepsilon\xi_1, \dots, Y_n + \varepsilon\xi_n$

Then, we pass to the limit as $\varepsilon \rightarrow 0$. This should be done carefully and is left to the reader.

So, assume that Σ_t is non-singular for all $t \in [0, 1]$. For $\Sigma = ((r_{jk}))_{j,k=1,\dots,n}$ positive definite and non-singular we denote by $\varphi_\Sigma(x)$, $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ the density of the centered normal distribution in \mathbb{R}^n with covariance Σ .

We have the identity

$$(2.3) \quad \frac{\partial \varphi_\Sigma}{\partial r_{jk}} = \frac{\partial^2 \varphi_\Sigma}{\partial x_j \partial x_k} \quad (j, k = 1, \dots, n, j < k).$$

To prove (2.3) we use the inversion formula for the Fourier transform and the form of the Fourier transform of the Normal distribution in \mathbb{R}^n :

$$\varphi_\Sigma(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \exp \left[-i \langle x, z \rangle - \frac{1}{2} \langle z, \Sigma z \rangle \right] dz.$$

In this equality we may differentiate under the integral sign either with respect to r_{jk} or with respect to x_j . This can be justified using dominated convergence, since the non-singularity of Σ_t implies the existence of a positive constant c such that $\langle z, \Sigma z \rangle \geq c \|z\|^2$ for all $z \in \mathbb{R}^n$. (2.3) follows.

Using this, we can compute the derivative $F'(t)$:

$$\begin{aligned} F'(t) &= \frac{d}{dt} \int_{x_h \leq a_h, h=1,\dots,n} \varphi_{\Sigma_t}(x) dx \\ &= \int_{x_h \leq a_h, h=1,\dots,n} \left[\sum_{1 \leq j < k \leq n} \frac{\partial \varphi_{\Sigma_t}}{\partial r_{jk}^t}(x) \frac{dr_{jk}^t}{dt} \right] dx \\ &= \int_{x_h \leq a_h, h=1,\dots,n} \left[\sum_{1 \leq j < k \leq n} \frac{\partial^2 \varphi_{\Sigma_t}(x)}{\partial x_j \partial x_k} (r_{jk}^Y - r_{jk}^X) \right] dx \\ &= \sum_{1 \leq j < k \leq n} (r_{jk}^Y - r_{jk}^X) \int_{x_h \leq a_h, h=1,\dots,n} \prod_{\substack{h=1 \\ h \neq j,k}}^n dx_h \int_{-\infty}^{a_j} \int_{-\infty}^{a_k} \frac{\partial^2 \varphi_{\Sigma_t}(x)}{\partial x_j \partial x_k} dx_j dx_k. \end{aligned}$$

using (2.3). r_{jk}^t stands for $(1-t)r_{jk}^X + tr_{jk}^Y$, the element (j, k) of Σ_t and dx stands for dx_1, \dots, dx_n .

In each term of this equality we integrate twice (first in x_k and second in x_j) obtaining

$$F'(t) = \sum_{1 \leq j < k \leq n} (r_{jk}^Y - r_{jk}^X) \int_{x_h \leq a_h, h=1,\dots,n} \varphi_{\Sigma_t}(\tilde{x}_{j,k}) \prod_{\substack{h=1 \\ h \neq j,k}}^n dx_h.$$

where $\tilde{x}_{j,k} = (\tilde{x}_1, \dots, \tilde{x}_n)$, $\tilde{x}_h = x_h$ for $h \neq j, k$, $\tilde{x}_h = a_h$ for $h = j, k$. Now majorizing the integral above by the same integral but with integration over all \mathbb{R} for x_h , $h \neq j, k$ we get

$$-F'(t) \leq \sum_{1 \leq j < k \leq n} (r_{jk}^X - r_{jk}^Y)^+ \varphi(a_j, a_k; r_{j,k}^t),$$

where $\varphi(u, v; \rho)$ is the joint density at the point (u, v) of two jointly Gaussian random variables with zero expectation, variance 1 and covariance ρ . Now standard algebra shows that

$$\varphi(a_j, a_k; r_{j,k}^t) \leq \frac{1}{2\pi} \frac{1}{\sqrt{1 - (r_{jk}^t)^2}} \exp\left(-\frac{a_j^2 + a_k^2}{2(1 + \rho_{jk})}\right)$$

As a consequence we get

$$F(0) - F(1) \leq \frac{1}{2\pi} \sum_{1 \leq j < k \leq n} (r_{jk}^X - r_{jk}^Y)^+ \exp\left(-\frac{a_j^2 + a_k^2}{2(1 + \rho_{jk})}\right) \int_0^1 \frac{1}{\sqrt{1 - (r_{jk}^t)^2}} dt$$

Now, on account of the form of r_{jk}^t , changing variables in the integral we have, whenever $r_{jk}^Y \leq r_{jk}^X$

$$(r_{jk}^X - r_{jk}^Y) \int_0^1 \frac{1}{\sqrt{1 - (r_{jk}^t)^2}} dt = \int_{r_{jk}^Y}^{r_{jk}^X} \frac{1}{\sqrt{1 - w^2}} dw = \arcsin(r_{jk}^X) - \arcsin(r_{jk}^Y).$$

□

COROLLARY 2.2. *Let $\{X(t) : t \in T\}$, $\{Y(t) : t \in T\}$ be separable centered Gaussian processes with almost surely bounded paths, defined on a topological space T .*

Let us assume that

$$E(X(t)^2) = E(Y(t)^2) \text{ for all } t \in T$$

$$E((Y(t) - Y(s))^2) \leq E((X(t) - X(s))^2) \text{ for all } s, t \in T$$

Then, for each $x \in \mathbb{R}$:

$$P\{\sup_{t \in T} X(t) \leq x\} \leq P\{\sup_{t \in T} Y(t) \leq x\}$$

We will say that $\sup_{t \in T} X(t)$ is stochastically greater or equal to $\sup_{t \in T} Y(t)$.

PROOF. Because of the separability, it is enough to prove the result for finite T . This follows immediately from the Plackett-Slepian version of the Normal comparison Lemma. □

EXAMPLE 2.1. *Let T be a positive number. Consider the three centered Gaussian stationary processes $X(t)$, $Y(t)$ and $Z(t)$, $t \in [0, T]$, with respective covariances*

$$\Gamma_X(t) := \exp(-t^2/2), \quad \Gamma_Y(t) := \exp(-|t|), \quad \Gamma_Z(t) := (1 - |t|)^+.$$

See figure 2.1. $X(t)$ is called the "stationary process with Gaussian covariance", $Y(t)$ is the Ornstein-Uhlenbeck process, and $Z(t)$ is the "Slepian process". Then it is easy to check that

$$\Gamma_Z(t) \leq \Gamma_Y(t) \leq \Gamma_X(t) \text{ for } |t| \leq 2.$$

So let $T \leq 2$, then

$$\sup_{t \in T} X(t) \stackrel{s}{\leq} \sup_{t \in T} Y(t) \stackrel{s}{\leq} \sup_{t \in T} Z(t)$$

where $\stackrel{s}{\leq}$ is the stochastic order.

We finish this section stating without proof two related results having important applications.

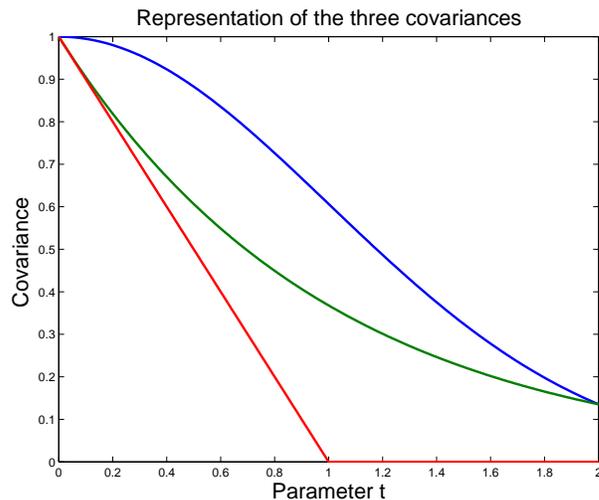


FIGURE 2.1. Representation of the three covariances

THEOREM 2.3 (Li and Shao, 2002). *Use the notations of Lemma 2.1, assume $n \geq 3$, and that*

$$r_{ij}^X \geq r_{ij}^Y \geq 0 \text{ for all } 1 \leq i, j \leq n$$

Then for $a \leq 0$ one has:

$$\begin{aligned} & \mathbb{P}\{Y_1 \leq a, \dots, Y_n \leq a\} \leq \mathbb{P}\{X_1 \leq a, \dots, X_n \leq a\} \\ & \leq \mathbb{P}\{Y_1 \leq a, \dots, Y_n \leq a\} \exp \left[\sum_{1 \leq i < j \leq n} \log \left(\frac{\pi - 2 \arcsin(r_{ij}^Y)}{\pi - 2 \arcsin(r_{ij}^X)} \right) \exp \left(-\frac{a^2}{(1 + r_{ij}^X)} \right) \right] \end{aligned}$$

THEOREM 2.4 (Sudakov, Fernique). *Let $\{X(t) : t \in T\}$, $\{Y(t) : t \in T\}$ be separable centered Gaussian processes with almost surely bounded paths, defined on a topological space T . Let us assume that*

$$\mathbb{E}((X(t) - X(s))^2) \leq \mathbb{E}((Y(t) - Y(s))^2) \text{ for all } t \text{ and } s$$

Then

$$\mathbb{E}(\sup_{t \in T} X(t)) \leq \mathbb{E}(\sup_{t \in T} Y(t))$$

A proof of this theorem can be found in Adler(1990). We will see later on that under the conditions of this theorem both expectations are finite (which is not evident at all!).

Notice that under the - more restrictive - conditions of the Corollary 2.2, the conclusion of the last Theorem follows immediately from it, since for any integrable real-valued random variable η one can express the expectation as:

$$\mathbb{E}(\eta) = \int_0^{+\infty} \mathbb{P}(\eta > x) dx - \int_{-\infty}^0 \mathbb{P}(\eta < -x) dx$$

Apply this to $\sup_{t \in T} X(t)$ and to $\sup_{t \in T} Y(t)$.

2. Ehrhard's inequality.

THEOREM 2.5. *Let γ_n be the standard Gaussian probability measure on \mathbb{R}^n .*

Then, for any pair A, B of Borel subsets of \mathbb{R}^n and any λ , $0 < \lambda < 1$,

$$(2.4) \quad \Phi^{-1}(\gamma_n(\lambda A + (1 - \lambda)B)) \geq \lambda \Phi^{-1}(\gamma_n(A)) + (1 - \lambda) \Phi^{-1}(\gamma_n(B))$$

holds true.

This theorem was proved by Ehrhard (1983) for convex A and B and by Latala (1996) when at least one of the two sets is convex. The proof in its general form is due to C. Borell (2003) and is the following.

PROOF. It suffices to prove (2.4) for compact A and B .

Let $0 < \varepsilon < 1$ and $0 < \delta < \varepsilon$. We put $\alpha := 1 - \varepsilon + \delta$ so that $\delta < \alpha < 1$, and $A_\varepsilon := A + \overline{B}_n(0; \varepsilon)$ where $\overline{B}_n(0; \varepsilon)$ is the closed ball in \mathbb{R}^n centered at the origin and having radius ε .

Let $f_1 : \mathbb{R}^n \rightarrow \mathbb{R}$ be a C^∞ -function such that $f_1|_A \equiv \alpha$, $f_1|_{A^c} \equiv \delta$. We have $\delta \leq f_1(x) \leq \alpha$ for all $x \in \mathbb{R}^n$.

In a similar way, define f_2 by changing, in the definition of f_1 , the set A by the set B , and f_3 , by changing in the definition of f_1 the set A by the set $\lambda A_\varepsilon + (1 - \lambda)B_\varepsilon$ and the minimum value δ by

$$\kappa = \max \left(\Phi \left[\lambda \Phi^{-1}(\alpha) + (1 - \lambda) \Phi^{-1}(\delta) \right], \Phi \left[\lambda \Phi^{-1}(\delta) + (1 - \lambda) \Phi^{-1}(\alpha) \right] \right).$$

Notice that $\kappa \rightarrow 0$ as $\delta \rightarrow 0$.

We will prove the inequality

$$(2.5) \quad \Phi^{-1} \left(\int_{\mathbb{R}^n} f_3 d\gamma_n \right) \geq \lambda \Phi^{-1} \left(\int_{\mathbb{R}^n} f_1 d\gamma_n \right) + (1 - \lambda) \Phi^{-1} \left(\int_{\mathbb{R}^n} f_2 d\gamma_n \right).$$

(2.4) follows from (2.5) by letting $\delta \rightarrow 0$ and $\varepsilon \rightarrow 0$, in this order.

Let us define, for $j = 1, 2, 3$ and $(t, x) \in [0, 1] \times \mathcal{R}^n$:

$$u_j(t, x) = \int_{\mathbb{R}^n} f_j(x + \sqrt{t}z) \gamma_n(dz)$$

Instead of (2.5) we will prove in fact the more general inequality

$$(2.6) \quad \Phi^{-1}(u_3(t, \lambda x + (1 - \lambda)y)) \geq \lambda \Phi^{-1}(u_1(t, x)) + (1 - \lambda) \Phi^{-1}(u_2(t, y)).$$

Putting $t = 1$, $x = y = 0$ in (2.6) we obtain (2.5).

So, our aim is to prove that

$$(2.7) \quad C(t, x, y) \geq 0 \quad \text{for all } t \in [0, 1], x, y \in \mathbb{R}^n,$$

where

$$(2.8) \quad C(t, x, y) = U_3(t, \lambda x + (1 - \lambda)y) - \lambda U_1(t, x) + (1 - \lambda)U_2(t, y)$$

with the notation $U_j = \Phi^{-1} \circ u_j$ ($j = 1, 2, 3$).

An instant reflection shows that the definitions of the functions f_j, u_j imply that the functions U_j , as functions of the space variable x , are C^∞ and the partial derivatives of all orders are bounded for $t \geq 0, x \in \mathbb{R}^n$.

Let us check that (2.7) holds true when $t = 0$ in which case it becomes

$$(2.9) \quad f_3(\lambda x + (1 - \lambda)y) \geq \Phi \left[\lambda \Phi^{-1}(f_1(x)) + (1 - \lambda) \Phi^{-1}(f_2(y)) \right].$$

If $x \notin A_\varepsilon$, the right-hand side of (2.9) is bounded above by $\Phi \left[\lambda \Phi^{-1}(\delta) + (1 - \lambda) \Phi^{-1}(\alpha) \right] \leq \kappa$ and κ is a lower bound of f_3 . Similarly if $y \notin B_\varepsilon$. So, it remains to prove (2.9) when $x \in A_\varepsilon$ and also $y \in B_\varepsilon$ in which case the left-hand side of (2.9) is equal to α and this is an upper-bound for the right-hand side of (2.9).

To show that (2.7) also holds true for all $t \in [0, 1]$, Borell's proof uses a method that reminds the maximum principle for parabolic equations. We denote ∇ and Δ respectively the gradient and the Laplace operators with respect to the space variables.

Firstly, u_j ($j = 1, 2, 3$) verifies the heat equation

$$\frac{\partial u_j}{\partial t} = \frac{1}{2} \Delta u_j \quad \text{on } [0, 1] \times \mathbb{R}^n,$$

and this implies, by a simple computation, that

$$(2.10) \quad \frac{\partial U_j}{\partial t} = \frac{1}{2} \Delta U_j - \frac{1}{2} U_j \|\nabla U_j\|^2 \quad \text{on } [0, 1] \times \mathbb{R}^n.$$

Using the identities (2.10) for $j = 1, 2, 3$, we can compute the value of the differential operator

$$\mathcal{L} = \frac{1}{2} \sum_{j=1}^n \left[\frac{\partial^2}{\partial x_j^2} + 2 \frac{\partial^2}{\partial x_j \partial y_j} + \frac{\partial^2}{\partial y_j^2} \right]$$

on the function $C(t, \cdot, \cdot)$. We obtain:

$$(2.11) \quad \mathcal{L}C(t, x, y) = \frac{\partial C}{\partial t}(t, x, y) + \frac{1}{2} \|(\nabla U_3)(t, \lambda x + (1 - \lambda)y)\|^2 C(t, x, y) + b(t, x, y)$$

where

$$(2.12) \quad b(t, x, y) = U_1 \langle \nabla_x C, \nabla U_1 + \nabla U_3 \rangle + U_2 \langle \nabla_y C, \nabla U_2 + \nabla U_3 \rangle.$$

In (2.12) U_1 and ∇U_1 are computed at (t, x) , U_2 and ∇U_2 are computed at (t, y) , U_3 and ∇U_3 at $(t, \lambda x + (1 - \lambda)y)$ and ∇_x, ∇_y denote the gradient with respect to x and y . The computation of derivatives to check (2.11) is left to the reader.

Let us suppose that (2.7) does not hold and show that this leads to a contradiction. We prove first that

$$(2.13) \quad \liminf_{\|x\| + \|y\| \rightarrow +\infty} \inf_{0 \leq t \leq 1} C(t, x, y) \geq 0.$$

Since A and B are bounded, one can find $a > 0$ such that if $\|w\| \geq a$, then $f_1(w) = f_2(w) = \delta$, $f_3(w) = \kappa$. Note that

$$u_1(t, x) = \mathbb{E} \left(f_1 \left(x + \sqrt{t} \xi \right) \right)$$

where ξ is standard normal in \mathbb{R}^n . Hence,

$$\begin{aligned} u_1(t, x) &= \mathbb{E} \left(f_1 \left(x + \sqrt{t} \xi \right) \mathbb{1}_{\{\|x + \sqrt{t} \xi\| \geq a\}} \right) + \mathbb{E} \left(f_1 \left(x + \sqrt{t} \xi \right) \mathbb{1}_{\{\|x + \sqrt{t} \xi\| < a\}} \right) \\ &= \delta + R, \end{aligned}$$

with $|R| \leq \mathbb{P}(\|x + \sqrt{t} \xi\| < a)$. Choose now x so that $\|x\| \geq 2a$, and we get for $0 \leq t \leq 1$:

$$\begin{aligned} |R| &\leq \mathbb{P} \left(\sqrt{t} \|\xi\| > \frac{\|x\|}{2} \right) \leq \mathbb{P} \left(\|\xi\| > \frac{\|x\|}{2} \right) \\ &= \sigma_{n-1} (S^{n-1}) \int_{\frac{\|x\|}{2}}^{+\infty} \rho^{n-1} \exp \left(-\frac{\rho^2}{2} \right) d\rho \leq C_1 \exp \left(-C_2 \|x\|^2 \right) \end{aligned}$$

where C_1, C_2 are positive constants. This shows that as $\|x\| \rightarrow +\infty$, $u_1(t, x)$ converges to δ , uniformly on $t \in [0, 1]$. A similar result holds for u_2 and u_3 , in the latter case replacing δ by κ . Going back to the definitions of $C(t, x, y)$ and κ , (2.13) follows.

On the other hand, (2.13) implies that if $C(t, x, y)$ takes a negative value somewhere in $[0, 1] \times \mathbb{R}^n \times \mathcal{R}^n$, then it has a minimum and since we also know that $C(0, x, y) \geq 0$ for any choice of x, y , choosing $\varepsilon > 0$ small enough we can assure that the function $C(t, x, y) + \varepsilon t$ will also have a negative minimum at some point $(\bar{t}, \bar{x}, \bar{y})$ with $0 < \bar{t} \leq 1$. Clearly, this implies that

$$\nabla_x C(\bar{t}, \bar{x}, \bar{y}) = 0, \quad \nabla_y C(\bar{t}, \bar{x}, \bar{y}) = 0, \quad \frac{\partial C}{\partial t}(\bar{t}, \bar{x}, \bar{y}) \leq -\varepsilon.$$

Denote by M the $(2n) \times (2n)$ matrix of second partial derivatives of the function $C(\bar{t}, \cdot, \cdot)$ computed at the point (\bar{x}, \bar{y}) , that is, if we rename the vector $(x_1, \dots, x_n, y_1, \dots, y_n)$ as (z_1, \dots, z_{2n}) , we have:

$$M = \left(\left(\frac{\partial^2 C}{\partial z_h \partial z_k} \Big|_{(t,x,y)=(\bar{t},\bar{x},\bar{y})} \right) \right)_{h,k=1,\dots,2n}.$$

Since there is a minimum of $C(\bar{t}, \cdot, \cdot)$ at the point (\bar{x}, \bar{y}) , M has to be semi-definite positive. One also has

$$\mathcal{L}C(\bar{t}, \bar{x}, \bar{y}) = \frac{1}{2} \sum_{j=1}^n \langle M \theta_j, \theta_j \rangle$$

where θ_j denotes the vector $(x_1, \dots, x_n, y_1, \dots, y_n)$ such that $x_k = y_k = \delta_{jk}$ ($k = 1, \dots, n$). So, $\mathcal{L}C(\bar{t}, \bar{x}, \bar{y}) \geq 0$.

However, putting $(t, x, y) = (\bar{t}, \bar{x}, \bar{y})$ in the right-hand side of (2.11), we see that it becomes strictly negative. This ends the proof. \square

3. Gaussian isoperimetric inequality.

One possible version of the classical isoperimetric inequality for Lebesgue measure in \mathbb{R}^n states that if A is a Borel subset of \mathbb{R}^n and the ball $B(0; r)$ has the same Lebesgue measure as A , then, for any $t > 0$, $\lambda_n(A_t) \geq \lambda_n(B(0; r + t))$ holds true. The notation for A_t is the one introduced at the beginning of the proof of Theorem 2.5.

In the mid-seventies, C. Borell (1975) and V.N. Sudakov and B.S. Tsirelson (1974) proved independently a similar property for Gaussian measures, which is the statement of the next theorem. The use of isoperimetric methods for Gaussian distributions seems to have started with the paper by Landau and Shepp (1970) in which they studied the tails of the distribution of the supremum of Gaussian processes. The results were improved by Marcus & Shepp (1972) in a paper in which they gave what seems to be the first published proof of (2.33). An independent and purely probabilistic proof of (2.33) is in Fernique's Lecture Notes, along with other connected results.

All this was well established around 1975, and will be sufficient for our uses in the next chapters. For the many interesting directions of the relationship between isoperimetry and Gaussian and related measures, the important references are the monographs by M. Ledoux (1996, 2001). A synthesis of known results and open problems on Gaussian and related inequalities and the relations with isoperimetry is in Latala's conference at ICM 2002.

THEOREM 2.6. *Let A be a Borel subset of \mathbb{R}^n and H a half-space in \mathbb{R}^n , such that $\gamma_n(A) = \gamma_n(H) = \Phi(a)$ some $a \in \mathbb{R}$.*

Then,

$$(2.14) \quad \gamma_n(A_t) \geq \gamma_n(H_t) = \Phi(a + t) \quad \text{for every } t > 0.$$

PROOF. Let $0 < \lambda < 1$. Applying Theorem 2.5:

$$(2.15) \quad \begin{aligned} \Phi^{-1}(\gamma_n(A_t)) &= \Phi^{-1}\left(\gamma_n\left(A + \overline{B_n(0; t)}\right)\right) \\ &= \Phi^{-1}\left[\gamma_n\left(\lambda \frac{1}{\lambda} A + (1 - \lambda) \frac{1}{1 - \lambda} \overline{B_n(0; t)}\right)\right] \\ &\geq \lambda \Phi^{-1}\left(\gamma_n\left(\frac{1}{\lambda} A\right)\right) + (1 - \lambda) \Phi^{-1}\left(\gamma_n\left(\frac{1}{1 - \lambda} \overline{B_n(0; t)}\right)\right). \end{aligned}$$

Let $\lambda \uparrow 1$ in (2.15). The first term in the right-hand-side tends to $\Phi^{-1}(\gamma_n(A)) = a$. To compute the limit of the second term, put $r = \frac{1}{1 - \lambda} \rightarrow +\infty$ as $\lambda \uparrow 1$ and $\gamma_n\left(\frac{1}{1 - \lambda} \overline{B_n(0; t)}\right) = \Phi(y)$, so that $y \rightarrow +\infty$ as $r \rightarrow +\infty$. Using the standard formulas for the Gaussian distribution,

$$\begin{aligned} 1 - \Phi(y) &\approx \frac{1}{\sqrt{2\pi}} \frac{1}{y} e^{-\frac{1}{2}y^2} \\ 1 - \gamma_n\left(\frac{1}{1 - \lambda} \overline{B_n(0; t)}\right) &\approx \frac{1}{2^{\frac{n}{2}-1} \Gamma\left(\frac{n}{2}\right)} (rt)^{n-2} e^{-\frac{1}{2}r^2 t^2} \end{aligned}$$

since both left-hand members are equal, we conclude that $\frac{y}{r} \rightarrow t$. Summing up,

$$\Phi^{-1}(\gamma_n(A_t)) \geq a + t.$$

\square

4. Inequalities for the tails of the distribution of the supremum.

Let $\mathcal{X} = \{X(t) : t \in T\}$ be a real-valued centered Gaussian process, $M_T(\omega) = \sup_{t \in T} X(t)(\omega)$ ($M_T(\omega)$ may have the value $+\infty$). We will assume in this section that there exists a countable subset \mathcal{D} of the parameter set T such that almost surely $M_T = \sup_{t \in \mathcal{D}} X(t)$. In particular, this condition holds true if \mathcal{X} is a separable process. We denote $\sigma^2(t) = \mathbb{E}(X^2(t))$.

This section contains two theorems with general bounds for the probability distribution of M_T .

The first one is the Borell, Sudakov, Tsirelson inequality that gives an exponential bound for $P(|M_T - \mu(M_T)| > x)$ where $\mu(\xi)$ denotes a median of the distribution of the real-valued random variable ξ . The proof is a consequence of the isoperimetric inequality (2.14).

The second one is similar, but instead of the median appears in the statement the expectation $E(M_T)$. We have included a proof due to Ibragimov, Sudakov and Tsirelson (1976), which is independent of the foregoing arguments. This proof is interesting by itself, since it is based upon Ito's formula so that it establishes a link between the theory of Gaussian processes and Stochastic Analysis.

In what follows, f is the function $f(x) = \sup_{1 \leq j \leq N} x^j$, where $x = (x^1, \dots, x^N)^T$ and

$$(2.16) \quad f_\varepsilon(x) = \int_{\mathbb{R}^N} \prod_{j=1}^N g_\varepsilon(x^j - y^j) f(y) dy$$

is a regularization of f by convolution. Here, $g : \mathbb{R} \rightarrow \mathbb{R}^+$ is a function of class C^∞ with support in the interval $[-1, 1]$, $\int_{-1}^1 g(r) dr = 1$ and $g_\varepsilon(r) = \frac{1}{\varepsilon} g\left(\frac{r}{\varepsilon}\right)$, $0 < \varepsilon < 1$.

We start with the following lemma:

LEMMA 2.7.

- (1) $\sum_{j=1}^N \left| \frac{\partial f_\varepsilon}{\partial x^j}(x) \right| = 1 \quad \forall x \in \mathbb{R}^N$.
- (2) Let $A = (a_{ij})_{i,j=1,\dots,N}$ be a real $N \times N$ matrix and set $B = AA^T = (b_{ij})_{i,j=1,\dots,N}$. The function $h_A(x) = f(Ax)$, $x \in \mathbb{R}^N$ satisfies the Lipschitz condition

$$|h_A(x) - h_A(y)| \leq \bar{b} \|x - y\| \quad \forall x, y \in \mathbb{R}^N,$$

where $\bar{b}^2 = \sup\{b_{ii} : i = 1, \dots, N\}$.

PROOF. To prove (1), let us compute the partial derivatives of f_ε :

$$\begin{aligned} \frac{\partial f_\varepsilon}{\partial x^j}(x) &= \int_{\mathbb{R}^{N-1}} \prod_{k=1, k \neq j}^N [g_\varepsilon(x^k - y^k) dy^k] \int_{\mathbb{R}} g'_\varepsilon(x^j - y^j) f(y) dy^j \\ &= \int_{\mathbb{R}^{N-1}} \prod_{k=1, k \neq j}^N [g_\varepsilon(x^k - y^k) dy^k] \int_{\mathbb{R}} g_\varepsilon(x^j - y^j) \mathbb{1}_{y^j > \sup_{k \neq j} y^k} dy^j \\ &= \int_{A_j} \prod_{j=1}^N g_\varepsilon(x^j - y^j) dy \end{aligned}$$

with $A_j = \{y^j > \sup_{k \neq j} y^k\}$. The second equality above comes from integration by parts in the inner integral.

Now it becomes plain that

$$(2.17) \quad \sum_{j=1}^N \left| \frac{\partial f_\varepsilon}{\partial x^j}(x) \right| = 1$$

since the sets A_j ($j = 1, \dots, N$) are a partition of \mathbb{R}^N , modulo Lebesgue measure.

To prove (2), notice that it suffices to show the result for the function h_ε instead of h , where $h_\varepsilon(x) = f_\varepsilon(Ax)$ and then pass to the limit as $\varepsilon \downarrow 0$. This will be achieved if we prove:

$$\|\nabla h_\varepsilon(x)\| \leq \bar{b} \quad \forall x \in \mathbb{R}^N.$$

In fact:

$$(2.18) \quad \|\nabla h_\varepsilon(x)\|^2 = \sum_{i=1}^N \sum_{k,k'=1}^N \frac{\partial f_\varepsilon}{\partial x^k} a_{ik} \frac{\partial f_\varepsilon}{\partial x^{k'}} a_{ik'} = \sum_{k,k'=1}^N \frac{\partial f_\varepsilon}{\partial x^k} b_{kk'} \frac{\partial f_\varepsilon}{\partial x^{k'}} \leq \bar{b}^2 \left[\sum_{j=1}^N \left| \frac{\partial f_\varepsilon}{\partial x^j} \right| \right]^2 = \bar{b}^2.$$

□

Let us now state and prove the first theorem.

THEOREM 2.8. *Assume that $P(M_T < \infty) = 1$.*

Then:

$$\sigma_T^2 = \sup_{t \in T} \sigma^2(t) < +\infty$$

and for every $u > 0$

$$(2.19) \quad P(|M_T - \mu(M_T)| > u) \leq e^{-\frac{1}{2} \frac{u^2}{\sigma_T^2}}.$$

Remark. We will prove the stronger inequality:

$$P(|M_T - \mu(M_T)| > u) \leq 2[1 - \Phi(u/\sigma_T)].$$

PROOF. Let us prove that $\sigma_T^2 < \infty$. In fact, if $\sigma_T^2 = \infty$ and $\{t_n\}$ is a sequence in T such that $\sigma^2(t_n) \rightarrow \infty$, it follows that for $u > 0$:

$$P(M_T > u) \geq P(X(t_n) > u) = \frac{1}{\sqrt{2\pi}\sigma(t_n)} \int_u^{+\infty} e^{-\frac{1}{2} \frac{y^2}{\sigma^2(t_n)}} dy = \frac{1}{\sqrt{2\pi}} \int_{\frac{u}{\sigma(t_n)}}^{+\infty} e^{-\frac{1}{2} y^2} dy \rightarrow \frac{1}{2}.$$

So, $P(M_T > u) \geq \frac{1}{2}$ for every $u > 0$. This implies $P(M_T = +\infty) \geq \frac{1}{2}$ which contradicts the hypothesis that $P(M_T = +\infty) = 0$.

One can also assume that $\sigma_T^2 > 0$ since this only excludes the trivial case that a.s. $X_t = 0$ for every $t \in T$.

On the other hand, due to the hypothesis that a.s. M_T is the supremum over a countable set of parameter values, a simple approximation argument shows that it is enough to prove (2.19) when the parameter T is finite, say it consists of N points. So, our aim is to prove the inequality:

$$(2.20) \quad P(|f(X) - \mu(f(X))| > u) \leq 2[1 - \Phi(u/\sigma)].$$

where $X = (X_1, \dots, X_N)^T$ is a centered Gaussian vector in \mathbb{R}^N , $\text{Var}(X) = E(XX^T) = V = ((V_{ij}))_{i,j=1,\dots,N}$ and $\sigma^2 = \sup_{1 \leq j \leq N} V_{jj}$.

Let $V^{\frac{1}{2}}$ be a square root of the variance matrix V , that is, $V^{\frac{1}{2}}(V^{\frac{1}{2}})^T = V$. Then, the random vectors X and $V^{\frac{1}{2}}\eta$, where η has a standard normal distribution in \mathbb{R}^N , have the same distribution, and our problem is to prove that

$$(2.21) \quad P(|h(\eta) - \mu(h(\eta))| > u) \leq 2[1 - \Phi(u/\sigma)]$$

where

$$h(x) = f(V^{\frac{1}{2}}x).$$

We denote $\bar{\mu}$ the median of the random variable $h(\eta)$. Notice that for any $x, y \in \mathbb{R}^N$, using Lemma 2.7, part (2) with h instead of h_A :

$$L_h = \sup \left\{ \frac{|h(x) - h(y)|}{\|x - y\|} : x, y \in \mathbb{R}^N, x \neq y \right\} \leq \sigma.$$

Since $u > 0$, it follows that

$$(2.22) \quad P(h(\eta) - \bar{\mu} > u\sigma) \leq P(h(\eta) - \bar{\mu} > uL_h).$$

Define $A := \{x \in \mathbb{R}^N, h(x) \leq \bar{\mu}\}$. Then:

$$A_u \subset \{y \in \mathbb{R}^N : h(y) \leq \bar{\mu} + uL_h\}$$

In fact, according to the definition of A_u , if $w \in A_u$, one can write $w = x + z$, where $h(x) \leq \bar{\mu}$, $\|z\| \leq u$. So,

$$h(w) \leq h(x) + uL_h.$$

It follows, using (2.22) that

$$(2.23) \quad \mathbb{P}(h(\eta) - \bar{\mu} > u\sigma) \leq \mathbb{P}(\eta \in A_u^C).$$

We use now the isoperimetric inequality (2.14). Since $\bar{\mu}$ is the median of the distribution of $h(\eta)$, one has $\mathbb{P}(\eta \in A) \geq 1/2$, so that

$$\mathbb{P}(\eta \in A_u) \geq \Phi(u).$$

So, (2.23) implies that

$$(2.24) \quad \mathbb{P}(h(\eta) - \bar{\mu} > u\sigma) \leq 1 - \Phi(u) \leq \frac{1}{2}e^{-\frac{1}{2}u^2},$$

where checking last inequality is an elementary computation.

A similar argument applies to $\mathbb{P}(h(\eta) - \bar{\mu} < -u\sigma)$. This proves (2.21) and finishes the proof of the theorem. \square

Let us now turn to the second general inequality.

THEOREM 2.9. *Assume that the process \mathcal{X} satisfies the same hypotheses as in Theorem 2.8.*

Then:

1) $\mathbb{E}(|M_T|) < \infty$.

2) For every $u > 0$ the inequality

$$(2.25) \quad \mathbb{P}(|M_T - \mathbb{E}(M_T)| > u) \leq 2 \exp\left(-\frac{1}{2} \frac{u^2}{\sigma_T^2}\right).$$

PROOF. We prove first the inequality for finite parameter set T , say having N points. The first part of the proof is exactly the same as in the previous proof, and we use the same notations as above.

Let $\{W(t) : t \geq 0\}$ be a Wiener process in \mathbb{R}^N , that is

$$W(t) = (W^1(t), \dots, W^N(t))^T, \quad t \geq 0$$

where W^1, \dots, W^N are real-valued independent Wiener processes. We want to prove that

$$(2.26) \quad \mathbb{P}(|h(W(1)) - \mathbb{E}(h(W(1)))| > u) \leq 2 \exp\left(-\frac{1}{2} \frac{u^2}{\sigma^2}\right)$$

for any $u > 0$. It suffices to prove (2.26) for the smooth function h_ε . Then, passing to the limit as $\varepsilon \downarrow 0$ gives the result. In what follows, for shortness, we have put h instead of h_ε .

Consider the function $H : \mathbb{R}^N \times (0, 1) \rightarrow \mathbb{R}$ defined by means of:

$$H(x, t) = \mathbb{E}(h(x + W(1-t))) = \int_{\mathbb{R}^d} h(x+y) p_{1-t}(y) dy = (h * p_{1-t})(x)$$

where

$$p_t(y) = \frac{1}{(2\pi t)^{\frac{N}{2}}} \exp\left(-\frac{\|y\|^2}{2t}\right), \quad t > 0$$

is the density of the random variable $W(t)$ in \mathbb{R}^N .

One can easily check that:

$$(2.27) \quad \begin{aligned} \frac{\partial p_t}{\partial t} &= \frac{1}{2} \sum_{j=1}^N \frac{\partial^2 p_t}{(\partial x^j)^2} \\ \frac{\partial H}{\partial t} &= -\frac{1}{2} \sum_{j=1}^N \frac{\partial^2 H}{(\partial x^j)^2} \end{aligned}$$

and that the function H has the boundary values:

$$H(x, t) \rightarrow h(y) \text{ as } t \uparrow 1 \text{ and } x \rightarrow y$$

$$H(x, t) \rightarrow \mathbb{E}(h(y + W(1))) \text{ as } t \downarrow 0 \text{ and } x \rightarrow y.$$

Let us apply Itô's formula, $0 < s < t < 1$:

$$\begin{aligned} H(W(t), t) - H(W(s), s) &= \int_s^t \sum_{j=1}^N \left[\frac{\partial H}{\partial x^j}(W(u), u) dW^j(u) + \frac{1}{2} \frac{\partial^2 H}{(\partial x^j)^2}(W(u), u) du \right] \\ &\quad + \int_s^t \frac{\partial H}{\partial t}(W(u), u) du \\ &= \int_s^t \sum_{j=1}^N \frac{\partial H}{\partial x^j}(W(u), u) dW^j(u), \end{aligned}$$

using (2.27). Take now limits as $t \uparrow 1$ and $s \downarrow 0$, obtaining

$$(2.28) \quad h(W(1)) - \mathbb{E}(h(W(1))) = Z(1),$$

where $\{Z(t) : t \geq 0\}$ is the martingale

$$Z(t) = \int_0^t \sum_{j=1}^N \frac{\partial H}{\partial x^j}(W(u), u) dW^j(u).$$

Let us prove that the quadratic variation of Z verifies:

$$(2.29) \quad [Z]_1 \leq \sigma^2$$

From the proof of Lemma 2.7, we know that

$$\sup_{y \in \mathbb{R}^N} \|\nabla h(y)\|^2 \leq \sigma^2.$$

So,

$$\begin{aligned} \sum_{j=1}^N \left(\frac{\partial H}{\partial x^j}(x, u) \right)^2 &= \sum_{j=1}^N \left[\mathbb{E} \left(\frac{\partial h}{\partial x^j}(x + W(1 - u)) \right) \right]^2 \\ &\leq \sum_{j=1}^N \mathbb{E} \left[\left(\frac{\partial h}{\partial x^j}(x + W(1 - u)) \right)^2 \right] \leq \sigma^2 \end{aligned}$$

and we obtain the bound (2.29) since:

$$[Z]_1 = \int_0^1 \sum_{j=1}^N \left(\frac{\partial H}{\partial x^j}(W(u), u) \right)^2 du \leq \sigma^2.$$

Now, for each $\theta \in \mathbb{R}$ we consider the exponential martingale (see Mc Kean, 1969):

$$Y(t) = e^{\theta Z(t) - \frac{1}{2} \theta^2 [Z]_t}, \quad 0 \leq t \leq 1,$$

which satisfies

$$\mathbb{E}(Y(t)) = 1 \text{ for every } t, \quad 0 \leq t \leq 1.$$

This, together with (2.29) imply that for every $\theta \in \mathbb{R}$, $\mathbb{E}(e^{\theta Z(1) - \frac{1}{2} \theta^2 \sigma^2}) \leq 1$, so that:

$$(2.30) \quad \mathbb{E} \left(e^{\theta Z(1)} \right) \leq e^{\frac{1}{2} \theta^2 \sigma^2}.$$

Write the left-hand side of (2.26) for $u > 0$ as

$$\mathbb{P}(|Z(1)| > u) = \mathbb{P}(Z(1) > u) + \mathbb{P}(Z(1) < -u).$$

For the first term use (2.30) with $\theta = \frac{u}{\sigma^2}$. Then:

$$\mathbb{P}(Z(1) > u) = \mathbb{P}(e^{\theta Z(1)} > e^\theta) \leq e^{-\theta u} \mathbb{E}(e^{\theta Z(1)}) \leq \exp \left(-\theta u + \frac{1}{2} \theta^2 \sigma^2 \right) = \exp \left(-\frac{1}{2} u^2 \sigma^{-2} \right).$$

A similar argument produces the same bound for the second term. This proves (2.26).

To finish, we must show that the result holds for infinite T . Due to the hypothesis, it suffices to consider the case when T is countable. Put $T = \{t_n\}_{n=1,2,\dots}$, $T_N = \{t_1, \dots, t_N\}$, $N \geq 1$. Clearly

$$M_{T_N} \uparrow M_T \quad , \quad \sigma_{T_N}^2 \uparrow \sigma_T^2 \quad \text{as } N \uparrow +\infty$$

and

$$0 \leq M_{T_N} - X(t_1) \uparrow M_T - X(t_1) \quad \text{as } N \uparrow +\infty.$$

Beppo Levi's Theorem implies that

$$E(M_{T_N}) \uparrow E(M_T) \quad \text{as } N \uparrow +\infty.$$

Since we already know that (2.25) holds true for T_N instead of T , it will be sufficient to prove that $E(M_T) < \infty$ to obtain (2.25), by letting $N \uparrow +\infty$.

Let us suppose that this were not true, that is, that $E(M_T) = +\infty$. Using the fact that a.s. the paths are bounded, choose x_0 large enough to have

$$P(M_T < x_0) > \frac{3}{4} \quad \text{and} \quad \exp\left(-\frac{x_0^2}{2\sigma_T^2}\right) < \frac{1}{4}.$$

Now, if $E(M_T) = +\infty$ using Beppo Levi's Theorem, we can choose N large enough so that

$$E(M_{T_N}) > 2x_0.$$

Then, if $\omega \in \{M_T < x_0\}$ one has $M_{T_N}(\omega) \leq M_T(\omega) < E(M_{T_N}) - x_0$ which implies that $|M_{T_N}(\omega) - E(M_{T_N})| > x_0$. Hence,

$$\frac{3}{4} < P(M_T < x_0) \leq P(|M_{T_N} - E(M_{T_N})| > x_0) \leq 2 \exp\left(-\frac{x_0^2}{2\sigma_{T_N}^2}\right) \leq 2 \exp\left(-\frac{x_0^2}{2\sigma_T^2}\right) < \frac{1}{2}.$$

which is a contradiction. This implies that $E(M_T) < \infty$ and we are done. \square

4.1. Some derived tail inequalities.

1.- The same arguments show that Theorems 2.8 and 2.9 have unilateral versions, namely for x greater than the mean $E(M_T)$ or the median $\mu(M_T)$ of the process.

$$(2.31) \quad P(M > u) \leq \exp\left(-\frac{(u - E(M_T))^2}{2}\right)$$

and

$$(2.32) \quad P(M > u) \leq \frac{1}{2} \exp\left(-\frac{(u - \mu(M_T))^2}{2}\right).$$

2.- A weaker form of the above inequalities is the following: under the same hypotheses of Theorem 2.8 or 2.9, for each $\varepsilon > 0$ there exists a positive constant C_ε such that for all $u > 0$:

$$(2.33) \quad P(|M_T| > u) \leq C_\varepsilon \exp\left(-\frac{1}{2} \frac{u^2}{\sigma_T^2 + \varepsilon}\right).$$

(2.33) is a consequence of (2.25) on account of $\frac{u - E\{M_T\}}{u} \rightarrow 1$ as $u \rightarrow +\infty$.

Grosso modo, this says that the tail of the distribution of the random variable M_T is bounded (except for a multiplicative constant) by the value of the centered normal density having variance larger than, and arbitrarily close to, σ_T^2 .

The problem with this kind of inequality is that, in general, the constant C_ε can grow (and tend to infinity) as ε decreases to zero. Even for fixed ε , the usual situation is that, in general, one can only have rough bounds for C_ε and this implies serious limitations for the use of these inequalities in Statistics and in other fields. We will return to this problem in some of the next chapters, with the aim of giving more accurate results on the value of the tails of the distribution of the supremum, at least for certain classes of Gaussian processes satisfying regularity assumptions.

3.- (2.31) and (2.25) show that one can do better than (2.33) since for example (2.25) has the form

$$P(|M_T| > u) \leq C \cdot \exp\left(-\frac{u^2}{2\sigma_T^2} + C_1 u\right).$$

The difficulty for using this inequality is that the positive constants C and C_1 depend on $E(M_T)$, which is finite but unknown. The problem of giving bounds on $E(M_T)$ will be addressed in the next section.

4.- Under the same hypotheses, if $\sigma^2 > \sigma_T^2$ one has:

$$(2.34) \quad E\left(\exp\left(\frac{M_T^2}{2\sigma^2}\right)\right) < \infty.$$

This is a direct consequence of (2.33)

and implies that all moments of M_T are finite, since for any positive integer k :

$$E(M_T^{2k}) \leq (2\sigma^2)^k k! E\left(\exp\frac{M_T^2}{2\sigma^2}\right).$$

A straightforward consequence (that we have already proved by direct means in the previous chapter) is that if $X = \{X(t) : t \in T\}$ is a Gaussian process defined on a compact separable topological space such that almost surely the paths $t \rightsquigarrow X(t)$ are continuous then, the mean $m(t)$ and the covariance function $r(s, t) := \text{Cov}(X(s), X(t))$ are continuous functions. In fact, it suffices to notice that each continuous path is bounded, so that the theorem can be applied, and one can use Lebesgue dominated convergence to prove continuity.

5.- To illustrate the fact that these inequalities are not usually significant from numerical point of view, let us consider the simplest case, given by the Wiener process (Brownian motion). Let $\{W(t) : t \in [0, 1]\}$ be the Brownian motion on the unit interval and M its maximum. It is well known, McKean(1969), that the reflection principle implies that the distribution of M is that of the absolute value of a standard normal variable. It implies that

$$E(M) = \sqrt{\frac{2}{\pi}} = 0.7989\dots$$

and that the median $\mu(M)$ satisfies $\mu(M) = 0.675\dots$ If we apply Borell's type inequality to the Wiener process (Brownian motion) (with the advantage that the mean and the median are known, which is of course exceptional), we get:

u	true values of $P(M_W > u)$	Borell's b. mean	Borell's b. median
2	0.045	0.4855	0.2077
3	0.0027	0.0885	0.0347
4	$6.33 \cdot 10^{-5}$	$5.93 \cdot 10^{-3}$	$1.98 \cdot 10^{-3}$
5	$5.73 \cdot 10^{-7}$	$1.46 \cdot 10^{-4}$	$4.32 \cdot 10^{-5}$

In this table we have taken unilateral versions of Borell's type inequality namely, (2.31) and (2.32). The inequality with the median is sharper but even in this very favorable case both inequalities are imprecise.

5. Dudley's inequality.

We now turn to obtaining a bound for $E(M_T)$. A classical result in this direction is the next theorem. The proof below is taken from Talagrand (1996).

Let $\{X(t) : t \in T\}$ be a stochastic process not necessarily Gaussian. As in Section 4.4 of Chapter 1 we define the canonical distance by $d(s, t) := \sqrt{E((X(t) - X(s))^2)}$ identifying as usual points s, t when $d(s, t) = 0$. We define the covering number $N_\varepsilon := N(T, d, \varepsilon)$ as in Definition 1.17, using this metric.

THEOREM 2.10 (Dudley). *With the preceding notations, assume that*

$$E(X(t)) = 0 \quad \text{for every } t \in T$$

$$(2.35) \quad P(|X(t) - X(s)| > u) \leq 2e^{-\frac{1}{2} \frac{u^2}{d^2(s,t)}} \quad \text{for all } s, t \in T, \quad u > 0$$

Then,

$$(2.36) \quad E(\sup_{t \in T} X(t)) \leq K \int_0^{+\infty} (\log N_\varepsilon)^{\frac{1}{2}} d\varepsilon$$

where K is a universal constant.

Let us make some remarks on the statement before giving the proof.

- It is clear from the definition that $0 < \varepsilon < \varepsilon'$ implies $N_\varepsilon \geq N_{\varepsilon'}$. Hence if $N_{\varepsilon'} = +\infty$ for some $\varepsilon' > 0$ then $N_\varepsilon = +\infty$ for all $\varepsilon < \varepsilon'$ and the integral in the right-hand side of (2.36) is $+\infty$. In particular, this is the case when $\text{diam}(T)$ (the diameter of T) is ∞ . On the other hand, if $\text{diam}(T) < \infty$ and $\varepsilon > \text{diam}(T)$, then $N_\varepsilon = 1$, hence the integral in the right-hand side of (2.36) is in fact an integral over a finite interval.
- Condition (2.35) is easily verified for Gaussian processes.

In fact, in this case, if $u > 0$:

$$P(|X(t) - X(s)| > u) = \sqrt{\frac{2}{\pi}} \int_{\frac{u}{d(s,t)}}^{+\infty} \frac{1}{u} \exp\left(-\frac{y^2}{2}\right) dy \leq \sqrt{\frac{2}{\pi}} \frac{d(s,t)}{u} \exp\left(-\frac{u^2}{2d^2(s,t)}\right).$$

If $\frac{d(s,t)}{u} \leq \sqrt{2\pi}$ (2.35) follows.

If $\frac{d(s,t)}{u} > \sqrt{2\pi}$, then $2 \exp\left(-\frac{u^2}{2d^2(s,t)}\right) > 2 \exp\left(-\frac{1}{4\pi}\right) > 1$ and (2.35) also holds true.

- Under the general conditions of the theorem, it is necessary to precise the meaning of $E(\sup_{t \in T} X(t))$ to avoid measurability problems. This will be, by definition equal to

$$\sup_{F \subset T, F \text{ finite}} E\left(\sup_{t \in F} X(t)\right)$$

It is easy to see that if the conditions of the previous section hold true, this coincides with the ordinary expectation of $\sup_{t \in T} X(t)$. This will be the situation in the processes we will deal with in the remaining of the book.

So, it will be sufficient to prove (2.36) when one replaces in the left-hand side $\sup_{t \in T} X(t)$ by $\sup_{t \in F} X(t)$ F a finite subset of T .

PROOF. According to the previous remark, it suffices to consider the case $\text{diam}(T) < \infty$, since otherwise the right-hand side of (2.36) is infinite. Let F be a finite subset of T . Choose any $t_0 \in F$ and fix it for the remaining of the proof. Since the process is centered we have

$$(2.37) \quad E\left(\sup_{t \in F} X(t)\right) = E\left(\sup_{t \in F} (X(t) - X(t_0))\right) = \int_0^{+\infty} P(\sup_{t \in F} (X(t) - X(t_0)) > x) dx$$

given that $\sup_{t \in F} (X(t) - X(t_0)) \geq 0$. Let j_0 be the (unique) integer such that

$$2^{-j_0} < \text{diam}(T) \leq 2^{-j_0+1}$$

We define the following sequence $\{E_j\}_{j=j_0, j_0+1, \dots}$ of subsets of T :

The first member of the sequence is $\{E_{j_0}\} = \{t_0\}$

For each integer $j \geq j_0 + 1$, take a set of $N_{2^{-j}}$ closed balls of radius 2^{-j} such that the union covers T (which exists, according to the definition of N_ε) and let E_j be the set of the centers of these balls. This implies that for each $t \in T$ and each $j \geq j_0 + 1$, one can define $\pi_j(t) \in E_j$ such that

$$d(t, \pi_j(t)) \leq 2^{-j}.$$

Put also $\pi_{j_0}(t) = t_0$ for every $t \in T$. Clearly:

$$d(\pi_j(t), \pi_{j-1}(t)) \leq d(\pi_j(t), t) + d(\pi_{j-1}(t), t) \leq 3 \cdot 2^{-j}, \quad j \geq j_0 + 2.$$

and

$$d(\pi_{j_0+1}(t), \pi_{j_0}(t)) \leq \text{diam}(T) \leq 2^{-j_0+1}$$

so that we have

$$(2.38) \quad d(\pi_j(t), \pi_{j-1}(t)) \leq 4 \cdot 2^{-j}, \quad j \geq j_0 + 1.$$

Let us prove that for each $t \in T$ one has

$$(2.39) \quad a.s. \quad X(t) - X(t_0) = \sum_{j=j_0+1}^{\infty} (X(\pi_j(t)) - X(\pi_{j-1}(t))).$$

In fact, using the hypothesis and (2.38), for $\alpha_j > 0$:

$$\mathbb{P}(|X(\pi_j(t)) - X(\pi_{j-1}(t))| \geq \alpha_j) \leq 2 \exp\left(-\frac{\alpha_j^2}{2d^2(\pi_j(t), \pi_{j-1}(t))}\right) \leq 2 \exp\left(-\frac{\alpha_j^2}{32 \times 2^{-2j}}\right).$$

Taking $\alpha_j = 2^{-j/2}$ and applying the Borel-Cantelli Lemma, it follows that almost surely the series in (2.39) converges. On the other hand,

$$\mathbb{E} \left[(X(\pi_j(t)) - X(t))^2 \right] = d^2(\pi_j(t), t) \rightarrow 0 \quad \text{as } j \rightarrow \infty.$$

This proves (2.39).

It follows from (2.39) that if $\{a_j\}_{j=j_0, j_0+1, \dots}$ is a sequence of positive numbers, and $u > 0$:

$$\begin{aligned} \mathbb{P}\left(\sup_{t \in F} (X(t) - X(t_0)) > u \sum_{j=j_0+1}^{\infty} a_j\right) \\ \leq \mathbb{P}(\exists t \in F \text{ and } \exists j > j_0 \text{ such that } X(\pi_j(t)) - X(\pi_{j-1}(t)) > u a_j). \end{aligned}$$

Use now that there are at most $N_{2^{-j}} N_{2^{-(j-1)}}$ points in the product set $E_j \times E_{j-1}$ which implies that

$$(2.40) \quad \mathbb{P}\left(\sup_{t \in F} (X(t) - X(t_0)) > u \sum_{j=j_0+1}^{\infty} a_j\right) \leq \sum_{j=j_0+1}^{\infty} N_{2^{-j}} N_{2^{-(j-1)}} 2 \exp\left(-\frac{u^2 a_j^2}{32 \times 2^{-2j}}\right).$$

Choosing

$$a_j = 4 \cdot 2^{-j+\frac{1}{2}} \left[\log(2^{j-j_0} N_{2^{-j}} N_{2^{-(j-1)}}) \right]^{\frac{1}{2}}$$

the expression in the right-hand side of (2.40) becomes, for $u \geq 1$:

$$\begin{aligned} 2 \sum_{j=j_0+1}^{\infty} N_{2^{-j}} N_{2^{-(j-1)}} (2^{j-j_0} N_{2^{-j}} N_{2^{-(j-1)}})^{-u^2} &\leq 2 \sum_{j=j_0+1}^{\infty} 2^{-(j-j_0)u^2} \\ &\leq 2 \cdot 2^{-u^2} \sum_{k=0}^{\infty} 2^{-k} = 4 \cdot 2^{-u^2}. \end{aligned}$$

If we denote

$$S = \sum_{j=j_0+1}^{\infty} a_j$$

then if $v/S \geq 1$ we get:

$$\mathbb{P}\left(\sup_{t \in F} (X(t) - X(t_0)) > v\right) \leq 4 e^{-\frac{v^2}{S^2}}$$

which implies

$$(2.41) \quad \begin{aligned} \mathbb{E} \left[\sup_{t \in F} (X(t) - X(t_0)) \right] &\leq S + \int_S^{+\infty} \mathbb{P}\left(\sup_{t \in F} (X(t) - X(t_0)) > v\right) dv \\ &\leq S + 4 \int_S^{+\infty} e^{-\frac{v^2}{S^2}} dv \leq 2.S \end{aligned}$$

by a simple computation. On the other hand:

$$\begin{aligned} S &= \sum_{j=j_0+1}^{\infty} a_j \leq \sum_{j=j_0+1}^{\infty} 4 \cdot 2^{-j+\frac{1}{2}} [(j-j_0) \log 2 + 2(\log N_{2^{-j}})]^{\frac{1}{2}} \\ &\leq \sum_{j=j_0+1}^{\infty} 4 \cdot 2^{-j+\frac{1}{2}} \left[(j-j_0)^{\frac{1}{2}} (\log 2)^{\frac{1}{2}} + 2^{\frac{1}{2}} (\log N_{2^{-j}})^{\frac{1}{2}} \right] = T_1 + T_2. \end{aligned}$$

For T_1 we have

$$\begin{aligned} T_1 &\leq 4 \cdot 2^{\frac{1}{2}} (\log 2)^{\frac{1}{2}} \sum_{j=j_0+1}^{\infty} 2^{-j} (j-j_0) = 16 \cdot \sqrt{2} (\log 2) 2^{-(j_0+1)} \\ &\leq 16\sqrt{2} \sum_{j=j_0+1}^{\infty} 2^{-(j+1)} (\log N_{2^{-j}})^{\frac{1}{2}}. \end{aligned}$$

because $N_{2^{-j}} \geq 2$ for $j \geq j_0 + 1$ given that the definition of j_0 implies that one needs at least two balls of radius $2^{-(j_0+1)}$ to cover T .

As for T_2 $T_2 \leq 16 \sum_{j=j_0+1}^{\infty} 2^{-(j+1)} (\log N_{2^{-j}})^{\frac{1}{2}}$.

Putting the two pieces together we obtain

$$S \leq 16(1 + \sqrt{2}) \sum_{j=j_0+1}^{\infty} 2^{-(j+1)} (\log N_{2^{-j}})^{\frac{1}{2}} \leq 16(1 + \sqrt{2}) \int_0^{2^{-(j_0+1)}} (\log N_\varepsilon)^{\frac{1}{2}} d\varepsilon$$

where the last inequality is a standard lower bound for the integral of the monotone decreasing function $\varepsilon \rightsquigarrow (\log N_\varepsilon)^{\frac{1}{2}}$ by Riemann's sums. This finishes the proof. \square

Exercises

EXERCISE 2.1. Give a direct geometric proof of the Plackett-Slepian Lemma without using Fourier transform methods.

Hint:

1) Prove the Lemma for $n = 2$ by means of a comparison of measures in the plane.

2) For general n , it suffices to prove that $P(X_1 \leq a_1, \dots, X_n \leq a_n)$ increases - in the broad sense - if ONE of the covariances r_{jk} ($j \neq k$) increases, say r_{12} . For that purpose, write $X = A\xi$, A a non-random supertriangular matrix and ξ standard normal in \mathbb{R}^n .

Then, reduce the problem to dimension 2 by means of conditioning on the values of ξ_3, \dots, ξ_n .

EXERCISE 2.2. Prove that a direct consequence of the Normal comparison lemma is that if X_1, \dots, X_n are standard normal variables with $\text{Cov}(X_i, X_j) = r_{ij}$, then for any reals numbers u_1, \dots, u_n

$$P\left(\prod_{j=1}^n \{X_j \leq u_j\}\right) - \prod_{j=1}^n P\{X_j \leq u_j\} \leq \frac{1}{4} \sum_{1 \leq j < k \leq n} |r_{ij}| \exp\left(-\frac{u_j^2 + u_k^2}{2(1 + |r_{ij}|)}\right),$$

EXERCISE 2.3. Give an example showing that in the Plackett-Slepian version of the Normal comparison lemma one can not withdraw the equality of variances condition.

Crossings and Rice formulas for 1-dimensional parameter processes

1. Rice Formulas

Let $f : I \rightarrow \mathbb{R}$ be a real-valued function defined on an interval I of the real line. We will denote:

$$C_u(f, I) := \{t \in I : f(t) = u\}$$

$$N_u(f, I) := \#C_u(f; I).$$

$C_u(f, I)$ is the set of roots of the equation $f(t) = u$ in the interval I and $N_u(f, I)$ the number of these roots, that may be finite or infinite. We will usually replace $C_u(f, I)$ by C_u (respectively $N_u(f, I)$ by N_u) in case there is no doubt about the function f and the interval I .

In a similar way, if f is differentiable we define:

$$U_u(f, I) := \#\{t \in I : f(t) = u, f'(t) > 0\}$$

$$D_u(f, I) := \#\{t \in I : f(t) = u, f'(t) < 0\}.$$

N_u (resp. U_u, D_u) will be called the “number of crossings” (resp. “up-crossings”, “down-crossings”) of the level u by the function f on the interval I .

Our interest will be focused on $N_u(X, I), U_u(X, I), D_u(X, I)$ when $X(\cdot)$ is a path of a stochastic process. Even though these random variables are important in a large variety of problems, their probability distributions are unknown except for a small number of trivial cases. The Rice formulas that we are going to study in this chapter provide certain expressions, having the form of integral formulas, for the moments of $N_u(X, I), U_u(X, I), D_u(X, I)$ and also some other related random variables.

Rice formulas for one-parameter stochastic processes have been used for a long time in various contexts, such as telecommunications and signal processing (Rice, 1944-1945), ocean waves (Longuet-Higgins, 1957, 1962a,b) and random mechanics (Krée & Soize, 1983).

Rigorous results and a systematic treatment of the subject in the case of Gaussian processes came in the 1960's with the works - among others - of Belayev (1966), Ito (1964), Cramér (1965,1966). A landmark in the subject was the book by Cramér and Leadbetter (1967). The simple proof we have included below for general - not necessarily stationary - Gaussian processes with C^1 -paths is given here for the first time. Formulas for wider classes of processes can be found for example in Marcus (1977), Adler (1981) and Wschebor (1985). The proof of Theorem 3.4 which contains Rice Formula for general processes, not necessarily Gaussian, is an adaptation from the last reference.

We will say that the real-valued function f defined on the interval $I = [t_1, t_2]$ satisfies hypothesis $H_{1,u}$ if:

- f is a function of class C^1 ;
- $f(t_1) \neq u, f(t_2) \neq u$;
- $\{t : t \in I, f(t) = u, f'(t) = 0\} = \emptyset$

LEMMA 3.1 (Kac's counting formula). *If f satisfies $H_{1,u}$, then*

$$(3.1) \quad N_u(f, I) = \lim_{\delta \rightarrow 0} \frac{1}{2\delta} \int_I \mathbf{1}_{\{|f(t)-u|<\delta\}} |f'(t)| dt$$

PROOF. The hypothesis $H_{1,u}$ implies that $N_u(f, I)$ is finite, say $N_u(f, I) = n$. If $n = 0$, the result is obvious, since the integrand in the right-hand side of (3.1) is identically zero if δ is small enough. If $n \geq 1$, put

$$C_u(f, I) = \{s_1, \dots, s_n\}.$$

Since $f'(s_j) \neq 0$ for every $j = 1, \dots, n$, if $\delta > 0$ is small enough the inverse image of the interval $(u - \delta, u + \delta)$ by the function f is the union of exactly n pairwise disjoint intervals J_1, \dots, J_n which contain respectively the points s_1, \dots, s_n . The restriction of f to each one of the intervals J_k ($k = 1, \dots, n$) is a diffeomorphism and one easily checks changing variables that

$$\int_{J_k} |f'(t)| dt = 2\delta$$

for each k . So, if $\delta > 0$ is small enough:

$$\frac{1}{2\delta} \int_I \mathbf{1}_{\{|f(t)-u|<\delta\}} |f'(t)| dt = \frac{1}{2\delta} \sum_{k=1}^n \int_{J_k} |f'(t)| dt = n$$

and we are done. \square

Remarks on the lemma. The lemma holds true for polygonal f , even though these are not C^1 . More precisely, let

$$t_1 = \tau_0 < \tau_1 < \dots < \tau_m = t_2$$

be a partition of the interval $[t_1, t_2]$ and f a function having the polygonal graph with vertices $(\tau_i, f(\tau_i))$, $i = 0, 1, \dots, m$. Then, if $f(\tau_i) \neq u$ for $i = 0, 1, \dots, m$, formula (3.1) holds true.

The proof is immediate, since formula (3.1) is satisfied for each partition interval and, under these hypotheses, is additive as a function of I .

Moreover, notice that if f is such a polygonal function, then the expression

$$\frac{1}{2\delta} \int_I \mathbf{1}_{\{|f(t)-u|<\delta\}} |f'(t)| dt$$

in the right-hand side of (3.1) is bounded by m . This is again simple, since the integral on each partition interval is bounded by 1 if it contains a crossing point and by 1/2 if it does not.

Some basic ideas. An informal presentation of Rice formula for the expectation $E(N_u(X, I))$ where X is a stochastic process, can be the following: replace the function f in (3.1) by the random path $X(\cdot)$, and take expectations in both sides. Then:

$$\begin{aligned} E(N_u(X, I)) &= \lim_{\delta \rightarrow 0} E\left(\frac{1}{2\delta} \int_I \mathbf{1}_{\{|X(t)-u|<\delta\}} |X'(t)| dt\right) \\ &= \int_I dt \lim_{\delta \rightarrow 0} \int_{u-\delta}^{u+\delta} E(|X'(t)| | X(t) = x) p_{X(t)}(x) dx = \int_I E(|X'(t)| | X(t) = u) p_{X(t)}(u) dt. \end{aligned}$$

We will pay attention in this chapter to the justification of these equalities, in what concerns the passages to the limit. It is easy to prove weak forms of Rice formula such as equality for almost every level u (see Exercise 3.8), or to give upper-bounds for $E(N_u(X, I))$ (see Exercise 3.9).

However, a formula for almost every u is not satisfactory for a number of uses. For example, if one is willing to compute the moments of the number of critical points or the number of local maxima of a random function, one has to count the number of points in which the derivative is equal to zero, and a formula of this kind, valid for almost every u is uninteresting, one needs it for $u = 0$. So, it is worth to spend some energy to prove an exact formula for each level u . In all cases, some hypotheses on the processes will be necessary (see a simple counter-example in which the formula fails to hold true in Exercise 3.3).

When the process $\{X(t) : t \in \mathbb{R}\}$ is Gaussian centered stationary with variance 1, Rice formula for the expectation takes the simple form:

$$(3.2) \quad E(N_u(X, I)) = \frac{\sqrt{\lambda_2}}{\pi} e^{-u^2/2} |I|;$$

$|I|$ denotes the length of the interval I . This formula is due to S.O. Rice, who stated it for the first time in a series of pioneering papers pointing to electrical engineering applications (1944, 1945).

1.1. Gaussian case. We start with a statement and proof for Gaussian processes.

THEOREM 3.2 (Gaussian Rice formula). *Let $\mathcal{X} = \{X(t) : t \in I\}$, I an interval in the real line, be a Gaussian process having C^1 -paths. Let k be a positive integer. We assume that for every k pairwise distinct points t_1, \dots, t_k in I the joint distribution of $X(t_1), \dots, X(t_k)$ does not degenerate.*

Then

(3.3)

$$\mathbb{E}(N_u^{[k]}) = \int_{I^k} \mathbb{E}(|X'(t_1) \dots X'(t_k)| | X(t_1) = u \dots X(t_k) = u) p_{X(t_1), \dots, X(t_k)}(u, \dots, u) dt_1 \dots dt_k,$$

where

$$N_u := N_u(X, I)$$

$$m^{[k]} := m(m-1)\dots(m-k+1) \text{ if } m, k \text{ are positive integers, } m \geq k \\ := 0 \text{ otherwise}$$

PROOF. *Step 1*

Let $k = 1$. With no loss of generality, we assume that $I = [0, 1]$. Define $X^{(n)}(t)$ as the dyadic polygonal approximation of $X(t)$. As in the proof of Theorem 1.6, $X^{(n)}(t) - X(t)$ tends uniformly to zero and is bounded by the random variable $2 \sup_{t \in I} |X(t)|$, which has finite moments of all orders,

because of the results of Chapter 2. Using dominated convergence, it follows that $\text{Var}(X^{(n)}(t))$ converges uniformly for $t \in I$ to $\text{Var}(X(t))$. So, for n large enough, $\text{Var}(X^{(n)}(t)) \geq b$ for some $b > 0$ and all $t \in [0, 1]$. (The reader might show this using the more elementary arguments in the section on the normal distribution of Chapter 1).

For such an n , a.s. the process $X^{(n)}(t)$ does not take the value u at the partition points $j \cdot 2^{-n}$ ($j = 0, 1, \dots, 2^n$), since the random variable $X(t)$ has a density for each $t \in I$. So using the remarks after Lemma 3.1, we obtain

$$(3.4) \quad N_u(X^{(n)}, I) = \lim_{\delta \rightarrow 0} \frac{1}{2\delta} \int_I \mathbf{1}_{\{|X^{(n)}(t) - u| < \delta\}} |X^{(n)'}(t)| dt \text{ a.s.}$$

and the expression next to the limit in the right-hand side of (3.4) is bounded by 2^n .

Applying dominated convergence as $\delta \rightarrow 0$, for fixed n :

$$(3.5) \quad \mathbb{E}(N_u(X^{(n)}, I)) = \lim_{\delta \rightarrow 0} \frac{1}{2\delta} \int_I \mathbb{E}(\mathbf{1}_{\{|X^{(n)}(t) - u| < \delta\}} |X^{(n)'}(t)|) dt \\ = \lim_{\delta \rightarrow 0} \int_I dt \frac{1}{2\delta} \int_{u-\delta}^{u+\delta} \mathbb{E}(|X^{(n)'}(t)| | X^{(n)}(t) = x) p_{X^{(n)}(t)}(x) dx,$$

where the conditional expectation is the one defined by means of Gaussian regression.

Since the process has continuous sample paths, its expectation $m(t)$ and covariance $r(s, t)$ are continuous (see Chapter 1, Section 4.3). On the other hand, the regression formulas show that $\mathbb{E}(|X^{(n)'}(t)| | X^{(n)}(t) = x) p_{X^{(n)}(t)}(x)$ is a continuous function of the pair (t, x) and thus, it is bounded for $t \in I$ and x in a neighborhood of u . This implies that we may pass the limit sign inside the integral in the right-hand side of (3.5), so that:

$$(3.6) \quad \mathbb{E}(N_u(X^{(n)}, I)) = \int_I \mathbb{E}(|X^{(n)'}(t)| | X^{(n)}(t) = u) p_{X^{(n)}(t)}(u) dt.$$

To finish the proof, let us take limits in both sides as $n \rightarrow +\infty$ in (3.6). By Ylvisaker's Theorem (Theorem 1.21), with probability one there exists no local extrema at the level u . An instant reflection shows this implies that a.s.,

$$N_u(X^{(n)}, I) \uparrow N_u(X; I)$$

so that the left-hand side of (3.6) tends to $E(N_u(X, I))$, using monotone convergence. On the other hand, as already mentioned, as $n \rightarrow +\infty$ the expectation and variance matrix of the pair $(X^{(n)}(t), X^{(n)'}(t))$ converge uniformly to those of $(X(t), X'(t))$ and this implies the convergence of the right-hand side of (3.6) to the corresponding expression for $X(t)$. this finishes the proof for $k = 1$.

Step 2. For $k > 1$, let us denote by

$$C_u^k := C_u \times \dots \times C_u, \quad \text{with } C_u = C_u(X, I)$$

the cartesian product of C_u k times by itself and

$$\mu(J) = \#(C_u^k \cap J)$$

the number of points of C_u^k belonging to J for each Borel subset J of I^k .

Let $D_k(I)$ the diagonal set of the cube I^k defined as

$$(3.7) \quad D_k(I) = \{(t_1, \dots, t_k) : t_j \in I \text{ for } j = 1, \dots, k \text{ and there exist } j, j', j \neq j' \text{ such that } t_j = t_{j'}\}$$

It is easy to check that

$$N_u^{[k]} = \mu(I^k \setminus D_k(I))$$

So, it suffices to prove that

$$(3.8) \quad E(\mu(J)) = \int_J A_{t_1, \dots, t_k}(u, \dots, u) dt_1 \dots dt_k$$

where

$$(3.9) \quad A_{t_1, \dots, t_k}(u_1, \dots, u_k) := E(|X'(t_1) \dots X'(t_k)| | X(t_1) = u_1, \dots, X(t_k) = u_k) p_{X(t_1), \dots, X(t_k)}(u_1, \dots, u_k)$$

for every compact rectangle $J = J_1 \times \dots \times J_k$ contained in $I^k \setminus D_k(I)$ (which amounts to saying that the closed intervals J_1, \dots, J_k are pairwise disjoint). In fact, if this is proved, then the two Borel measures

$$J \rightsquigarrow E(\mu(J))$$

$$J \rightsquigarrow \int_J A_{t_1, \dots, t_k}(u, \dots, u) dt_1 \dots dt_k$$

coincide on these rectangles, hence on all Borel subsets of $I^k \setminus D_k(I)$. So,

$$E(N_u^{[k]}) = \int_{I^k \setminus D_k(I)} A_{t_1, \dots, t_k}(u, \dots, u) dt_1 \dots dt_k.$$

This proves (3.3) since $D_k(I)$ has Lebesgue measure zero.

To end up, let us turn to the proof of (3.8). We use the same arguments as in Step 1: first, we prove the equality for the polygonal approximation using Kac's formula and second, a similar domination argument allows to pass to the limit as one refines the partition. \square

Remark. A byproduct of Rice formula for $k = 1$ in the Gaussian case is that, under the conditions of the theorem, $E(N_u)$ is finite. This follows from the fact that the right-hand side of (3.3) is finite when $k = 1$ since it is the integral of a bounded function on a bounded interval.

For $k > 1$ both sides in (3.3) can be infinite.

1.2. Non-Gaussian case. For general processes, Lemma 3.1 will still be useful to get an upper bound for $E(N_u(X, I))$ via Fatou's Lemma. The next result will be helpful in the opposite direction.

LEMMA 3.3. *Let f be a function that satisfies $H_{1,u}$ and let $0 < \varepsilon < \delta < \frac{t_2 - t_1}{2}$. Let ψ be a real-valued function of one real variable, of class C^1 , with support contained in $[-1, 1]$, $\psi(s) \geq 0$, $\int_{\mathbb{R}} \psi(s) ds = 1$. We define $\psi_\varepsilon(s) = \frac{1}{\varepsilon} \psi(\frac{s}{\varepsilon})$ and, for each locally integrable function g ,*

$$g_\varepsilon(t) = (\psi_\varepsilon * g)(t) = \int_{\mathbb{R}} \psi_\varepsilon(t - s) g(s) ds$$

the convolution of g with the approximation of unity ψ_ε .

Then

$$(3.10) \quad N_u(f; I) \geq \int_{I_{-\delta}} |g'_\varepsilon(t)| dt$$

where

$$g(t) = \mathbf{1}_{(u, +\infty)}(f(t)) \quad \text{and} \quad I_{-\delta} = [t_1 + \delta, t_2 - \delta].$$

PROOF. By a duality argument, it suffices to show that

$$(3.11) \quad N_u(f; I) \geq \int_{I_{-\delta}} v(t) g'_\varepsilon(t) dt$$

for every \mathcal{C}^1 function v with support in $I_{-\delta}$ and such that $\|v\|_\infty \leq 1$.

Let $\{h_m\}_{m=1,2,\dots}$ be a sequence of \mathcal{C}^1 functions that approximate the step function $\mathbf{1}_{x>u}$. More precisely, h_m is monotone increasing in the broad sense, and $h_m(x) = 0$ for $x \leq u$, $h_m(x) = 1$ for $x \geq u + \frac{1}{m}$.

Applying dominated convergence we obtain:

$$\begin{aligned} \int_{I_{-\delta}} v(t) g'_\varepsilon(t) dt &= \int_{I_{-\delta}} v(t) dt \int_{\mathbb{R}} \psi'_\varepsilon(t-s) \mathbf{1}_{(u, +\infty)}(f(s)) ds \\ &= \lim_{m \rightarrow \infty} \int_{I_{-\delta}} v(t) dt \int_{\mathbb{R}} \psi'_\varepsilon(t-s) h_m(f(s)) ds \end{aligned}$$

Integrate by parts, use Fubini's Theorem and observe that if $t \in I_{-\delta}$ and $t-s$ is in the support of ψ_ε , then s must be in I :

$$\begin{aligned} \int_{I_{-\delta}} v(t) g'_\varepsilon(t) dt &= \lim_{m \rightarrow \infty} \int_{I_{-\delta}} v(t) dt \int_{\mathbb{R}} \psi_\varepsilon(t-s) h'_m(f(s)) f'(s) ds \\ &= \lim_{m \rightarrow \infty} \int_{\mathbb{R}} h'_m(f(s)) f'(s) ds \int_{I_{-\delta}} v(t) \psi_\varepsilon(t-s) dt \\ &\leq \overline{\lim}_{m \rightarrow \infty} \int_I h'_m(f(s)) |f'(s)| ds. \end{aligned}$$

The hypothesis $H_{1,u}$ plus an argument similar to the one in Lemma 3.1 show that for m large enough

$$\int_I h'_m(f(s)) |f'(s)| ds = N_u(f; I)$$

This shows (3.11). □

Next, we are going to impose a certain number of hypotheses on the stochastic process $\mathcal{X} = \{X(t) : t \in I\}$ for which we will state and prove the Rice formulas. They are the following:

A1) The paths of \mathcal{X} are of class \mathcal{C}^1 .

A2k) Let k be a positive integer. For any choice of the k -tuples $(t_1, \dots, t_k), (t'_1, \dots, t'_k) \in I^k \setminus D_k(I)$, where $D_k(I)$ is the diagonal set defined in (3.7), the random vector $(X(t_1), \dots, X(t_k), X'(t'_1), \dots, X'(t'_k))$ has a density - in \mathbb{R}^{2k} - denoted by

$$(3.12) \quad p_{X(t_1), \dots, X(t_k), X'(t'_1), \dots, X'(t'_k)}(x_1, \dots, x_k, x'_1, \dots, x'_k)$$

We also define:

$$I_k(x_1, \dots, x_k) := \int_{I^k} A_{t_1, \dots, t_k}(x_1, \dots, x_k) dt_1 \dots dt_k$$

where $A_{t_1, \dots, t_k}(x_1, \dots, x_k)$ has already been defined in the proof of Rice formula in the Gaussian case.

Notice that in the general case, due to the conditional expectation, this function is only defined for almost every point (x_1, \dots, x_k) . We are assuming that it has a continuous version and it is this

version that appears in what follows.

These integrals may have the value $+\infty$, but are always well defined.

We will assume that the density (3.12) is a continuous function of (x_1, \dots, x_k) at the point (u, \dots, u) (the other variables remaining constant) and of (t_1, \dots, t_k) in $I^k \setminus D_k(I)$ (the other variables remaining constant).

We also assume that $p_{X(t)}(x)$ is continuous for $t \in I$ and x in a neighborhood of u .

A3k) The function $(t_1, \dots, t_k, x_1, \dots, x_k) \rightsquigarrow A_{t_1, \dots, t_k}(x_1, \dots, x_k)$ is assumed to be continuous for (t_1, \dots, t_k) in $I^k \setminus D_k(I)$ and x_1, \dots, x_k in a neighborhood of (u, \dots, u) .

A4k)

$$\int_{\mathbb{R}^3} |x'_1|^{k-1} |x'_2 - x'_3| p_{X(t_1), \dots, X(t_k), X'(t'_1), X'(t'_2), X'(t'_3)}(x_1, \dots, x_k, x'_1, x'_2, x'_3) dx'_1 dx'_2 dx'_3$$

tends to zero as $t'_2 - t_1 \rightarrow 0$ uniformly for (t_1, \dots, t_k) in a compact subset of $I^k \setminus D_k(I)$ and x_1, \dots, x_k in a neighborhood of (u, \dots, u) .

THEOREM 3.4 (Rice Formula). *If \mathcal{X} satisfies A1), A2k), A3k), and A4k), then:*

(3.13)

$$\mathbb{E}(N_u^{[k]}) = \int_{I^k} \mathbb{E}(|X'(t_1) \dots X'(t_k)| | X(t_1) = u \dots X(t_k) = u) p_{X(t_1), \dots, X(t_k)}(u, \dots, u) dt_1 \dots dt_k,$$

PROOF. Using the same arguments as in the proof of Theorem 3.2, it is sufficient to prove that

$$(3.14) \quad \mathbb{E}(\mu(J)) = \int_J A_{t_1, \dots, t_k}(u, \dots, u) dt_1 \dots dt_k$$

for every compact rectangle $J = J_1 \times \dots \times J_k$ contained in $I^k \setminus D_k(I)$ (for $k = 1$, we put $D_k(I) = \emptyset$).

First we use Lemma 3.1. It is easy to check that almost surely the paths of the process satisfy hypothesis $H_{1,u}$.

$$\mu(J) = \prod_{i=1}^k N_u(X, J_i) = \lim_{\delta \rightarrow 0} \frac{1}{(2\delta)^k} \prod_{i=1}^k \left[\int_{J_i} \mathbf{1}_{\{|X(t_i) - u| < \delta\}} |X'(t_i)| dt_i \right]$$

By Fatou's Lemma, the definition of $A_{t_1, \dots, t_k}(x_1, \dots, x_k)$ and hypothesis A3k), we obtain:

$$(3.15) \quad \begin{aligned} \mathbb{E}(\mu(J)) &\leq \liminf_{\delta \rightarrow 0} \frac{1}{(2\delta)^k} \int_J dt_1 \dots dt_k \int_{u-\delta}^{u+\delta} \dots \int_{u-\delta}^{u+\delta} A_{t_1, \dots, t_k}(x_1, \dots, x_k) dx_1 \dots dx_k \\ &= \int_J A_{t_1, \dots, t_k}(u, \dots, u) dt_1 \dots dt_k. \end{aligned}$$

The converse inequality is somewhat more complicated. We apply Lemma 3.3 to each one of the intervals J_1, \dots, J_k . We have:

$$\mathbb{E}(\mu(J)) = \mathbb{E}\left(\prod_{i=1}^k N_u(X, J_i)\right) \geq \mathbb{E}\left(\prod_{i=1}^k \left[\int_{(J_i)_{-\delta}} |g'_\varepsilon(t_i)| dt_i \right]\right)$$

where $g(t) = 1_{(u, +\infty)}(X(t))$.

Define the sequence of functions $\{h_m\}_{m=1,2,\dots}$ as in the proof of Lemma 3.3. Using dominated convergence, Fubini's Theorem and integration by parts provide

$$(3.16) \quad \mathbb{E}(\mu(J)) \geq \lim_{m \rightarrow \infty} \int_{J_{-\delta}} dt_1 \dots dt_k \mathbb{E}\left(\prod_{i=1}^k \left| \int_{\mathbb{R}} \psi_\varepsilon(t_i - s_i) h'_m(X(s_i)) X'(s_i) ds_i \right|\right)$$

with $J_{-\delta} = (J_1)_{-\delta} \times \dots \times (J_k)_{-\delta}$.

where we have put

$$\begin{cases} \tau_\ell = t_\ell & \text{when } \ell = 1, \dots, i-1 \\ \tau_\ell = s_\ell & \text{when } \ell = i+1, \dots, k \end{cases}$$

Now, if we choose m large enough, since the integrand in the right-hand side of (3.19) is zero when $|x_h - u| \geq \frac{1}{m}$ for some $h = 1, \dots, k$, we can apply hypothesis A4k) and the inner integral in (3.19) is uniformly small if $|s_i - t_i|$ is small.

This shows that the second member in (3.19) tends to zero as $\varepsilon \downarrow 0$ and on account of (3.16), (3.17) and (3.18) we obtain:

$$\mathbb{E}(\mu(J)) \geq \int_{J-\delta} A_{t_1, \dots, t_k}(u, \dots, u) dt_1 \dots dt_k.$$

The converse inequality to (3.15) follows by making $\delta \downarrow 0$. \square

2. Variants and Examples.

(1) Another Form of Rice formulas.

Under the hypotheses of Theorem 3.4 one can write Rice formula for the k -factorial moment of crossings also in the form:

$$\begin{aligned} \mathbb{E}(N_u(N_u - 1) \dots (N_u - k + 1)) &= \int_{I^k} dt_1 \dots dt_k \int_{\mathbb{R}^k} \left[\prod_{i=1}^k |x'_i| \right] \\ &\quad \cdot p_{X(t_1), \dots, X(t_k), X'(t_1), \dots, X'(t_k)}(u, \dots, u, x'_1, \dots, x'_k) dx'_1 \dots dx'_k \end{aligned}$$

(2) Factorial moments and ordinary moments.

A simple remark is that the ordinary moments of the random variable N_u , i.e. $\mathbb{E}(N_u^k)$ are linear combinations of the factorial moments $\mathbb{E}(N_u^{[j]})$, $j = 1, \dots, k$ with fixed integer coefficients depending only on k and j , and conversely. Hence, one can express $\mathbb{E}(N_u^k)$ as linear combinations of multiple integrals of the form that appear in Rice formulas.

(3) First moments.

In the case $k = 1$ we have the formula

$$(3.20) \quad \mathbb{E}(N_u) = \int_I dt \int_{\mathbb{R}} |x'| p_{X(t), X'(t)}(u, x') dx'$$

Essentially the same proof that we have given above for the Rice formula in the general non-Gaussian case, works when $k = 1$ under slightly weaker hypotheses (and easier to check, especially A4k)). The reader may check that (3.20) holds true if:

- i) $(t, x) \rightsquigarrow p_{X(t)}(x)$ is continuous for $t \in I$, x in a neighborhood of u .
- ii) $(t, x, x') \rightsquigarrow p_{X(t), X'(t)}(x, x')$ is continuous for $t \in I$ and x in a neighborhood of u and $x' \in \mathbb{R}$.
- iii) $\mathbb{E}(\omega_{X'}(I, \delta)) \rightarrow 0$ as $\delta \rightarrow 0$, where $\omega_{X'}(I, \delta)$ denotes the modulus of continuity of $X'(\cdot)$.

(4) Gaussian stationary processes.

This is an important case, in which we have mentioned the simple classical formula (3.2). Suppose the process \mathcal{X} is centered Gaussian stationary with \mathcal{C}^1 paths and covariance $\Gamma(t) = \mathbb{E}(X(s)X(s+t))$ normalized by $\Gamma(0) = 1$. It is clear that $\Gamma'(0) = \mathbb{E}(X(t)X'(t)) = 0$, since Γ has a maximum at the point $t = 0$. Given that the joint distribution of $X(t), X'(t)$ is Gaussian, this implies that for each t , $X(t), X'(t)$ are independent random variables. Hence:

$$p_{X(t), X'(t)}(u, x') = p_{X(t)}(u) \cdot p_{X'(t)}(x') = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \frac{1}{\sqrt{2\pi\lambda_2}} e^{-\frac{1}{2}\frac{x'^2}{\lambda_2}}$$

(notice that $\lambda_2 = -\Gamma''(0) = \mathbb{E}([X'(t)]^2)$). Substituting into (3.20) we get (3.2).

Formula (3.2) remains valid if we only require the Gaussian centered stationary process to have continuous paths (see Exercise 3.2).

(5) **General Gaussian processes.**

Verifying hypotheses A1), A2k), A3k) and Ak4) for non-Gaussian processes can be a non-trivial task. For Gaussian processes, this approach is tractable, as shown by the next proposition, that we include to see how the verification of the general hypotheses can be performed in this case. Of course, this has a limited interest, since the direct approach for Gaussian processes, as we have seen, is simpler and permits to deduce Rice formulas under weaker conditions.

PROPOSITION 3.5. *If \mathcal{X} is a real-valued centered Gaussian process defined on a compact interval I of the real line, has C^1 paths and the densities in A1), A2k), A3k) and Ak4) do not degenerate for a given k , then A1), A2k), A3k) and Ak4) are verified.*

PROOF. Let us recall that the functions $(s, t) \rightsquigarrow E(X(s)X(t))$, $(s, t) \rightsquigarrow E(X(s)X'(t))$, $(s, t) \rightsquigarrow E(X'(s)X'(t))$ are continuous, so that the densities, since they do not degenerate, are also continuous. We have

$$(3.21) \quad A_{t_1, \dots, t_k}(x_1, \dots, x_k) = E \left\{ \prod_{i=1}^k |X'(t_i)| / X(t_1) = x_1, \dots, X(t_k) = x_k \right\} p_{X(t_1), \dots, X(t_k)}(x_1, \dots, x_k)$$

and the expression in hypothesis Ak4) is

$$(3.22) \quad E \left\{ |X'(t'_1)|^{k-1} |X'(t'_2) - X'(t_1)| / X(t_1) = x_1, \dots, X(t_k) = x_k \right\} p_{X(t_1), \dots, X(t_k)}(x_1, \dots, x_k)$$

If (t_1, \dots, t_k) varies in a compact subset of $I^k \setminus D_k(I)$ and (x_1, \dots, x_k) in a neighborhood of u , the density $p_{X(t_1), \dots, X(t_k)}(x_1, \dots, x_k)$ is continuous and bounded.

We want to get rid of the conditional expectation in both expressions (3.21) and (3.22). For this purpose we use linear regression (see Chapter 1) and write

$$X'(t) = X'(t) - \sum_{h=1}^k c_h(t)X(t_h) + \sum_{h=1}^k c_h(t)X(t_h)$$

and choose $c_h(t)$ ($h = 1, \dots, k$) in such a way that

$$Y(t) = X'(t) - \sum_{h=1}^k c_h(t)X(t_h)$$

be orthogonal to the components of the random vector $(X(t_1), \dots, X(t_k))$.

Denote by $\Gamma(s, t)$ the covariance of the given process \mathcal{X} , $\Sigma = ((\Gamma(t_h, t_\ell))_{h, \ell=1, \dots, k})$ and $\gamma(t) = (\Gamma_1(t, t_1), \dots, \Gamma_1(t, t_k))^T$, where we have used Γ_1 for the partial derivative of Γ with respect to the first variable.

The orthogonality condition is

$$\gamma(t) = \Sigma \cdot c(t) \quad \text{with} \quad c(t) = (c_1(t), \dots, c_k(t))^T$$

so that

$$c(t) = \Sigma^{-1} \gamma(t)$$

The non-degeneracy hypothesis implies that the function

$$(t, t_1, \dots, t_k) \rightsquigarrow c(t) \text{ is continuous for } (t, t_1, \dots, t_k) \in I \times I^k \setminus D_k(I).$$

The conditional expectations in (3.21) and (3.22) become respectively the unconditional expectations:

$$E \left(\prod_{i=1}^k \left| Y(t_i) + \sum_{h=1}^k c_h(t_i)x_h \right| \right)$$

$$E \left(\left| Y(t'_1) + \sum_{h=1}^k c_h(t'_1)x_h \right|^{k-1} \left| Y(t'_2) + \sum_{h=1}^k c_h(t'_2)x_h - Y(t_1) - \sum_{h=1}^k c_h(t_1)x_h \right| \right)$$

To verify hypotheses A3k) and A4k) and finish the proof, one can now pass to the limit under the expectation, using dominated convergence on account of the integrability of the moments of the supremum of Gaussian process. This is left to the reader. \square

(6) **Stationary Gaussian Processes, non-degeneracy condition.**

Let us consider a Gaussian process defined on an interval of the real line, and t_1, \dots, t_n n distinct parameter values. It is in general non-trivial to check whether the random variables $X(t_1), \dots, X(t_n)$ have a non-degenerate joint distribution. Similar difficulties appear if one is willing to prove non-degeneracy of the joint distribution of the process and its derivative.

However, in the stationary case, one can give the following sufficient condition for non-degeneracy to hold: the support of the spectral measure μ^X of the process has some accumulation point. In particular, this happens if μ^X is not purely atomic, that is, if there does not exist a countable subset A of the reals, such that $\mu^X(A^C) = 0$. (See Exercises 3.4 and 3.5)

(7) **Finiteness of Moments of Crossings.**

It may happen that Rice formula (3.13) holds true but both sides are infinite. In a certain number of applications one wants to know whether $E(N_u^k)$ is finite, but is not so much interested in its value. On the other hand, the standard situation is that to compute the right-hand side of (3.13) or even to obtain good upper bounds for it, can be a very complicated or actually untractable problem.

From a numerical point of view, the general question of efficient procedures to compute approximately the moments of the number of crossings remains widely open. We will come back to this subject in the next two chapters, adding some more or less recent results. This will be done in the context of relating crossings to the distribution of the maximum of a stochastic process, even though it has an independent interest.

Finiteness of moments of crossings of Gaussian processes has been considered by Belayev (1966) Miroshin (1977), Cuzick (1975). For stationary Gaussian processes, the sufficient condition for finiteness of the variance of $N_0(X, I)$ in Exercise 3.6 below is in Cramér and Leadbetter's book, where an explicit formula for the variance is also given. Geman (1972) proved that this condition is also necessary for finiteness at the level $u = 0$. In a recent paper, Kratz and León (2006) proved that the same condition is necessary and sufficient for any level u and also for the number of crossings with some differentiable curves.

For non-Gaussian processes, sufficient conditions have been given in Besson and Wschebor (1983). The next theorem gives sufficient conditions, that are reasonably easy to check in specific cases, to be able to assure the finiteness of $E(N_u^k)$. It is taken from Nualart and Wschebor (1991).

THEOREM 3.6. *Let m be a positive real number. Consider a real-valued stochastic process $\mathcal{X} = \{X(t) : t \in I\}$ defined on an interval I of the real line, with paths of class C^{p+1} , $p > 2m$.*

We assume also that for each $t \in I$ the random variable $X(t)$ has a density and that for some $\eta > 0$

$$C = \sup \{p_{X(t)}(x) : t \in I, x \in [u - \eta, u + \eta]\} < \infty$$

Then

$$(3.23) \quad E(N_u^m) \leq C_{p,m} [1 + C + E(\|X^{(p+1)}\|_\infty)]$$

where $C_{p,m}$ is a constant depending only on p, m and the length of the interval I .

PROOF. With no loss of generality we assume that $I = [0, 1]$.
Using a standard bound for the expectation of non-negative random variables:

$$\begin{aligned} \mathbf{E}(N_u^m) &\leq \sum_{k=0}^{\infty} \mathbf{P}(N_u^m \geq k) \leq C_{p,m} + C_{p,m} \sum_{k > p^m} \mathbf{P}(N_u^m \geq (p+1)^m k) \\ &= C_{p,m} + C_{p,m} \sum_{k > p^m} \mathbf{P}(N_u \geq (p+1)k^{\frac{1}{m}}). \end{aligned}$$

Our aim is to give an upper bound for this sum.

We have the inclusion of events (use Rolle's Theorem, $|J|$ denotes here the length of the interval J):

$$\begin{aligned} (3.24) \quad &\left\{ N_u(X, I) \geq (p+1)k^{\frac{1}{m}} \right\} \\ &\subset \left\{ \exists \text{ an interval } J \subset I, |J| = k^{-\frac{1}{m}}, N_u(X, J) \geq (p+1) \right\} \\ &\subset \left\{ \begin{array}{l} \exists \text{ an interval } J \subset I, |J| = k^{-\frac{1}{m}}, \text{ and points } \tau_1, \dots, \tau_p \in J \\ \text{such that } X^{(j)}(\tau_j) = 0 \text{ for } j = 1, \dots, p \end{array} \right\} \end{aligned}$$

Let $\{\varepsilon_k\}_{k=1,2,\dots}$ be a sequence of positive real numbers (that we will choose afterwards) and denote for $k = 1, 2, \dots$:

$$A_k = \left\{ N_u(X, I) \geq (p+1)k^{\frac{1}{m}} \right\} \cap \left\{ \omega_{X^{(p)}}(I, k^{-\frac{1}{m}}) < \varepsilon_k \right\}$$

(as usual, $\omega_f(I, \delta)$ denotes the continuity modulus).

Let us consider the random open set in I

$$V_k = \left\{ t : |X'(t)| < \varepsilon_k k^{-\frac{p-1}{m}} \right\}.$$

We show that

$$A_k \subset \{N_u(X, V_k) \geq (p+1)\}.$$

To prove this it suffices to prove that if A_k occurs, then the random interval J that appears in (3.24) is contained in V_k since in that case, the number of roots of the equation $X(t) = u$ that belong to V_k will be greater or equal than $N_u(X, J) \geq (p+1)$.

Suppose that A_k occurs and $t \in J$. Then,

$$\left| X^{(p)}(t) \right| = \left| X^{(p)}(t) - X^{(p)}(\tau_p) \right| < \varepsilon_k$$

given that $X^{(p)}(\tau_p) = 0$, the definition of A_k and $|J| = k^{-\frac{1}{m}}$.

Similarly, using the Mean Value Theorem,

$$\left| X^{(p-1)}(t) \right| = \left| X^{(p-1)}(t) - X^{(p-1)}(\tau_{p-1}) \right| < \varepsilon_k k^{-\frac{1}{m}}.$$

In the same way we can increase step by step the order of the derivative, and obtain

$$|X'(t)| < \varepsilon_k k^{-\frac{p-1}{m}}$$

which shows that $t \in V_k$.

It follows that

$$(3.25) \quad \mathbf{P}(N_u \geq (p+1)k^{\frac{1}{m}}) \leq \mathbf{P}(\omega_{X^{(p)}}(I, k^{-\frac{1}{m}}) \geq \varepsilon_k) + \mathbf{P}(A_k).$$

The first term in (3.25) is bounded by

$$\mathbf{P}(\|X^{(p+1)}\|_{\infty} \geq k^{\frac{1}{m}} \varepsilon_k) \leq k^{-\frac{1}{m}} \varepsilon_k^{-1} \mathbf{E}(\|X^{(p+1)}\|_{\infty}).$$

As for the second term in (3.25),

$$\mathbf{P}(A_k) \leq \mathbf{P}(N_u(X, V_k) \geq (p+1)) \leq \frac{1}{p+1} \mathbf{E}(N_u(X, V_k)).$$

One can check as an exercise that Lemma 3.1 holds true, *mutatis mutandis*, whenever the set in which the function is defined on an open set in the real line - as is the case of V_k - instead of an interval. Hence, a.s.

$$N_u(X, V_k) = \lim_{\delta \rightarrow 0} \frac{1}{2\delta} \int_{V_k} \mathbf{1}_{\{|X(t)-u| < \delta\}} |X'(t)| dt$$

and applying Fatou's Lemma and the definition of the set V_k :

$$\begin{aligned} \mathbb{E}(N_u(X, V_k)) &\leq \liminf_{\delta \rightarrow 0} \frac{1}{2\delta} \mathbb{E} \left\{ \int_{V_k} \mathbf{1}_{\{|X(t)-u| < \delta\}} |X'(t)| dt \right\} \\ &\leq \liminf_{\delta \rightarrow 0} \frac{1}{2\delta} \mathbb{E} \left\{ \int_I \mathbf{1}_{\{|X(t)-u| < \delta\}} \varepsilon_k k^{-\frac{p-1}{m}} dt \right\} \\ &\leq \varepsilon_k k^{-\frac{p-1}{m}} \liminf_{\delta \rightarrow 0} \frac{1}{2\delta} \int_I dt \int_{u-\delta}^{u+\delta} p_{X(t)}(x) dx \leq C \varepsilon_k k^{-\frac{p-1}{m}}. \end{aligned}$$

Replacing in (3.25) and then substituting in the upper bound for $\mathbb{E}(N_u^m)$ we obtain:

$$\mathbb{E}(N_u^m) \leq C_{p,m} + C_{p,m} \left[C \sum_{k=1}^{\infty} \varepsilon_k k^{-\frac{p-1}{m}} + \mathbb{E}(\|X^{(p+1)}\|_{\infty}) \sum_{k=1}^{\infty} k^{-\frac{1}{m}} \varepsilon_k^{-1} \right].$$

Choosing

$$\varepsilon_k = k^{\beta - \frac{1}{m}} \quad \text{with} \quad 1 < \beta < \frac{p}{m} - 1$$

which is possible since $\frac{p}{m} > 2$, the two series converge and we have the statement of the theorem, with some new constant $C_{p,m}$. \square

COROLLARY 3.7. *If \mathcal{X} is Gaussian with \mathcal{C}^{∞} -paths and $\text{Var}(X(t)) \geq a > 0$ for $t \in I, x \in \mathbb{R}$, then*

$$\mathbb{E}(N_u^m) < \infty$$

for every $u \in \mathbb{R}$ and every $m = 1, 2, \dots$

PROOF. Using the results in Chapter 2, we know that $\mathbb{E}(\|X^{(p+1)}\|_{\infty}) < \infty$ for every $p = 1, 2, \dots$ and $p_{X(t)}(x) \leq \frac{1}{\sqrt{2\pi a}}$ for $t \in I, x \in \mathbb{R}$. \square

(8) Variations on Rice formulas.

In applications one frequently needs a certain number of variants of formula (3.13). We give here some examples.

a) If instead of all crossings we consider only up-crossings or down-crossings, under the same hypotheses as in Theorem 3.4 we obtain the following formulas:

$$\begin{aligned} (3.26) \quad \mathbb{E}(U_u^{[k]}) &= \mathbb{E}(U_u(U_u - 1) \dots (U_u - k + 1)) \\ &= \int_{I^k} \mathbb{E} \left(\prod_{i=1}^k X'^+(t_i) / X(t_1) = u, \dots, X(t_k) = u \right) p_{X(t_1), \dots, X(t_k)}(u, \dots, u) dt_1 \dots dt_k \end{aligned}$$

$$\begin{aligned} (3.27) \quad \mathbb{E}(D_u^{[k]}) &= \mathbb{E}(D_u(D_u - 1) \dots (D_u - k + 1)) \\ &= \int_{I^k} \mathbb{E} \left(\prod_{i=1}^k X'^-(t_i) / X(t_1) = u, \dots, X(t_k) = u \right) p_{X(t_1), \dots, X(t_k)}(u, \dots, u) dt_1 \dots dt_k \end{aligned}$$

Proofs are exactly the same.

b) If instead of counting crossings we count "marked crossings" that is, points t such that $X(t) = u$ and in which some other event is happening, we obtain various Rice-type formulas.

For example, let $\{Y(t) : t \in I\}$ be a second stochastic process, a, b extended real numbers (they may take the values $-\infty$ or $+\infty$), $a < b$ and define

$$N_u(X, Y; I, a, b) = \# \{t : t \in I, X(t) = u, a < Y(t) \leq b\}$$

Then

$$(3.28) \quad E(N_u(X, Y; I, a, b)) = \int_I dt \int_a^b dy \int_{\mathbb{R}} |x'| p_{X(t), X'(t), Y(t)}(u, x', y) dx'$$

Formulas similar to (3.13) can be written for the factorial moments of $N_u(X, Y; I, a, b)$. We leave to the reader establishing hypotheses and giving proofs. These follow the same lines as in Theorem 3.4.

A typical application is the computation of the moments of

$$M(X, I, a, b) = \# \{t : t \in I, X(\cdot) \text{ has a local maximum at } t, a < X(t) \leq b\}$$

where $a < b$. Under general conditions (the statement of which is left to the reader), almost surely,

$$M(X, I, a, b) = \# \{t : t \in I, X'(t) = 0, X''(t) < 0, a < X(t) \leq b\}.$$

It means that a.s. $M(X, I, a, b)$ is the number of down-crossings of the level 0 by the stochastic process $\{X'(t) : t \in I\}$ in which the process $X(t)$ itself takes values between a and b . So, we may apply (3.28) with X' instead of X , X instead of Y and down-crossings instead of crossings, to get:

$$E(M(X, I, a, b)) = \int_I dt \int_a^b dx \int_{-\infty}^0 |x''| p_{X(t), X'(t), X''(t)}(x, 0, x'') dx''.$$

Again, similar formulas hold true for higher factorial moments, under appropriate hypotheses on the process.

c) If ξ is a bounded random variable (one can relax this condition) and if the stochastic process $\{X(t) : t \in I\}$ satisfies the hypotheses of Theorem 3.4, then one has the more general equality:

$$(3.29) \quad E(\xi N_u^{[k]}) = \int_{I^k} E(\xi \prod_{i=1}^k |X'(t_i)| | X(t_1) = \dots = X(t_k) = u) \cdot p_{X(t_1), \dots, X(t_k)}(u, \dots, u) dt_1 \dots dt_k$$

The proof is left to the reader.

Exercises

EXERCISE 3.1. [Kac counting formula] Prove the following inequality, related to Lemma 3.1. Assume that

- $f : I \rightarrow \mathbb{R}$, $I = [t_1, t_2]$ is an absolutely continuous function,
- $f(t_1) \neq u$, $f(t_2) \neq u$,

Then:

$$N_u(f; I) \leq \liminf_{\delta \rightarrow 0} \frac{1}{2\delta} \int_I \mathbf{1}_{\{|f(t)-u| < \delta\}} |f'(t)| dt$$

EXERCISE 3.2. Prove that formula (3.2) is always true in the following sense: Let $\{X(t) : t \in I\}$ be a centered Gaussian stationary process with continuous paths, defined on a compact interval I of the real line, normalized by $r(0) = 1$. Then, if λ_2 is finite (3.2) holds true and if λ_2 is infinite, then $E\{N_u\} = +\infty$.

(This means that the remaining hypotheses are not necessary in this case).

EXERCISE 3.3. (A simple example in which Rice formula does not hold) Let $X(t) = \xi t$, $t \in [-1, 1]$, where ξ is a standard normal random variable. Show that $E(N_u^X)$ can be computed by means of Rice formula if $u \neq 0$, but that the formula fails to hold for $u = 0$.

EXERCISE 3.4. a) Prove that if the process $\{X(t) : t \in \mathbb{R}\}$ is Gaussian stationary, has \mathcal{C}^1 paths and the support of the spectral measure has an accumulation point, then the set of random variables $X(t_1), \dots, X(t_n)$ has a joint non-degenerate distribution for any choice of the distinct parameter values t_1, \dots, t_n .

b) Deduce that under the hypotheses in a), Rice formulas can be applied on any compact interval I and for any $k = 1, 2, \dots$

Hint for a): With no loss of generality, one can assume that the process is centered. Denote $Y = (X(t_1), \dots, X(t_n))^T$. The variance of the Gaussian vector Y is:

$$\Lambda = E(Y Y^T)$$

The aim is to prove that the quadratic form $F(z) = z^T \Lambda z$, $z \in \mathbb{R}^n$ is positive definite. To prove it, show that $F(z)$ can be written in terms of the spectral measure μ of the process X by means of the formula:

$$F(z) = \int_{\mathbb{R}} \left| \sum_{k=1}^n e^{it_k x} z_k \right|^2 \mu(dx).$$

Conclude that $F(z) > 0$ whenever $z \neq 0$, using that the function $x \rightsquigarrow \sum_{k=1}^n e^{it_k x} z_k$ is analytic.

EXERCISE 3.5. Assume that the process $X(t) : t \in \mathbb{R}$ verifies the hypotheses of part a) of the previous exercise, and moreover, that its paths are \mathcal{C}^k -functions, k an integer, $k \geq 1$. Then, for any choice of distinct parameter values t_1, \dots, t_n , the joint distribution of the random variables

$$X(t_1), \dots, X(t_n), X'(t_1), \dots, X'(t_n), \dots, X^{(k)}(t_1), \dots, X^{(k)}(t_n)$$

does not degenerate. (Hint: use the same method as in the previous exercise)

EXERCISE 3.6. Let $\{X(t) : t \in \mathbb{R}\}$ be a centered Gaussian stationary process. Assume the Geman Condition :

$$G1 \quad \Gamma(t) = E(X(s)X(s+t)) \neq \pm 1 \text{ for } t > 0$$

$$G2 \quad \Gamma(\tau) = 1 - \frac{\lambda_2 \tau^2}{2} + \theta(\tau) \text{ with}$$

$$\int \frac{\theta'(\tau)}{\tau^2} d\tau \text{ converges at } \tau = 0^+.$$

Prove that this condition is sufficient to have

$$E \left\{ [N_0(X, I)]^2 \right\} < \infty$$

for any bounded interval I .

EXERCISE 3.7. a) Let $f : J \rightarrow \mathbb{R}$ be a \mathcal{C}^1 -function, J an interval in the reals, and δ a positive number. Prove that:

$$\frac{1}{2\delta} \int_J \mathbf{1}_{\{|f(t)-u|\}} |f'(t)| dt \leq N_0(X', J) + 1.$$

b) Let $\{X(t) : t \in I\}$, I a compact interval in the reals, be a stochastic process with \mathcal{C}^1 -paths. Let k be an integer, $k \geq 1$, and $u \in \mathbb{R}$.

Assume that the function $A_{t_1, \dots, t_k}(x_1, \dots, x_k)$ is a continuous function of its $2k$ arguments, when $(t_1, \dots, t_k) \in I^k \setminus D_k(I)$ and x_1, \dots, x_k are in some neighborhood of u .

Prove that if

$$E \left([N_0(X', I)]^k \right) < \infty$$

then, Rice formula (3.3) for the factorial moments of $N_0(X, I)$ holds true.

EXERCISE 3.8. (a) (So-called ‘‘Banach formula’’). Let f be a continuous function $\mathbb{R} \rightarrow \mathbb{R}$. The total variation $TV(f, I)$ of f over an interval I is defined as

$$TV(f, I) = \sup \sum_{k=0}^{m-1} |f(t_{k+1}) - f(t_k)|$$

where $I = [a, b]$, $a = t_0 < t_1 < \dots < t_m = b$ is a partition of I and the sup is taken over all possible partitions of I .

Prove that

$$(3.30) \quad \int_{-\infty}^{+\infty} N_u(f, I) \, du = TV(f, I).$$

Both sides can be finite or infinite.

(Hint: for each partition $t_0 < t_1 < \dots < t_m$ of the interval I , put $L_k(u) = 1$ if $N_u(f, [t_k, t_{k+1}]) \geq 1$ and $L_k(u) = 0$ otherwise, $k = 0, 1, \dots, m-1$. Show that

$$\int_{-\infty}^{+\infty} \sum_{k=0}^{m-1} L_k(u) \, du = \sum_{k=0}^{m-1} (M_k - m_k)$$

where M_k (resp. m_k) is the maximum (resp. minimum) of the function f on t_k, t_{k+1} . Use this equality to prove (3.30).

(b) Assume furthermore that f is also absolutely continuous. Prove that for any bounded Borel-measurable function $g : \mathbb{R} \rightarrow \mathbb{R}$, one has:

$$(3.31) \quad \int_{-\infty}^{+\infty} N_u(f, I) \, g(u) \, du = \int_I |f'(t)| g(f(t)) \, dt.$$

(c) Prove that if f is absolutely continuous, for every bounded Borel-measurable function $h(t, u)$

$$(3.32) \quad \int_{\mathbb{R}} \sum_{t \in I: f(t)=u} h(t, u) \, du = \int_I |f'(t)| h(t, f(t)) \, dt.$$

(d) Let $\{X(t) : t \in I\}$ be a real-valued stochastic process with absolutely continuous paths, I a compact interval in the real line. Assume that for each $t \in I$, the distribution of the random variable $X(t)$ has a density $p_{X(t)}(\cdot)$. Prove that

$$(3.33) \quad \mathbb{E}(N_u(X, I)) = \int_I \mathbb{E}(|X'(t)| | X(t) = u) \, p_{X(t)}(u) \, dt$$

for almost every $u \in \mathbb{R}$.

(Hint: Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be continuous with compact support. Apply (3.31), replacing f by $X(\cdot)$ and take expectations in both sides).

EXERCISE 3.9. [Upper-bound part of Rice formula] Using Fatou's lemma show that if $\{X_t : t \in I\}$ is a process with C^1 paths and such that:

- (a) for every $t \in I$, X_t admits a density p_{X_t} ,
- (b) the conditional expectation

$$\mathbb{E}(|X'_t| | X_t)$$

is well defined,

- (c) the function

$$\int_0^T \mathbb{E}(X'_t | X_t = x) p_{X_t}(x) \, dt$$

is continuous as a function of x at the point $x = u$. Then

$$\mathbb{E}(N_u(X; I)) \leq \int_I \mathbb{E}(|X'_t| | x_t = u) p_{X_t}(u) \, dt.$$

EXERCISE 3.10. (Rice formulas for χ^2 -processes) Let $\{X(t) : t \in I\}$ be a centered Gaussian process with C^1 paths and values in \mathbb{R}^d . Assume that for each $t \in I$, $\text{Var}(X(t)) = Id$, then the process $Y(t) := \|X(t)\|^2$ is called a χ^2 process. Adapt the proof of Theorem 3.2 to this process.

EXERCISE 3.11. Let $\psi : \mathbb{R} \rightarrow \mathbb{R}$ be a C^1 -function taking non-negative values, with support contained in the interval $[-1, 1]$, such that $\int_{-\infty}^{+\infty} \psi(t) dt = 1$. For $\varepsilon > 0$ we put:

$$\psi_\varepsilon(t) := \frac{1}{\varepsilon} \psi\left(\frac{t}{\varepsilon}\right).$$

If $f : \mathbb{R} \rightarrow \mathbb{R}$ is any locally integrable function, we define the regularized function f^ε :

$$f^\varepsilon(t) := \int_{\mathbb{R}} f(t-s) \psi_\varepsilon(s) ds.$$

Let $\{W(t) : t \in \mathbb{R}\}$ be a Wiener process defined on some probability space (Ω, \mathcal{A}, P) (see Exercise 1.10 for the definition on the hole line).

(a) Prove that for each $t \in \mathbb{R}$, the distribution of the random variable $[W^\varepsilon]'(t)$ is centered normal with variance $\frac{\|\psi\|_2^2}{\varepsilon}$, where $\|\psi\|_2$ is the norm of the function ψ in $L^2(\Omega, \mathcal{A}, P)$.

(b) Prove that for each $u \in \mathbb{R}$, and $I = [t_1, t_2]$, $0 < t_1 < t_2$:

$$(3.34) \quad \lim_{\varepsilon \rightarrow 0} \sqrt{\frac{\pi\varepsilon}{2}} \frac{1}{\|\psi\|_2} N_u(W^\varepsilon, I) = L^W(u, I),$$

where convergence in (3.34) takes place in $L^2(\Omega, \mathcal{A}, P)$.

Hint: use the definition of the local time $L^W(u, I)$ of the Wiener process given in Exercise 1.12 and use Rice Formula to estimate:

$$\mathbb{E}\left(\left[\sqrt{\frac{\pi\varepsilon}{2}} \frac{1}{\|\psi\|_2} N_u(W^\varepsilon, I) - \frac{1}{2\delta} \int_I \mathbf{1}_{|W(t)-u|<\delta} dt\right]^2\right).$$

(c) Prove that convergence in (3.34) holds true in $L^p(\Omega, \mathcal{A}, P)$, for any $p > 0$.

EXERCISE 3.12. Let $\{X(t) : t \in R\}$ be a one-parameter centered Gaussian stationary process with covariance function Γ , $\Gamma(0) = 1$ and finite fourth spectral moment λ_4 (see Chapter 1 for the notation).

We denote $M_{u_1, u_2}(I)$ the number of local maxima of the process in the interval I , with values lying in the interval $[u_1, u_2]$.

a) Use Rice formula to express the expectation

$$\mathbb{E}(M_{u_1, u_2}(I)) = |I| \int_{u_1}^{u_2} f(x) dx$$

where f is a certain function and $|I|$ denotes the length of the interval I .

b) The function

$$g(x) = \frac{f(x)}{\int_{-\infty}^{+\infty} f(x) dx}$$

is called the “density of the values of the local maxima per unit length”.

Give a heuristic interpretation of the function g and verify that, under the above hypotheses, one has:

$$g(x) = \frac{1}{\sqrt{2\pi}} \left[a \exp\left(-\frac{x^2}{2a^2}\right) + \sqrt{1-a^2} x \exp\left(-\frac{x^2}{2}\right) \int_{-\infty}^{\frac{x\sqrt{1-a^2}}{a}} \exp\left(-\frac{u^2}{2}\right) du \right]$$

where

$$a = \sqrt{\frac{\lambda_4 - \lambda_2^2}{\lambda_4}}.$$

c) Conclude from b) that if $a \rightarrow 0$, then

$$g(x) \rightarrow x \exp\left(-\frac{x^2}{2}\right) \text{ if } x > 0, \text{ and } g(x) \rightarrow 0 \text{ if } x < 0.$$

(This is the Rayleigh density of the values of the maxima for "narrow spectrum" Gaussian processes).

Remark. Exercise 3.11 is the beginning of the far-reaching subject of the approximation of local time by means of functionals defined on some smoothing of the paths, which has applications in statistical inference of continuous parameter random processes. In this exercise, only the Wiener process and so-called first order approximations are considered, with the aim of applying Rice formulas in the computations required in the proof. For related work on this subject, the interested reader can see, for example, Azaïs (1989), Wschebor (1992), Berzin and León (2005), Azaïs and Wschebor (1996,1997), Jacod (1998,2000), Perera and Wschebor (1998,2002) and references therein. A review paper on this subject is Wschebor (2006).

Some Statistical Applications

This chapter contains two independent subjects.

In the first part, we use Rice formulas to obtain bounds for the tails of the distribution of the maximum of one-parameter Gaussian processes having regular paths. We also include some results on the asymptotic behavior of the tails of the supremum on a fixed interval, when the level tends to infinity.

The second part is a quite detailed account of two examples of statistical applications of the distribution of the maximum, one to genetics and the other one to the study of mixtures of distributions.

These two examples share the common trait that statistical inference is performed in the presence of nuisance parameters, which are not identifiable under the null hypothesis .

In standard classical situations in hypothesis testing, the critical region consists of the event that some function of the empirical observations (the “test statistic”) falls into some subset of its value space. The computation of the probabilities which are relevant for the test, for large samples, follows from some weak limit theorem allowing to obtain the asymptotic law of the test statistic, as the sample size goes to infinity. Typically, this will be a normal distribution or some function of the normal distribution. In certain more complicated situations, as it happens when nuisance parameters are present and are not identifiable under the null hypothesis, it turns out that a reasonable choice consists in using as test statistic the supremum of a process indexed by the possible values of the nuisance parameter. Then, when passing to the limit as the sample size grows, instead of limiting distributions of finite-dimensional valued random variables, we have to deal with a limit process - typically Gaussian - and the relevant random variable to compute probabilities becomes its supremum.

The literature on this subject has been growing during the last years, including applications to Biology, Econometrics and in general, stochastic models which include nuisance parameters, of which hidden Markov chains have become a quite popular example. The interested reader can see, for example, Andrews and Ploberger, (1994), Hansen (1996), Dacunha-Castelle and Gassiat (1997, 1999), Gassiat (2002), Azaïs, Gassiat and Mercadier (2006,2008) and references therein. The first example in this chapter is extracted from Azaïs and Cierco-Ayrolles (2002) and the second from Delmas (2001, 2003 a).

The theory includes two parts: first, limit theorems allowing to find the asymptotic behavior of certain stochastic processes and second, computing (or obtaining bounds) for the distribution of the supremum of the limiting process or its absolute value. For this second part, a common practice consists in simulating the paths and approximate the tails of the distribution of its supremum using Monte Carlo. This is not what we will be doing here. Our aim is a better understanding of the behavior of this distribution and for that purpose, we will use the results of the first part of this chapter. Since these concern only one-parameter Gaussian processes, the models in the examples of this chapter have one-dimensional nuisance parameter and the asymptotic law of the relevant processes is Gaussian.

1. Elementary bounds for $P\{M > u\}$

In this chapter, if it is not stated otherwise, $\mathcal{X} = \{X(t) : t \in \mathbb{R}\}$ is a real-valued centered Gaussian process with continuously differentiable sample paths and covariance function $r(s, t)$. Let us recall that for $T > 0$, we denote $M_T = \sup_{t \in [0, T]} X(t)$.

We begin with some bounds for the tails of the distribution of M_T which are a first approximation to the relationship between Rice formulas and the distribution of the maximum, a subject that we will frequently come across in this book. These bounds are based on variants of the following simple inequality:

We use throughout the notations: $\nu_k := \mathbb{E}(U_u^{[k]})$, $\tilde{\nu}_k := \mathbb{E}(U_u^{[k]} \mathbf{1}_{X(0) < u})$ for $k = 1, 2, \dots$, where $U_u := U_u^X([0, T])$ is the number of up-crossings on the interval $[0, T]$. Then

$$(4.1) \quad \mathbb{P}\{M_T > u\} \leq \mathbb{P}\{X(0) > u\} + \mathbb{P}\{U_u > 0\} \leq \mathbb{P}\{X(0) > u\} + \nu_1,$$

We precise the bounds under certain hypotheses for the process, in the next proposition.

PROPOSITION 4.1. (a) Assume that $\text{Var}(X(t)) \equiv 1$. Then:

$$(4.2) \quad \mathbb{P}\{M_T > u\} \leq \frac{e^{-u^2/2}}{2\pi} \int_0^T \sqrt{r_{11}(t, t)} dt + 1 - \Phi(u)$$

(b) Assume in addition that the covariance function is of class \mathcal{C}^4 , that $r(s, t) \neq 1$ for $s, t \in [0, T]$, $s \neq t$, and $r_{11}(s, s) > 0$ for all $s \in [0, T]$. Then, if $u > 0$:

$$\mathbb{P}\{M_T > u\} = \frac{e^{-u^2/2}}{2\pi} \int_0^T \sqrt{r_{11}(t, t)} dt + 1 - \Phi(u) + O(\phi(u(1 + \delta)))$$

for some positive real number δ .

Remarks.

1.- The bound (4.2) is sometimes called the Davies bound (Davies, 1977).

2.- Part (b) of Proposition 4.1 was originally proved by Piterbag in 1981 (see also his 1996 book) for centered stationary Gaussian processes.

In this case, under the hypotheses:

- $\lambda_8 < \infty$,
- the joint distribution of $X(s), X'(s), X''(s), X(t), X'(t), X''(t)$ does not degenerate for distinct $s, t \in [0, T]$,
- $\Gamma'(t) < 0$ for $0 < t \leq T$,

Azaïs, Bardet and Wschebor (2001) have been able to describe more precisely the complementary term in the asymptotic expansion given in part (b). The result is that, as $u \rightarrow +\infty$:

$$\mathbb{P}(M_T > u) = 1 - \Phi(u) + \sqrt{\frac{\lambda_2}{2\pi}} T \phi(u) - \frac{3\sqrt{3}(\lambda_4 - \lambda_2^2)^{\frac{9}{2}}}{2\pi\lambda_2^{\frac{9}{2}}(\lambda_2\lambda_6 - \lambda_4^2)} \frac{T}{u^5} \phi\left(\sqrt{\frac{\lambda_4}{\lambda_4 - \lambda_2^2}} u\right) [1 + o(1)].$$

This asymptotic behavior is already in Piterbag's paper of 1981, but only for sufficiently small T . It appears to be the only result known at present giving a precise description of the second order term for the asymptotics of $\mathbb{P}(M_T > u)$ as $u \rightarrow +\infty$. We are not going to prove it here, since it requires quite long and complicated computations. We refer the interested reader to the two above mentioned papers.

PROOF OF PROPOSITION 4.1 : (a) It is clear that:

$$\{M_T > u\} = \{X(0) \geq u\} \cup \{X(0) < u, U_u > 0\} \text{ a.s.}$$

where the convention is that if A and B are two events in a probability space, " $A = B$ a.s." means that $\mathbb{P}\{A \Delta B\} = 0$.

Using Markov's inequality, it follows that

$$\mathbb{P}\{M_T > u\} \leq 1 - \Phi(u) + \mathbb{E}(U_u).$$

Now we apply the Rice formula (Theorem 3.2) to compute $\mathbb{E}(U_u)$:

$$\mathbb{E}(U_u) = \int_0^T \mathbb{E}(X'^+(t) | X(t) = u) p_{X(t)}(u) dt = \phi(u) \int_0^T \mathbb{E}(X'^+(t)) dt = \frac{\phi(u)}{\sqrt{2\pi}} \int_0^T \sqrt{r_{11}(t, t)} dt$$

since $X(t)$ and $X'(t)$ are independent. This proves (a).

For (b), we have the lower bound:

$$(4.3) \quad \begin{aligned} P\{M_T > u\} &= 1 - \Phi(u) + P\{U_u > 0\} - P\{U_u > 0; X(0) > u\} \\ &\geq 1 - \Phi(u) + \nu_1 - \frac{\nu_2}{2} - P\{U_u > 0, X(0) > u\}. \end{aligned}$$

The result will be obtained as soon as we prove that the last two terms in the right-hand side are $O(\phi(u(1 + \delta)))$, $\delta > 0$.

We have:

$$(4.4) \quad \begin{aligned} P\{X(0) > u, U_u > 0\} &\leq P\{X(0) > u, X(T) > u\} + P\{X(0) > u, X(T) < u, U_u > 0\} \\ &\leq P\{X(0) > u, X(T) > u\} + P\{D_u > 1\}. \end{aligned}$$

Remark that since $r(0, T) \neq \pm 1$,

$$\begin{aligned} P\{X(0) > u; X(T) > u\} &\leq P\{X(0) + X(T) > 2u\} \\ &= \int_{2u}^{+\infty} \frac{1}{2\sqrt{\pi}} \frac{1}{\sqrt{1 + r(0, T)}} \exp\left[-\frac{1}{4(1 + r(0, T))} x^2\right] dx = O(\phi(u(1 + \delta))) \end{aligned}$$

Let us look at the second term in the right-hand side of inequality (4.4).

It is clear that:

$$\mathbf{1}_{D_u > 1} \leq \frac{1}{2} D_u (D_u - 1)$$

which implies

$$P\{D_u > 1\} \leq \frac{1}{2} E(D_u(D_u - 1)) = \frac{1}{2} E(U_u(U_u - 1)) = \frac{1}{2} \nu_2,$$

where the penultimate equality follows reverting the time in the interval $[0, T]$, that is, changing t into $T - t$.

So, our aim is to show that $\nu_2 = O(\phi(u(1 + \delta)))$. We have,

$$(4.5) \quad \nu_2 = \int_0^T \int_0^T A_{s,t}^+(u, u) ds dt.$$

where

$$\begin{aligned} A_{s,t}^+(u, u) &:= E(X'^+(s)X'^+(t) | X(s) = X(t) = u) p_{X(s), X(t)}(u, u) \\ p_{X(s), X(t)}(u, u) &= \frac{1}{2\pi} (1 - r^2(s, t))^{-1/2} \exp\left[\frac{-u^2}{1 + r(s, t)}\right]. \end{aligned}$$

Conditionally on $\mathcal{C} := \{X(s) = X(t) = u\}$, the random variables $X'(s), X'(t)$ have a joint Gaussian distribution with expectations and variances given by the following formulas, which can be obtained using regression formulas:

$$\begin{aligned} E(X'(s) | \mathcal{C}) &= \frac{r_{10}(s, t)}{1 + r(s, t)} u, \\ E(X'(t) | \mathcal{C}) &= \frac{r_{10}(t, s)}{1 + r(s, t)} u, \\ \text{Var}(X'(s) | \mathcal{C}) &= r_{11}(s, s) - \frac{r_{10}^2(s, t)}{1 - r^2(s, t)}, \\ \text{Var}(X'(t) | \mathcal{C}) &= r_{11}(t, t) - \frac{r_{10}^2(t, s)}{1 - r^2(s, t)}. \end{aligned}$$

Our hypotheses imply that

$$\begin{aligned} \mathbb{E}(X'^+(s)X'^+(t)|\mathcal{C}) &\leq \frac{1}{2}\mathbb{E}((X'(s))^2 + (X'(t))^2|\mathcal{C}) \\ &\leq \frac{1}{2}(r_{11}(s, s) + r_{11}(t, t)) + \frac{u^2}{2} \frac{r_{10}^2(s, t) + r_{10}^2(t, s)}{[1 + r(s, t)]^2}. \end{aligned}$$

So, for fixed $\gamma > 0$, one has:

$$(4.6) \quad \iint_{s, t \in [0, T]: |s-t| \geq \gamma} A_{s, t}^+(u, u) ds dt \leq (L_1 u^2 + L_2) \iint_{s, t \in [0, T]: |s-t| \geq \gamma} \exp\left[-L_3 \frac{u^2}{2}\right] ds dt$$

where L_1, L_2, L_3 are positive constants, $L_3 > 1$, since $1 + r(s, t) < 2$ for all $s, t \in [0, T], s \neq t$. This shows that

$$\iint_{s, t \in [0, T]: |s-t| \geq \gamma} A_{s, t}^+(u, u) ds dt = O(\phi(u(1 + \delta)))$$

Let us now look at the double integral near the diagonal $\{s, t \in [0, T] : s = t\}$. We take into account that $\text{Var}(X(t))$ is constant, and $r_{10}(t, t) = 0$. A Taylor expansion in the expressions for the conditional expectations as s, t approach the same value t^* permits to show that:

$$(4.7) \quad \mathbb{E}(X'(s)|\mathcal{C}) = [r_{11}(t^*, t^*) + A(t-s)] \frac{(t-s)u}{2}$$

$$(4.8) \quad \mathbb{E}(X'(t)|\mathcal{C}) = [r_{11}(t^*, t^*) + B(t-s)] \frac{(s-t)u}{2}$$

where A, B are bounded functions of the pair s, t .

A similar expansion for the conditional variances shows that:

$$\text{Var}(X'(s)|\mathcal{C}), \text{Var}(X'(t)|\mathcal{C}) \leq L(s-t)^2,$$

for some positive constant L . So,

$$\begin{aligned} (4.9) \quad &\iint_{s, t \in [0, T]: |s-t| < \gamma} A_{s, t}^+(u, u) ds dt \\ &= \iint_{s, t \in [0, T]: |s-t| < \gamma} \mathbb{E}(X'^+(s)X'^+(t)|\mathcal{C}) p_{X(s), X(t)}(u, u) ds dt \\ &\leq \iint_{s, t \in [0, T]: |s-t| < \gamma} \left[\mathbb{E}((X'^+(s))^2|\mathcal{C}) \mathbb{E}((X'^+(t))^2|\mathcal{C}) \right]^{\frac{1}{2}} \frac{1}{2\pi\sqrt{1-r^2(s, t)}} \exp\left[-\frac{u^2}{1+r(s, t)}\right] ds dt. \end{aligned}$$

To bound the conditional expectations in the integrand we use the following inequalities that the reader can easily check. Let Z be a real-valued random variable with normal distribution having parameters μ, σ^2 . Then:

$$(4.10) \quad \mathbb{E}((Z^+)^2) \leq \sigma^2 + \mu^2 \text{ if } \mu > 0,$$

$$(4.11) \quad \mathbb{E}((Z^+)^2) \leq (\sigma^2 + \mu^2) \left[1 - \Phi\left(-\frac{\mu}{\sigma}\right)\right] + \mu\sigma\phi\left(\frac{\mu}{\sigma}\right) \text{ if } \mu < 0.$$

Using the expressions for the conditional expectations, one can see that if $|t-s|$ is sufficiently small, there exists a positive constant D such that

$$\frac{|\mathbb{E}(X'^+(s)|\mathcal{C})|}{\sqrt{\text{Var}(X'^+(s)|\mathcal{C})}} \geq D u.$$

A similar inequality holds for t instead of s .

Now, from the expansions of the conditional expectations, it follows that if $|s-t|$ is small enough, $s \neq t$, $\mathbb{E}(X'(s)|\mathcal{C})$ and $\mathbb{E}(X'(t)|\mathcal{C})$ have opposite signs, so that we can apply to each one of them one of the inequalities (4.10), (4.11). It follows that:

$$\mathbb{E}((X'^+(s))^2|\mathcal{C})\mathbb{E}((X'^+(t))^2|\mathcal{C}) \leq K_1 |s-t|^4 \left[1 - \Phi(-K_2 u) + \exp\left(-\frac{1}{2}K_3 u^2\right)\right].$$

A Taylor expansion around a point in the diagonal shows that if $|s - t|$ is small enough:

$$1 - r(s, t) \geq K_4(s - t)^2$$

for some positive constant K_4 . It follows that if γ is small enough, one has

$$\iint_{s, t \in [0, T]: |s - t| < \gamma} A_{s, t}^+(u, u) ds dt \leq K_5 \exp(-K_6 u^2) \exp\left(-\frac{u^2}{2}\right) \iint_0^T |s - t| ds dt,$$

where K_5, K_6 are new positive constants. It is clear that the right-hand side is $O(\phi(u(1 + \delta)))$, $\delta > 0$. \square

One can obtain the same kind of asymptotic expansion given in Proposition 4.1 in an easier way when the process is also stationary. This is the next proposition; the hypotheses are less demanding and the result for the error term is weaker.

PROPOSITION 4.2 (Stationary processes). *Let $\{X(t) : t \in [0, T]\}$ be a centered stationary Gaussian process. We assume*

G1 $\Gamma(t) = E(X(s)X(s+t)) \neq \pm 1$ for $0 < t \leq T$

G2 the Geman condition : the integral

$$\int \frac{\theta'(\tau)}{\tau^2} d\tau \text{ converges at } \tau = 0^+,$$

where $\theta(\tau)$ is defined by means of: $\Gamma(\tau) = 1 - \frac{\lambda_2 \tau^2}{2} + \theta(\tau)$.

Then, as $u \rightarrow +\infty$

$$(4.12) \quad P\{M_T > u\} = \sqrt{\frac{\lambda_2}{2\pi}} T \phi(u) [1 + o(1)].$$

Remark. Conditions G1 and G2, as already mentioned, are necessary and sufficient for finiteness of the second moment of the crossings (Kratz and León, 2006).

PROOF. : As in the preceding proof, we have

$$(4.13) \quad \nu_2 := E(U_u^{[2]}) = \int_0^T 2(T - \tau) E(X'^+(0)X'^+(\tau) | X(0) = X(\tau) = u) \frac{1}{2\pi} \frac{1}{\sqrt{1 - \Gamma^2(\tau)}} \exp\left[-\frac{u^2}{1 + \Gamma(\tau)}\right] d\tau$$

and

$$E(X'^+(0) | X(0) = X(\tau) = u) = \frac{-\Gamma'(\tau)u}{1 + \Gamma(\tau)} = -E(X'^+(\tau) | X(0) = X(\tau) = u).$$

A standard regression shows that

$$\sigma^2(\tau) := \text{Var}(X'(0) | X(0) = X(\tau) = u) = \text{Var}(X'_\tau / X(0) = X(\tau) = u) = \frac{\lambda_2(1 - \Gamma^2(\tau)) - \Gamma'^2(\tau)}{1 - \Gamma^2(\tau)}.$$

Using inequality $a^+ b^+ \leq (a + b)^2 / 4$, we get

$$(4.14) \quad \nu_2 \leq T \int_0^T \frac{\sigma^2(\tau)}{\sqrt{1 - \Gamma^2(\tau)}} \frac{1}{2\pi} \exp\left[-\frac{u^2}{1 + \Gamma(\tau)}\right] d\tau.$$

Since $\theta(\tau), \theta'(\tau), \theta''(\tau) \geq 0$, an elementary expansion shows that

$$\sigma^2(\tau)(1 - \Gamma^2(\tau)) = \lambda_2(1 - \Gamma^2(\tau)) - \Gamma'^2(\tau) \leq 2\lambda_2\tau\theta'(\tau),$$

$$1 - \Gamma^2(\tau) \approx \lambda_2\tau^2.$$

So that

$$\nu_2 \leq T(\text{const}) \int_0^T \frac{\theta'(\tau)}{\tau^2} \exp\left[-\frac{u^2}{1 + \Gamma(\tau)}\right] d\tau = o(\phi(u)),$$

on account of hypotheses G1 and G2. \square

2. More detailed computation of the first two moments

We return to the lower bound for $P\{M_T > u\}$ for stationary centered Gaussian processes. In what follows, we also assume that the distribution of the triplet $(X(s), X(t), X'(s))$ does not degenerate for $s \neq t$.

We will use in the remaining of this chapter the following inequality, which is a slight modification of (4.3). The proof is immediate and left to the reader.

$$(4.15) \quad P(X(0) > u) + \tilde{\nu}_1 - \frac{\nu_2}{2} \leq P(M_T > u) \leq P(X(0) > u) + \tilde{\nu}_1.$$

Our goal is to give as simple as possible formulas for the quantities involved in (4.15).

Let us introduce or recall some notation that will be used in the rest of this section. We set:

- $\bar{\nu}_1 := \nu_1 - \tilde{\nu}_1$. For large values of u , $\bar{\nu}_1$ and $\tilde{\nu}_1$ are worth being distinguished, since they tend to zero at different exponential rates as $u \rightarrow +\infty$.
- $\mu(t) = E(X'(0)|X(0) = X(t) = u) = -\frac{\Gamma'(t)}{1+\Gamma(t)}u$,
- $\sigma^2(t) = \text{Var}(X'(0)|X(0) = X(t) = u) = \lambda_2 - \frac{\Gamma'^2(t)}{1-\Gamma^2(t)}$,
- $\rho(t) = \text{Cor}(X'(0), X'(t)|X(0) = X(t) = u) = \frac{-\Gamma''(t)(1-\Gamma^2(t)) - \Gamma(t)\Gamma'^2(t)}{\lambda_2(1-\Gamma^2(t)) - \Gamma'^2(t)}$,
- $k(t) = \sqrt{\frac{1+\rho(t)}{1-\rho(t)}}$,
- $b(t) = \frac{\mu(t)}{\sigma(t)}$.

In what follows, the variable t will be omitted whenever there is no confusion, so that we will be writing $\Gamma, \Gamma', \mu, \sigma, \rho, k, b$ instead of $\Gamma(t), \Gamma'(t), \mu(t), \sigma(t), \rho(t), k(t), b(t)$.

LEMMA 4.3. *Let (X, Y) be a random vector in \mathbb{R}^2 having centered normal distribution with variance matrix $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, $|\rho| \neq 1$. Then $\forall a \in \mathbb{R}^+$:*

$$\begin{aligned} P(X > a, Y > -a) &= \frac{1}{\pi} \arctan\left(\sqrt{\frac{1+\rho}{1-\rho}}\right) - 2 \int_0^a \phi(x) \left[\Phi\left(\sqrt{\frac{1+\rho}{1-\rho}}x\right) - \frac{1}{2} \right] dx \\ &= 2 \int_a^{+\infty} \left[\Phi\left(\sqrt{\frac{1+\rho}{1-\rho}}x\right) - \frac{1}{2} \right] \phi(x) dx \end{aligned}$$

PROOF. We first give an integral expression for $P(X > a, Y > a)$.

Put $\rho = \cos \theta$, $\theta \in (0, \pi)$, and use the orthogonal decomposition $Y = \rho X + \sqrt{1-\rho^2}Z$.

Then $\{Y > a\} = \left\{ Z > \frac{a - \rho X}{\sqrt{1-\rho^2}} \right\}$. Thus:

$$P\{X > a, Y > a\} = \int_a^{+\infty} \phi(x) \left[1 - \Phi\left(\frac{a - \rho x}{\sqrt{1-\rho^2}}\right) \right] dx = \iint_{\mathcal{D}} \phi(x)\phi(z) dx dz,$$

where \mathcal{D} is the domain located between the two half straight lines starting from the point $(a, a\sqrt{\frac{1-\rho}{1+\rho}})$ and with angles $\theta - \frac{\pi}{2}$ and $\frac{\pi}{2}$.

Using the symmetry with respect to the straight line with angle $\frac{\theta}{2}$ passing through the origin, we get:

$$(4.16) \quad P\{X > a, Y > a\} = 2 \int_a^{+\infty} \phi(x) \left[1 - \Phi\left(\sqrt{\frac{1-\rho}{1+\rho}}x\right) \right] dx.$$

Now,

$$P\{X > a, Y > -a\} = 1 - \Phi(a) - P\{X > a, (-Y) > a\}.$$

Applying (4.16) to the pair of random variables $(X, -Y)$ yields

$$\begin{aligned} P(X > a, Y > -a) &= 1 - \Phi(a) - 2 \int_a^{+\infty} \phi(x) \left[1 - \Phi \left(\sqrt{\frac{1+\rho}{1-\rho}} x \right) \right] dx \\ &= 2 \int_a^{+\infty} \left[\Phi \left(\sqrt{\frac{1+\rho}{1-\rho}} x \right) - \frac{1}{2} \right] \phi(x) dx. \end{aligned}$$

Now, using polar coordinates, it is easy to establish that

$$\int_0^{+\infty} \left[\Phi(kx) - \frac{1}{2} \right] \phi(x) dx = \frac{1}{2\pi} \arctan(k)$$

which proves the lemma. \square

PROPOSITION 4.4. *Let $\{X(t) : t \in \mathbb{R}\}$ be a centered stationary Gaussian process satisfying the conditions above in this section.*

Then,

$$(i) \quad \bar{\nu}_1 = \phi(u) \int_0^T \left(\sqrt{\frac{\lambda_2}{2\pi}} \left[1 - \Phi \left(\sqrt{\frac{1-\Gamma}{1+\Gamma}} \frac{\sqrt{\lambda_2}}{\sigma} u \right) \right] + \phi \left(\sqrt{\frac{1-\Gamma}{1+\Gamma}} u \right) \left[1 - \Phi(b) \right] \frac{\Gamma'}{\sqrt{1-\Gamma^2}} \right) dt$$

$$(ii) \quad \nu_2 = \int_0^T 2(T-t) \frac{1}{\sqrt{1-\Gamma^2(t)}} \phi^2 \left(\frac{u}{\sqrt{1+\Gamma(t)}} \right) [T_1(t) + T_2(t) + T_3(t)] dt$$

with:

$$T_1(t) = \sigma^2(t) \sqrt{1-\rho^2(t)} \phi(b(t)) \phi(k(t)b(t)),$$

$$T_2(t) = 2(\sigma^2(t)\rho(t) - \mu^2(t)) \int_{b(t)}^{+\infty} \left[\Phi(k(t)x) - \frac{1}{2} \right] \phi(x) dx,$$

$$T_3(t) = 2\mu(t)\sigma(t) \left[\Phi(k(t)b(t)) - \frac{1}{2} \right] \phi(b(t)).$$

(iii) A second expression for $T_2(t)$ is:

$$(4.17) \quad T_2(t) = (\sigma^2(t)\rho(t) - \mu^2(t)) \left[\frac{1}{\pi} \arctan(k(t)) - 2 \int_0^{b(t)} \left[\Phi(k(t)x) - \frac{1}{2} \right] \phi(x) dx \right].$$

Remarks.

- (1) The formula in Lemma 4.3 is analogous to (2.10.4) in Cramér and Leadbetter, (1967), p. 27, that is:

$$P\{X > a, Y > -a\} = \Phi(a)[1 - \Phi(a)] + \int_0^\rho \frac{1}{2\pi\sqrt{1-z^2}} \exp\left(-\frac{a^2}{1-z}\right) dz.$$

The formula here is easier to prove and more adapted to numerical applications because, as $t \rightarrow 0$, $\rho(t) \rightarrow -1$, while the integrand in Cramér and Leadbetter's formula tends to infinity.

- (2) These formulas allow to compute ν_2 at the cost of a double integral with finite limits. This implies a significant reduction of complexity with respect to the original form. The form (4.17) is more adapted to effective computation, because it involves an integral on a bounded interval.

PROOF. OF (i).

Conditionally on $\{X(0) = x, X(t) = u\}$, $X'(t)$ is Gaussian with:

- mean $m(t) = \frac{\Gamma'(t)(x-\Gamma(t)u)}{1-\Gamma^2(t)}$,
- variance $\sigma^2(t)$, see the proof of Proposition 4.2.

If Z is a Gaussian random variable with mean m and variance σ^2 , then

$$\mathbb{E}(Z^+) = \sigma\phi\left(\frac{m}{\sigma}\right) + m\Phi\left(\frac{m}{\sigma}\right).$$

These two remarks yield $\bar{v}_1(u, T) = I_1 + I_2$, with:

- $I_1 = \int_0^T dt \int_u^{+\infty} \sigma\phi\left(\frac{\Gamma'(x - ru)}{(1 - \Gamma^2)\sigma}\right) p_{X(0), X(t)}(x, u) dx$
- $I_2 = \int_0^T dt \int_u^{+\infty} \frac{\Gamma'(x - ru)}{(1 - \Gamma^2)} \Phi\left(\frac{\Gamma'(x - ru)}{(1 - r^2)\sigma}\right) p_{X(0), X(t)}(x, u) dx.$

$$I_1 \text{ can be written as: } I_1 = \phi(u) \int_0^T \frac{\sigma^2}{\sqrt{2\pi\lambda_2}} \left[1 - \Phi\left(\frac{\sqrt{\lambda_2}}{\sigma} \sqrt{\frac{1 - \Gamma}{1 + \Gamma}} u\right)\right] dt.$$

Integrating by parts in I_2 leads to

$$I_2 = \phi(u) \int_0^T \frac{\Gamma'}{\sqrt{1 - \Gamma^2}} \phi\left(\sqrt{\frac{1 - \Gamma}{1 + \Gamma}} u\right) [1 - \Phi(b)] + \frac{\Gamma'^2}{\sqrt{2\pi\lambda_2}(1 - \Gamma^2)} \left[1 - \Phi\left(\frac{\sqrt{\lambda_2}}{\sigma} \sqrt{\frac{1 - \Gamma}{1 + \Gamma}} u\right)\right] dt.$$

Since $\sigma^2 + \Gamma'^2/(1 - \Gamma^2) = \lambda_2$, we obtain:

$$\bar{v}_1 = \sqrt{\frac{\lambda_2}{2\pi}} \phi(u) \int_0^T \left[1 - \Phi\left(\frac{\sqrt{\lambda_2}}{\sigma} \sqrt{\frac{1 - \Gamma}{1 + \Gamma}} u\right)\right] dt + \phi(u) \int_0^T \frac{\Gamma'}{\sqrt{1 - \Gamma^2}} \phi\left(\sqrt{\frac{1 - \Gamma}{1 + \Gamma}} u\right) [1 - \Phi(b)] dt.$$

PROOF OF (ii).

We set:

- $v(x, y) = \frac{(x-b)^2 - 2\rho(x-b)(y+b) + (y+b)^2}{2(1 - \rho^2)}$
- for $(i, j) \in \{(0, 0); (1, 0); (0, 1); (1, 1); (2, 0); (0, 2)\}$

$$J_{ij} = \int_0^{+\infty} \int_0^{+\infty} \frac{x^i y^j}{2\pi\sqrt{1 - \rho^2}} \exp(-v(x, y)) dy dx.$$

We have

$$(4.18) \quad J_{10} - \rho J_{01} - (1 + \rho)bJ_{00} = (1 - \rho^2) \int_0^{+\infty} \left(\int_0^{+\infty} \frac{\partial}{\partial x} v(x, y) \frac{\exp(-v(x, y))}{2\pi\sqrt{1 - \rho^2}} dx \right) dy \\ = (1 - \rho^2) [1 - \Phi(kb)] \phi(b).$$

Replacing x by y and b by $-b$ in (4.18) yields

$$(4.19) \quad J_{01} - \rho J_{10} + (1 + \rho)bJ_{00} = (1 - \rho^2) \Phi(kb) \phi(b).$$

In the same way, multiplying the integrand by y , we get

$$(4.20) \quad J_{11} - \rho J_{02} - (1 + \rho)bJ_{01} = (1 - \rho^2)^{3/2} [\phi(kb) - kb[1 - \Phi(kb)]] \phi(b).$$

Now, multiplying the integrand by x leads to

$$(4.21) \quad J_{11} - \rho J_{20} + (1 + \rho)bJ_{10} = (1 - \rho^2)^{3/2} [\phi(kb) + kb\Phi(kb)] \phi(b).$$

So,

$$J_{20} - \rho J_{11} - (1 + \rho)bJ_{10} = (1 - \rho^2) \int_0^{+\infty} \int_0^{+\infty} x \frac{\partial}{\partial x} v(x, y) \frac{\exp(-v(x, y))}{2\pi\sqrt{1 - \rho^2}} dx dy.$$

Then, integrating by parts

$$(4.22) \quad J_{20} - \rho J_{11} - (1 + \rho)bJ_{10} = (1 - \rho^2) J_{00}.$$

Multiplying equation (4.22) by ρ and adding (4.21) gives:

$$J_{11} = -bJ_{10} + \rho J_{00} + \sqrt{1 - \rho^2} [\phi(kb) + kb\Phi(kb)] \phi(b).$$

Multiplying equation (4.19) by ρ and adding equation (4.18) yields:

$$J_{10} = bJ_{00} + [1 - \Phi(kb) + \rho\Phi(kb)] \phi(b).$$

Using the previous lemma, $J_{00} = 2 \int_b^{+\infty} [\Phi(kx) - \frac{1}{2}] \phi(x) dx$.

Gathering the various pieces:

$$J_{11} = J_{11}(b, \rho) = \sqrt{1 - \rho^2} \phi^2 \left(\frac{b}{\sqrt{1 - \rho}} \right) \phi(b) + 2(\rho - b^2) \int_b^{+\infty} [\Phi(kx) - \frac{1}{2}] \phi(x) dx + 2b[\Phi(kb) - \frac{1}{2}] \phi(b).$$

The final result is obtained taking into account that

$$\mathbb{E} \left((X'_0)^+ (X'_t)^+ | X_0 = X_t = u \right) = \sigma^2(t) J_{11}(b(t), \rho(t)).$$

PROOF OF (iii)

Expression (4.17) follows from the second expression of J_{00} . \square

The numerical computation of ν_2 has some difficulties, related to the behavior of the integrand near the diagonal. Since $\nu_2 = \iint_0^T A_{s,t}^+(u, u) ds dt$, one can use the next proposition to describe the function $A_{s,t}^+(u, u)$ when $|t - s|$ is small. For the proof, that we are not going to give here, one can use Maple to compute its Taylor expansion as a function of $t - s$ in a neighborhood of 0.

PROPOSITION 4.5 (Azaïs Cierco-Ayrolles and Croquette (1999)). *Let the Gaussian stationary process $\{X(t) : t \in \mathbb{R}\}$ satisfy the above hypotheses in this section. Assume moreover that λ_8 is finite.*

(a) As $t \rightarrow 0$:

$$A_{0,t}^+(u, u) = \frac{1}{1296} \frac{(\lambda_2 \lambda_6 - \lambda_4)^{3/2}}{(\lambda_4 - \lambda_2^2)^{1/2} \pi^2 \lambda_2^2} \exp \left(-\frac{1}{2} \frac{\lambda_4}{\lambda_4 - \lambda_2^2} u^2 \right) t^4 + O(t^5).$$

(b) There exists $T_0 > 0$ such that, for every $T, 0 < T < T_0$

$$\begin{aligned} \bar{\nu}_1 &= \frac{27}{4\sqrt{\pi}} \frac{(\lambda_4 - \lambda_2^2)^{11/2}}{\lambda_2^5 (\lambda_2 \lambda_6 - \lambda_4)^{3/2}} \phi \left(\sqrt{\frac{\lambda_4}{\lambda_4 - \lambda_2^2}} u \right) u^{-6} \left(1 + O\left(\frac{1}{u}\right) \right) \\ \nu_2 &= \frac{3\sqrt{3}T}{\pi} \frac{(\lambda_4 - \lambda_2^2)^{9/2}}{\lambda_2^{9/2} (\lambda_2 \lambda_6 - \lambda_4^2)} \phi \left(\sqrt{\frac{\lambda_4}{\lambda_4 - \lambda_2^2}} u \right) u^{-5} \left(1 + O\left(\frac{1}{u}\right) \right) \end{aligned}$$

as $u \rightarrow +\infty$.

We will consider the same question on the behavior of the integrand $A_{s,t}^+(u, u)$ near the diagonal for non-stationary Gaussian processes in Chapter 5.

3. Maximum of the absolute value

We denote $M_T^* = \sup_{t \in [0, T]} |X(t)|$. The following inequality for $\mathbb{P}\{M_T^* > u\}$ is elementary, the proof being left to the reader:

$$\begin{aligned} (4.23) \quad \mathbb{P}(|X(0)| > u) + \mathbb{E}(U_u \mathbf{1}_{|X(0)| < u}) + \mathbb{E}(D_{-u} \mathbf{1}_{|X(0)| < u}) - \frac{1}{2} \mathbb{E}((U_u + D_{-u})^2) \\ \leq \mathbb{P}(M_T^* > u) \leq \mathbb{P}(|X(0)| > u) + \mathbb{E}(U_u \mathbf{1}_{|X(0)| < u}) + \mathbb{E}(D_{-u} \mathbf{1}_{|X(0)| < u}) \end{aligned}$$

Delmas (2001) has given tables with formulas to compute the terms appearing in (4.15) and (4.23), as well as extensions to non-stationary processes. We refer to it for a comprehensive list of useful formulas. Here, we only describe the result for the distribution of M_T^* in the case of centered stationary Gaussian processes, normalized by $\Gamma(0) = 1$. The terms are included in the following table.

For $u > 0$

$$\mathbb{P}(|X(0)| > u) = 2[1 - \Phi(u)]; \quad U_u \stackrel{D}{=} D_{-u},$$

and for $b > 0$, we use the already defined notations:

$$\sigma^2 J_{11}(b, \rho) := T_1(b, \rho, \sigma^2) + T_2(b, \rho, \sigma^2) + T_3(b, \rho, \sigma^2) \text{ with:}$$

$$\begin{aligned}
T_1(b, \rho, \sigma^2) &:= \sigma^2 \sqrt{1 - \rho^2} \phi(b) \phi(kb) \\
T_2(b, \rho, \sigma^2) &:= (\sigma^2 \rho - \sigma^2 b^2) \left[\frac{1}{\pi} \arctan(k) - 2 \int_0^b \phi(x) \Psi(kx) dx \right] \\
T_3(b, \rho, \sigma^2) &:= 2b\sigma^2 \left[\Phi(kb) - \frac{1}{2} \right] \phi(b).
\end{aligned}$$

Then:

Centered stationary case	
$E[U_u^X[0, T] \mathbf{1}_{\{X(0) > u\}}]$	$F(\Gamma, \Gamma') := \phi(u) \int_0^T \frac{\sqrt{\lambda_2}}{\sqrt{2\pi}} \left(1 - \Phi \left[\frac{u}{\sigma} \sqrt{\lambda_2} \sqrt{\frac{1-\Gamma}{1+\Gamma}} \right] \right) + \frac{\Gamma'}{\sqrt{1-\Gamma^2}} \phi \left(u \sqrt{\frac{1-\Gamma}{1+\Gamma}} \right) \left(1 - \Phi \left[-\frac{u\Gamma'}{\sigma(1+\Gamma)} \right] \right) dt$
$E[U_u^X[0, T] \mathbf{1}_{\{X(0) \leq u\}}]$	$\frac{T\sqrt{\lambda_2}}{2\pi} \exp\left(-\frac{u^2}{2}\right) - F(\Gamma, \Gamma')$
$E[U_u^X[0, T] \mathbf{1}_{\{ X(0) \leq u\}}]$	$\frac{T\sqrt{\lambda_2}}{2\pi} \exp\left(-\frac{u^2}{2}\right) - F(\Gamma, \Gamma') - F(-\Gamma, -\Gamma')$
$E[U_u^X[0, T]^2]$	$2 \int_0^T (T-t) \sigma^2 J_{11} \left(-\frac{\Gamma' u}{\sigma(1+\Gamma)}, \rho \right) \frac{\phi^2\left(\frac{u}{\sqrt{1+\Gamma}}\right)}{\sqrt{1-\Gamma^2}} dt$
$E[U_u^X[0, T] D_{-u}^X[0, T]]$	$2 \int_0^T (T-t) \sigma^2 J_{11} \left(-\frac{\Gamma' u}{\sigma(1-\Gamma)}, -\rho \right) \frac{\phi^2\left(\frac{u}{\sqrt{1-\Gamma}}\right)}{\sqrt{1-\Gamma^2}} dt$

4. Application to quantitative gene detection

We study a back-cross population: $A \times (A \times B)$, where A and B are purely homozygous lines and we address the problem of detecting a gene influencing some quantitative trait (i.e. which is able to be measured) on a given chromosome. The trait is observed on n individuals and we denote by Y_k , $k = 1, \dots, n$ the observations, which we will assume to be independent.

The mechanism of genetics, or more precisely of meiosis, implies that among the two chromosomes of each individual, one is purely inherited from A while the other (the “recombined” one) consists of parts originated from A and parts originated from B , due to crossing-overs. Using the Haldane (1919) distance and modelling, each chromosome will be represented by a segment $[0, L]$. The distance on $[0, L]$ is called the genetic distance (which is measured in Morgans).

Now, let us describe the mathematical model that we will be using. We assume that the “recombined” chromosome starts at the left-end point $t = 0$ with probability $1/2$ with some part originated from A or B and then switches from one value to another at every location of a crossing-over. We model the crossing-over points by a standard Poisson process, independent of the fact that the chromosome starts with A or B .

The influence of the putative gene (often called QTL for Quantitative Trait Locus by geneticists) on the quantitative trait is represented by a classical linear model:

$$(4.24) \quad Y_k = \mu + G_k(t_0) a/2 + \varepsilon_k \quad (k = 1, \dots, n),$$

where

- μ is the general mean.
- t_0 is the location of the gene on the chromosome ,

- $G_k(t)$ is the genotypic composition of the individual k at location t on the chromosome, $t \in [0, L]$. In a back-cross crossing scheme it can only take two values (AB or AA) that are denoted +1 and -1.
- ε_k , $k = 1, \dots, n$ are independent errors that are assumed to be independent, identically distributed with zero mean and finite fourth moment. We denote by σ^2 their common variance.

Formula (4.24) implies that the gene effect is a . Our problem is to test the null hypothesis of absence of gene influence, that is, $a = 0$.

The main statistical difficulties come from the following facts:

- The putative location t_0 of the gene is unknown;
- The functions $G_k(t)$ can not be fully observed. However, some “genetic markers” are available, allowing to know the genetic composition at some fixed locations t_1, \dots, t_M on $[0, L]$.

Since the number of genetic markers available on a given species becomes larger and larger and since considering the limit model where the number of markers tends to infinity permits to construct a test which is free from the markers positions, we will assume that the number M_n of markers as well as their locations depend on the number of observations n . More precisely, we use now a local asymptotic framework in which:

- The number n of observed individuals tends to infinity;
- The number of genetic markers M_n tends to infinity with n , their locations being denoted by $t_{i,n}$; $i = 1, \dots, M_n$;
- The size a of the gene effect is small, satisfying the contiguity condition $a = \delta n^{-1/2}$, where δ is some constant;
- The observation is

$$\{(Y_k^n, G_k(t_{1,n}), \dots, G_k(t_{M_n,n})), k = 1, \dots, n\}.$$

Notice that since we have set $a = \delta n^{-1/2}$, the distribution of the observation depends on n so that the observations are denoted by Y_k^n instead of Y_k .

If the true position t_0 of the gene were known and would coincide with a marker location, a natural test would be to make a “comparison of means”, computing the statistics :

$$(4.25) \quad S_n(t_0) := \frac{\sum_{k=1}^n Y_k^n \mathbf{1}_{[G_k(t_0)=1]}}{\sum_{k=1}^n \mathbf{1}_{[G_k(t_0)=1]}} - \frac{\sum_{k=1}^n Y_k^n \mathbf{1}_{[G_k(t_0)=-1]}}{\sum_{k=1}^n \mathbf{1}_{[G_k(t_0)=-1]}}$$

In case of Gaussian observations, the test based on this statistics is equivalent to the likelihood ratio test.

Since t_0 is unknown, the calculation of the quantity given by formula (4.25) should be performed at each location $t \in [0, L]$. To do this, we compute S_n at the points $t_{1,n}, \dots, t_{M_n,n}$, make a linear interpolation between two consecutive marker positions, and extend by a constant to the left of the first marker position and similarly to the right of the last one, so that the function thus obtained is continuous. We denote by $\{S_n(t) : t \in [0, L]\}$ the random process obtained in this form, which has continuous polygonal paths.

Our aim is to find a normalization of the process S_n in order to get weak convergence. Before that, let us make a short parenthesis to give an overview on this subject, which includes some general and useful results.

4.1. Quick reminder on weak convergence of stochastic processes. Let us consider real-valued processes defined on $[0, 1]$ (extensions to the whole line do not present essential difficulties). We need the following function spaces:

- $C = C([0, 1])$ the space of continuous functions $[0, 1] \rightarrow \mathbb{R}$, equipped with the topology of uniform convergence generated by the distance:

$$d_u(f, g) = \|f - g\|_\infty := \sup_{t \in [0, 1]} |f(t) - g(t)|$$

- $D = D([0, 1])$ the set of càd-làg functions (functions which are right-continuous and have left-hand limits at each point), equipped with the Skorohod topology generated by the distance

$$d_s(f, g) := \inf_h \{ \sup \{ \|h - Id\|_\infty, \|f - g \circ h\|_\infty \} \},$$

where the infimum is taken over all strictly increasing continuous mappings $h : [0, 1] \rightarrow [0, 1]$, $h(0) = 0$, $h(1) = 1$, Id is the identity mapping, i.e. $Id(x) = x$ for all $x \in [0, 1]$ and “ \circ ” denotes composition.

C is a Polish space (this means, a complete separable metric space). D is not complete, but with a natural change of the metric, which does not modify the topology, it becomes complete (see Billingsley, 1999 for details). A possible modification consists in replacing d_s by d_s^0 defined by:

$$d_s^0(f, g) := \inf_h \{ \sup \{ \|h\|^\circ, \|f - g \circ h\|_\infty \} \},$$

where the inf is over the same class of functions as above, and

$$\|h\|^\circ = \sup_{s < t} \left| \log \frac{h(t) - h(s)}{t - s} \right|.$$

We will usually denote, without further reference, with the same letters C and D the analogous spaces of functions defined on some interval $[t_1, t_2]$ other than $[0, 1]$.

Next, we give two definitions that are basic in what follows: weak convergence and tightness.

DEFINITION 4.6. *Let E be a Polish space and \mathcal{E} the σ -algebra of subsets of E generated by the open sets. The sequence of probability measures $\{P_n\}_{n=1,2,\dots}$ on (E, \mathcal{E}) is said to converge weakly to the measure P defined on the same measurable space (this is denoted by $P_n \Rightarrow P$ as $n \rightarrow +\infty$) if for every continuous and bounded function $f : E \rightarrow \mathbb{R}$,*

$$\int_E f dP_n \rightarrow \int_E f dP \text{ as } n \rightarrow +\infty.$$

DEFINITION 4.7. *Let E be a Polish space. A collection of probability measures \mathcal{F} on (E, \mathcal{E}) is said to be tight if for any $\varepsilon > 0$ there exists K_ε , a compact subset of E , such that for all $P \in \mathcal{F} : P(K_\varepsilon^c) \leq \varepsilon$.*

Let (Ω, \mathcal{A}, P) be a probability space, E a separable metric space and $Y : \Omega \rightarrow E$ a random variable, that is $Y^{-1}(F) \in \mathcal{A}$ for any choice of $F \in \mathcal{E}$. The image measure P^Y on (E, \mathcal{E}) , defined as

$$P^Y(F) = P[Y^{-1}(F)]$$

for every $F \in \mathcal{E}$ is called the probability distribution of the random variable Y .

A sequence of random variables $\{Y_n\}_{n=1,2,\dots}$ with values in E is said to converge weakly to the random variable Y , if

$$P^{Y_n} \Rightarrow P^Y \text{ as } n \rightarrow +\infty.$$

that is, if the sequence of distributions converges weakly to the distribution of Y . In that case, we will write $Y_n \Rightarrow Y$.

The next three theorems contain the results that we will be using in the statistical applications of this chapter. For proofs, examples and a general understanding of the subject, the reader is referred to Billingsley’s classical book (1999).

THEOREM 4.8 (Prohorov). *Let E is a Polish space. From every tight sequence of probability measures on (E, \mathcal{E}) one can extract a weakly convergent subsequence. If a sequence of probability measures on (E, \mathcal{E}) is tight then, it is weakly convergent if and only if all weakly convergent subsequences have the same limit.*

THEOREM 4.9. *Let $\{X_n\}_{n=1,2,\dots}$ be a sequence of random variables with values in the spaces C or D . $X_n(t)$ denotes the value of X_n at $t \in [0, 1]$.*

Then, $X_n \Rightarrow X$ as $n \rightarrow +\infty$, if and only if,

- $\{X_n\}_{n=1,2,\dots}$ is tight (in the corresponding space C or D) and

- for any choice of $k = 1, 2, \dots$, and distinct $t_1, \dots, t_k \in [0, 1]$, the random vector $(X_n(t_1), \dots, X_n(t_k))$ converges weakly to $(X(t_1), \dots, X(t_k))$ in \mathbb{R}^k , as $n \rightarrow +\infty$.

In case convergence takes place, if $g : C \rightarrow \mathbb{R}$ (respectively $g : D \rightarrow \mathbb{R}$) is a continuous function with respect to the topology of C (resp. D), then the sequence of real-valued random variables $\{g(X_n)\}_{n=1,2,\dots}$ converges in distribution to the random variable $g(X)$.

Moreover, if $X_n \Rightarrow X$ in D , and X has continuous paths, then weak convergence holds true in C (see Exercise 4.3).

The above statement contains the standard procedure, in a large set of statistical problems, to prove the existence of weak limits of stochastic processes and compute the limit distributions whenever they exist. One has to check tightness and finite-dimensional weak convergence. If possible, one also wants to identify the limit measure (the distribution of X) and if the ‘‘observable’’ quantity in which one is interested in is $g(X_n)$, this also allows to find its limit distribution.

The next theorem gives a sufficient condition based upon upper-bounds for moments of increments, to verify tightness and prove weak convergence.

THEOREM 4.10. *Let $\{X_n(t) : t \in [0, 1]\}$, $n = 1, 2, \dots$ be a sequence of random processes and $\{X(t) : t \in [0, 1]$ a process with sample paths in C (resp. D) satisfying*

- (1) *for any choice of $k = 1, 2, \dots$, and distinct $t_1, \dots, t_k \in [0, 1]$, the sequence of random vectors $(X_n(t_1), \dots, X_n(t_k))$ converges weakly to $(X(t_1), \dots, X(t_k))$ in \mathbb{R}^k , as $n \rightarrow +\infty$.*
- (2) *-if the sample paths are in C : there exist three positive constants, α, β, γ such that for all $s, t \in [0, 1]$:*

$$\mathbb{E}|X_n(s) - X_n(t)|^\alpha \leq \beta|s - t|^{1+\gamma}$$

-if the sample paths are in D : there exist three positive constants, α, β, γ such that for all $t_1, t, t_2 \in [0, 1]$, $t_1 \leq t \leq t_2$:

$$(4.26) \quad \mathbb{E}\left(|X_n(t_1) - X_n(t)|^\alpha |X_n(t_2) - X_n(t)|^\alpha\right) \leq \beta|t_2 - t_1|^{1+\gamma}.$$

Then, $X_n \Rightarrow X$ as $n \rightarrow +\infty$ in C (resp. in D).

The property (2) in addition to the tightness of the family of probability distributions $X_n(0)$ is a sufficient condition for the tightness of the sequence $\{X_n\}_{n=1,2,\dots}$.

We will use also the *Skorohod embedding* (see for example Dudley 1989 p. 324) that states that if X_n and X are random variables taking values in a separable metric space (E, d) and if X_n converges weakly to X , then we can construct a representation of these variables, say $\{Y_n\}_{n=1,2,\dots}, Y$ defined on some new probability space with values in E , so that X_n and Y_n have the same distribution for every $n = 1, 2, \dots$ (as well as Y and X) and:

$$d(Y_n, Y) \rightarrow 0 \text{ almost surely as } n \rightarrow +\infty.$$

4.2. Weak convergence of the detection process. We go back to the genetic model.

THEOREM 4.11. *a) $\text{Var}(S_n(t)) \approx \frac{4\sigma^2}{n}$*

b) Since $\text{Var}(Y_k^n) \approx \sigma^2$ we can estimate the parameter σ^2 by the empirical variance: $\hat{\sigma}_n^2$. We define the normalized process

$$(4.27) \quad X_n(t) := \frac{\sqrt{n}}{2\hat{\sigma}_n} S_n(t).$$

Assume that

$$\max_{i=0,\dots,M_n} (t_{i+1,n} - t_{i,n}) \rightarrow 0 \text{ as } n \rightarrow +\infty$$

where we have used the convention: $t_{0,n} = 0$; $t_{M_n+1,n} = L$.

Then, the normalized process $X_n(t)$ converges weakly in the space of continuous functions on $[0, L]$ to a Gaussian process $\mathcal{X} = \{X(t) : t \in [0, L]\}$ with:

$$\begin{aligned} \mathbb{E}(X(t)) &= \frac{\delta}{2\sigma} \exp(-2|t_0 - t|) \\ (4.28) \quad \text{Cov}(X(t), X(t+h)) &= \exp(-2|h|). \end{aligned}$$

\mathcal{X} is an Ornstein-Uhlenbeck process with a change of scale and a deterministic drift (see Exercise 4.6 at the end of this chapter).

The proof of Theorem 4.11 is based on the following lemma.

LEMMA 4.12. Let $\{\eta_k\}_{k=1,2,\dots}$ be centered random variables with common variance σ^2 and finite fourth order moment μ_4 . Suppose that the collection of random variables and processes $\eta_1, G_1(\cdot), \dots, \eta_n, G_n(\cdot), \dots$ are independent. Here $G_k(t), t \in [0, L]$ is the genotypic composition of the k -th individual, which follows the switching Poisson model that has been described above. Then, as $n \rightarrow +\infty$:

(a) the processes

$$Z_n(t) := n^{-1/2} \sum_{k=1}^n \eta_k G_k(t)$$

converge weakly in the space D to a stationary Gaussian process with zero mean and covariance $h \rightsquigarrow \sigma^2 \exp(-2|h|)$.

(b) The processes

$$\tilde{Z}_n(t) := n^{-1} \sum_{k=1}^n G_k(t_0) G_k(t)$$

converge uniformly (for $t \in [0, L]$) in probability to the function $\exp(-2|t_0 - t|)$.

PROOF. (a) Independence implies that

$$\mathbb{E}(\eta_k G_k(t)) = 0.$$

If Z is a random variable having the Poisson distribution with parameter λ , one easily checks that:

$$\mathbb{P}(Z \text{ even}) - \mathbb{P}(Z \text{ odd}) = \exp(-2\lambda).$$

This implies that

$$\mathbb{E}(G_k(t) G_k(t')) = \exp(-2|t - t'|), \quad \mathbb{E}(\eta_k^2 G_k(t) G_k(t')) = \sigma^2 \exp(-2|t - t'|).$$

To prove part (a) of the lemma, we apply Theorem 4.10. The convergence of the finite-dimensional distributions follows from a standard application of the multivariate central limit theorem and is left to the reader.

We now prove the moment condition (4.26). Let $t_1 < t < t_2$ be in $[0, L]$:

$$\begin{aligned} (4.29) \quad \mathbb{E}[(Z_n(t) - Z_n(t_1))^2 (Z_n(t) - Z_n(t_2))^2] &= \frac{1}{n^2} \sum_{1 \leq k_1, k_2, k_3, k_4 \leq n} \\ &\mathbb{E}\left(\eta_{k_1} \dots \eta_{k_4} (G_{k_1}(t) - G_{k_1}(t_1)) (G_{k_2}(t) - G_{k_2}(t_1)) (G_{k_3}(t) - G_{k_3}(t_2)) (G_{k_4}(t) - G_{k_4}(t_2))\right). \end{aligned}$$

The independence implies that, as soon as one index k_i is different from the other three, the expectation in the corresponding term of the sum vanishes. Hence,

$$\begin{aligned}
(4.30) \quad & \mathbb{E}[(Z_n(t) - Z_n(t_1))^2 (Z_n(t) - Z_n(t_2))^2] \\
&= \frac{1}{n^2} \sum_{1 \leq k_1 \neq k_2 \leq n} \mathbb{E}(\eta_{k_1}^2 \eta_{k_2}^2 (G_{k_1}(t) - G_{k_1}(t_1))^2 (G_{k_2}(t) - G_{k_2}(t_2))^2) \\
&+ \frac{1}{n^2} \sum_{1 \leq k_1 \neq k_2 \leq n} \mathbb{E}(\eta_{k_1}^2 \eta_{k_2}^2 (G_{k_1}(t) - G_{k_1}(t_1))(G_{k_1}(t) - G_{k_1}(t_2))(G_{k_2}(t) - G_{k_2}(t_1))(G_{k_2}(t) - G_{k_2}(t_2))) \\
&\quad + \frac{1}{n^2} \sum_{1 \leq k \leq n} \mathbb{E}(\eta_k^4 (G_k(t) - G_k(t_1))^2 (G_k(t) - G_k(t_2))^2) \\
&\quad \leq (\text{const})\sigma^4 |t - t_1| |t - t_2| + (\text{const})\mu_4 |t - t_1| |t - t_2| \leq (\text{const})(t_2 - t_1)^2.
\end{aligned}$$

where we have used that:

- almost surely, $(G_k(t) - G_k(t_1))(G_k(t) - G_k(t_2))$ vanishes, except if the number of occurrences in intervals $[t_1, t]$ and $[t, t_2]$ are both odd. In that case it takes the value -4. In conclusion its expectation is non-positive
- $(G_k(t) - G_k(t_2))^2$ depends only on the parity of the number of occurrences of the Poisson process on $[t, t_2]$, and is independent of $(G_k(t) - G_k(t_1))$.

$$\mathbb{E}(G_k(t) - G_k(t'))^2 = 2(1 - \exp(-2|t - t'|)) \leq (\text{const})|t - t'|.$$

The inequality (4.30) implies (4.26). This proves (a).

As for assertion (b), we write

$$\begin{aligned}
(4.31) \quad \tilde{Z}_n(t) &:= n^{-1} \sum_{k=1}^n G_k(t_0) G_k(t) \\
&= n^{-1} \sum_{k=1}^n \left(G_k(t_0) G_k(t) - \exp(-2|t - t_0|) \right) + \exp(-2|t - t_0|) \\
&= n^{-1} \sum_{k=1}^n T_k(t) + \exp(-2|t - t_0|) = \tilde{Z}_{n,1}(t) + \exp(-2|t - t_0|).
\end{aligned}$$

(b) follows from the result in Exercise 4.4, if we can prove that $\tilde{Z}_{n,1}$ tends weakly to zero in the space D as $n \rightarrow +\infty$. Because of the strong law of large numbers, for each t , $\tilde{Z}_{n,1}(t)$ tends to zero almost surely. This obviously implies that a.s. the k -tuple $(\tilde{Z}_{n,1}(t_1), \dots, \tilde{Z}_{n,1}(t_k))$ converges to $(0, \dots, 0)$ as $n \rightarrow +\infty$.

So, to apply Theorem 4.10 it suffices to check (4.26).

This can be done much in the same way we did with formula (4.30) except that: (i) the normalizing constant is now n^{-4} , (ii) $G_k(t)$ is replaced by $T_k(t)$ (iii) the variables η_k are absent. To conclude we have to check that:

- $\mathbb{E}(T_k(t) - T_k(t'))^2 \leq \mathbb{E}(G_k(t_0)(G_k(t) - G_k(t'))^2) = \mathbb{E}(G_k(t) - G_k(t'))^2 \leq (\text{const})|t - t'|$
- $\mathbb{E}((T_k(t) - T_k(t_1))(T_k(t) - T_k(t_2))) = \text{Cov}(G_k(t_0)(G_k(t) - G_k(t_1)), G_k(t_0)(G_k(t) - G_k(t_2)))$
 $= \mathbb{E}((G_k(t) - G_k(t_1))(G_k(t) - G_k(t_2))) - (e^{-2|t_0-t|} - e^{-2|t_0-t_1|})(e^{-2|t_0-t|} - e^{-2|t_0-t_2|})$
 $\leq -(e^{-2|t_0-t|} - e^{-2|t_0-t_1|})(e^{-2|t_0-t|} - e^{-2|t_0-t_2|}) \leq (\text{const})(t_1 - t_2)^2$
- If Z and T are two random variables :

$$\mathbb{E}[(Z - \mathbb{E}(Z))^2 (T - \mathbb{E}(T))^2] \leq 4\mathbb{E}(Z^2 T^2) + 12\mathbb{E}(Z^2)\mathbb{E}(T^2).$$

Applying this with $Z - E(Z) = T_k(t) - T_k(t_1)$; $T - E(T) = T_k(t) - T_k(t_2)$ we get

$$\begin{aligned} E[(T_k(t) - T_k(t_1))^2 (T_k(t) - T_k(t_2))^2] &\leq 16 E(G_k(t) - G_k(t_1))^2 E(G_k(t) - G_k(t_2))^2 \\ &\leq (\text{const})(t_1 - t_2)^2. \end{aligned}$$

Summing up:

$$E[(\tilde{Z}_{n,1}(t_1) - \tilde{Z}_{n,1}(t))^2 (\tilde{Z}_{n,1}(t_2) - \tilde{Z}_{n,1}(t))^2] \leq (\text{const})(t_1 - t_2)^2 n^{-2}$$

□

PROOF OF THEOREM 4.11:

Step 1: Convergence in case of full genetic information. In a first step, we assume that the genetic information $G_k(t)$ is available at every location of the chromosome, so that no interpolation is needed.

We define several auxiliary random processes:

$$X_{n,1}(t) := \frac{\sqrt{n}}{2\sigma} S_n(t).$$

It is clear that if the process $\{X_{n,1}(t) : t \in [0, L]\}$ converges weakly, the same holds true for the process $\{X_n(t) : t \in [0, L]\}$, with the same limit.

$$X_{n,2}(t) := \frac{1}{\hat{\sigma}_n \sqrt{n}} \sum_{k=1}^n (Y_k^n - \bar{Y}_n) G_k(t).$$

which can actually be computed (\bar{Y}_n is the mean of the sample), and

$$X_{n,3}(t) := \frac{1}{\sigma \sqrt{n}} \sum_{k=1}^n (Y_k^n - \mu) G_k(t).$$

which cannot be actually computed.

The convergence in probability of $\hat{\sigma}_n^2$ implies that if $X_{n,3}(t)$ converges weakly then it is also the case for the process

$$X_{n,4}(t) := \frac{1}{\hat{\sigma}_n \sqrt{n}} \sum_{k=1}^n (Y_k^n - \mu) G_k(t).$$

We have:

$$X_{n,4}(t) - X_{n,2}(t) = \frac{\bar{Y}_n - \mu}{\hat{\sigma}_n \sqrt{n}} \sum_{k=1}^n G_k(t).$$

Since under the conditions of Lemma 4.12 the law of the process $\{\sum_{k=1}^n G_k(t) : t \in [0, L]\}$ is the same as the one of $\{\sum_{k=1}^n \eta_k G_k(t) : t \in [0, L]\}$, we can apply part (a) of this lemma and Exercise 4.4 to deduce that the process $\{n^{-1/2} \sum_{k=1}^n G_k(t) : t \in [0, L]\}$ is stochastically uniformly bounded. Apply now the law of large numbers to the sequence of means $\{\bar{Y}_n\}_{n=1,2,\dots}$ to deduce that:

$$\sup_{t \in [0, L]} |X_{n,4}(t) - X_{n,2}(t)|$$

tends to zero in probability.

Let us see that also

$$(4.32) \quad \sup_{t \in [0, L]} |X_{n,3}(t) - X_{n,1}(t)| \Rightarrow 0.$$

In fact, some algebra permits to check that:

$$\begin{aligned} X_{n,3}(t) - X_{n,1}(t) &= \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^n (Y_k^n - \mu) G_k(t) - \frac{\sqrt{n}}{2\sigma} \sum_{k=1}^n Y_k^n \left(\frac{\mathbb{1}_{G_k(t)=1}}{\nu_n(t)} - \frac{\mathbb{1}_{G_k(t)=-1}}{n - \nu_n(t)} \right) \\ &= \frac{\sqrt{n}}{2\sigma} \sum_{k=1}^n (Y_k^n - \mu) \left[\frac{2}{n} G_k(t) - \left(\frac{\mathbb{1}_{G_k(t)=1}}{\nu_n(t)} - \frac{\mathbb{1}_{G_k(t)=-1}}{n - \nu_n(t)} \right) \right] \\ &= \frac{f_n(t) - 1/2}{\sigma\sqrt{n}f_n(t)[1 - f_n(t)]} \sum_{k=1}^n [G_k(t_0) \frac{\delta}{\sqrt{n}} + \varepsilon_k] \left[\frac{1}{2} + (1/2 - f_n(t)) G_k(t) \right] \end{aligned}$$

where we have put

$$\nu_n(t) = \#\{k : 1 \leq k \leq n, G_k(t) = 1\}$$

and

$$f_n(t) = \frac{1}{n} \nu_n(t) = \frac{1}{2n} \sum_{k=1}^n G_k(t) + \frac{1}{2}.$$

(4.32) follows using Lemma 4.12.

So weak convergence of $X_{n,3}$ implies the convergence of the other processes. The above proves also that the variance of $X_n(t)$ is equivalent to that of $X_{n,3}(t)$ which tends to 1, thus proving assertion a) in the statement of the theorem.

The model implies that at every location t and for every individual k , $P\{G_k(t) = 1\} = 1/2$, and

$$E(\varepsilon_k G_k(t)) = 0; \quad \text{Var}(\varepsilon_k G_k(t)) = \sigma^2.$$

To finish this part, let us turn to the process $\{X_{n,3}(t) : t \in [0, L]\}$. Set

$$X_{n,3}(t) := \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^n Y_k^n G_k(t) = \frac{\delta}{2n\sigma} \sum_{k=1}^n G_k(t_0) G_k(t) + \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^n \varepsilon_k G_k(t) = X_{n,5}(t) + X_{n,6}(t).$$

By Lemma 4.12, $X_{n,5}(t)$ converges uniformly to the function $\frac{\delta}{2\sigma} \exp(-2|t - t_0|)$ and $X_{n,6}$ converges weakly to the Ornstein-Uhlenbeck with a scale change, having covariance (4.28).

Step 2: Convergence in case of partial genetic information. Using the Skorohod embedding technique the weak convergence of the process $X_{n,3}(t)$ towards the limit process $X(t)$ can be represented by an almost sure convergence in some probability space. Since $X(t)$ has continuous sample paths, the convergence is also true in the uniform topology:

$$\|X_{n,3}(\cdot) - X(\cdot)\|_\infty \rightarrow 0 \text{ a.s.}$$

Now let \mathcal{D}_n be the operator on D that consists of (a) discretization at the locations $d_{i,n}$, $i = 1, \dots, M_n$ followed by (b) linear interpolation between two marker positions and extending by a constant before the first and after the last marker. This operator is a contraction for the uniform norm. We can deduce that:

$$\begin{aligned} (4.33) \quad \|\mathcal{D}_n[X_{n,3}(\cdot)] - X(\cdot)\|_\infty &\leq \|\mathcal{D}_n[X_{n,3}(\cdot)] - \mathcal{D}_n[X(\cdot)]\|_\infty + \|\mathcal{D}_n[X(\cdot)] - X(\cdot)\|_\infty \\ &\leq \|X_{n,3}(\cdot) - X(\cdot)\|_\infty + \|\mathcal{D}_n[X(\cdot)] - X(\cdot)\|_\infty \rightarrow 0 \end{aligned}$$

as $n \rightarrow +\infty$ and we are done. \square

4.3. Smoothing the detection test process. In the remaining of the study of this example, we will combine the previous theoretical results with a series of practical recipes to be able to answer the relevant questions.

The original problem is to test the null hypothesis $\delta = 0$ against $\delta \neq 0$. The classical approach would be to use the test statistic $T_n = \sup_{t \in [0, L]} |X_n(t)|$ which corresponds to a likelihood ratio test in the case of Gaussian observations. This is inconvenient because of two reasons:

- (1) The limit process has irregular sample paths (non-differentiable), and the distribution of its supremum is known only when $\delta = 0$ (Delong, 1981) and for certain lengths of the observation interval. In the other cases, for $\delta = 0$ we can use asymptotic bounds for the maximum of “ α -regular processes” that are due to Pickands (see, for example, Leadbetter Lindgren and Rootzen, 1983). But to compute the power of the test, that is for $\delta \neq 0$ the only available method is Monte-Carlo.
- (2) It does not take into account that the presence of a gene at t_0 modifies the expectation of the limit process in a neighborhood of t_0 .

Given these two reasons, we will smooth the paths of the detection process $\{X_n(t) : t \in [0, L]\}$ by means of convolution with a regular kernel, that we will take to be a centered Gaussian kernel having variance ε^2 , which we denote ϕ_ε . Let $\{X_n^\varepsilon(t) : t \in [0, L]\}$ be the smoothed process, defined as $X_n^\varepsilon(t) = (X_n * \phi_\varepsilon)(t)$.

We consider the test statistic $T_n^\varepsilon = \sup_{t \in [0, L]} |X_n^\varepsilon(t)|$. The reader can check that the limit of $(X_n^\varepsilon(t))_{t \in [0, L]}$ is the smoothed version of the previous limit process (or consult Billingsley’s book, 1999), and compute the mean $m^\varepsilon(\cdot)$ and the covariance $r^\varepsilon(\cdot)$ of the new process:

(4.34)

$$m^\varepsilon(t) = \frac{\delta}{2\sigma} \left\{ \exp[2(-t_0 + \varepsilon^2 + t)] \Phi\left(\frac{t_0 - 2\varepsilon^2 - t}{\varepsilon}\right) + \exp[2(t_0 + \varepsilon^2 - t)] \left[1 - \Phi\left(\frac{t_0 + 2\varepsilon^2 - t}{\varepsilon}\right)\right] \right\}$$

$$r^\varepsilon(t) = \exp(2(2\varepsilon^2 - t)) \Phi\left(\frac{t - 4\varepsilon^2}{(2\varepsilon^2)^{\frac{1}{2}}}\right) + \exp(2(2\varepsilon^2 + t)) \left[1 - \Phi\left(\frac{t + 4\varepsilon^2}{(2\varepsilon^2)^{\frac{1}{2}}}\right)\right].$$

To compute approximately the power of the test based on this process we use the basic inequalities (4.23). Of course, the use of crossings is feasible for the regularized process but it is not for the original one, that has non-differentiable paths. For the determination of the threshold, it turns out (on the basis of numerical simulation), that the lower-bound in it - which uses second order factorial moment of crossings - is more accurate than the upper one. So we determine thresholds using the lower-bound.

4.4. Simulation study. On the basis of a Monte-Carlo experiment one can evaluate the quality of the proposed method under a variety of conditions. Especially,

- the relationship between the value of the smoothing parameter and the validity of the asymptotic approximation for a reasonable number of markers and individuals in the sample,
- the sharpness of the bounds given by the inequality (4.23) for various values of the smoothing parameter.

Table 4.1 displays empirical levels for smoothed and unsmoothed procedures with thresholds computed under the asymptotic distribution.

- For the unsmoothed process ($\varepsilon = 0$), the threshold is calculated using Table II of Delong (1981). For this reason, the chromosome length, 0.98 M (Morgan) has been chosen to correspond to an entry of DeLong’s table, and to be close to lengths encountered for several vegetable species.
- For the smoothed process, we used the lower bound in inequality (4.23).

Simulations have been performed for two values of the smoothing parameter and three markers densities: a marker every each icM with $i = 1, 2, 7$. The number of individuals is equal to 500; the crossing-overs are simulated according to a standard Poisson process; the simulation has 10^4 realizations; 5% confidence interval for the empirical levels associated to the theoretical ones are indicated.

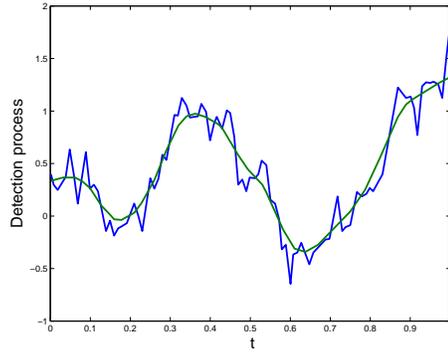


FIGURE 4.1. A realization of the detection process and its smoothing with $\varepsilon^2 = 10^{-3}$. There are 100 markers on a chromosome of size 1 Morgan and 500 individuals are observed.

	Nominal level of the test								
	10%			5%			1%		
5% confidence interval for the emp. level threshold $\varepsilon = 0$	9.41-10.59			4.57-5.43			0.80-1.19		
threshold $\varepsilon^2 = 10^{-2}$	2.74			3.01			3.55		
threshold $\varepsilon^2 = 10^{-3}$	2.019			2.276			2.785		
marker density	1 cM	2 cM	7 cM	1 cM	2 cM	7 cM	1 cM	2 cM	7 cM
emp. level $\varepsilon = 0$	7.37	6.67	4.99	3.91	3.42	2.4	0.77	0.67	0.43
emp. level $\varepsilon^2 = 10^{-2}$	12.17	12.17	11.82	6.75	6.69	6.53	1.76	1.72	1.77
emp. level $\varepsilon^2 = 10^{-3}$	10.84	10.66	9.71	5.63	5.55	5.02	1.34	1.32	1.04

TABLE 4.1. Threshold and empirical level (in %) of test using the unsmoothed detection test process ($\varepsilon = 0$) $(X_n(d))_{t \in [0, L]}$ and the smoothed detection process $(X_n^\varepsilon(t))_{d \in [0, L]}$. The chromosome length is equal to 0.98 M, and the number of individuals is equal to 500. The second line of the table gives a confidence interval for the empirical proportion related to the nominal level over 10^4 simulations.

Table 4.1 presents the power associated to the detection test in the case of a gene of size $\delta = 6$ located at the position $t_0 = 0.4$. The length of the chromosome is 1 M, calculations are made under the asymptotic distribution, using a test with nominal level equal to 5%.

- For the unsmoothed detection test process, the threshold is calculated via DeLong's table and the power using Monte-Carlo with 10^4 simulations.
- For the smoothed process, the threshold is calculated as above using the lower bound in (4.23). The power of the test is calculated in three ways: 1) using the upper bound in (4.23), 2) using the lower bound in (4.23), 3) by a Monte-Carlo method.

Summing up, let us add some final comments on this study of the genetic model (4.24).

- Table 4.1 clearly indicates that the unsmoothed procedure is very conservative.
- The empirical level given by the smoothed procedure is close to the nominal value. For $\varepsilon^2 = 10^{-3}$, it is nearly inside the confidence interval.
- Table 4.2 shows clearly that smoothing at size $\varepsilon^2 = 10^{-2}$ instead of 10^{-3} does not imply a sizable loss of the power computed with the asymptotic distribution.
- It is also clear from Table 4.2 that at the size $\varepsilon^2 = 10^{-2}, 10^{-3}$, the lower bound being almost exact.
- The use of the asymptotic test after smoothing with the window size $\varepsilon^2 = 10^{-3}$ and thresholds and powers computed by means of the lower bound in (4.23), has a number

	$\varepsilon^2 = 10^{-2}$	$\varepsilon^2 = 10^{-3}$	unsmoothed process
5% threshold	2.281	2.599	3.02
$P(Y(0) > u)$	15.70	9.81	-
$E((U_u + D_{-u}) \mathbb{1}_{ Y(0) \leq u})$	55.57	72.30	-
$\frac{\nu_2}{2}$	1.43	13.06	-
$\frac{E(U_u(U_u-1))}{2}$	1.43	13.06	-
$\frac{E(D_{-u}(D_{-u}-1))}{2}$	$7.84 \cdot 10^{-6}$	$3.40 \cdot 10^{-4}$	-
$E(U_u D_{-u})$	$2.54 \cdot 10^{-5}$	$4.10 \cdot 10^{-3}$	-
lower bound	69.84	69.05	-
upper bound	71.27	82.11	-
empirical power	71.37 ± 0.88	72.53 ± 0.87	68.99 ± 0.91

TABLE 4.2. Power in % associated to the detection test in the case of a gene of size $\delta = 6$, located at a distance $t_0 = 0.4$ from the origin of a chromosome of length 1 M. The value of σ is equal to 1. In the smoothed procedure, the 5% level, the upper and lower bounds are calculated using the S⁺ program previously mentioned. For the unsmoothed process, the 5% level is given by DeLong's table. The empirical powers are calculated over 10^4 simulations and the corresponding 95 % confidence intervals are given.

	chromosome length in Morgans					
	0.75	1	1.5	2	2.5	3
1% level	3.059	3.133	3.239	3.315	3.375	3.423
5 % level	2.516	2.599	2.721	2.809	2.878	2.934
10% level	2.239	2.328	2.458	2.553	2.626	2.687

TABLE 4.3. Thresholds calculated using the lower bound in (4.23) for different values of level α and various chromosome lengths. The smoothing parameter is equal to 10^{-3} .

of advantages. The thresholds corresponding to levels α equal to 1%, 5 % and 10% and for certain chromosome lengths are given in Table 4.3. For other cases and power calculations, the S⁺ program developed by Cierco-Ayrolles, Croquette and Delmas (2003) can be used (see also Chapter 9).

5. Mixtures of Gaussian distributions

A classical problem in statistical inference is the one of deciding, on the basis of a sample, whether a population should be considered homogeneous or a mixture of various different ones. We are going to address this question, but only when very simple possible models for the description of the population are present. This already leads to mathematical problems having a certain complexity and to the use of techniques that are directly related to our main subjects.

Our framework is restricted to Gaussian mixtures and we will consider the following hypothesis testing situations:

- (1) The simple mixture model

$$(4.35) \quad \begin{cases} H_0 : Y \sim N(0, 1) \\ H_1 : Y \sim pN(0, 1) + (1 - p)N(\mu, 1) \quad p \in [0, 1], \quad \mu \in \mathcal{M} \subset \mathbb{R} \end{cases}$$

We mean the following: assume that we are measuring a certain magnitude Y on each individual of a population. Under the null hypothesis H_0 , Y has a Normal (0,1) distribution for each individual. Under the alternative hypothesis H_1 each individual can be considered to have been chosen at random with probability p in a population in which

the magnitude Y is Normal $(0,1)$ and with probability $1 - p$ in a population in which Y is Normal $(\mu, 1)$. The purpose of the game is to make a decision about which one is the underlying true situation, on the basis of the observation of a sample.

The foregoing explanation applies to the other two cases that we describe next, *mutatis mutandis*.

(2) The test of one population against two, variance known.

$$(4.36) \quad \begin{cases} H_0 : Y \sim N(\mu_0, 1) \mu_0 \in \mathcal{M} \\ H_1 : Y \sim pN(\mu_1, 1) + (1 - p)N(\mu_2, 1) p \in [0, 1], \mu_1, \mu_2 \in \mathcal{M} \subset \mathbb{R} \end{cases}$$

(3) The test of one population against two, variance unknown.

$$(4.37) \quad \begin{cases} H_0 : Y \sim N(\mu, \sigma^2) \mu \in \mathcal{M} \subset \mathbb{R}, \sigma^2 \in \Sigma \subset \mathbb{R}^+ \\ H_1 : Y \sim pN(\mu_1, \sigma^2) + (1 - p)N(\mu_2, \sigma^2) p \in [0, 1], \mu_1, \mu_2 \in \mathcal{M} \subset \mathbb{R}, \sigma^2 \in \Sigma \subset \mathbb{R}^+. \end{cases}$$

These problems appear in many different kinds of applications and of course, the statistical methods apply quite independently of the field generating the problem. However, to persist in biological questions considered in the previous section, let us choose again genetics to show a possible meaning of the third case above.

Let us consider a quantitative trait on a given population, for example, the yield per unit surface in a plant breeding experiment. A reasonable model consists in assuming that such a complex trait is influenced by a large number of genes having each one of them, small effects. Assuming independence or almost-independence between the effects of the different genes, a heuristic application of the Central Limit Theorem leads to a Gaussian distribution for the trait. That corresponds to the null hypothesis in formula (4.37).

Suppose now that a mutation appears in the population, introducing a new allele that has, alone, a non negligible effect on the trait. Let G be the new allelic form and g the old one and suppose that one form is dominant, for example G (this means that Gg is equivalent to GG). Then the distribution of the trait in the considered population can be modelled by the general hypothesis in formula (4.37) with $(1 - p)$ being the frequency of individuals gg .

So, rejection of H_0 is associated to the detection of the existence of a new gene, and the purpose of the hypothesis testing is to take this decision on the basis of the observation of the value of the trait in a sample of the population. Of course, if H_0 is rejected, understanding the location of the gene will require, for example, some genetic markers information and the techniques of Section 4.

To perform a test of the type we have described above in (4.35), (4.36) or (4.37), there exist two main classical techniques:

- (1) a test based on moments: in a first approximation, expectation, variance, order-three moment.
- (2) a test based on likelihood ratio.

The asymptotic distribution of the likelihood ratio test was first established by Gosh and Sen (1985) under a strong separation hypothesis : for example $|\mu_1 - \mu_2| > \varepsilon > 0$ for model (4.36). Without this hypothesis but still assuming that the set of the means \mathcal{M} is compact and that the variance σ^2 is bounded away from zero, the asymptotic distribution has been studied by Dacunha-Castelle and Gassiat (1997, 1999) (see also Gassiat (2002) and Azaïs Gassiat and Mercadier (2007) for a simpler proof). See also Azaïs Gassiat and Mercadier (2006) for further developments and a discussion of the behavior when \mathcal{M} is not compact or is large in some sense.

On the other hand, moment based tests do not demand compactness assumptions and have invariance properties. Since the distribution of the Likelihood Ratio Test (LRT from now on) is related to that of the maximum of a rather regular Gaussian process, we will use a method based upon Rice formulas to address the following problems:

- (1) Is the power of the LRT test much influenced by the size of the interval(s) in which the parameters are supposed to be?
- (2) Is it true, as generally believed but without proof, that the LRT test is more powerful than the moment tests? Notice that Azaïs Gassiat and Mercadier (2006) have proven

that theoretically, the power of the LRT is smaller than the one of moment tests when the parameter set is very large.

Our aim here is to show how Rice formulas on crossings can be used to perform the computations required by the results below on the hypothesis testing problems on mixtures. We are not including proofs of the asymptotic statements, since they would lead us far away from the main subjects of this book. The interested reader can find them in the references above.

5.1. Simple mixture.

THEOREM 4.13 (Asymptotic distribution of the LRT). *Suppose that \mathcal{M} is a bounded interval that contains zero and define the local asymptotic alternative:*

$$\mu = \mu_0 \in \mathcal{M}, \quad (1 - p) = \frac{\delta}{\sqrt{n}},$$

for fixed μ_0 and δ , under this alternative the LRT of H_0 against H_1 has the distribution of the random variable

$$(4.38) \quad \frac{1}{2} \sup_{t \in \mathcal{M}} \{Z^2(t)\},$$

where $Z(\cdot)$ is a Gaussian process with mean

$$m(t) = \frac{\delta(e^{t\mu_0} - 1)}{\sqrt{e^{t^2} - 1}}$$

and covariance function

$$r(s, t) = \frac{e^{st} - 1}{\sqrt{e^{s^2} - 1}\sqrt{e^{t^2} - 1}}.$$

A direct application of the regularity results of Chapter 1 shows that the process $Z(\cdot)$ has C^∞ paths on $(0, +\infty)$ and on $(-\infty, 0)$, but has a discontinuity at $t = 0$ where it has right and left limits. Since

$$m(0^-) = -m(0^+) = -\delta\mu_0, \quad r(0^-, 0^+) = -1$$

it follows that a.s. $Z(0^-) = -Z(0^+)$.

We assume, for simplicity, that $\mathcal{M} = [-T, T]$ and set $M_T^* := \sup_{t \in [-T, T]} |Z(t)|$. Then, we have the following inequalities, which are analogous to (4.23).

Let

$$\xi := U_u[0, T] + D_{-u}[0, T] + D_u[-T, 0] + U_{-u}[0, T].$$

Then:

$$(4.39) \quad \mathbb{P}(|Z(0)| > u) + \mathbb{E}(\xi \mathbf{1}_{|Z(0)| < u}) - \frac{1}{2} \mathbb{E}(\xi^{[2]}) \leq \mathbb{P}(M_T^* > u) \leq \mathbb{P}(|Z(0)| > u) + \mathbb{E}(\xi \mathbf{1}_{|Z(0)| < u}).$$

and we can use the results in Section 2 to compute upper and lower bounds for $\mathbb{P}(M_T^* > u)$. The deduced critical values for the test are shown in Table 4.4.

This table shows that the Rice method is very precise. It also shows that the critical values depend heavily on the size of the interval \mathcal{M} .

As for the power, we give some examples in Table 4.5. More examples can be found in Delmas (2001) (2003 a).

In Table 4.5 we can see that the power is affected by the size of \mathcal{M} . For example for $\mu_0 = 1$ $\delta = 1$ the power varies from 16% to 23%.

Let us now compare the LRT test with two tests based on moments:

Size	Nominal level of the test		
	1%	5%	10%
T=0.5	3.6368	2.1743	1.5710
T=1	3.9015	2.3984	1.7707-1.7713
T=2	4.3432-4.3438	2.7908-2.7942	2.1300-2.1378
T=3	4.6784-4.6798	3.1002-3.1075	2.4209-2.4350
T=5	5.1357-5.1384	3.5353-3.5478	2.8390-2.8635
T=10	5.7903-5.7940	4.1736-4.1921	3.4641-3.5013
T=15	6.1837-6.1879	4.5621-4.5828	3.8480-3.8908
T=20	6.4657-6.4703	4.8414-4.8636	4.1253-4.1712
T=30	6.8658-6.8705	5.2392-5.2627	4.5212-4.5698
T=50	7.3725-7.3774	5.7443-5.7688	5.0245-5.0758

TABLE 4.4. Critical values or thresholds for the LRT test for simple Gaussian mixture: upper-bound and lower-bound when they differ significantly.

Size	Location	$\delta = 1$	$\delta = 3$
$T = 0.5$	$\mu_0 = 0.25$	5.79	11.84
	$\mu_0 = 0.5$	8.31	34.89
$T = 1$	$\mu_0 = 0.25$	5.79	11.48
	$\mu_0 = 0.5$	8.40	34.59
	$\mu_0 = 1$	23.73-23.74	96.62-96.63
$T = 2$	$\mu_0 = 0.25$	5.64-5.69	10.18-10.24
	$\mu_0 = 1$	22.93-23.09	95.99-96.19
	$\mu_0 = 2$	100	100
$T = 3$	$\mu_0 = 0.25$	5.49-5.57	9.17-9.28
	$\mu_0 = 1$	20.21-20.50	94.76-95.17
	$\mu_0 = 1.5$	74.04-75.02	100
	$\mu_0 = 3$	100	100
$T = 5$	$\mu_0 = 0.25$	5.31-5.44	8.02-8.20
	$\mu_0 = 1$	16.76-17.18	92.66-93.53
	$\mu_0 = 1.5$	68.34-69.99	99.99-100
	$\mu_0 = 2.5$	100	100

TABLE 4.5. Power in % of the LRT at the level 5% as a function of the size of \mathcal{M} and of δ and μ_0 . Again, when they differ significantly the upper-bound and lower-bound are given.

- (1) The “ \bar{X}_n ” test, which is based on the fact the mean is zero under H_0 and that under the local alternative

$$(4.40) \quad \sqrt{n}\bar{X}_n \xrightarrow{\mathcal{D}} N(\delta\mu_0, 1), \text{ as } n \text{ tends to } +\infty.$$

- (2) The “ S_n^2 ” test, which is based on the fact the variance is one under H_0 and that under the local alternative

$$(4.41) \quad \sqrt{\frac{n}{2}}(S_n^2 - 1) \xrightarrow{\mathcal{D}} N\left(\frac{\delta\mu_0^2}{\sqrt{2}}, 1\right), \text{ as } n \text{ tends to } +\infty.$$

Since the S_n^2 test has one-side rejection region while the \bar{X}_n has two-side rejection region, it is not straightforward to see which one of these two tests is more powerful. Moreover, the answer depends on the level. A general fact is that for large μ_0 the S_n^2 test is more powerful. Comparisons of the three tests are presented in Figure 4.2.

The main point, which can be seen on Figure 4.2, is that **the likelihood ratio test is not uniformly more powerful.**

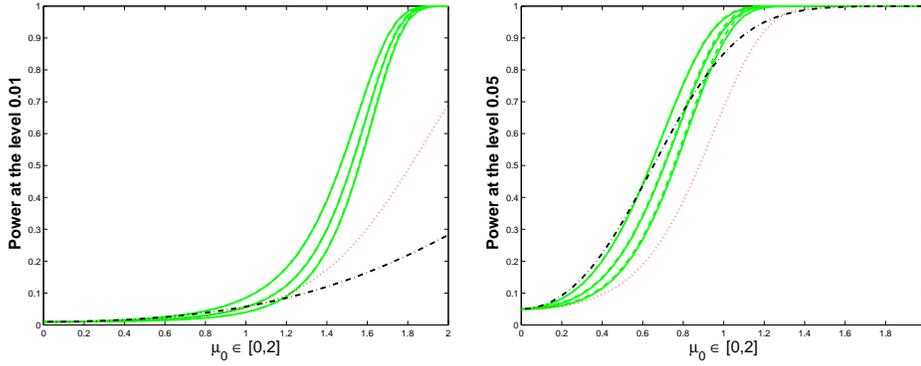


FIGURE 4.2. Variation of the power of the three tests as a function of μ_0 . The power is displayed in black (dashed-dotted) for of the \bar{X}_n test, in red (dotted) for the S_n^2 test and in green for the LRT. For the last one, the upper bound is in dashed line and the lower-bound in solid line (they almost coincide) and from top to bottom we have the power for $T = 2, 5$ and 10 . In the picture on the left, the level is 0.01 and $\delta^2 = 1$; in the one on the right, the level is 0.01 and $\delta^2 = 3$

Figure 4.2, left, is rather typical in the sense that the situations for which the LRT is not optimal (for example $T = 10$, $\mu_0 = 1$) correspond to very small power. They are actually uninteresting.

Figure 4.2, right, corresponds to a deliberate choice of a situation where the LRT behaves badly. For example, for $\mu_0 = 0.6$ and $T = 10$ the lack of power of the LRT as compared to the \bar{X}_n test is important.

5.2. One population against two, σ^2 known.

THEOREM 4.14 (Asymptotic distribution of the LRT). *Suppose that \mathcal{M} is a bounded interval that now will be chosen of the form $\mathcal{M} = [0, T]$. We define the local asymptotic alternative in model by means of (4.36) :*

$$(1-p) = \frac{\delta^2}{\sqrt{n}}, \quad \mu_1 - \mu_0 = \frac{\alpha}{\sqrt{n}}, \quad \mu_2 = \mu_{2,0} \neq \mu_0 \in (0, T)$$

for fixed $\mu_0, \mu_{2,0}$ and δ .

Then, under this alternative, the LRT of H_0 against H_1 has the limit distribution given by (4.38) where now $Z(\cdot)$ is a Gaussian process with mean

$$m(t) = \frac{\delta^2(e^{(\mu_0 - \mu_{2,0})(\mu_0 - t)} - 1 - (\mu_0 - \mu_{2,0})(\mu_0 - t))}{\sqrt{e^{(\mu_0 - t)^2} - 1 - (\mu_0 - t)^2}}.$$

and covariance function

$$r(s, t) = \frac{e^{(s - \mu_0)(t - \mu_0)} - 1 - (s - \mu_0)(t - \mu_0)}{\sqrt{e^{(s - \mu_0)^2} - 1 - (s - \mu_0)^2} \sqrt{e^{(t - \mu_0)^2} - 1 - (t - \mu_0)^2}}.$$

One should notice that these functions are of class C^∞ and so are the sample paths of the process. As a consequence the Rice method applies directly. But a new problem arises since the null hypothesis is composite (it consists of more than one distribution) and the distribution of the LRT statistic under H_0 is not free from the parameter μ_0 , which is unknown in practical applications. Table 4.6 illustrates the variation of the threshold as a function of μ_0 and T .

We can clearly see in Table 4.6 that the value taken by μ_0 does not matter very much. So the LRT can be used in practical situations. As in the preceding subsection, Figure 4.3 compares the power of the LRT test with the S_n^2 test based on the fact that the variance is one under H_0 . Of course, the \bar{X}_n test cannot be performed since the expectation is not yet fixed under H_0 .

size	position	Level of the test		
		1%	5%	10%
T=1	$\mu_0=0.5$	3.1524	1.6925	1.0991
	$\mu_0=0.25$	3.1561	1.6956	1.1018
	$\mu_0=0+$	3.1676	1.7049	1.1097
T=2	$\mu_0=1$	3.4921-3.4924	1.9779-1.9782	1.3487-1.3490
	$\mu_0=0.5$	3.5137-3.5139	1.9967-1.9970	1.3658-1.3661
	$\mu_0=0$	3.5712	2.0475-2.0477	1.4120-1.4121
T=4	$\mu_0=2$	4.0424-4.0439	2.4733-2.4772	1.8048-1.8106
	$\mu_0=1$	4.0904-4.0915	2.5189-2.5220	1.8486-1.8532
	$\mu_0=0+$	4.1876-4.1884	2.6108-2.6141	1.9362-1.9420
T=6	$\mu_0=3$	4.4663-4.4688	2.8738-2.8826	2.1869-2.2018
	$\mu_0=1.5$	4.4863-4.4882	2.8934-2.9009	2.2059-2.2194
	$\mu_0=0+$	4.5756-4.5776	2.9797-2.9879	2.2897-2.3048
T=10	$\mu_0=5$	5.0092-5.0126	3.3994-3.4137	2.6975-2.7252
	$\mu_0=2.5$	5.0100-5.0132	3.4001-3.4142	2.6981-2.7254
	$\mu_0=0+$	5.0746-5.0778	3.4636-3.4779	2.7608-2.7886

TABLE 4.6. Variation of the threshold of the test as a function of the level, the size T of the interval and the position μ_0 . Upper-bound and lower-bound are given when they differ significantly.

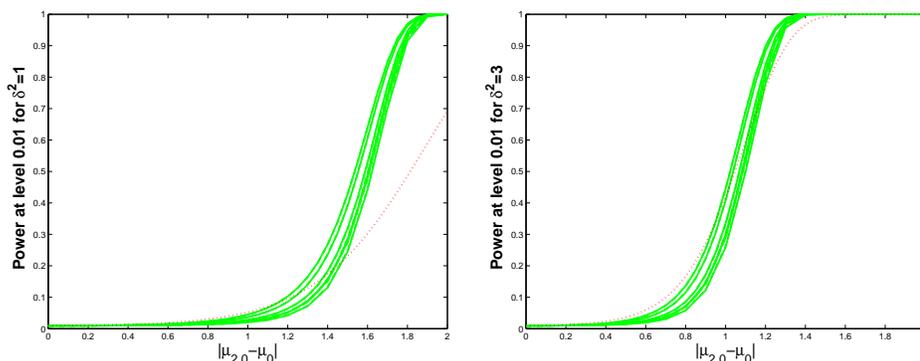


FIGURE 4.3. Variation of the power of the LRT and the S_n^2 test as a function of $\mu_{2,0} - \mu_0$ at the level 1%. The power is displayed in red (dotted) for S_n^2 test and in green for the LRT. For this last test the upper-bound, is in dashed line and the lower-bound in solid line. From top to bottom we have the power for $T = 4, 10$ and 15 . Left for $\delta^2 = 1$, right for $\delta^2 = 3$. Some lines are superposed because the upper-bound and the lower-bound are numerically equal. The difference between upper- and lower-bound is due to both the inequality (4.23) and to the variation of the nuisance parameter μ_0 inside \mathcal{M} .

We observe roughly the same phenomenon as in the case of the simple mixture problem.

5.3. One population against two, σ^2 unknown.

THEOREM 4.15 (asymptotic distribution of the LRT). *Suppose that \mathcal{M} is bounded interval and that $\Sigma = [S_1, S_2], 0 < S_1 < S_2 < +\infty$ and define the local asymptotic alternative in model (4.37) :*

$$1 - p = \frac{\delta^2}{\sqrt{n}}, \quad \mu_1 - \mu_0 = \frac{\alpha}{\sqrt{n}}, \quad \sigma^2 - \sigma_0^2 = \frac{\beta}{\sqrt{n}}, \quad \mu_2 = \mu_{2,0} \neq \mu_0$$

for fixed $\mu_0, \mu_{2,0}, \alpha, \beta$ and δ .

	Level of the test		
	1%	5%	10%
$\nu = [-2; 2]$	4.1533	2.6209	1.9755
$\nu = [-5; 5]$	5.0390	3.4514	2.7688
$\nu = [-8; 8]$	5.5168	3.9183	3.2296
$\nu = [-1; 3]$	4.2076	2.6687	2.0180
$\nu = [-2.5; 7.5]$	5.0413	3.4532	2.7707
$\nu = [-4; 12]$	5.5168	3.9183	3.2296
$\nu = [0; 4]$	4.3092	2.7594	2.1002
$\nu = [0; 10]$	5.1242	3.5330	2.8477
$\nu = [0; 16]$	5.5708	3.9714	3.2814

TABLE 4.7. Upper-bound for the threshold of the LRT test as a function of the level and of the set of parameter after the change of variable (4.42)

Then under this alternative the LRT of H_0 against H_1 has the limit distribution given by (4.38) where **now** $Z(\cdot)$ is a Gaussian process with mean

$$m(t) = \frac{\delta^2 \left[\exp\left(\frac{(t-\mu_0)(\mu_{2,0}-\mu_0)}{\sigma_0^2}\right) - 1 - \frac{(t-\mu_0)(\mu_{2,0}-\mu_0)}{\sigma_0^2} - \frac{(t-\mu_0)^2(\mu_{2,0}-\mu_0)^2}{2\sigma_0^4} \right]}{\sqrt{\exp\left(\frac{(t-\mu_0)^2}{\sigma_0^2}\right) - 1 - \frac{(t-\mu_0)^2}{\sigma_0^2} - \frac{(t-\mu_0)^4}{2\sigma_0^4}}}$$

and covariance function

$$r(s, t) = \frac{\exp\left(\frac{(s-\mu_0)(t-\mu_0)}{\sigma_0^2}\right) - 1 - \frac{(s-\mu_0)(t-\mu_0)}{\sigma_0^2} - \frac{(s-\mu_0)^2(t-\mu_0)^2}{2\sigma_0^4}}{\sqrt{\exp\left(\frac{(s-\mu_0)^2}{\sigma_0^2}\right) - 1 - \frac{(s-\mu_0)^2}{\sigma_0^2} - \frac{(s-\mu_0)^4}{2\sigma_0^4}} \sqrt{\exp\left(\frac{(t-\mu_0)^2}{\sigma_0^2}\right) - 1 - \frac{(t-\mu_0)^2}{\sigma_0^2} - \frac{(t-\mu_0)^4}{2\sigma_0^4}}}$$

The process $Z(\cdot)$ of Theorem 4.15 is in fact a random function of $\nu := \frac{t-\mu_0}{\sigma_0}$

$$(4.42) \quad Z(t) = T\left(\frac{t-\mu_0}{\sigma_0}\right)$$

and the process $T(\cdot)$ satisfies a.s. $T(0^-) = -T(0^+)$ as we found in the simple mixture model. The method based on crossings can be applied in much the same manner. Table 4.7 displays the variation of the threshold as a function of the level and of the size of the parameter set $\nu = \frac{t-\mu_0}{\sigma_0}$.

The intervals of variation for the parameter ν that are displayed in Table 4.7 correspond, for example, to $t \in \mathcal{M} = [5, 15]$ and several values of σ_0 and μ_0 . The table is divided into three “great rows” corresponding to $\mu_0 = 10, 7, 5$. Each “great row” is divided into three rows corresponding to $\sigma_0 = 5/2, 1, 5/8$.

The threshold depends now heavily on the form of the null hypothesis. If we know that (μ_0, σ_0) are close to some prior value, we can take the threshold corresponding to this prior value. In the others cases, a classical choice is to take the highest value of the threshold. This in general leads to an important loss of power. Another possibility would be to perform a “plug-in”, that is, to take for μ_0 and σ_0^2 the values computed from an empirical estimation. The behavior of such a procedure does not seem to have been studied yet.

In this paragraph the LRT is compared with a moment test based on the difference between the rough estimator of variance S_n^2 and a robust estimator (Caussinus and Ruiz, 1995). Figure 4.4 displays the powers of the two tests.

Summing up, for the last two models, the distribution of the LRT is not free under H_0 , since it depends on the position of the true mean μ_0 with respect of the interval in which the means are supposed to be. When σ^2 is known, this dependence is not heavy, so that LRT can be performed in

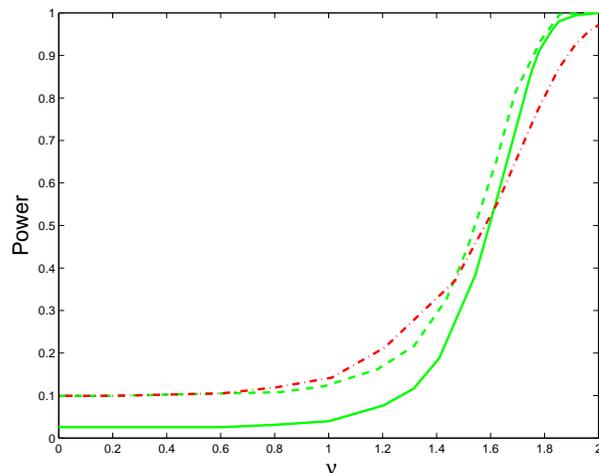


FIGURE 4.4. Variation of the power of the LRT and moment test at the level 10% as a function of $\nu = \frac{t-\mu_0}{\sigma_0}$. The power is displayed in red (dashed-dotted) for the moment test and in green for the LRT. For this last test the upper-bound, in dotted line, corresponds to the best choice of the nuisance parameter μ_0/σ_0 and the lower-bound, in solid line, corresponds to the worst choice.

practice without introducing a prior value or an estimation for μ_0 . The situation is more complex when σ is unknown. In any case, LRT appears to be non-uniformly optimal but in most of the relevant situations, more powerful than moment tests. It remains the best choice in practice.

Exercises

EXERCISE 4.1. (Exact formula for the distribution of M_T for the sine-cosine process) : the simplest periodic Gaussian process.

Let the stochastic process $\mathcal{X} = \{X(t) : t \in \mathbb{R}\}$ be defined as:

$$X(t) := \xi_1 \cos(\omega t) + \xi_2 \sin(\omega t),$$

where ξ_1 and ξ_2 are two independent standard normal random variables and ω is a real number, $\omega \neq 0$.

(a) Show that we can write the process \mathcal{X} as:

$$X(t) = Z \cos(\omega(t + \theta))$$

where Z and θ are independent random variables, having respectively the distributions:

- square root of a χ_2^2 for Z
- uniform on $[0, 2\pi]$ for θ .

(b) Show that the covariance of \mathcal{X} is $r(s, t) = \Gamma(t-s) = \cos(\omega(t-s))$, and its spectral measure is $\frac{1}{2}(\delta_\omega + \delta_{-\omega})$ where δ_x denotes the unit atom at the point x .

Prove that for $u > 0$

$$(4.43) \quad \text{For } T \leq \frac{\pi}{\omega} : \mathbb{P}\{M_T > u\} = 1 - \Phi(u) + \frac{T\omega}{2\pi} e^{-\frac{u^2}{2}}$$

$$(4.44) \quad \text{For } T \geq \frac{2\pi}{\omega} : \mathbb{P}\{M_T > u\} = e^{-\frac{u^2}{2}} = \mathbb{P}\{|Z| > u\}$$

$$(4.45) \quad \text{For } \frac{\pi}{\omega} \leq T < \frac{2\pi}{\omega} : \mathbb{P}\{M_T > u\} = 1 - \Phi(u) + \frac{T\omega}{2\pi} e^{-\frac{u^2}{2}}$$

$$(4.46) \quad - \int_{\pi}^{T/\omega} \frac{1}{2\pi} \exp\left[-\frac{u^2(1-\cos(t))}{\sin^2(t)}\right] dt.$$

EXERCISE 4.2. Let $\{X(t) : t \in \mathbb{R}\}$ be a centered stationary Gaussian process having covariance function $\Gamma(t-s) = E(X(s)X(t))$, normalized by $\Gamma(0) = 1$.

Assume that Γ satisfies the hypothesis A2 in Proposition 4.2 and denote, as usual, $\lambda_2 = -\Gamma''(0)$ the second spectral moment. Let $\tau = \inf\{t > 0 : \Gamma(t) = 1\}$.

Exclude the trivial case in which $\lambda_2 = 0$. Let $M_T = \max_{t \in [0, T]} X(t)$.

(a) Prove that $\tau > 0$.

(b) Show that

$$P(X(t) = X(t + \tau) \text{ for all } t \in \mathbb{R}) = 1.$$

(that is, almost surely the paths are periodic functions having period τ).

(c) Prove that if $T < \tau$, then the conclusion of Proposition 4.2 holds true.

(d) Let $T \geq \tau$. Prove that as $u \rightarrow +\infty$,

$$P(M_T > u) = \tau \sqrt{\frac{\lambda_2}{2\pi}} \phi(u) + O(\phi(u(1+\delta)))$$

for some $\delta > 0$.

Hint: Prove and use the following equalities:

- $M_T = M_\tau$.
- $P(M_\tau > u) = P(X(t) > u \text{ for all } t \in [0, \tau]) + P(U_u(X, [0, \tau]) \geq 1)$.

EXERCISE 4.3. Suppose that f_n is a sequence of functions in D that converges to f for the Skorohod topology. Prove that if f is continuous then in fact the convergence holds true for the uniform distance.

EXERCISE 4.4. Let $\{Z_n\}_{n=1,2,\dots}$ be a sequence of random variables with values in the space D . Assume that $Z_n \Rightarrow 0$ as $n \rightarrow +\infty$ (which means that the distribution of Z_n tends to the unit measure at the identically zero function).

Then, for each $\varepsilon > 0$:

$$P\left(\sup_{t \in [0,1]} |Z_n(t)| \geq \varepsilon\right) \rightarrow 0$$

as $n \rightarrow +\infty$, that is, $\sup_{t \in [0,1]} |Z_n(t)|$ tends to zero in probability.

EXERCISE 4.5. Prove relations (4.40) and (4.41). Hint: use the Central Limit Theorem under Lindeberg's condition.

EXERCISE 4.6. Consider the function $\Gamma : \mathbb{R} \rightarrow \mathbb{R}^+$ defined as:

$$\Gamma(t) = \exp(-|t|).$$

(a) Show that Γ is the covariance of a stationary centered Gaussian process $\{X(t) : t \in \mathbb{R}\}$ (which is called the "Ornstein-Uhlenbeck" process).

(b) Compute the spectral density of the process.

(c) Study the Hölder properties of the paths.

(d) Show that the process is Markovian, that is, if $t_1 < t_2 < \dots < t_k$, k a positive integer $k \geq 2$, then,

$$P(X(t_k) \in B | X(t_1) = x_1, \dots, X(t_{k-1}) = x_{k-1}) = P(X(t_k) \in B | X(t_{k-1}) = x_{k-1})$$

for any Borel set B and any choice of x_1, \dots, x_{k-1} . (The conditional probability is to be interpreted in the sense of Gaussian regression).

(e) Let $\{W(t) : t \geq 0\}$ be a Wiener process. Show that

$$X(t) = \exp(-t) W(\exp(2t)), \quad t \in \mathbb{R}$$

is an Ornstein-Uhlenbeck process.

EXERCISE 4.7. Prove the statements about the properties of the process $\{Z(t) : t \in \mathbb{R} \setminus \{0\}\}$ in Subsection 5.1.

The Rice Series

Let $\mathcal{X} = \{X(t) : t \in [0, T]\}$ be a one-parameter stochastic process with real values and let us denote by $M_T := \sup_{t \in T} X(t)$ its supremum. In this chapter we will continue to study the distribution of the random variable M_T , that is, the function $F_{M_T}(u) := P(M_T \leq u)$, $u \in \mathbb{R}$ and we will express this distribution by means of a series (the “Rice series”) whose terms contain the factorial moments of the number of up-crossings. The underlying ideas have been known for a long time (Rice (1944-1945), Slepian (1962), Miroschin (1974)). The results in this chapter are taken from Azais & Wschebor (2002). We have included some numerical computations that have been performed with the help of A. Croquette and C. Delmas.

The main result in this chapter is to prove the convergence of the Rice series in a general framework instead of considering only some particular processes. This provides a method that can be applied to a large class of stochastic processes.

A typical situation is given by Theorem 5.6, that states that if a stationary Gaussian process has a covariance with a Taylor expansion at zero that is absolutely convergent at $t = 2T$, then $F_{M_T}(u)$ can be computed by means of the Rice series. On the other hand, even though Theorems 5.1 and 5.7 below do not refer specifically to Gaussian processes, in practice, for the time being we are able to apply them to the numerical computation of $F_{M_T}(u)$ only in Gaussian cases.

The section “Numerical aspects of Rice series” includes a comparison between the complexities of the computation of $F_{M_T}(u)$ using the Rice series versus Monte-Carlo method, in the relevant case of a general class of stationary Gaussian processes. It shows that the use of Rice series is a priori better. More important is the fact that the Rice series is self-controlling for the numerical errors. This implies that the a posteriori number of computations can be much smaller than the one required by simulation. In fact, in relevant cases for standard bounds for the error, the actual computation is performed with a few terms of the Rice series.

As examples we give tables for $F_{M_T}(u)$ for a number of Gaussian processes. When the length of the interval T increases, one needs an increasing number of terms in the Rice’s series not to surpass a given bound for the error. For small values of T and large values of the level u one can use the so-called “Davies bound” (4.2) or more accurately, the first term in the Rice series which is in fact the upper-bound in Inequality (4.15).

As T increases, for moderate values of u the “Davies bound” is far from the true value and one requires the computation of several terms (see Figures 5.1 to 5.4). Numerical results are shown in the case of four Gaussian stationary processes for which no closed formula is known. The same examples will be considered in Chapter 9 which is devoted to the “Record Method”.

One of the key points is the numerical computation of the factorial moments of up-crossings by means of Rice integral formulas. The main difficulty is the precise description of the behavior of the integrands appearing in these formulas near the diagonal, which is again a subject that is interesting on its own (see Belayev (1966), Cuzick (1975)). Even though this is an old subject, it remains widely open. We have included a section “Computation of Moments” in which we give some partial answers that are helpful to improve the numerical methods and have also some theoretical interest.

The extension to processes with non-smooth trajectories can be done by smoothing the paths by means of a deterministic device, applying the previous methods to the regularized process and estimating the error as a function of the smoothing band-width. The last section of the present chapter contains these type of results, that have been included even though they do not seem to have at present practical uses for the actual computation of the distribution of the maximum.

1. The Rice Series

We recall the following notations:

- $U_u = U_u(X, [0, T])$ is the number of up-crossings of the level u by the function $X(\cdot)$ on the interval $[0, T]$.
- $\tilde{\nu}_m := \mathbb{E}(U_u^{[m]} \mathbf{1}_{\{X(0) \leq u\}})$ ($m = 1, 2, \dots$).
- $\nu_m := \mathbb{E}(U_u^{[m]})$ ($m = 1, 2, \dots$).

$\tilde{\nu}_m$ is the factorial moment of the number of up-crossings, when starting below u at $t = 0$.

The Rice formula to compute $\tilde{\nu}_m$, whenever it holds true is the following:

$$\tilde{\nu}_m = \int_{[0, T]^m} dt_1 \dots dt_m \int_{-\infty}^u \mathbb{E}(X'^+(t_1) \dots X'^+(t_m) | X(0) = x, X(t_1) = \dots = X(t_m) = u) \cdot p_{X(0), X(t_1), \dots, X(t_m)}(x, u, \dots, u) dx$$

This section contains two main results. The first is Theorem 5.1 that requires the process to have \mathcal{C}^∞ paths and contains a general condition enabling to compute $F_{M_T}(u)$ as the sum of a series. The second is Theorem 5.6, which illustrates the same situation for Gaussian stationary processes. As for Theorem 5.7, it contains upper and lower bounds on $F_{M_T}(u)$ for processes with \mathcal{C}^k paths, verifying some additional conditions.

THEOREM 5.1. *Assume that a.s. the paths of the stochastic process \mathcal{X} are of class \mathcal{C}^∞ and that the density $p_{X_{T/2}}(\cdot)$ is bounded by some constant D .*

(i) *If there exists a sequence of positive numbers $\{c_k\}_{k=1,2,\dots}$ such that:*

$$(5.1) \quad \gamma_k := \mathbb{P}\left(\|X^{(2k-1)}\|_\infty \geq c_k \cdot T^{-(2k-1)}\right) + \frac{Dc_k}{2^{2k-1}(2k-1)!} = o(2^{-k}) \quad (k \rightarrow \infty)$$

then :

$$(5.2) \quad 1 - F_{M_T}(u) = \mathbb{P}(X(0) > u) + \sum_{m=1}^{\infty} (-1)^{m+1} \frac{\tilde{\nu}_m}{m!}$$

(ii) *In formula (5.2) the error when one replaces the infinite sum by its m_0 -th partial sum is bounded by $\gamma_{m_0+1}^*$ where:*

$$\gamma_m^* := \sup_{k \geq m} (2^{k+1} \gamma_k).$$

We will call the series in the right-hand side of (5.2) the "Rice Series". For the proof we will assume, with no loss of generality that $T = 1$. We start with the following Lemma on the remainder for polynomial interpolation (Davis 1975, Th. 3.1.1). It is a standard tool in numerical analysis.

LEMMA 5.2. *a) Let $f : I \rightarrow \mathbb{R}$, $I = [0, 1]$, be a function of class \mathcal{C}^k , k a positive integer, t_1, \dots, t_k , k points in I and let $P(t)$ be the - unique - interpolation polynomial of degree $k-1$ such that $f(t_i) = P(t_i)$ for $i = 1, \dots, k$, taking into account possible multiplicities. Then, for $t \in I$:*

$$f(t) - P(t) = \frac{1}{k!} (t - t_1) \dots (t - t_k) f^{(k)}(\xi)$$

where

$$\min(t_1, \dots, t_k, t) \leq \xi \leq \max(t_1, \dots, t_k, t).$$

b) *If f is of class \mathcal{C}^k and has k zeros in I (taking into account possible multiplicities), then:*

$$|f(1/2)| \leq \frac{1}{k! 2^k} \|f^{(k)}\|_\infty.$$

PROOF. Fix $t \in I, t \neq t_1, \dots, t_k$ and define

$$F(v) = f(v) - P(v) - \frac{(v - t_1) \dots (v - t_k)}{(t - t_1) \dots (t - t_k)} [f(t) - P(t)]$$

Clearly F has at least the $k + 1$ zeros t_1, \dots, t_k, t , so that by Rolle's Theorem, there exists ξ , $\min(t_1, \dots, t_k, t) \leq \xi \leq \max(t_1, \dots, t_k, t)$ such that $F^{(k)}(\xi) = 0$. This gives a). b) is a simple consequence of a), since in this case the interpolating polynomial vanishes. \square

The next combinatorial lemma plays the central role in what follows. A proof is given in Lindgren (1972), similar to the one we include here.

LEMMA 5.3. *Let ξ be a non-negative integer-valued random variable having finite moments of all orders. Let $k, m, M (k \geq 0, m \geq 1, M \geq 1)$ be integers and denote:*

$$p_k := P(\xi = k); \quad \mu_m := E(\xi^{[m]}); \quad S_M := \sum_{m=1}^M (-1)^{m+1} \frac{\mu_m}{m!}$$

Then

(i) For each M :

$$(5.3) \quad S_{2M} \leq \sum_{k=1}^{2M} p_k \leq \sum_{k=1}^{\infty} p_k \leq S_{2M+1}$$

(ii) The sequence $\{S_M\}_{M=1,2,\dots}$ has a finite limit if and only if $\mu_m/m! \rightarrow 0$ as $m \rightarrow \infty$, and in that case:

$$(5.4) \quad P(\xi \geq 1) = \sum_{k=1}^{\infty} p_k = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{\mu_m}{m!}.$$

PROOF. (ii) is an immediate consequence of (i). As for (i) denote by $\binom{k}{m}$ the binomial numbers and write

$$(5.5) \quad S_M = \sum_{m=1}^M (-1)^{m+1} \sum_{k=m}^{\infty} \binom{k}{m} p_k = \sum_{k=1}^{\infty} p_k B_{k,M}$$

with

$$(5.6) \quad B_{k,M} := \sum_{m=1}^{k \wedge M} (-1)^{m+1} \binom{k}{m}$$

It is clear that $B_{k,M} = 1$ if $k \leq M$.

If $k > M$, we have two cases:

1.- $k \geq 2M$.

Note that $\binom{k}{m}$ increases with m if $1 \leq m \leq \frac{k}{2}$. It follows that $B_{k,M} \geq k$ if M is odd and $B_{k,M} \leq -\frac{k}{2}$ if M is even, since $B_{k,M} \leq \binom{k}{1} - \binom{k}{2} \leq -k/2$, given that $k \geq 2M \geq 4$.

2.- $M < k < 2M$.

Check that in this case:

$$B_{k,M} = 1 + (-1)^{k+1} \sum_{h=0}^{k-M-1} (-1)^{h+1} \binom{k}{h} = 1 + (-1)^{k+1} (B_{k,k-M-1} - 1).$$

with the convention $B_{k,0} = 0$.

Since $k > 2(k - M - 1)$, if $0 < k - M - 1 < k$, we can apply the first case and it turns out that

$$\begin{aligned} k - M - 1 \text{ odd} &\Rightarrow B_{k,k-M-1} \geq k \\ k - M - 1 \text{ even} &\Rightarrow B_{k,k-M-1} \leq -k/2. \end{aligned}$$

Finally if $k = M + 1$, $B_{k,M} = 2$ when M is odd and $B_{k,M} = 0$ if M is even.

Summing up the two cases, if $k > M$, we have $B_{k,M} > 1$ if M is odd and $B_{k,M} \leq 0$ if M is even.

So that from

$$S_M = \sum_{k=1}^M p_k + \sum_{k=M+1}^{\infty} p_k B_{k,M}$$

one gets (i). This proves the lemma. \square

Remark. A by-product of Lemma 5.3 that will be used in the sequel is the following: if in (5.4) one substitutes the infinite sum by the M -partial sum, the absolute value $\mu_{M+1}/(M+1)!$ of the first neglected term is an upper-bound for the error in the computation of $P(\xi \geq 1)$.

LEMMA 5.4. *With the same notations as in Lemma 5.3 we have the equality:*

$$E(\xi^{[m]}) = m \sum_{k=m}^{\infty} (k-1)^{[m-1]} P(\xi \geq k) \quad (m = 1, 2, \dots).$$

PROOF. Check the identity

$$j^{[m]} = m \sum_{k=m-1}^{j-1} (k)^{[m-1]}$$

for each pair of integers j, m . So,

$$\begin{aligned} E(\xi^{[m]}) &= \sum_{j=m}^{\infty} j^{[m]} P(\xi = j) = \sum_{j=m}^{\infty} P(\xi = j) m \sum_{k=m}^j (k-1)^{[m-1]} = \\ &= m \sum_{k=m}^{\infty} (k-1)^{[m-1]} P(\xi \geq k). \end{aligned}$$

\square

LEMMA 5.5. *Suppose that a.s. the paths of the process \mathcal{X} are of class C^∞ and that the density $p_{X_{1/2}}(\cdot)$ is bounded by the constant D . Then for any sequence $\{c_k\}_{k=1,2,\dots}$ of positive numbers, one has*

$$(5.7) \quad E((U_u)^{[m]}) \leq m \sum_{k=m}^{\infty} (k-1)^{[m-1]} \left[P(\|X^{(2k-1)}\|_\infty \geq c_k) + \frac{Dc_k}{2^{2k-1}(2k-1)!} \right],$$

PROOF. Because of Lemma 5.4 it is enough to prove that $P(U_u \geq k)$ is bounded by the expression in brackets in the right-hand side of (5.7). We have

$$P(U_u \geq k) \leq P(\|X^{(2k-1)}\|_\infty \geq c_k) + P(U_u \geq k, \|X^{(2k-1)}\|_\infty < c_k).$$

Because of Rolle's theorem:

$$\{U_u \geq k\} \subset \{N_u(X; I) \geq 2k-1\}.$$

Applying Lemma 5.2 to the function $X(\cdot) - u$ and replacing in its statement k by $2k-1$, we obtain:

$$\{U_u \geq k, \|X^{(2k-1)}\|_\infty < c_k\} \subset \{|X_{1/2} - u| \leq \frac{c_k}{2^{2k-1}(2k-1)!}\}.$$

The remainder is plain. \square

PROOF OF THEOREM 5.1. Using Lemma 5.5 and the hypothesis we obtain:

$$\frac{\nu_m}{m!} \leq \frac{1}{m!} \sum_{k=m}^{\infty} k^{[m]} \gamma_m^* 2^{-(k+1)} = \frac{\gamma_m^*}{m!} 2^{-(m+1)} \left[\left(\frac{1}{1-x} \right)^{(m)} \Big|_{x=1/2} \right] = \gamma_m^*$$

Since $\tilde{\nu}_m \leq \nu_m$ we can apply Lemma 5.3 to the random variable $\xi = U_u 1_{\{X_0 \leq u\}}$ and the result follows from $\gamma_m^* \rightarrow 0$. \square

One can replace condition “ $p_{X(T/2)}(x) \leq D$ for all x ” by “ $p_{X(T/2)}(x) \leq D$ for x in some neighbourhood of u ”. In this case, the statement of Theorem 5.1 holds if one adds in (ii) that the error is bounded by $\gamma_{m_0+1}^*$ for m_0 large enough. The proof is similar.

Also, one can substitute the one-dimensional density $p_{X(T/2)}(\cdot)$ by $p_{X(t)}(\cdot)$ for some other $t \in [0, T]$, introducing into the bounds the corresponding modifications.

The application of Theorem 5.1 requires an adequate choice of the sequence $\{c_k, k = 1, 2, \dots\}$ which depends on the available description of the process X . The whole procedure will have some practical interest for the computation of $P(M > u)$ only if we get appropriate bounds for the

quantities γ_m^* and the factorial moments $\tilde{\nu}_m$ can be actually computed by means of Rice formulas (or by some other procedure). The next Theorem shows how this can be done in the case of a general class of Gaussian stationary processes.

THEOREM 5.6. *Let $\mathcal{X} = \{X(t) : t \in \mathbb{R}\}$ be Gaussian, centered and stationary, with covariance Γ normalized by $\Gamma(0) = 1$. Assume that Γ has a Taylor expansion at the origin which is absolutely convergent at $t = 2T$. Then, the conclusion of Theorem 5.1 holds true, so that the Rice series converges and $F_{M_T}(u)$ can be computed by means of (5.2).*

PROOF. Again we assume, with no loss of generality, that $T = 1$. Notice that the hypothesis implies that the spectral moments λ_k exist and are finite for every $k = 0, 1, 2, \dots$. We will obtain the result assuming:

$$H_1 : \lambda_{2k} \leq C_1(k!)^2.$$

It is easy to verify that if Γ has a Taylor expansion at zero which is absolutely convergent at $t = 2$, then H_1 holds true. (In fact, both conditions are only slightly different, since H_1 implies that the Taylor expansion of Γ at zero is absolutely convergent in $\{|t| < 2\}$). Let us check that the hypotheses of Theorem 5.1 are satisfied. First, $p_{X(1/2)}(x) \leq D = (2\pi)^{-1/2}$. Second, let us show a sequence $\{c_k\}$ that satisfies (5.1). We have

$$(5.8) \quad \begin{aligned} \mathbb{P}(\|X^{(2k-1)}\|_\infty \geq c_k) &\leq \mathbb{P}(|X^{(2k-1)}(0)| \geq c_k) + 2\mathbb{P}(U_{c_k}(X^{(2k-1)}, I) \geq 1) \\ &\leq \mathbb{P}(|Z| \geq c_k \lambda_{4k-2}^{-1/2}) + 2\mathbb{E}(U_{c_k}(X^{(2k-1)}, I)), \end{aligned}$$

where Z is standard normal. One easily checks that $\{X^{(2k-1)}(t); t \in \mathbb{R}\}$ is a Gaussian stationary centered process with covariance function $-\Gamma^{(4k-2)}(\cdot)$. So, we can use Rice formula for the expectation of the number of up-crossings of a stationary centered Gaussian process to compute the second term in the right-hand side of (5.8).

Using the inequality $1 - \Phi(x) \leq (1/x)\phi(x)$ valid for $x > 0$, one gets:

$$(5.9) \quad \mathbb{P}(\|X^{(2k-1)}\|_\infty \geq c_k) \leq \left[\sqrt{\frac{2}{\pi}} \frac{\lambda_{4k-2}^{1/2}}{c_k} + (1/\pi) \left(\frac{\lambda_{4k}}{\lambda_{4k-2}} \right)^{1/2} \right] \exp\left(-\frac{c_k^2}{2\lambda_{4k-2}}\right)$$

Choose

$$\begin{aligned} c_k &:= (B_1 k \lambda_{4k-2})^{1/2} \quad \text{if } \frac{\lambda_{4k}}{\lambda_{4k-2}} \leq B_1 k \\ c_k &:= (\lambda_{4k})^{1/2} \quad \text{if } \frac{\lambda_{4k}}{\lambda_{4k-2}} > B_1 k. \end{aligned}$$

Using hypothesis H_1), if $B_1 > 1$:

$$\mathbb{P}(\|X^{(2k-1)}\|_\infty \geq c_k) \leq \left[\sqrt{\frac{2}{\pi}} + \frac{1}{\pi} (B_1 k)^{1/2} \right] e^{-\frac{B_1 k}{2}}.$$

Finally, choosing $B_1 := 4 \log(2)$:

$$\gamma_k \leq \sqrt{\frac{2}{\pi}} (1 + 2(C_1^{1/2} + 1)k) 2^{-2k} \quad (k = 1, 2, \dots),$$

so that (5.1) is satisfied. As a by product, notice that

$$(5.10) \quad \gamma_m^* \leq \sqrt{\frac{8}{\pi}} (1 + 2(C_1^{1/2} + 1)m) 2^{-m} \quad (m = 1, 2, \dots).$$

□

REMARK 5.1. *For Gaussian processes if one is willing to use Rice formulas to compute the factorial moments $\tilde{\nu}_m$, it is enough to verify that the distribution of*

$$X(0), X(t_1), \dots, X(t_m)$$

is non-degenerate for any choice of non-zero distinct $t_1, \dots, t_m \in I$. For stationary Gaussian processes a simple sufficient condition on the spectral measure to have non-degeneracy was given in Chapter 3 (see Exercises 3.4 and 3.5).

If instead of requiring the paths of the process \mathcal{X} to be of class \mathcal{C}^∞ , one relaxes this condition up to a certain order of differentiability, one can still get upper and lower bounds for $P(M > u)$, as stated in the next Theorem.

THEOREM 5.7. *Let $X = \{X(t) : t \in I\}$ be a real -valued stochastic process. Suppose that $p_{X(t)}(x)$ is bounded for $t \in I$, $x \in \mathbb{R}$ and that the paths of \mathcal{X} are of class \mathcal{C}^{p+1} . Then*

$$\text{if } 2K + 1 < p/2 : P(M > u) \leq P(X(0) > u) + \sum_{m=1}^{2K+1} (-1)^{m+1} \frac{\tilde{\nu}_m}{m!}$$

and

$$\text{if } 2K < p/2 : P(M > u) \geq P(X(0) > u) + \sum_{m=1}^{2K} (-1)^{m+1} \frac{\tilde{\nu}_m}{m!}.$$

Notice that all the moments in the above formulas are finite.

The proof is a straightforward application of Lemma 5.3 and Theorem 3.6. When the level u is high, a first approximation is given by Proposition 4.1 which shows that only the **first** term in the Rice series takes part in the equivalent of $1 - F_{M_T}(u)$ as $u \rightarrow +\infty$.

2. Computation of Moments.

An efficient numerical computation of the factorial moments of crossings is associated to a fine description of the behavior as the k -tuple (t_1, \dots, t_k) approaches the diagonal $D_k(I)$, of the integrands

$$A_{t_1, \dots, t_k}^+(u, \dots, u) = E\left(X'^+(t_1) \dots X'^+(t_k) | X(t_1) = \dots = X(t_k) = u\right) p_{X(t_1), \dots, X(t_k)}(u, \dots, u),$$

$$(5.11) \quad \tilde{A}_{t_1, \dots, t_k}^+(u, \dots, u) = \int_{-\infty}^u$$

$$E\left(X'^+(t_1) \dots X'^+(t_k) | X(0) = x, X(t_1) = \dots = X(t_k) = u\right) p_{X(0), X(t_1), \dots, X(t_k)}(x, u, \dots, u) dx.$$

We recall that $A_{t_1, \dots, t_k}^+(u, \dots, u)$ and $\tilde{A}_{t_1, \dots, t_k}^+(u, \dots, u)$ appear respectively in Rice formulas for the k^{th} -factorial moment of up-crossings and the k^{th} -factorial moment of up-crossings with the additional condition that $X(0) \leq u$.

If the process is Gaussian, stationary and satisfies a certain number of regularity conditions, we have seen in Proposition 4.5 that:

$$(5.12) \quad A_{s,t}^+(u, u) \approx \frac{1}{1296} \frac{(\lambda_2 \lambda_6 - \lambda_4^2)^{3/2}}{(\lambda_4 - \lambda_2^2)^{1/2} \pi^2 \lambda_2^2} \exp\left(-\frac{1}{2} \frac{\lambda_4}{\lambda_4 - \lambda_2^2} u^2\right) (t-s)^4,$$

as $t-s \rightarrow 0$.

(5.12) can be extended to non-stationary Gaussian processes obtaining an equivalence of the form:

$$(5.13) \quad A_{s,t}^+(u, u) \approx J(\tilde{t})(t-s)^4 \quad \text{as } s, t \rightarrow \tilde{t}$$

where $J(\tilde{t})$ is a continuous non-zero function of \tilde{t} depending on u , which can be expressed in terms of the mean and covariance functions of the process and its derivatives. We give a proof of an equivalence of the form (5.13) in the next proposition.

One can profit of this equivalence to improve the numerical methods to compute $\tilde{\nu}_2$ (the second factorial moment of the number of up-crossings with the restriction $X(0) \leq u$). Equivalence formulas such as (5.12) or (5.13) can be used to avoid numerical degeneracies near the diagonal $D_2(I)$. Notice that even in case the process \mathcal{X} is stationary at the departure, under conditioning

on $X(0)$, the process that must be taken into account in the computation of the factorial moments of up-crossings for the Rice series (5.2) will be non-stationary, so that equivalence (5.13) is the appropriate tool for our main purpose here.

PROPOSITION 5.8. *Suppose that \mathcal{X} is a Gaussian process with C^5 paths and that for each $t \in I$ the joint distribution of $X(t), X'(t), X^{(2)}(t), X^{(3)}(t)$ does not degenerate. Then (5.13) holds true.*

PROOF. We give the general scheme of the proof and leave to the reader the detailed computations. Denote by $\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}$ a two-dimensional random vector having as probability distribution the conditional distribution of $\begin{pmatrix} X'(s) \\ X'(t) \end{pmatrix}$ given $X(s) = X(t) = u$.

One has:

$$(5.14) \quad A_{s,t}^+(u, u) = \mathbb{E}(\xi_1^+ \xi_2^+) p_{X(s), X(t)}(u, u)$$

Put $\tau = t - s$ and check the following Taylor expansions around the point s :

$$(5.15) \quad \mathbb{E}(\xi_1) = m_1\tau + m_2\tau^2 + L_1\tau^3$$

$$(5.16) \quad \mathbb{E}(\xi_2) = -m_1\tau + m_2'\tau^2 + L_2\tau^3$$

$$(5.17) \quad \text{Var}(\xi) = \begin{pmatrix} a\tau^2 + b\tau^3 + c\tau^4 + \rho_{11}\tau^5 & -a\tau^2 - \frac{b+b'}{2}\tau^3 + d\tau^4 + \rho_{12}\tau^5 \\ -a\tau^2 - \frac{b+b'}{2}\tau^3 + d\tau^4 + \rho_{12}\tau^5 & a\tau^2 + b'\tau^3 + c'\tau^4 + \rho_{22}\tau^5 \end{pmatrix}$$

where $m_1, m_2, m_2', a, b, c, d, a', b', c'$ are continuous functions of s and $L_1, L_2, \rho_{11}, \rho_{12}, \rho_{22}$ are bounded functions of s and t . (5.15), (5.16) and (5.17) follow directly from the regression formulas of the pair $\begin{pmatrix} X'(s) \\ X'(t) \end{pmatrix}$ on the condition $X(s) = X(t) = u$.

Notice that (as in Belayev, 1966 or Azais and Wschebor, 2002)

$$\text{Var}(\xi_1) = \frac{\det[\text{Var}(X(s), X(t), X'(s))^T]}{\det[\text{Var}(X(s), X(t))^T]} = \frac{\det[\text{Var}(X(s), X'(s), X(t) - X(s) - (t-s)X'(s))^T]}{\det[\text{Var}(X(s), X(t) - X(s))^T]}$$

A direct computation gives:

$$(5.18) \quad \det[\text{Var}(X(s), X(t))^T] \approx \tau^2 \det[\text{Var}(X(s), X'(s))^T]$$

$$\text{Var}(\xi_1) \approx \frac{1}{4} \frac{\det[\text{Var}(X(s), X'(s), X^{(2)}(s))^T]}{\det[\text{Var}(X(s), X'(s))^T]} \tau^2$$

where \approx denotes equivalence as $\tau \rightarrow 0$. So,

$$a = \frac{1}{4} \frac{\det[\text{Var}(X(s), X'(s), X^{(2)}(s))^T]}{\det[\text{Var}(X(s), X'(s))^T]}$$

which is a continuous non-vanishing function for $s \in I$. Notice that the coefficient of τ^3 in the Taylor expansion of $\text{Cov}(\xi_1, \xi_2)$ is equal to $-\frac{b+b'}{2}$. This follows either by direct computation or taking into account that $\det[\text{Var}(\xi)]$ is a symmetric function of the pair s, t . Put

$$\Delta(s, t) = \det[\text{Var}(\xi)]$$

The behavior of $\Delta(s, t)$ as $s, t \rightarrow \tilde{t}$ can be obtained from

$$\Delta(s, t) = \frac{\det[\text{Var}(X(s), X(t), X'(s), X'(t))^T]}{\det[\text{Var}(X(s), X(t))^T]}$$

applying the more general Proposition 5.9 below, which provides an equivalent for the numerator (or use Lemma 4.3, p.76 in Piterbarg (1996) which is sufficient in the present case). We get:

$$(5.19) \quad \Delta(s, t) \approx \overline{\Delta}(\tilde{t})\tau^6$$

where

$$\overline{\Delta}(\tilde{t}) = \frac{1}{144} \frac{\det[\text{Var}(X(\tilde{t}), X'(\tilde{t}), X^{(2)}(\tilde{t}), X^{(3)}(\tilde{t}))^T]}{\det[\text{Var}(X(\tilde{t}), X'(\tilde{t}))^T]}$$

The non degeneracy hypothesis implies that $\bar{\Delta}(\tilde{t})$ is continuous and non zero. One has:

$$(5.20) \quad \mathbb{E}(\xi_1^+ \xi_2^+) = \frac{1}{2\pi [\Delta(s, t)]^{1/2}} \int_0^{+\infty} \int_0^{+\infty} xy \exp \left[-\frac{1}{2\Delta(s, t)} F(x, y) \right] dx dy$$

where

$$F(x, y) = \text{Var}(\xi_2)(x - \mathbb{E}(\xi_1))^2 + \text{Var}(\xi_1)(y - \mathbb{E}(\xi_2))^2 - 2\text{Cov}(\xi_1, \xi_2)(x - \mathbb{E}(\xi_1))(y - \mathbb{E}(\xi_2))$$

Substituting the expansions (5.15), (5.16), (5.17) in the integrand of (5.20) and making the change of variables $x = \tau^2 v$, $y = \tau^2 w$, we get, as $s, t \rightarrow \tilde{t}$:

$$(5.21) \quad \mathbb{E}(\xi_1^+ \xi_2^+) \approx \frac{\tau^5}{2\pi [\bar{\Delta}(\tilde{t})]^{1/2}} \int_0^{+\infty} \int_0^{+\infty} vw \exp \left[-\frac{1}{2\bar{\Delta}(\tilde{t})} \bar{F}(v, w) \right] dv dw$$

$\bar{\Delta}(\tilde{t})$ can also be expressed in terms of the functions a, b, c, d, a', b', c' :

$$\bar{\Delta}(\tilde{t}) = ac' + ca' + 2ad - \left(\frac{b - b'}{2} \right)^2$$

and

$$\bar{F}(v, w) = a(v - m_2 + w - m'_2)^2 + m_1^2(c + c' + 2d) - m_1(b - b')(v + w - m_2 - m'_2)$$

The functions $a, b, c, d, b', c', m_1, m_2$ which appear in these formulas are all evaluated at the point \tilde{t} . Replacing (5.21) and (5.18) into (5.14) one gets (5.13). \square

For $k \geq 3$, the general behavior of the functions $\tilde{A}_{t_1, \dots, t_k}(u, \dots, u)$ and $A_{t_1, \dots, t_k}^+(u, \dots, u)$ when (t_1, \dots, t_k) approaches the diagonal is not known. Proposition 5.10 below, even though it contains some restrictions (it requires $\mathbb{E}(X(t)) = 0$ and $u = 0$) can be applied to improve the efficiency in the computation of the k^{th} -factorial moments by means of a Monte-Carlo method, via the use of important sampling. More precisely, this proposition can be used when computing the integral of $A_{t_1, \dots, t_k}^+(u, \dots, u)$ over I^k in the following way: instead of choosing at random the point (t_1, t_2, \dots, t_k) in the cube I^k with a uniform distribution, we should do it with a probability law having a density proportional to the function $\prod_{1 \leq i < j \leq k} (t_j - t_i)^4$. For the proof of Proposition 5.10 we will use the following auxiliary one, which has its own interest.

PROPOSITION 5.9. *Let $\mathcal{X} = \{X(t) : t \in I\}$ be a Gaussian process defined on the compact interval I of the real line, k an integer $k \geq 2$, and $t_1, \dots, t_k \in I$. When the paths of the process \mathcal{X} are of class \mathcal{C}^1 , we denote:*

$$D_l(t) = \det \left[\text{Var} \left(X(t), X'(t), \dots, X^{(l)}(t) \right)^T \right].$$

(i) *If the paths of the process \mathcal{X} are of class \mathcal{C}^{k-1} and $t_1, t_2, \dots, t_k \rightarrow t^*$, then*

$$(5.22) \quad \det [\text{Var}(X(t_1), X(t_2), \dots, X(t_k))^T] \approx \frac{1}{[2! \dots (k-1)!]^2} \left[\prod_{1 \leq i < j \leq k} (t_j - t_i)^2 \right] \cdot D_{k-1}(t^*).$$

(ii) *If the paths of \mathcal{X} are of class \mathcal{C}^{2k-1} and $t_1, t_2, \dots, t_k \rightarrow t^*$, then*

$$(5.23) \quad \Delta = \det [\text{Var}(X(t_1), X'(t_1), \dots, X(t_k), X'(t_k))^T] \approx \frac{1}{[2! \cdot 3! \dots (2k-1)!]^2} \prod_{1 \leq i < j \leq k} (t_j - t_i)^8 D_{2k-1}(t^*)$$

PROOF. We prove (ii). The proof of (i) can be done along the same lines as the one of (ii). It is in fact simpler and left to the reader.

With no loss of generality, we may assume that t_1, t_2, \dots, t_k are pairwise different. Suppose $f : I \rightarrow \mathbb{R}$ is a function of class \mathcal{C}^{2m-1} , $m \geq 2$, $m \geq 1$. We use the following notations for interpolating polynomials: $P_m(t; f)$ is the polynomial of degree $2m - 1$ such that

$$P_m(t_j; f) = f(t_j) \quad \text{and} \quad P'_m(t; f) = f'(t_j) \quad \text{for } j = 1, \dots, m.$$

$Q_m(t; f)$ is the polynomial of degree $2m - 2$ such that

$$Q_m(t_j; f) = f(t_j) \quad \text{for } j = 1, \dots, m; \quad Q'_m(t; f) = f'(t_j) \quad \text{for } j = 1, \dots, m - 1.$$

From Lemma 5.2 we know that

$$(5.24) \quad f(t) - P_m(t; f) = \frac{1}{(2m)!} (t - t_1)^2 \dots (t - t_m)^2 f^{(2m)}(\xi)$$

$$(5.25) \quad f(t) - Q_m(t; f) = \frac{1}{(2m-1)!} (t - t_1)^2 \dots (t - t_{m-1})^2 (t - t_m) f^{(2m-1)}(\eta)$$

where

$$\xi = \xi(t_1, t_2, \dots, t_m, t), \eta = \eta(t_1, t_2, \dots, t_m, t)$$

and

$$\min(t_1, t_2, \dots, t_m, t) \leq \xi, \eta \leq \max(t_1, t_2, \dots, t_m, t).$$

The function

$$g(t) = f^{(2m-1)}(\eta(t_1, t_2, \dots, t_m, t)) = \frac{(2m-1)! [f(t) - Q_m(t; f)]}{(t - t_1)^2 \dots (t - t_{m-1})^2 (t - t_m)}$$

is differentiable at the point $t = t_m$ and differentiating in (5.25):

$$(5.26) \quad f'(t_m) - Q'_m(t_m; f) = \frac{1}{(2m-1)!} (t_m - t_1)^2 \dots (t_m - t_{m-1})^2 f^{(2m-1)}(\eta(t_1, t_2, \dots, t_m, t_m))$$

Put

$$\xi_m = \xi(t_1, t_2, \dots, t_m, t_m), \eta_m = \eta(t_1, t_2, \dots, t_m, t_m).$$

Since $P_m(t; f)$ is a linear functional of

$$(f(t_1), \dots, f(t_m), f'(t_1), \dots, f'(t_m))$$

and $Q_m(t; f)$ is a linear functional of

$$(f(t_1), \dots, f(t_m), f'(t_1), \dots, f'(t_{m-1}))$$

with coefficients depending (in both cases) only on t_1, t_2, \dots, t_m, t , it follows that:

$$\begin{aligned} \Delta &= \det \left[\text{Var}(X(t_1), X'(t_1), X(t_2) - P_1(t_2; X), X'(t_2) - Q'_2(t_2, X), \right. \\ &\quad \left. \dots, X(t_k) - P_{k-1}(t_k; X), X'(t_k) - Q'_k(t_k; X)) \right]^T = \\ &= \det \left[\text{Var}(X(t_1), X'(t_1), \frac{1}{2!} (t_2 - t_1)^2 X^{(2)}(\xi_1), \frac{1}{3!} (t_2 - t_1)^2 X^{(3)}(\eta_2), \right. \\ &\quad \left. \dots, \frac{1}{(2k-2)!} (t_k - t_1)^2 \dots (t_k - t_{k-1})^2 X^{(2k-2)}(\xi_{k-1}), \frac{1}{(2k-1)!} (t_k - t_1)^2 \dots (t_k - t_{k-1})^2 X^{(2k-1)}(\eta_{k-1})) \right]^T \\ &= \frac{\tilde{\Delta}}{[2! \dots (2k-1)!]^2} \prod_{1 \leq i < j \leq k} (t_j - t_i)^8 \end{aligned}$$

with

$$\begin{aligned} \tilde{\Delta} &= \det \left[\text{Var}(X(t_1), X'(t_1), X^{(2)}(\xi_1), X^{(3)}(\eta_2), \dots, X^{(2k-2)}(\xi_{k-1}), X^{(2k-1)}(\eta_{k-1})) \right]^T \rightarrow \\ &\rightarrow \det \left[\text{Var}(X(t^*), X'(t^*), \dots, X^{(2k-1)}(t^*)) \right]^T = D_{2k-1}(t^*) \end{aligned}$$

as $t_1, t_2, \dots, t_k \rightarrow t^*$. This proves (5.23). \square

PROPOSITION 5.10. *Suppose that \mathcal{X} is a centered Gaussian process with C^{2k-1} paths and that for each pairwise distinct values of the parameter $t_1, t_2, \dots, t_k \in I$ the joint distribution of $(X(t_h), X'(t_h), \dots, X^{(2k-1)}(t_h), h = 1, 2, \dots, k)$ is non-degenerate. Then, as $t_1, t_2, \dots, t_k \rightarrow t^*$:*

$$A_{t_1, \dots, t_k}^+(0, \dots, 0) \approx J_k(t^*) \prod_{1 \leq i < j \leq k} (t_j - t_i)^4$$

where $J_k(t)$ is a continuous non-zero function of t .

PROOF. For k distinct values t_1, t_2, \dots, t_k , let $Z = (Z_1, \dots, Z_k)^T$ be a random vector having the conditional distribution of $(X'(t_1), \dots, X'(t_k))^T$ given $X(t_1) = X(t_2) = \dots = X(t_k) = 0$. The (Gaussian) distribution of Z is centered and we denote its covariance matrix by Σ . Also put:

$$\Sigma^{-1} = \frac{1}{\det(\Sigma)} (\sigma^{ij})_{i,j=1,\dots,k}$$

σ^{ij} being the cofactor of the position (i, j) in the matrix Σ . Then, one can write:

$$(5.27) \quad A_{t_1, \dots, t_k}^+(0, \dots, 0) = \mathbb{E} \{ Z_1^+ \dots Z_k^+ \} \cdot p_{X(t_1), \dots, X(t_k)}(0, \dots, 0)$$

and

$$(5.28) \quad A_{t_1, \dots, t_k}^+(0, \dots, 0) = \frac{1}{(2\pi)^{\frac{k}{2}} (\det(\Sigma))^{\frac{1}{2}}} \int_{(R^+)^k} x_1 \dots x_k \exp \left[-\frac{F(x_1, \dots, x_k)}{2 \cdot \det(\Sigma)} \right] dx_1 \dots dx_k$$

where

$$F(x_1, \dots, x_k) = \sum_{i,j=1}^k \sigma^{ij} x_i x_j.$$

Letting $t_1, t_2, \dots, t_k \rightarrow t^*$ and using (5.23) and (5.22) we get:

$$\begin{aligned} \det(\Sigma) &= \frac{\det[\text{Var}(X(t_1), X'(t_1), \dots, X(t_k), X'(t_k))^T]}{\det[\text{Var}(X(t_1), \dots, X(t_k))^T]} \\ &\approx \frac{1}{[k! \dots (2k-1)!]^2} \left[\prod_{1 \leq i < j \leq k} (t_j - t_i)^6 \right] \cdot \frac{D_{2k-1}(t^*)}{D_{k-1}(t^*)}. \end{aligned}$$

We consider now the behaviour of the σ^{ij} ($i, j = 1, \dots, k$). Let us first look at σ^{11} . Using the same method as above, now applied to the cofactor of the position $(1, 1)$ in Σ , one has:

$$\begin{aligned} \sigma^{11} &= \frac{\det[\text{Var}(X(t_1), X(t_2), \dots, X(t_k), X'(t_2), \dots, X'(t_k))^T]}{\det[\text{Var}(X(t_1), \dots, X(t_k))^T]} \\ &\approx \frac{[2! \dots (2k-2)!]^{-2} \left[\prod_{2 \leq i < j \leq k} (t_j - t_i)^8 \right] \left[\prod_{2 \leq h \leq k} (t_1 - t_h)^4 \right] D_{2k-2}(t^*)}{[2! \dots (k-1)!]^{-2} \left[\prod_{1 \leq i < j \leq k} (t_j - t_i)^2 \right] D_{k-1}(t^*)} \\ &= [k! \dots (2k-2)!]^{-2} \left[\prod_{2 \leq i < j \leq k} (t_j - t_i)^6 \right] \left[\prod_{2 \leq h \leq k} (t_1 - t_h)^2 \right] \frac{D_{2k-2}(t^*)}{D_{k-1}(t^*)} \end{aligned}$$

A similar computation holds for σ^{ii} , $i = 2, \dots, k$.

Consider now σ^{12} . One has:

$$\begin{aligned} \sigma^{12} &= - \frac{\det[\mathbb{E} \{ (X(t_1), X(t_2), \dots, X(t_k), X'(t_2), \dots, X'(t_k))^T \cdot (X(t_1), X(t_2), \dots, X(t_k), X'(t_1), X'(t_3), \dots, X'(t_k)) \}]}{\det[\text{Var}(X(t_1), \dots, X(t_k))^T]} \\ &= \frac{\det[\mathbb{E} \{ (X(t_2), X'(t_2), \dots, X(t_k), X'(t_k), X(t_1))^T \cdot (X(t_1), X'(t_1), X(t_3), X'(t_3), \dots, X(t_k), X'(t_k), X(t_2)) \}]}{\det[\text{Var}(X(t_1), \dots, X(t_k))^T]} \\ &\approx \frac{1}{[k! \dots (2k-2)!]^2} \left[\prod_{3 \leq i < j \leq k} (t_j - t_i)^6 \right] \left[\prod_{3 \leq h \leq k} (t_1 - t_h)^4 (t_2 - t_h)^4 \right] (t_2 - t_1)^2 \cdot \frac{D_{2k-2}(t^*)}{D_{k-1}(t^*)} \end{aligned}$$

A similar computation applies to all the cofactors σ^{ij} , $i \neq j$.

Perform in the integral in (5.28) the change of variables

$$x_j = \left[\prod_{i=1, i \neq j}^{i=k} (t_i - t_j)^2 \right] \cdot y_j \quad j = 1, \dots, k$$

and the integral becomes:

$$\left[\prod_{1 \leq i < j \leq k} (t_j - t_i)^8 \right] \int_{(R^+)^k} y_1 \dots y_k \exp \left[-\frac{1}{2 \cdot \det(\Sigma)} G(y_1, \dots, y_k) \right] dy_1 \dots dy_k$$

where

$$G(y_1, \dots, y_k) = \sum_{i,j=1}^k \sigma^{ij} \left[\prod_{h=1, h \neq i}^{h=k} (t_h - t_i)^2 \right] \left[\prod_{h=1, h \neq j}^{h=k} (t_h - t_j)^2 \right] y_i y_j.$$

so that, as $t_1, t_2, \dots, t_k \rightarrow t^*$

$$\frac{G(y_1, \dots, y_k)}{\det(\Sigma)} \approx [(2k-1)!]^2 \frac{D_{2k-2}(t^*)}{D_{2k-1}(t^*)} \left(\sum_{i=1}^{i=k} y_i \right)^2.$$

Now, passage to the limit under the integral sign in (5.28), which is easily justified by application of the Lebesgue Theorem, leads to

$$E \{ Z_1^+ \dots Z_k^+ \} \approx \frac{1}{(2\pi)^{\frac{k}{2}}} k! \dots (2k-1)! \left[\prod_{1 \leq i < j \leq k} |t_j - t_i|^5 \right] \left(\frac{D_{k-1}(t^*)}{D_{2k-1}(t^*)} \right)^{\frac{1}{2}} I_k(\alpha^*)$$

where $I_k(\alpha)$, $\alpha > 0$ is

$$I_k(\alpha) = \int_{(R^+)^k} y_1 \dots y_k \exp \left[-\frac{\alpha}{2} \left(\sum_{i=1}^{i=k} y_i \right)^2 \right] dy_1 \dots dy_k = \frac{1}{\alpha^k} I_k(1)$$

and

$$\alpha^* = [(2k-1)!]^2 \frac{D_{2k-2}(t^*)}{D_{2k-1}(t^*)}$$

Replacing into (5.27) one gets the result with

$$J_k(t) = \frac{2! \dots (2k-2)!}{[2\pi(2k-1)!]^{2k-1}} \frac{I_k(1)}{[D_{2k-1}(t)]^{\frac{1}{2}}} \left[\frac{D_{2k-1}(t)}{D_{2k-2}(t)} \right]^k$$

This finishes the proof. \square

3. Numerical aspects of Rice Series.

Let us compare the numerical computation based upon Theorem 5.1 with the Monte-Carlo method based on the simulation of the paths. We do this for stationary Gaussian processes that satisfy the hypotheses of Theorem 5.6 and also the non-degeneracy condition that ensures that one is able to compute the factorial moments of crossings by means of Rice formulas.

Suppose that we want to compute $P(M > u)$ with an error bounded by δ , where $\delta > 0$ is a given positive number.

To proceed by simulation, we discretize the paths by means of a uniform partition $\{t_j := j/n, j = 0, 1, \dots, n\}$. Denote

$$M^{(n)} := \sup_{0 \leq j \leq n} X(t_j).$$

Using Taylor's formula at the time where the maximum M of $X(\cdot)$ occurs, one gets :

$$0 \leq M - M^{(n)} \leq \|X''\|_{\infty} / (2n^2).$$

It follows that

$$0 \leq P(M > u) - P(M^{(n)} > u) = P(M > u, M^{(n)} \leq u) \leq P(u < M \leq u + \|X''\|_{\infty} / (2n^2)).$$

Let us admit that the distribution of M has a locally bounded density (see Ylvisaker's Theorem 1.22). The above suggests that a number of $n = (\text{const}) \delta^{-1/2}$ points is required if one wants the error $P(M > u) - P(M^{(n)} > u)$ to be bounded by δ .

On the other hand, to estimate $P(M^{(n)} > u)$ by Monte-Carlo with a **mean square error** smaller than δ , we require the simulation of $N = (\text{const}) \delta^{-2}$ Gaussian n -tuples $(X_{t_1}, \dots, X_{t_n})$

from the distribution determined by the given stationary process. Performing each simulation demands $(const) n \log(n)$ elementary operations (see for example Dietrich and Newsam, 1997 for this computational point). Summing up, the total mean number of elementary operations required to get a **mean square error** bounded by δ in the estimation of $P(M > u)$ has the form $(const) \delta^{-5/2} \log(1/\delta)$.

Suppose now that we apply Theorem 5.1 to a Gaussian stationary centered process verifying the hypotheses of Theorem 5.6 and the non-degeneracy condition. The bound for γ_m^* in Equation (5.10) implies that computing a partial sum with $(const) \log(1/\delta)$ terms assures that the tail in the Rice series is bounded by δ . If one computes each $\tilde{\nu}_m$ by means of a Monte-Carlo method for the multiple integrals appearing in the Rice formulas, then the number of elementary operations for the whole procedure will have the form $(const) \delta^{-2} \log(1/\delta)$. Hence, this is better than simulation as δ tends to zero.

As usual, for given $\delta > 0$, the value of the generic constants decides the comparison between both methods, and these are very difficult to estimate for a general class of processes.

More important is the fact that the enveloping property of the Rice series implies that the actual number of terms required by the application of Theorem 5.1 can be **much** smaller than the one resulting from the a priori bound for γ_m^* . More precisely, suppose that we have obtained each numerical approximation $\tilde{\nu}_m^*$ of $\tilde{\nu}_m$ with a precision η

$$|\tilde{\nu}_m^* - \tilde{\nu}_m| \leq \eta,$$

and that we stop when

$$(5.29) \quad \frac{\tilde{\nu}_{m_0+1}^*}{(m_0+1)!} \leq \eta.$$

Then, it follows that

$$\left| \sum_{m=1}^{\infty} (-1)^{m+1} \frac{\tilde{\nu}_m}{m!} - \sum_{m=1}^{m_0} (-1)^{m+1} \frac{\tilde{\nu}_m^*}{m!} \right| \leq (e+1)\eta.$$

Putting $\eta = \delta/(e+1)$, we get the desired bound. In other words one can profit of the successive numerical approximations of $\tilde{\nu}_m$ to determine a new m_0 which turns out to be - in certain interesting examples - much smaller than the one deduced from the a priori bound on γ_m^* .

Next, we give the results of the evaluation of $P(M_T > u)$ using up to three terms in the Rice series in a certain number of typical cases. We compare these results with the classical evaluation given by Proposition 4.1 For fixed T and high level u this bound is sharp. But when both T and u are fixed, the situation becomes essentially different and using more than one term of the Rice series supplies a remarkable improvement in the computation.

We consider several stationary centered Gaussian processes listed in the following table, where the covariances and the corresponding spectral densities are indicated.

process	covariance	spectral density
X_1	$\Gamma_1(t) = \exp(-t^2/2)$	$f_1(x) = (2\pi)^{-1/2} \exp(-x^2/2)$
X_2	$\Gamma_2(t) = (\cosh(t))^{-1}$	$f_2(x) = (2 \cosh((\pi x)/2))^{-1}$
X_3	$\Gamma_3(t) = (3^{1/2}t)^{-1} \sin(3^{1/2}t)$	$f_3(x) = 12^{-1/2} \mathbf{1}_{\{-\sqrt{3} < x < \sqrt{3}\}}$
X_4	$\Gamma_4(t) = e^{- \sqrt{5}t } (\frac{\sqrt{5}}{3} t ^3 + 2t^2 + \sqrt{5} t + 1)$	$f_4(x) = \frac{10^4}{\sqrt{5}\pi} (5 + x^2)^{-4}$

In all cases, $\lambda_0 = \lambda_2 = 1$ to be able to compare the various results. Notice that Γ_1 and Γ_3 have analytic extensions to the whole plane, so that Theorem 5.6 applies to the processes X_1 and X_3 . On the other hand, even though all spectral moments of the process X_2 are finite, Theorem 5.6 applies only for a length less than $\pi/4$ since the meromorphic extension of $\Gamma_2(\cdot)$ has poles at the points $i\pi/2 + k\pi i$, k an integer. With respect to $\Gamma_4(\cdot)$ notice that it is obtained as the convolution $\Gamma_5 * \Gamma_5 * \Gamma_5 * \Gamma_5$ where $\Gamma_5(t) := e^{-|t|}$ is the covariance of the Ornstein-Uhlenbeck process, plus a change of scale to get $\lambda_0 = \lambda_2 = 1$. The process X_4 has $\lambda_6 < \infty$ and $\lambda_8 = \infty$ and its paths are C^3 . For the processes X_2 and X_4 we apply Theorem 5.7 to compute $F(T, u)$.

u	Length of the time interval T				
	1	4	6	8	10
-2	0.99	1.00	1.00	1.00	1.00
	0.99	1.00	1.00	1.00	1.00
	1.00	1.00	1.00	1.00	1.00
	0.99	1.00	1.00	1.00	1.00
-1	0.93	1.00	1.00	0.99-1.00	0.98-1.00
	0.93	0.99	1.00	0.99-1.00	0.98-1.00
	0.93	1.00	1.00	1.00	0.99
	0.93	1.00	1.00	0.99-1.00	0.98-1.00
0	0.65	0.90	0.95	0.95-0.99	0.90-1.00
	0.65	0.89	0.94-0.95	0.93-0.99	0.87-1.00
	0.656	0.919	0.97	0.98-0.99	0.92-1.00
	0.65	0.89	0.94-0.95	0.94-0.99	0.88-1.00
1	0.25	0.49	0.61	0.69-0.70	0.74-0.77
	0.25	0.48	0.58	0.66-0.68	0.70-0.76
	0.26	0.51	0.62	0.71	0.76-0.78
	0.25	0.48	0.59	0.67-0.69	0.72-0.77
2	0.04	0.11	0.15	0.18	0.22
	0.04	0.11	0.14	0.18	0.21
	0.04	0.11	0.15	0.19	0.22
	0.04	0.11	0.14	0.18	0.22
3	0.00	0.01	0.01	0.02	0.02
	0.00	0.01	0.01	0.02	0.02
	0.00	0.01	0.01	0.02	0.02
	0.00	0.01	0.01	0.02	0.02

TABLE 5.1. Values of $P(M > u)$ for the different processes. Each cell contains, from top to bottom, the values corresponding to stationary centered Gaussian processes with covariances $\Gamma_1, \Gamma_2, \Gamma_3$ and Γ_4 respectively. The calculation uses three terms of the Rice series for the upper-bound and two terms for the lower-bound. Both are rounded up to two decimals and when they differ, both are displayed.

Table 5.1 contains the results for $T = 1, 4, 6, 8, 10$ and the values $u = -2, -1, 0, 1, 2, 3$ using three terms of the Rices series. A single value is given when a precision of 10^{-2} is met; otherwise the lower-bound and the upper-bound given by two or three terms of the Rices series respectively, are displayed. The calculation uses a deterministic evaluation of the first two moments $\tilde{\nu}_1$ and $\tilde{\nu}_2$ using program written by Cierco-Ayrolles, Croquette and Delmas (2003) and a Monte-Carlo evaluation of $\tilde{\nu}_3$. In fact, for simpler and faster calculation, ν_3 has been evaluated instead of $\tilde{\nu}_3$ providing a slightly weaker bound.

In addition Figures 5.1 to 5.4 show the behavior of 4 bounds: from the highest to the lowest

- The Davies' bound (D) defined by Proposition 4.1
- One, three, or two terms of the Rice series ($R1, R3, R2$ in the sequel) that is

$$P(X(0) > u) + \sum_{m=1}^K (-1)^{m+1} \frac{\tilde{\nu}_m}{m!}$$

with $K = 1, 3$ or 2 .

Notice that the bound D differs from $R1$ due to the difference between $\tilde{\nu}_1$ and ν_1 . These bounds are evaluated for $T = 4, 6, 8, 10, 15$ and also for $T = 20$ and $T = 40$ when they fall in the range $[0, 1]$. Between these values, ordinary spline interpolation has been performed.

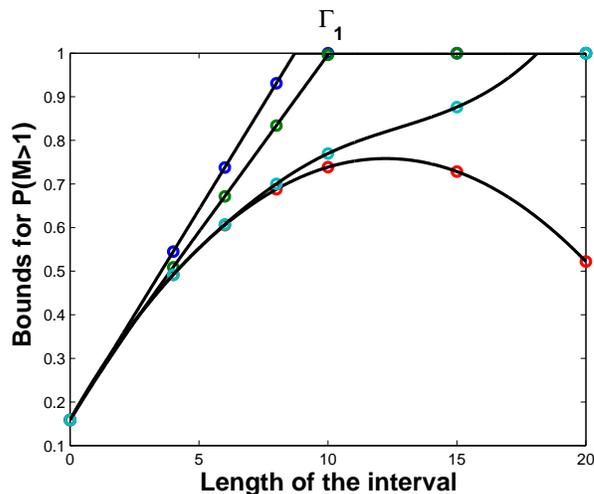


FIGURE 5.1. For the process with covariance Γ_1 and the level $u = 1$, representation of the three upper-bounds D , $R1$, $R3$ and the lower-bound $R2$ (from top to bottom) as a function of the length T of the interval

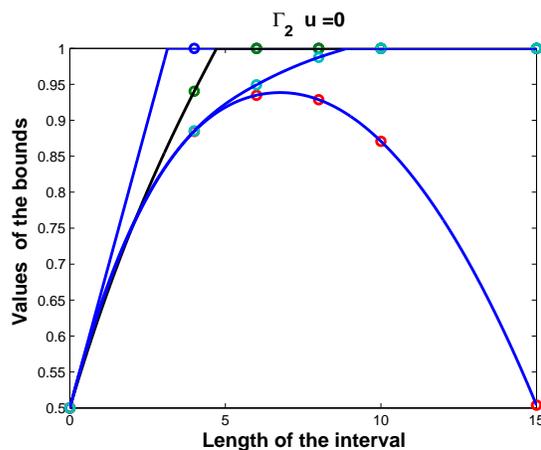


FIGURE 5.2. For the process with covariance Γ_2 and the level $u = 0$, representation of the three upper-bounds D , $R1$, $R3$ and the lower-bound $R2$ (from top to bottom) as a function of the length T of the interval

We illustrate the complete detailed calculation in three cases. They correspond to zero and positive levels u . For u negative, it is easy to check that the Davies bound is often greater than 1, thus non informative.

- For $u = 0$, $T = 6$, $\Gamma = \Gamma_1$, we have $P(X(0) > u) = 0.5$, $\tilde{\nu}_1 = 0.955$, $\tilde{\nu}_1 = 0.602$, $\tilde{\nu}_2/2 = .150$, $\tilde{\nu}_3/6 = 0.004$, so that:

$$D = 1.455, R1 = 1.103, R3 = 0.957, R2 = 0.953$$

$R2$ and $R3$ give a rather good evaluation of the probability. The Davies bound gives no information.

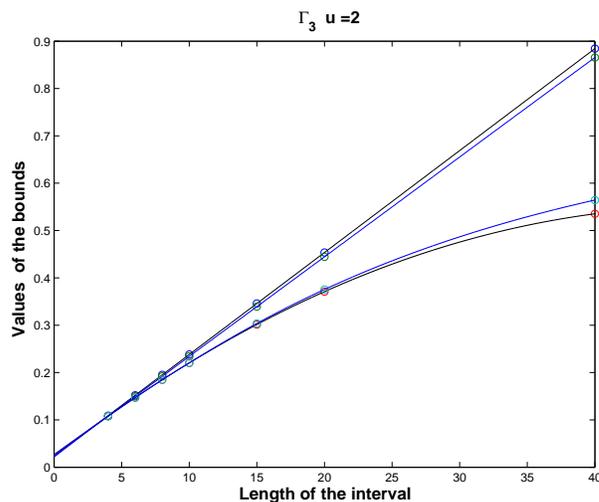


FIGURE 5.3. For the process with covariance Γ_3 and the level $u = 2$, representation of the three upper-bounds D, R1, R3 and the lower-bound R2 (from top to bottom) as a function of the length T of the interval

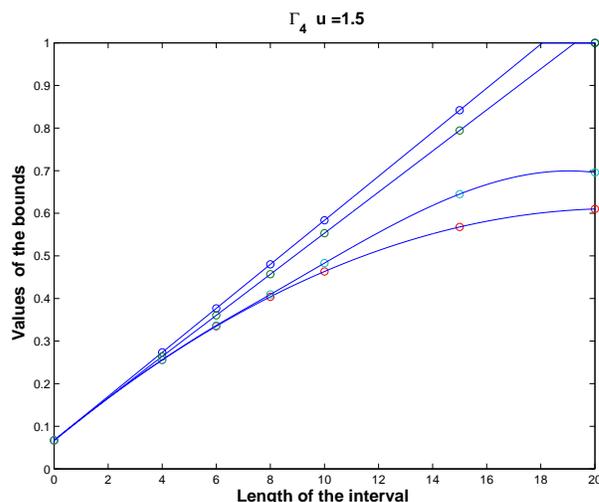


FIGURE 5.4. For the process with covariance Γ_4 and the level $u = 1.5$, representation of the three upper-bounds D, R1, R3 and the lower-bound R2 (from top to bottom) as a function of the length T of the interval

- For $u = 1.5$, $T = 15$, $\Gamma = \Gamma_2$, we have $P(X_0 > u) = 0.067$, $\nu_1 = 0.517$, $\tilde{\nu}_1 = 0.488$, $\tilde{\nu}_2/2 = 0.08$, $\nu_3/6 = 0.013$, so that:

$$D = 0.584, \quad R1 = 0.555, \quad R3 = 0.488, \quad R2 = 0.475$$

In this case the Davies bound is not sharp and a very clear improvement is provided by $R2$ and $R3$.

- For $u = 2$, $T = 10$, $\Gamma = \Gamma_3$, we have $P(X_0 > u) = 0.023$, $\tilde{\nu}_1 = 0.215$, $\nu_1 = 0.211$, $\tilde{\nu}_2/2 = 0.014$, $\nu_3/6 = 3.10^{-4}$, so that:

$$D = 0.238, \quad R1 = 0.234, \quad R3 = 0.220, \quad R2 = 0.220$$

In this case the Davies bound is rather sharp.

As a conclusion, these numerical results show that it is worth using several terms of the Rice series. In particular the first three terms are relatively easy to compute and provide a good evaluation of the distribution of M under a rather broad set of conditions.

4. Processes with Continuous Paths

This subsection is devoted to a modification of Theorem 5.1 to include processes which do not have sufficiently differentiable paths. This is done using a regularization of the paths by convolution with a deterministic approximation of unity. For simplicity, we will limit ourselves to the case of Gaussian kernels. Other kernels can be employed in a similar way.

Suppose $\mathcal{X} = \{X(t) : t \in [0, 1]\}$, is a stochastic process with continuous paths. Let ε be a positive real number, we define

$$(5.30) \quad X^\varepsilon(t) := (\phi_\varepsilon * X(\cdot))(t) = \int_{-\infty}^{+\infty} \phi_\varepsilon(t-s)X(s)ds,$$

where

$$\phi_\varepsilon(t) := (2\pi)^{-1/2}(\varepsilon)^{-1}e^{-t^2/2\varepsilon^2}, \quad t \in \mathbb{R}$$

and in (5.30) we have extended $X(\cdot)$ by $X(0)$ (respectively $X(1)$) for $t \leq 0$ (respectively $t \geq 1$). Denote by $M^\varepsilon, \tilde{\nu}_m^\varepsilon, \dots$ the analogous to $M, \tilde{\nu}_m, \dots$ for the process $\mathcal{X}^\varepsilon = \{X^\varepsilon(t), t \in [0, 1]\}$ instead of \mathcal{X} .

THEOREM 5.11. *With the above notations, suppose that the following conditions are satisfied:*

- a) $p_{X^\varepsilon(1/2)}(x)$ is bounded by a constant D_1 for ε small enough.
- b) $E(\|X\|_\infty) < \infty$.
- c) *The distribution of M has no atoms.*

Then:

(i)

$$(5.31) \quad P(M > u) = P(X(0) > u) + \lim_{\varepsilon \rightarrow 0} \sum_{m=1}^{\infty} (-1)^{m+1} \frac{\tilde{\nu}_m^\varepsilon}{m!}$$

(ii) *In formula (5.31) the error, when one replaces the limit by a given ε ($0 < \varepsilon < \varepsilon_0 := e^{-2}$) and the infinite sum by the m_0 partial sum is bounded by:*

$$(5.32) \quad [32D_1E(\|X\|_\infty)]^{1/2} \Psi_{m_0+1}^{*,\varepsilon} + P(|X(0) - u| < \eta) + P(u < M \leq u + \eta) + \\ + P(\omega_X(\delta(\varepsilon)) \geq \eta/2) + P(\|X\|_\infty > \frac{\sqrt{2\pi\eta}}{8\varepsilon})$$

for each $\eta > 0$, Where

$$\delta(\varepsilon) := \varepsilon(2\log(1/\varepsilon))^{1/2} \\ \Psi_m^{*,\varepsilon} := \sup_{k \geq m} \left([(2k-1)!]^{1/2} \varepsilon^{2k-1} \right)^{-1/2}.$$

Note: If one wishes the bound for the error in Formula (5.32) to be smaller than some positive number, proceed according to the following steps:

- 1) choose $\eta > 0$ so that the second and third terms are small;
- 2) with that value of η , choose $\varepsilon > 0$, so that the fourth and fifth terms are small;
- 3) choose m_0 large enough to make the first term small.

PROOF. Consider the events

$$E_1 := \{|X(0) - u| < \eta\}, E_2 := \{u < M \leq u + \eta\}, \\ E_3 := \{\omega_X(\delta(\varepsilon)) \geq \eta/2\}, E_4 := \{\|X\|_\infty > \frac{\sqrt{2\pi\eta}}{4\varepsilon}\} \\ E := E_1 \cup E_2 \cup E_3 \cup E_4.$$

Observe that if $\omega \notin E$ and $\varepsilon < \varepsilon_0$, then

(5.33)

$$|X^\varepsilon(t) - X(t)| \leq \int_{-\infty}^{+\infty} \phi_\varepsilon(t-s)|X(s) - X(t)|ds \leq \omega_X(\delta(\varepsilon)) + 2\|X\|_\infty \int_{|t-s|>\delta(\varepsilon)} \phi_\varepsilon(t-s)ds < \eta$$

Using this relation one gets:

$$\begin{aligned} \mathbb{P}(M > u, X(0) \leq u) &\leq \mathbb{P}(M > u + \eta, X(0) \leq u - \eta, E^c) + \mathbb{P}(E) \leq \\ &\leq \mathbb{P}(M^\varepsilon > u, X^\varepsilon(0) < u) + \mathbb{P}(E) \leq \mathbb{P}(U_u^\varepsilon \geq 1, X^\varepsilon(0) < u) + \mathbb{P}(E). \end{aligned}$$

Also

$$\begin{aligned} \mathbb{P}(U_u^\varepsilon \geq 1, X^\varepsilon(0) \leq u, E^c) &\leq \mathbb{P}(U_u^\varepsilon \geq 1, X^\varepsilon(0) \leq u, X(0) \leq u - \eta, E^c) \leq \\ &\leq \mathbb{P}(M > u, X(0) \leq u - \eta) \leq \mathbb{P}(M > u, X(0) \leq u). \end{aligned}$$

Summing up:

$$\begin{aligned} \mathbb{P}(X(0) > u) + \mathbb{P}(U_u^\varepsilon \geq 1, X^\varepsilon(0) \leq u) - \mathbb{P}(E) &\leq \mathbb{P}(M > u) \leq \\ &\leq \mathbb{P}(X(0) > u) + \mathbb{P}(U_u^\varepsilon \geq 1, X^\varepsilon(0) \leq u) + \mathbb{P}(E). \end{aligned}$$

To compute $\mathbb{P}(U_u^\varepsilon \geq 1, X^\varepsilon(0) \leq u)$ we apply the same method as in the proof of Theorem 5.1. For that purpose, we need to show that the process X^ε satisfies the conditions for an appropriate choice of the sequence $\{c_k; k = 1, 2, \dots\}$. Denoting by $\bar{H}_k(s)$, the k -th modified Hermite polynomial (see Section 1 of Chapter 8), we have:

$$\begin{aligned} |X^{\varepsilon(k)}(t)| &\leq \varepsilon^{-(k+1)}\|X\|_\infty \int_{-\infty}^{+\infty} |\phi^{(k)}((t-s)/\varepsilon)|ds = \varepsilon^{-k}\|X\|_\infty \int_{-\infty}^{+\infty} |\phi^{(k)}(u)|du \\ &= \varepsilon^{-k}\|X\|_\infty k! \int_{-\infty}^{+\infty} |\bar{H}_k(s)|\phi(s)ds \leq \varepsilon^{-k}\|X\|_\infty k! \left(\int_{-\infty}^{+\infty} (\bar{H}_k(s))^2 \phi(s)ds \right)^{1/2} \\ &= \varepsilon^{-k}\|X\|_\infty (k!)^{1/2}. \end{aligned}$$

So,

$$\gamma_k^\varepsilon = \mathbb{P}(\|X^{\varepsilon(2k-1)}\|_\infty \geq c_k) + \frac{D_1 c_k}{2^{2k-1}(2k-1)!} \leq \frac{((2k-1)!)^{1/2}}{\varepsilon^{2k-1} c_k} E(\|X\|_\infty) + \frac{D_1 c_k}{2^{2k-1}(2k-1)!}$$

Choosing

$$c_k := \left[\frac{((2k-1)!)^{3/2} E(\|X\|_\infty)}{(\varepsilon/2)^{2k-1} D_1} \right]^{1/2}.$$

we obtain

$$\gamma_k^\varepsilon \leq 2^{-k} \left[\frac{8D_1 E(\|X\|_\infty)}{\varepsilon^{2k-1} ((2k-1)!)^{1/2}} \right]^{1/2}.$$

Hence,

$$\gamma_m^{\varepsilon*} = \sup_{k \geq m} (2^{k+1} \gamma_k^\varepsilon) \leq [32D_1 E(\|X\|_\infty)]^{1/2} \Psi_m^{*,\varepsilon}.$$

The rest follows as in the proof of Theorem 5.1. \square

Remarks and examples

1) Conditions a), b) and c) in Theorem 5.11 are usually not trivial to check and the a priori estimation of the error can be a hard problem. Moreover, when this can be actually done, the validity of Rice formulas and the feasibility of the method still remains a problem if one is willing to use Theorem 5.11 as a tool for numerical computation. For a given error, smaller ε implies larger m_0 and the usefulness of Theorem 5.11 for numerical applications is still doubtful. The bound in (5.32) shows that a priori we require at least $m_0 \approx (1/2)\varepsilon^{-2}$ terms in the sum as $\varepsilon \rightarrow 0$.

2) Let \mathcal{X} be a Gaussian process with continuous paths and

$$m(t) := \mathbb{E}(X(t)) ; \sigma^2(t) := \text{Var}(X(t)) > 0$$

be the (continuous) mean and variance of $X(t)$. Condition a) in Theorem 5.11 follows easily together with bounds on D_1 and $P(E_1)$.

Condition b) is well-known from the classical inequalities for Gaussian processes that we have considered in Chapter 2. These inequalities also imply a priori bounds for $P(E_4)$.

Condition c) follows from Ylvisaker's Theorem 1.22).

A priori bounds on $P(E_2)$ follow from bounds on the density of the distribution of the random variable M , a subject that we will consider again for certain classes of Gaussian processes in Chapter 7 .

$P(E_3)$ can be bounded using the classical methods to study the modulus of continuity of a stochastic process, as in Chapter 1.

3) Theorem 5.11 can be applied to one-dimensional diffusions satisfying certain assumptions. The reader who is not familiar with stochastic differential equations, is referred - for example - to Ikeda and Watanabe's book (1981).

Let $\{X(t) : t \geq 0\}$ be the strong solution of the stochastic differential equation

$$dX(t) = \sigma(t, X(t))dW(t) + b(t, X(t))dt ; X(0) = x_0,$$

where $\{W(t) : t \geq 0\}$ stands for the standard Wiener process , $\sigma, b : \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}$ are continuous, $\frac{\partial \sigma}{\partial x}, \frac{\partial b}{\partial x}$ are continuous and bounded and $x_0 \in \mathbb{R}$. We also assume that

$$\sigma(t, x) \geq \sigma_0 > 0 , t \in \mathbb{R}^+, x \in \mathbb{R}.$$

The methods employed in Azaïs (1989) or Nualart and Wschebor (1991) (stochastic calculus and Malliavin calculus respectively) permit to prove that $p_{X^\varepsilon(t)}$ exists and is a bounded function for $t \in [\delta, 1]$ for each $\delta > 0, 0 < \varepsilon < \varepsilon_0(\delta)$. Condition b) is standard and well known. Condition c), can be proved as in Nualart and Vives (1988) using stochastic calculus of variations.

Hence, Theorem 5.11 can be used to obtain formula (5.31) for $P(M_\delta > u)$, $M_\delta := \max_{\delta \leq t \leq 1} X(t)$ and bounds having the form (5.32) for the error. Adding an elementary bound on the local oscillation $P(\max_{0 \leq t \leq \delta} |X(t) - x_0| \leq \eta)$, one is able to get $P(M > u)$ with a controlled error. An obstacle to have an actual numerical computation for $P(M > u)$ is the lack of a good description of the joint densities of $X^\varepsilon(t), X^{\varepsilon'}(t)$ at the k -tuple (t_1, \dots, t_k) to be used in Rice formulas. This problem does not seem to have a satisfactory solution until now.

Rice formulas for random fields

In this chapter we start to study random fields, that is, random functions defined on multi-dimensional parameter sets. More precisely, the random fields that we will consider throughout are defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and have the form $\mathcal{X} = \{X(t) : t \in S\}$, where S is a subset of Euclidean space \mathbb{R}^d and the function $X(\cdot)$ takes values in some Euclidean space $\mathbb{R}^{d'}$, $d' \leq d$. We will require the paths $t \rightsquigarrow X(t)$ to be smooth functions and in some situations, we will also ask the domain S to have some geometric structure.

Our main interest lies in the random level sets $C_u(X, S) = \{t \in S : X(t) = u\}$ for each $u \in \mathbb{R}^{d'}$.

We first consider the case in which $d = d'$, in which generally speaking, for each $u \in \mathbb{R}^{d'}$ the set $C_u(X, S)$ will be a locally finite random set, and the main question is about the number of points belonging to it and lying in a subset T of S . We will denote this random number by $N_u = N_u(X, T)$, as in the one-dimensional case. The first half of this chapter is devoted to prove Rice formulas for the moments of N_u .

When $d' < d$, the random level set $C_u(X, S)$ will have of course a more complicated geometry and counting the number of points is no more interesting. Generally speaking, one expects the typical level set to be a $(d - d')$ -dimensional differentiable manifold. We will prove Rice formulas for the moments of the geometric measure of the level set.

These and related formulas have been used by Longuet-Higgins in the 50's and 60's (see for example Longuet-Higgins, 1957). A systematic treatment seems to have started with the book of R.A. Adler (1981), followed by the papers of Adler and Aronowich (1985), Adler et al. (1993) Adler and Samorodnisky (1997) and Adler (2000,2004). A proof of Rice formula for the expectation of the geometric measure of level sets of real-valued random fields is in Wschebor (1982), and was followed by various extensions and higher moments (see Wschebor, 1983, 1985). See also Cabaña (1985) where proofs for Rice formulas are given. First moments for functionals describing the geometry of $C_u(X, S)$ can be found in a recent paper by Bürgisser (2006).

The case $d' > d$ is uninteresting, since in a natural situation, for fixed u , $C_u(X, S)$ will be almost surely empty.

1. Random fields from \mathbb{R}^d to \mathbb{R}^d .

Our next task is to prove Rice Formulas for Gaussian random fields, that is, for the moments of the number of roots N_u , which is self-contained and uses only elementary arguments. It is published here for the first time and follows the proof in Azaïs and Wschebor (2006). We will also consider formulas for the moments of the total weight (as in Theorem 6.4 below) when random weights are put in each root.

1.1. The Area formula.

PROPOSITION 6.1 (Area formula). *Let f be a \mathcal{C}^1 function defined on an open subset U of \mathbb{R}^d taking values in \mathbb{R}^d . Assume that the set of critical values of f has zero Lebesgue measure.*

Let $g: \mathbb{R}^d \rightarrow \mathbb{R}$ be continuous and bounded. Then

$$(6.1) \quad \int_{\mathbb{R}^d} g(u) N_u(f, B) du = \int_B |\det(f'(t))| g(f(t)) dt.$$

for any Borel subset B of U , whenever the integral in the right-hand side is well defined.

Remarks on the statement.

1.- The hypothesis that the set of critical values of f has zero Lebesgue measure (that will be a.s. satisfied in our case) is unnecessary, since it is implied by the fact that f is a \mathcal{C}^1 -function. (This is a special case of Sard's Lemma).

2.- The result of proposition 6.1 is true under the weaker hypothesis that the function f verifies a Lipschitz condition (see Federer, Th. 3.2.5, 1969).

3.- Using standard extension arguments the continuous function g can be replaced by the indicator function of a Borel set T . Formula (6.1) can then be rewritten as

$$(6.2) \quad \int_{\mathbb{R}^d} \sum_{t \in f^{-1}(u)} h(t, u) du = \int_{\mathbb{R}^d} |\det(f'(t))| h(t, f(t)) dt,$$

where h is the function $(t, u) \rightsquigarrow \mathbf{1}_{t \in T} g(u)$. Again by a standard approximation argument (6.2) holds true for every bounded Borel-measurable function h such that the right-hand side of (6.2) is well-defined.

PROOF OF PROPOSITION 6.1

First notice that, due to standard extension arguments, it suffices to prove (6.1) for non-negative g and for T a compact parallelotope contained in U . Second, since T is a compact parallelotope, since f is \mathcal{C}^1 , the set $f(\partial T)$ of boundary values of f has Lebesgue measure zero.

Next we define an auxiliary function $\delta(u)$ for $u \in \mathbb{R}^d$ in the following way:

- If u is neither a critical value nor a boundary value and $n := N_u(f, T)$ is non zero, we denote by $x^{(1)}, \dots, x^{(n)}$ the roots of $f(x) = u$ belonging to T . Using the local inversion theorem, we know that there exists some $\delta > 0$ and n neighborhoods U_1, \dots, U_n of $x^{(1)}, \dots, x^{(n)}$ such that:

- (1) f is a \mathcal{C}^1 diffeomorphism $U_i \rightarrow B(u; \delta)$, the ball centered at u with radius δ .
- (2) U_1, \dots, U_n are pairwise disjoint and included in T .
- (3) if $t \notin \bigcup_{i=1}^n U_i$, then $f(t) \notin B(u; \delta)$.

The compactness implies that n is finite.

In this case, we define

$$\delta(u) := \sup\{\delta > 0 : (1), (2), (3) \text{ hold true for all } \delta' \leq \delta\}.$$

- If u is a critical value or a boundary value we set $\delta(u) := 0$.
- If $N_u(f, T) = 0$, we put

$$\delta(u) := \sup\{\delta > 0 : f(T) \cap B(u; \delta) = \emptyset\}.$$

It is clear that in this case $\delta(u) > 0$.

The function $\delta(u)$ is Lipschitz. In fact, let u be a value of f which is not a critical value nor a boundary value, if u' belongs to $B(u; \delta(u))$, then $B(u'; \delta(u) - \|u' - u\|) \subset B(u; \delta(u))$ and as a consequence $\delta(u') \geq \delta(u) - \|u' - u\|$. Exchanging the roles of u and u' , we get

$$|\delta(u') - \delta(u)| \leq \|u - u'\|.$$

The Lipschitz condition is easily checked in the other two cases.

Let \mathcal{F} be a real-valued monotone continuous function defined on \mathbb{R}^+ such that

$$(6.3) \quad \mathcal{F} \equiv \begin{cases} 0 & \text{on } [0, 1/2], \\ 1 & \text{on } [1 + \infty) \end{cases}$$

Let $\delta(u) > 0$ and $0 < \delta < \delta(u)$. Using the change of variable formula we have

$$\int_T |\det(f'(t))| \mathbf{1}_{\|f(t)-u\|<\delta} dt = \sum_{i=1}^n \int_{U_i} |\det(f'(t))| dt = V(\delta)n,$$

where $V(\delta)$ is the volume of the ball with radius δ in \mathbb{R}^d . Thus, we have an exact counter for $N_u(f, T)$ when it is non-zero, which obviously holds true also when $N_u(f, T) = 0$ for $\delta < \delta(u)$.

Let $g : \mathbb{R}^d \rightarrow \mathbb{R}$ continuous, bounded and non-negative and $\delta_0 > 0$. For every $\delta' < \delta_0/2$ we have:

$$\int_{\mathbb{R}^d} g(u) N_u(f, T) \mathcal{F} \left(\frac{\delta(u)}{\delta_0} \right) du = \int_{\mathbb{R}^d} g(u) \mathcal{F} \left(\frac{\delta(u)}{\delta_0} \right) du \frac{1}{V(\delta')} \int_T |\det(f'(t))| \mathbf{1}_{\|f(t)-u\|<\delta'} dt$$

Applying Fubini's Theorem we see that the expression above is equal to:

$$A_{\delta_0, \delta'} := \int_T |\det(f'(t))| dt \frac{1}{V(\delta')} \int_{B(f(t); \delta')} \mathcal{F} \left(\frac{\delta(u)}{\delta_0} \right) g(u) du.$$

$A_{\delta_0, \delta'}$ in fact does not depend on δ' so it is equal to its limit as $\delta' \rightarrow 0$ which is, because of the continuity of the function $u \rightsquigarrow \mathcal{F} \left(\frac{\delta(u)}{\delta_0} \right) g(u)$, equal to

$$\int_T |\det(f'(t))| \mathcal{F} \left(\frac{\delta(f(t))}{\delta_0} \right) g(f(t)) dt.$$

Let δ_0 tend to zero and use monotone convergence. For the left-hand side, we take into account that the set of critical values and the set of boundary values have measure zero. For the right-hand side, we use the definition of \mathcal{F} , that the boundary of T has Lebesgue measure zero and the integrand is zero if t is a critical point of f . \square

1.2. Rice formulas for Gaussian random fields.

1.2.1. Main results.

THEOREM 6.2 (Rice formula for the expectation). *Let $Z : U \rightarrow \mathbb{R}^d$ be a random field, U an open subset of \mathbb{R}^d and $u \in \mathbb{R}^d$ a fixed point. Assume that:*

- (i) Z is Gaussian,
- (ii) almost surely the function $t \rightsquigarrow Z(t)$ is of class \mathcal{C}^1 ,
- (iii) for each $t \in U$, $Z(t)$ has a non degenerate distribution (i.e. $\text{Var}(Z(t)) \succ 0$),
- (iv) $\mathbb{P}\{\exists t \in U, Z(t) = u, \det(Z'(t)) = 0\} = 0$.

Then, for every Borel set B contained in U , one has

$$(6.4) \quad \mathbb{E}(N_u(Z, B)) = \int_B \mathbb{E}(|\det(Z'(t))| | Z(t) = u) p_{Z(t)}(u) dt.$$

If B is compact, both sides in (6.4) are finite.

THEOREM 6.3 (Rice formula for the k -th moment). *Let $k, k \geq 2$ be an integer. Assume the same hypotheses as in Theorem 6.2 except for (iii) that is replaced by*

- (iii') for $t_1, \dots, t_k \in U$ distinct values of the parameter, the distribution of

$$(Z(t_1), \dots, Z(t_k))$$

does not degenerate in $(\mathbb{R}^d)^k$.

Then for every Borel set B contained in U , one has

$$(6.5) \quad \mathbb{E} \left[(N_u(Z, B)) (N_u(Z, B) - 1) \dots (N_u(Z, B) - k + 1) \right] \\ = \int_{B^k} \mathbb{E} \left(\prod_{j=1}^k |\det(Z'(t_j))| \mid Z(t_1) = \dots = Z(t_k) = u \right) \\ p_{Z(t_1), \dots, Z(t_k)}(u, \dots, u) dt_1 \dots dt_k,$$

where both sides may be infinite.

Remark: With the same proof as that of Theorem 6.3 and under the same conditions we have for distinct u_1, \dots, u_k

$$\mathbb{E} \left[(N_{u_1}(Z, B)) (N_{u_2}(Z, B)) \dots (N_{u_k}(Z, B)) \right] \\ = \int_{B^k} \mathbb{E} \left(\prod_{j=1}^k |\det(Z'(t_j))| \mid Z(t_1) = u_1, \dots, Z(t_k) = u_k \right) \\ p_{Z(t_1), \dots, Z(t_k)}(u_1, \dots, u_k) dt_1 \dots dt_k.$$

THEOREM 6.4 (Expected number of weighted roots.). *Let Z be a random field that verifies the hypotheses of Theorem 6.2. Assume that for each $t \in U$ one has another random field $Y^t : W \rightarrow \mathbb{R}^n$, where W is some topological space, verifying the following conditions:*

- a) $Y^t(w)$ is a measurable function of (ω, t, w) and almost surely, $(t, w) \rightsquigarrow Y^t(w)$ is continuous.
- b) For each $t \in U$ the random process $(s, w) \rightsquigarrow (Z(s), Y^t(w))$ defined on $U \times W$ is Gaussian.

Moreover, assume that $g : U \times \mathcal{C}(W, \mathbb{R}^n) \rightarrow \mathbb{R}$ is a bounded function, which is continuous when one puts on $\mathcal{C}(W, \mathbb{R}^n)$ the topology of uniform convergence on compact sets. Then, for each compact subset I of U , one has

$$(6.6) \quad \mathbb{E} \left(\sum_{t \in I, Z(t)=u} g(t, Y^t) \right) = \int_I \mathbb{E} (|\det(Z'(t))| g(t, Y^t) \mid Z(t) = u) \cdot p_{Z(t)}(u) dt.$$

PROOF OF THEOREM 6.2

Let $\mathcal{F} : \mathbb{R}^+ \rightarrow [0, 1]$ be the function defined in (6.3), For m, n positive integers and $x \geq 0$, define :

$$(6.7) \quad F_m(x) := \mathcal{F}(mx) ; G_n(x) := 1 - \mathcal{F}(x/n).$$

A standard extension argument says that it is enough to prove the theorem when B is a compact rectangle included in U . So assume B satisfies this condition. Let us introduce some more notations:

- $\Delta(t) := |\det(Z'(t))|$ ($t \in U$)
- For n, m positive integers and $u \in \mathbb{R}^d$:

$$(6.8) \quad C_u^m(B) := \sum_{s \in B: Z(s)=u} F_m(\Delta(s)).$$

$$(6.9) \quad Q_u^{n,m}(B) := C_u^m(B) G_n(C_u^m(B)).$$

In (6.8) when the summation index set is empty, we put $C_u^m(B) = 0$. Let $g : \mathbb{R}^d \rightarrow \mathbb{R}$ be continuous with compact support . We apply the area formula (6.2) for the function

$$h(t, u) = F_m(\Delta(t)) G_n(C_u^m(B)) g(u) \mathbf{1}_{t \in B}$$

to get

$$\int_{\mathbb{R}^d} g(u) Q_u^{n,m}(B) du = \int_B \Delta(t) F_m(\Delta(t)) G_n(C_{Z(t)}^m(B)) g(Z(t)) dt.$$

Taking expectations in both sides provides

$$\int_{\mathbb{R}^d} g(u) \mathbb{E}(Q_u^{n,m}(B)) du = \int_{\mathbb{R}^d} g(u) du \int_B \mathbb{E}[\Delta(t) F_m(\Delta(t)) G_n(C_u^m(B)) | Z(t) = u] p_{Z(t)}(u) dt.$$

Since this equality holds for any g continuous with bounded support, it follows that

$$(6.10) \quad \mathbb{E}(Q_u^{n,m}(B)) = \int_B \mathbb{E}[\Delta(t) F_m(\Delta(t)) G_n(C_u^m(B)) | Z(t) = u] p_{Z(t)}(u) dt,$$

for almost every $u \in \mathbb{R}^d$.

Notice that the hypotheses imply that if J is a subset of U , $\lambda_d(J) = 0$, then $\mathbb{P}\{N_u(Z, J) = 0\} = 1$ for each $u \in \mathbb{R}^d$. In particular, almost surely, there are no roots of $Z(t) = u$ in the boundary of the rectangle B .

Let us prove that the left-hand side of (6.10) is a continuous function of u . Fix $u \in \mathbb{R}^d$. Outside the compact set

$$\{t \in B : \Delta(t) \geq 1/2m\},$$

the contribution to the sum (6.8) defining $C_v^m(B)$ is zero, for any $v \in \mathbb{R}^d$. Using the local inversion theorem, the number of points $t \in B$ such that $Z(t) = u$; $\Delta(t) \geq 1/2m$, say k , is finite. Almost surely, all these points are interior to B .

If k is non-zero, $Z(t)$ is locally invertible in k neighborhoods $V_1, \dots, V_k \subset B$ around these k points. For v in some (random) neighborhood of u , there is exactly one root of $Z(s) = v$ in each V_1, \dots, V_k and the contribution to $C_v^m(B)$ of these points can be made arbitrarily close to the one corresponding to $v = u$. Outside the union of V_1, \dots, V_k , $Z(t) - u$ is bounded away from zero in B , so that the contribution to $C_v^m(B)$ vanishes if v is sufficiently close to u .

Therefore the function $v \rightsquigarrow Q_v^{n,m}$ is a.s. continuous at $v = u$. On the other hand, it is obvious from its definition that $Q_v^{n,m}(B) \leq n$ and an application of the Lebesgue dominated convergence theorem implies the continuity of $\mathbb{E}(Q_u^{n,m}(B))$ as a function of u .

Let us now write the Gaussian regression formulas for fixed $t \in B$:

$$(6.11) \quad \begin{aligned} Z(s) &= a^t(s)Z(t) + Z^t(s) \\ Z'(s) &= (a^t)'(s)Z(t) + (Z^t)'(s), \end{aligned}$$

where $'$ denotes the derivative with respect to s and the pair $(Z^t(s), (Z^t)'(s))$ is independent from $Z(t)$ for all $s \in U$.

Then, we write the conditional expectation on the right-hand side of (6.10) as the unconditional expectation :

$$(6.12) \quad \mathbb{E}[\Delta_u^t(t) F_m(\Delta_u^t(t)) G_n(\tilde{C}_u^m(B))],$$

where we use the notations

$$\begin{aligned} \Delta_u^t(s) &:= |\det(Z_u^t)'(s)| \\ Z_u^t(s) &:= a^t(s)u + Z^t(s) \\ \tilde{C}_u^m(B) &:= \sum_{s \in B, Z_u^t(s)=u} F_m(\Delta_u^t(s)). \end{aligned}$$

Now, observe that (6.10) implies that for almost every $u \in \mathbb{R}^d$ one has the inequality

$$(6.13) \quad \mathbb{E}(Q_u^{n,m}(B)) \leq \int_B \mathbb{E}[\Delta(t) | Z(t) = u] p_{Z(t)}(u) dt,$$

which is in fact true for all $u \in \mathbb{R}^d$ since both sides are continuous functions of u .

The remainder of the proof consists in proving an inequality in the opposite sense. Let us fix n, m, u and t . Let K be the compact set

$$K := \{s \in B : \Delta_u^t(s) \geq 1/4m\}$$

If v varies in a sufficient small (random) neighborhood of u , the points outside K do not contribute to the sum defining $\tilde{C}_v^m(B)$.

Let k be the almost surely finite number of roots of $Z_u^t(s) = u$ lying in the set K . Assume that k does not vanish and denote these roots by $\bar{s}_1, \dots, \bar{s}_k$. Consider the equation

$$(6.14) \quad Z_v^t(s) - v = 0.$$

in a neighborhood of each one of the pairs $\tilde{s} = \bar{s}_i$, $v = u$. Applying the Implicit Function Theorem, one can find k pairwise disjoint open sets V_1, \dots, V_k such that if v is sufficiently close to u , equation (6.14) has exactly one root $s_i = s_i(v)$ in V_i , $i = 1, \dots, k$. These roots vary continuously with v and $s_i(u) = \bar{s}_i$. On the other hand on the compact set $K \setminus (V_1 \cup \dots \cup V_k)$ the quantity $\|Z_u^t(s) - u\|$ is bounded away from zero so $\|Z_v^t(s) - v\|$ does not vanishes if v is sufficiently close to u . As a conclusion, we have that

$$\limsup_{v \rightarrow u} \tilde{C}_v^m(B) \leq \tilde{C}_u^m(B)$$

where the inequality arises from the fact that some of the points $s_i(v)$ may not belong to B and hence, don't contribute to the sum defining $\tilde{C}_v^m(B)$. Now since (6.10) holds for a.e. u , one can find a sequence $\{u_N, N = 1, 2, \dots\}$ converging to u such that (6.10) holds true for $u = u_N$ and all $N = 1, 2, \dots$. Using the continuity - already proved - of the function $u \rightsquigarrow E(Q_u^{n,m}(B))$, Fatou's Lemma and the fact that G_n is non-increasing, we have :

$$\begin{aligned} E(Q_u^{n,m}(B)) &= \lim_{N \rightarrow +\infty} E(Q_{u_N}^{n,m}(B)) \\ &= \lim_{N \rightarrow +\infty} \int_B E[\Delta_{u_N}^t(t) F_m(\Delta_{u_N}^t(t) G_n(\tilde{C}_{u_N}^m(B)))] p_{Z(t)}(u_N) dt \\ &\geq \int_B E[\Delta_u^t(t) F_m(\Delta_u^t(t) G_n(\tilde{C}_u^m(B)))] p_{Z(t)}(u) dt. \end{aligned}$$

Since $\tilde{C}_u^m(B)$ is a.s. finite, we can now pass to the limit as $n \rightarrow +\infty$, $m \rightarrow +\infty$ (in that order) and applying Beppo-Levi's Theorem, conclude the proof. \square

PROOF OF THEOREM 6.3: For each $\delta > 0$, define the domain

$$D_{k,\delta}(B) = \{(t_1, \dots, t_k) \in B^k, \|t_i - t_j\| \geq \delta \text{ if } i \neq j, i, j = 1, \dots, k\}$$

and the process \tilde{Z}

$$(t_1, \dots, t_k) \in D_{k,\delta}(B) \rightsquigarrow \tilde{Z}(t_1, \dots, t_k) = (Z(t_1), \dots, Z(t_k)).$$

It is clear that \tilde{Z} satisfies the hypotheses of Theorem 6.2 for every value $(u, \dots, u) \in (\mathbb{R}^d)^k$. So,

$$(6.15) \quad E \left[N_{(u, \dots, u)}(\tilde{Z}, D_{k,\delta}(B)) \right] \\ = \int_{D_{k,\delta}(B)} E \left(\prod_{j=1}^k |\det(Z'(t_j))| \mid Z(t_1) = \dots = Z(t_k) = u \right) p_{Z(t_1), \dots, Z(t_k)}(u, \dots, u) dt_1 \dots dt_k$$

To finish, let $\delta \downarrow 0$, and take into account that $(N_u(Z, B))(N_u(Z, B) - 1) \dots (N_u(Z, B) - k + 1)$ is the monotone limit of $N_{(u, \dots, u)}(\tilde{Z}, (D_{k,\delta}(B)))$, and that the diagonal $D_k(B) = \{(t_1, \dots, t_k) \in B^k, t_i = t_j \text{ for some pair } i, j, i \neq j\}$ has zero Lebesgue measure in $(\mathbb{R}^d)^k$. \square

PROOF OF THEOREM 6.4: The proof is essentially the same. It suffices to consider instead of $C_u^m(B)$ the quantity

$$(6.16) \quad C_u^m(I) := \sum_{s \in I: Z(s)=u} \mathcal{F}_m(\Delta(s)) \cdot g_s(s, Y^s).$$

□

1.2.2. Sufficient conditions for hypothesis (iv) in Theorem 6.2.

These conditions are given by the following proposition:

PROPOSITION 6.5. *Let $Z : U \rightarrow \mathbb{R}^d$, U a compact subset of \mathbb{R}^d be a random field with paths of class C^1 and $u \in \mathbb{R}^d$. Assume that*

- $p_{Z(t)}(x) \leq C$ for all $t \in U$ and x in some neighborhood of u .
- at least one of the two following hypotheses is satisfied:
 - a) a.s. $t \rightsquigarrow Z(t)$ is of class C^2
 - b)

$$\alpha(\delta) = \sup_{t \in U, x \in V(u)} \mathbb{P}\{|\det(Z'(t))| < \delta | Z(t) = x\} \rightarrow 0$$

as $\delta \rightarrow 0$, where $V(u)$ is some neighborhood of u .

Then (iv) holds true.

PROOF. Assume with no loss of generality that $I = [0, 1]^d$ and that $u = 0$. Put $G_I = \{\exists t \in I, Z(t) = 0, \det(Z'(t)) = 0\}$

PROOF UNDER CONDITION a) (Cucker and Wschebor, 2003).

For each integer N consider I as a union of cubes of sides $1/N$ with sides parallel to the axis. We denote these cubes C_1, \dots, C_{N^d} . In a similar way, we consider each face at the boundary of the cube C_r as a union of $(d-1)$ -dimensional cubes of sides $1/N^2$. We denote these cubes D_{rs} , $s = 1, \dots, 2dN^{d-1}$. In each D_{rs} fix a point τ_{rs}^* , for instance, the center.

We denote $Z = (Z_1, \dots, Z_d)^T$ and $t = (t_1, \dots, t_d)^T$.

For a given $\eta > 0$, choose $B > 0$ large enough so that $\mathbb{P}\{F_B\} < \eta$, where F_B is the event:

$$F_B = \left\{ \left[\sup \left[\left| \frac{\partial Z_i}{\partial t_j}(t) \right|, \left| \frac{\partial^2 Z_i}{\partial t_j \partial t_h}(t) \right| \right] : i, j, h = 1, \dots, d; t \in [0, 1]^d \right] > B \right\}.$$

Clearly:

$$G_I = \bigcup_{r=1}^{r=N^d} \{\exists \tau_r \in C_r, v \in \mathbb{R}^d, \text{ such that } Z(\tau_r) = 0, \|v\| = 1, Z'(\tau_r)v = 0\} = \bigcup_{r=1}^{r=N^d} G_r$$

Assume $G_r \cap F_B^C$ is non-empty. Denote $\tilde{\tau}_{rs}$ an intersection point with the boundary of C_r , of the straight line through τ_r which is parallel to v . Consider the Taylor expansion of Z_i at the point τ_r , evaluated at the point $\tilde{\tau}_{rs}$:

$$\begin{aligned} Z_i(\tilde{\tau}_{rs}) &= Z_i(\tau_r) + \sum_{j=1}^d \frac{\partial Z_i}{\partial t_j}(\tau_r)(\tilde{\tau}_{rs,j} - \tau_{r,j}) \\ &\quad + \sum_{j,h=1}^d \frac{\partial^2 Z_i}{\partial t_j \partial t_h}(\tau_r + \theta(\tilde{\tau}_{rs} - \tau_r))(\tilde{\tau}_{rs,j} - \tau_{r,j})(\tilde{\tau}_{rs,h} - \tau_{r,h}), \end{aligned}$$

with $0 < \theta < 1$. Since the first two terms in this sum are equal to zero, we deduce that $|Z_i(\tilde{\tau}_{rs})| \leq K_d B N^{-2}$ for all $i = 1, \dots, d$, where K_d is some constant depending only on the dimension.

Since the diameter of each D_{rs} is bounded by a constant depending only on the dimension, times N^{-2} , it follows that $\|Z(\tau_{rs}^*)\| \leq KN^{-2}$ for some constant K depending only on d and B .

So,

$$\begin{aligned} \mathbb{P}\{G_I\} &\leq \mathbb{P}\{F_B\} + \mathbb{P}\{\exists r \leq N^d, s \leq N^{d-1}, \text{ s.t. } \|Z(\tau_{rs}^*)\| \leq KN^{-2}\} \\ &\leq \eta + \sum_{r=1}^{N^d} \sum_{s=1}^{2dN^{d-1}} \mathbb{P}\{\|Z(\tau_{rs}^*)\| \leq KN^{-2}\} \\ &\leq \eta + N^{2d-1} C K_1 N^{-2d} \end{aligned}$$

where K_1 is a constant depending on d and B , using the hypotheses on the boundedness of the density. The remainder is plain. This proves (iv) under condition a)

PROOF UNDER CONDITION b) (Azaïs and Wschebor, 2005)

Choose $\varepsilon > 0$, $\eta > 0$; there exists a positive number M such that

$$\mathbb{P}(E_M) = \mathbb{P}\left\{\sup_{t \in I} \|Z'(t)\| > M\right\} \leq \varepsilon.$$

Denote by ω_{\det} the modulus of continuity of $|\det(X'(\cdot))|$ and choose N large enough so that

$$\mathbb{P}(F_{N,\eta}) = \mathbb{P}\left\{\omega_{\det}\left(\frac{\sqrt{d}}{N}\right) \geq \eta\right\} \leq \varepsilon.$$

Consider the partition of I used in part a) into N^d small cubes. Let τ_r^* be the center of C_r . Then

$$(6.17) \quad \mathbb{P}(G_I) \leq \mathbb{P}(E_M) + \mathbb{P}(F_{N,\eta}) + \sum_r \mathbb{P}\left(G_{C_r} \cap E_M^c \cap F_{N,\eta}^c\right)$$

When the event in the r -th term occurs, we have:

$$|Z_j(\tau_r^*)| \leq \frac{M}{N} \sqrt{d}, \quad j = 1, \dots, d$$

and

$$|\det(Z'(\tau_r^*))| < \eta.$$

So, if N is chosen sufficiently large so that $V(0)$ contains the ball centered at 0 with radius $\frac{M\sqrt{d}}{N}$, one has:

$$\mathbb{P}(G_I) \leq 2\varepsilon + N^d \left(\frac{2M}{N} \sqrt{d}\right)^d C \alpha(\eta)$$

Since ε and η are arbitrarily small, the result follows. \square

1.3. Maxima and critical points on a smooth manifold. Let us write Rice formula for the first moment in two special cases that will appear various times in the remaining of this book. These correspond to the number of local maxima and the number of critical points of a real-valued random field.

Assume $\{X(t) : t \in W\}$ is a real-valued random field defined on the open subset W of \mathbb{R}^d and such that $Z(t) = X'(t)$ satisfies the hypothesis of Theorem 6.2. Let S be a Borel subset of W and $u \in \mathbb{R}$. The following quantities are well defined and measurable: $M_{u,1}(X, S)$, the number of local maxima and $M_{u,2}(X, S)$, the number of critical points of $X(\cdot)$ belonging to S in which the function $X(\cdot)$ takes a value bigger than u .

We also introduce the following notation: for each real symmetric matrix M we put $\delta^1(M) := |\det(M)| \mathbf{1}_{M < 0}$ and $\delta^2(M) := |\det(M)|$.

Then, we have the following formulas for the expectation ($k = 1, 2$):

$$(6.18) \quad \mathbb{E}(M_{u,k}^X(S)) = \int_S ds \int_u^{+\infty} \mathbb{E}(\delta^k(X''(s)) | X(s) = x, X'(s) = 0) p_{X(s), X'(s)}(x, 0) dx.$$

Similar expressions are obtained when extending the statements of Theorems 6.3 and 6.4 to this case.

Let W be a \mathcal{C}^2 -manifold of dimension d . We suppose that W is orientable, that is, there exists an atlas $((U_i, \phi_i); i \in I)$ such that for any pair of intersecting charts (U_i, ϕ_i) , (U_j, ϕ_j) , the Jacobian of the map $\phi_i \circ \phi_j^{-1}$ is positive.

We consider a Gaussian random field with real values and \mathcal{C}^2 paths $X = \{X(t) : t \in W\}$ defined on the manifold W . In this subsection, our aim is to write Rice Formulas for this kind of random fields under various geometric settings for W . More precisely we will consider three cases: first, when W is a manifold without any additional structure on it; second, when W has a Riemannian metric; third, when it is embedded in Euclidean space. We will use these formulas in the next chapters but they have an interest in themselves (see Taylor and Adler (2003) for other details or similar results).

We will denote the derivative along the manifold by $DX(t)$ to distinguish it from the free derivative in \mathbb{R}^d and we will assume that, in every chart, the pair $X(t)$ and $DX(t)$ has a non-degenerate joint distribution and that hypothesis (iv) of Theorem 6.2 is verified.

1.3.1. Abstract manifold.

PROPOSITION 6.6. For $k = 1, 2$ the quantity which is expressed in every chart ϕ with coordinates s_1, \dots, s_d as

$$(6.19) \quad \int_u^{+\infty} dx \mathbb{E}(\delta^k(Y''(s)) | Y(s) = x, Y'(s) = 0) p_{Y(s), Y'(s)}(x, 0) \wedge_{i=1}^d ds_i,$$

where $Y(s)$ is the process X written in the chart : $Y = X \circ \phi^{-1}$, defines a d -form Ω^k on \dot{W} and for every Borel set $S \subset \dot{W}$

$$\int_S d\Omega^k = \mathbb{E}(M_{u,k}^X(S)).$$

PROOF. Note that a d -form is a measure on \dot{W} whose image in each chart is absolutely continuous with respect to Lebesgue measure $\wedge_{i=1}^d ds_i$. To prove that (6.19) defines a d -form, it is sufficient to prove that its density with respect to $\wedge_{i=1}^d ds_i$, satisfies locally the change-of-variable formula. Let (U_1, ϕ_1) , (U_2, ϕ_2) two intersecting charts and set

$$U_3 := U_1 \cap U_2 ; Y_1 := X \circ \phi_1^{-1} ; Y_2 := X \circ \phi_2^{-1} ; H := \phi_2 \circ \phi_1^{-1}.$$

Denote by s_i^1 and s_i^2 , $i = 1, \dots, d$ the coordinates in each chart. We have

$$\begin{aligned} \frac{\partial Y_1}{\partial s_i^1} &= \sum_{i'} \frac{\partial Y_2}{\partial s_i^2} \frac{\partial H_{i'}}{\partial s_i^1} \\ \frac{\partial^2 Y_1}{\partial s_i^1 \partial s_j^1} &= \sum_{i', j'} \frac{\partial^2 Y_2}{\partial s_i^2 \partial s_j^2} \frac{\partial H_{i'}}{\partial s_i^1} \frac{\partial H_{j'}}{\partial s_j^1} + \sum_{i'} \frac{\partial Y_2}{\partial s_i^2} \frac{\partial^2 H_{i'}}{\partial s_i^1 \partial s_j^1}. \end{aligned}$$

Thus at every point

$$Y_1'(s^1) = (H'(s^1))^T Y_2'(s^2),$$

$$p_{Y_1(s^1), Y_1'(s^1)}(x, 0) = p_{Y_2(s^2), Y_2'(s^2)}(x, 0) |\det(H'(s^1))|^{-1}$$

and at a singular point

$$Y_1''(s^1) = (H'(s^1))^T Y_2''(s^2) H'(s^1).$$

On the other hand, by the change of variable formula

$$\wedge_{i=1}^d ds_i^1 = |\det(H'(s^1))|^{-1} \wedge_{i=1}^d ds_i^2.$$

Replacing in the integrand in (6.19), one checks the desired result.

To prove the second part again it suffices to prove it locally for an open subset S included in a unique chart. Let $((S, \phi))$ be a chart and let again $Y(s)$ be the process written in this chart. It suffices to prove that

$$(6.20) \quad \mathbb{E}(M_{u,k}^X(S)) = \int_{\phi(S)} d\lambda(s) \int_u^{+\infty} dx \mathbb{E}(\delta^k(Y''(s)) | Y(s) = x, Y'(s) = 0) p_{Y(s), Y'(s)}(x, 0).$$

Since $M_{u,k}^X(S)$ is equal to $M_{u,k}^Y\{\phi(S)\}$, the result is a direct consequence of Theorem 6.4. \square

1.3.2. *Riemannian manifold.* The form in (6.19) is independent of the parameterization but the terms inside the integrand are not. It is possible to give an expression that consists of three terms independent of the parameterization in the case when W is a Riemannian manifold. When such a Riemannian metric is not given, it is always possible to use the metric g induced by the process itself (see Taylor and Adler, 2002) by setting

$$g(s)(Y, Z) = \mathbb{E}\left(\left(Y(X)\right)\left(Z(X)\right)\right),$$

Y, Z being two tangent vectors belonging to the tangent space $T(s)$ at $s \in W$. $Y(X)$, (resp. $Z(X)$) denotes the action of the tangent vector Y (resp. Z) on the function X . This metric leads to very simple expressions for centered variance-1 Gaussian processes.

The main point is that at a singular point of X the second order derivative D^2X does not depend on the parameterization since it defines locally the Taylor expansion of the function X . Given the Riemannian metric g_s the second differential can be represented by an endomorphism that will be denoted $\nabla^2 X(s)$.

$$(D^2X)(s)[Y, Z] = g_s(\nabla^2 X(s)Y, Z)$$

This endomorphism is independent of the parameterization and of course its determinant. So in a chart

$$(6.21) \quad \det(\nabla^2 X(s)) = \det(D^2X(s)) \det(g_s)^{-1},$$

and $\nabla^2 X(s)$ is negative definite if and only if $D^2X(s)$ is. Hence

$$\delta^k(\nabla^2 X(s)) = \delta^k(D^2X(s)) \det(g_s)^{-1}; \quad (k = 1, 2).$$

We turn now to the density in (6.19). The gradient at some location s is defined as the unique vector $\nabla X(s) \in T(s)$ such that $g_s(\nabla X(s), Y) = DX(s)[Y]$. In a chart the vector of coordinates of the gradient in the basis $\partial x_i, i = 1, d$ is given by $(g_s)^{-1}DX(s)$ where $DX(s)$ is now the vector of coordinates of the derivative in the basis $dx^i, i = 1, d$. The joint density at $(x, 0)$ of $(X(s), \nabla X(s))$ is intrinsic only if expressed in an orthonormal basis of the tangent space. In that case the vector of coordinates is given by

$$\widetilde{\nabla X}(s) = (g_s)^{1/2} \nabla X(s) = (g_s)^{-1/2} DX$$

By the change of variable formula :

$$p_{X(s), \widetilde{\nabla X}(s)}(x, 0) = p_{X(s), DX(s)}(x, 0) \sqrt{\det(g_s)}$$

Reminding that the Riemannian volume Vol satisfies

$$dVol = \sqrt{\det(g_s)} \wedge_{i=1}^d ds_i^2$$

we can rewrite expression (6.19) as

$$(6.22) \quad \int_u^{+\infty} dx \mathbb{E}(\delta^k(\nabla^2 X(s)) | X(s) = x, \nabla X(s) = 0) p_{X(s), \nabla X(s)}(x, 0) dVol$$

where we have omitted the tilde above $\nabla X(s)$ for simplicity. This is the Riemannian expression.

1.3.3. *Embedded manifold.* In most practical applications, W is naturally embedded in an Euclidean space say \mathbb{R}^m . In this case we look for an expression for (6.22) as a function of the natural derivative on \mathbb{R}^m . The manifold is equipped with the metric induced by the Euclidean metric in \mathbb{R}^m . Considering the form (6.22), clearly the Riemannian volume is just the geometric measure σ on W .

Following Milnor (1965), we assume that the process $X(t)$ is defined on an open neighbourhood of W so that the ordinary derivatives $X'(s)$ and $X''(s)$ are well defined for $s \in W$. Denoting the projector onto the tangent and normal spaces by $P_{T(s)}$ and $P_{N(s)}$, we have.

$$\nabla X(s) = P_{T(s)}(X'(s)).$$

The next formula is well-known and gives the expression of the second differential form at a singular point

$$(6.23) \quad Y, Z \in T(s) \rightsquigarrow X''(s)[Y, Z] + \langle \mathbb{I}[Y, Z], X'(s) \rangle,$$

where \mathbb{I} is the second fundamental form of W embedded in \mathbb{R}^m than can be defined in our simple case by

$$Y, Z \in T(s) \rightsquigarrow P_{N(s)}(D_X Y).$$

The determinant of the bilinear form given by (6.23), expressed in an orthonormal basis, gives the value of $\det(\nabla^2 X(s))$. As a conclusion we get the expression of every terms involved in (6.22).

1.3.4. *Examples:*

Codim. 1: with a given orientation we get

$$\nabla^2 X = X''_T + \mathbb{I}.X'_N$$

where X''_T is the tangent projection of the second derivative and X'_N the normal component of the gradient.

Sphere: When W is a sphere of radius $r > 0$ in \mathbb{R}^{d+1} oriented towards the inside

$$(6.24) \quad \nabla^2 X = X''_T + rI_d X'_N$$

Curve: When the manifold is a curve parameterized by arc length

$$(6.25) \quad \mathbb{E}(M_u^k(W)) = \int_u^{+\infty} dx \int_0^L dt \\ \mathbb{E}(\delta^k(X''_T(t) + C(t)X'_N(t)) | X(t) = x, X'_T(t) = 0) p_{X(t), X'_T(t)}(x, 0),$$

Where $C(t)$ is the curvature at location t and $X'_N(t)$ is the derivative taken in the direction of the normal to the curve at point t .

1.4. Extensions to certain Non-Gaussian random fields. .

It is easy to adapt the above proofs of Rice formulas to certain classes of Gaussian-related random fields which do not need to be Gaussian. We exemplify this with the statement of Theorem 6.2, but the same holds true, mutatis mutandis, with the other theorems.

To be precise, the conclusion of Theorem 6.2 remains valid if we replace the hypotheses (i),(ii),(iii),(iv) by the following (we keep the same notations as in the statement of the theorem):

- (i) $Z(t) = H[Y(t)]$ for $t \in W$ where:
- $\{Y(t) : t \in W\}$ is a Gaussian random field having values in \mathbb{R}^n and \mathcal{C}^1 -paths and such that for each $t \in W$ the distribution of $Y(t)$ does not degenerate,
 - $H : \mathbb{R}^n \rightarrow \mathbb{R}^d$ is a \mathcal{C}^1 -function.

(ii) for each $t \in W$, $Z(t)$ has a density $p_{Z(t)}(x)$ which is a continuous function of the pair $(t, x) \in W \times \mathbb{R}^d$.

(iii) $P\{\exists t \in W, Z(t) = u, \det(Z'(t)) = 0\} = 0$.

Notice that these hypotheses imply that one must have $n \geq d$. The only change to be introduced in the proof is that instead of performing the regression on $Z(t)$ one should do it on $Y(t)$.

As for the validity of Rice formulas for more general Non-Gaussian random fields, a careful analysis of the proof of Theorem 6.2, shows that in fact Gaussianity plays a role only in assuring the continuity (as functions of u) of the conditional expectation and the density which appear in the integrands in the right-hand sides of (6.4) and (6.10). The continuity of the density is obvious and that of the conditional expectation is a consequence of the possibility of using regression to get rid of the conditioning, which is a specifically Gaussian property. Otherwise, the proof is independent of the nature of the law of the random field $\{Z(t) : t \in W\}$. From the standpoint of applications, one must consider also that if the random field is non-Gaussian, the actual computation of the conditional expectation can be hard or impossible to be actually performed, and the interest of the formula remains limited.

On account of this, next we state as a theorem Rice formula for the expectation of the number or roots of non-Gaussian random fields. Similar expressions hold true for higher moments and for weighted roots as well as for random fields parameterized on manifolds. The proof follows strictly that of Theorem 6.2, except for the points just mentioned.

THEOREM 6.7. *Let $Z : W \rightarrow \mathbb{R}^d$ be a random field, W an open subset of \mathbb{R}^d and $u \in \mathbb{R}^d$ a fixed point. Assume that:*

- (i) *Almost surely the function $t \rightsquigarrow Z(t)$ is of class \mathcal{C}^1 .*
- (ii) *For each $t \in W$, $Z(t)$ has a density $p_{Z(t)}(\cdot)$ and the function $(t, x) \rightsquigarrow p_{Z(t)}(x)$ is continuous for $t \in U$ and x is some neighborhood of u .*
- (iii) *Let $\alpha : \mathcal{C}^1(U) \rightarrow \mathbb{R}$ be a real-valued functional defined on $\mathcal{C}^1(W)$, which is continuous in the sense that if $\{f_n\}_{n=1,2,\dots}$ is a sequence of functions in $\mathcal{C}^1(W)$ such that $f_n \rightarrow f$, $f'_n \rightarrow f'$ as $n \rightarrow \infty$, uniformly on the compact subsets of W , then $\alpha(f_n) \rightarrow \alpha(f)$.*

Our assumption is that for such a functional α there exists a version of the conditional expectation

$$E(\alpha(Z) | Z(t) = x)$$

which is continuous as a function of the pair (t, x) , for $t \in W$ and x in some neighborhood of u .

(iv) $P\{\exists t \in W, Z(t) = u, \det(Z'(t)) = 0\} = 0$

Then, for every Borel set B contained in W , one has

$$(6.26) \quad E(N_u(Z, B)) = \int_B E(|\det(Z'(t))| | Z(t) = u) p_{Z(t)}(u) dt.$$

If B is compact, both sides are finite.

2. Random fields from \mathbb{R}^d to $\mathbb{R}^{d'}$, $d > d'$

We will follow a similar method to the one of the previous section, but a certain number of new problems arise.

For f a \mathcal{C}^1 -function defined on W and u a regular value of f , we will denote $\sigma_u(f, T)$ the $(d-d')$ geometric measure of the intersection of the set T with the level set $C_u(f, U) = \{t \in U : f(t) = u\}$. Note that since u is a regular value at each point $t \in C_u(f, U)$ the jacobian matrix $f'(t)$ is of full rank d' . Thus we can choose a subset γ of $\{1, \dots, d\}$ of size d' such that the matrix

$\{\frac{\partial f_i}{\partial t_j}, i = 1, \dots, d', j \in \gamma\}$ is invertible. For simplicity, and without loss of generality, we will assume that $\gamma^c = \{1, \dots, (d - d')\}$. Using the implicit function theorem we know that there exist a neighbourhood V_t of $t_1, \dots, t_{d-d'}$ and a function $g: \mathbb{R}^{d-d'} \rightarrow \mathbb{R}^{d'}$ (that depend on t) such that $s_1, \dots, s_{d-d'}, g(s_1, \dots, s_{d-d'})$ is a local parameterization of the level set $C_u(f, U)$. This defines a chart and proves that the level set is a C^1 manifold of dimension $d - d'$.

2.1. We start with three statements for Gaussian random fields that are analogous to those of Theorems 6.2, 6.3 and 6.4.

THEOREM 6.8 (Rice formula for the expectation of the geometric measure of the level set). *Let $Z: W \rightarrow \mathbb{R}^{d'}$ be a random field, W an open subset of \mathbb{R}^d and $u \in \mathbb{R}^{d'}$ a fixed point. Assume that:*

(i) Z is Gaussian.

(ii) Almost surely the function $t \rightsquigarrow Z(t)$ is of class C^1 .

(iii) For each $t \in W$, $Z(t)$ has a non degenerate distribution (i.e. $\text{Var}(Z(t)) \succ 0$).

(iv) $P\{\exists t \in W, Z(t) = u, Z'(t) \text{ does not have full rank}\} = 0$

Then, for every Borel set B contained in W , one has

$$(6.27) \quad E(\sigma_u(Z, B)) = \int_B E\left([\det(Z'(t)(Z'(t))^T)]^{1/2} | Z(t) = u\right) p_{Z(t)}(u) dt.$$

If B is compact, both sides in (6.27) are finite.

THEOREM 6.9 (Rice formula for the k -th moment). *Let $k, k \geq 2$ be an integer. Assume the same hypotheses as in Theorem 6.8 except for (iii) that is replaced by*

(iii') *for distinct values $t_1, \dots, t_k \in W$ of the parameter, the distribution of*

$$(Z(t_1), \dots, Z(t_k))$$

does not degenerate in $(\mathbb{R}^{d'})^k$.

Then for every Borel set B contained in W and levels u_1, \dots, u_k one has

$$(6.28) \quad E\left(\prod_{j=1}^k \sigma_{u_j}(Z, B)\right) = \int_{B^k} E\left(\prod_{j=1}^k [\det(Z'(t_j)(Z'(t_j))^T)]^{1/2} | Z(t_1) = u_1, \dots, Z(t_k) = u_k\right) p_{Z(t_1), \dots, Z(t_k)}(u_1, \dots, u_k) dt_1 \dots dt_k,$$

where both members may be infinite.

The same kind of result holds true for integrals over the level set, as stated in the next theorem

THEOREM 6.10 (Expected integral on the level set). *Let Z be a random field that verifies the hypotheses of Theorem 6.8. Assume that for each $t \in W$ one has another random field $Y^t: V \rightarrow \mathbb{R}^n$, where V is some topological space, verifying the following conditions:*

a) $Y^t(v)$ is a measurable function of (ω, t, v) and almost surely, $(t, v) \rightsquigarrow Y^t(v)$ is continuous.

b) For each $t \in W$ the random process $(s, v) \rightsquigarrow (Z(s), Y^t(v))$ defined on $W \times V$ is Gaussian.

Moreover, assume that $g: W \times C(V, \mathbb{R}^n) \rightarrow \mathbb{R}$ is a bounded function, which is continuous when one puts on $C(V, \mathbb{R}^n)$ the topology of uniform convergence on compact sets. Then, for each compact subset I of W , one has

$$(6.29) \quad E\left(\int_{I \cap Z^{-1}(u)} g(t, Y^t) \sigma_u(Z, dt)\right) = \int_I E([\det(Z'(t)(Z'(t))^T)]^{1/2} g(t, Y^t) | Z(t) = u) \cdot p_{Z(t)}(u) dt.$$

2.2. Remark on hypothesis (iv) of Theorems 6.8, 6.9 and 6.10. .

Let us give sufficient conditions to assure that hypothesis (iv) holds true, i.e., that with probability one the given level u is not a critical value of the random field. They are more restrictive than the ones for $d = d'$, and based upon the following proposition, which is a generalization of Bulinskaya's Lemma 1.20 (see Exercise 6.4 for a proof).

PROPOSITION 6.11. *Let $\mathcal{Y} = \{Y(t) : t \in W\}$ be a random field with values in \mathbb{R}^{m+k} and W an open subset of \mathbb{R}^d . m and k are positive integers. Let $u \in \mathbb{R}^{m+k}$ and I a subset of W .*

We assume that \mathcal{Y} satisfies the following conditions:

- *the paths $t \rightsquigarrow Y(t)$ are of class \mathcal{C}^1 ,*
- *for each $t \in W$, the random vector $Y(t)$ has a density and there exists a constant C such that*

$$p_{Y(t)}(x) \leq C$$

for $t \in I$ and x in some neighborhood of u ,

- *the Hausdorff dimension of I is smaller or equal than m .*

Then, almost surely, there is no point $t \in I$ such that $Y(t) = u$

This implies the following:

PROPOSITION 6.12. *Let $\mathcal{Z} = \{Z(t) : t \in W\}$ be a random field, W an open subset of \mathbb{R}^d , with values in $\mathbb{R}^{d'}$. Let $u \in \mathbb{R}^{d'}$.*

We assume:

- *the paths of \mathcal{Z} are of class \mathcal{C}^2 ,*
- *for each $t \in W$, the pair $(Z(t), Z'(t))$ has a joint density $p_{Z(t), Z'(t)}(x, x')$ in $\mathbb{R}^{d'} \times \mathbb{R}^{d \cdot d'}$, which is bounded for (t, x') varying in a compact subset of $W \times \mathbb{R}^{d \cdot d'}$ and x in some neighborhood of u .*

Then, (iv) holds true.

PROOF. . Apply Proposition 6.11 to the random field

$$Y(t, \lambda) = (Z(t) : (Z'(t))^T \lambda)$$

defined for $(t, \lambda) \in W \times S^{d'-1}$ with values in $\mathbb{R}^{d'} \times \mathbb{R}^d$. □

2.3. Scheme of the proofs of Theorems 6.8, 6.9, 6.10. .

We are not going to give full proofs of these theorems, since as we have already mentioned, they follow the same lines of the analogous ones for $d' = d$. We will limit ourselves to point out the differences between both situations.

First, we need a proposition replacing the Area Formula 6.1 for non-random functions. We state it for \mathcal{C}^1 functions, since this will be sufficient for our purposes.

PROPOSITION 6.13 (Co-area formula). *Let f be a \mathcal{C}^1 function defined on an open subset W of \mathbb{R}^d taking values in $\mathbb{R}^{d'}$. Assume that the set of critical values of f has zero Lebesgue measure. Let $g: \mathbb{R}^{d'} \rightarrow \mathbb{R}$ be continuous and bounded.*

Then,

$$(6.30) \quad \int_{\mathbb{R}^{d'}} g(u) \sigma_u(f, B) du = \int_B [\det (f'(t)(f'(t))^T)]^{1/2} g(f(t)) dt.$$

for any Borel subset B of W , whenever the integral in the right-hand side is well defined.

Remarks on formula (6.30).

1. Clearly, this extends the Area formula, since if $d = d'$, $\sigma_u(f, T)$ is the number points of the level set on T , i.e., $N_u(f, T)$.

2. For a proof of Proposition 6.13, under more general conditions, we refer the reader to Federer's book (1969).

3. The remark 3. after Proposition 6.1 applies here in the same way. This means that if we replace the function $g(u)$ in (6.30) by any measurable function $h(t, u)$, one obtains the weighted co-area formula:

$$(6.31) \quad \int_{\mathbb{R}^{d'}} du \int_{\mathbb{R}^d} h(t, u) \sigma_u(f, dt) = \int_{\mathbb{R}^d} [\det (f'(t)(f'(t))^T)]^{1/2} h((t, f(t))) dt,$$

whenever the right-hand side is well defined.

Second, let us now enumerate the changes required in the proof of Theorem 6.2 to obtain Theorem 6.8.

- Replace $\Delta(t)$ by

$$\bar{\Delta}(t) = [\det (Z'(t)(Z'(t))^T)]^{1/2}.$$

- Whenever u is not a critical value of $Z(\cdot)$, instead of $C_u^m(B)$ and $Q_u^{n,m}(B)$ that were defined in (6.8) and (6.9), we put respectively:

$$(6.32) \quad c_u^m(B) = \int_B F_m(\bar{\Delta}(s)) \sigma_u(Z, ds)$$

and

$$(6.33) \quad q_u^{n,m}(B) = c_u^m(B) G_n(c_u^m(B)).$$

- Instead of (6.10), we have:

$$(6.34) \quad E(q_u^{n,m}(B)) = \int_B E[\bar{\Delta}(t) F_m(\bar{\Delta}(t)) G_n(c_u^m(B)) | Z(t) = u] p_{Z(t)}(u) dt,$$

which holds true for almost every $u \in \mathbb{R}^{d'}$. This follows from the weighted co-area formula (6.31).

To finish one performs two additional steps: 1) proving that both hand-sides in equality (6.34) are continuous functions of u , so that equality holds true for all $u \in \mathbb{R}^{d'}$, and 2) passing to the limit as $n \rightarrow \infty$, $m \rightarrow \infty$, in that order.

On the left-hand side of (6.34) the first step follows from the continuity provided by the Implicit Function Theorem and the obvious inequality $q_u^{n,m}(B) \leq n$. The second step by monotone convergence.

For the remaining of the proof, we proceed as in the one of Theorem 6.2.

Third, let us now consider the formula (6.28) for the higher moments of $\sigma_u(Z, B)$ in Theorem 6.9. Define (as in the case $d' = d$) the random process

$$\tilde{Z}(t_1, \dots, t_k) = (Z(t_1), \dots, Z(t_k)).$$

with parameter set $W^k \subset (\mathbb{R}^d)^k$ and values in $(\mathbb{R}^{d'})^k$.

If u_1, \dots, u_k are regular values of Z , then $(u_1, \dots, u_k) \in (\mathbb{R}^{d'})^k$ is a regular value of \tilde{Z} and

$$\sigma_{(u_1, \dots, u_k)}(\tilde{Z}, B_1 \times \dots \times B_k) = \sigma_{u_1}(B_1) \dots \sigma_{u_k}(B_k)$$

for any choice of the Borel subsets B_1, \dots, B_k of W . Following the same reasoning as in the proof of Theorem 6.3, one only needs to prove that the measure of the diagonal set

$$D_k(I) = \{(t_1, \dots, t_k) \in I^k, t_i = t_j \text{ for some pair } i, j, i \neq j\}$$

that is, $\sigma_{(u_1, \dots, u_k)}(\tilde{Z}, D_k(I))$ vanishes for any rectangle $I \subset W$. To see this, notice that $\tilde{Z}^{-1}(u_1, \dots, u_k)$ is a differentiable manifold with dimension $k(d-d')$ which carries the geometric measure $\sigma_{(u_1, \dots, u_k)}(\tilde{Z}, \cdot)$ and its intersection with $D_k(I)$ is a finite union of sub-manifolds having dimension smaller or equal to $((k-1)(d-d'))$, so its geometric measure is zero.

One should notice that when $u_1 = \dots = u_k$ there is a difference between the case $d = d'$ and the case $d > d'$, since in the first one, the diagonal charges a positive geometric measure. In fact, in this case, all the manifolds are 0-dimensional and the argument in the previous paragraph does not work. That is why when $d = d'$ one actually gets the integral formula for the *factorial* moments of the number of roots. The difference between ordinary and factorial moments of order k is the expectation of the measure carried by the diagonal $D_k(B)$.

Fourth. The proof of Theorem 6.10 does not require any new ingredients with respect to the one of 6.4.

Exercises

EXERCISE 6.1. (a) Assume that Z_1, Z_2 are \mathbb{R}^d -valued random fields defined on compact subsets I_1, I_2 of \mathbb{R}^d and suppose that (Z_i, I_i) ($i = 1, 2$) satisfy the hypotheses of Theorem 6.2 and that for every $s \in I_1$ and $t \in I_2$, the distribution of $(Z_1(s), Z_2(t))$ does not degenerate.

Prove that for each pair $u_1, u_2 \in \mathbb{R}^d$:

$$(6.35) \quad \mathbb{E}(N_{u_1}^{Z_1}(I_1)N_{u_2}^{Z_2}(I_2)) \\ = \int_{I_1 \times I_2} \mathbb{E}(|\det(Z_1'(t_1))||\det(Z_2'(t_2))||Z_1(t_1) = u_1, Z_2(t_2) = u_2) p_{Z_1(t_1), Z_2(t_2)}(u_1, u_2) dt_1 dt_2.$$

(b) Extend (a) to higher moments.

EXERCISE 6.2. Let $\mathcal{X} = \{X(t) : t \in \mathbb{R}^d\}$ be a real-valued centered Gaussian random field. We denote

$$r(s, t) = \mathbb{E}(X(s)X(t)), \quad s, t \in \mathbb{R}^d$$

its covariance. We assume that the process is stationary, in the sense that $r(s, t) = \Gamma(s - t)$ for all $s, t \in \mathbb{R}^d$.

(a) Let the function Γ be continuous. Prove Bochner's Theorem (see Chapter 1), that is, there exists a unique Borel measure on \mathbb{R}^d , say μ such that for all $\tau \in \mathbb{R}^d$:

$$\Gamma(\tau) = \int_{\mathbb{R}^d} \exp[i\langle \tau, x \rangle] \mu(dx)$$

μ is called the spectral measure of the random field \mathcal{X}

(b) Denote

$$\Lambda_2 = \int_{\mathbb{R}^d} \|x\|^2 \mu(dx)$$

which can be finite or infinite.

Prove that Γ is twice differentiable at the origin if and only if Λ_2 is finite, and in this case Γ is a \mathcal{C}^2 -function and its partial derivatives can be computed by means of the formula:

$$\frac{\partial^2 \Gamma}{\partial \tau_j \partial \tau_k}(\tau) = - \int_{\mathbb{R}^d} x_j x_k \exp[i\langle \tau, x \rangle] \mu(dx)$$

for $j, k = 1, \dots, d$, with the notation $\tau = (\tau_1, \dots, \tau_d)^T$, $x = (x_1, \dots, x_d)^T$.

(c) Prove that if the field has \mathcal{C}^1 sample paths then $\Lambda_2 < \infty$. Let I is a Borel subset of \mathbb{R}^d , then:

$$\mathbb{E}(\sigma_u(X, I)) = \lambda_d(I) \cdot \phi(u) \cdot \mathbb{E}(\|\xi\|),$$

where λ_d denotes Lebesgue measure in \mathbb{R}^d , $\phi(u)$ is the standard normal density and ξ is a $N[0, \Lambda_2]$ Gaussian random variable with values in \mathbb{R}^d .

EXERCISE 6.3. Let $\mathcal{X} = \{X(t) : t \in \mathbb{R}^d\}$ be a random field defined as:

$$X(t) = X_1^2(t) + \dots + X_m^2(t),$$

where $\{X_k(t) : t \in \mathbb{R}^d\}_{k=1, \dots, m}$ are m independent random fields, each one of them being centered Gaussian stationary with covariance function Γ (see the previous exercise).

Prove that for each Borel subset I of \mathbb{R}^d , one has:

$$\mathbf{E}(\sigma_u(X, I)) = 2 \sqrt{u} \lambda_d(I) \cdot \chi_m^2(u) \cdot \mathbf{E}(\|\xi\|),$$

where ξ is as in the previous exercise and $\chi_m^2(u)$ is the χ^2 density with m degrees of freedom, that is, the density of the random variable $\|\eta\|^2$, where η is standard normal in \mathbb{R}^m .

EXERCISE 6.4. Prove Propositions 6.11 and 6.12.

Regularity of the Distribution of the Maximum

In this chapter, except in Theorem 7.4, we will only consider Gaussian processes and our purpose is to give an account of what is known on the regularity of the probability distribution of the supremum. The main classical result is Tsirelson's Theorem (1975). We begin with a statement of this theorem, as it can be found in Lifshits' book (1995), to which we refer also for the proof.

THEOREM 7.1 (Tsirelson). *Let $\{X(t), t \in T\}$ be a real-valued bounded Gaussian process defined on a countable parameter set T .*

Then, the distribution F_M of the random variable $M = \sup_{t \in T} X(t)$ has the following properties:

- (1) *It is continuous on \mathbb{R} , except at most at one point: the left limit of its support, that is*

$$u_0 := \inf\{u : F_M(u) > 0\}.$$

- (2) *It is absolutely continuous on the half-line $(u_0, +\infty)$.*
 (3) *It is differentiable on $(u_0, +\infty)$ except for at most a countable set E .*
 (4) *The derivative F' is positive and continuous on $(u_0, +\infty) \setminus E$. At each point of E , the derivative F' has left and right limits and jumps downwards.*
 (5) *For each $u > u_0$, F' has finite variation on $[u, +\infty)$*
 (6) *F' is the density of M on $(u_0, +\infty)$.*

Further improvements are in Weber (1985), Lifshits (1995), Diebolt and Posse (1996) and references therein. One should notice that in this statement of Tsirelson's Theorem, the parameter set is countable. This says that the same result holds true for separable bounded Gaussian processes, since in this case, almost surely the distribution of the supremum coincides with the one of the supremum on some countable non-random set.

Our aim in this chapter is to go beyond these regularity properties of the distribution of M , at the cost of imposing a certain number of conditions on the process. In fact, we will require the parameter set to have a certain geometric structure and the paths of the process to have a certain regularity. This will permit to exploit the analytic properties of the paths to obtain results about the distribution function F_M .

The theorems we present are much stronger in the case of one-parameter processes, than in the case of random fields. In the first case we are able to give a considerable extension of Tsirelson's-type properties. For example, we prove that if a Gaussian process defined on a compact interval T of the real line has C^∞ -paths and its law satisfies a quite general non-degeneracy condition, then the distribution of its maximum is a function of class C^∞ . For multiparameter processes (random fields) much less is known and we will only prove results on the first derivative of F_M .

For one-parameter processes the main results are taken from Azaïs and Wschebor (2001). The proofs here are simpler than the original version due to some technical improvements that we present in Section 1, where we start with an implicit formula for the density of the maximum of a Gaussian random field defined on a subset of \mathbb{R}^d . This will be our main tool in this chapter, and we will see later on that it is also useful as a tool to study the asymptotic properties of the tails of the distribution of the maximum. Its proof is extracted from Azaïs and Wschebor (2007).

1. The implicit formula for the density of the maximum

Assumptions and notations. $\mathcal{X} = \{X(t) : t \in S\}$ denotes a real-valued Gaussian field defined on the parameter set S . We assume that S satisfies the hypotheses:

A1 :

- S is a compact subset of \mathbb{R}^d
- S is the disjoint union of S_d, S_{d-1}, \dots, S_0 , where S_j is an orientable C^3 manifold of dimension j without boundary. The S_j 's will be called faces. Let S_{d_0} , $d_0 \leq d$ be the non empty face having largest dimension. σ_j denotes the j -dimensional geometric measure on S_j .
- We will assume that each S_j has an atlas such that the second derivatives of the inverse functions of all charts (viewed as diffeomorphisms from an open set in \mathbb{R}^j to S_j) are bounded by a fixed constant. For $t \in S_j$, we denote L_t the maximum curvature of S_j at the point t . It follows that L_t is bounded for $t \in S$.

Notice that the decomposition $S = S_d \cup \dots \cup S_0$ is not unique.

Concerning the random field we make the following assumptions A2-A5

A2 : \mathcal{X} is defined on an open set containing S and has C^2 paths

A3 : for every $t \in S$ the distribution of $(X(t), X'(t))$ does not degenerate; for every $s, t \in S$, $s \neq t$, the distribution of $(X(s), X(t))$ does not degenerate.

A4 : Almost surely the maximum of $X(t)$ on S is attained at a single point.

For $t \in S_j$, $X'_j(t)$ $X'_{j,N}(t)$ denote respectively the derivative along S_j and the normal derivative. Both quantities are viewed as vectors in \mathbb{R}^d , and the density of their distribution will be expressed respectively with respect to an orthonormal basis of the tangent space $T_{t,j}$ of S_j at the point t , or its orthogonal complement $N_{t,j}$. $X''_j(t)$ will denote the second derivative of X along S_j , at the point $t \in S_j$ and will be viewed as a matrix expressed in an orthogonal basis of $T_{t,j}$. Similar notations will be used for any function defined on S_j .

A5 : Almost surely, for every $j = 1, \dots, d$ there is no point t in S_j such that $X'_j(t) = 0$, $\det(X''_j(t)) = 0$.

The fundamental property that we will use is the representation of the density of the maximum given in the next theorem.

THEOREM 7.2. *Let $M = \max_{t \in S} X(t)$. Under assumptions A1 to A5, the distribution of M has the density*

$$(7.1) \quad p_M(x) = \sum_{t \in S_0} \mathbb{E}(\mathbf{1}_{A_x} | X(t) = x) p_{X(t)}(x) + \sum_{j=1}^d \int_{S_j} \mathbb{E}(|\det(X''_j(t))| \mathbf{1}_{A_x} | X(t) = x, X'_j(t) = 0) p_{X(t), X'_j(t)}(x, 0) \sigma_j(dt),$$

where $A_x = \{M \leq x\}$.

Remarks: This equality is stated in terms of the density, but it is obvious that one also obtains an exact (implicit) formula for the distribution of the maximum on integrating once both sides of (7.1).

One can replace $|\det(X''_j(t))|$ in the conditional expectation by $(-1)^j \det(X''_j(t))$, since under the conditioning and whenever the event $\{M \leq x\}$ holds true, $X''_j(t)$ is negative semi-definite.

PROOF. Let $N_j(u)$, $j = 0, \dots, d$ be the number of global maxima of $X(\cdot)$ on S that belong to S_j and are larger than u . From the hypotheses it follows that a.s. $\sum_{j=0, \dots, d} N_j(u)$ is equal to 0 or to 1, so that

$$(7.2) \quad \mathbb{P}\{M > u\} = \sum_{j=0, \dots, d} \mathbb{P}\{N_j(u) = 1\} = \sum_{j=0, \dots, d} \mathbb{E}(N_j(u)).$$

The proof will be finished as soon as we show that each term in (7.2) is the integral over $(u, +\infty)$ of the corresponding term in (7.1). This is self-evident for $j = 0$.

Let us consider the term $j = d$. We apply the weighted Rice formula (6.6) of Chapter 6, as follows:

- Z is the random field X' defined on S_d .
- For each $t \in S_d$, put $W = S$ and $Y^t : S \rightarrow \mathbb{R}^2$ defined as:

$$Y^t(w) := (X(w) - X(t), X(t)).$$

Notice that the second coordinate in the definition of Y^t does not depend on w .

- In the place of the function g , we take for each $n = 1, 2, \dots$ the function g_n defined as follows:

$$g_n(t, f_1, f_2) = g_n(f_1, f_2) = (1 - \mathcal{F}_n(\sup_{w \in S} f_1(w))) \cdot (1 - \mathcal{F}_n(u - f_2(\bar{w}))),$$

where \bar{w} is any point in W and for n a positive integer and $x \geq 0$, we define as in formula 6.7 :

$$(7.3) \quad \mathcal{F}_n(x) := \mathcal{F}(nx) ; \quad \text{with } \mathcal{F}(x) = 0 \text{ if } 0 \leq x \leq 1/2, \quad \mathcal{F}(x) = 1 \text{ if } x \geq 1,$$

and \mathcal{F} monotone non-decreasing and continuous.

It is easy to check that all the requirements in Theorem 6.4 are satisfied, so that, for the value 0 instead of u in formula (6.6) we get:

$$(7.4) \quad \mathbb{E}\left(\sum_{t \in S_d, X'(t)=0} g_n(Y^t)\right) = \int_{S_d} \mathbb{E}(|\det(X''(t))| g_n(Y^t) | X'(t) = 0) \cdot p_{X'(t)}(0) \lambda_d(dt).$$

Notice that the formula holds true for each compact subset of S_d in the place of S_d , hence for S_d itself by monotone convergence.

Let now $n \uparrow \infty$ in (7.4). Clearly $g_n(Y^t) \downarrow \mathbb{1}_{X(s)-X(t) \leq 0, \forall s \in S} \cdot \mathbb{1}_{X(t) \geq u}$. The passage to the limit does not present any difficulty since $0 \leq g_n(Y^t) \leq 1$ and the sum in the left-hand side is bounded by the random variable $N_0^{X'}(\bar{S}_d)$, which is in L^1 because of Rice Formula. We get

$$\mathbb{E}(N_d(u)) = \int_{S_d} \mathbb{E}(|\det(X''(t))| \mathbb{1}_{X(s)-X(t) \leq 0, \forall s \in S} \mathbb{1}_{X(t) \geq u} | X'(t) = 0) \cdot p_{X'(t)}(0) \lambda_d(dt)$$

Conditioning on the value of $X(t)$, we obtain the desired formula for $j = d$.

The proof for $1 \leq j \leq d-1$ is essentially the same, but one must take care of the parameterization of the manifold S_j . One can first establish locally the formula on a chart of S_j , using local coordinates.

It can be proved as in Proposition 6.6 (the only modification is due to the term $\mathbb{1}_{A_x}$) that the quantity written in some chart as

$$\mathbb{E}(\det(Y''(s)) \mathbb{1}_{A_x} | Y(s) = x, Y'(s) = 0) p_{Y(s), Y'(s)}(x, 0) ds,$$

where the random field $Y(s)$ is \mathcal{X} written in some chart of S_j , i.e. $Y(s) = X(\phi^{-1}(s))$, defines a j -form, that is, a measure on S_j that does not depend on the parameterization and which has a density with respect to the Lebesgue measure ds in every chart. It can be proved that the integral of this j -form on S_j gives the expectation of $N_j(u)$.

To get formula (7.1) it suffices to consider locally around a precise point $t \in S_j$ the chart ϕ given by the projection on the tangent space at t . In this case we obtain that at t

- ds is in fact $\sigma_j(dt)$
- $Y'(s)$ is isometric to $X'_j(t)$

where $s = \phi(t)$.

This finishes the proof. □

2. One parameter processes

In this section we restrict the scope of our study to Gaussian processes defined on a compact interval of the line. Without loss of generality, we will assume this interval to be $[0, 1]$. As announced this will enable us to obtain deeper results on the regularity of the distribution of M . The statement of the main theorem is the following:

Let $\mathcal{X} = \{X(t) : t \in [0, 1]\}$ be a stochastic process with real values. It is said to satisfy the hypothesis H_k , k a positive integer, if:

- (1) \mathcal{X} is Gaussian;
- (2) a.s. \mathcal{X} has \mathcal{C}^k sample paths;
- (3) For every integer $n \geq 1$ and any set t_1, \dots, t_n of distinct parameter values, the distribution of the random vector:

$$X(t_1), \dots, X(t_n), X'(t_1), \dots, X'(t_n), \dots, X^{(k)}(t_1), \dots, X^{(k)}(t_n)$$

is non-degenerate.

We denote $m(t)$ and $r(s, t)$ the mean and covariance functions of X and use the notation $r_{ij} := \frac{\partial^{i+j}}{\partial s^i \partial t^j} r$ ($i, j = 0, 1, \dots$) for the derivatives, whenever they exist.

It is in general a non-trivial task to verify condition (3). However, for stationary Gaussian processes a simple sufficient condition on the spectral measure which implies (3) is given in Exercise 3.5.

THEOREM 7.3. *Assume that \mathcal{X} satisfies H_{2k} . Denote by $F(u) = P(M \leq u)$ the distribution function of $M = \max_{t \in [0, 1]} X(t)$.*

Then, F is of class \mathcal{C}^k and its successive derivatives can be computed by repeated application of Lemma 7.7 below.

Theorem 7.3 for random processes with one parameter appears to be a considerable extension of Theorem 7.1. For example, it implies that if the process is Gaussian with \mathcal{C}^∞ paths and satisfies the non-degeneracy condition for every $k=1, 2, \dots$, then the distribution of the maximum is \mathcal{C}^∞ . The same methods we will be using in the proof, provide also bounds for the successive derivatives. The asymptotic behaviour as their argument tends to $+\infty$ will be considered in Chapter 8, when we will study a certain number of asymptotic methods related to the distribution of the maximum.

Before proceeding to the proof of Theorem 7.3, which turns out to be quite long and presents a number of technical difficulties, let us make a parenthesis to state two theorems on the density of the maximum, which are easier to prove and provide simple inequalities for the density of the maximum. The first one, Theorem 7.4 refers to general, not necessarily Gaussian processes. The second one, Theorem 7.5 concerns Gaussian processes. As applications, one gets upper and lower bounds for the density of M under conditions which otherwise have required complicated calculations and unnecessary restrictions.

THEOREM 7.4. *Assume that the process $\mathcal{X} = \{X(t) : t \in [0, 1]\}$ has \mathcal{C}^2 -paths, that for each $t \in [0, 1]$, the triplet $(X(t), X'(t), X''(t))$ admits a joint density and $X'(t)$ has a bounded density $p_{X'(t)}(\cdot)$. We also assume that the function*

$$I(x, z) := \int_0^1 \mathbb{E}(X''(t) | X(t) = x, X'(t) = z) p_{X(t), X'(t)}(x, z) dt$$

is uniformly continuous in z for (x, z) in some neighborhood of $(u, 0)$. Then the distribution of M admits a density $p_M(\cdot)$ satisfying a.e.

$$(7.5) \quad p_M(u) \leq P(X'(0) < 0 | X(0) = u) p_{X(0)}(u) + P(X'(1) > 0 | X(1) = u) p_{X(1)}(u) \\ + \int_0^1 \mathbb{E}(X''(t) | X(t) = x, X'(t) = 0) p_{X(t), X'(t)}(x, 0) dt$$

PROOF. Let $u \in \mathbb{R}$ and $h > 0$. We have

$$\begin{aligned} \mathbb{P}(M \leq u) - \mathbb{P}(M \leq u - h) &= \mathbb{P}(u - h < M \leq u) \leq \\ &\leq \mathbb{P}(u - h < X(0) \leq u, X'(0) < 0) + \mathbb{P}(u - h < X(1) \leq u, X'(1) > 0) + \mathbb{P}(M_{u-h,u}^+ > 0), \end{aligned}$$

where $M_{u-h,u}^+ = M_{u-h,u}^+(0, 1)$, since if $u - h < M \leq u$, then either the maximum occurs in the interior of the interval $[0, 1]$ or at 0 or 1, with the derivative taking the indicated sign. Notice that

$$\mathbb{P}(M_{u-h,u}^+ > 0) \leq \mathbb{E}(M_{u-h,u}^+).$$

Using Proposition 1.20, with probability 1, $X'(\cdot)$ has no tangencies at the level 0, thus an upper bound for this expectation follows from Kac's formula:

$$M_{u-h,u}^+ = \lim_{\delta \rightarrow 0} \frac{1}{2\delta} \int_0^1 \mathbb{1}_{\{X(t) \in [u-h, u]\}} \mathbb{1}_{\{X'(t) \in [-\delta, \delta]\}} \mathbb{1}_{\{X''(t) < 0\}} |X''(t)| dt \quad a.s.$$

which together with Fatou's lemma imply:

$$\mathbb{E}(M_{u-h,u}^+) \leq \liminf_{\delta \rightarrow 0} \frac{1}{2\delta} \int_{-\delta}^{\delta} dz \int_{u-h}^u I(x, z) dx = \int_{u-h}^u I(x, 0) dx.$$

Combining this bound with the preceding one, we get

$$\begin{aligned} \mathbb{P}(M \leq u) - \mathbb{P}(M \leq u - h) &\leq \int_{u-h}^u [\mathbb{P}(X'(0) < 0 | X(0) = x) p_{X(0)}(x) \\ &\quad + \mathbb{P}(X'(1) > 0 | X(1) = x) p_{X(1)}(x) + I(x, 0)] dx, \end{aligned}$$

which gives the result. \square

In spite of the simplicity of the proof, already in the case of Gaussian processes, this theorem provides, under quite general conditions, an upper bound for the density which is hard to improve (see for example Diebolt and Posse (1996)). If we consider Gaussian, centered process with unit variance, by means of a deterministic time change, one can also assume that the process has "unit speed" (which means that $\text{Var}(X'(t)) \equiv 1$). This transforms the interval $[0, 1]$ into an interval having length say L . Then, one can prove (see Exercise 7.2) that (7.5) reduces to:

$$(7.6) \quad p_M(u) \leq p^+(u) := \varphi(u) \left[1 + (2\pi)^{-1/2} \int_0^L C(t) \varphi(u/C(t)) + u \Phi(u/C(t)) dt \right],$$

with $C(t) := \sqrt{r_{22}(t, t) - 1}$.

As $u \rightarrow +\infty$,

$$(7.7) \quad p^+(u) = \varphi(u) \left[1 + Lu(2\pi)^{-1/2} + (2\pi)^{-1/2} u^{-2} \int_0^L C^3(t) \varphi(u/C(t)) dt \right] + O(u^{-4} \varphi(u/C^+)),$$

with $C^+ := \sup_{t \in [0, L]} C(t)$.

The following theorem is a direct consequence of Theorem 7.2. The only point which is new is the continuity of the density which will be proved later on.

THEOREM 7.5. *Suppose that X is a Gaussian process with C^2 paths and such for all $s, t, s \neq t \in [0, 1]$, $X(s), X(t), X'(t)$ and $X(t), X'(t), X''(t)$ admit a joint density. Then M has a **continuous density** p_M given for every u by*

$$(7.8) \quad \begin{aligned} p_M(u) &= \mathbb{P}(M \leq u | X(0) = u) p_{X(0)}(u) + \mathbb{P}(M \leq u | X(1) = u) p_{X(1)}(u) \\ &\quad + \int_0^1 \mathbb{E}(X''^-(t) \mathbb{1}_{M \leq u} | X(t) = u, X'(t) = 0) p_{X(t), X'(t)}(u, 0) dt \end{aligned}$$

Using (7.8) one can obtain sharper upper-bounds than those produced by (7.5). See Exercise 7.1.

We now turn to the proofs of our main results.

2.1. Proofs. We start with an auxiliary technical lemma.

LEMMA 7.6. (a) Let $\mathcal{Z} = \{Z(t) : t \in [0, 1]\}$ be a **centered** stochastic process satisfying H_k ($k \geq 2$) and t a point in $(0, 1)$. Define the Gaussian processes $Z^0(s), Z^1(s), Z^t(s)$ by means of the orthogonal decompositions:

$$(7.9) \quad Z(s) = a^0(s)Z(0) + sZ^0(s) \quad s \in (0, 1].$$

$$(7.10) \quad Z(s) = a^1(s)Z(1) + (1-s)Z^1(s) \quad s \in [0, 1).$$

$$(7.11) \quad Z(s) = b^t(s)Z(t) + c^t(s)Z'(t) + \frac{(s-t)^2}{2}Z^t(s) \quad s \in [0, 1] \quad s \neq t.$$

Then, the processes Z^0, Z^1, Z^t can be extended continuously at $s = 0, s = 1, s = t$ respectively so that they satisfy $H_{k-1}, H_{k-1}, H_{k-2}$ respectively.

Notice that in fact the functions a^0, a^1, b^t and c^t are the ordinary regression coefficients.

(b) Let f be any function of class \mathcal{C}^k . When there is no ambiguity on the process \mathcal{Z} , we will define f^0, f^1, f^t in the same manner, putting f instead of Z in (7.9), (7.10), (7.11), but still keeping the regression coefficients corresponding to \mathcal{Z} . Then f^0, f^1, f^t can be extended by continuity in the same way to functions in $\mathcal{C}^{k-1}, \mathcal{C}^{k-1}, \mathcal{C}^{k-2}$ respectively.

(b') As a consequence, if \mathcal{Z} is a process satisfying H_k which is not centered, we can define Z^0, Z^1, Z^t using (7.9), (7.10), (7.11) applied separately to the centered process $t \rightarrow Z(t) - \mathbb{E}(Z(t))$ and to the mean $\mathbb{E}(Z(t))$ and summing up the two components. In fact we obtain again (7.9), (7.10), (7.11)

(c) Let m be a positive integer, suppose $Z(t)$ satisfies H_{2m+1} and t_1, \dots, t_m belong to $[0, 1]$. Denote by $Z^{t_1, \dots, t_m}(s)$ the process obtained by repeated application of the operation of part (a) of this lemma, that is

$$Z^{t_1, \dots, t_m}(s) = (Z^{t_1, \dots, t_{m-1}})^{t_m}(s).$$

Denote by s_1, \dots, s_p ($p \leq m$) the ordered p -tuple of the elements of t_1, \dots, t_m that belong to $(0, 1)$ (i.e. they are neither 0 nor 1). Then, a.s. the application:

$$(s_1, \dots, s_p, s) \rightsquigarrow (Z^{t_1, \dots, t_m}(s), (Z^{t_1, \dots, t_m})'(s))$$

is continuous.

PROOF. (a) and (b) follow in a direct way, by computing the regression coefficients $a^0(s), a^1(s), b^t(s), c^t(s)$, substituting into formulas (7.9), (7.10), (7.11) and using the arguments above. We prove now (c) which is a consequence of the following:

Suppose $Z(t_1, \dots, t_k)$ is a Gaussian field with \mathcal{C}^p sample paths ($p \geq 2$) defined on $[0, 1]^k$ with no degeneracy in the same sense that in the definition of hypothesis $H_k(3)$ for one-parameter processes. Then the Gaussian fields defined by means of:

$$\begin{aligned} Z^0(t_1, \dots, t_k) &= (t_k)^{-1} (Z(t_1, \dots, t_{k-1}, t_k) - a^0(t_1, \dots, t_k)Z(t_1, \dots, t_{k-1}, 0)) \quad \text{for } t_k \neq 0, \\ Z^1(t_1, \dots, t_k) &= (1 - t_k)^{-1} (Z(t_1, \dots, t_{k-1}, t_k) - a^1(t_1, \dots, t_k)Z(t_1, \dots, t_{k-1}, 1)) \quad \text{for } t_k \neq 1, \\ \tilde{Z}(t_1, \dots, t_k, t_{k+1}) &= 2(t_{k+1} - t_k)^{-2} (Z(t_1, \dots, t_{k-1}, t_{k+1}) - b(t_1, \dots, t_k, t_{k+1})Z(t_1, \dots, t_k) - \\ &\quad - c(t_1, \dots, t_k, t_{k+1}) \frac{\partial Z}{\partial t_k}(t_1, \dots, t_k)) \quad \text{for } t_{k+1} \neq t_k \end{aligned}$$

can be extended to $[0, 1]^k$ (respectively $[0, 1]^k, [0, 1]^{k+1}$) into fields with paths in \mathcal{C}^{p-1} (respectively $\mathcal{C}^{p-1}, \mathcal{C}^{p-2}$). In the above formulas,

- $a^0(t_1, \dots, t_k)$ is the regression coefficient of $Z(t_1, \dots, t_k)$ on $Z(t_1, \dots, t_{k-1}, 0)$,
- $a^1(t_1, \dots, t_k)$ is the regression coefficient of $Z(t_1, \dots, t_k)$ on $Z(t_1, \dots, t_{k-1}, 1)$,
- $b(t_1, \dots, t_k, t_{k+1}), c(t_1, \dots, t_k, t_{k+1})$ are the regression coefficients of $Z(t_1, \dots, t_{k-1}, t_{k+1})$ on the pair $\left(Z(t_1, \dots, t_k), \frac{\partial Z}{\partial t_k}(t_1, \dots, t_k) \right)$.

Let us prove the statement about $\tilde{Z}(t_1, \dots, t_k, t_{k+1})$. The other two are simpler. Suppose for the moment that $\tilde{Z}(t_1, \dots, t_k, t_{k+1})$ is centered. Denote by V the subspace of $L^2(\Omega, \mathfrak{F}, P)$

generated by the pair $\left(Z(t_1, \dots, t_k), \frac{\partial Z}{\partial t_k}(t_1, \dots, t_k) \right)$. Denote by Π_{V^\perp} the version of the orthogonal projection of $L^2(\Omega, \mathfrak{F}, P)$ on the orthogonal complement of V , which is:

$$\Pi_{V^\perp}(Y) := Y - \left[bZ(t_1, \dots, t_k) + c \frac{\partial Z}{\partial t_k}(t_1, \dots, t_k) \right],$$

where b and c are the regression coefficients of Y on the pair

$$Z(t_1, \dots, t_k), \frac{\partial Z}{\partial t_k}(t_1, \dots, t_k).$$

If $\{Y(\theta) : \theta \in \Theta\}$ is a random field with continuous paths and such that $\theta \rightarrow Y(\theta)$ is continuous in $L^2(\Omega, \mathfrak{F}, P)$, then a.s.

$$(\theta, t_1, \dots, t_k) \rightarrow \Pi_{V^\perp}(Y_\theta)$$

is continuous.

From the definition:

$$\tilde{Z}(t_1, \dots, t_k, t_{k+1}) = 2(t_{k+1} - t_k)^{-2} \Pi_{V^\perp}(Z(t_1, \dots, t_{k-1}, t_{k+1})).$$

On the other hand, by Taylor's formula:

$$Z(t_1, \dots, t_{k-1}, t_{k+1}) = Z(t_1, \dots, t_k) + (t_{k+1} - t_k) \frac{\partial Z}{\partial t_k}(t_1, \dots, t_k) + R_2(t_1, \dots, t_k, t_{k+1})$$

with

$$R_2(t_1, \dots, t_k, t_{k+1}) = \int_{t_k}^{t_{k+1}} \frac{\partial^2 Z}{\partial t_k^2}(t_1, \dots, t_{k-1}, \tau) (t_{k+1} - \tau) d\tau$$

so that

$$(7.12) \quad \tilde{Z}(t_1, \dots, t_k, t_{k+1}) = \Pi_{V^\perp} \left[2(t_{k+1} - t_k)^{-2} R_2(t_1, \dots, t_k, t_{k+1}) \right].$$

It is clear that the paths of the random field \tilde{Z} are $p-1$ times continuously differentiable for $t_{k+1} \neq t_k$. Relation (7.12) shows that they have a continuous extension to $[0, 1]^{k+1}$ with $\tilde{Z}(t_1, \dots, t_k, t_k) = \Pi_{V^\perp} \left(\frac{\partial^2 Z}{\partial t_k^2}(t_1, \dots, t_k) \right)$. In fact,

$$\begin{aligned} \Pi_{V^\perp} \left(2(s_{k+1} - s_k)^{-2} R_2(s_1, \dots, s_k, s_{k+1}) \right) &= \\ &= 2(s_{k+1} - s_k)^{-2} \int_{s_k}^{s_{k+1}} \Pi_{V^\perp} \left(\frac{\partial^2 Z}{\partial t_k^2}(s_1, \dots, s_{k-1}, \tau) \right) (s_{k+1} - \tau) d\tau. \end{aligned}$$

A.s. the integrand is a continuous function of the parameters therein so that, a.s.:

$$\tilde{Z}(s_1, \dots, s_k, s_{k+1}) \rightarrow \Pi_{V^\perp} \left(\frac{\partial^2 Z}{\partial t_k^2}(t_1, \dots, t_k) \right) \text{ when } (s_1, \dots, s_k, s_{k+1}) \rightarrow (t_1, \dots, t_k, t_k)$$

This proves (c) in case \tilde{Z} is centered.

It remains to consider the case when \tilde{Z} is purely deterministic, say $\tilde{Z}(t_1, \dots, t_k) = f(t_1, \dots, t_k)$. Making t_{k+1} tend to t_k in the regression equation, we see that

$$(7.13) \quad b(t_1, \dots, t_k, t_k) = 1 \quad c(t_1, \dots, t_k, t_k) = 0$$

$$(7.14) \quad \frac{\partial b}{\partial t_k}(t_1, \dots, t_k, t_k) = 0 \quad \frac{\partial c}{\partial t_k}(t_1, \dots, t_k, t_k) = 1$$

So that $b(t_1, \dots, t_k, t_{k+1}) = 1 + O((t_k - t_{k+1})^2)$ and $c(t_1, \dots, t_k, t_{k+1}) = t_{k+1} - t_k + O((t_k - t_{k+1})^2)$ and

$$\tilde{f}(t_1, \dots, t_k, t_{k+1}) = 2(t_{k+1} - t_k)^{-2} [f(t_1, \dots, t_{k-1}, t_{k+1}) - Z(t_1, \dots, t_k) - (t_{k+1} - t_k) \frac{\partial Z}{\partial t_k}(t_1, \dots, t_k)] + O(1)$$

The result is now a simple consequence of the Taylor formula. In the same way, when $p \geq 3$, we obtain the continuity of the partial derivatives of \tilde{Z} up to the order $p-2$. \square

PROOF OF THEOREM 7.5.

We apply Theorem 7.2. It is easy to check that whenever the parameter set is a compact interval in the line, it is not necessary the process to be defined in a neighborhood of S , since this assumption is in fact used to define the derivative at the boundary of S . In the case of an interval, we simply use one-sided derivatives at the extremes.

The conditions required for (7.8) to hold true are fulfilled.

Set $\beta(t) \equiv 1$. Then:

- For $t \neq 0, 1$, under the condition $X(t) = u, X'(t) = 0$, the event $\{M \leq u\}$ can be written $\{\forall s \in [0, 1], X^t(s) \leq \beta^t(s)u\}$ with the notation of Lemma 7.6. This event will be denoted $A_u(X^t, \beta^t)$.
- For $t \neq 0, 1$, under the condition $X(t) = u, X'(t) = 0, X''(t)$ is equal to $X^t(t)$
- under the condition $X(0) = u$, the event $\{M \leq u\}$ is equal to $A_u(X^0, \beta^0)$.
- under the condition $X(1) = u$, the event $\{M \leq u\}$ is equal to $A_u(X^1, \beta^1)$.

We prove, for example, that

$$P\{A_u(X^t, \beta^t)\}$$

is a continuous function of u . Let $h > 0$. We have the inequalities:

$$(7.15) \quad \begin{aligned} & |P\{A_u(X^t, \beta^t)\} - P\{A_{u-h}(X^t, \beta^t)\}| \leq P(\{A_u(X^t, \beta^t)\} \setminus \{A_{u-h}(X^t, \beta^t)\}) + P(\{A_{u-h}(X^t, \beta^t)\} \setminus \{A_u(X^t, \beta^t)\}) \\ & \leq P\left\{ \sup_{s \in [0,1]} (X^t(s) - \beta^t(s)u) \in [-h\|\beta^t\|_\infty, 0] \right\} + P\left\{ \sup_{s \in [0,1]} (X^t(s) - \beta^t(s)u) \in [0, h\|\beta^t\|_\infty] \right\} \end{aligned}$$

Now, apply Ylvisaker's Theorem 1.21 to prove that the expression above tends to zero as $h \rightarrow 0$, which proves the continuity of $P\{A_u(X^t, \beta^t)\}$. Similar arguments can be applied to prove that each one of the three terms in the right-hand side of (7.8) is a continuous function of u . \square

Our next lemma is the basic technical tool to prove the fundamental Theorem 7.3.

LEMMA 7.7. *Suppose $\mathcal{Z} = \{Z(t) : t \in [0, 1]\}$ is a stochastic process that verifies H_2 . Define:*

$$F_v(u) = E\{\xi_v \cdot \mathbf{1}_{A_u}\}$$

where

$$A_u = A_u(Z, \beta) = \{Z(t) \leq \beta(t)u \text{ for all } t \in [0, 1]\},$$

$\beta(\cdot)$ is a real valued C^2 function defined on $[0, 1]$,

$\xi_v = G(Z(t_1) - \beta(t_1)v), \dots, Z(t_m) - \beta(t_m)v)$ for some positive integer $m, t_1, \dots, t_m \in [0, 1], v \in \mathbb{R}$ and some C^∞ function $G : \mathbb{R}^m \rightarrow \mathbb{R}$ having at most polynomial growth at ∞ , that is, $|G(x)| \leq C(1 + \|x\|^p)$ for some positive constants C, p and all $x \in \mathbb{R}^m$.

Then, for each $v \in \mathbb{R}$, F_v is of class C^1 and its derivative is a continuous function of the pair (u, v) that can be written in the form:

$$(7.16) \quad \begin{aligned} F'_v(u) &= \beta(0)E\{\xi_{v,u}^0 \cdot \mathbf{1}_{A_u(Z^0, \beta^0)}\} \cdot p_{Z(0)}(\beta(0) \cdot u) \\ &\quad + \beta(1)E\{\xi_{v,u}^1 \cdot \mathbf{1}_{A_u(Z^1, \beta^1)}\} \cdot p_{Z(1)}(\beta(1) \cdot u) \\ &\quad - \int_0^1 \beta(t)E\{\xi_{v,u}^t (Z_t^t - \beta^t(t) \cdot u) \mathbf{1}_{A_u(Z^t, \beta^t)}\} \cdot p_{Z(t), Z'(t)}(\beta(t) \cdot u, \beta'(t) \cdot u) dt, \end{aligned}$$

PROOF.

$$A_u \setminus A_{u-h} = A_u \cap \left[u - h < \sup_{t: \beta(t) > 0} \frac{Z(t)}{\beta(t)} \leq u \right].$$

The set $B^+ := \{t \in [0, 1] : \beta(t) > 0\}$ is open in $[0, 1]$, so it is a countable union of disjoint open intervals: $B^+ = \bigcup_n (a_n, b_n)$ (The reader may notice that intervals having the form $[0, b_n)$ or $(a_n, 1]$ may be present). By a monotone convergence argument,

$$E(\xi_v \mathbf{1}_{A_u \setminus A_{u-h}}) = \lim_{N \rightarrow +\infty} \lim_{\varepsilon \rightarrow 0} E[\xi_v \mathbf{1}_{\{u-h < \sup_{t \in B_N^+} Y(t) \leq u\}}],$$

where

$$Y(t) := \frac{Z(t)}{\beta(t)}, \quad B_\varepsilon^N := \cup_{n=1}^N (a_n + \varepsilon; b_n - \varepsilon)$$

and ε is small enough so that B_ε^N is well-defined. Since $\beta(t)$ is bounded away from zero on B_ε^N , the conditions of Theorem 7.2 are fulfilled by the process $Y(t)$. Putting the weights ξ_v at every global maxima, we get

(7.17)

$$\begin{aligned} \mathbb{E}(\xi_v \mathbf{1}_{A_u \setminus A_{u-h}}) &= \lim_{N \rightarrow +\infty} \lim_{\varepsilon \rightarrow 0} \sum_{n=1}^N \int_{a_n - \varepsilon}^{b_n + \varepsilon} dt \int_{u-h}^u dx \mathbb{E}[\xi_v | Y''(t) | \mathbf{1}_{A_x} | Y(t) = x, Y'(t) = 0] p_{Y(t), Y'(t)}(x, 0) \\ &+ \sum_{n=1}^N \int_{u-h}^u \left[\mathbb{E}[\xi_v \mathbf{1}_{A_x} | Y(a_n + \varepsilon) = x] p_{Y(a_n + \varepsilon)}(x) + \mathbb{E}[\xi_v \mathbf{1}_{A_x} | Y(b_n - \varepsilon) = x] p_{Y(b_n - \varepsilon)}(x) \right] dx \end{aligned}$$

Changing variables in each integral, we get:

$$\begin{aligned} \mathbb{E}[\xi_v \mathbf{1}_{A_u \setminus A_{u-h}}] &= \lim_{N \rightarrow +\infty} \lim_{\varepsilon \rightarrow 0} \sum_{n=1}^N \int_{a_n - \varepsilon}^{b_n + \varepsilon} dt \int_{u-h}^u dx \\ (7.18) \quad &\beta(t) \mathbb{E}[\xi_v | Z''(t) - \beta''(t)x | \mathbf{1}_{A_x} | Z(t) = \beta(t)x, Z'(t) = \beta'(t)x] p_{Z(t), Z'(t)}(\beta(t).x, \beta'(t)x) \\ &+ \sum_{n=1}^N \int_{u-h}^u \left[\beta(a_n + \varepsilon) \mathbb{E}[\xi_v \mathbf{1}_{A_x} | Z(a_n + \varepsilon) = \beta(a_n + \varepsilon)x] p_{Z(a_n + \varepsilon)}(\beta(a_n + \varepsilon).x) \right. \\ &\left. + \beta(b_n - \varepsilon) \mathbb{E}[\xi_v \mathbf{1}_{A_x} | Z(b_n - \varepsilon) = \beta(b_n - \varepsilon)x] p_{Z(b_n - \varepsilon)}(\beta(b_n - \varepsilon).x) \right] dx \end{aligned}$$

We get, as in the proof of Theorem 7.8:

(7.19)

$$\begin{aligned} \mathbb{E}[\xi_v \mathbf{1}_{A_u \setminus A_{u-h}}] &= \lim_{N \rightarrow +\infty} \lim_{\varepsilon \rightarrow 0} \sum_{n=1}^N \int_{a_n - \varepsilon}^{b_n + \varepsilon} dt \int_{u-h}^u dx \beta(t) \mathbb{E}[\xi_{v,u}^t | Z^t(t) - \beta^t(t).x | \mathbf{1}_{A_x(Z^t, \beta^t)}] p_{Z(t), Z'(t)}(\beta(t).x, \beta'(t)x) \\ &+ \sum_{n=1}^N \int_{u-h}^u \left[\beta(a_n + \varepsilon) \mathbb{E}[\xi_v \mathbf{1}_{A_x} | Z(a_n + \varepsilon) = \beta(a_n + \varepsilon)x] p_{Z(a_n + \varepsilon)}(\beta(a_n + \varepsilon).x) \right. \\ &\left. + \beta(b_n - \varepsilon) \mathbb{E}[\xi_v \mathbf{1}_{A_x} | Z(b_n - \varepsilon) = \beta(b_n - \varepsilon)x] p_{Z(b_n - \varepsilon)}(\beta(b_n - \varepsilon).x) \right] dx \end{aligned}$$

In the present form we can see that the conditional expectations and the densities appearing in the right-hand side of this formula are bounded, so excepting the case where $a_n = 0$ or $b_n = 1$, the contribution of the points $a_n + \varepsilon$ and $b_n - \varepsilon$ tends to zero as $\varepsilon \rightarrow 0$ since $\beta(a_n) = \beta(b_n) = 0$. Letting $\varepsilon \rightarrow 0$ and N tend to $+\infty$ in that order we obtain:

(7.20)

$$\begin{aligned} \mathbb{E}(\xi_v \mathbf{1}_{A_u \setminus A_{u-h}}) &= \int_{B^+} \beta(t) dt \int_{u-h}^u dx \mathbb{E}[\xi_{v,x}^t | Z^t(t) - \beta^t(t).x | \mathbf{1}_{A_x(Z^t, \beta^t)}] p_{Z(t), Z'(t)}(\beta(t).x, \beta'(t)x) \\ &+ \int_{u-h}^u dx (\beta(0))^+ \mathbb{E}[\xi_{v,x}^0 | \mathbf{1}_{A_x(Z^0, \beta^0)}] p_{Z(0)}(\beta(0).x) \\ &+ \int_{u-h}^u dx (\beta(1))^+ \mathbb{E}[\xi_{v,x}^1 | \mathbf{1}_{A_x(Z^1, \beta^1)}] p_{Z(1)}(\beta(1).x). \end{aligned}$$

Lemma 7.7 shows that the integrand is a continuous function of x , so that, taking into account also of the sign of $Z^t(t) - \beta^t(t).u$ inside the expectation, we obtain:

(7.21)

$$\begin{aligned} \lim_{h \downarrow 0} \frac{1}{h} \mathbb{E}(\xi_v \mathbf{1}_{A_u \setminus A_{u-h}}) &= - \int_{B^+} \beta(t) \mathbb{E}[\xi_{v,u}^t | Z^t(t) - \beta^t(t).u | \mathbf{1}_{A_u(Z^t, \beta^t)}] p_{Z(t), Z'(t)}(\beta(t).u, \beta'(t)x) dt \\ &+ (\beta(0))^+ \mathbb{E}[\xi_{v,u}^0 | \mathbf{1}_{A_u(Z^0, \beta^0)}] p_{Z(0)}(\beta(0).u) + (\beta(1))^+ \mathbb{E}[\xi_{v,u}^1 | \mathbf{1}_{A_u(Z^1, \beta^1)}] p_{Z(1)}(\beta(1).u). \end{aligned}$$

Similar computations - that we will not perform here - show an analogous result for

$$\lim_{h \downarrow 0} \frac{1}{h} \mathbb{E} \{ \xi_v \cdot \mathbf{1}_{A_{u-h} \setminus A_u} \}$$

This finishes the proof of the lemma. \square

PROOF OF THEOREM 7.3.

We proceed by induction on k . We will give some details for the first two derivatives including some implicit formulas that will illustrate the procedure for general k .

We introduce the following additional notations. Put $Y(t) := X(t) - \beta(t)u$ and define, on the interval $[0, 1]$, the processes $X^0, X^1, X^t, Y^0, Y^1, Y^t$, and the functions $\beta^0, \beta^1, \beta^t$, as in Lemma 7.6. Notice that the regression coefficients corresponding to the processes X and Y are the same, so that anyone of them may be used to define the functions $\beta^0, \beta^1, \beta^t$. One can easily check that

$$Y^0(s) = X^0(s) - \beta^0(s)u$$

$$Y^1(s) = X^1(s) - \beta^1(s)u$$

$$Y^t(s) = X^t(s) - \beta^t(s)u.$$

For $t_1, \dots, t_m \in [0, 1]$, $m \geq 2$, we define by induction the stochastic processes $X^{t_1, \dots, t_m} = (X^{t_1, \dots, t_{m-1}})^{t_m}$, $Y^{t_1, \dots, t_m} = (Y^{t_1, \dots, t_{m-1}})^{t_m}$ and the function $\beta^{t_1, \dots, t_m} = (\beta^{t_1, \dots, t_{m-1}})^{t_m}$, applying Lemma 7.6 for the computations at each stage.

With the aim of somewhat reducing the size of the formulas we will express the successive derivatives in terms of the processes Y^{t_1, \dots, t_m} instead of X^{t_1, \dots, t_m} . The reader must keep in mind that for each m -tuple t_1, \dots, t_m the results depend on u through the expectation of the stochastic process Y^{t_1, \dots, t_m} . Also, for a stochastic process Z we will use the notation

$$A(Z) = A_0(Z, \beta) = \{Z(t) \leq 0 : \text{for all } t \in [0, 1]\}$$

First derivative

Suppose that X satisfies H_2 . We apply formula (7.16) in Lemma 7.7 for $\xi \equiv 1$, $Z = X$ and $\beta(\cdot) \equiv 1$ obtaining for the first derivative:

(7.22)

$$F'(u) = \mathbb{E} [\mathbf{1}_{A(Y^0)}] p_{Y(0)}(0) + \mathbb{E} [\mathbf{1}_{A(Y^1)}] p_{Y(1)}(0) - \int_0^1 \mathbb{E} [Y^{t_1}(t_1) \mathbf{1}_{A(Y^{t_1})}] p_{Y(t_1), Y'(t_1)}(0, 0) dt_1.$$

This expression is exactly the expression in (7.8) with the notational changes just mentioned and after taking profit of the fact that the process is Gaussian, via regression on the condition in each term. Notice that according to the definition of the Y -processes:

$$\mathbb{E} [\mathbf{1}_{A(Y^0)}] = \mathbb{E} [\mathbf{1}_{A_u(X^0, \beta^0)}]$$

$$\mathbb{E} [\mathbf{1}_{A(Y^1)}] = \mathbb{E} [\mathbf{1}_{A_u(X^1, \beta^1)}]$$

$$\mathbb{E} [Y^{t_1}(t_1) \mathbf{1}_{A(Y^{t_1})}] = \mathbb{E} [Y^{t_1}(t_1) \mathbf{1}_{A_u(X^{t_1}, \beta^{t_1})}].$$

Second derivative.

Suppose that X satisfies H_4 . Then, X^0, X^1, X^{t_1} satisfy H_3, H_3, H_2 respectively. Therefore Lemma 7.7 applied to these processes can be used to show the existence of $F''(u)$ and to compute a similar formula, except for the necessity of justifying differentiation under the integral sign in

the third term. We get the expression:

$$(7.23) \quad F''(u) = -\mathbb{E}[\mathbb{1}_{A(Y^0)}]p_{Y(0)}^{(1)}(0) - \mathbb{E}[\mathbb{1}_{A(Y^1)}]p_{Y(1)}^{(1)}(0) \\ + \int_0^1 \mathbb{E}[Y^{t_1}(t_1)\mathbb{1}_{A(Y^{t_1})}]p_{Y(t_1),Y'(t_1)}^{(1,0)}(0,0)dt_1 \\ + p_{Y(0)}(0) \left\{ \beta^0(0)\mathbb{E}[\mathbb{1}_{A(Y^{0,0})}]p_{Y^0(0)}(0) + \beta^0(1)\mathbb{E}[\mathbb{1}_{A(Y^{0,1})}]p_{Y^0(1)}(0) \right\} \\ - \int_0^1 \beta^0(t_2)\mathbb{E}[Y^{0,t_2}(t_2)\mathbb{1}_{A(Y^{0,t_2})}]p_{Y^0(t_2),(Y^0)'(t_2)}(0,0)dt_2 + \\ + p_{Y(1)}(0) \left\{ \beta^1(0)\mathbb{E}[\mathbb{1}_{A(Y^{1,0})}]p_{Y^1(0)}(0) + \beta^1(1)\mathbb{E}[\mathbb{1}_{A(Y^{1,1})}]p_{Y^1(1)}(0) \right\} \\ - \int_0^1 \beta^1(t_2)\mathbb{E}[Y^{1,t_2}(t_2)\mathbb{1}_{A(Y^{1,t_2})}]p_{Y^1(t_2),(Y^1)'(t_2)}(0,0)dt_2 - \\ - \int_0^1 p_{Y(t_1),Y'(t_1)}(0,0) \left\{ \begin{array}{l} -\beta^{t_1}(t_1)\mathbb{E}[\mathbb{1}_{A(Y^{t_1})}] + \beta^{t_1}(0)\mathbb{E}[Y^{t_1,0}(t_1)\mathbb{1}_{A(Y^{t_1,0})}]p_{Y^{t_1}(0)}(0) + \\ + \beta^{t_1}(1)\mathbb{E}[Y^{t_1,1}(t_1)\mathbb{1}_{A(Y^{t_1,1})}]p_{Y^{t_1}(1)}(0) - \\ - \int_0^1 \beta^{t_1}(t_2)\mathbb{E}[Y^{t_1,t_2}(t_1)Y^{t_1,t_2}(t_2)\mathbb{1}_{A(Y^{t_1,t_2})}]p_{Y^{t_1}(t_2),(Y^{t_1})'(t_2)}(0,0)dt_2 \end{array} \right\} dt_1,$$

In this formula $p_{Y(t_0)}^{(1)}$, $p_{Y(t_1)}^{(1)}$ and $p_{Y(t_1),Y'(t_1)}^{(1,0)}(0,0)$ stand respectively for the derivative of $p_{Y(t_0)}(\cdot)$, the derivative of $p_{Y(t_1)}(\cdot)$ and the derivative with respect to the first variable of $p_{Y(t_1),Y'(t_1)}(\cdot, \cdot)$. To validate the above formula, notice the following points:

- The first two lines are obtained by differentiating with respect to u , the densities $p_{Y(0)}(0) = p_{X(0)}(-u)$, $p_{Y(1)}(0) = p_{X(1)}(-u)$, $p_{Y(t_1),Y'(t_1)}(0,0) = p_{X(t_1),X'(t_1)}(-u,0)$.
- Lines 3 and 4 come from the application of Lemma 7.7 to differentiate $\mathbb{E}[1_{A(Y^0)}]$. The lemma is applied with $Z = X^0$, $\beta = \beta^0$, $\xi = 1$.
- Similarly, lines 5 and 6 contain the derivative of $\mathbb{E}[1_{A(Y^1)}]$.
- The remaining corresponds to differentiate the function

$$\mathbb{E}[Y^{t_1}(t_1)\mathbb{1}_{A(Y^{t_1})}] = \mathbb{E}[(X^{t_1}(t_1) - \beta^{t_1}(t_1)u)\mathbb{1}_{A_u(X^{t_1},\beta^{t_1})}]$$

in the integrand of the third term in (7.22). The first term in line 7 comes from the simple derivative

$$\frac{\partial}{\partial v}\mathbb{E}[(X^{t_1}(t_1) - \beta^{t_1}(t_1)v)\mathbb{1}_{A_u(X^{t_1},\beta^{t_1})}] = -\beta^{t_1}(t_1)\mathbb{E}(\mathbb{1}_{A(Y^{t_1})}).$$

The other terms are obtained by applying Lemma 7.7 to compute

$$\frac{\partial}{\partial u}\mathbb{E}[(X^{t_1}(t_1) - \beta^{t_1}(t_1)v)\mathbb{1}_{A_u(X^{t_1},\beta^{t_1})}],$$

putting $Z = X^{t_1}$, $\beta = \beta^{t_1}$, $\xi = X^{t_1}(t_1) - \beta^{t_1}(t_1)v$.

- Finally, differentiation under the integral sign is valid since because of Lemma 7.6, the derivative of the integrand is a continuous function of (t_1, t_2, u) due to regularity and non-degeneracy of the Gaussian distributions involved, and the application of Ylvisaker's Theorem.

General case.

With the above notation, given the m -tuple t_1, \dots, t_m of elements of $[0, 1]$ we will call the processes $Y, Y^{t_1}, Y^{t_1, t_2}, \dots, Y^{t_1, \dots, t_{m-1}}$ the "ancestors" of Y^{t_1, \dots, t_m} . In the same way we define the ancestors of the function β^{t_1, \dots, t_m} .

Assume the following induction hypothesis: If X satisfies H_{2k} then F is k times continuously differentiable and $F^{(k)}$ is the sum of a finite number of terms belonging to the class D_k which consists of all expressions of the form:

$$(7.24) \quad \int_0^1 \dots \int_0^1 ds_1 \dots ds_p Q(s_1, \dots, s_p) \mathbb{E}[\xi \mathbb{1}_{A(Y^{t_1, \dots, t_m})}] K_1(s_1, \dots, s_p) K_2(s_1, \dots, s_p)$$

where:

- $1 \leq m \leq k$.
- $t_1, \dots, t_m \in [0, 1]$, $m \geq 1$.

- s_1, \dots, s_p , $0 \leq p \leq m$, are the elements in $\{t_1, \dots, t_m\}$ that belong to $(0, 1)$ (that is, which are neither “0” nor “1”). When $p = 0$ no integral sign is present.
- $Q(s_1, \dots, s_p)$ is a polynomial in the variables s_1, \dots, s_p .
- ξ is a product of values of Y^{t_1, \dots, t_m} at some locations belonging to $\{s_1, \dots, s_p\}$.
- $K_1(s_1, \dots, s_p)$ is a product of values of some ancestors of β^{t_1, \dots, t_m} at some locations belonging to the set $\{s_1, \dots, s_p\} \cup \{0, 1\}$.
- $K_2(s_1, \dots, s_p)$ is a sum of products of densities and derivatives of densities of the random variables $Z(\tau)$ at the point 0, or the pairs $(Z(\tau), Z'(\tau))$ at the point $(0, 0)$ where $\tau \in \{s_1, \dots, s_p\} \cup \{0, 1\}$ and the process Z is some ancestor of Y^{t_1, \dots, t_m} .

Notice that K_1 does not depend on u but K_2 is a function of u .

It is clear that the induction hypothesis is verified for $k = 1$. Assume that it is true up to the integer k and that X satisfies H_{2k+2} . Then $F^{(k)}$ can be written as a sum of terms of the form (7.24). Consider a term of this form and observe that the variable u may appear in three locations:

- (1) In ξ , where differentiation is simple given its product form, the fact that $\frac{\partial}{\partial u} Y^{t_1, \dots, t_q}(s) = -\beta^{t_1, \dots, t_q}(s)$, $q \leq m$, $s \in \{s_1, \dots, s_p\}$ and the boundedness of moments allowing to differentiate under the integral and expectation signs.
- (2) In $K_2(s_1, \dots, s_p)$ which is clearly C^∞ as a function of u . Its derivative with respect to u takes the form of a product of functions of the types $K_1(s_1, \dots, s_p)$ and $K_2(s_1, \dots, s_p)$ defined above.
- (3) In $1_{A(Y^{t_1, \dots, t_m})}$. Lemma 7.7 shows that differentiation produces 3 terms depending upon the processes $Y^{t_1, \dots, t_m, t_{m+1}}$ with t_{m+1} belonging to $(0, 1) \cup \{0, 1\}$. Each term obtained in this way belongs to D_{k+1} .

The proof is achieved by taking into account that, as in the computation of the second derivative, Lemma 7.6 implies that the derivatives of the integrands are continuous functions of u that are bounded as functions of $(s_1, \dots, s_p, t_{m+1}, u)$ if u varies in a bounded set. \square

The statement and proof of Theorem 7.3 can not, of course, be used to obtain explicit expressions for the derivatives of the distribution function F . However, the implicit formula for $F^{(k)}(u)$ as sum of elements of D_k can be transformed into explicit upper-bounds if one replaces everywhere the indicator functions $\mathbb{1}_{A(Y^{t_1, \dots, t_m})}$ by 1 and the functions $\beta^{t_1, \dots, t_m}(\cdot)$ by their absolute value.

On the other hand, Theorem 7.3 permits to have the exact asymptotic behaviour of $F^{(k)}(u)$ as $u \rightarrow +\infty$ in case $\text{Var}(X_t)$ is constant. Even though the number of terms in the formula increases rapidly with k , there is exactly one term that is dominant. It turns out that as $u \rightarrow +\infty$, $F^{(k)}(u)$ is equivalent to the k -th derivative of the equivalent of $F(u)$. We will come back to this point in Chapter 8.

3. Continuity of the density of the maximum of random fields

Let $S \subset \mathbb{R}^d$, $d > 1$ and $\mathcal{X} = \{X(t) : t \in U\}$ a real-valued random field defined on some open neighborhood U of S . We assume that S and \mathcal{X} satisfy the assumptions A1 to A5 of Section 1. We know from Theorem 7.2 that the probability distribution of the maximum $M = \max_{t \in S} X(t)$ has a density p_M which verifies equality (7.1).

As we did in Theorem 7.5, corresponding to the one-dimensional case, we will now prove that p_M is continuous. The problem is more difficult here, since the equivalent of Lemma 7.6 is harder.

The result is the following :

THEOREM 7.8. *In addition to the above assumptions, for every $s \in S$, $t \in S_j$ $j = 0, 1, \dots, d_0$, $s \neq t$, the joint distributions of the triplets $(X(t), X'(t), X''(t))$, $(X(s), X(t), X'_j(t))$ do not degenerate. Then the density p_M given by (7.1) is continuous.*

PROOF. For $t \in S_j$ and $s \in S$, we define the normalization

$$n(t, s) := \|(s - t)_{j,N}\| + \frac{1}{2}\|s - t\|^2.$$

where $(s - t)_{j,N}$ is the normal component of $(s - t)$, i.e. its orthogonal projection onto the subspace $N_{t,j}$ (see Section 1 for the notation).

For $s \neq t, t \in S_j$, define $X^t(s)$ by means of the Gaussian regression:

$$(7.25) \quad X(s) = a^t(s)X(t) + \langle b^t(s), X'(t) \rangle + n(t, s)X^t(s),$$

The reader should notice that $X^t(s)$ is not the same as before, since the present normalization is different. Let us consider the expression in the right-hand side of (7.1). A dominated convergence argument shows that it is enough to prove the continuity of the integrands appearing in that formula. Such an integrand has the form:

$$\mathbb{E}(|\det(X_j''(t))| \mathbf{1}_{A_x} | X(t) = x, X_j'(t) = 0) p_{X(t), X_j'(t)}(x, 0) \sigma_j(dt)$$

The density $p_{X(t), X_j'(t)}(x, 0)$ is clearly a continuous function of x . The regression of $X_j''(t)$ on the condition does not present any problem and the only remaining point is to check that

$$\mathbb{P}(A_x | X(t) = x, X_j'(t) = 0)$$

is a continuous function of x . We write this as an unconditional probability, using the regression formula 7.25, so that it becomes:

$$\mathbb{P}\{Y^t(s) \leq \gamma^t(s) x \text{ for all } s \in S, s \neq t\},$$

where

$$Y^t(s) = X^t(s) + \frac{\langle b^t(s), X'_{j,N}(t) \rangle}{n(t, s)}$$

and

$$\gamma^t(s) = \frac{1 - a^t(s)}{n(t, s)}$$

On the other hand, from the regression formula it follows easily that the \mathcal{C}^2 -functions a^t and b^t , verify:

$$a^t(t) = 1, (a^t)'(t) = 0, b^t(t) = 0, (b^t)'(t) = Id,$$

where Id is the identity in \mathbb{R}^d .

The reader can check that there exist positive constants K, c and $\bar{\gamma}$ such that for all $s \in S, s \neq t$, one has:

$$(7.26) \quad \begin{aligned} \mathbb{E}(Y^t(s)) &\leq K \\ \text{Var}(Y^t(s)) &\geq c \\ |\gamma(s, t)| &\leq \bar{\gamma} \end{aligned}$$

Let us denote $C_x := \{Y^t(s) \leq \gamma^t(s) x \text{ for all } s \in S, s \neq t\}$ and let $h > 0$. Clearly:

$$|\mathbb{P}\{C_x\} - \mathbb{P}\{C_{x-h}\}| \leq \mathbb{P}\{C_x \setminus C_{x-h}\} + \mathbb{P}\{C_{x-h} \setminus C_x\},$$

and it suffices to prove that both terms in the right-hand side are small if h is small. Let us do it for the first one, the second is similar. We have:

$$(7.27) \quad \mathbb{P}\{C_x \setminus C_{x-h}\} \leq \mathbb{P}\{-h \cdot \bar{\gamma} < \sup_{s \in S, s \neq t} [Y^t(s) - \gamma(s, t) \cdot x] \leq 0\}.$$

To finish, observe that the random field $\{Y^t(s) - \gamma(s, t) \cdot x : s, t \in S, s \neq t\}$ verifies the hypothesis of Theorem 1.22. So, the distribution of its supremum has no atom in \mathbb{R} . This shows that the right-hand side of (7.27) tends to zero as $h \rightarrow 0$, and we are done. \square

Remark.

The last proof exhibits the main technical difference between one-parameter processes and multidimensional parameter ones, in what concerns the study of the regularity of the distribution of the maximum. The random field $\{X^t(s) : s \in S, s \neq t\}$ is constructed after Gaussian regression and re-normalization. In the multidimensional parameter case, it does not have a limit as $s \rightarrow t$, the paths present a “helix behavior” as one approaches the point t .

In the one-dimensional parameter case, this new process can be extended continuously to $s = t$, and the extension also preserves a part of the regularity of the original process. So, we are able to iterate the procedure, re-normalize again, and continue in this way. This is the basis of the proof of Theorem 7.3 that we are unable to reproduce for general Gaussian random fields.

Exercises

EXERCISE 7.1. (a) Check the following inequalities:

$$\begin{aligned} \mathbb{P}(M \leq u | X_0 = u) &= \mathbb{P}(M \leq u, X'(0) < 0 | X(0) = u) \\ &\geq \mathbb{P}(X'(0) < 0 | X(0) = u) - \mathbb{E}(U_u[0, 1] \mathbf{1}_{\{X'(0) < 0\}} | X(0) = u). \end{aligned}$$

$$\begin{aligned} \mathbb{P}(M \leq u | X(1) = u) &= \mathbb{P}(M \leq u, X'(1) > 0 | X(1) = u) \geq \\ &\geq \mathbb{P}(X'(1) > 0 | X(1) = u) - \mathbb{E}(D_u[0, 1] \mathbf{1}_{\{X'(1) > 0\}} | X(1) = u) \end{aligned}$$

If $x'' < 0$:

$$\begin{aligned} \mathbb{P}(M \leq u | X(t) = u, X'(t) = 0, X''(t) = x'') &\geq \\ &\geq 1 - \mathbb{E}([D_u([0, t]) + U_u([t, 1])] | X(t) = u, X'(t) = 0, X''(t) = x''). \end{aligned}$$

(b) Using the inequalities in (a) prove the following lower bound for the density of the maximum: (7.28)

$$\begin{aligned} p_M(u) &\geq \mathbb{P}(M \leq u | X(0) = u) p_{X(0)}(u) + \mathbb{P}(M \leq u | X(1) = u) p_{X(1)}(u) \\ &+ \int_0^1 \mathbb{E}(X''^-(t) \mathbf{1}_{M \leq u} | X(t) = x, X'(t) = 0) p_{X(t), X'(t)}(x, 0) dt \\ &- \int_0^1 ds \int_{-\infty}^0 dx' \int_0^{+\infty} x'_s p_{X(s), X'(s), X(0), X'(0)}(u, x'_s, u, x') dx'_s - \\ &- \int_0^1 dt \int_{-\infty}^0 |x''| \left[\int_0^t ds \int_{-\infty}^0 |x'| p_{X(s), X'(s), X(t), X'(t), X''(t)}(u, x', u, 0, x') dx' \right. \\ &\quad \left. + \int_t^1 ds \int_0^{+\infty} x' p_{X(s), X'(s), X(t), X'(t), X''(t)}(u, x', u, 0, x'') dx' \right] dx''. \end{aligned}$$

EXERCISE 7.2. (a) Prove inequality (7.6).

(b) Prove (7.7) which gives an asymptotic bound for $p_M(u)$ as $u \rightarrow +\infty$.

EXERCISE 7.3. Let $\{X(s, t) : s, t \in \mathbb{R}\}$ be a real-valued two-parameter Gaussian, centered stationary isotropic random field with covariance $\Gamma, \Gamma(0) = 1$. Assume that its spectral measure μ is absolutely continuous with respect to Lebesgue measure in \mathbb{R}^2 with density

$$\mu(dx, dy) = f(\rho) ds dt, \quad \rho = (x^2 + y^2)^{\frac{1}{2}},$$

so that

$$2\pi \int_0^{+\infty} \rho f(\rho) d\rho = 1.$$

Assume further that $J_k = \int_0^{+\infty} \rho^k f(\rho) d\rho < \infty$, for $1 \leq k \leq 5$.

For short, denote by $X, X_s, X_t, X_{ss}, X_{st}, X_{tt}$ the values of X and the first and second partial derivatives at the point (s, t) and $X' = (X_s, X_t)^T$ and X'' the matrix of second order partial

derivatives.

Let $S = \{(s, t) : s^2 + t^2 \leq 1\}$ be the closed disk of radius 1 centered at the origin and $M = \max_{(s,t) \in S} X(s, t)$.

(a) Prove that

- X' is independent of X and X'' , and has variance $\pi J_3 I_d$
- X''_{st} is independent of X , X' , X''_{ss} and X''_{tt} and has variance $\frac{\pi}{4} J_5$
- Conditionally on $X = u$, the random variables X''_{ss} and X''_{tt} have
 - expectation: $-\pi J_3$
 - variance: $\frac{3\pi}{4} J_5 - (\pi J_3)^2$
 - covariance: $\frac{\pi}{4} J_5 - (\pi J_3)^2$.

(b) Prove that

$$p_M(u) \leq (I_1 + I_2)\varphi(u),$$

where I_1, I_2 are computed by the formulas:

$$I_1 = \frac{1}{4\sqrt{2\pi}J_3} \int_0^\infty [(\alpha^2 + a^2 - c^2 x^2)\Phi(a - cx) + [2a\alpha - \alpha^2(a - cx)]\varphi(a - cx)] x\varphi(x) dx,$$

with

$$\begin{aligned} a &= 2\pi J_3 u \quad , \quad c = \sqrt{\frac{\pi J_5}{4}}, \\ I_2 &= \sqrt{\frac{2}{J_3}} \left[\left(\frac{3\pi}{4} J_5 - (\pi J_3)^2 \right)^{\frac{1}{2}} \varphi(bu) + \pi J_3 u \Phi(bu) \right], \\ b &= \frac{\pi J_3}{\left(\frac{3\pi}{4} J_5 - (\pi J_3)^2 \right)^{\frac{1}{2}}}. \end{aligned}$$

The tail of the distribution of the maximum

Let $\mathcal{X} = \{X(t) : t \in S\}$ be a real-valued random field defined on some parameter set S and $M := \sup_{t \in S} X(t)$ its supremum. In this chapter, we present recent results which allow, in certain cases, to give more precise approximations of the tails of the distribution of the random variable M . We will be especially interested in the approximation of $P(M > u)$ for large u , but we also give results that can be used for all u .

For Gaussian processes, a number of fundamental results have been presented in Chapter 2, that we have called the basic inequalities. These are essential for the development of most of the mathematical theory. However, in a wide number of applications, the general situation is that these inequalities are not good enough, one reason being that they depend on certain constants that one is unable to estimate or for which estimations differ substantially from the true values. Some refinements of the same subject, either for certain classes of processes or touching special topics concerning the computation of the distribution of the maximum, have also been considered in Chapters 3, 4, 5 and 9.

Since the 1990's several methods have been introduced with the aim of obtaining more precise results than those arising from the classical theory, at least under certain restrictions on the process \mathcal{X} . These results are interesting both from the standpoint of the mathematical theory and of their use in significant applications. The restrictions on \mathcal{X} include the requirement that the domain S have some finite-dimensional geometrical structure and the paths of the random field, a certain regularity.

Examples of these contributions are the double sum method by Piterbarg (1996); the Euler-Poincaré Characteristic (EPC) approximation, Taylor, Takemura and Adler (2005), Adler and Taylor's book(2007); the tube method , Sun (1993) and also the methods contained in the previous chapters just mentioned. We refer to these books and papers for an account of these results.

This chapter is divided into two parts. In the first part, we consider two special topics which concern only the tails of the distribution of one-parameter Gaussian processes:

In section 1 we will look at the asymptotic behavior of the successive derivatives of the distribution of the maximum and related questions, using the methods of Chapter 7.

In section 2 we again use similar tools to study the tails of the distribution of the maximum of certain unbounded Gaussian processes, that is, processes for which the probability q that the supremum is finite is smaller than 1, and we are willing to understand the speed at which $P(M \leq u)$ approaches the limiting value q as $u \rightarrow +\infty$. This section opens what seems to be an unexplored subject until now.

The second part, starting in Section 3, is the main body of the chapter. It is based upon Theorem 7.2 of Chapter 7, allowing to express the density p_M of F_M by means of a general formula. Even though this is an exact formula, it is only implicit as an expression for the density, since the relevant random variable M appears in the right-hand side. However, it can be usefully employed for our purposes.

First, one can use Theorem 7.2 to obtain bounds for $p_M(u)$ and thus for $P\{M > u\}$ for **every** u by means of replacing some indicator function in (7.1) by the condition that the normal derivative is “extended outward” (see below for the precise meaning). This will be called the

“direct method”. Of course, this may be interesting whenever the expression one obtains can be handled, which is the actual situation when the random field has a law which is stationary and isotropic. For this family of random fields, our method relies on the application of some known results on the spectrum of random matrices, which we will use without proof and one can find in Mehta’s book (2004).

Second, one can use Theorem 7.2 to study the asymptotics of $P\{M > u\}$ as $u \rightarrow +\infty$. More precisely, one wants to write, whenever it is possible

$$(8.1) \quad P\{M > u\} = A(u) \exp\left(-\frac{1}{2} \frac{u^2}{\sigma^2}\right) + B(u)$$

where $A(u)$ is a known function having polynomially bounded growth as $u \rightarrow +\infty$, $\sigma^2 = \sup_{t \in S} \text{Var}(X(t))$ and $B(u)$ is an error bounded by a centered Gaussian density with variance σ_1^2 , $\sigma_1^2 < \sigma^2$. We will call the first (respectively the second) term in the right-hand side of (8.1) the “first (resp second) order approximation of $P\{M > u\}$.”

First order approximation has been considered by Taylor, Takemura and Adler (2005) and also Adler and Taylor (2007) by means of the expectation of the EPC of the excursion set $E_u := \{t \in S : X(t) > u\}$. This works for large values of u . The same authors have considered the second order approximation, that is, how fast does the difference between $P\{M > u\}$ and the expected EPC tend to zero when $u \rightarrow +\infty$.

We will address the same question both for the direct method and the EPC approximation method. Our results on the second order approximation only speak about the size of the variance of the Gaussian bound. More precise results are only known in the special case where S is a compact interval of the real line, the Gaussian process \mathcal{X} is stationary and satisfies a certain number of additional requirements. We have stated this special result without proof in Chapter 4. It is due to Piterbarg (1981) when the domain interval is small enough and to Azais, Bardet and Wschebor (2002) in its general form. See the remarks after Proposition 4.1.

The first order approximation is computed for the direct method in Theorem 8.8 in the case of stationary isotropic random fields defined on a polyhedron, from which a new upper bound for $P\{M > u\}$ for all real u follows.

As for second order approximation, Theorem 8.10 is the first result here in this direction. It gives a rough bound for the error $B(u)$ as $u \rightarrow +\infty$, in case the maximum variance is attained at some strict subset of the face in S having the largest dimension. In Theorem 8.12 we consider random fields with constant variance. This is close to Theorem 4.3 in Taylor, Takemura and Adler (2005). In Theorem 8.15, S is convex, the random field is stationary and isotropic and we are able to compute the exact asymptotic rate for the second order approximation as $u \rightarrow +\infty$ corresponding to the direct method.

In all cases, the second order approximation for the direct method provides an upper bound for the one arising from the EPC method.

From technical point of view, the proofs of the results about the supremum of random fields contained in Section 3 to 5 require a minimum of elementary differential geometry which does not go beyond the definitions of embedded differentiable manifold, differentiation of functions defined on it (as in Chapter 6) and curvature. The reader can consult any introductory book on the subject.

1. One-dimensional parameter: asymptotic behavior of the derivatives of F_M

THEOREM 8.1. *Let \mathcal{X} be a stochastic process with parameter set $[0, 1]$ verifying the hypotheses H_{2k} of Section 2 in Chapter 7. We also assume that $E(X(t)) = 0$ and $\text{Var}(X(t)) = 1$.*

Then, as $u \rightarrow +\infty$,

$$(8.2) \quad F^{(k)}(u) \approx (-1)^{k-1} \frac{u^k}{2\pi} e^{-u^2/2} \int_0^1 \sqrt{r_{11}(t, t)} dt.$$

PROOF. We use the notations and results of Chapter 7.

To prove the result for $k = 1$ notice first that under the hypothesis of the theorem, one has $r(t, t) = 1$, $r_{01}(t, t) = 0$, $r_{02}(t, t) = -r_{11}(t, t)$.

An elementary computation of the regression (7.11) replacing Z by X , shows that:

$$b^t(s) = r(s, t), \quad c^t(s) = \frac{r_{01}(s, t)}{r_{11}(t, t)}$$

and

$$\beta^t(s) = 2 \frac{1 - r(s, t)}{(t - s)^2}$$

since we start with $\beta(t) = 1$ for every t .

This shows that for every $t \in [0, 1]$ one has $\inf_{s \in [0, 1]} (\beta^t(s)) > 0$ because of the non-degeneracy condition and $\beta^t(t) = -r_{02}(t, t) = r_{11}(t, t) > 0$. The expression for F' becomes:

$$(8.3) \quad F'(u) = \varphi(u)L(u),$$

where

$$L(u) = L_1(u) + L_2(u) + L_3(u),$$

$$L_1(u) = P(A_u(X^0, \beta^0)),$$

$$L_2(u) = P(A_u(X^1, \beta^1)),$$

$$L_3(u) = - \int_0^1 \mathbb{E} \left\{ (X_t^t - \beta^t(t)u) \mathbb{1}_{A_u(X^t, \beta^t)} \right\} \frac{dt}{(2\pi r_{11}(t, t))^{1/2}}.$$

Since for each $t \in [0, 1]$ the process X^t is bounded it follows that

$$a.s. \mathbb{1}_{A_u(X^t, \beta^t)} \rightarrow 1 \text{ as } u \rightarrow +\infty.$$

A dominated convergence argument shows now that $L_3(u)$ is equivalent to

$$- \frac{u}{(2\pi)^{1/2}} \int_0^1 \frac{r_{02}(t, t)}{(r_{11}(t, t))^{1/2}} dt = \frac{u}{(2\pi)^{1/2}} \int_0^1 \sqrt{r_{11}(t, t)} dt.$$

Since $L_1(u), L_2(u)$ are bounded by 1, (8.2) follows for $k = 1$.

For $k \geq 2$, write

$$(8.4) \quad F^{(k)}(u) = \varphi^{(k-1)}(u)L(u) + \sum_{h=2}^{h=k} \binom{k-1}{h-1} \varphi^{(k-h)}(u)L^{(h-1)}(u)$$

As $u \rightarrow +\infty$, for each $j = 0, 1, \dots, k-1$, $\varphi^{(j)}(u) \simeq (-1)^j u^j \varphi(u)$ so that the first term in (8.4) is equivalent to the expression in (8.2). Hence, to prove the theorem it suffices to show that the successive derivatives of the function L are bounded. In fact, we prove the stronger inequality

$$(8.5) \quad |L^{(j)}(u)| \leq l_j \varphi\left(\frac{u}{a_j}\right), \quad j = 1, \dots, k-1$$

for some positive constants $l_j, a_j, j = 1, \dots, k-1$.

We first consider the function L_1 . One has:

$$\beta^0(s) = \frac{1 - r(s, 0)}{s} \text{ for } 0 < s \leq 1, \beta^0(0) = 0,$$

$$(\beta^0)'(s) = \frac{-1 + r(s, 0) - s.r_{10}(s, 0)}{s^2} \text{ for } 0 < s \leq 1, (\beta^0)'(0) = \frac{1}{2}r_{11}(0, 0).$$

The derivative $L'_1(u)$ becomes

$$L'_1(u) = \beta^0(1)E \{ \mathbf{1}_{A_u(X^{0,1}, \beta^{0,1})} \} p_{X(1)}^0(\beta^0(1)u) - \int_0^1 \beta^0(t)E \{ (X(t)^{0,t} - \beta^{0,t}(t)u) \mathbf{1}_{A_u(X^{0,t}, \beta^{0,t})} \} p_{X^0(t), (X^0)'(t)}(\beta^0(t)u, (\beta^0)'(t)u) dt.$$

Notice that $\beta^0(1)$ is non-zero so that the first term is bounded by a constant times a non-degenerate Gaussian density. Even though $\beta^0(0) = 0$, the second term is also bounded by a constant times a non-degenerate Gaussian density because the joint distribution of the pair $(X^0(t), (X^0)'(t))$ is non-degenerate and the pair $(\beta^0(t), (\beta^0)'(t)) \neq (0, 0)$ for every $t \in [0, 1]$.

Applying a similar argument to the successive derivatives we obtain (8.5) with L_1 instead of L .

The same follows with no changes for

$$L_2(u) = P(A_u(X^1, \beta^1)).$$

For the third term

$$L_3(u) = - \int_0^1 E \{ (X^t(t) - \beta^t(t)u) \mathbf{1}_{A_u(X^t, \beta^t)} \} \frac{dt}{(2\pi r_{11}(t, t))^{1/2}}$$

we proceed similarly, taking into account $\beta^t(s) \neq 0$ for every $s \in [0, 1]$.

So (8.5) follows and we are done. \square

1.1. A refinement. It is possible to refine the result obtaining the exact first order approximation and a bound for the second order approximation, not only for the tail of the distribution, but also for its derivatives. That is, the following refinement allows us to write the successive derivatives of F_M in an analogous form to (8.1) in which the second term in the right-hand side is bounded by a Gaussian density having a variance which is smaller than the supremum of the variance of the process.

Let $n = 0, 1, 2, \dots$. We will use repeatedly the Hermite polynomials, defined as

$$(8.6) \quad H_n(x) := e^{x^2} \left(-\frac{\partial}{\partial x} \right)^n e^{-x^2}$$

and the modified Hermite polynomials:

$$(8.7) \quad \bar{H}_n(x) := e^{x^2/2} \left(-\frac{\partial}{\partial x} \right)^n e^{-x^2/2}.$$

For the properties of the Hermite polynomials we refer to Mehta's book (2004).

THEOREM 8.2. *Suppose that \mathcal{X} satisfies the hypotheses of the theorem with $k \geq 2$. For $j = 1, \dots, k$, one has:*

$$(8.8) \quad F^{(j)}(u) = (-1)^{j-1} \bar{H}_{j-1}(u) \left[1 + (2\pi)^{-1/2} u \int_0^1 (r_{11}(t, t))^{1/2} dt \right] \varphi(u) + \rho_j(u) \varphi(u)$$

where

$$|\rho_j(u)| \leq C_j \exp(-\delta u^2)$$

with C_1, C_2, \dots positive constants and $\delta > 0$ does not depend on j .

PROOF. The proof of (8.8) is a slight modification of the one of the previous theorem.

Notice first that from the above computation of $\beta^0(s)$ it follows that 1) if $X^0(0) < 0$, then if u is large enough $X^0(s) - \beta^0(s) \cdot u \leq 0$ for all $s \in [0, 1]$ and 2) if $X^0(0) > 0$, then $X^0(0) - \beta^0(0) \cdot u > 0$ so that:

$$L_1(u) = P(X^0(s) - \beta^0(s) \cdot u \leq 0 \text{ for all } s \in [0, 1]) \uparrow \frac{1}{2} \text{ as } u \uparrow +\infty.$$

On account of (8.5) this implies that if $u \geq 0$:

$$0 \leq \frac{1}{2} - L_1(u) = \int_u^{+\infty} L'_1(v) dv \leq D_1 \exp(-\delta_1 u^2)$$

with D_1, δ_1 positive constants.

$L_2(u)$ is similar. Finally:

$$(8.9) \quad L_3(u) = - \int_0^1 \mathbb{E} \{ X^t(t) - \beta^t(t)u \} \frac{dt}{(2\pi r_{11}(t, t))^{1/2}} - \\ - \int_0^1 \mathbb{E} \left\{ (X^t(t) - \beta^t(t)u) \mathbb{I}_{(A_u(X^t, \beta^t))^c} \right\} \frac{dt}{(2\pi r_{11}(t, t))^{1/2}}$$

The first term in (8.9) is equal to:

$$(2\pi)^{-1/2} \cdot u \cdot \int_0^1 (r_{11}(t, t))^{1/2} dt.$$

As for the second term in (8.9) denote $\beta_{\#} = \inf_{s, t \in [0, 1]} \beta^t(s) > 0$ and let $u > 0$. Then:

$$\mathbb{P}((A_u(X^t, \beta^t))^c) \leq \mathbb{P}(\exists s \in [0, 1] \text{ such that } X^t(s) > \beta_{\#} \cdot u) \leq D_3 \exp(-\delta_3 u^2)$$

with D_3, δ_3 are positive constants, the last inequality being a consequence of the basic inequality 2.33.

The remainder follows in the same way as the proof of Theorem 8.1. \square

2. An Application to Unbounded Processes.

As we have seen in Chapter 2, the tails of the probability distribution of the supremum of a Gaussian processes which a.s. has bounded paths have a nice behavior for large values of the argument, since they are bounded by the tails of the distribution of one Gaussian variable. As we have seen, this has important consequences to handle the supremum of a Gaussian processes, via the exponential bound for the tails of its probability distribution.

In this section we consider a more unexplored subject, that is, the behavior of large values of the supremum of Gaussian processes in case they are unbounded. One should observe that if a separable centered Gaussian process $\{Z(t) : t \in T\}$ is unbounded then $\mathbb{P}(\{\sup |Z(t)| < \infty, t \in T\}) = 0$, but this does not imply that $q = \mathbb{P}(\{\sup Z(t) < \infty, t \in T\}) = 0$. [Notice that $\{\sup Z(t) < +\infty, t \in T\}$ is **not** a subspace of the space of possible paths, so that the Gaussian 0 or 1 law can not be applied to it].

We will see that the methods of Chapter 7, which a priori had the purpose of studying the regularity of the distribution of M_T , are useful to understand the asymptotic behavior of an interesting family of one-parameter unbounded processes. In fact the tail of the probability distribution of M_T , having total mass q strictly smaller than 1, that is:

$$q - \mathbb{P}(M_T \leq u),$$

which tends to zero as $u \rightarrow +\infty$, is not necessarily bounded by a Gaussian tail and can be estimated using the differentiation theorems above.

Let $\mathcal{X} = \{X(t) : t \in [0, 1]\}$ satisfy the hypothesis H_2 of Chapter 7. We also assume that the process is centered.

Let $\beta : [0, 1] \rightarrow \mathbb{R}^+$ be a continuous function that vanishes only at $t = 0$, twice continuously differentiable for $t \in (0, 1]$. We are going to study the behavior as $u \rightarrow +\infty$ of the function:

$$F(u) = \mathbb{P}(X(t) \leq \beta(t) \cdot u \text{ for all } t \in [0, 1])$$

which is obviously the (defective) distribution function of the supremum over $(0, 1]$ of the unbounded Gaussian process:

$$Z_t = \frac{X(t)}{\beta(t)}$$

that has exploding paths at the only point $t = 0$.

Clearly,

$$a.s. \quad \lim_{u \rightarrow +\infty} \mathbb{I}_{\{X(t) \leq \beta(t) \cdot u \text{ for all } t \in [0, 1]\}} = \mathbb{I}_{\{X(0) < 0\}}$$

so that:

$$q = F(+\infty) = \frac{1}{2}.$$

THEOREM 8.3. *Assume that if $t \downarrow 0$ one has $\beta(t) \approx C t^\alpha$, $\beta'(t) \approx C \alpha t^{\alpha-1}$, $\beta''(t) \approx C \alpha(\alpha-1) t^{\alpha-2}$, where C, α are positive constants.*

Then, we have

(i) *If $0 < \alpha \leq 1$,*

$$\frac{1}{2} - F(u) \leq a \cdot \exp(-b \cdot u^2) \quad \text{for } u > 0$$

where a, b are positive constants.

(ii) *If $\alpha > 1$,*

$$(8.10) \quad \frac{1}{2} - F(u) \approx \frac{K}{u^{1/(\alpha-1)}} \quad \text{when } u \rightarrow +\infty$$

where K is a positive constant depending on $C, \alpha, \text{Var}\{X(0)\}, \text{Var}\{X'(0)\}$ that we will obtain explicitly in the proof.

PROOF. With no loss of generality, we may assume that $\text{Var}\{X(t)\} = 1$ for all $t \in [0, 1]$. If this is not the case, it suffices to replace the process by $X(t)/[\text{Var}\{X(t)\}]^{\frac{1}{2}}$ and the function $\beta(t)$ by $\beta(t)/[\text{Var}\{X(t)\}]^{\frac{1}{2}}$.

To study the behavior of

$$\frac{1}{2} - F(u) = \int_u^{+\infty} F'(x) dx$$

when $u \rightarrow +\infty$, we write $F'(x)$ using Lemma 7.7 that is, on account of $\beta(0) = 0$:

(8.11)

$$F'(x) = \beta(1)P(X^1(t) \leq \beta^1(t) \cdot x \ \forall t \in [0, 1]) p_{X(1)}(\beta(1) \cdot x)$$

$$(8.12) \quad - \int_0^1 \beta(t) E \left\{ [X^t(t) - \beta^t(t)x] \mathbf{1}_{\{X^t(s) \leq \beta^t(s)x \ \forall s \in [0, 1]\}} \right\} p_{X(t), X'(t)}(\beta(t)x, \beta'(t)x) dt.$$

A simple computation shows that in the present case the process X^1 and the function β^1 are the continuous extensions to $[0, 1]$ of

$$X^1(t) = \frac{X(t) - r(t, 1)X(1)}{1 - t}, \quad \beta^1(t) = \frac{\beta(t) - r(t, 1)\beta(1)}{1 - t}$$

defined for $t \in [0, 1)$.

In the same way, for $t \in [0, 1]$, X^t, β^t are the continuous extensions to $[0, 1]$ of

$$X^t(s) = \frac{2}{(s-t)^2} \left[X(s) - r(s, t)X(t) - \frac{r_{10}(s, t)}{r_{11}(t, t)} X'(t) \right]$$

$$\beta^t(s) = \frac{2}{(s-t)^2} \left[\beta(s) - r(s, t)\beta(t) - \frac{r_{10}(s, t)}{r_{11}(t, t)} \beta'(t) \right]$$

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

In fact the proof of the formula (8.11) requires the function $\beta(\cdot)$ to be of class \mathcal{C}^2 on the whole interval $[0, 1]$, which need not to be our case. The proof that (8.11) holds true for $x > 0$ is left to the reader, and is a consequence of the hypotheses on $\beta(\cdot)$ near $t = 0$.

It is easily seen that the first term in the right-hand side of (8.11) is bounded above by $a_1 \cdot \exp(-b_1 \cdot x^2)$, where a_1 and b_1 are positive constants.

As for the second term, the density in the integrand is:

$$(8.13) \quad p_{X(t), X'(t)}(\beta(t)x, \beta'(t)x) = \frac{1}{2\pi [\text{Var}\{X'(t)\}]^{\frac{1}{2}}} \exp \left[-\frac{x^2}{2} \left(\beta^2(t) + \frac{(\beta'(t))^2}{\text{Var}\{X'(t)\}} \right) \right]$$

and

$$(8.14) \quad \beta^t(s) = \frac{2}{(s-t)^2} [\beta(s) - \beta(t) - (s-t)\beta'(t)] - r_{20}(t + \theta_1(s-t), t)\beta(t) - \frac{r_{12}(t, t + \theta_2(s-t))}{r_{11}(t, t)} \beta'(t) \quad (0 < \theta_1, \theta_2 < 1)$$

so that

$$\beta^t(t) = \beta^{\alpha}(t) - r_{20}(t, t)\beta(t) - \frac{r_{12}(t, t)}{r_{11}(t, t)}\beta'(t)$$

si $0 < t \leq 1$.

Assume now hypothesis (i). Using the conditions on the function $\beta(\cdot)$, one can verify that if $x > 1$, the absolute value of the integrand in the right-hand side of (8.11) is bounded by $a_2 \cdot \exp(-b_2 \cdot x^2)$, a_2 and b_2 positive constants. This proves the first part of the theorem.

If condition (ii) holds, take $\varepsilon, 0 < \varepsilon < 1$ and use the splitting:

$$\frac{1}{2} - F(u) = \frac{1}{2} - F_\varepsilon(u) + G_\varepsilon(u)$$

where

$$F_\varepsilon(u) = P(X(t) \leq \beta(t) \cdot u \text{ for all } t \in [0, \varepsilon])$$

and

$$0 \leq G_\varepsilon(u) \leq P(\sup_{t \in [\varepsilon, 1]} \frac{X(t)}{\beta(t)} > u) \leq g_1 \exp(-g_2 \cdot u^2),$$

with g_1, g_2 positive constants depending on ε , the last inequality resulting from an application of (2.33) in Chapter 2 to the bounded process $\frac{X(t)}{\beta(t)}$ for $t \in [\varepsilon, 1]$.

Hence, to finish the proof of (ii) it suffices to show (ii) when one replaces $F(u)$ by $F_\varepsilon(u)$, where $\varepsilon > 0$ is fixed and small enough.

With that purpose, we apply the same formula (8.11) mutatis mutandis, that is changing the interval $[0, 1]$ by $[0, \varepsilon]$. We obtain:

$$(8.15) \quad F'_\varepsilon(x) = \beta(\varepsilon)P(X^1(t) \leq \beta^1(t) \cdot x \text{ for all } t \in [0, \varepsilon]) p_{X(\varepsilon)}(\beta(\varepsilon) \cdot x) \\ - \int_0^\varepsilon \beta(t) E \{ [X^t(t) - \beta^t(t)x] \mathbf{1}_{\{X^t(s) \leq \beta^t(s)x \text{ for all } s \in [0, \varepsilon]\}} \} \\ \cdot p_{X(t), X'(t)}(\beta(t)x, \beta'(t)x) dt.$$

In the integral in the right-hand side of (8.15) make the change of variable $z = t^{\alpha-1}x$. Then, $t(x) = \left(\frac{z}{x}\right)^{\frac{1}{\alpha-1}} \rightarrow 0$ when $x \rightarrow +\infty$ for each $z > 0$ fixed.

For an adequate choice of ε , we have for each $z > 0$:

$$a.s. \quad \mathbf{1}_{\{X^{t(x)}(s) \leq \beta^{t(x)}(s)x \text{ for all } s \in [0, \varepsilon]\}} \rightarrow 1 \text{ when } x \rightarrow +\infty.$$

This follows from the statement:

$$(8.16) \quad \text{for } z > 0 \text{ fixed, } \beta^{t(x)}(s)x \rightarrow +\infty \text{ when } x \rightarrow +\infty \text{ uniformly in } s \in [0, \varepsilon].$$

To prove (8.16) use (8.14) and the form of $\beta(t)$ near $t = 0$. We have (elementary checking):

-if $1 < \alpha < 2$ and L is any positive real number, one can choose $\varepsilon > 0$ small enough so that:

$$(8.17) \quad \beta^t(s) \cdot x \geq L \cdot x - C_1 \cdot t \cdot z - C_2 \cdot z \quad \text{for all } s, t \in [0, \varepsilon],$$

and C_1, C_2 are some positive constants.

- if $\alpha \geq 2$ one can choose small enough $\varepsilon > 0$ in such a way that:

$$(8.18) \quad \beta^t(s) \cdot x \geq C'_1 \cdot \frac{z}{t} - C'_2 \cdot t \cdot z - C'_3 \cdot z \quad \text{for all } s, t \in]0, \varepsilon],$$

and C'_1, C'_2, C'_3 are some positive constants. Either (8.17) or (8.18) imply (8.16).

To find the equivalent of the right-hand side of (8.15) is now an exercise (apply dominated convergence). We get:

$$F'_\varepsilon(x) \approx \left(\frac{\sigma}{\alpha}\right)^{\alpha/(\alpha-1)} \frac{1}{\sqrt{2\pi}} C^{1/(1-\alpha)} I_{\alpha/(\alpha-1)} x^{\alpha/(1-\alpha)} \text{ when } x \rightarrow +\infty,$$

where $\sigma = [\text{Var} \{X'(0)\}]^{\frac{1}{2}}$ and I_a is defined, for $a \geq 0$ by means of:

$$I_a = \int_0^{+\infty} y^a \frac{\exp(-y^2/2)}{\sqrt{2\pi}} dy.$$

Finally, (8.10) follows integrating once. One also obtains the value of the constant

$$K = (\alpha - 1) \left(\frac{\sigma}{\alpha} \right)^{\alpha/(\alpha-1)} \frac{1}{\sqrt{2\pi}} C^{1/(1-\alpha)} I_{\alpha/(\alpha-1)}$$

□

3. A general bound for p_M

We need to introduce some further notations.

For t in S_j , $j \leq d_0$ we define $\mathcal{C}_{t,j}$ as the closed convex cone generated by the set of directions:

$$\{\lambda \in \mathbb{R}^d : \|\lambda\| = 1 ; \exists s_n \in S, (n = 1, 2, \dots) \text{ such that } s_n \rightarrow t, \frac{t - s_n}{\|t - s_n\|} \rightarrow \lambda \text{ as } n \rightarrow +\infty\},$$

whenever this set is non-empty and $\mathcal{C}_{t,j} = \{0\}$ if it is empty. We will denote by $\widehat{\mathcal{C}}_{t,j}$ the dual cone of $\mathcal{C}_{t,j}$, that is:

$$\widehat{\mathcal{C}}_{t,j} := \{z \in \mathbb{R}^d : \langle z, \lambda \rangle \geq 0 \text{ for all } \lambda \in \mathcal{C}\}.$$

Notice that these definitions easily imply that $T_{t,j} \subset \mathcal{C}_{t,j}$ and $\widehat{\mathcal{C}}_{t,j} \subset N_{t,j}$. Remark also that for $j = d_0$, $\widehat{\mathcal{C}}_{t,j} = N_{t,j}$.

We will say that the function $X(\cdot)$ has an “extended outward” derivative at the point t in S_j , $j \leq d_0$ if $X'_{j,N}(t) \in \widehat{\mathcal{C}}_{t,j}$.

THEOREM 8.4. *Under assumptions A1 to A5 of Chapter 7, Section 1, one has :*

(a) $p_M(x) \leq \bar{p}(x)$ where

$$(8.19) \quad \bar{p}(x) := \sum_{t \in S_0} \mathbb{E}(\mathbf{1}_{X'(t) \in \widehat{\mathcal{C}}_{t,0}} | X(t) = x) p_{X(t)}(x) + \sum_{j=1}^{d_0} \int_{S_j} \mathbb{E}(|\det(X''_j(t))| \mathbf{1}_{X'_{j,N}(t) \in \widehat{\mathcal{C}}_{t,j}} | X(t) = x, X'_j(t) = 0) p_{X(t), X'_j(t)}(x, 0) \sigma_j(dt).$$

(b) $\mathbb{P}\{M > u\} \leq \int_u^{+\infty} \bar{p}(x) dx.$

PROOF. (a) follows from Theorem 7.2 and the observation that if $t \in S_j$, one has $\{M \leq X(t)\} \subset \{X'_{j,N}(t) \in \widehat{\mathcal{C}}_{t,j}\}$. (b) is an obvious consequence of (a). □

The actual interest of this Theorem depends on the feasibility of computing $\bar{p}(x)$. It turns out that this can be done in some relevant cases, as we will see in the next section. The results can be compared with the approximation of $\mathbb{P}\{M > u\}$ by means of $\int_u^{+\infty} p^E(x) dx$ given by Adler and Taylor (2007) and Taylor, Takemura and Adler (2005), where

$$(8.20) \quad p^E(x) := \sum_{t \in S_0} \mathbb{E}(\mathbf{1}_{X'(t) \in \widehat{\mathcal{C}}_{t,0}} | X(t) = x) p_{X(t)}(x) + \sum_{j=1}^{d_0} (-1)^j \int_{S_j} \mathbb{E}(\det(X''_j(t)) \mathbf{1}_{X'_{j,N}(t) \in \widehat{\mathcal{C}}_{t,j}} | X(t) = x, X'_j(t) = 0) p_{X(t), X'_j(t)}(x, 0) \sigma_j(dt).$$

Under certain conditions, $\int_u^{+\infty} p^E(x) dx$ is the expected value of the EPC of the excursion set E_u . The advantage of $p^E(x)$ over $\bar{p}(x)$ is that one can have nice expressions for it (see Exercise 8.1). Conversely $\bar{p}(x)$ has the obvious advantage that it is an upper-bound of the true density $p_M(x)$. Hence, upon integrating once, it provides an upper-bound for the tail probability, **for every u value.**

We can go a bit farther, and show that $[\bar{p}(x) + p^E(x)]/2$ is also an upper bound for $p^M(x)$. This follows easily upon the following observation: Denote by E_j the event that the random linear operator $X''_j(t)$ is non-negative definite. In the formula giving $p^M(x)$ (see Theorem 7.2 in Chapter 7), one can replace inside the conditional expectation $|\det(X''_j(t))| \mathbf{1}_{A_x}$ by $(-1)^j \det(X''_j(t)) \mathbf{1}_{A_x} \mathbf{1}_{E_j}$,

since under the conditioning E_j occurs, and obviously, by $\frac{|\det(X_j''(t))| + (-1)^j \det(X_j''(t))}{2} \mathbb{1}_{A_x} \mathbb{1}_{E_j}$. Replacing $\mathbb{1}_{A_x} \mathbb{1}_{E_j}$ by $\mathbb{1}_{X'_{j,N}(t) \in \hat{c}_{t,j}}$ we get for $p^M(x)$ the more precise upper bound $[\bar{p}(x) + p^E(x)]/2$.

Under additional conditions, these upper bounds both provide good first order approximations for $p_M(x)$ as $x \rightarrow \infty$ as we will see in the remaining of this chapter. In the special case in which the random field \mathcal{X} is centered and has a law that is invariant under orthogonal linear transformations and translations, we give in the next section a procedure to compute $\bar{p}(x)$.

4. Computing $\bar{p}(x)$ for stationary isotropic Gaussian fields

For one-parameter centered Gaussian process having constant variance and satisfying certain regularity conditions, a bound for $p_M(x)$ has been given by inequality (7.6) of Chapter 7. In the two-dimensional parameter case, Mercadier (2005) has given a bound for $P\{M > u\}$, using a method especially suited to dimension 2 that will be presented in Chapter 9. When the parameter is one or two-dimensional, these bounds are sharper than the ones below which apply to any dimension.

We will assume that the process \mathcal{X} is centered Gaussian, with a covariance function that can be written as

$$(8.21) \quad \mathbb{E}(X(s)X(t)) = \rho(\|s - t\|^2),$$

where $\rho : \mathbb{R}^+ \rightarrow \mathbb{R}$ is of class \mathcal{C}^4 . Without loss of generality, we assume that $\rho(0) = 1$. Assumption (8.21) is equivalent to saying that the law of \mathcal{X} is invariant under orthogonal linear transformations and translations of the underlying parameter space \mathbb{R}^d .

We will also assume that the set S has a polyhedral shape. More precisely we assume that each $S_j (j = 1, \dots, d)$ is a union of subsets of affine manifolds of dimension j in \mathbb{R}^d .

For the proof of Theorem 8.8 below, which gives an expression for the bound $\bar{p}(x)$ of the density, we need some auxiliary computational lemmas.

The first lemma is elementary and the proof is left to the reader. Here, and in the remainder of this chapter, we use the abridged notation: $\rho' := \rho'(0)$, $\rho'' := \rho''(0)$

LEMMA 8.5. *Under the conditions above, for each $t \in U$, $i, i', k, k', j = 1, \dots, d$:*

- (1) $\mathbb{E}\left(\frac{\partial X}{\partial t_i}(t) \cdot X(t)\right) = 0$,
- (2) $\mathbb{E}\left(\frac{\partial X}{\partial t_i}(t) \cdot \frac{\partial X}{\partial t_k}(t)\right) = -2\rho' \delta_{ik}$ and $\rho' < 0$,
- (3) $\mathbb{E}\left(\frac{\partial^2 X}{\partial t_i \partial t_k}(t) \cdot X(t)\right) = 2\rho' \delta_{ik}$, $\mathbb{E}\left(\frac{\partial^2 X}{\partial t_i \partial t_k}(t) \cdot \frac{\partial X}{\partial t_j}(t)\right) = 0$
- (4) $\mathbb{E}\left(\frac{\partial^2 X}{\partial t_i \partial t_k}(t) \cdot \frac{\partial^2 X}{\partial t_{i'} \partial t_{k'}}(t)\right) = 24\rho'' [\delta_{ii'} \cdot \delta_{kk'} + \delta_{i'k} \cdot \delta_{ik'} + \delta_{ik} \delta_{i'k'}]$,
- (5) $\rho'' - \rho'^2 \geq 0$
- (6) *If $t \in S_j$, the conditional distribution of $X_j''(t)$ given $X(t) = x, X_j'(t) = 0$ is the same as the unconditional distribution of the random matrix*

$$Z + 2\rho' x I_j,$$

where $Z = (Z_{ik} : i, k = 1, \dots, j)$ is a symmetric $j \times j$ matrix with centered Gaussian entries, independent of the pair $(X(t), X'(t))$ such that, for $i \leq k, i' \leq k'$ one has :

$$\mathbb{E}(Z_{ik} Z_{i'k'}) = 4[2\rho'' \delta_{ii'} + (\rho'' - \rho'^2)] \delta_{ik} \delta_{i'k'} + 4\rho'' \delta_{ii'} \cdot \delta_{kk'} (1 - \delta_{ik}) .$$

Our second lemma is the following:

LEMMA 8.6. *Let*

$$(8.22) \quad J_n(x) := \int_{-\infty}^{+\infty} e^{-y^2/2} H_n(z) dy, \quad n = 0, 1, 2, \dots$$

where z stands for the linear form $z = ay + bx$ and a, b are real parameters that satisfy $a^2 + b^2 = 1/2$. Then

$$J_n(x) = (2b)^n \sqrt{2\pi} \bar{H}_n(x).$$

PROOF. From the definitions of H_n and \bar{H}_n we get

$$\sum_{n=0}^{\infty} \frac{(w)^n}{n!} H_n(z) = e^{-w^2+2wz} \quad \sum_{n=0}^{\infty} \frac{(w)^n}{n!} \bar{H}_n(z) = e^{-\frac{w^2}{2}+wz},$$

using the Taylor expansion of $e^{(z-w)^2}$ and $e^{\frac{(z-w)^2}{2}}$ in w around 0. Therefore

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{(w)^n}{n!} J_n(x) &= \int_{\mathbb{R}} e^{-\frac{y^2}{2}-w^2+2w(ay+bx)} dy \\ &= e^{2wbx-2(bw)^2} \int_{\mathbb{R}} e^{\frac{(y-2wa)^2}{2}} dy \\ &= \sqrt{2\pi} \sum_{n=0}^{\infty} \frac{(2bw)^n}{n!} \bar{H}_n(x). \end{aligned}$$

This implies $J_n(x) = (2b)^n \sqrt{2\pi} \bar{H}_n(x)$. \square

We also need the integrals

$$I_n(v) = \int_v^{+\infty} e^{-t^2/2} H_n(t) dt.$$

They are computed in the next lemma, which can be proved easily, using standard properties of Hermite polynomials. This is also left to the reader.

LEMMA 8.7. (a)

$$(8.23) \quad \begin{aligned} I_n(v) &= 2e^{-v^2/2} \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} 2^k \frac{(n-1)!!}{(n-1-2k)!!} H_{n-1-2k}(v) \\ &\quad + \mathbf{1}_{\{n \text{ even}\}} 2^{\frac{n}{2}} (n-1)!! \sqrt{2\pi} \bar{\Phi}(x) \end{aligned}$$

(b)

$$(8.24) \quad I_n(-\infty) = \mathbf{1}_{\{n \text{ even}\}} 2^{\frac{n}{2}} (n-1)!! \sqrt{2\pi}$$

We are now ready to state and prove the announced expression for $\bar{p}(x)$:

THEOREM 8.8. *Assume that the random field \mathcal{X} is centered Gaussian, satisfies conditions A1-A5 of Chapter 7, Section 1, and has a covariance having the form (8.21) which verifies the regularity conditions of the beginning of this section. Moreover, let S have polyhedral shape. Then, $\bar{p}(x)$ can be expressed by means of the formula:*

$$(8.25) \quad \bar{p}(x) = \varphi(x) \left\{ \sum_{t \in S_0} \hat{\sigma}_0(t) + \sum_{j=1}^{d_0} \left[\left(\frac{|\rho'|}{\pi} \right)^{j/2} \bar{H}_j(x) + R_j(x) \right] g_j \right\},$$

where

- g_j is a geometric parameter of the face S_j defined by

$$(8.26) \quad g_j = \int_{S_j} \hat{\sigma}_j(t) \sigma_j(dt),$$

where $\hat{\sigma}_j(t)$ is the normalized solid angle of the cone $\hat{\mathcal{C}}_{t,j}$ in $N_{t,j}$, that is:

$$(8.27) \quad \hat{\sigma}_j(t) = \frac{\sigma_{d-j-1}(\hat{\mathcal{C}}_{t,j} \cap \mathcal{S}^{d-j-1})}{\sigma_{d-j-1}(\mathcal{S}^{d-j-1})} \text{ for } j = 0, \dots, d-1,$$

$$(8.28) \quad \hat{\sigma}_d(t) = 1.$$

Notice that for convex or other usual polyhedra $\hat{\sigma}_j(t)$ is constant for $t \in S_j$, so that g_j is equal to this constant multiplied by the j -dimensional geometric measure of S_j .

- For $j = 1, \dots, d$,

$$(8.29) \quad R_j(x) = \left(\frac{2\rho''}{\pi|\rho'|}\right)^{\frac{j}{2}} \frac{\Gamma((j+1)/2)}{\pi} \int_{-\infty}^{+\infty} T_j(v) \exp\left(-\frac{y^2}{2}\right) dy$$

where

$$(8.30) \quad v := -(2)^{-1/2}((1-\gamma^2)^{1/2}y - \gamma x) \quad \text{with } \gamma := |\rho'|(\rho'')^{-1/2},$$

$$(8.31) \quad T_j(v) := \left[\sum_{k=0}^{j-1} \frac{H_k^2(v)}{2^k k!} \right] e^{-v^2/2} - \frac{H_j(v)}{2^j (j-1)!} I_{j-1}(v),$$

where I_n is given in the previous lemma.

For the proof of the theorem, we still need some additional ingredients from random matrices theory.

The random $n \times n$ real random matrix G_n is said to have the GOE distribution (for ‘‘Gaussian Orthogonal Ensemble’’), if it is symmetric, has centered Gaussian entries $g_{ik}, i, k = 1, \dots, n$ satisfying $E(g_{ii}^2) = 1$, $E(g_{ik}^2) = 1/2$ if $i < k$ and the random variables: $\{g_{ik}, 1 \leq i \leq k \leq n\}$ are independent.

Following Mehta (2004), we denote $q_n(\nu)$ the density of eigenvalues of $n \times n$ GOE matrices at the point ν , that is, $q_n(\nu)d\nu$ is the probability of G_n having an eigenvalue in the interval $(\nu, \nu+d\nu)$. One has the formula:

$$(8.32) \quad \begin{aligned} e^{\nu^2/2} q_n(\nu) &= e^{-\nu^2/2} \sum_{k=0}^{n-1} c_k^2 H_k^2(\nu) \\ &+ 1/2 (n/2)^{1/2} c_{n-1} c_n H_{n-1}(\nu) \left[\int_{-\infty}^{+\infty} e^{-y^2/2} H_n(y) dy - 2 \int_{\nu}^{+\infty} e^{-y^2/2} H_n(y) dy \right] \\ &+ \mathbf{1}_{\{n \text{ odd}\}} \frac{H_{n-1}(\nu)}{\int_{-\infty}^{+\infty} e^{-y^2/2} H_{n-1}(y) dy}, \end{aligned}$$

where $c_k := (2^k k! \sqrt{\pi})^{-1/2}, k = 0, 1, \dots$. The proof can be found in Mehta’s book, ch. 7.

We will use the following remark due to Fyodorov (2006), that we state as a lemma.

LEMMA 8.9. *Let G_n be a GOE $n \times n$ matrix. Then, for $\nu \in \mathbb{R}$ one has:*

$$(8.33) \quad E(|\det(G_n - \nu I_n)|) = 2^{3/2} \Gamma((n+3)/2) \exp(\nu^2/2) \frac{q_{n+1}(\nu)}{n+1},$$

PROOF. Denote by ν_1, \dots, ν_n the eigenvalues of G_n . It is well-known (Mehta (2004), Kendall et al. (1983)) that the joint density f_n of the n -tuple of random variables (ν_1, \dots, ν_n) is given by the formula

$$f_n(\nu_1, \dots, \nu_n) = k_n \exp\left(-\frac{\sum_{i=1}^n \nu_i^2}{2}\right) \prod_{1 \leq i < k \leq n} |\nu_k - \nu_i|, \quad \text{with } k_n := (2\pi)^{-n/2} (\Gamma(3/2))^n \left(\prod_{i=1}^n \Gamma(1+i/2)\right)^{-1}$$

Then,

$$\begin{aligned} \mathbb{E}(|\det(G_n - \nu I_n)|) &= \mathbb{E}\left(\prod_{i=1}^n |\nu_i - \nu|\right) \\ &= \int_{\mathbb{R}^n} \prod_{i=1}^n |\nu_i - \nu| k_n \exp\left(-\frac{\sum_{i=1}^n \nu_i^2}{2}\right) \prod_{1 \leq i < k \leq n} |\nu_k - \nu_i| d\nu_1, \dots, d\nu_n \\ &= e^{\nu^2/2} \frac{k_n}{k_{n+1}} \int_{\mathbb{R}^n} f_{n+1}(\nu_1, \dots, \nu_n, \nu) d\nu_1, \dots, d\nu_n = e^{\nu^2/2} \frac{k_n}{k_{n+1}} \frac{q_{n+1}(\nu)}{n+1}. \end{aligned}$$

The remainder is plain. \square

PROOF OF THEOREM 8.8:

We use the definition (8.19) given in Theorem 8.4 and the moment computations of Lemma 8.5 which imply that:

$$(8.34) \quad p_{X(t)}(x) = \varphi(x)$$

$$(8.35) \quad p_{X(t), X'_j(t)}(x, 0) = \varphi(x) (2\pi)^{-j/2} (-2\rho')^{-j/2}$$

$$(8.36) \quad X'(t) \text{ is independent of } X(t)$$

$$(8.37) \quad X'_{j,N}(t) \text{ is independent of } (X''_j(t), X(t), X'_j(t)).$$

Since the distribution of $X'(t)$ is centered Gaussian with variance $-2\rho' I_d$, it follows that :

$$\mathbb{E}(\mathbb{1}_{X'(t) \in \widehat{C}_{t,0}} | X(t) = x) = \widehat{\sigma}_0(t) \quad \text{if } t \in S_0,$$

and if $t \in S_j, j \geq 1$:

$$\begin{aligned} (8.38) \quad \mathbb{E}(|\det(X''_j(t))| \mathbb{1}_{X'_{j,N}(t) \in \widehat{C}_{t,j}} | X(t) = x, X'_j(t) = 0) \\ = \widehat{\sigma}_j(t) \mathbb{E}(|\det(X''_j(t))| | X(t) = x, X'_j(t) = 0) \\ = \widehat{\sigma}_j(t) \mathbb{E}(|\det(Z + 2\rho' x I_j)|). \end{aligned}$$

In the formula above, $\widehat{\sigma}_j(t)$ is the normalized solid angle defined in the statement of the theorem and the random $j \times j$ real matrix Z has the distribution of Lemma 8.5 .

A standard moment computations shows that Z has the same distribution as the random matrix:

$$\sqrt{8\rho''} G_j + 2\sqrt{\rho'' - \rho'^2} \xi I_j,$$

where G_j is a $j \times j$ GOE random matrix, ξ is standard normal in \mathbb{R} and independent of G_j . So, for $j \geq 1$ one has

$$\mathbb{E}(|\det(Z + 2\rho' x I_j)|) = (8\rho'')^{j/2} \int_{-\infty}^{+\infty} \mathbb{E}(|\det(G_j - \nu I_j)|) \varphi(y) dy,$$

where ν is given by (8.30).

For the conditional expectation in (8.19) plug the last expression into (8.38) and use (8.33), (8.32) and Lemma 8.7. For the density in (8.19) use (8.35). Then, after some algebra, Lemma 8.6 gives (8.25). \square

Remarks on the theorem.

- The “principal term” is

$$(8.39) \quad \varphi(x) \left\{ \sum_{t \in S_0} \widehat{\sigma}_0(t) + \sum_{j=1}^{d_0} \left[\left(\frac{|\rho'|}{\pi} \right)^{j/2} \overline{H}_j(x) \right] g_j \right\},$$

which is the product of a standard normal density times a polynomial with degree d_0 . Integrating once, we get -in our special case- the formula for the expectation of the EPC of the excursion set given in Adler and Taylor (2007).

- The “complementary term”

$$(8.40) \quad \varphi(x) \sum_{j=1}^{d_0} R_j(x) g_j,$$

can be computed by means of a formula, as it follows from the statement of the theorem. These formulas will be in general quite unpleasant due to the complicated form of $T_j(v)$. However, for low dimensions they are simple. For example:

$$(8.41) \quad T_1(v) = \sqrt{2\pi}[\varphi(v) - v(1 - \Phi(v))],$$

$$(8.42) \quad T_2(v) = 2\sqrt{2\pi}\varphi(v),$$

$$(8.43) \quad T_3(v) = \sqrt{\frac{\pi}{2}}[3(2v^2 + 1)\varphi(v) - (2v^2 - 3)v(1 - \Phi(v))].$$

- Second order asymptotics for $p_M(x)$ as $x \rightarrow +\infty$ will be mainly considered in the following sections. However, we can state already that the complementary term (8.40) is equivalent, as $x \rightarrow +\infty$, to

$$(8.44) \quad \varphi(x) g_{d_0} K_{d_0} x^{2d_0-4} e^{-\frac{1}{2}\frac{\gamma^2}{3-\gamma^2}x^2},$$

where the constant K_j , $j = 1, 2, \dots$ is given by:

$$(8.45) \quad K_j = 2^{3j-2} \frac{\Gamma(\frac{j+1}{2})}{\sqrt{\pi}(2\pi\gamma)^{j/2}(j-1)!} \rho''_{j/4} \left(\frac{\gamma}{3-\gamma^2}\right)^{2j-4}.$$

We are not going to go through this calculation, which is completely elementary but requires some extra work which is left to the reader. An outline of it is the following. Replace the Hermite polynomials in the expression for $T_j(v)$ given by (8.31) by the well-known expansion (see again Mehta’s book):

$$(8.46) \quad H_j(v) = j! \sum_{i=0}^{[j/2]} (-1)^i \frac{(2v)^{j-2i}}{i!(j-2i)!}$$

and $I_{j-1}(v)$ by means of the formula in Lemma 8.7.

Evaluating the term of highest degree in the polynomial part, one proves that, as $v \rightarrow +\infty$, $T_j(v)$ is equivalent to

$$(8.47) \quad \frac{2^{j-1}}{\sqrt{\pi}(j-1)!} v^{2j-4} e^{-\frac{v^2}{2}}.$$

Using now the definition of $R_j(x)$ and changing variables in the integral in (8.29), one gets for $R_j(x)$ the equivalent:

$$(8.48) \quad K_j x^{2j-4} e^{-\frac{1}{2}\frac{\gamma^2}{3-\gamma^2}x^2}.$$

In particular, the equivalent of (8.40) is given by the highest order non-vanishing term in the sum.

- Consider now the case in which S is the sphere \mathcal{S}^{d-1} and the process satisfies the same conditions as in the theorem. Even though the theorem can not be applied directly, it is possible to deal with this example to compute $\bar{p}(x)$, only performing some minor changes. In this case, only the term that corresponds to $j = d - 1$ in (8.19) does not vanish, $\widehat{C}_{t,d-1} = N_{t,d-1}$, so that $\mathbf{1}_{X'_{d-1,N}(t) \in \widehat{C}_{t,d-1}} = 1$ for each $t \in \mathcal{S}^{d-1}$ and one can use invariance under the orthogonal group to obtain:

$$(8.49) \quad \bar{p}(x) = \varphi(x) \frac{\sigma_{d-1}(\mathcal{S}^{d-1})}{(2\pi)^{(d-1)/2}} \mathbb{E}(|\det(Z + 2\rho'xI_{d-1}) + (2|\rho'|)^{1/2}\eta I_{d-1}|)$$

where Z is a $(d-1) \times (d-1)$ centered Gaussian matrix with the covariance structure of Lemma 8.5 and η is a standard normal real random variable, independent of Z . (8.49) follows from the fact that the normal derivative at each point is centered Gaussian with variance $2|\rho'|$ and independent of the tangential derivative. So, we apply the previous computation, replacing x by $x + (2|\rho'|)^{-1/2}\eta$ and obtain the expression:

$$(8.50) \quad \bar{p}(x) = \varphi(x) \frac{2\pi^{d/2}}{\Gamma(d/2)} \int_{-\infty}^{+\infty} \left[\left(\frac{|\rho'|}{\pi} \right)^{(d-1)/2} \bar{H}_{d-1}(x + (2|\rho'|)^{-1/2}y) + R_{d-1}(x + (2|\rho'|)^{-1/2}y) \right] \varphi(y) dy.$$

5. Asymptotics as $x \rightarrow +\infty$

In this section we will consider the errors in the direct and the EPC methods (see Section 3) for large values of the argument x . These errors are:

$$(8.51) \quad \bar{p}(x) - p_M(x) = \sum_{t \in S_0} \mathbb{E}(\mathbf{1}_{X'(t) \in \hat{C}_{t,0}} \cdot \mathbf{1}_{M > x} | X(t) = x) p_{X(t)}(x) \\ + \sum_{j=1}^{d_0} \int_{S_j} \mathbb{E}(|\det(X_j''(t))| \mathbf{1}_{X_{j,N}'(t) \in \hat{C}_{t,j}} \cdot \mathbf{1}_{M > x} | X(t) = x, X_j'(t) = 0) p_{X(t), X_j'(t)}(x, 0) \sigma_j(dt).$$

$$(8.52) \quad p^E(x) - p_M(x) = \sum_{t \in S_0} \mathbb{E}(\mathbf{1}_{X'(t) \in \hat{C}_{t,0}} \cdot \mathbf{1}_{M > x} | X(t) = x) p_{X(t)}(x) \\ + \sum_{j=1}^{d_0} (-1)^j \int_{S_j} \mathbb{E}(\det(X_j''(t)) \mathbf{1}_{X_{j,N}'(t) \in \hat{C}_{t,j}} \cdot \mathbf{1}_{M > x} | X(t) = x, X_j'(t) = 0) p_{X(t), X_j'(t)}(x, 0) \sigma_j(dt).$$

It is clear that for every real x ,

$$|p^E(x) - p_M(x)| \leq \bar{p}(x) - p_M(x)$$

so that the upper bounds for $\bar{p}(x) - p_M(x)$ will automatically be upper bounds for $|p^E(x) - p_M(x)|$.

Our next theorem gives sufficient conditions allowing to ensure that the error

$$\bar{p}(x) - p_M(x)$$

is bounded by a constant times a Gaussian density having strictly smaller variance than the maximum variance of the given process \mathcal{X} . In this theorem, we assume that the maximum of the variance is not attained in $S \setminus S_{d_0}$. This excludes constant variance or some other stationary-like condition. This kind of processes will be considered later on in Theorem 8.12. For parameter dimension $d_0 > 1$, a result of this type for non-constant variance processes is Theorem 3.3 of Taylor, Takemura and Adler (2005).

THEOREM 8.10. *Assume that the process \mathcal{X} satisfies conditions A1 -A5 of Chapter 7, Section 1. With no loss of generality, we assume that $\max_{t \in S} \text{Var}(X(t)) = 1$. In addition, we will assume that the set S_v of points $t \in S$ where the variance of $X(t)$ attains its maximal value is contained in S_{d_0} ($d_0 > 0$) the non-empty face having largest dimension and that no point in S_v is a boundary point of $S \setminus S_{d_0}$. Then, there exist some positive constants C, δ such that for every $x > 0$.*

$$(8.53) \quad |p^E(x) - p_M(x)| \leq \bar{p}(x) - p_M(x) \leq C\varphi(x(1 + \delta)),$$

where $\varphi(\cdot)$ is the standard normal density.

PROOF. Let W be an open neighborhood of the compact subset S_v of S such that $\text{dist}(W, (S \setminus S_{d_0})) > 0$ where dist denote the Euclidean distance in \mathbb{R}^d . For $t \in S_j \cap W^c$, the density

$$p_{X(t), X'_j(t)}(x, 0)$$

can be written as the product of the density of $X'_j(t)$ at the point 0, times the conditional density of $X(t)$ at the point x given that $X'_j(t) = 0$, which is Gaussian with some bounded expectation and a conditional variance which is smaller than the unconditional variance, hence, bounded by some constant smaller than 1. Since the conditional expectations in (8.51) are uniformly bounded by some constant, due to the bounds on the moments of the Gaussian law (see Chapter 2), one can deduce that:

$$(8.54) \quad \bar{p}(x) - p_M(x) = \int_{W \cap S_{d_0}} \mathbb{E}(|\det(X''_{d_0}(t))| \mathbf{1}_{X'_{d_0, N}(t) \in \hat{C}_{t, d_0}} \cdot \mathbf{1}_{M > x} | X(t) = x, X'_{d_0}(t) = 0) \\ \cdot p_{X(t), X'_{d_0}(t)}(x, 0) \sigma_{d_0}(dt) + O(\varphi((1 + \delta_1)x)),$$

as $x \rightarrow +\infty$, for some $\delta_1 > 0$. Our following task is to choose W such that one can assure that the first term in the right hand-side of (8.54) has the same form as the second, with a possibly different positive constant δ_1 .

To do this, for $s \in S$ and $t \in S_{d_0}$, let us write the Gaussian regression formula of $X(s)$ on the pair $(X(t), X'_{d_0}(t))$:

$$(8.55) \quad X(s) = a^t(s)X(t) + \langle b^t(s), X'_{d_0}(t) \rangle + \frac{\|t - s\|^2}{2} X^t(s).$$

where the regression coefficients $a^t(s), b^t(s)$ are respectively real-valued and \mathbb{R}^{d_0} -valued.

From now onwards, we will only be interested in those $t \in W$. In this case, since W does not contain boundary points of $S \setminus S_{d_0}$, it follows that

$$\hat{C}_{t, d_0} = N_{t, d_0} \quad \text{and} \quad \mathbf{1}_{X'_{d_0, N}(t) \in \hat{C}_{t, d_0}} = 1.$$

Moreover, whenever $s \in S$ is close enough to t it has to belong to S_{d_0} . For each t , $\{X^t(s) : s \in S\}$ is a ‘‘helix process’’ (compare with the Chapter 7, Section 3 where a similar normalization has been introduced).

Let us prove that, almost surely, the paths of the real-valued random field $\{X^t(s) : t \in W \cap S_{d_0}, s \in S\}$ are bounded. With no loss of generality (take a chart) we may assume that s varies in a closed ball in \mathbb{R}^{d_0} containing t as an interior point, and remove the subscript d_0 for the derivative. Let us write the Taylor expansion of $X(\cdot)$ around the point t :

$$(8.56) \quad X(s) = X(t) + \langle s - t, X'(t) \rangle + \|s - t\|^2 \int_0^1 \langle v, X''((1 - \alpha)t + \alpha s)(1 - \alpha)v \rangle d\alpha$$

where $v = \frac{s-t}{\|s-t\|}$, and $\langle \cdot, \cdot \rangle, \|\cdot\|$ denote respectively Euclidean scalar product and norm in \mathbb{R}^{d_0} .

For each α perform the Gaussian regression of $\langle v, X''((1 - \alpha)t + \alpha s)(1 - \alpha)v \rangle$ on the pair $(X(t), X'(t))$, that is:

$$(8.57) \quad \langle v, X''((1 - \alpha)t + \alpha s)(1 - \alpha)v \rangle = \tilde{a}^t(s, \alpha, v)X(t) + \langle \tilde{b}^t(s, \alpha, v), X'(t) \rangle + \tilde{X}^t(s, \alpha, v)$$

where the notation is as above, mutatis mutandis. From the regression formulas it follows that $\tilde{a}^t(s, \alpha), \tilde{b}^t(s, \alpha)$ are uniformly bounded, independently of s, t, α, v . Comparing (8.55) with (8.57), it follows that

$$X^t(s) = 2 \int_0^1 \tilde{X}^t(s, \alpha, v)(1 - \alpha) d\alpha$$

and the a.s. boundedness of $X^t(s), t \in W \cap S_{d_0}, s \in S$ follows (The reader may check that for $d_0 > 1$, even though $\lim_{r \downarrow 0} X^t(t + rv)$ exists for each $v \in S^{d_0-1}$, it may depend on v , so that the function $X^t(\cdot)$ can be discontinuous at $s = t$).

Let us go back to formula (8.54).

Conditionally on $X(t) = x, X'_{d_0}(t) = 0$ the event $\{M > x\}$ can be written as

$$\{X^t(s) > \beta^t(s) x, \text{ for some } s \in S\}$$

where

$$(8.58) \quad \beta^t(s) = \frac{2(1 - a^t(s))}{\|t - s\|^2}.$$

Our next goal is to prove that if we can choose W in such a way that

$$(8.59) \quad \inf\{\beta^t(s) : t \in W \cap S_{d_0}, s \in S, s \neq t\} > 0,$$

then we are done. In fact, apply the Cauchy-Schwarz inequality to the conditional expectation in (8.54). Under the conditioning, the elements of $X''_{d_0}(t)$ are the sums of affine functions of x with bounded coefficients plus centered Gaussian variables with bounded variances, hence, the absolute value of the conditional expectation is bounded by an expression of the form

$$(8.60) \quad (Q(t, x))^{1/2} \left(\mathbb{P} \left(\sup_{s \in S \setminus \{t\}} \frac{X^t(s)}{\beta^t(s)} > x \right) \right)^{1/2},$$

where $Q(t, x)$ is a polynomial in x of degree $2d_0$ with bounded coefficients. For each $t \in W \cap S_{d_0}$, the second factor in (8.60) is bounded by

$$\left(\mathbb{P} \left(\sup \left\{ \frac{X^t(s)}{\beta^t(s)} : t \in W \cap S_{d_0}, s \in S, s \neq t \right\} > x \right) \right)^{1/2}.$$

Now, we apply to the bounded separable Gaussian process

$$\left\{ \frac{X^t(s)}{\beta^t(s)} : t \in W \cap S_{d_0}, s \in S, s \neq t \right\}$$

the basic inequality 2.33 of Chapter 2, which gives the bound

$$\mathbb{P} \left(\sup \left\{ \frac{X^t(s)}{\beta^t(s)} : t \in W \cap S_{d_0}, s \in S, s \neq t \right\} > x \right) \leq C_2 \exp(-\delta_2 x^2),$$

for some positive constants C_2, δ_2 and any $x > 0$. Also, the same argument above for the density $p_{X(t), X'_{d_0}(t)}(x, 0)$ shows that it is bounded by a constant times the standard normal density. To finish, it suffices to replace these bounds in the first term at the right-hand side of (8.54).

It remains to choose W for (8.59) to hold true. Consider the auxiliary process

$$(8.61) \quad Y(s) := \frac{X(s)}{\sqrt{r(s, s)}}, \quad s \in S.$$

Clearly, $\text{Var}(Y(s)) = 1$ for all $s \in S$. We set

$$r^Y(s, s') := \text{Cov}(Y(s), Y(s')), \quad s, s' \in S.$$

Let us assume that $t \in S_v$. Since the function $s \rightsquigarrow \text{Var}(X(s))$ attains its maximum value at $s = t$, it follows that $X(t), X'_{d_0}(t)$ are independent. This implies that in the regression formula (8.55) the coefficients are easily computed and $a^t(s) = r(s, t)$ which is strictly smaller than 1 if $s \neq t$, because of the non-degeneracy condition.

Then

$$(8.62) \quad \beta^t(s) = \frac{2(1 - r(s, t))}{\|t - s\|^2} \geq \frac{2(1 - r^Y(s, t))}{\|t - s\|^2}.$$

Since $r^Y(s, s) = 1$ for every $s \in S$, the Taylor expansion of $r^Y(s, t)$ as a function of s , around $s = t$ takes the form:

$$(8.63) \quad r^Y(s, t) = 1 + \langle s - t, r_{20, d_0}^Y(t, t)(s - t) \rangle + o(\|s - t\|^2),$$

where the notation is self-explanatory.

Also, using that $\text{Var}(Y(s)) = 1$ for $s \in S$, we easily obtain:

$$(8.64) \quad -r_{20, d_0}^Y(t, t) = \text{Var}(Y'_{d_0}(t)) = \text{Var}(X'_{d_0}(t))$$

where the last equality follows by differentiation in (8.61) and putting $s = t$. (8.64) implies that $-r_{20, d_0}^Y(t, t)$ is uniformly positive definite on $t \in S_v$, meaning that its minimum eigenvalue has a strictly positive lower bound. This, on account of (8.62) and (8.63), already shows that

$$(8.65) \quad \inf\{\beta^t(s) : t \in S_v, s \in S, s \neq t\} > 0,$$

The foregoing argument also shows that

$$(8.66) \quad \inf\{-\tau(a^t)''_{d_0}(t)\tau : t \in S_v, \tau \in \mathcal{S}^{d_0-1}, s \neq t\} > 0,$$

since whenever $t \in S_v$, one has $a^t(s) = r(s, t)$ so that

$$(a^t)''_{d_0}(t) = r_{20, d_0}(t, t).$$

To end up, assume there is no neighborhood W of S_v satisfying (8.59). In that case using a compactness argument, one can find two convergent sequences $\{s_n\} \subset S$, $\{t_n\} \subset S_{d_0}$, $s_n \rightarrow s_0$, $t_n \rightarrow t_0 \in S_v$ such that

$$\beta^{t_n}(s_n) \rightarrow \ell \leq 0.$$

ℓ may be $-\infty$.

$t_0 \neq s_0$ is not possible, since it would imply

$$\ell = 2 \frac{(1 - a^{t_0}(s_0))}{\|t_0 - s_0\|^2} = \beta^{t_0}(s_0),$$

which is strictly positive.

If $t_0 = s_0$, on differentiating in (8.55) with respect to s along S_{d_0} we get:

$$X'_{d_0}(s) = (a^t)'_{d_0}(s)X(t) + \langle (b^t)'_{d_0}(s), X'_{d_0}(t) \rangle + \frac{\partial_{d_0} \|t - s\|^2}{\partial s} X^t(s),$$

where $(a^t)'_{d_0}(s)$ is a column vector of size d_0 and $(b^t)'_{d_0}(s)$ is a $d_0 \times d_0$ matrix. Then, one must have $a^t(t) = 1$, $(a^t)'_{d_0}(t) = 0$. Thus

$$\beta^{t_n}(s_n) = -u_n^T (a^{t_0})''_{d_0}(t_0) u_n + o(1),$$

where $u_n := (s_n - t_n)/\|s_n - t_n\|$. Since $t_0 \in S_v$ we may apply (8.66) and the limit ℓ of $\beta^{t_n}(s_n)$ cannot be non-positive. \square

A straightforward application of Theorem 8.10 is the following

COROLLARY 8.11. *Under the hypotheses of Theorem 8.10, there exists positive constants C, δ such that, for every $u > 0$:*

$$0 \leq \left| \int_u^{+\infty} p^E(x) dx - \text{P}(M > u) \right| \leq \int_u^{+\infty} \bar{p}(x) dx - \text{P}(M > u) \leq C \text{P}(\xi > u),$$

where ξ is a centered Gaussian variable with variance $1 - \delta$

The precise order of approximation of $\bar{p}(x) - p_M(x)$ or $p^E(x) - p_M(x)$ as $x \rightarrow +\infty$ remains in general an open problem, even if one only asks for the constants σ_d^2, σ_E^2 respectively which govern the second order asymptotic approximation and which are defined by means of

$$(8.67) \quad \frac{1}{\sigma_d^2} := \lim_{x \rightarrow +\infty} -2x^{-2} \log [\bar{p}(x) - p_M(x)]$$

and

$$(8.68) \quad \frac{1}{\sigma_E^2} := \lim_{x \rightarrow +\infty} -2x^{-2} \log |p^E(x) - p_M(x)|$$

whenever these limits exist. In general, we are unable to compute the limits (8.67) or (8.68) or even to prove that they actually exist or differ. In the remainder of this chapter, we will give some lower-bounds for the \liminf as $x \rightarrow +\infty$. This is already interesting since it gives some upper-bounds for the speed of approximation for $p_M(x)$ either by means of $\bar{p}(x)$ or $p^E(x)$. A more precise result is Theorem 8.15 below, where we are able to prove the existence of the limit and compute σ_d^2 when \mathcal{X} is centered Gaussian, defined on a convex parameter set, and has a law which is invariant under isometries and translations of \mathbb{R}^d .

For the next theorem we need an additional condition on the parameter set S . For S verifying A1 we define

$$(8.69) \quad \kappa(S) = \sup_{0 \leq j \leq d_0} \sup_{t \in S_j} \sup_{s \in S, s \neq t} \frac{\text{dist}((t-s), \mathcal{C}_{t,j})}{\|s-t\|^2}$$

where dist is the Euclidean distance in \mathbb{R}^d .

In Exercise 8.2 at the end of this chapter, it is proved that $\kappa(S) < \infty$ for various relevant classes of parameter sets:

- S is convex (in which case, in fact $\kappa(S) = 0$),
- S is a C^3 manifold, with or without boundary
- S verifies a certain kind of local convexity condition which is precisely described in the exercise.

However, $\kappa(S) < \infty$ can fail in general. A simple example showing what is going on is the following: take an orthonormal basis of \mathbb{R}^2 and put

$$S = \{(\lambda, 0) : 0 \leq \lambda \leq 1\} \cup \{(\mu \cos \theta, \mu \sin \theta) : 0 \leq \mu \leq 1\}$$

where $0 < \theta < \pi$, that is, S is the boundary of an angle of size θ . One easily checks that $\kappa(S) = +\infty$.

THEOREM 8.12. *Let \mathcal{X} be a stochastic process on S satisfying conditions A1-A5 in Section 1 of Chapter 7. Suppose in addition that $\text{Var}(X(t)) = 1$ for all $t \in S$ and that $\kappa(S) < +\infty$.*

Then

$$(8.70) \quad \liminf_{x \rightarrow +\infty} -2x^{-2} \log [\bar{p}(x) - p_M(x)] \geq 1 + \inf_{t \in S} \frac{1}{\sigma_t^2 + \bar{\lambda}(t)\kappa_t^2}$$

with

$$\sigma_t^2 := \sup_{s \in S \setminus \{t\}} \frac{\text{Var}(X(s) | X(t), X'(t))}{(1 - r(s, t))^2}$$

and

$$(8.71) \quad \kappa_t := \sup_{s \in S \setminus \{t\}} \frac{\text{dist}(-\Lambda_t^{-1} r_{01}(s, t), \mathcal{C}_{t,j})}{1 - r(s, t)},$$

where

- $\Lambda_t := \text{Var}(X'(t))$
- $\bar{\lambda}(t)$ is the maximum eigenvalue of Λ_t
- in (8.71), j is such that $t \in S_j$, ($j = 0, 1, \dots, d_0$).

The quantity in the right hand side of (8.70) is strictly bigger than 1.

Remark. In formula (8.70) it may happen that the denominator in the right-hand side is identically zero, in which case we put $+\infty$ for the infimum. This is the case of the one-parameter process $X(t) = \xi \cos t + \eta \sin t$ where ξ, η are independent standard normal random variables (the sine-cosine process), and S is an interval having length strictly smaller than π .

PROOF. Let us first prove that $\sup_{t \in S} \kappa_t < \infty$.

For each $t \in S$, let us write the Taylor expansions

$$\begin{aligned} r_{01}(s, t) &= r_{01}(t, t) + r_{11}(t, t)(s - t) + O(\|s - t\|^2) \\ &= \Lambda_t(s - t) + O(\|s - t\|^2) \end{aligned}$$

where O is uniform on $s, t \in S$, and

$$1 - r(s, t) = (s - t)^T \Lambda_t(s - t) + O(\|s - t\|^2) \geq L_2 \|s - t\|^2,$$

where L_2 is some positive constant. It follows that for $s \in S$, $t \in S_j$, $s \neq t$, one has:

$$(8.72) \quad \frac{\text{dist}\left(-\Lambda_t^{-1} r_{01}(s, t), \mathcal{C}_{t,j}\right)}{1 - r(s, t)} \leq L_3 \frac{\text{dist}\left((t - s), \mathcal{C}_{t,j}\right)}{\|s - t\|^2} + L_4,$$

where L_3 and L_4 are positive constants. So,

$$\frac{\text{dist}\left(-\Lambda_t^{-1} r_{01}(s, t), \mathcal{C}_{t,j}\right)}{1 - r(s, t)} \leq L_3 \kappa(S) + L_4.$$

which implies $\sup_{t \in S} \kappa_t < \infty$.

With the same notations as in the proof of Theorem 8.10, using (7.1) and (8.19), one has:

$$(8.73) \quad \bar{p}(x) - p_M(x) = \varphi(x)B(x)$$

where

$$\begin{aligned} B(x) &:= \sum_{t \in S_0} \mathbb{E}(\mathbf{1}_{X'_j(t) \in \hat{\mathcal{C}}_{t,0}} \mathbf{1}_{M > x} | X(t) = x) + \sum_{j=1}^{d_0} (2\pi)^{-j/2} \\ &\int_{S_j} \mathbb{E}(|\det(X''_j(t))| \mathbf{1}_{X'_{j,N}(t) \in \hat{\mathcal{C}}_{t,j}} \mathbf{1}_{M > x} | X(t) = x, X'_j(t) = 0) [\det(\text{Var}(X'_j(t)))]^{-1/2} \sigma_j(dt). \end{aligned}$$

Proceeding in a similar way to that of the proof of Theorem 8.10, an application of the Hölder inequality to the conditional expectation in each term in the right-hand side of (8.73) shows that the desired result will follow as soon as we prove that:

$$(8.74) \quad \liminf_{x \rightarrow +\infty} -2x^{-2} \log \mathbb{P}(\{X'_{j,N} \in \hat{\mathcal{C}}_{t,j}\} \cap \{M > x\} | X(t) = x, X'_j(t) = 0) \geq \frac{1}{\sigma_t^2 + \lambda(t)\kappa_t^2},$$

for each $j = 0, 1, \dots, d_0$, where the \liminf has some uniformity in t .

Let us write the Gaussian regression of $X(s)$ on the pair $(X(t), X'(t))$

$$X(s) = a^t(s)X(t) + \langle b^t(s), X'(t) \rangle + R^t(s).$$

Since $X(t)$ and $X'(t)$ are independent, one easily computes :

$$\begin{aligned} a^t(s) &= r(s, t) \\ b^t(s) &= \Lambda_t^{-1} r_{01}(s, t). \end{aligned}$$

Hence, conditionally on $X(t) = x$, $X'_j(t) = 0$, the events

$$\{M > x\} \quad \text{and} \quad \{R^t(s) > (1 - r(s, t))x - r_{01}^T(s, t)\Lambda_t^{-1} X'_{j,N}(t) \text{ for some } s \in S\}$$

coincide.

Denote by $(X'_{j,N}(t) | X'_j(t) = 0)$ the regression of $X'_{j,N}(t)$ on $X'_j(t) = 0$. So, the probability in (8.74) can be written as

$$(8.75) \quad \int_{\hat{\mathcal{C}}_{t,j}} \mathbb{P}\left\{s^t(s) > x - \frac{r_{01}^T(s, t)\Lambda_t^{-1} x'}{1 - r(s, t)} \text{ for some } s \in S\right\} p_{X'_{j,N}(t) | X'_j(t)=0}(x') dx'$$

where

- $\zeta^t(s) := \frac{R^t(s)}{1 - r(s, t)}$
- dx' is the Lebesgue measure on $N_{t,j}$. Recall that $\widehat{C}_{t,j} \subset N_{t,j}$.

If $-\Lambda_t^{-1}r_{01}(s, t) \in \mathcal{C}_{t,j}$ one has

$$-r_{01}^T(s, t)\Lambda_t^{-1}x' \geq 0$$

for every $x' \in \widehat{C}_{t,j}$, because of the definition of $\widehat{C}_{t,j}$.

If $-\Lambda_t^{-1}r_{01}(s, t) \notin \mathcal{C}_{t,j}$, since $\mathcal{C}_{t,j}$ is a closed convex cone, we can write

$$-\Lambda_t^{-1}r_{01}(s, t) = z' + z''$$

with $z' \in \mathcal{C}_{t,j}$, $z' \perp z''$ and $\|z''\| = \text{dist}(-\Lambda_t^{-1}r_{01}(s, t), \mathcal{C}_{t,j})$.

So, if $x' \in \widehat{C}_{t,j}$:

$$\frac{-r_{01}^T(s, t)\Lambda_t^{-1}x'}{1 - r(s, t)} = \frac{z'^T x' + z''^T x'}{1 - r(s, t)} \geq -\kappa_t \|x'\|$$

using that $z'^T x' \geq 0$ and the Cauchy-Schwarz inequality. It follows that in any case, if $x' \in \widehat{C}_{t,j}$ the expression in (8.75) is bounded by

$$(8.76) \quad \int_{\widehat{C}_{t,j}} \mathbb{P}\left(\zeta^t(s) > x - \kappa_t \|x'\| \text{ for some } s \in S\right) p_{X'_{j,N}(t)|X'_j(t)=0}(x') dx'.$$

To obtain a bound for the probability in the integrand of (8.76) we will use the classical inequality for the tail of the distribution of the supremum of a Gaussian process with bounded paths.

The Gaussian process $(s, t) \rightsquigarrow \zeta^t(s)$, defined on $(S \times S) \setminus \{s = t\}$ has continuous paths. As the pair (s, t) approaches the diagonal of $S \times S$, $\zeta^t(s)$ may not have a limit but, almost surely, one can prove that it is bounded using a similar argument to the one in the proof of Theorem 8.10 for ‘‘helix processes’’, that is, Taylor expansion followed by Gaussian regression.

We set

- $m^t(s) := \mathbb{E}(\zeta^t(s))$ ($s \neq t$)
- $m := \sup_{s, t \in S, s \neq t} |m^t(s)|$
- $\mu := \mathbb{E}(|\sup_{s, t \in S, s \neq t} [\zeta^t(s) - m^t(s)]|)$.

The almost sure boundedness of the paths of $\zeta^t(s)$ implies that $m < \infty$ and $\mu < \infty$. Applying the basic inequality (2.25) of Theorem 2.9 to the centered process $s \rightsquigarrow \zeta^t(s) - m^t(s)$ defined on $S \setminus \{t\}$, we get whenever $x - \kappa_t \|x'\| - m - \mu > 0$:

$$\begin{aligned} & \mathbb{P}\{\zeta^t(s) > x - \kappa_t \|x'\| \text{ for some } s \in S\} \\ & \leq \mathbb{P}\{\zeta^t(s) - m^t(s) > x - \kappa_t \|x'\| - m \text{ for some } s \in S\} \\ & \leq 2 \exp\left(-\frac{(x - \kappa_t \|x'\| - m - \mu)^2}{2\sigma_t^2}\right). \end{aligned}$$

The Gaussian density in the integrand of (8.76) is bounded by

$$(2\pi\lambda_j(t))^{\frac{j-d}{2}} \exp\left(-\frac{\|x' - m'_{j,N}(t)\|^2}{2\bar{\lambda}_j(t)}\right)$$

where $\lambda_j(t)$ and $\bar{\lambda}_j(t)$ are respectively the minimum and maximum eigenvalue of $\text{Var}(X'_{j,N}(t)|X'_j(t))$ and $m'_{j,N}(t)$ is the conditional expectation $\mathbb{E}(X'_{j,N}(t)|X'_j(t) = 0)$. Notice that $\lambda_j(t)$, $\bar{\lambda}_j(t)$, $m'_{j,N}(t)$ are bounded, $\lambda_j(t)$ is bounded below by a positive constant and $\bar{\lambda}_j(t) \leq \bar{\lambda}(t)$.

Replacing into (8.76) we have the bound :

$$(8.77) \quad \begin{aligned} & \mathbb{P}(\{X'_{j,N} \in \widehat{C}_{t,j}\} \cap \{M > x\} | X(t) = x, X'_j(t) = 0) \\ & \leq (2\pi\lambda_j(t))^{\frac{i-d}{2}} 2 \int_{\widehat{C}_{t,j} \cap \{x - \kappa_t \|x'\| - m - \mu > 0\}} \exp\left(-\left(\frac{(x - \kappa_t \|x'\| - m - \mu)^2}{2\sigma_t^2} + \frac{\|x' - m'_{j,N}(t)\|^2}{2\bar{\lambda}(t)}\right)\right) dx' \\ & \quad + \mathbb{P}\left(\|X'_{j,N}(t) | X'_j(t) = 0\| \geq \frac{x - m - \mu}{\kappa_t}\right), \end{aligned}$$

where it is understood that the second term in the right-hand side vanishes if $\kappa_t = 0$. Let us consider the first term in the right-hand side of (8.77). We have:

$$\begin{aligned} & \frac{(x - \kappa_t \|x'\| - m - \mu)^2}{2\sigma_t^2} + \frac{\|x' - m'_{j,N}(t)\|^2}{2\bar{\lambda}(t)} \\ & \geq \frac{(x - \kappa_t \|x'\| - m - \mu)^2}{2\sigma_t^2} + \frac{(\|x'\| - \|m'_{j,N}(t)\|)^2}{2\bar{\lambda}(t)} \\ & = [A(t)\|x'\| + B(t)(x - m - \mu) + C(t)]^2 + \frac{(x - m - \mu - \kappa_t \|m'_{j,N}(t)\|)^2}{2\sigma_t^2 + 2\bar{\lambda}(t)\kappa_t^2}, \end{aligned}$$

where the last inequality is obtained after some algebra, $A(t), B(t), C(t)$ are bounded functions and $A(t)$ is bounded below by some positive constant.

So, the first term in the right-hand side of (8.77) is bounded by :

$$(8.78) \quad \begin{aligned} & 2 \cdot (2\pi\lambda_j)^{\frac{i-d}{2}} \exp\left(-\frac{(x - m - \mu - \kappa_t \|m'_{j,N}(t)\|)^2}{2\sigma_t^2 + 2\bar{\lambda}(t)\kappa_t^2}\right) \\ & \int_{\mathbb{R}^{d-j}} \exp\left(-[A(t)\|x'\| + B(t)(x - m - \mu) + C(t)]^2\right) dx' \\ & \leq L|x|^{d-j-1} \exp\left(-\frac{(x - m - \mu - \kappa_t \|m'_{j,N}(t)\|)^2}{2\sigma_t^2 + 2\bar{\lambda}(t)\kappa_t^2}\right) \end{aligned}$$

where L is some constant. The last inequality follows easily using polar coordinates.

Consider now the second term in the right-hand side of (8.77). Using the form of the conditional density $p_{X'_{j,N}(t) | X'_j(t)=0}(x')$, it follows that it is bounded by

$$(8.79) \quad \begin{aligned} & \mathbb{P}\left\{\|(X'_{j,N}(t) | X'_j(t) = 0) - m'_{j,N}(t)\| \geq \frac{x - m - \mu - \kappa_t \|m'_{j,N}(t)\|}{\kappa_t}\right\} \\ & \leq L_1|x|^{d-j-2} \exp\left(-\frac{(x - m - \mu - \kappa_t \|m'_{j,N}(t)\|)^2}{2\bar{\lambda}(t)\kappa_t^2}\right) \end{aligned}$$

where L_1 is some constant. Putting together (8.78) and (8.79) with (8.77), we obtain (8.74). \square

The following two corollaries are straightforward consequences of Theorem 8.12:

COROLLARY 8.13. *Under the hypotheses of Theorem 8.12 one has*

$$\liminf_{x \rightarrow +\infty} -2x^{-2} \log |p^E(x) - p_M(x)| \geq 1 + \inf_{t \in S} \frac{1}{\sigma_t^2 + \bar{\lambda}(t)\kappa_t^2}.$$

COROLLARY 8.14. *Let \mathcal{X} a stochastic process on S satisfying A1 -A5. Suppose in addition that $\mathbb{E}(X(t)) = 0$, $\mathbb{E}(X^2(t)) = 1$, $\text{Var}(X'(t)) = I_d$ for all $t \in S$.*

Then

$$\liminf_{u \rightarrow +\infty} -2u^{-2} \log \left| \mathbb{P}(M > u) - \int_u^{+\infty} p^E(x) dx \right| \geq 1 + \inf_{t \in S} \frac{1}{\sigma_t^2 + \kappa_t^2}.$$

and

$$p^E(x) = \left[\sum_{j=0}^{d_0} (-1)^j (2\pi)^{-j/2} g_j \bar{H}_j(x) \right] \varphi(x).$$

where g_j is given by (8.26) and $\bar{H}_j(x)$ has been defined in Section 4.

The proof follows directly from Theorem 8.12 the definition of $p^E(x)$ and the calculation of

$$(8.80) \quad E(|\det(X_j''(t))| | X(t) = x, X_j'(t) = 0),$$

which is detailed in Exercise 8.6.

6. Examples

1) A simple application of Theorem 8.10 is the following. Let \mathcal{X} be a one parameter real-valued centered Gaussian process with regular paths, defined on the interval $[0, T]$ and satisfying an adequate non-degeneracy condition. Assume that the variance $v(t)$ has a unique maximum, say 1 at the interior point t_0 , and $k = \min\{j : v^{(2j)}(t_0) \neq 0\} < \infty$. Notice that $v^{(2k)}(t_0) < 0$. Then, one can obtain the equivalent of $p_M(x)$ as $x \rightarrow \infty$ which is given by:

$$(8.81) \quad p_M(x) \simeq \frac{1 - v''(t_0)/2}{k C_k^{1/k}} E\left(|\xi|^{1/k - 1}\right) x^{1-1/k} \varphi(x),$$

where ξ is a standard normal random variable and $C_k = -\frac{1}{(2k)!} v^{(2k)}(t_0) + \frac{1}{4} [v''(t_0)]^2 \mathbf{1}_{k=2}$. The proof is a direct application of the Laplace method, and the reader can prove it in Exercise 8.3 at the end of this chapter.

Integrating the density from u to $+\infty$, one gets the corresponding bound for $P\{M > u\}$. An independent proof under somewhat weaker hypotheses can be found in Piterbarg (1996).

2) Let the process \mathcal{X} be centered and satisfy A1-A5. Assume that the law of the process is isotropic and stationary, so that the covariance has the form (8.21) and verifies the regularity condition of Section 4. To simplify somewhat the computations, with no loss of generality we add the normalization $\rho' = \rho'(0) = -1/2$. One can easily check that

$$(8.82) \quad \sigma_t^2 = \sup_{s \in S \setminus \{t\}} \frac{1 - \rho^2(\|s - t\|^2) - 4\rho'^2(\|s - t\|^2)\|s - t\|^2}{[1 - \rho(\|s - t\|^2)]^2}$$

Furthermore if

$$(8.83) \quad \rho'(x) \leq 0 \text{ for } x \geq 0,$$

one can show that the sup in (8.82) is attained as $\|s - t\| \rightarrow 0$ and is independent of t . Its value is

$$(8.84) \quad \sigma_t^2 = 12\rho'' - 1.$$

The proof is elementary (see Exercise 8.4 at the end of this chapter).

Let S be a convex set. For $t \in S_j$, $s \in S$:

$$(8.85) \quad \text{dist}(-r_{01}(s, t), \mathcal{C}) = \text{dist}(-2\rho'(\|s - t\|^2)(t - s), \mathcal{C}).$$

The convexity of S implies that $(t - s) \in \mathcal{C}$. Since \mathcal{C} is a convex cone and $-2\rho'(\|s - t\|^2) \geq 0$, one can conclude that $-r_{01}(s, t) \in \mathcal{C}$ so that the distance in (8.85) is equal to zero. Hence,

$$\kappa_t = 0 \text{ for every } t \in S$$

and an application of Theorem 8.12 gives the inequality

$$(8.86) \quad \liminf_{x \rightarrow +\infty} -\frac{2}{x^2} \log [\bar{p}(x) - p_M(x)] \geq 1 + \frac{1}{12\rho'' - 1}.$$

A direct consequence is that the same inequality holds true when replacing $\bar{p}(x) - p_M(x)$ by $|p^E(x) - p_M(x)|$ in (8.86). The bound for the EPC method has been obtained by other methods

in Taylor and Adler (2003).

Our next theorem improves (8.86). In fact, under the same hypotheses, it says that the liminf is an ordinary limit and the sign \geq is an equality sign.

THEOREM 8.15. *Assume that \mathcal{X} is centered, satisfies hypotheses A1-A5, the covariance has the form (8.21) with $\rho'(0) = -1/2$, $\rho'(x) \leq 0$ for $x \geq 0$. Let S be a convex set, and $d_0 = d \geq 1$. Then*

$$(8.87) \quad \lim_{x \rightarrow +\infty} -\frac{2}{x^2} \log [\bar{p}(x) - p_M(x)] = 1 + \frac{1}{12\rho'' - 1}.$$

Remark Notice that since S is convex, the added hypothesis that the maximum dimension d_0 such that S_j is not empty is equal to d is not an actual restriction. In fact, in this case, S is a subset of some affine manifold having dimension d_0 , that is, the smallest one containing S , and a simple change of parameter allows to consider S as a subset of \mathbb{R}^{d_0} .

PROOF. In view of (8.86), it suffices to prove that

$$(8.88) \quad \limsup_{x \rightarrow +\infty} -\frac{2}{x^2} \log [\bar{p}(x) - p_M(x)] \leq 1 + \frac{1}{12\rho'' - 1}.$$

Using (7.1) and the definition of $\bar{p}(x)$ given by (8.19), one has the inequality

$$(8.89) \quad \bar{p}(x) - p_M(x) \geq (2\pi)^{-d/2} \varphi(x) \int_{S_d} \mathbb{E}(|\det(X''(t))| \mathbf{1}_{M>x} | X(t) = x, X'(t) = 0) \sigma_d(dt),$$

where the right-hand side is a lower bound since it only contains the term corresponding to the largest dimension. We have already replaced the density $p_{X(t), X'(t)}(x, 0)$ by its explicit expression using the law of the process. Under the condition $\{X(t) = x, X'(t) = 0\}$ if $v_0^T X''(t) v_0 > 0$ for some $v_0 \in S^{d-1}$, a Taylor expansion around the point t implies that $M > x$. It follows that

$$(8.90) \quad \begin{aligned} \mathbb{E}(|\det(X''(t))| \mathbf{1}_{M>x} | X(t) = x, X'(t) = 0) \\ \geq \mathbb{E}(|\det(X''(t))| \mathbf{1}_{\sup_{v \in S^{d-1}} v^T X''(t) v > 0} | X(t) = x, X'(t) = 0). \end{aligned}$$

We now apply Lemma 8.5 which describes the conditional distribution of $X''(t)$ given $X(t) = x, X'(t) = 0$. Using the notations of this lemma, we may write the right-hand side of (8.90) as :

$$\mathbb{E}(|\det(Z - xId)| \mathbf{1}_{\sup_{v \in S^{d-1}} v^T Z v > x}),$$

which is obviously bounded below by

$$(8.91) \quad \begin{aligned} \mathbb{E}(|\det(Z - xId)| \mathbf{1}_{Z_{11} > x}) \\ = \int_x^{+\infty} \mathbb{E}(|\det(Z - xId)| | Z_{11} = y) (2\pi)^{-1/2} \sigma^{-1} \exp\left(-\frac{y^2}{2\sigma^2}\right) dy, \end{aligned}$$

where $\sigma^2 := \text{Var}(Z_{11}) = 12\rho'' - 1$. The conditional distribution of Z given $Z_{11} = y$ is easily deduced from Lemma 8.5. It can be represented by the random $d \times d$ real symmetric matrix

$$\tilde{Z} := \begin{pmatrix} y & Z_{12} & \dots & \dots & Z_{1d} \\ & \xi_2 + \alpha y & Z_{23} & \dots & Z_{2d} \\ & & \ddots & & \\ & & & & \xi_d + \alpha y \end{pmatrix},$$

where the random variables $\{\xi_2, \dots, \xi_d, Z_{ik}, 1 \leq i < k \leq d\}$ are independent centered Gaussian with

$$\text{Var}(Z_{ik}) = 4\rho'' \quad (1 \leq i < k \leq d); \quad \text{Var}(\xi_i) = \frac{16\rho''(8\rho'' - 1)}{12\rho'' - 1} \quad (i = 2, \dots, d); \quad \alpha = \frac{4\rho'' - 1}{12\rho'' - 1}$$

Notice that $0 < \alpha < 1$.

Choose now $\alpha_0 > 0$ such that $(1 + \alpha_0)\alpha < 1$. The expansion of $\det(\tilde{Z} - xId)$ shows that if $x(1 + \alpha_0) \leq y \leq x(1 + \alpha_0) + 1$ and x is large enough, then

$$\mathbb{E}(|\det(\tilde{Z} - xId)|) \geq L \alpha_0 (1 - \alpha(1 + \alpha_0))^{d-1} x^d,$$

where L is some positive constant. This implies that

$$\frac{1}{\sqrt{2\pi}\sigma} \int_x^{+\infty} \exp\left(-\frac{y^2}{2\sigma^2}\right) \mathbb{E}(|\det(\tilde{Z} - xId)|) dy \geq \frac{L}{\sqrt{2\pi}\sigma} \int_{x(1+\alpha_0)}^{x(1+\alpha_0)+1} \exp\left(-\frac{y^2}{2\sigma^2}\right) \alpha_0 (1 - \alpha(1 + \alpha_0))^{d-1} x^d dy$$

for x large enough. On account of (8.89), (8.90), (8.91), we conclude that for x large enough,

$$\bar{p}(x) - p_M(x) \geq L_1 x^d \exp\left[-\frac{x^2}{2} + \frac{(x(1 + \alpha_0) + 1)^2}{2\sigma^2}\right].$$

for some new positive constant L_1 . Since α_0 can be chosen arbitrarily small, this implies (8.88). \square

3) Consider the same processes of Example 2, but now defined on the non-convex set $\{a \leq \|t\| \leq b\}$, $0 < a < b$. The same calculations as above show that $\kappa_t = 0$ if $a < \|t\| \leq b$ and

$$\kappa_t = \max \left\{ \sup_{z \in [2a, a+b]} \frac{-2\rho'(z^2)z}{1 - \rho(z^2)}, \sup_{\theta \in [0, \pi]} \frac{-2a\rho'(2a^2(1 - \cos\theta))(1 - \cos\theta)}{1 - \rho(2a^2(1 - \cos\theta))} \right\},$$

for $\|t\| = a$.

4) Let us keep the same hypotheses as in Example 2 but without assuming that the covariance is decreasing as in (8.83). The variance is still given by (8.82) but κ_t is not necessarily equal to zero. More precisely, relation (8.85) shows that

$$\kappa_t \leq \sup_{s \in S \setminus \{t\}} 2 \frac{\rho'(\|s - t\|^2)^+ \|s - t\|}{1 - \rho(\|s - t\|^2)}$$

The normalization: $\rho' = -1/2$ implies that the process \mathcal{X} is “identity speed”, that is $\text{Var}(X'(t)) = I_d$ so that $\bar{\lambda}(t) = 1$. An application of Theorem 8.12 gives

$$(8.92) \quad \liminf_{x \rightarrow +\infty} -\frac{2}{x^2} \log [\bar{p}(x) - p_M(x)] \geq 1 + 1/Z_\Delta.$$

where

$$Z_\Delta := \sup_{z \in (0, \Delta]} \frac{1 - \rho^2(z^2) - 4\rho'^2(z^2)z^2}{[1 - \rho(z^2)]^2} + \max_{z \in (0, \Delta]} \frac{4[\rho'(z^2)^+ z]^2}{[1 - \rho(z^2)]^2},$$

and Δ is the diameter of S .

Let us show on a numerical example that all these quantities can be actually computed. Suppose that $d = 2$ and let us consider the covariance $r(s, t)$ defined as follows: τ is the Fourier transform of the probability measure on \mathbb{R}^2 having density

$$\frac{1}{\pi} \exp\left(-\frac{\|z\|^2}{2}\right) \left[1 - \exp\left(-\frac{\|z\|^2}{2}\right)\right],$$

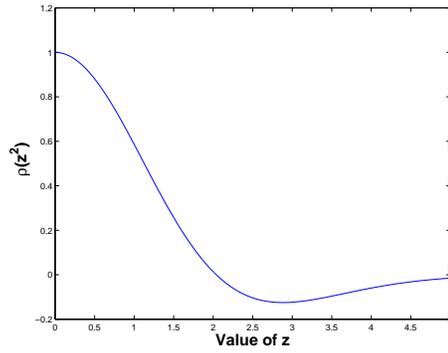
We put

$$r(s, t) := \tau\left(\sqrt{\frac{2}{3}}(s - t)\right).$$

One easily verifies that, with our previous notation,

$$r(s, t) = \rho(\|s - t\|^2) = \text{with } \rho(z) = 2\left(e^{-z/3} - \frac{e^{-z/6}}{2}\right), \quad z \geq 0.$$

Check that our conditions are satisfied (the change of scale has been chosen so that $\rho'(0) = -1/2$)

FIGURE 8.1. Representation of the function $\rho(z^2)$

Numerically we find that

$$\frac{2\rho'(z^2)+z}{1-\rho(z^2)},$$

vanishes for z in the interval $[0; 2.884\dots]$ and attains its maximum value: $0.0689\dots$ for $z = 3.7\dots$. On the other hand

$$\sup_{z \in (0, \Delta]} \frac{1 - \rho^2(z^2) - 4\rho'^2(z^2)z^2}{[1 - \rho(z^2)]^2}$$

is always attained at $Z = 0^+$ and takes the constant value $4/3$.

As a consequence, for a diameter of S smaller than $2.884\dots$, the bound for the exponent $1+1/Z_\Delta$ takes the value $7/4 = 1.75$, and takes the minimum value of 1.7473 for a diameter greater or equal to $3.7\dots$

5) Suppose that

- the process $\mathcal{X} = X(t) : t \in \mathbb{R}^d$ is stationary with covariance having the form

$$\Gamma(t_1, \dots, t_d) = \prod_{i=1, \dots, d} \Gamma_i(t_i)$$

where $\Gamma_1, \dots, \Gamma_d$ are d covariance functions on \mathbb{R} which are monotone, positive on $[0, +\infty)$ and of class C^4 ,

- S is a rectangle

$$S = \prod_{i=1, \dots, d} [a_i, b_i], a_i < b_i.$$

Then, adding an appropriate non-degeneracy condition, conditions A2-A5 are fulfilled and Theorem 8.12 applies

Clearly

$$-r_{0,1}(s, t) = \begin{bmatrix} \Gamma'_1(s_1 - t_1)\Gamma_2(s_2 - t_2)\dots\Gamma_d(s_d - t_d) \\ \vdots \\ \Gamma_1(s_1 - t_1)\dots\Gamma_{d-1}(s_{d-1} - t_{d-1})\cdot\Gamma'_d(s_d - t_d) \end{bmatrix}$$

belongs to $\mathcal{C}_{t,j}$ for every $s \in S$. As a consequence $\kappa_t = 0$ for all $t \in S$. On the other hand, standard regressions formulas show that

$$\frac{\text{Var}(X(s)|X(t), X'(t))}{(1 - r(s, t))^2} = \frac{1 - \Gamma_1^2 \dots \Gamma_d^2 - \Gamma_1'^2 \Gamma_2^2 \dots \Gamma_d^2 - \dots - \Gamma_1^2 \dots \Gamma_{d-1}^2 \Gamma_d'^2}{(1 - \Gamma_1 \dots \Gamma_d)^2},$$

where Γ_i stands for $\Gamma_i(s_i - t_i)$. Computation and maximisation of σ_t^2 should be performed numerically in each particular case.

Exercises

EXERCISE 8.1. For $n = 1, 2, \dots$, let G_n be an $n \times n$ GOE random matrix (see Section 4). Define:

$$D_n(\lambda) = (-2)^n \mathbb{E}(\det(G_n - \lambda I_n)), \quad (\lambda \in \mathbb{R})$$

a) Prove that

- $D_1(\lambda) = 2\lambda$
- $D'_n(\lambda) = 2nD_{n-1}(\lambda)$ ($n = 2, 3, \dots$)

b) Prove that

$$D_n(0) = (-1)^{n/2} \frac{n!}{(n/2)!}$$

if n is even and $D_n(0) = 0$ if n is odd.

c) Using a), b) and the fact that the Hermite polynomials satisfy the same relations, conclude that

$$D_n(\lambda) = H_n(\lambda) \quad \forall \lambda \in \mathbb{R}, \quad n = 1, 2, \dots$$

d) Prove that under the conditions of Theorem 8.8 of this Chapter,

$$p^E(x) = \varphi(x) \left\{ \sum_{t \in S_0} \widehat{\sigma}_0(t) + \sum_{j=1}^{d_0} \left(\frac{|\rho'_j|}{\pi} \right)^{j/2} \overline{H}_j(x) g_j \right\}.$$

(Hint: Mimic the proof of Theorem 8.8 and take into account c)).

EXERCISE 8.2. Let $\kappa(S)$ be defined in (8.69):

Prove that:

- a) If S is convex, then $\kappa(S) = 0$.
- b) If S is a C^3 -manifold, with or without boundary, then $\kappa(S) < \infty$.
- c) Assume that S verifies the following condition:

For every $t \in S$ there exists an open neighborhood V of t in \mathbb{R}^d and a C^3 diffeomorphism $\psi : V \rightarrow B(0, r)$ (where $B(0, r)$ denotes the open ball in \mathbb{R}^d centered at 0 and having radius r , $r > 0$) such that

$$\psi(V \cap S) = C \cap B(0, r), \quad \text{where } C \text{ is a convex cone.}$$

Then, $\kappa(S) < \infty$.

EXERCISE 8.3. Prove the equivalence (8.81) for one-parameter processes given in Example 1 of Section 5 of this chapter.

EXERCISE 8.4. Prove that formula (8.84) under the conditions of Example 2) of this chapter.

EXERCISE 8.5. Perform the details of the computations appearing in the Examples 3,4,5 of this chapter.

EXERCISE 8.6. Computation of $\mathbb{E}(\det(X''(t)) | X(t) = x, X'(t) = 0)$. Let $X(t)$ $t \in \mathbb{R}^d$ be a Gaussian field with real values and such that

- It has C^2 paths
- $\mathbb{E}(X(t)) = 0$
- $\text{Var}(X(t))$ is constant and non singular. Without loss of generality we suppose that $\text{Var}(X(t)) = I_d$.

a) Prove that $X(t)$ and $X'(t)$ are independent.

b) Prove that $\mathbb{E}(X_{ij} X_{kl})$ is a symmetric function of (i, j, k, l) : (it is invariant by permutation).

c) Admit or prove the following classical result: let Y_1, \dots, Y_n be n centered jointly Gaussian variables, then

- if $n = 2m + 1$; $\mathbb{E}(Y_1 \dots Y_n) = 0$,

- if $n = 2m$;

$$E(Y_1 \cdots Y_n) = \sum E(Y_{i_1} Y_{i_2}) \cdots E(Y_{i_{2m-1}} Y_{i_{2m}})$$

where the sum is over the $\frac{(2m)!}{m!2^m}$ ways of grouping pairwise the $2m$ variables.

Hint: compute moments using the characteristic function or see Adler (1981) pp 108-109.

4 Let Δ be a $n \times n$ centered Gaussian matrix with entries Δ_{ij} . Suppose that

$$E(\Delta_{ij} \Delta_{kl}) = \mathcal{E}(i, j, k, l) - \mathbf{1}_{i=j} \mathbf{1}_{k=l},$$

where \mathcal{E} is a symmetric function of (i, j, k, l) . Prove that

- if n is odd $E(\det(\Delta)) = 0$.
- if $n = 2m$

$$E(\det(\Delta)) = \frac{(-1)^m (2m)!}{m! 2^m}.$$

Hint : develop the determinant using permutations and signature and use the formula above to see that the part corresponding to \mathcal{E} vanishes.

5 Let $D = \Delta - xI_n$ where Δ is as in 4. Prove that

$$E(\det(D)) = (-1)^n \sum_{j=0}^{[n/2]} (-1)^j \binom{n}{2j} \frac{(2j)!}{j! 2^j} x^{n-2j} = (-1)^n \overline{H}_n(x).$$

6 Conclude.

This result, due to Delmas (2001), extends to the non-stationary case, Lemma 11.7.1 of Adler and Taylor (2007).

The record method

This chapter presents a very efficient method for the numerical computation the distribution of the maximum of a stochastic process or a random field with two-dimensional parameter . It is based mainly on a paper by Mercadier (2006) and the Matlab toolbox MAGP (Mercadier 2005), which uses the routine Rind of the Matlab package WAFO (WAFO-group, 2000).

1. Smooth processes with one dimensional parameter

1.1. Main result. The basic idea is the following: let $\{X(t), t \in \mathbb{R}\}$ be a real valued stochastic process with almost surely absolutely continuous sample paths, and suppose that we are looking for an expression of

$$1 - F_M(u) = P\{M > u\}.$$

We denote $M_T = \sup\{X(t) : 0 \leq t \leq T\}$ and $M = M_1$. Instead of looking at all crossings of the level u , we will look only at those crossings that are “record times” . The set \mathcal{R} of “record times” is defined by

$$\mathcal{R} := \{t \in [0, 1] : X(s) < X(t), \forall s \in [0, t)\},$$

with the convention that 0 is always in \mathcal{R} . We have the following trivial identity

$$(9.1) \quad P\{M \geq u\} = P\{X(0) > u\} + P\{\exists t \in \mathcal{R} : X(t) = u\}.$$

The number of “record times” t such that $X(t) = u$ is equal to 0 or 1. The second term in the right hand side of (9.1) is equal to the expectation of

$$\mathcal{R}(u) := N_u(X, \mathcal{R}) = \#\{t \in \mathcal{R} : X(t) = u\}.$$

On this idea is based the following result:

THEOREM 9.1 (Rychlik’s formula). *Let $X = X(t), t \in [0, 1]$ be a real valued stochastic process with almost surely absolutely continuous sample paths such that for almost all $t \in [0, 1]$ $X(t)$ admits a density $p_{X(t)}$ and $E(|X'(t)|) < \infty$. Then for every $u \in \mathbb{R}$*

$$(9.2) \quad P\{M > u\} = P\{X(0) \geq u\} + \lim_{\delta \rightarrow 0} \frac{1}{\delta} \int_u^{u+\delta} dx \int_0^1 E(X'(t)^+ \mathbf{1}_{t \in \mathcal{R}} | X(t) = x) p_{X(t)}(x) dt.$$

Remarks:- A first version of this formula under stronger conditions is due to Rychlik (1990). The present version is due to Mercadier(2006)

- The limit in (9.2) is in fact a manner of choosing a convenient version of the conditional expectation. For Gaussian processes under certain conditions the usual conditional distributions defined by the regression formulas is convenient, see Corollary 9.2.

- The expression (9.2) seems at first sight worthless since its right hand side does not seem to be simpler than the left hand side. But in the next section we will deduce from formula (9.2) some upper bounds that are sharp.

COROLLARY 9.2. *Suppose that in addition to the conditions of Theorem 9.1 the process $X(t)$ is Gaussian with C^1 paths and satisfies*

- For $s, t \in [0, 1], s < t$, the distribution of $(X(s), X(t))$ does not degenerate.
- For $t \in [0, 1]$, the distribution of $(X(t), X'(t))$ does not degenerate.

Then

$$(9.3) \quad P\{M > u\} = P\{X(0) > u\} + \int_0^1 E(X'(t)^+ \mathbf{1}_{t \in \mathcal{R}} | X(t) = u) p_{X(t)}(u) dt.$$

PROOF OF THEOREM 9.1 The main idea is that the number $\mathcal{R}(u)$ of record points taking a particular value u is equal to 0 or 1 and that

$$\mathbb{P}\{M \geq u\} = \mathbb{P}\{X(0) \geq u\} + \mathbb{E}(\mathcal{R}(u)).$$

To compute the expectation $\mathbb{E}(\mathcal{R}(u))$, we use the Banach formula (see Exercise 3.8). Let $g(u)$ be a continuous bounded function, we have

$$\int_{\mathbb{R}} g(u) \mathcal{R}(u) du = \int_0^1 |X'(t)| g(X(t)) \mathbb{1}_{t \in \mathcal{R}} dt.$$

Taking expectations in both sides gives:

$$\int_{\mathbb{R}} g(u) \mathbb{E}[\mathcal{R}(u)] du = \int_{\mathbb{R}} du g(u) \int_0^1 dt \mathbb{E}[|X'(t)| \mathbb{1}_{t \in \mathcal{R}} | X(t) = u] p_{X(t)}(u)$$

showing that the two function of u :

$$\mathbb{E}[\mathcal{R}(u)] \text{ and } \int_0^1 \mathbb{E}[|X'(t)| \mathbb{1}_{t \in \mathcal{R}} | X(t) = u] p_{X(t)}(u) dt$$

are u -almost surely equal. From this we deduce that the two functions $\mathbb{P}\{M > u\}$, and $\mathbb{P}\{X(0) > u\} + \mathbb{E}[\mathcal{R}(u)]$ are u -almost surely equal. The result follows because $\mathbb{P}\{M > u\}$ is càd-làg. \square

The proof of the corollary is left to the reader.

The next proposition is an easy consequence of Rychlik's formula, which is a version under weaker hypotheses of Theorem 8.12 for random fields.

PROPOSITION 9.3. *Let $\{X(t) : 0 \leq t \leq T\}$ be a Gaussian process that satisfies*

- *it is twice differentiable in quadratic mean,*
- *for all $t \in [0, T]$, $\mathbb{E}(X(t)) = 0$, $\text{Var}(X(t)) = 1$,*
- *$\text{Var}(X'(t))$ is bounded away from zero. Without loss of generality we can use the “unit speed transformation” and suppose that $\text{Var}(X'(t)) = 1$,*
- *for all $s \neq t$, $r(s, t) < 1$.*

Then for every $\delta > 0$ there exists some constant C_δ such that

$$(9.4) \quad 0 \leq 1 - \Phi(u) + T \sqrt{\frac{2}{\pi}} \varphi(u) - \mathbb{P}\{M_T > u\} \leq C_\delta \exp \left[- \left(\left(1 + \frac{1}{Z} \right) \frac{u^2(1-\delta)}{2} \right) \right],$$

where

$$Z := \sup_{0 \leq s < t \leq T} \left[\frac{\text{Var}(X(s) | X(t), X'(t))}{(1 - r(s, t))^2} + \frac{(r_{0,1}^+(s, t))^2}{(1 - r(s, t))^2} \right] < +\infty.$$

PROOF. We use a method that has been employed in the context of random fields, in the proof of Theorem 8.12. Clearly the expression in (9.4) is bounded by

$$\int_0^1 \mathbb{E}(X'(t)^+ \mathbb{1}_{t \notin \mathcal{R}} | X(t) = u) p_{X(t)}(u) dt.$$

An application of the Hölder inequality shows that it sufficient to give bounds to

$$\mathbb{P}\{\exists s : s < t, X(s) \geq u | X(t) = u, X'(t) > 0\}.$$

For that purpose we write the regression of $X(s)$ on $(X(t), X'(t))$:

$$X(s) = r(s, t)X(t) + r_{0,1}(s, t)X'(t) + R^t(s).$$

The three terms in the right hand side above are independent. Under the condition $\{X(t) = u, X'(t) > 0\}$ the event $X(s) \geq u$ can be written:

$$\frac{R^t(s)}{1 - r(s, t)} + \frac{r_{0,1}(s, t)}{1 - r(s, t)} X'(t) \geq u.$$

It is obvious that the left hand side in the inequality above is smaller than

$$Y^t(s) := \frac{R^t(s)}{1-r(s,t)} + \frac{r_{0,1}^+(s,t)}{1-r(s,t)}X'(t).$$

Suppose for the moment that $Y^t(s)$ is bounded and that Z is finite, then using the Landau-Shepp-Fernique inequality (2.33), we know that for every $\delta > 0$ there exists some constant C'_δ such that

$$\mathbb{P}\{\exists s : s < t, X(s) \geq u | X(t) = u, X'(t) > 0\} \leq C'_\delta \exp\left[-\frac{u^2(1-\delta)}{2Z}\right].$$

The rest of the proof is plain.

It remains to prove that $Y^t(s)$ is bounded and that its maximal variance is finite. The variance is the sum of the variance of the two terms. The variance can become infinite or $Y^t(s)$ can become infinite, only for s tending to t . Using Taylor's formula at t :

$$X(s) = X(t) + (s-t)X'(t) + \frac{(s-t)^2}{2}Q(s),$$

where $Q(s)$ is an integral remainder. It is easy to see that

$$1-r(s,t) \approx \frac{(s-t)^2}{2}, \quad r_{0,1}(s,t) \approx (s-t).$$

and that $\lim_{s \rightarrow t} R^t(s)$ is just the projection in $L^2(\Omega)$ of $Q(s)$ onto the orthogonal complement of the linear subspace generated by $X(t)$, $X'(t)$, so that we can conclude that it is a.s. finite and has finite variance. \square

1.2. Numerical application. The exact implicit formula (9.3) can be turned into an explicit upper-bound by means of a discretization of the condition $\{X(s) < X(t), \forall s \in [0, t]\}$. One convenient way is to use the points $\{kt/n, k = 0, \dots, n-1\}$ to get

$$(9.5) \quad \mathbb{P}\{M > u\} \leq \mathbb{P}\{X(0) > u\} + \int_0^1 \mathbb{E}(X'(t)^+ \mathbf{1}_{X(0), \dots, X(t(n-1)/n) < u} | X(t) = u) p_{x(t)}(u) dt.$$

On the other hand the time discretization provides the trivial lower-bound

$$(9.6) \quad \mathbb{P}\{M > u\} \geq 1 - \mathbb{P}\{X(0), \dots, X((n-1)/n) \leq u\}.$$

The main point is that, when the process is Gaussian, the integrals that appear in (9.5) and (9.6) can be computed using the MAGP tool-box. All details are given the web-page of Mercadier (2005). The program is able to perform such calculations for n up to 100.

The precision of the computations of MAGP has been evaluated by Mercadier in two ways:

First, comparing the lower bounds (9.6) with the exact theoretical value for the ‘‘sine-cosine’’ process (i.e. the centered Gaussian process with covariance $\Gamma(t) = \cos t$) given by Berman(1971 b) and Delmas (2003 b) (See Exercise 4.1).

Second, comparing lower and upper-bounds with results in other chapters of this book . For example Figure 9.1 compares the lower and upper-bound of this chapter with the lower-bounds and upper-bounds given by two or three terms in the Rice series for the centered stationary process with covariance $\gamma_1(t) = e^{-t^2/2}$ (see Chapter 5).

From these comparisons it appears that one can trust the result from MAGP up to 10^{-3} .

Another question is the precision of the estimation, which can be measured by the difference between (9.5) and (9.6).

We will consider that the estimation given by (9.5) and (9.6) is ‘‘numerically significant’’ if it corresponds to an absolute error smaller than 10^{-2} and to a relative error smaller than 10^{-1} . We will concentrate ourselves on the case of stationary centered Gaussian processes with variance 1 and ‘‘unit speed’’ ($\text{Var}(X'(t)) = 1$). The result depends of course on T and u . The larger u (or the smaller T), the better the results.

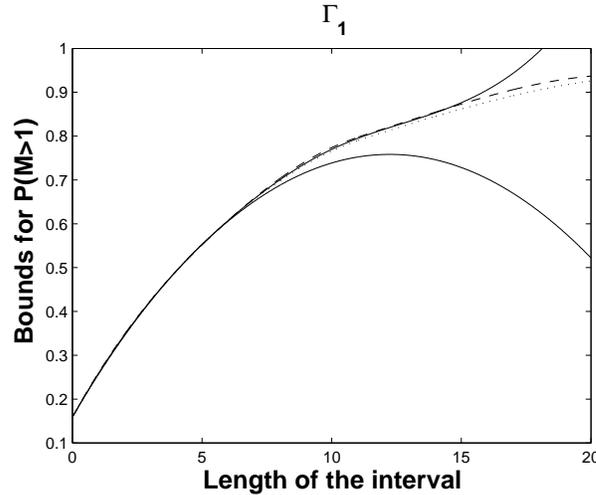


FIGURE 9.1. Comparison of the present bounds with those of Chapter 5. From top to bottom : Upper-bound from Chapter 5, upper-bound (9.5), lower-bound (9.6), lower-bound from Chapter 5 (M stands for M_T).

It happens that for levels $u \geq 1$ the result is numerical significant for times intervals of sizes 20 to 25 in “unit-speed measure”. See Mercadier (2006) for more details.

One can check in Figure 9.1, that the record method implemented by MAGP is, at this stage, the most efficient for numerically computing the distribution of the maximum.

2. Non-smooth Gaussian processes

When the process has non-differentiable paths, one way is to use smoothing, as in Chapter 5. Another way is Durbin's formula (Durbin, 1985), based on the pseudo-derivative defined as the normalized increment $(X(t) - X(s))/(t - s)$. Mercadier has found that this method is very unstable. A better way is to use the time discretization and the lower-bound (9.6).

It remains to give bounds on the discretization error. This will be done for a process defined on $[0, 1]$, discretized at the points $k/n, k = 0, 1, \dots, n$ which has the same irregularity as the Wiener process (Brownian motion), i.e. it satisfies the law of iterated logarithm (LIL), for fixed t , almost surely:

$$(9.7) \quad -1 = \liminf_{s \rightarrow 0} \frac{X(t+s) - X(t)}{\sqrt{2s \log(\log(1/s))}} \text{ and } \limsup_{s \rightarrow 0} \frac{X(t+s) - X(t)}{\sqrt{2s \log(\log(1/s))}} = 1.$$

Generalizations to other local behaviors can be performed using similar tools, mutatis mutandis, that is, changing this oscillation by the one of the processes being considered.

Our method is based on the following heuristic approximations (which may not be actually verified by the paths)

- (1) The instant t^* where the maximum is attained satisfies the \liminf part of the LIL.
- (2) The maximum of the discretized process is attained among the point k/n at the point t_n^* which is the nearest to t^* .
- (3) $|t_n^* - t^*|$ has a uniform distribution among the possible value in $[0, 1/(2n)]$.

With all these approximations we get

$$M - M_n \simeq \sqrt{2Z \log(\log(1/Z))},$$

where Z is uniformly distributed over $[0, 1/2n]$ which amounts to saying that:

$$(9.8) \quad \mathbb{P}\{M > u\} \simeq \int_0^{1/2n} \mathbb{P}\{M_n > u - \sqrt{2z \log(\log(1/z))}\} 2n \, dz$$

Notice that the right hand side of (9.8) is easy to compute numerically using MAGP.

Results are shown in Figure 9.2 for the maximum of processes defined on an interval $[0, T]$ and parametrized by the length T . They refer to the Ornstein-Uhlenbeck process (that is, the centered stationary Gaussian process with covariance $\Gamma(t) := \exp(-t)$, $t \geq 0$). These results are compared with the exact value from Delong's (1981) paper. The figure suggests that the approximation (9.8) is very good.

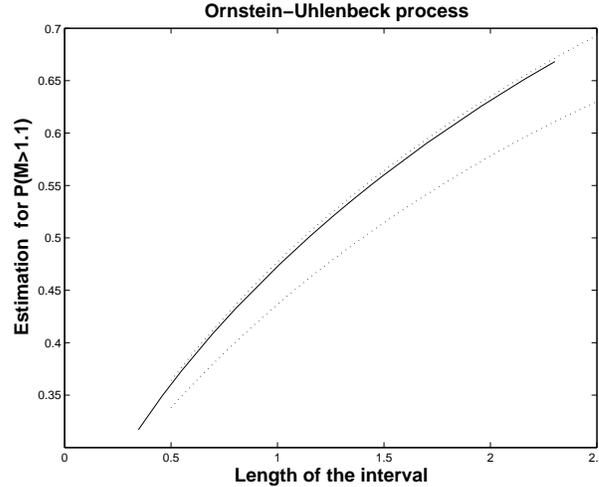


FIGURE 9.2. Ornstein-Uhlenbeck process : Approximation (9.8) (Top) , lower-bound (9.6) (bottom) and exact value (solid line) for $\mathbb{P}\{M_T > 1.1\}$ as a function of T .

3. Two-parameter Gaussian processes

3.1. Main result. This section is based on similar ideas to those of Section 1, adapted to two-parameter processes. We consider a continuous random field $\mathcal{X} = \{X(t) : t \in S\}$ with real values and defined on a compact subset S of \mathbb{R}^2 and $M_S := \max_{t \in S} X(t)$.

Let us consider now a “rather high” level u and a realization such that $\{M_S > u\}$.

Let us suppose that the probability that the process remains above the level u for all $t \in S$ can be neglected. Then, the event $\{M_S > u\}$ is almost equivalent to

“the level curve $\mathcal{C}_u := \{t \in S : X(t) = u\}$ is not empty”.

More precisely, let us choose a particular direction (say South) and to consider the point at the southern extremity of \mathcal{C}_u (which is in general unique). To do so denote by \prec the lexicographic order on \mathbb{R}^2 , that is:

$$s = (s_1, s_2) \prec t = (t_1, t_2) \Leftrightarrow \{s_2 < t_2\} \text{ or } \{s_2 = t_2; s_1 < t_1\}.$$

We define the “lexicographic past” $\mathcal{L}(t)$ of a point $t \in S$ as

$$\mathcal{L}(t) := \{s \in S : s \prec t\}.$$

A point $t \in S$ will be called a “record point” if for all points $s \in \mathcal{L}(t)$: $X(s) < X(t)$. We denote by \mathcal{R} the set of record points. Obviously there is at most one record point where the process $X(t)$ takes a particular value u , and this point is (in general) at the southern extremity of the level curve.

Eventually the event $\{M > u\}$ is almost equivalent to

$$\{\text{The number of record points on } \mathcal{C}_u \text{ is 1 and not 0}\}.$$

We assume the following hypotheses:

- (A0) The set S is compact, convex, the parameterization $\rho : [0, L] \rightarrow \partial S$ of the boundary ∂S by its arc length is of class \mathcal{C}^1 , except perhaps at a finite number of points where ρ is only continuous. Moreover we will assume that $\rho(0)$ is the point of ∂S which is minimal with respect to \leq .
- (A1) The sample paths of the random field $Z := (X, X_{10})$ are almost surely continuously differentiable.
- (A2) For $t \in S$, the distribution of $Z(t)$ does not degenerate.
- (A3) For every $w \in \mathbb{R}^2$ there is almost surely no point $t \in S$ such that $Z(t) = w$ and $\det(Z'(t)) = 0$.

Recall the notation $X_{ij}(t_1, t_2) := \frac{\partial^{i+j}}{\partial t_1^i \partial t_2^j} X(t_1, t_2)$ ($i, j = 0, 1, \dots$).

THEOREM 9.4. *Let S be a subset of \mathbb{R}^2 satisfying (A0). Let $X(t)$ be a real valued Gaussian process defined on some neighborhood of S and satisfying assumptions (A1-A3). Then for every real u :*

$$(9.9) \quad \begin{aligned} \mathbb{P}\{M > u\} &= \mathbb{P}\{Y(0) > u\} + \int_0^L \mathbb{E}(|Y'(\ell)| \mathbf{1}_{X(s) < u, \forall s \in \mathcal{L}(\rho(\ell))} | Y(\ell) = u) p_{Y(\ell)}(u) d\ell \\ &\quad + \int_S \mathbb{E}(|X_{20}(t)^- X_{01}(t)^+| \mathbf{1}_{X(s) < u, \forall s \in \mathcal{L}(t)} | X(t) = u, X_{01}(t) = 0) p_{X(t), X_{01}(t)}(u, 0) dt, \end{aligned}$$

where: $Y(\ell) := X(\rho(\ell))$.

Sufficient conditions for (A3) are given in Proposition 6.5.

PROOF. The proof is very close to that of Theorem 7.2 and will be sketched.

Assume that the event $\{M > u\} \cap \{Y(0) < u\}$ occurs. Then, \mathcal{C}_u is non empty and compact. The point τ which is minimal for \leq on \mathcal{C}_u is uniquely determined. We want to prove that τ is a record point.

- If τ and $\rho(0)$ have the same second coordinate, $\mathcal{L}(\tau)$ is reduced to the segment $I := [\rho(0); \tau)$. The value of the process $X(t)$ in $\rho(0)$ is less than u and by definition, the process $X(t)$ cannot take the value u on I . $X(t)$ cannot take a value larger than u on I because of the Intermediate Value Theorem. As a consequence τ is a record point.
- If τ and $\rho(0)$ have distinct second coordinate then $\rho(0)_2 < \tau_2$. On $\mathcal{L}(\tau)$, $X(t)$ cannot take the value u . Suppose that there exists $\tilde{\tau}$ in $\mathcal{L}(\tau)$ such that $X(\tilde{\tau}) > u$. The whole segment $[\rho(0), \tilde{\tau}]$ is in S and thus in $\mathcal{L}(\tau)$ and by the Intermediate Value Theorem, there is a point on this segment where $X(t)$ takes the value u which is not possible.

Adding the trivial case $\{Y(0) > u\}$, we have proved that, on the event $Y(0) \neq u$, which has probability 1, almost surely $\{M > u\}$ is the disjoint union of the two events

“ $Y(0) > u$ ” and “there exist exist a record point with value u ”.

The event “there exist exist a record point with value u ” can be split into two disjoint ones depending on whether τ belongs to ∂S or \dot{S} . Since there is at most one record point, these two cases are disjoint and their probabilities are equal to the expectation of the number of record points in ∂S and \dot{S} .

Let us consider a non-increasing function $\mathcal{F}, \mathbb{R} \rightarrow \mathbb{R}$ satisfying:

$$\mathcal{F}(x) = 1 \text{ if } x < -1/2, \quad \mathcal{F}(x) = 0 \text{ if } x \geq 0.$$

Then

$$\mathcal{F}_n(x) := \mathcal{F}(nx) \uparrow \mathbf{1}_{x < 0} \text{ as } n \uparrow +\infty.$$

We compute first the expectation of number of zeroes of $Z(t)$ on S with the weights $\mathcal{F}_n(\sup_{s \in \mathcal{L}(t)} X(s) - X(t))$ using the Rice formula (Theorem 6.4) and then we pass to the monotone limit as $n \rightarrow \infty$.

For the boundary of S

$$\mathbb{E}(\#\{\ell \in [0, L] : \rho(\ell) \in \mathcal{R}, Y(\ell) = u\})$$

we have to use the same kind of proof after splitting $(0, L)$ into a finite number of sub-intervals in which $\ell \rightsquigarrow Y(\ell)$ is \mathcal{C}^1 . We remark that almost surely $Y(\cdot)$ does not takes the value u at the extremities of these intervals. Summing up, we get the result. \square

THEOREM 9.5 (Bounds). *Let S be a subset of \mathbb{R}^2 satisfying:*

(A'0) : *S is compact, S and its complement are connected, the parameterization $\rho : [0, L] \rightarrow \partial S$ of the boundary ∂S by its arc length is of class \mathcal{C}^1 except perhaps at a finite number of points where ρ is only continuous.*

Let \mathcal{X} be a real valued Gaussian process defined on some neighborhood of S , satisfying assumptions (A1) and (A2). Then, for every real u , using the notations of Theorem 9.9:

$$(9.10) \quad \mathbb{P}\{M > u\} \leq \mathbb{P}\{Y(0) > u\} + \int_0^L \mathbb{E}(|Y'(\ell)| | Y(\ell) = u) p_{Y(\ell)}(u) d\ell \\ + \int_S \mathbb{E}(|X_{20}(t)^- X_{01}(t)^+| | X(t) = u, X_{01}(t) = 0) p_{X(t), X_{01}(t)}(u, 0) dt,$$

PROOF. Let M_∂ be the maximum of $X(t)$ on ∂S . One has:

$$(9.11) \quad \mathbb{P}\{M > u\} = \mathbb{P}\{M_\partial > u\} + \mathbb{P}(\{M > u\} \cap \{M_\partial < u\}) \\ \leq \mathbb{P}\{Y(0) > u\} + \mathbb{P}\{U_u^Y([0, L] > 0\} + \mathbb{P}\{\exists t \in \dot{S} : X(t) = u, X_{10}(t) = 0, X_{20} < 0, X_{01} > 0\}.$$

The last inequality is due to the fact that if $M > u$ and $M_\partial < u$, the level curve \mathcal{C}_u is contained in the interior of S . There exists at least one point on this curve with minimal second coordinate. It follows that:

$$\mathbb{P}\{M > u\} \leq \mathbb{P}\{Y(0) > u\} + \mathbb{E}\{U_u^Y([0, L])\} + \mathbb{E}(\#\{t \in \dot{S} : X(t) = u, X_{10}(t) = 0, X_{20} < 0, X_{01} > 0\})$$

It suffices to apply the Rice formula (Theorem 6.2), remarking that under condition $X(t) = u, X_{10}(t) = 0$ we have : $\det(Z'(t)) = X_{20}(t) \cdot X_{01}(t)$. \square

3.2. Numerical application. As in Section 1.2 the exact formula (9.9) can be transformed into an explicit upper-bound by discretizing the condition $\mathbf{1}_{X(s) < u, \forall s \in \mathcal{L}(t)}$. For simplicity we limit now our attention to the case where S is the square $[0, T]^2$ and the process X is “standardized” i.e. centered, stationary, isotropic, with variance one and “identity speed”. In that case

- The two terms in (9.9) corresponding to the edges $0 \times [0, T]$ and $1 \times [0, T]$ are equal and equal to

$$\int_0^T \mathbb{E}((X_{10}(v, 0))^+ \mathbf{1}_{X(s) < u, \forall s \in \mathcal{L}(v, 0)} | X(v, 0) = u) \phi(u) dv.$$

- The term corresponding to the edge $[0, T] \times 1$ vanishes: indeed if there is a record point τ on this edge, the derivative X_{01} must vanish. Because of Rice formula (or using Bulinskaya’s Lemma 1.20) the expectation of the number of such points is zero.

If in formula (9.9), we replace the whole lexicographic past $\mathcal{L}(t)$ of a point t by its intersection $\mathcal{L}_n(t)$ with a grid :

$$\mathcal{L}_n(t) := \mathcal{L}(t) \cap \left\{ \left(\frac{kt}{n}, \frac{lt}{n} \right), k = 0, \dots, n, l = 0, \dots, n \right\}.$$

we get the upper-bound:

$$(9.12) \quad \begin{aligned} \mathbb{P}\{M > u\} \leq & \bar{\Phi}(u) + \varphi(u) \left[2 \int_0^T \mathbb{E}(X_{01}^+(v, 0) \mathbf{1}_{X(s) < u, \forall s \in \mathcal{L}_n(v, 0)} | X(v, 0) = u) dv \right. \\ & + \int_0^T \mathbb{E}(X_{10}^+(v, 0) \mathbf{1}_{X(s) < u, \forall s \in \mathcal{L}_n(v, 0)} | X(v, 0) = u) dv \\ & \left. + \frac{1}{\sqrt{2\pi}} \iint_{[0, T]^2} \mathbb{E}(X_{20}(t)^- X_{01}(t)^+ \mathbf{1}_{X(s) < u, \forall s \in \mathcal{L}_n(t)} | X(t) = u, X_{01}(t) = 0) dt \right]. \end{aligned}$$

The main point is that this upper-bound can be computed by MAGP.

In the other direction, we get a lower-bound using discretization :

$$(9.13) \quad \mathbb{P}\{M > u\} \geq \mathbb{P}\left\{ \max \left(X\left(\frac{kt}{n}, \frac{lt}{n}\right), k = 0, \dots, n, l = 0, \dots, n \right) > u \right\}.$$

Figure 9.3 shows these bounds compared with equivalent in Adler (1981) and the equivalent given by the Euler-Poincaré Characteristic method.

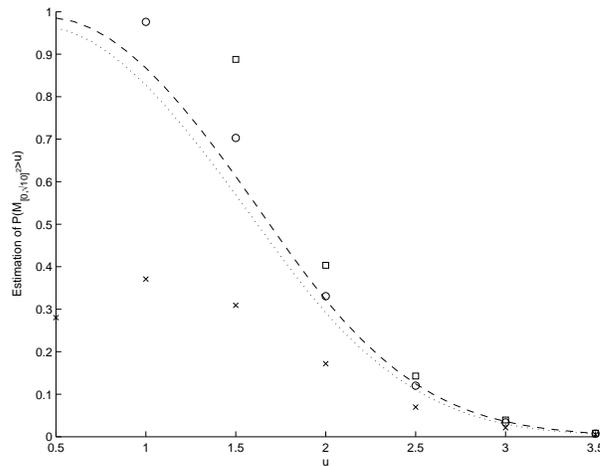


FIGURE 9.3. $\mathbb{P}\{M > u\}$ for the random field with Gaussian covariance on $[0, \sqrt{10}]^2$ from top to bottom : the upper bound (9.10), the Euler characteristic approximation, the upper-bound (9.12) , the lower-bound (9.12) , the equivalent by Adler(1981).

Exercises

EXERCISE 9.1. Prove Corollary 9.2 using Ylvisaker’s Theorem and the regression method of Proposition 9.3

EXERCISE 9.2. Suppose that S is the square $[0, T]^2$. Suppose that the process $X(t)$ is “standardized” then show that the upper bound (9.10) takes the form

$$\mathbb{P}\{M > u\} \leq \bar{\Phi}(u) + \sqrt{\frac{2}{\pi}} T \phi(u) + \frac{T^2 u}{(2\pi)} \left[c\phi(u/c) + u\Phi(u/c) \right] \phi(u),$$

with $c := \sqrt{\text{Var}(X_{20}) - 1}$

Show that the difference between this bound and the equivalent given by EPC method :

$\bar{\Phi}(u) + \sqrt{\frac{2}{\pi}} \phi(u) + \frac{T^2 u}{(2\pi)} \phi(u)$ is bounded by

$$\frac{T^2 u}{((2\pi)^{3/2})} c^3 u^{-2} \phi\left(u \sqrt{\frac{1+c^2}{c^2}}\right) \left[c\phi(u/c) + u\Phi(u/c) \right] \phi(u).$$

Asymptotic methods for infinite time horizon

A contrario to Chapter 8 this chapter considers asymptotic results for intervals of time with size tending to infinity. Let us consider the crossings of the level u by a stationary Gaussian process. In Section 1 the level u tends to infinity jointly with the size so that the expectation of the number of crossings remains constant. In that case it is proven that, under weak hypotheses, the asymptotic distribution of the number of crossings is Poisson. This implies that the maximum of the process converges, after renormalization, to a Gumbel distribution.

In Section 2 the level u is fixed and the number of crossings tends to infinity. Under some conditions, this number of crossings satisfies a central limit theorem. The main tool is in fact an elementary presentation of Wiener Chaos decomposition .

1. Poisson character of “high” up-crossings

In this section we give a proof of the following theorem, originally due to Volkonskii and Rozanov (1959,1961).

THEOREM 10.1. *Let $\mathcal{X} = \{X(t) : t \in \mathbb{R}\}$ be a zero-mean stationary Gaussian process with covariance $\Gamma(\tau) = E(X(t)X(t + \tau))$ satisfying the following conditions:*

- $\Gamma(0) = 1$.
- $\lambda_2 < \infty$
- $\Gamma(\tau) \log(\tau) \rightarrow 0$ as $\tau \rightarrow \infty$ (Berman's condition)
- If one writes $\Gamma(\tau) = 1 - \lambda_2 \tau^2/2 + \theta(\tau)$ then, for some $\delta > 0$ the integral

$$\int_0^\delta \frac{\theta'(\tau)}{\tau^2} d\tau$$

is convergent (this is Geman's condition, that we have already mentioned in Proposition 4.2).

Set

$$C_u := E(U_u(X; [0, 1])) = \sqrt{\lambda_2} \frac{\exp(-u^2/2)}{2\pi}$$

and define, for $t \in \mathbb{R}^+$

$$(10.1) \quad R_u(t) := U_u(X; [0, C_u^{-1}t])$$

Then, as $u \rightarrow +\infty$ the family of point processes $\{R_u(t) : t \geq 0\}$ converges weakly in the Skorohod space to a standard Poisson process.

Remark. Notice that Berman's condition implies that $|\Gamma(\tau)| < 1$ for all $\tau \neq 0$ (See Feller, 1966, Ch. XV).

This theorem has a direct interest for modelling phenomena depending on time, like pollution levels, floods or other situations in which up-crossings of a threshold by a certain stochastic process imply the occurrence of a relevant event. This theorem is the mathematical explanation of a standard procedure which consists in using the Poisson process as a model to represent the sequence of these random time points, whenever the threshold and the size of the time window are

large enough. Of course, beyond the statement of the theorem, the fitting of such a model to empirical data is a problem of statistical nature which should be appropriately considered in each case.

In the formulation above, the process is Gaussian and stationary, has some local regularity given by Geman's condition and some mixing (asymptotic independence for distant values of the parameter) given by Berman's condition. In fact, the mild "Geman condition" is not needed in its full generality (see Leadbetter et al. (1983) for a longer proof without this condition and also for various extensions, including to non-Gaussian processes).

We are using here only some elementary well-known properties of point processes defined on the half-line $[0, +\infty)$ (see for example Neveu, 1977). We will indistinctly call "point process" the random set of points $\Psi = \{t_k\}$, the random measure μ_Ψ having a unit atom at each one of these points, which we will assume to be almost surely locally finite and the càd-làg version of its cumulative distribution function, that is, for $t \geq 0$, $F_\Psi(t) = \#\{k : t_k \leq t\}$. Weak convergence of these point processes is to be understood as weak convergence of the stochastic process $F_\Psi(\cdot)$ in the Skorohod space, see Section 4.1 of Chapter 4. The two properties that we will use without proving them are the following:

PROPOSITION 10.2 (Rényi, 1967).

(a) Let $T > 0$. The family of point processes $\{F_{\Psi_n}(\cdot)\}_{n=1,2,\dots}$ is tight in the space $D[0, T]$ if and only if the sequence of distributions of the (integer-valued) random variables $\{F_{\Psi_n}(T)\}_{n=1,2,\dots}$ is tight on the line.

(b) Assume that a point process as above verifies the following conditions, for any subset B of $[0, +\infty)$ which is a union of intervals:

$$(10.2) \quad \mathbb{E}(\mu_\Psi(B)) \leq \lambda(B)$$

and

$$(10.3) \quad \mathbb{P}\{\mu_\Psi(B) = 0\} = \exp(-\lambda(B)).$$

Then, Ψ is a standard Poisson process.

The next useful corollary of Theorem 10.1 states that, after re-normalization, the maximum M_T of the process converges weakly to a Gumbel distribution:

COROLLARY 10.3. Under the conditions of Theorem 10.1, setting

$$M_T := \sup_{t \in [0, T]} X_t, \quad a_T := (2 \log T)^{1/2}$$

$$b_T := (2 \log T)^{1/2} + \log \left(\frac{\lambda_2}{2\pi} \right) (2 \log T)^{-1/2}$$

then as $T \rightarrow +\infty$

$$\mathbb{P}\{a_T(M_T - b_T) \leq x\} \rightarrow \exp(-e^{-x}).$$

Both the theorem and its corollary can be extended, under mild conditions, to constant variance non-stationary processes (see Azaïs and Mercadier, 2003).

PROOF OF THEOREM 10.1 : Without loss of generality we may assume that $\lambda_2 = 1$. Set $\rho(t) := \sup_{s>t} |\Gamma(s)|$. If it is not otherwise specified, all limits in this proof are for $u \rightarrow +\infty$. Let T_0 be such that $\rho(T_0) < 1/3$.

We break the proof into several steps.

Step 1: Let $T(u)$ be increasing as a function of u , and tend to infinity in a controlled manner, meaning that $T(u) = O((C_u)^{-1/2})$. Let us prove that

$$(10.4) \quad \mathbb{P}\{M_{T(u)} > u\} = \sqrt{\frac{\lambda_2}{2\pi}} T(u) \phi(u) (1 + o(1)).$$

Let u be large enough so that $T(u) > T_0$. We use the computation in the proof of Proposition 4.2 of Chapter 4.

For short, denote ν_2 the second factorial moment of the up-crossings on an interval of length $T(u)$. It suffices to prove that

$$\nu_2 = T(u) \phi(u) o(1)$$

For given $\varepsilon > 0$, choose $\delta > 0$ so that

$$\int_0^\delta \frac{\theta'(\tau)}{\tau^2} d\tau < \varepsilon.$$

Then,

$$\nu_2 \leq T(u) \phi(u) [\varepsilon + T_0 \phi(\delta_1 u) + (T(u) - T_0) \phi^{1/2+\delta_2}(u)]$$

where $\delta_1, \delta_2 > 0$, which implies (10.4).

Step 2: It is sufficient to prove the result for $t \in [0, T]$ and each $T \in \mathbb{R}^+$. We use Proposition 10.2 (a) and observe that by construction $\mathbb{E}(R_u(T)) = T$ which obviously implies tightness of the probability distributions of the random variables $\{R_u(T)\}_{u>0}$. So, the family of processes $R_u(\cdot)$ is tight and if we denote by $R(\cdot)$ a limit process of the family, what remains is to show that $R(\cdot)$ is a standard Poisson process. For this goal, we use the characterization of Poisson processes given by Proposition 10.2 (b).

Clearly Fatou's lemma implies that the process $R(\cdot)$ satisfies (10.2). The remainder of the proof is based upon a discretization argument to verify condition (10.3). We do this first when B is a unique interval. Due to the stationarity of the process, we may assume that $B = (0, L]$.

Set $J_u = (0, C_u^{-1}L]$. Consider the partition of J_u into $n = n(u)$ intervals I_1, \dots, I_n of equal length $C_u^{-1}L/n$, where $n-1$ is the integer part of $C_u^{-1/2}L$. Then each interval I_i ($i = 1, \dots, n$) satisfies the result of step 1.

We now divide each interval I_i into intervals of length $q = q(u)$ such that $q \cdot u \rightarrow 0$, but sufficiently slowly in a way that will be precised later on. In Step 3 we will prove the following intermediate results:

$$(10.5) \quad \mathbb{P}\{U_u(J_u) = 0\} = \mathbb{P}\{M_{J_u} < u\} + o(1).$$

- For all intervals $A = A_u \subset J_u$ having the form $A_u = [aq, bq]$, a, b integers,

$$(10.6) \quad 0 \leq \mathbb{P}\{X(kq) \leq u : \forall k; kq \in A_u\} - \mathbb{P}\{M_{A_u} < u\} = \lambda(A_u) C_u o(1),$$

with $M_{A_u} := \sup_{t \in A_u} X(t)$, where the $o(1)$ is uniform over all sequences of intervals with length bounded below by some positive number.

$$(10.7) \quad \mathbb{P}\{X(kq) \leq u : \forall k; kq \in D_u\} = \prod_{i=1}^n \mathbb{P}\{X(kq) \leq u : \forall k; kq \in I_i\} + o(1)$$

A direct consequence of these equivalences is that

$$(10.8) \quad \begin{aligned} \mathbb{P}\{U_u(J_u) = 0\} &= \prod_{i=1}^n \left(1 - C_u^{1/2}(1 + o(1))\right) = \left(1 - C_u^{1/2}(1 + o(1))\right)^{C_u^{-1/2}L(1+o(1))} v + o(1) \\ &= \exp(-L)(1 + o(1)). \end{aligned}$$

Use now that $R([0, L])$ is the weak limit of $U_u(J_u)$ through a sub-sequence of u 's tending to $+\infty$ to conclude that (10.3) holds true when B is a single interval.

Step 3

We prove (10.5).

Clearly, $\mathbb{P}\{M_{J_u} < u\} \leq \mathbb{P}\{U_u(J_u) = 0\}$. On the other hand:

$$\mathbb{P}\{U_u(J_u) = 0\} = \mathbb{P}\{U_u(J_u) = 0; X(0) < u\} + \mathbb{P}\{U_u(J_u) = 0; X(0) \geq u\} \leq \mathbb{P}\{M_{J_u} < u\} + 1 - \Phi(u).$$

Consider now (10.6). We define $U_u^q(A_u)$ as the number of up-crossings of the size- q discretization of X on the interval A_u . More precisely

$$U_u^q(A_u) := \#\{k \in \mathbf{Z} : kq \in A_u, X(kq) > u, X((k-1)q) < u\}$$

Then

$$0 \leq \mathbb{P}\{X(kq) \leq u : \forall k; kq \in A_u\} - \mathbb{P}\{M_{A_u} < u\} \leq \mathbb{P}\{U_u(A_u) - U_u^q(A_u) > 0\} \leq \mathbb{E}(U_u(A_u) - U_u^q(A_u)).$$

An upper-bound for the right-hand side is given in the following auxiliary lemma, of which we give a proof after finishing the one of the theorem.

LEMMA 10.4. *Let A_u, q and $U_u^q(A_u)$ be defined as above. Then,*

$$\mathbb{E}[U_u(A_u) - U_u^q(A_u)] = o(\lambda(A_u)C_u),$$

with $o(\lambda(A_u)C_u)$ uniform over the considered class of intervals.

We now prove (10.7). We use the Li and Shao's Normal Comparison Lemma, lemma 2.1. Let Σ be the variance matrix of $X(kq)$ for kq in J_u and Σ' the variance matrix obtained by setting to zero the extra-diagonal blocks of Σ with respect to the partition I_1, \dots, I_n .

Using the Li-Shao inequality in both senses, we obtain:

$$(10.9) \quad \left| \mathbb{P}\{X(kq) \leq u : \forall k; kq \in J_u\} - \prod_{i=1}^n \mathbb{P}\{X(kq) \leq u : \forall k; kq \in I_i\} \right| \leq \frac{1}{4} \sum_{i,j=1,\dots,N, i < j} |\Sigma_{ij}| \exp\left(-\frac{u^2}{1 + |\Sigma_{ij}|}\right),$$

where N is the number of discretization points in J_u . Using stationarity and remarking that the function $r \rightsquigarrow r \exp\left(-\frac{u^2}{1+r}\right)$ is increasing on $(0, 1)$ we obtain that the right-hand side in formula (10.9) is bounded above by

$$(10.10) \quad \frac{1}{4} \sum_{l=1, N} k(l) \rho(lq) \exp\left(-\frac{u^2}{1 + \rho(lq)}\right),$$

where $k(l)$ is the difference of the number of occurrences of the quantity $\Gamma(lq)$ between Σ and Σ' . It is easy to see that

$$(10.11) \quad k(l) = l(n-1) \quad \text{for } lq < C_u^{-1}L/n$$

$$(10.12) \quad k(l) \leq C_u^{-1}L/q \quad \text{for every } q.$$

So, using T_0 already introduced in Step 1, and the monotonicity of $r \rightsquigarrow r \exp\left(-\frac{u^2}{1+r}\right)$ we can bound the terms in the sum (10.10) in the following way:

for $0 < l < T_0/q$ we use (10.11) and $\rho(lq) \leq 1$

for $T_0/q \leq l < \frac{C_u^{-1}L}{nq}$ we use (10.11) and $\rho(lq) \leq \rho(T_0)$

for $\frac{C_u^{-1}L}{nq} \leq l$ we use (10.12) and the fact that $t \rightsquigarrow \frac{C_u^{-1}L}{nq}$ is non decreasing,

to conclude that the expression in formula (10.10) is bounded by

$$(10.13) \quad (\text{const}) \sum_{0 < l < T_0/q} l.n \exp\left(-\frac{u^2}{2}\right) + (\text{const}) \sum_{T_0/q \leq l < \frac{C_u^{-1}L}{nq}} l.n \exp\left(-\frac{u^2}{1 + \rho(T_0)}\right) \\ + (\text{const}) \frac{C_u^{-1}}{q^2} \int_{C_u^{-1}L/n}^{C_u^{-1}L} \rho(t) \exp\left(-\frac{u^2}{1 + \rho(t)}\right) dt = I_1 + I_2 + I_3,$$

It is easy to see that

$$I_1 \leq (\text{const}) C_u^{-1/2} q^{-2} \exp(-u^2/2) \\ I_2 \leq (\text{const}) \left(\frac{C_u^{-1}L}{nq}\right)^2 n \exp\left(-\frac{u^2}{1 + \rho(T_0)}\right) \leq q^{-2} (\text{const}) \exp\left(\frac{3u^2}{4}\right) \exp\left(-\frac{u^2}{1 + \rho(T_0)}\right)$$

and that these quantities tend to zero as soon as q does not go to zero faster than some power of u .

As for I_3 ,

$$I_3 \leq (\text{const}) C_u q^{-2} \int_{C_u^{-1}L/n}^{C_u^{-1}L} \rho(t) \exp(u^2 \rho(t)) dt \\ \leq (\text{const}) (qu)^{-2} \int_0^L u^2 \rho(t C_u^{-1}) \exp(u^2 \rho(t C_u^{-1})) dt,$$

after a change of variables. Since $\rho(z) \log(z) \rightarrow 0$, $\rho(z) \leq (\text{const}) (\log(z))^{-1}$ and $u^2 \rho(t C_u^{-1})$ is bounded and converges pointwise to zero, the dominated convergence theorem implies that

$$(10.14) \quad \int_0^L u^2 \rho(t C_u^{-1}) \exp(u^2 \rho(t C_u^{-1})) dt \rightarrow 0.$$

Now it suffices to impose that $(qu)^{-2}$ grows to infinity slower than (10.14) and than some power of u to prove that I_1, I_2 and I_3 tend to zero.

Step 4: Let $D = \cup_{i=1,p} D_i = \cup_{i=1,p} (a_i, b_i]$ be a union of disjoint intervals. We can apply the same arguments as above, discretization and normal comparison lemma, to show that

$$\mathbb{P}\{M_D < u\} = \prod_{i=1,p} \mathbb{P}\{M_{D_i} < u\},$$

that gives the result. The proof is simpler in the sense that terms like I_1 and I_2 in formula 10.13 are not present. The only modification is to use a monotone convergence argument in order to replace the extremes of the intervals: $(C_u^{-1}a_i, C_u^{-1}b_i]$ by multiples of q . \square

PROOF OF LEMMA 10.4: The stationarity of the process implies that

$$\mathbb{E}(U_u(A)) = \lambda(A) C_u$$

and since A is supposed to be of the form $(kq, hq]$ $k, h \in \mathbb{N}$

$$\mathbb{E}(U_u^q(A)) = \left(\frac{\lambda(A)}{q}\right) \mathbb{P}\{X(0) < u < X(q)\}.$$

Remark that

$$J_q(u) := \mathbb{P}\{X(0) < u < X(q)\} = \mathbb{P}\{|Y_1 - u| < \frac{q}{2} Y_2\},$$

with $Y_1 := \frac{X(0)+X(q)}{2}$, $Y_2 := \frac{X(q)-X(0)}{q}$ are two independent Gaussian variables with respective variances: $\sigma_1^2 := \frac{1+r(q)}{2}$, $\sigma_2^2 := \frac{2(1-r(q))}{q^2}$.

We want to prove that $J_q(u) \approx qC_u$. With this goal we compute

$$\begin{aligned}
 (10.15) \quad (C_u q)^{-1} J_q(u) &= (C_u q)^{-1} \int_0^{+\infty} dy_2 \frac{\varphi(y_2/\sigma_2)}{\sigma_2} \int_{u-(qy_2)/2}^{u+(qy_2)/2} dy_1 \frac{\varphi(y_1/\sigma_1)}{\sigma_1} \\
 &= \int_0^{+\infty} \frac{y_2}{\sigma_2^2} \exp\left(-\frac{y_2^2}{2\sigma_2^2}\right) \left[\frac{\sigma_2}{2\sigma_1\sqrt{\lambda_2}} \int_{-1}^1 \exp\left(-\frac{u^2}{2\sigma_1^2}(1-\sigma_1^2) - \frac{uqsy_2}{2\sigma_1^2} - \frac{q^2s^2y_2^2}{8\sigma_1^2}\right) ds \right] dy_2.
 \end{aligned}$$

since $1 - \sigma_1^2 \simeq \frac{\lambda_2 q^2}{4}$, $\sigma_1 \simeq 1$, $\sigma_2^2 \simeq \lambda_2$ we see that pointwise in s and y_2

$$-\frac{u^2}{2\sigma_1^2}(1-\sigma_1^2) - \frac{uqsy_2}{2\sigma_1^2} + \frac{q^2y_2^2}{8\sigma_1^2} \rightarrow 0.$$

On the other hand, the integrand in the last term of (10.15) is bounded by

$$(10.16) \quad (const)y_2 \exp\left(-\frac{y_2^2}{2\sigma_2^2}\right) \exp\left(-\frac{u^2}{2\sigma_1^2}(1-\sigma_1^2) + \frac{uqy_2}{2\sigma_1^2}\right)$$

For

$$y_2 > \frac{2uq\sigma_2^2}{\sigma_1^2}$$

the exponent in formula (10.16) is bounded by $-\frac{y_2^2}{4\sigma_2^2}$ so that the integral for y_2 in $[\frac{uq\sigma_2^2}{\sigma_1^2}, +\infty]$ tends, by the dominated convergence theorem, to

$$\int_0^{+\infty} \frac{y_2}{\lambda_2} \exp\left(-\frac{y_2^2}{2\lambda_2}\right) dy_2$$

The remaining integral can be bounded by

$$(const) \int_0^{\frac{2uq\sigma_2^2}{\sigma_1^2}} y_2 \exp\frac{uqy_2}{2\sigma_1^2} dy_2 \leq (const) \int_0^{\frac{2uq\sigma_2^2}{\sigma_1^2}} y_2 \exp\frac{u^2q^2\sigma_2^2}{\sigma_1^4} dy_2.$$

Since $uq \rightarrow 0$, we see that this integral tends to zero. □

PROOF OF COROLLARY 10.3: Set

$$\tau = \exp(-x), \quad u^2 = 2(\log T + x + \log\left(\frac{\sqrt{\lambda_2}}{2\pi}\right)).$$

We have $TC_u = \tau$. By Theorem 10.1

$$P\{M_T < u\} \approx P\{R_u(\tau) = 0\} \approx \exp(-\tau).$$

Remarking that

$$u = \frac{x}{a_t} + b_t + o(a_t^{-1})$$

we get the result. □

1.1. Extensions to random fields. There exist a series of extensions of the Volkonskii-Rozanov Theorem (Theorem 10.1). We consider here only two of them and do not give proofs. Both refer to real-valued d -parameter random fields, with $d > 1$.

The first one consists in studying, instead of the number of up-crossings of a high level, as was done above, the geometric measure of the inverse image of a high level, that is, to replace the 0-dimensional measure by the $(d - 1)$ -geometric measure, under an adequate normalization. This has been done in Wschebor (1986) using Rice formulas for random fields and is the subject of Exercise 10.2.

The second is based upon the remark that whenever some mixing-like condition is present, one can expect that the point process of local maxima above a high level have a Poisson behavior under a similar renormalization to the one of the Volkonskii-Rozanov Theorem, adapted to the multiparameter case.

We restrict ourselves to quote here Theorem 14.1 of Piterbarg's book (1996), which implies a consequence which is near to Corollary 10.3 above:

THEOREM 10.5 (Piterbarg). *Suppose that the real-valued Gaussian, centered stationary random field $\{X(t) : t \in \mathbb{R}^d\}$ satisfies the following conditions:*

- the covariance $\Gamma(t) = E(X(s)X(s+t))$ verifies $\Gamma(0) = 1$ and $\Gamma(t) \rightarrow 0$ as $\|t\| \rightarrow \infty$,
- the process is three times differentiable, in the mean square sense,
- there exist $C > 0$, $\alpha > 1$ and $\delta > 0$ such that:

$$\lambda_d(t \in [0, T]^d; \Gamma(t) \log^\alpha(\|t\|) > C) = O(T^{d(1-\delta)}) \quad T \rightarrow +\infty.$$

Then if $M_T = \max_{t \in [0, T]^d} X(t)$

$$P\{((M_T - l_T)l_T) < x\} \rightarrow \exp(-\exp(-x)) \quad T \rightarrow \infty,$$

where l_T is the largest solution in l of the equation

$$T^d \sqrt{\det(\Lambda)} l^{d-1} \exp(-l^2/2) = (2\pi)^{(d+1)/2}.$$

and $\Lambda = \text{Var}(X'(0))$.

2. Central limit theorem for non-linear functionals

2.1. Ergodic processes. Let $\mathcal{Y} = \{Y(t) : t \in \mathbb{R}\}$, be a real-valued stochastic process defined on some probability space (Ω, \mathcal{A}, P) . The process is said to be strictly stationary if for any choice of the positive integer k and $t_1, \dots, t_k, t \in \mathbb{R}$, the joint distribution (in \mathbb{R}^k) of $Y(t_1+t), \dots, Y(t_k+t)$ does not depend on t . Clearly, if the process is Gaussian, it is strictly stationary if and only if, for any choice of $\tau \in \mathbb{R}$, the expectation and covariance $E(Y(t))$ and $\text{Cov}(Y(t), Y(t+\tau))$ do not depend on t .

We will assume some mild regularity of the paths of the process \mathcal{Y} such as that almost surely, they are Riemann-integrable on every bounded interval. For example, if the paths are a.s. càd-làg, this follows easily. In fact, this condition can be replaced by some more general measurability condition without affecting what follows. We denote $\sigma(\mathcal{Y})$ the smallest σ -algebra with respect to which all the functions $Y(t) : \Omega \rightarrow \mathbb{R}, t \in \mathbb{R}$ are measurable. Clearly, $\sigma(\mathcal{Y}) \subset \mathcal{A}$.

Let \mathcal{Y} be a real-valued strictly stationary process and η a random variable defined on (Ω, \mathcal{A}, P) , which is also $\sigma(\mathcal{Y})$ -measurable. For $t \in \mathbb{R}$ one can define the random variable $\theta_t(\eta)$ which is the image of η under the translation of size t , in the following natural way: if η has the form

$$(10.17) \quad \eta = g(Y(t_1), \dots, (t_k)),$$

where $g : \mathbb{R}^k \rightarrow \mathbb{R}$ is Borel-measurable, we define

$$\theta_t(\eta) = g(Y(t_1+t), \dots, (t_k+t))$$

For general η , we approximate it in probability by means of cylindrical functions having the form (10.17) and commute the limit with the translation, using the strict stationarity of the process. A $\sigma(\mathcal{Y})$ -measurable random variable η is "invariant" if for every $t \in \mathbb{R}$, almost surely $\theta_t(\eta) = \eta$. The stochastic process \mathcal{Y} is called "ergodic" when η is invariant if and only if it is almost surely constant. A famous theorem due to Maruyama (1949) states that if \mathcal{Y} is a stationary Gaussian process, it is ergodic, if and only if its spectral measure has no atoms.

Assume now that \mathcal{Y} is a strictly stationary stochastic process and $E(|Y(t)|) < +\infty$ (notice that this expectation does not depend on t). Then, the classical *Birkhoff-Khintchine ergodic theorem* says that, almost surely, as $T \rightarrow +\infty$ the time average,

$$\frac{1}{T} \int_0^T Y(t) dt$$

converges to an invariant random variable with finite expectation, I_∞ with $E(I_\infty) = E(Y(0))$. If the process is also ergodic, then this random variable is almost surely constant and equal to

$E(Y(0))$. This corresponds to the usual statement that for strictly stationary ergodic processes, “one can replace time averages by space averages”.

A similar result holds true for the time average

$$\frac{1}{T} \int_{-T}^0 Y(t) dt.$$

For proofs, the reader can consult, for example, again Cramér and Leadbetter’s book, or Brown (1976).

2.2. Non-linear functionals. Let us now turn to the main subject of this section. Let $\mathcal{X} = \{X(t) : t \in \mathbb{R}\}$ be a centered real-valued stationary Gaussian process. Without loss of generality, we assume that $\text{Var}(X(t)) = 1 \forall t \in \mathbb{R}$. We want to consider functionals having the form:

$$(10.18) \quad T_t := 1/t \int_0^t F(X(s)) ds,$$

where F is some function in $L^1(\phi(x)dx)$.

Set $\mu := E(F(Z))$, Z being a standard normal variable. The ergodic theorem implies that a.s. the expression in (10.18) has an invariant limit as $t \rightarrow +\infty$, which is also $\sigma(\mathcal{X})$ -measurable. So, if the spectral measure of the process has no atoms, because of Maruyama’s Theorem, this limit is a.s. constant and equal to μ . Our aim is to compute the speed of convergence and establish for it a central limit theorem.

We will assume furthermore that the function F is in $L^2(\phi(x)dx)$. For the statement of the next result, which is not hard to prove, we need the following additional definition. The Gaussian process $\{X(t) : t \in \mathbb{R}\}$ is called “ m -dependent” if $\text{Cov}(X(s), X(t)) = 0$ whenever $|t - s| > m$.

THEOREM 10.6 (Hoeffding and Robins, 1948). *With the notations and hypotheses above, if the process $X(t)$ is m dependent, then*

$$\sqrt{t} \left(1/t \int_0^t F(X(s)) ds - \mu \right) \rightarrow N(0, \sigma^2) \text{ in distribution as } t \rightarrow +\infty,$$

where σ^2 is the variance of $F(Z)$. Z stands for a standard normal variable.

Our aim is to extend this result to processes which are not m -dependent. The proof we present follows Berman (1992 b) with a generalization, due to Kratz and León (2001) (see also León, 2006), to functions F in (10.18) having an Hermite rank not necessarily equal to 1. For $\varepsilon > 0$, we will approximate the given process \mathcal{X} by a new one \mathcal{X}_ε which is $1/\varepsilon$ -dependent and estimate the error.

We need to recall some facts and prove some auxiliary ones before stating and proving the main results. $\bar{H}_n(x)$ denote the modified Hermite’s polynomials of degree n , orthogonal w.r.t. the standard Gaussian measure that have been already defined in Chapter 8 . Recall, that \bar{H}_n can be defined by means of the identity:

$$\exp(tx - t^2/2) = \sum_{n=0}^{\infty} \bar{H}_n(x) \frac{t^n}{n!}.$$

Since F is in $L^2(\phi(x) dx)$, it can be written as

$$F_n = \sum_{n=0}^{\infty} a_n \bar{H}_n(x),$$

with

$$a_n = \frac{1}{n!} \int_{-\infty}^{\infty} F(x) \bar{H}_n(x) \phi(x) dx,$$

and the norm of F in $L^2(\phi(x)dx)$ satisfies

$$\|F\|_2^2 = \sum_{n=0}^{\infty} a_n^2 n!.$$

The Hermite rank of F is defined as the smallest n such that $a_n \neq 0$. For our purpose, we can assume that this rank greater or equal than 1.

A useful standard tool to perform computations with Hermite polynomials is Mehler's formula which we state and prove with an extension (see León and Ortega, 1989).

LEMMA 10.7 (Mehler's formula). (a) Let (X, Y) be a centered Gaussian vector $E(X^2) = E(Y^2) = 1$ and $\rho = E(XY)$. Then,

$$E(\overline{H}_j(X)\overline{H}_k(Y)) = \delta_{j,k}\rho^j.$$

(b) Let (X_1, X_2, X_3, X_4) be a centered Gaussian vector with variance matrix

$$\Sigma = \begin{pmatrix} 1 & 0 & \rho_{13} & \rho_{14} \\ 0 & 1 & \rho_{23} & \rho_{24} \\ \rho_{13} & \rho_{23} & 1 & 0 \\ \rho_{14} & \rho_{24} & 0 & 1 \end{pmatrix}$$

Then, if $r_1 + r_2 = r_3 + r_4$,

$$E(\overline{H}_{r_1}(X_1)\overline{H}_{r_2}(X_2)\overline{H}_{r_3}(X_3)\overline{H}_{r_4}(X_4)) = \sum_{(d_1, d_2, d_3, d_4) \in Z} \frac{r_1!r_2!r_3!r_4!}{d_1!d_2!d_3!d_4!} \rho_{13}^{d_1} \rho_{14}^{d_2} \rho_{23}^{d_3} \rho_{24}^{d_4},$$

where Z is the set of d_i 's satisfying : $d_i \geq 0$;

$$(10.19) \quad d_1 + d_2 = r_1 ; d_3 + d_4 = r_2 ; d_1 + d_3 = r_3 ; d_2 + d_4 = r_4.$$

If $r_1 + r_2 \neq r_3 + r_4$ the expectation is equal to zero.

Notice that the four equations in (10.19) are not independent, and that the set Z is finite and contains, in general, more than one 4-tuple.

PROOF. We give a proof of (b), (a) being a particular case. We have

$$(10.20) \quad E\left(\prod_{i=1}^4 \exp(t_i X_i - \frac{1}{2}t^2)\right) = \exp(\rho_{13}t_1t_3 + \rho_{14}t_1t_4 + \rho_{23}t_2t_3 + \rho_{24}t_2t_4)$$

First, we have by definition

$$\exp(tx - \frac{1}{2}t^2) = \sum_{q=0}^{\infty} \frac{t^q \overline{H}_q(x)}{q!}.$$

So, the left hand side of (10.20) is equal to

$$\sum_{r_i=0}^{\infty} \frac{t_1^{r_1} t_2^{r_2} t_3^{r_3} t_4^{r_4}}{r_1! r_2! r_3! r_4!} E(\overline{H}_{r_1}(X_1)\overline{H}_{r_2}(X_2)\overline{H}_{r_3}(X_3)\overline{H}_{r_4}(X_4)).$$

Second, the right hand side of (10.20) is equal to

$$(10.21) \quad \sum_{r=0}^{\infty} \frac{1}{r!} (\rho_{13}t_1t_3 + \rho_{14}t_1t_4 + \rho_{23}t_2t_3 + \rho_{24}t_2t_4)^r \\ = \sum_{r=0}^{\infty} \sum_{d_1+d_2+d_3+d_4=r} \frac{1}{d_1!d_2!d_3!d_4!} t_1^{d_1+d_2} t_2^{d_3+d_4} t_3^{d_1+d_3} t_4^{d_2+d_4}.$$

Identifying both sides it follows that the expectation $E(\overline{H}_{r_1}(X_1)\overline{H}_{r_2}(X_2)\overline{H}_{r_3}(X_3)\overline{H}_{r_4}(X_4))$ is zero if $r_1 + r_2 \neq r_3 + r_4$. In the other cases, the monomial of degree (r_1, r_2, r_3, r_4) in the right hand side of (10.20) corresponds to $r = (r_1 + r_2 + r_3 + r_4)/2$ and it can be found in a unique term in the sum $\sum_{r=0}^{\infty}$. The result follows. \square

As an additional hypothesis, we will assume that the process \mathcal{X} has a spectral density $f(\lambda)$. \mathcal{X} has the following spectral representation:

$$(10.22) \quad X(t) = \sqrt{2} \int_0^\infty [\cos(t\lambda) \sqrt{f(\lambda)} dW_1(\lambda) + \sin(t\lambda) \sqrt{f(\lambda)} dW_2(\lambda)],$$

where W_1 and W_2 are two independent Wiener processes (Brownian motions). Indeed, using isometry properties of the stochastic integral, it is easy to see that the process given by (10.22) is centered, Gaussian and with covariance

$$\begin{aligned} \Gamma(t) = \mathbb{E}(X(s)X(s+t)) &= 2 \int_0^\infty \cos(\lambda s) \cos(\lambda(t+s)) f(\lambda) d\lambda + 2 \int_0^\infty \sin(\lambda s) \sin(\lambda(t+s)) f(\lambda) d\lambda \\ &= 2 \int_0^\infty \cos(\lambda t) f(\lambda) d\lambda \end{aligned}$$

Define now the function $\psi(\cdot)$ as the convolution $\mathbf{1}_{[-\frac{1}{2}, \frac{1}{2}]} * \mathbf{1}_{[-\frac{1}{2}, \frac{1}{2}]}$. This function is even, non negative, $\psi(0) = 1$, has support included in $[-1, 1]$ and a non-negative Fourier transform. Set $\psi_\varepsilon(\cdot) := \frac{1}{\varepsilon} \psi(\varepsilon \cdot)$ and let $\widehat{\psi}_\varepsilon$ be its Fourier transform. Define

$$(10.23) \quad X^\varepsilon(t) := \sqrt{2} \int_0^\infty [\cos(t\lambda) \sqrt{f * \widehat{\psi}_\varepsilon(\lambda)} dW_1(\lambda) + \sin(t\lambda) \sqrt{f * \widehat{\psi}_\varepsilon(\lambda)} dW_2(\lambda)],$$

where the convolution must be understood after prolonging f as an even function on \mathbb{R} . The covariance function Γ_ε of $X^\varepsilon(t)$ satisfies $\Gamma_\varepsilon(t) = \Gamma(t)\psi(\varepsilon t)$. This implies that the process $X^\varepsilon(t)$ is $\frac{1}{\varepsilon}$ -dependent. We have the following proposition:

PROPOSITION 10.8. *Let \mathcal{X} be a centered stationary Gaussian process with spectral density $f(\lambda)$ and covariance function Γ with $\Gamma^\ell \in L^1(\mathbb{R})$, ℓ positive integer. Let $X_\varepsilon(t)$ be defined by (10.23). Then*

$$(10.24) \quad \lim_{\varepsilon \rightarrow 0} \lim_{t \rightarrow \infty} \mathbb{E} \left[\frac{1}{\sqrt{t}} \int_0^t (\overline{H}_\ell(X(s)) - \overline{H}_\ell(X^\varepsilon(s))) ds \right]^2 = 0.$$

PROOF. Using Mehler's formula and the change of variables $\tau = s_1 - s_2$:

$$\begin{aligned} &\mathbb{E} \left[\frac{1}{\sqrt{t}} \int_0^t (\overline{H}_\ell(X(s)) - \overline{H}_\ell(X^\varepsilon(s))) ds \right]^2 \\ &= 2! \left(\int_0^t (1 - \tau/t) (\Gamma^\ell(\tau) + \Gamma_\varepsilon^\ell(\tau) - 2\rho_\varepsilon^\ell(\tau)) d\tau \right) \\ &= 2! \left(\int_0^t (1 - \tau/t) (\Gamma_\varepsilon^\ell(\tau) - \Gamma^\ell(\tau)) d\tau + 2 \int_0^t (1 - \tau/t) (\Gamma^\ell(\tau) - \rho_\varepsilon^\ell(\tau)) d\tau \right), \end{aligned}$$

where $\rho_\varepsilon(\tau) := \mathbb{E}[X(0)X^\varepsilon(\tau)]$.

Since $|\Gamma_\varepsilon(\tau)|^\ell \leq |\Gamma(\tau)|^\ell$, we see that the first term tends to zero, as t tends to infinity and then ε tend to zero, on applying the dominated convergence theorem.

As for the second, we have

$$\int_0^t (1 - \tau/t) [\Gamma^\ell(\tau) - \rho_\varepsilon^\ell(\tau)] d\tau = \int_0^t (1 - \tau/t) d\tau \int_{-\infty}^{+\infty} \cos(\lambda\tau) [f^{*(\ell)}(\lambda) - g_\varepsilon^{*(\ell)}(\lambda)] d\lambda.$$

where g_ε is the spectral density $\lambda \rightsquigarrow \sqrt{f(\lambda)} \sqrt{f * \widehat{\psi}_\varepsilon(\lambda)}$ and $g_\varepsilon^{*(\ell)}$ denotes the convolution of g_ε , ℓ times with itself.

Using Fubini's Theorem:

$$(10.25) \quad \begin{aligned} &\int_0^t (1 - \tau/t) d\tau \int_{-\infty}^{+\infty} \cos(\lambda\tau) (f^{*(\ell)}(\lambda) - g_\varepsilon^{*(\ell)}(\lambda)) d\lambda \\ &= \int_{-\infty}^{+\infty} \frac{1 - \cos \lambda t}{t\lambda^2} (f^{*(\ell)}(\lambda) - g_\varepsilon^{*(\ell)}(\lambda)) d\lambda = \int_{-\infty}^{+\infty} \frac{1 - \cos \lambda}{\lambda^2} (f^{*(\ell)}(\frac{\lambda}{t}) - g_\varepsilon^{*(\ell)}(\frac{\lambda}{t})) d\lambda, \end{aligned}$$

When ℓ is equal to 1, the function f and thus g_ε are bounded and continuous and the dominated convergence theorem implies that the limit as $t \rightarrow \infty$ of (10.25) is equal to

$$[f(0) - g_\varepsilon(0)] \int_{-\infty}^{+\infty} \frac{1 - \cos \lambda}{\lambda^2} d\lambda.$$

This quantity tends to zero as $\varepsilon \rightarrow 0$.

When $\ell > 1$ we first prove that $g_\varepsilon^{*\ell}$ is bounded by $\|f * \psi_\varepsilon\|_\infty$. We have:

$$(10.26) \quad g_\varepsilon^{*2}(\lambda) = \int_{-\infty}^{+\infty} g_\varepsilon(\lambda - \lambda_1) g_\varepsilon(\lambda_1) d\lambda_1 \\ \leq \|f * \psi_\varepsilon\|_\infty \int_{-\infty}^{+\infty} f(\lambda - \lambda_1) f(\lambda_1) d\lambda_1 \leq \|f * \psi_\varepsilon\|_\infty,$$

because of the Cauchy-Schwarz inequality.

For $k > 2$, we use induction. Clearly:

$$\int_{-\infty}^{+\infty} (f * \hat{\psi}_\varepsilon)(\lambda) d\lambda = \Gamma_\varepsilon(0) = \Gamma(0)\psi(0) = 1,$$

so that

$$(10.27) \quad g_\varepsilon^{*(k)}(\lambda) \leq \|g_\varepsilon^{*(k-1)}\|_\infty \int_{-\infty}^{+\infty} ((f * \psi_\varepsilon)(\lambda) f(\lambda))^{\frac{1}{2}} d\lambda \leq \|g_\varepsilon^{*(k-1)}\|_\infty \leq \|f * \psi_\varepsilon\|_\infty.$$

Now $g_\varepsilon(\cdot - \frac{\lambda}{t})$ converges to $g_\varepsilon(\cdot)$ in $L^1(\mathbb{R})$, as $t \rightarrow +\infty$. This is nothing more than the continuity of the translation. The duality between $L^1(\mathbb{R})$ and $L^\infty(\mathbb{R})$ implies that $g_\varepsilon^{*(k)}(\frac{\lambda}{t}) \rightarrow g_\varepsilon^{*(k)}(0)$. Using (10.25) and (10.27)

$$\int_0^{+\infty} (1 - \tau/t) \Gamma^\ell(\tau) - \rho_\varepsilon^\ell(\tau) d\tau \rightarrow (f^{*\ell}(0) - g_\varepsilon^{*\ell}(0)) \int_{-\infty}^{+\infty} \frac{1 - \cos \lambda}{\lambda^2} d\lambda,$$

as $t \rightarrow +\infty$.

Fatou's Lemma and the definition of g_ε imply that

$$(10.28) \quad \liminf_{\varepsilon \rightarrow 0} g_\varepsilon^{*(\ell)}(0) \geq f^{*(\ell)}(0).$$

On the other hand

$$(10.29) \quad g_\varepsilon^{*\ell}(0) = \int_{-\infty}^{\infty} g_\varepsilon(\lambda_{\ell-1}) g_\varepsilon^{*(\ell-1)}(\lambda_{\ell-1}) d\lambda_{\ell-1} \\ = \int_{\mathbb{R}^{\ell-1}} g_\varepsilon(\lambda_{\ell-1}) g_\varepsilon(\lambda_{\ell-1} - \lambda_{\ell-2}) \dots g_\varepsilon(\lambda_2 - \lambda_1) g_\varepsilon(\lambda_1) d\lambda_1, \dots, d\lambda_{\ell-1} \\ \leq \left[\int_{\mathbb{R}^{\ell-1}} (f * \hat{\psi}_\varepsilon)(\lambda_{\ell-1}) (f * \hat{\psi}_\varepsilon)(\lambda_{\ell-1} - \lambda_{\ell-2}) \dots (f * \hat{\psi}_\varepsilon)(\lambda_2 - \lambda_1) (f * \hat{\psi}_\varepsilon)(\lambda_1) d\lambda_1, \dots, d\lambda_{\ell-1} \right]^{1/2} \\ \times \left[\int_{\mathbb{R}^{\ell-1}} f(\lambda_{\ell-1}) f(\lambda_{\ell-1} - \lambda_{\ell-2}) \dots f(\lambda_2 - \lambda_1) f(\lambda_1) d\lambda_1, \dots, d\lambda_{\ell-1} \right]^{1/2}, \\ = [(f * \hat{\psi}_\varepsilon)^{*(\ell)}(0)]^{1/2} [f^{*(\ell)}(0)]^{1/2} \rightarrow f^{*(\ell)}(0) \text{ as } \varepsilon \rightarrow 0,$$

using the Cauchy-Schwarz inequality and the continuity of $f^{*(\ell)}$ since Γ^ℓ is in L^1 .

Summing up, (10.28) and (10.29) imply $\lim_{\varepsilon \rightarrow 0} g_\varepsilon^{*(\ell)}(0) = f^{*(\ell)}(0)$, and we are done. \square

THEOREM 10.9. *Let \mathcal{X} be a Gaussian process satisfying the hypotheses of Proposition 10.8 and F a function in $L^2(\phi(x)dx)$ with Hermite rank $\ell \geq 1$. Then, as $t \rightarrow +\infty$,*

$$\sqrt{t}T_t = \frac{1}{\sqrt{t}} \int_0^t F(X(s)) ds \rightarrow N(0, \sigma^2(F)) \text{ in distribution}$$

where

$$\sigma^2(F) := 2 \sum_{k=\ell}^{\infty} a_k^2 k! \int_0^{\infty} \Gamma^k(s) ds.$$

PROOF. Define $F_M := \sum_{n=\ell}^M a_n \bar{H}_n(x)$ and $T_t^M := \frac{1}{t} \int_0^t F_M(X(s)) ds$. Let $M = M(\delta) > \ell$ such that

$$2 \sum_{k=M+1}^{\infty} a_k^2 < \delta.$$

Using Mehler's formula, we get

$$t \operatorname{Var}(T_t - T_t^M) = 2 \sum_{k=M}^{\infty} c_k^2 k! \int_0^t \left(1 - \frac{s}{t}\right) \Gamma^k(s) ds \leq 2 \sum_{k=M}^{\infty} c_k^2 k! \int_0^{\infty} |\Gamma|^k(s) ds < \delta \int_0^{\infty} |\Gamma|^\ell(s) ds.$$

Since δ is arbitrary, we only need to prove the asymptotic normality for T_t^M . Let us introduce

$$T_t^{M,\varepsilon} = \frac{1}{t} \int_0^t F_M(X^\varepsilon(s)) ds,$$

where $X^\varepsilon(t)$ has been defined in (10.23). By Proposition 10.8 recalling that for $k \geq \ell$, Γ^k is in $L^1(\mathbb{R})$ since Γ^ℓ is, we obtain:

$$\lim_{\varepsilon \rightarrow 0} \lim_{t \rightarrow \infty} t \operatorname{Var}(T_t^M - T_t^{M,\varepsilon}) = 0.$$

Now Theorem 10.6 for m -dependent sequences implies that $\sqrt{t} T_t^{M,\varepsilon}$ is asymptotically normal. Notice that

$$\sigma_{M,\varepsilon} := \lim_{t \rightarrow \infty} t \operatorname{Var}(T_t^{M,\varepsilon}) = 2 \sum_{k=0}^M a_k^2 k! \int_0^{\frac{1}{\varepsilon}} \Gamma_\varepsilon^k(s) ds$$

and that $\sigma_{M,\varepsilon} \rightarrow \sigma^2(F)$ when $\varepsilon \rightarrow 0$ and $M \rightarrow \infty$, giving the result. \square

2.3. Hermite expansion for crossings of regular processes. Let \mathcal{X} be a centered stationary Gaussian process. With no loss of generality for our purposes, we assume that $\Gamma(0) = -\Gamma''(0) = 1$ and $\Gamma(t) \neq \pm 1$ for $t \neq 0$. We also assume Geman's Condition of Proposition 4.2:

$$\Gamma(t) = 1 - t^2/2 + \theta(t) \quad \text{with} \quad \int \frac{\theta'(t)}{t^2} dt \text{ converges at } 0^+.$$

We define the following expansions

$$(10.30) \quad x^+ = \sum_{k=0}^{\infty} a_k \bar{H}_k(x), \quad x^- = \sum_{k=0}^{\infty} b_k \bar{H}_k(x), \quad |x| = \sum_{k=0}^{\infty} c_k \bar{H}_k(x),$$

We have $a_1 = 1/2$, $b_1 = -1/2$, $c_1 = 0$ and using (8.7) and integration by parts for $k > 2$:

$$a_k = \frac{1}{k!} \int_0^{+\infty} x \bar{H}_k(x) \varphi(x) dx = \frac{1}{k! \sqrt{2\pi}} \bar{H}_{k-2}(0).$$

The classical properties of Hermite polynomials easily imply that for positive k :

$$\begin{aligned} a_{2k+1} &= b_{2k+1} = c_{2k+1} = 0 \\ a_{2k} &= b_{2k} = \frac{(-1)^{k+1}}{\sqrt{2\pi} 2^k k! (2k-1)} \\ c_{2k} &= 2a_{2k}. \end{aligned}$$

We have the following Hermite expansion for the number of up-crossings:

THEOREM 10.10. *Under the conditions above,*

$$U_u := U_u(X, [0, T]) = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} d_j(u) a_k \int_0^T H_j(X(s)) H_k(X'(s)) ds \text{ a.s.}$$

where $d_j(u) = \frac{1}{j!} \phi(u) \bar{H}_j(u)$ and a_k is defined by (10.30). We have similar results, replacing a_k by b_k or c_k , for the number $D(u, [0, T])$ of down-crossings and for the total number of crossings $N(u, [0, T])$.

PROOF. Let $g(\cdot) \in L^2(\phi(x)dx)$ and define the functional

$$T_g^+(t) = \int_0^t g(X(s)) X'^+(s) ds.$$

The convergence of the Hermite expansion implies that a.s.

$$(10.31) \quad T_g^+(t) = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} g_j a_k \int_0^t H_j(X(s)) H_k(X'(s)) ds,$$

where the g'_j s are the coefficients of the Hermite expansion of g . Using that for each s , $X(s)$ and $X'(s)$ are independent, we get:

$$(10.32) \quad \mathbb{E} \left[\int_0^t \left[g(X(s)) |X'(s)| - \sum_{j,k \geq 0: k+j \leq Q} g_j a_k H_j(X(s)) H_k(X'(s)) \right] ds \right]^2 \leq (\text{const}) t^2 \sum_{j,k \geq 0: k+j \geq Q} j! g_j^2 k! a_k^2,$$

On the other hand, using Proposition 4.2,

$$\nu_2(u, T) = \mathbb{E}(U_u([0, T])(U_u([0, T]) - 1)) = \int_0^T 2(T - \tau) A_{0,\tau}^+(u, u) d\tau,$$

with

$$A_{0,\tau}^+(u, u) = \mathbb{E}(X'^+(0) X'^+(\tau) | X(0) = X(\tau) = u) p_{X(0), X(\tau)}(u, u) \leq \frac{\theta'(\tau)}{\tau^2},$$

For every T , $\nu_2(u, T)$ is a bounded continuous function of u and the same holds true for $\mathbb{E}(U_u^2)$. Let us now define

$$U_u^\delta := \frac{1}{2\delta} \int_0^T \mathbb{1}_{|X(t)-u| \leq \delta} X'^+(t) dt$$

In our case, hypotheses $H_{1,u}$ of Lemma 3.1 are a.s. satisfied. This lemma can be easily extended to up-crossings, showing that

$$U_u^\delta \rightarrow U_u \text{ a.s. as } \delta \rightarrow 0.$$

By Fatou's Lemma

$$\mathbb{E}((U_u)^2) \leq \liminf_{\delta \rightarrow 0} \mathbb{E}((U_u^\delta)^2)$$

To obtain an inequality in the opposite sense, we use the so-called Banach formula (3.31) (see Exercise 3.8). To do that, notice that this formula remains valid if one replaces in the left-hand side the total number of crossings by the up-crossings and in the right-hand side $|f'(t)|$ by $f'^+(t)$. So, on applying it to the random path $X(\cdot)$, we see that:

$$U_u^\delta = \frac{1}{2\delta} \int_{u-\delta}^{u+\delta} U_x dx.$$

and using Jensen's inequality,

$$\limsup_{\delta \rightarrow 0} \mathbb{E}((U_u^\delta)^2) \leq \limsup_{\delta \rightarrow 0} \frac{1}{2\delta} \int_{u-\delta}^{u+\delta} \mathbb{E}((U_x)^2) dx = \mathbb{E}((U_u)^2)$$

So, $\mathbb{E}((U_u^\delta)^2) \rightarrow \mathbb{E}((U_u)^2)$ and since the random variables involved are non-negative, a standard argument of passage to the limit based upon Fatou's Lemma shows that $U_u^\delta \rightarrow U_u$ in L^2 . We now apply (10.31) to U_u^δ .

$$(10.33) \quad U_u^\delta = \sum_{j,k=0}^{\infty} d_j^\delta(u) a_k \zeta_{jk},$$

where $d_j^\delta(u)$ are the Hermite coefficients of the function $x \rightsquigarrow \frac{1}{\delta} \mathbf{1}_{\|x-u\| \leq \delta}$ and

$$\zeta_{jk} = \int_0^T H_j(X(s)) H_k(X'(s)) ds.$$

Notice that

$$(10.34) \quad d_j^\delta(u) \rightarrow \frac{1}{j!} \phi(u) \bar{H}_j(u) = d_j(u).$$

On the other hand, let us denote by \mathcal{S}_q the closed linear subspace of the L^2 of the probability space, generated by the random variables $\{\zeta_{jk} : j, k \geq 0, j+k=q\}$.

A direct application of Mehler's formula's Lemma 10.7, part (b), plus Fubini's Theorem, shows that the subspaces $\{\mathcal{S}_q\}_{q=0,1,\dots}$ are pairwise orthogonal. So, we may rewrite (10.33) in the form:

$$(10.35) \quad U_u^\delta = \sum_{q=0}^{\infty} \gamma_q^\delta$$

where

$$\gamma_q^\delta = \sum_{j+k=q} d_j^\delta(u) a_k \zeta_{jk} \rightarrow \gamma_q := \sum_{j+k=q} d_j(u) a_k \zeta_{jk}.$$

For every integer $Q > 0$, $\sum_{q=0}^Q \gamma_q^\delta$ is equal to $\Pi_Q(U_u^\delta)$, where Π_Q is the orthogonal projector on the space generated by the first Q spaces \mathcal{S}_Q . Using the convergence of U_u^δ and the continuity of the projection:

$$\Pi_Q(U_u^\delta) \rightarrow \Pi_Q(U_u).$$

On the other hand

$$\Pi_Q(U_u^\delta) \rightarrow \sum_{q=0}^Q \gamma_q \text{ as } \delta \rightarrow 0.$$

This implies that:

$$(10.36) \quad U = \sum_{q=0}^{\infty} \sum_{j+k=q} d_j(u) a_k \zeta_{jk}.$$

□

THEOREM 10.11. *Let $\mathcal{X} = \{X(t) : t \in \mathbb{R}\}$ be a centered stationary Gaussian process verifying the conditions at the beginning of this subsection. Furthermore, let us assume that:*

$$\int_0^{+\infty} |\Gamma(t)| dt, \int_0^{+\infty} |\Gamma'(t)| dt, \int_0^{+\infty} |\Gamma''(t)| dt < \infty.$$

Let $\{g_k\}_{k=0,1,2,\dots}$ a sequence of coefficients which satisfies $\sum_0^{+\infty} g_k^2 k! < \infty$. Put:

$$F_t := \frac{1}{\sqrt{t}} \sum_{k,j \geq 0} g_j a_k \int_0^t \bar{H}_j(X(s)) \bar{H}_k(X'(s)) ds$$

where a_k has been defined in (10.30). Then

$$F_t - \mathbb{E}(F_t) \rightarrow N(0, \sigma^2) \text{ in distribution as } t \rightarrow +\infty$$

where

$$0 < \sigma^2 = \sum_{q=1}^{\infty} \sigma^2(q) < \infty,$$

and

$$\sigma^2(q) := \sum_{k=0}^q \sum_{k'=0}^q a_k a_{k'} g_{q-k} g_{q-k'} \int_0^{+\infty} \mathbb{E}[\overline{H}_{q-k}(X(0))\overline{H}_k(X'(0))\overline{H}_{q-k'}(X(s))\overline{H}_{k'}(X'(s))] ds.$$

The integrand in the right-hand side of this formula can be computed using Lemma 10.7. Similar results exist, *mutatis mutandis*, for the sequences $\{b_k\}$ and $\{c_k\}$.

A consequence is

COROLLARY 10.12. *If the process \mathcal{X} satisfies the conditions of Theorem 10.11 then, as $T \rightarrow +\infty$*

$$\begin{aligned} \frac{1}{\sqrt{t}} \left(U_u([0, T]) - T \frac{e^{-u^2/2}}{2\pi} \right) &\rightarrow N(0, \sigma_1^2) \text{ in distribution} \\ \frac{1}{\sqrt{t}} \left(N_u([0, T]) - T \frac{e^{-u^2/2}}{\pi} \right) &\rightarrow N(0, \sigma_2^2) \text{ in distribution,} \end{aligned}$$

where σ_1^2 and σ_2^2 are finite and positive .

Remark The result of Theorem 10.11 is in fact true under weaker hypotheses namely $\int_0^{+\infty} |\Gamma(t)| dt < \infty$, $\int_0^{+\infty} |\Gamma''(t)| dt < \infty$, see Theorem 1 of Kratz and León (2001) or Kratz (2006). Our stronger hypotheses make it possible to make a proof self-contained and rather short.

PROOF OF THE THEOREM

Since Γ is integrable, the process \mathcal{X} admits a spectral density. The hypotheses and the Riemann-Lebesgue Lemma imply that:

$$\Gamma^{(i)}(t) \rightarrow 0 \quad i = 0, 1, 2 \quad \text{as } t \rightarrow +\infty.$$

Hence, we can choose T_0 so that for $t \geq T_0$

$$(10.37) \quad \overline{\Gamma}(t) := \sup\{|\Gamma(t)|, |\Gamma'(t)|, |\Gamma''(t)|\} \leq 1/4.$$

Step1 In this step we prove that one can choose Q large enough so that F_t can be replaced with an arbitrarily small error (in the L^2 sense) by

$$F_t^Q := \frac{1}{\sqrt{t}} \sum_{q=0}^Q G_t^q \quad \text{with } G_t^q := \sum_{k=0}^q g_{q-k} a_k \int_0^t \overline{H}_{q-k}(X(s)) \overline{H}_k(X'(s)) ds.$$

Let us consider

$$(10.38) \quad \frac{1}{t} \mathbb{E}((G_t^q)^2) = \frac{1}{t} \sum_{k, k'=0}^q g_{q-k} a_k g_{q-k'} a_{k'} \int_0^t dt_1 \int_0^t \mathbb{E}(\overline{H}_{q-k}(X(t_1)) \overline{H}_k(X'(t_1)) \overline{H}_{q-k'}(X(t_2)) \overline{H}_{k'}(X'(t_2))) dt_2.$$

To give an upper-bound for this quantity we split it into two parts.

The part corresponding to $|t_1 - t_2| \geq T_0$ is bounded, using Lemma 10.7, by

$$(10.39) \quad \begin{aligned} &(\text{const}) \sum_{k, k'=0}^q |g_{q-k}| |a_k| |g_{q-k'}| |a_{k'}| \int_{T_0}^t \sum_{(d_1, d_2, d_3, d_4) \in Z} \frac{k!(q-k)!k'!(q-k')!}{d_1!d_2!d_3!d_4!} |\Gamma(s)|^{d_1} |\Gamma'(s)|^{d_2+d_3} |\Gamma''(s)|^{d_4} \\ &\leq (\text{const}) \sum_{k, k'=0}^q |g_{q-k}| |a_k| |g_{q-k'}| |a_{k'}| \int_{T_0}^t \sum_{(d_1, d_2, d_3, d_4) \in Z} \frac{k!(q-k)!k'!(q-k')!}{d_1!d_2!d_3!d_4!} \left(\frac{1}{4}\right)^{(q-1)} \overline{\Gamma}(t), \end{aligned}$$

where Z is as in Lemma 10.7, setting $r_1 = q - k$, $r_2 = k$, $r_3 = q - k'$, $r_4 = k'$.

Remarking that $\sup_d \frac{1}{d!(k-d)!} \leq \frac{2^k}{k!}$ it follows that $\frac{k!(q-k)!k'!(q-k')!}{d_1!d_2!d_3!d_4!}$ in (10.39) is bounded above by $2^q(k'!(q-k')!)$ or $2^q(k!(q-k)!)$ depending on the way we group terms. As a consequence

it is also bounded above by $2^q \sqrt{(k')!(q-k')!(k)!(q-k)!}$ and the right-hand side of (10.39) is bounded above by

$$(10.40) \quad (\text{const}) \sum_{k,k'=0}^q |g_{q-k}| |a_k| |g_{q-k'}| |a_{k'}| q 2^{-q} \sqrt{(k')!(q-k')!(k)!(q-k)!} \int_0^{+\infty} \bar{\Gamma}(t) dt$$

$$\leq (\text{const}) \sum_{k,k'=0}^q |g_{q-k}| |a_k| |g_{q-k'}| |a_{k'}| \sqrt{(k')!(q-k')!(k)!(q-k)!}$$

where we have used that the number of terms in Z is bounded by q .

On the other hand, the integration region in (10.38) corresponding to $|t_1 - t_2| \leq T_0$ can be covered by at most $[t/T_0]$ squares of size $2T_0$. Using Jensen's inequality as we did for the proof of (10.32) we obtain:

$$(10.41) \quad \mathbb{E} \left((G_{2T_0}^q)^2 \right) \leq (\text{const}) T_0^2 \sum_{k=0}^q (q-k)! k! g_{q-k}^2 a_k^2.$$

Finally,

$$\frac{1}{t} \mathbb{E} \left((G_t^q)^2 \right) \leq (\text{const}) \sum_{k=0}^q (q-k)! k! g_{q-k}^2 a_k^2,$$

which is the general term of a convergent series. This proves also that σ^2 is finite.

Step 2 Let us prove that $\sigma^2 > 0$. It is sufficient to prove that $\sigma^2(2) > 0$. Recall that $a_1 = 0$ so that

$$(10.42) \quad \sigma^2(2) = a_0^2 g_2^2 \int_0^{+\infty} \mathbb{E}(\bar{H}_2(X(0))\bar{H}_2(X(s))) ds$$

$$+ a_2^2 g_0^2 \int_0^{+\infty} \mathbb{E}(\bar{H}_2(X'(0))\bar{H}_2(X'(s))) ds$$

$$+ 2a_0 g_2 a_2 g_0 \int_0^{+\infty} \mathbb{E}(\bar{H}_2(X(0))\bar{H}_2(X'(s))) ds.$$

Using the Mehler formula

$$(10.43) \quad \sigma^2(2) = 2a_0^2 g_2^2 \int_0^{+\infty} \Gamma^2(s) ds + 2a_2^2 g_0^2 \int_0^{+\infty} (\Gamma''(s))^2 ds + 4a_0 g_2 a_2 g_0 \int_0^{+\infty} (\Gamma'(s))^2 ds$$

$$= \int_{-\infty}^{+\infty} (\lambda^4 a_0^2 g_2^2 + \lambda^2 2a_0 g_2 a_2 g_0 + a_0^2 g_2^2) f^2(\lambda) d\lambda$$

$$= \int_{-\infty}^{+\infty} (\lambda^2 a_2 g_0 + a_0 g_2)^2 f^2(\lambda) d\lambda > 0.$$

Step 3 Set

$$F_t^{Q,\varepsilon} := \frac{1}{\sqrt{t}} \sum_{q=0}^Q G_t^{q,\varepsilon},$$

with

$$G_t^{q,\varepsilon} = \sum_{k=0}^q g_{q-k} a_k \int_0^t \bar{H}_{q-k}(X^\varepsilon(s)) \bar{H}_k((X^\varepsilon)'(s)) ds.$$

In this step, we prove that F_t^Q can be replaced, with an arbitrarily small error if ε is small enough, by $F_t^{Q,\varepsilon}$. Since the expression of F_t^Q involves only a finite number of terms having the form:

$$K_{q-k,k}^0 := \frac{1}{\sqrt{t}} \int_0^t \bar{H}_{q-k}(X(s)) \bar{H}_k(X'(s)) ds$$

if ε is small enough, one can replace with an arbitrarily small error by

$$K_{q-k,k}^\varepsilon := \frac{1}{\sqrt{t}} \int_0^t \overline{H}_{q-k}(X^\varepsilon(s)) \overline{H}_k((X^\varepsilon)'(s)) ds.$$

For that purpose we study

$$(10.44) \quad \begin{aligned} \mathbb{E}(K_{q-k,k}^0 - K_{q-k,k}^\varepsilon)^2 &= 2 \int_0^t \frac{t-s}{t} \mathbb{E} \left[\overline{H}_{q-k}(X(0)) \overline{H}_k(X'(0)) \overline{H}_{q-k}(X(s)) \overline{H}_k(X'(s)) \right] \\ &\quad + \mathbb{E} \left[\overline{H}_{q-k}(X^\varepsilon(0)) \overline{H}_k((X^\varepsilon)'(0)) \overline{H}_{q-k}(X^\varepsilon(s)) \overline{H}_k((X^\varepsilon)'(s)) \right] \\ &\quad - 2 \mathbb{E} \left[\overline{H}_{q-k}(X(0)) \overline{H}_k(X'(0)) \overline{H}_{q-k}(X^\varepsilon(s)) \overline{H}_k((X^\varepsilon)'(s)) \right] ds. \end{aligned}$$

Consider the computation of terms of the kind

$$(10.45) \quad \int_0^t \frac{t-s}{t} \mathbb{E} \left[\overline{H}_{q-k}(Y_1(0)) \overline{H}_k(Y_1'(0)) \overline{H}_{q-k}(Y_2(s)) \overline{H}_k(Y_2'(s)) \right] ds$$

where the processes $Y_1(t)$ and $Y_2(t)$ are chosen among $\{X(t), X^\varepsilon(t)\}$. It suffices to prove that all these terms have the same limit, as $t \rightarrow +\infty$ and then $\varepsilon \rightarrow 0$ whatever the choice is.

Applying Lemma 10.7, the expectation of (10.45) is equal to

$$\int_0^t \frac{t-s}{t} \sum_{d_1, \dots, d_4 \in Z} \frac{(q-k)!^2 k!^2}{d_1! d_2! d_3! d_4!} (\rho(s))^{d_1} (\rho'(s))^{d_2} (-\rho'(s))^{d_3} (-\rho''(s))^{d_4} ds,$$

where $\rho(\cdot)$ is the covariance function between the processes Y_1 and Y_2 and Z is defined as in Lemma 10.7. Again, since the number of terms in Z is finite, it suffices to prove that

$$\lim_{\varepsilon \rightarrow 0} \lim_{t \rightarrow \infty} \int_0^t \frac{t-s}{t} (\rho(s))^{d_1} (\rho'(s))^{d_2+d_3} (\rho''(s))^{d_4} ds,$$

where (d_1, \dots, d_4) is chosen in Z , does not depend on the way to choose Y_1 and Y_2 . ρ is the Fourier transform of (say) $g(\lambda)$ whig is taken among $f(\lambda)$; $f * \widehat{\psi}_\varepsilon(\lambda)$ or $\sqrt{f(\lambda)} \sqrt{f * \widehat{\psi}_\varepsilon(\lambda)}$. Define $\overline{g}(\lambda) = i\lambda g(\lambda)$ and $\overline{\overline{g}}(\lambda) = -\lambda^2 g(\lambda)$. Then $(\rho(s))^{d_1} (\rho'(s))^{d_2+d_3} (\rho''(s))^{d_4}$ is the Fourier transform of the function

$$h(\lambda) = g^{*d_1}(\lambda) * \overline{g}^{*(d_2+d_3)}(\lambda) \overline{\overline{g}}^{*d_4}(\lambda).$$

The continuity and boundedness of f imply that all the functions above are bounded and continuous. The same reasoning that led to (10.25) shows that

$$\int_0^t \frac{t-s}{t} \rho(s)^{d_1} \rho'(s)^{d_2+d_3} (\rho''(s))^{d_4} ds = \int_{-\infty}^{+\infty} \frac{1 - \cos \lambda}{\lambda^2} h\left(\frac{\lambda}{t}\right) d\lambda,$$

As $t \rightarrow 0$, the right-hand side converges, using dominated convergence, to

$$\int_{-\infty}^{+\infty} \frac{1 - \cos \lambda}{\lambda^2} h(0) d\lambda.$$

The continuity of f now gives the result, as in Proposition 10.8. □.

PROOF OF COROLLARY 10.12:

Some attention must be payed to the fact that the coefficients

$$d_j(u) = \frac{1}{j!} \phi(u) \overline{H}_j(u)$$

do not satisfy $\sum_j = 0^\infty j! d_j^2(u) < \infty$. They only satisfy the relation

$$(10.46) \quad j! d_j^2(u) \text{ is bounded}$$

First, considering the bound given by the right hand side of (10.40), we can improve it by reintroducing the factor $q2^{-q}$ that had been bound by 1. We get that in its new expression this right hand side is bounded by

$$\begin{aligned} & (\text{const})q2^{-q} \sum_{k,k'=0}^q |d_{q-k}(u)||a_k||d_{q-k'}(u')||a_{k'}|\sqrt{(k')!(q-k')!(k)!(q-k)!} \\ & \leq (\text{const})q2^{-q} \sum_{k=0}^q (d_{q-k}(u))^2 a_k^2(k)!(q-k)! \\ & \leq (\text{const})q2^{-q} \sum_{k=0}^q a_k^2 k! \leq (\text{const})q2^{-q}. \end{aligned}$$

Second we have to replace the bound (10.41). Since the series in 10.36 is convergent $E\left(\left(G_{2T_0}^q\right)^2\right)$ is the term of a convergent series and this is enough to conclude. \square

2.4. Extensions to random fields. Some of these results can be extended to real-valued random fields, to obtain convergence for the geometric measure of level sets corresponding to a fixed height u , as well as some related functionals defined on them, when the observation window grows to the whole space. More precisely, Iribarren (1989) contains a Central Limit Theorem for integrals on the level set, under some regularity and mixing conditions. The main tool are the formulas (6.9) and (6.10). This asymptotic result has been used when $d = 2$ by Cabaña (1987) to provide a method to test the isotropy of the law of the random field, on the basis of the observation of level sets. The original idea is simple and fruitful: a deformation in the domain which breaks isotropy is reflected in the length of the level sets, and this can be used to estimate anisotropy. The same idea is used in Wschebor (1985, Chapter 3) for general $d \geq 1$. Some extensions can be found in Kratz and León (2001).

Exercises

EXERCISE 10.1. Prove Theorem 10.6. Hint: partition the interval $[0, t]$ into $2n - 1$ intervals, (n being a function of t) $I_1, J_1, \dots, J_{n-1}, I_n$, the J_i 's being of size m .

EXERCISE 10.2. Let $\{X(t) : t \in \mathbb{R}^d\}$, $d \geq 2$ be a real-valued, centered Gaussian, stationary random field with paths of class \mathcal{C}^4 and covariance:

$$\Gamma(t) = E(X(s)X(s+t)), \quad s, t \in \mathbb{R}^d.$$

We assume the normalization $\Gamma(0) = 1$ and that $-\Gamma''(0) = \text{Var}(X'(t))$ is positive definite;

For each $u \in \mathbb{R}$, denote by $\sigma(T, u)$ the $(d - 1)$ -dimensional geometric measure of the intersection of the inverse image of u with the window $T \subset \mathbb{R}^d$.

a) Prove that for each Borel subset T of \mathbb{R}^d , one has:

$$E(\sigma(T, u)) = \lambda_d(T)\phi(u)E(\|\xi\|),$$

where ξ is a centered random vector with values in \mathbb{R}^d , $\text{Var}(\xi) = -\Gamma''(0)$.

b) Put (see the notation in a)):

$$c(u) = \phi(u)E(\|\xi\|).$$

Under the additional (mixing-type) hypothesis that for $i = 0, 1, 2$:

$$(\log \|t\|)^{1-\frac{i}{2}}\Gamma^{(i)}(t) \rightarrow 0 \text{ as } \|t\| \rightarrow +\infty,$$

prove that for each bounded Borel set $T \subset \mathbb{R}^d$, as $u \rightarrow +\infty$, one has:

$$(10.47) \quad \mu((c(u))^{-1/d}T, u) \rightarrow \lambda_d(T)$$

in probability.

(Hint: Apply Rice' formula for $k = 2$).

Geometric characteristics of random sea-waves

In this chapter we will consider extremely simplified representations of a very complicated phenomenon and our presentation will not go into the actual fluid dynamics and numerical problems. We will only consider a set of limited questions which interest oceanographers, at least since the 1950's, say since the founding papers of M.S. Longuet-Higgins and collaborators.

The random sea surface will be modelled using a special Gaussian stationary model that appears as a limit of the superposition of infinitely many elementary sea-waves obeying to the Euler model.

For the defined random surface, we consider some geometrical characteristics like wave-length, crests, length and speed of contours. The various Rice formulas are used to compute expectation or Palm distribution (see definition below) of such quantities. Some numerical applications are presented and a brief description of some non-Gaussian models is given in Section 5.

1. Gaussian model for infinitely deep sea

Let us consider a moving incompressible fluid (the water of the sea) in a domain of infinite depth. If one writes the Euler equations, after some approximations one can show that a class of solutions describing the sea level $W(t, x, y)$, where t is the time variable and x, y are space variables, is given by

$$(11.1) \quad W(t, x, y) = f \cos(\lambda_t t + \lambda_x x + \lambda_y y + \theta),$$

where f and θ and the amplitude and the phase and the pulsations $\lambda_t, \lambda_x, \lambda_y$ are some parameters that satisfy the so-called Airy relation

$$(11.2) \quad \kappa = \frac{\lambda_t^2}{g} \quad \text{with } \kappa^2 := \lambda_x^2 + \lambda_y^2$$

where g is the acceleration of gravity. In what follows, we assume that units have been chosen so that $g = 1$.

For a suitable random choice of f and θ namely, independent, f having Rayleigh distribution (see Exercise 3.12) and θ uniform in $[0, 1\pi]$, $W(t, x, y)$ is an elementary Gaussian field called the sine-cosine process, because it can be written in the form:

$$(11.3) \quad W(t, x, y) = \xi_1 \sin(\lambda_t t + \lambda_x x + \lambda_y y) + \xi_2 \cos(\lambda_t t + \lambda_x x + \lambda_y y)$$

where ξ_1 and ξ_2 are two independent standard normal random variables.

Since the Euler equation is linear, a finite sum of elementary waves having the form (11.3) is again a solution. The limit of such a sum as the number of elementary waves tends to infinity is, using the results in Chapter 1, a stationary random field having the particularity that its spectral measure $F(d\lambda_t, d\lambda_x, d\lambda_y)$ lies in the surface defined by the Airy relation. This surface is a paraboloid having circular sections for constant t .

This will be our basic model. It is an approximation which can be valid only over short periods of time (about 1 hour) and over short geographical areas (several kilometers). It is also understood that very long-period phenomena, like tide and surge, have been removed, so that we will also assume that the process is centered.

The symmetry of the distribution implied by the Gaussian hypothesis (i.e. that the random fields $W(t, x, y)$ and $-W(t, x, y)$ have the same law), is considered by certain authors as a drawback for an adequate representation of the true behavior of the sea level. We will present in Section 5 of this chapter an extension, intending to take this problem into account.

The covariance function of the process, that is:

$$\Gamma(\Delta t, \Delta x, \Delta y) = E\{W(t, x, y)W(t + \Delta t, x + \Delta x, y + \Delta y)\}$$

is the Fourier transform of the Borel measure $F(d\lambda_t, d\lambda_x, d\lambda_y)$.

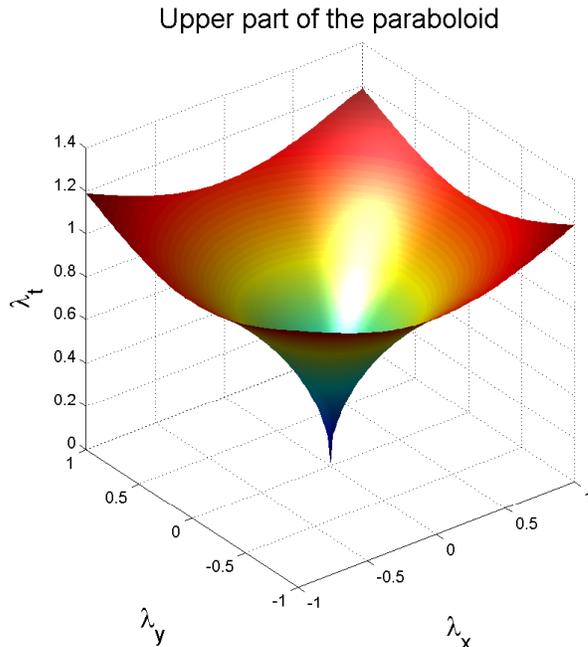


FIGURE 11.1. Representation of the surface on which the spectral measure lies .

Since the spectral measure is symmetric with respect to 0, the lower half of the paraboloid can be removed for our calculations. If we keep only the polar variable κ and α where α is the angle of the vector (λ_x, λ_y) with the x -axis, we can write

$$(11.4) \quad \Gamma(\Delta t, \Delta x, \Delta y) = \int_0^{+\infty} \int_0^{2\pi} \cos(\sqrt{\kappa}\Delta t + \kappa \cos \alpha \Delta x + \kappa \sin \alpha \Delta y) \tilde{G}(d\kappa, d\alpha).$$

Here \tilde{G} is the measure obtained by expressing in polar coordinates the projection of the spectral measure (after removing the lower part) onto the plane (λ_x, λ_y) . Notice that this measure does not need to be symmetric, in the sense that it may not be invariant under the transformation $(\kappa, \alpha) \rightarrow (\kappa, \alpha + \pi)$.

A standard form to write the spectral representation of the covariance is a slight modification of (11.4). Put $\omega = \lambda_t$ (the pulsation) and make the change of variables $\omega = \lambda_t = \sqrt{\kappa}$. Then:

$$(11.5) \quad \Gamma(\Delta t, \Delta x, \Delta y) = \int_0^{+\infty} \int_0^{2\pi} \cos(\omega\Delta t + \omega^2 \cos(\alpha\Delta x) + \omega^2 \sin(\alpha\Delta y)) G(d\omega, d\alpha)$$

G is the spectral measure of the random wave in the so-called sense of “wave community”. It is a non-negative measure expressed in m^2/s which is a unit of power. G is called the “directional power spectrum”. More details on wave modelling can be found, for example, in Kinsman (1965) or Ochi (1998).

2. Some geometric characteristics of waves

For the time being, the observation of the sea level is performed by indirect methods. As far as the authors know, registration of the height as a function of the three variables t, x, y is not available and measurements are often limited to the spectrum, in the sense of (11.5), computed as solutions of certain inverse problems. So, a very important question is to deduce from these spectrums some information on the geometry of the waves.

2.1. Time waves. Suppose for the moment that the location (x, y) is fixed and we consider the level $W(t) = W(t, x, y)$ as a function of the time variable only. The length and the height of waves can be defined in various ways. The definitions given by Lindgren and Rychlik (1995) are the following:

Let t_3 be a down-crossing of zero “chosen at random” (this notion will be defined precisely later on, using the Palm distribution) and consider (see figure 11.2):

- t_1 , the last up-crossing of zero preceding t_3
- the point (t_2, w_2) where the maximum between t_1 and t_3 is attained
- t_5 the first up-crossing of zero following t_3
- the point (t_4, w_4) where the minimum between t_3 and t_5 is attained.

Then the wave is defined as the part of the curve between t_1 and t_5 , its length is $L = t_5 - t_1$, its height is $H = w_2 - w_4$ its half length can be also defined as $t_5 - t_3$ or $t_3 - t_1$.

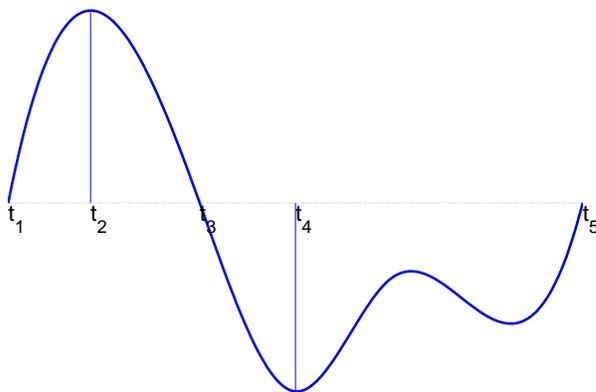


FIGURE 11.2. Remarkable points in the definition of wave-length and wave-height.

Other definitions exist, based on local extrema (again, see Lindgren and Rychlik (1995) and references therein).

DEFINITION 11.1 (Palm distribution). Let $\{T_i, M_i\}_{i=1,2,\dots}$ be a stationary marked point process. This means that $\{T_i\}_{i=1,2,\dots}$ is a point process on the real line (see Chapter 10) and to each point T_i is attached a random variable M_i , “the mark” that takes its values in some measurable space E . Then, for every measurable subset B of E , the Palm distribution of B is given by:

$$(11.6) \quad \mathcal{P}(B) := \frac{\mathbb{E}(\#\{T_i \in [0, T] : M_i \in B\})}{\mathbb{E}(\#\{T_i \in [0, T]\})}$$

Because of the stationarity, this quantity does not depend of the value of $T > 0$.

If the process is defined on the real line and is ergodic (see Chapter 10) then, almost surely one has:

$$\mathcal{P}(B) = \lim_{T \rightarrow \infty} \frac{(\#\{T_i \in [0, T] : M_i \in B\})}{\#\{T_i \in [0, T]\}}$$

so that the Palm measure can be estimated in a consistent way as $T \rightarrow +\infty$ by means of the quotient in the right-hand side of this formula, on the basis of the observation of the point process in the window $[0, T]$.

On the other hand, when applied to random waves, according to the definition given by (11.6), the Palm measure can be computed using one-parameter weighted Rice formula. A basic example is the following:

PROPOSITION 11.2 (Rychlik (1987)). *Let $\{X(t), t \in \mathbb{R}\}$ be a centered stationary Gaussian process satisfying the conditions of Theorem 6.2. The density of the Palm distribution of the half wave period $T_5 - T_3$ is*

$$p_{T_5 - T_3}(\tau) = (\text{const}) p_{X(0), X(\tau)}(0, 0) \mathbb{E}(X'(0)X'(\tau) \mathbf{1}_{X(s) \leq 0, \forall s \in [0, \tau]} | X(0) = X(\tau) = 0)$$

See Exercise 11.1 which contains a hint for the proof.

3. Level curves, crests and velocities for space waves

Let $\mathcal{Z} = \{Z(x, y) : (x, y) \in \mathbb{R}^2\}$ be real-valued 2-parameter centered stationary Gaussian process with differentiable paths. The part of the level curve corresponding to level u contained in the Borel set S , is:

$$\mathcal{C}_u(Z, S) = \{(x, y) \in S : Z(x, y) = u\}.$$

Its mean length $\mathbb{E}(\mathcal{L}(\mathcal{C}_u(Z, S)))$ is given by Rice formula for random fields (Theorem 6.8):

THEOREM 11.3. *Assume that the process \mathcal{Z} satisfies the conditions of Theorem 6.8. Then, with the notations above*

$$(11.7) \quad \mathbb{E}(\mathcal{L}(\mathcal{C}_u(Z, S))) = \lambda_2(S) p_Z(u) \mathbb{E}(\|Z'(0, 0)\|) = \sqrt{\frac{2}{\pi}} \lambda_2(S) p_Z(u) \sqrt{\gamma_2} \mathcal{E}(k)$$

where Σ is the variance matrix of $Z'(0, 0)$, $\gamma_2 > \gamma_1$ are its eigenvalues; $k^2 := (1 - \frac{\gamma_1}{\gamma_2})$; $\mathcal{E}(k) := \int_0^{\pi/2} (1 - k^2 \sin^2 \theta)^{1/2} d\theta$ is the elliptic integral of the first kind and p_Z is the density of $Z(x, y)$.

PROOF. Applying Theorem 6.8 we get

$$\mathbb{E}(\mathcal{L}(\mathcal{C}_u(Z, S))) = \int_S \mathbb{E}(\|Z'(x, y)\| | Z(x, y) = u) p_{Z(x, y)}(u) dx dy = \lambda_2(S) p_Z(u) \mathbb{E}\|Z'(0, 0)\|,$$

because of stationarity. This prove the first relation.

As for the second, after, diagonalization of $\text{Var}(Z'(0, 0))$, $\|Z'(0, 0)\|$, can be represented by $\|\sqrt{\gamma_1} \xi_1 + \sqrt{\gamma_2} \xi_2\|$, where ξ_1 and ξ_2 are two independent standard normal variables. Passing to polar coordinates we have

$$\begin{aligned} \mathbb{E}(\|Z'(0, 0)\|) &= \int_0^{+\infty} d\rho \int_0^{2\pi} \sqrt{\gamma_1 \sin^2(\theta) + \gamma_2 \cos^2(\theta)} \rho e^{-\rho^2/2} d\theta \\ &= \sqrt{\frac{2}{\pi}} \sqrt{\gamma_2} \int_0^{\pi/2} \sqrt{\cos^2(\theta) + (\gamma_1/\gamma_2) \sin^2(\theta)} d\theta = \sqrt{\frac{2}{\pi}} \sqrt{\gamma_2} \int_0^{\pi/2} \sqrt{1 + (\frac{\gamma_1}{\gamma_2} - 1) \sin^2(\theta)} d\theta. \end{aligned}$$

□

Remarks:

- One can find this formula already in Longuet-Higgins (1957, formula (2.3.13))
- Formula (11.7) gives a generalization to every level u of Corrsin's formula (1955). This formula was established for $u = 0$ in a different manner. It says that

$$\frac{\mathbb{E}[\mathcal{L}(\mathcal{C}_u(Z, S))]}{\lambda_2(S)} = \frac{1}{4} \int_0^{2\pi} \tilde{\mathbb{E}}(N_u^\theta) d\theta$$

where $\tilde{E}(N_u^\theta)$ is the expectation per unit of space of the number of crossings in the direction θ . By Rice's formula

$$\tilde{E}(N_u^\theta) = \sqrt{\frac{2m_{2,\theta}}{\pi}} p_Z(u)$$

where $m_{2,\theta}$ is the second spectral moment in the direction θ . Without loss of generality we can assume that the direction in the plane has been chosen to diagonalize the variance matrix of Z' . Then

$$\Sigma(\theta) = \begin{pmatrix} \gamma_2 & 0 \\ 0 & \gamma_1 \end{pmatrix}$$

and $m_{2,\theta} = \sqrt{\gamma_2}(1 - (1 - \frac{\gamma_1}{\gamma_2}) \sin^2 \theta)^{1/2}$ so that the right-hand side in (11.7), for S having Lebesgue measure equal to 1 is equal to:

$$p_Z(u) \sqrt{\frac{2}{\pi}} \sqrt{\gamma_2} \int_0^{\pi/2} (1 - (1 - \frac{\gamma_1}{\gamma_2}) \sin^2 \theta)^{1/2} d\theta = p_Z(u) \sqrt{\frac{2}{\pi}} \sqrt{\gamma_2} \mathcal{E}(k).$$

3.1. Length of a Crest. A crest is defined as a local maximum in a given direction, say θ , of the sea surface modelled as in Section 1.

We define first a static crest at a fixed time (say $t = 0$) as

$$C^s(S, \theta) := \{(x, y) \in S; W'_\theta(x, y) = 0; W''_\theta(x, y) < 0\}$$

where W'_θ ; W''_θ are respectively the first and second derivatives of the field $W(x, y)$ in the θ direction of the (x, y) plane at point $(x, y, 0)$. Since θ is the direction of a straight line it can be chosen in $[0, \pi)$.

It is also possible to define moving crest as

$$C^m(S, T, \theta) := \{(z \cos \theta, z \sin \theta, t) \in S \times [0, T]; W'_\theta(z \cos \theta, z \sin \theta, t) = 0; W''_{\theta\theta}(z \cos \theta, z \sin \theta, t) < 0\}.$$

See Azañs León and Ortega (2005) for more details.

PROPOSITION 11.4. *To simplify the presentation, and without loss of generality, we assume that $\theta = 0$. Let us define the spectral moments*

$$m_{ijk} = \int_0^\infty \int_0^{2\pi} (\omega^2 \cos(\alpha))^i (\omega^2 \sin(\alpha))^j \omega^k G(d\omega, d\alpha),$$

where G has been defined in (11.5). Set $m_{ij} = m_{ij0}$, with definitions above and if the process W'_θ satisfies the conditions of Theorem 6.8. Then

$$E(\mathcal{L}(C^s(S, \theta))) = \frac{\lambda_2(S) \sqrt{\gamma_2}}{2\pi(a_{11})^{1/2}} \mathcal{E}(k),$$

where $k = \sqrt{1 - \frac{\gamma_1}{\gamma_2}}$; $a_{11} = E[W'_x(0, 0, 0)^2] = m_{20}$ and $\gamma_2 > \gamma_1$ are the eigenvalues of Σ the variance matrix of the gradient of W'_x

$$\Sigma(\theta) = \begin{pmatrix} m_{40} & m_{31} \\ m_{31} & m_{22} \end{pmatrix}.$$

PROOF. Denote by $S(x, y)$ the process $W(x, y, 0)$ and let $Z(x, y) = \frac{\partial S}{\partial x}(x, y)$. Then $C^s_{S,0}$ can be written as

$$C^s_{S,0} = \{(x, y) \in S; Z(x, y) = 0; Z'_x(x, y) < 0\}$$

where Z'_x stands for $\frac{\partial Z}{\partial x}(x, y)$. Thus

$$E(\mathcal{L}(C^s_{S,0})) = E \int_{\mathcal{C}_0(Z, S)} \mathbf{1}_{\{Z'_x(x, y) < 0\}} d\sigma$$

where $\mathcal{C}_0(Z, S) = \{(x, y) \in S : Z(x, y) = 0\}$. Since Z' and $-Z'$ have the same distribution,

$$E(\mathcal{L}(C^s_{S,0})) = E \int_{\mathcal{C}_0(Z, S)} \mathbf{1}_{\{Z'_x(x, y) \geq 0\}} d\sigma = \frac{1}{2} E(\mathcal{L}(\mathcal{C}_0(Z, S))).$$

Applying Theorem 11.3 we get

$$\mathbb{E}(\mathcal{L}(C_{S,0}^s)) = \frac{1}{2\pi} \frac{\lambda_2(S)}{(a_{11})^{1/2}} \sqrt{\gamma_2} \mathcal{E}(k)$$

with $a_{11} = \text{Var}(Z(x, y))$ and $\gamma_2 > \gamma_1$ are the eigenvalues of the variance matrix of Z' . □

Remark. When W is an elementary wave of the form

$$W(x, y) = \xi_1 \cos(\lambda_x x + \lambda_y y) + \xi_2 \sin(\lambda_x x + \lambda_y y),$$

where ξ_1, ξ_2 are two standard normal variables, direct computations on the sine-cosine process show that

$$(11.8) \quad \mathbb{E}(\mathcal{L}(C_{S,0}^s)) = \frac{\lambda_2(S) \sqrt{\lambda_x^2 + \lambda_y^2}}{2\pi}$$

Thus, the length of the crest is a non-linear functional of the spectrum.

3.2. Velocity of contours. In this section we give a more rigorous basis to some heuristic considerations of Longuet-Higgins (1957). Other approaches to the same problem have been proposed by Podgórski et al. (2000) and Baxevari et al. (2003) where several notions of velocity are introduced, including the one used here, called “velocity in the direction of the gradient”. Our results are different in the sense that we look at the two components of the gradient while the cited authors express their results in terms of the joint distribution of the modulus and the angle.

Speed of crossings Let us fix y (say $y = 0$). We want to study the speed of a crossing of a given level u chosen “at random” among all the crossings. Define S_0 as the section of S in the direction of the x axis. Using stationarity, it is always possible to suppose that $S_0 = [0, M]$ for some value M . Also by stationarity we can look at the speed of the sea at time 0. A crossing is a point x such that

$$(11.9) \quad W(x, 0, 0) = u$$

The expectation of the number of crossings N_u is given by Rice’s formula

$$D := \mathbb{E}(N_u) = M \sqrt{\frac{2m_{200}}{\pi}} p_Z(u)$$

The speed of such crossings can be computed using the implicit function theorem. From (11.9) we get that

$$(11.10) \quad C_x(x) := \frac{dx}{dt} = -\frac{W'_t(x, 0, 0)}{W'_x(x, 0, 0)}.$$

The mean number of crossings with speed C_x in the interval $[\alpha_1, \alpha_2]$ ($\alpha_1 < \alpha_2$) can also be computed using a Rice formula. If the spectral measure S defined in (11.5) is not reduced to a unit atom, then:

$$(11.11) \quad N := \mathbb{E}(N_u \mathbf{1}_{C_x \in [\alpha_1, \alpha_2]}) = \int_{\alpha_1}^{\alpha_2} dc \int_0^M dx \int_{-\infty}^{\infty} |x'| p_{W, W'_x, C_x}(u, x', c) dx'$$

where p_{W, W'_x, C_x} is the joint density of $(W(x, 0, 0), W'_x(x, 0, 0), C_x(x, 0, 0))$ which does not depend on x because of stationarity. As the values of the process and its derivative at a given point are independent random variables, we get:

$$N = M p_W(u) \int_{\alpha_1}^{\alpha_2} dc \int_{-\infty}^{\infty} |x'| p_{W'_x, C_x}(x', c) dx'$$

The probability of a crossing chosen at random to have a speed in the range $[\alpha_1, \alpha_2]$ is therefore $\frac{N}{D}$. Divide now by $\alpha_2 - \alpha_1$, let both α_1 and α_2 tend to a common limit c , and we get that the distribution of the speed of the crossing is given by

$$\tilde{p}_{C_x}(c) = \sqrt{\frac{\pi}{2m_{200}}} \int_{-\infty}^{\infty} |x'| p_{W'_x, C_x}(x', c) dx'$$

where

$$p_{W'_x, W'_t}(x', t') = \frac{\Delta^{-1/2}}{2\pi} \exp\left\{-\frac{1}{2\Delta}(m_{002}x'^2 - 2m_{101}x't' + m_{200}t'^2)\right\}$$

with

$$\Delta = \det \begin{pmatrix} m_{200} & m_{101} \\ m_{101} & m_{002} \end{pmatrix}.$$

Making the change of variables $c = -\frac{t'}{x'}$ we get

$$\tilde{p}_{C_x}(c) = \frac{1}{2}\Delta(m_{002} + 2m_{101}c + m_{200}c^2)^{-3/2}(m_{200})^{-1/2}$$

This is (with a slightly different notation) formula (2.5.14) in Longuet-Higgins (1957) where it is shown that it can also be written as

$$\tilde{p}_{C_x}(c) = \frac{1}{2}\Delta m_{200}^{-2}((c - \hat{c}) + \Delta m_{200}^{-2})^{-3/2}$$

showing that this distribution is symmetric around its mean value $\hat{c} = -\frac{m_{101}}{m_{200}}$. An important point is that this speed does not depend on the level.

3.3. Velocity of level curves. To define the normal velocity of a level curve we fix a point $P = (0, x_0, y_0)$ such that $W(0, x_0, y_0) = u$ and consider:

- The level surface in time and space

$$C_1 := \{(t, x, y) : W(t, x, y) = u\} \text{ for } (t, x, y) \text{ in some neighbourhood of } (0, x_0, y_0)$$

- The level curve at fixed time

$$C_2 := \{(x, y) : W(0, x, y) = u\} \text{ for } (x, y) \text{ in some neighbourhood of } (x_0, y_0).$$

In an infinitesimal interval of time the point P moves to $P' = P + dt \vec{v}$ where:

- \vec{v} is in the tangent space to C_1 . So, \vec{v} is orthogonal to the gradient of W , that is, (W'_t, W'_x, W'_y) (the derivatives are computed at the point P),
- the t -coordinate of \vec{v} is equal to 1,
- Define $\vec{V} = (V_x, V_y)$, as the orthogonal projection of \vec{v} onto the x, y -plane. \vec{V} is the “normal velocity to the curve” if it is orthogonal to C_2 at the point P .

Then, V_x and V_y satisfy the following equations

$$\begin{aligned} W'_t + V_x W'_x + V_y W'_y &= 0 \\ V_x W'_y - V_y W'_x &= 0 \end{aligned}$$

And it is easy to deduce that

$$V_x = -\frac{W'_t W_x}{(W'_x)^2 + (W'_y)^2} \quad V_y = -\frac{W'_t W_y}{(W'_x)^2 + (W'_y)^2}$$

Following Longuet-Higgins it is simpler to obtain first the distribution of (K_x, K_y) with $K_x = -W'_t/W'_x$ and $K_y = -W'_t/W'_y$ and then pass to the distribution of the velocity using the change of variables formula. As in the preceding proof, we consider two intervals $[\alpha_1, \alpha_2]$, $\alpha_1 < \alpha_2$ and $[\alpha_3, \alpha_4]$, $\alpha_3 < \alpha_4$, for $t = 0$ and define

$$D := \mathbb{E}(\mathcal{L}(\mathcal{C}_u(Z, S))) = |S|p(u)\mathbb{E}\|W_{xy}(0, 0, 0)\|$$

where $Z(x, y) = W(0, x, y)$ and W_{xy} is the gradient limited to the variables x and y and

$$(11.12) \quad N := \mathbb{E} \left[\int_{\mathcal{C}_u(Z, S)} \mathbb{1}_{K_x \in [\alpha_1, \alpha_2]} \mathbb{1}_{K_y \in [\alpha_3, \alpha_4]} d\sigma \right].$$

This expectation can be computed using Rice formula for integrals on a level set (Theorem 6.10 of Chapter 6), as soon as the process $W(t, x, y)$ satisfies the hypotheses of Theorem 6.8.

$$\begin{aligned}
N &= \lambda_2(S)p_W(u)\mathbb{E}[\|W_{xy}\|\mathbf{1}_{K_x \in [\alpha_1, \alpha_2]}\mathbf{1}_{K_y \in [\alpha_3, \alpha_4]}] \\
&= \lambda_2(S)p_W(u) \int_{\mathbf{R}^3} \sqrt{x'^2 + y'^2} \mathbf{1}_{\{-x'/t' \in [\alpha_1, \alpha_2]\}} \mathbf{1}_{\{-y'/t' \in [\alpha_3, \alpha_4]\}} p_{W'_x, W'_y, W'_t}(x', y', t') dx' dy' dt'
\end{aligned}$$

Making the change of variables $k_x = -x'/t'$, $k_y = -y'/t'$, $t' = t'$ with $dx' dy' dt' = t'^2 dk_x dk_y dt'$, after some calculations we get

$$\begin{aligned}
N &= 4\lambda_2(S)p_W(u)\pi^{-2}\Delta_2^{-1/2} \\
&\quad \times \int_{\alpha_1}^{\alpha_2} \int_{\alpha_3}^{\alpha_4} dk_x dk_y \sqrt{k_x^2 + k_y^2} [\mu_{11}k_x^2 + 2\mu_{12}k_x k_y - 2\mu_{13}k_x + \mu_{22}k_y^2 - 2\mu_{23}k_y + \mu_{33}]^{-2}
\end{aligned}$$

where Δ_2 and μ_{ij} are respectively, the determinant, and, the entries of the inverse matrix, of

$$(11.13) \quad \begin{pmatrix} m_{200} & m_{110} & m_{101} \\ m_{110} & m_{020} & m_{011} \\ m_{101} & m_{011} & m_{002} \end{pmatrix}.$$

Letting α_1 and α_2 tend to k_x and α_3 and α_4 tend to k_y we get the joint density of K_x and K_y :

$$\begin{aligned}
\tilde{p}_{K_x, K_y}(k_x, k_y) &= \lim_{\substack{\alpha_1, \alpha_2 \rightarrow k_x \\ \alpha_3, \alpha_4 \rightarrow k_y}} \frac{1}{(\alpha_2 - \alpha_1)(\alpha_4 - \alpha_3)} \frac{N}{D} \\
&= \frac{1}{\pi} (\gamma_2)^{-1/2} \Delta_2^{-1/2} (\mathcal{E}(k))^{-1} \sqrt{k_x^2 + k_y^2} [\mu_{11}k_x^2 + 2\mu_{12}k_x k_y - 2\mu_{13}k_x + \mu_{22}k_y^2 - 2\mu_{23}k_y + \mu_{33}]^{-2}
\end{aligned}$$

where $k = \sqrt{1 - \frac{\gamma_1}{\gamma_2}}$, as before but γ_1 and γ_2 are the eigenvalues of the matrix

$$\begin{pmatrix} m_{200} & m_{110} \\ m_{110} & m_{020} \end{pmatrix}.$$

Again we find the same result as in Longuet-Higgins (1957, equation 2.6.21).

We look now at the distribution of the velocity

$$\vec{V} = (V_x, V_y) = \left(\frac{K_x}{K_x^2 + K_y^2}, \frac{K_y}{K_x^2 + K_y^2} \right)$$

so that:

$$dK_x dK_y = (V_x^2 + V_y^2)^{-2} dV_x dV_y. \text{ As a consequence}$$

$$\begin{aligned}
\tilde{p}_{V_x, V_y}(v_x, v_y) &= \frac{1}{\pi} (\gamma_2)^{-1/2} \Delta_2^{-1/2} (\mathcal{E}(k))^{-1} (v_x^2 + v_y^2)^{7/2} \\
&\quad \times [\mu_{11}v_x^2 + 2\mu_{12}v_x v_y - 2\mu_{13}v_x(v_x^2 + v_y^2) + \mu_{22}v_y^2 - 2\mu_{23}v_y(v_x^2 + v_y^2) + \mu_{33}(v_x^2 + v_y^2)^2]^{-2}.
\end{aligned}$$

Velocity of crests: Since the distributions of W and $-W$ are the same, the mean velocity of a crest is the mean speed of the zero level set for the process W'_θ . Thus the same result holds changing the meaning of the moments in matrix (11.13).

4. Real Data

In this section we present a numerical application from Azaïs León and Ortega (2005). We consider two directional spectra kindly provided by M. Prevosto from Ifremer in France, depicted in figures 11.3

We now compare the geometric characteristics of the random seas corresponding to these spectrums.

Figure 11.4 shows the expected length of static crests along directions, showing a maximum at approximately 1.3 rad. It is interesting to observe that, in accordance with theoretical results, this direction is orthogonal to the direction for the maximum integral of the spectrum, which is the most probable direction for the waves.

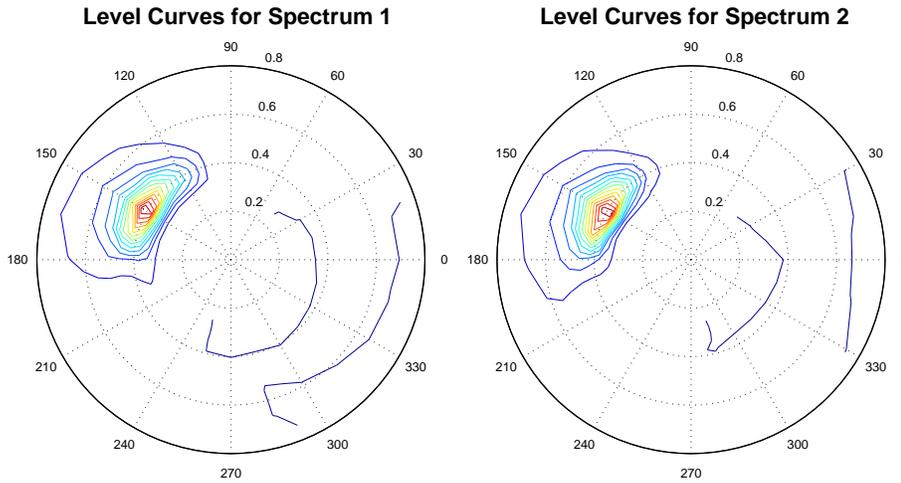


FIGURE 11.3. Representation of the two spectrums

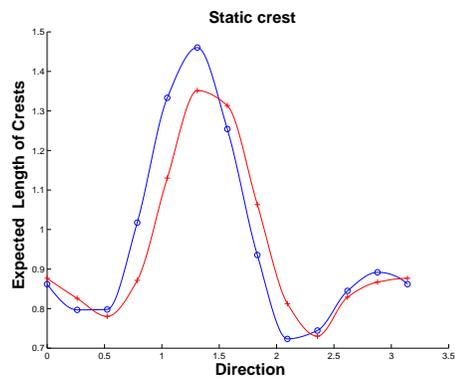


FIGURE 11.4. Expected length of static crests per unit area; circles: spectrum1; cross spectrum2

Figures 11.5 show the level curves for probability densities of the velocity of a level contour of $W(x, y, t)$. Both graphs show a clear asymmetry as predicted by Longuet-Higgins (1957). The distributions are clearly different though the spectra differ only slightly.

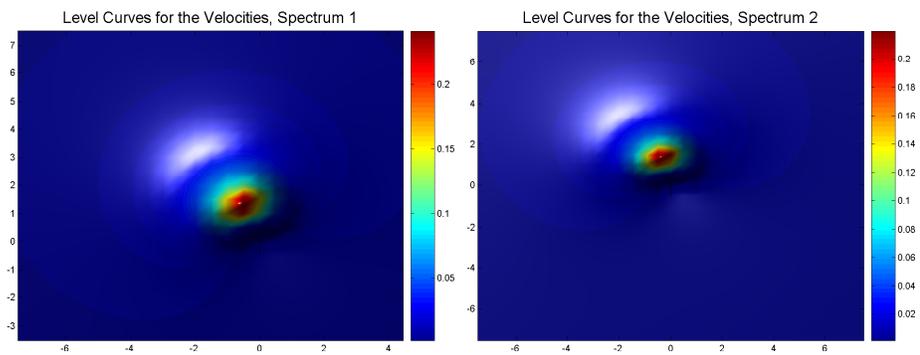


FIGURE 11.5. Representation of the distributions of the velocities of contours

5. Generalizations of the Gaussian model

The crest-trough symmetry of the Gaussian model does not correspond exactly to reality, especially in the critical situation of very high waves. In practice it is often observed that the crests are peaked and narrow while the troughs are wide and shallow. This can affect significantly the distribution of the slopes of the waves, as well as the extremal behavior. These are two important issues to study wave slamming on ships or offshore structures. Unfortunately, the description and understanding of non-Gaussian models appears to be very difficult, so that effective generalizations are based on “nearly-Gaussian” models. We will sketch here two of them: the transformed Gaussian models and the Lagrange models. We limit ourselves to one or two-dimensional models.

5.1. Transformed Gaussian models. Let us consider the elevation $W(t)$ of the sea as a function of time. We assume that it follows a model having the form:

$$(11.14) \quad W(t) = \mu + G(X(t)),$$

where $X(t)$ is a stationary Gaussian process and $G(\cdot)$ is a “nice function”. Such an equation has several advantages: firstly, computations are tractable because they can be conducted on the Gaussian process $X(t)$, second the transformation (11.14) modifies the extremal behavior as it can be seen in Azaïs et al. (2007). Notice that this is not the case with the Lagrange model in the next section.

The function G can be

- a polynomial, in which case it is convenient to use a low degree polynomial (say 4) and represent it in the Hermite basis (see Azaïs et al., 2007, and references therein). The estimation of $G(\cdot)$ is based on the marginal density of the process $W(t)$ and uses the method of moments.
- Non-parametric, as in Rychlik et al. (1997). In this case the function G can be estimated by the intensity of crossings. We give an example of extreme situations corresponding to a registration of the Camilla hurricane in 1969. Figure 11.6 shows a small discard from normality in the high levels.

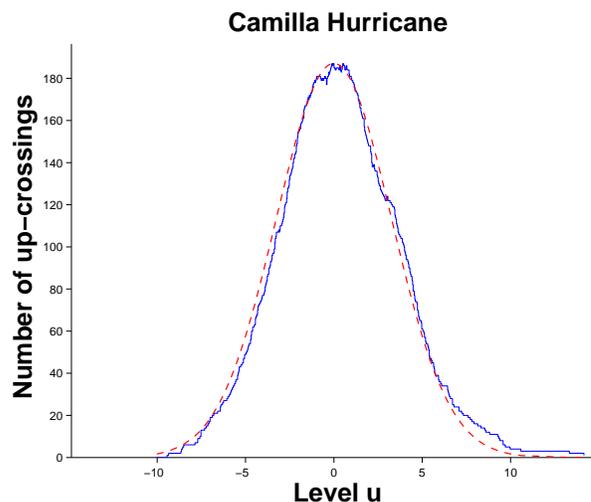


FIGURE 11.6. Number of up-crossings of a level for data from the Camilla hurricane. In dotted line is given the expectation under the Gaussian model.

5.2. Lagrange models. We return to the equations of an incompressible fluid. Under less crude assumptions than for the Euler equation we obtain the *Lagrange model* which has as a main characteristic that water particles have a circular movement around a mean position.

Random models issued from that model are described in a paper by Lindgren (2006). The sea surface (depending on t, x) is described as a parametric surface depending on t (the time) and a dummy parameter u which is close to the location. It represents the mean position around which particles are moving. The sea surface is written as:

$$(t, u) \rightsquigarrow (X(t, u), W(t, u)),$$

where $W(t, u)$ is the height of the sea at the location $X(t, u)$ at time t . The two random fields $X(t, u), W(t, u)$ are jointly Gaussian and described by the stochastic integrals:

$$W(t, u) = \int_{\mathbb{R}} \exp(i(\kappa(\lambda)u - \lambda t)) d\xi(\lambda)$$

$$W(t, u) = u + \int_{\mathbb{R}} i \frac{\cosh(\kappa(\lambda)h)}{\sinh(\kappa(\lambda)h)} \exp(i(\kappa(\lambda)u - \lambda t)) d\xi(\lambda),$$

where h is the water depth, $\kappa(\lambda)$ is defined (up to the sign, the choice of which defines different kinds of waves) by the relation $\lambda^2 = |\kappa| \tanh(|\kappa|h)$ and ξ is a complex spectral process with orthogonal increment satisfying $d\xi(-\lambda) = d\xi(\lambda)$.

This model is essentially used to compute distribution of steepness of waves that differ significantly from the Euler model.

Exercises

EXERCISE 11.1. Prove Proposition 11.2. Let F_n a continuous approximation of the function $\mathbf{1}_{>0}$ as defined for example in 6.7, define $Y^t(s) = X(s)$ and

$$g(t, Y^t) = F_n \left[\sup_{s \in [t, t+\tau]} X(s) \right]$$

and use a monotone convergence argument.

EXERCISE 11.2. Prove formula (11.8) by a direct computation.

Consider now the case of a spectrum G with two atoms, which is the sum of two spectra of elementary waves. Show that the length of the crest is not a linear function of the spectrum.

EXERCISE 11.3. Prove formula (11.11). Prove first that if the spectral measure G is not restricted to a Dirac measure, the joint distribution of the derivatives W_t and W_x does not degenerate.

Second, replacing the indicator function $\mathbf{1}_{[\alpha_1, \alpha_2]}$ by a continuous approximation, prove (11.11) using Theorem 6.4. Then conclude.

EXERCISE 11.4. Prove formula (11.12) using Theorem 6.10 and the same kind of approximation than in Exercise 11.3.

EXERCISE 11.5. Give a detailed version of the last argument of Section 3.3 concerning the velocity of crests.

Systems of random equations

In this chapter we are going to use Rice formula to study the number of real roots of a system of random equations. Our emphasis is on polynomial systems, even though we will also give some results on non-polynomial ones.

Let us consider m polynomials in m variables with real coefficients

$$X_i(t) = X_i(t_1, \dots, t_m), \quad i = 1, \dots, m.$$

We use the notation

$$(12.1) \quad X_i(t) := \sum_{\|j\| \leq d_i} a_j^{(i)} t^j,$$

where $j := (j_1, \dots, j_m)$ is a multi-index of non-negative integers, $\|j\| := j_1 + \dots + j_m$, $j! := j_1! \dots j_m!$, $t = (t_1, \dots, t_m) \in \mathbb{R}^m$, $t^j := t_1^{j_1} \dots t_m^{j_m}$, $a_j^{(i)} := a_{j_1, \dots, j_m}^{(i)}$. The degree of the i -th polynomial is d_i and we assume that $d_i \geq 1 \forall i$.

We denote $N^X(V)$ the number of roots of the system of equations

$$(12.2) \quad X_i(t) = 0, \quad i = 1, \dots, m.$$

lying in the Borel subset V of \mathbb{R}^m . We denote $N^X = N^X(\mathbb{R}^m)$.

Let us randomize the coefficients of the system. In the case of one equation in one variable, a certain number of results on the probability distribution of the number of roots have been known for a long time, starting in the thirties with the work of Bloch and Polya (1932) and Littlewood and Offord (1938, 1939) and especially, of Marc Kac (1943). We are not going to consider this special subject here, see for example the book by Bharucha-Reid and Sambandham (1986).

Instead, when $m > 1$ little is known on the distribution of the random variables $N^X(V)$ or N^X , even for simple choices of the probability law on the coefficients. This appears to be quite different and much harder than one equation only, and it is this case that we will consider in this chapter. In fact, we will be especially interested in large systems, in the sense that $m \gg 1$. In the last 15 years some initial progress has been made in the understanding of the distributional properties of the number of roots. The first important result in this context is the Shub-Smale Theorem (1993), in which the authors computed by means of a simple formula the expectation of N^X when the coefficients are Gaussian, centered independent random variables with certain specified variances (see Theorem 12.1 below). Extensions of their work, including new results for one polynomial in one variable, can be found in the review paper by Edelman and Kostlan (1995). See also Kostlan (2002).

There is of course the curiosity about the number of roots, for example, being able to answer the question whether the system has no real roots, i.e. $N^X = 0$, or, in the random case, what can one say about $P(N^X = 0)$ or $P(N^X > n)$ where n is some meaningful integer for the underlying problem. More deeply, the study of the number of roots is associated to natural questions in Numerical Analysis and Complexity Theory. Generally speaking, the complexity in solving a system of equations numerically is naturally related to the number of roots. So, understanding of the mean (or probabilistic) behavior of an algorithm with respect to a family of problems of this sort is associated to the distribution of the random variable N^X . On the other hand, the condition number of a system of equations, which measures in this case the difficulty for an

algorithm to separate roots, is related to analogous problems, and plays a central role in complexity computations. We are not going to pursue this subject here, the interested reader can consult the book by Blum, Cucker, Shub and Smale (1998).

It is obvious that the distribution of the number of roots will depend on the probability law that we put on the coefficients of the system. So, the first question is what conditions should we require to this law. As we said above, only a restricted family of distributions has been considered until now. The Shub-Smale distribution on the coefficients is invariant under the orthogonal group of the underlying space \mathbb{R}^m and is related to the H. Weyl L^2 -structure in the space of polynomial systems (see also the book by Blum et al (1998) on this subject).

In section 2 we review some results which extend the computation of the expectation to some other probability laws on the coefficients, which have a centered Gaussian law that is invariant under the orthogonal group of \mathbb{R}^m . This allows to extend substantially the family of examples and to show that the behavior of the expectation of the number of roots can be very different from the one in the Shub-Smale Theorem.

We have also included some recent asymptotic results for variances, but only for the Shub-Smale model with equal degrees (that we call Kostlan-Shub-Smale). The main tool is Rice formula to compute the factorial moments of the number of zeros of a random field (see Theorem 6.3) and the asymptotics is for large systems, meaning by that $m \rightarrow +\infty$. We are only giving some brief sketch of the proofs, which turn out to require lengthy calculations, at least when using the available methods. At present, a major open problem is to show weak convergence of some renormalization of N^X , under the same asymptotics.

In section 3 we consider “smooth analysis”, that is, we start with a non-random system, perturb it with some noise, and the question is what can we say about the number of roots of the perturbed system, under some reasonable hypotheses on the relationship between “signal” and “noise”. Here again, we are only able to give results having some interest when the number m of equations and unknowns become large.

Finally, in Section 4 we consider random systems having a probability law which is invariant under translations as well as orthogonal transformations of the underlying Euclidean space. This implies that the system is non-polynomial and the expectation of N^X is infinite in non-trivial cases. So, one has to localize and consider $N^X(V)$ for subsets V of \mathbb{R}^m having finite Lebesgue measure. These systems are interesting by themselves and under some general conditions, one can use similar methods to compute the expected number of roots per unit volume, as well as to understand the behavior of the variance as the the number of unknowns m tends to infinity, which turns out to be strikingly opposite to the one in the Kostlan-Shub-Smale model for polynomial systems.

All the above concerns “square” systems. We have not included results on random systems having less equations than unknowns. If the system has n equations and m unknowns with $n < m$, generically the set of solutions will be $(m - n)$ -dimensional, and the description of the geometry becomes more complicated (and more interesting) than for $m = n$. A recent contribution to the calculation of the expected value of certain parameters describing the geometry of the (random) set of solutions is in P. Bürgisser (2006).

1. The Shub-Smale model

We say that (12.2) is a Shub-Smale system, if the coefficients

$$\{a_j^{(i)} : i = 1, \dots, m; \|j\| \leq d_i\}$$

are centered independent Gaussian random variables, such that

$$(12.3) \quad \text{Var}(a_j^{(i)}) = \binom{d_i}{j} = \frac{d_i!}{j!(d_i - \|j\|)!}$$

1.1. Expectation of N^X .

THEOREM 12.1 (Shub-Smale(1993)). *Let the system (12.2) be a Shub-Smale system. Then,*

$$(12.4) \quad \mathbb{E}(N^X) = \sqrt{D}$$

where $D = d_1 \dots d_m$ is the Bézout number of the polynomial system.

PROOF. . For $i = 1, \dots, m$, let \tilde{X}_i denote the homogeneous polynomial of degree d_i in $m + 1$ variables associated to X_i , that is:

$$\tilde{X}_i(t_0, t_1, \dots, t_m) = \sum_{\sum_{h=0}^m j_h = d_i} a_{j_1, \dots, j_m}^{(i)} t_0^{j_0} t_1^{j_1} \dots t_m^{j_m},$$

Y_i denotes the restriction of \tilde{X}_i to the unit sphere S^m in \mathbb{R}^{m+1} . It is clear that

$$(12.5) \quad N^X = \frac{1}{2} N^Y(S^m)$$

A simple computation using (12.3) and the independence of the coefficients, shows that $\tilde{X}_1, \dots, \tilde{X}_m$ are independent Gaussian centered random fields, with covariances given by:

$$(12.6) \quad r^{\tilde{X}_i}(t, t') = E(\tilde{X}_i(t)\tilde{X}_i(t')) = \langle t, t' \rangle^{d_i}, \quad t, t' \in \mathbb{R}^{m+1}, \quad i = 1, \dots, m.$$

Here $\langle \cdot, \cdot \rangle$ denotes the usual scalar product in \mathbb{R}^{m+1} .

For $E(N^Y(S^m))$ we apply Rice formula to the random field Y defined on the parameter set S^m :

$$(12.7) \quad E(N^Y(S^m)) = \int_{S^m} E(|\det(Y'(t))| | Y(t) = 0) \frac{1}{(2\pi)^{m/2}} \sigma_m(dt),$$

where $\sigma_m(dt)$ stands for the m -dimensional geometric measure on S^m . (12.7) follows easily from the fact that for each $t \in S^m$, the random variables $Y_1(t), \dots, Y_m(t)$ are i.i.d. standard normal.

Since $E(Y_i^2(t)) = 1$ for all $t \in S^m$, on differentiating under the expectation sign, we see that for each $t \in S^m$, $Y(t)$ and $Y'(t)$ are independent, and the condition can be erased in the conditional expectation in the right-hand side of (12.7).

Since the law of $Y'(t)$ is invariant under the orthogonal group of \mathbb{R}^{m+1} it suffices to compute the integrand at one point of the sphere. Denote the canonical basis of \mathbb{R}^{m+1} by $\{e_0, e_1, \dots, e_m\}$. Then:

$$(12.8) \quad \begin{aligned} E(N^Y(S^m)) &= \sigma_m(S^m) \frac{1}{(2\pi)^{m/2}} E(|\det(Y'(e_0))|) \\ &= \frac{2\pi^{(m+1)/2}}{\Gamma(\frac{m+1}{2})} \frac{1}{(2\pi)^{m/2}} E(|\det(Y'(e_0))|). \end{aligned}$$

To compute the probability law of $Y'(e_0)$, let us write it as an $m \times m$ matrix with respect to the orthonormal basis e_1, \dots, e_m of the tangent space to S^m at e_0 . This matrix is

$$\left(\left(\frac{\partial \tilde{X}_i}{\partial t_j}(e_0) \right) \right)_{i,j=1, \dots, m}$$

and

$$E\left(\frac{\partial \tilde{X}_i}{\partial t_j}(e_0) \frac{\partial \tilde{X}_{i'}}{\partial t_{j'}}(e_0) \right) = \delta_{ii'} \frac{\partial^2 r^{\tilde{X}_i}}{\partial t_j \partial t_{j'}} \Big|_{t=t'=e_0} = d_i \delta_{ii'} \delta_{jj'},$$

The last equality follows computing derivatives of the function $r^{\tilde{X}_i}$ given by (12.6). So,

$$(12.9) \quad \det(Y'(e_0)) = \sqrt{D} \det(G)$$

where G is an $m \times m$ matrix with i.i.d. standard normal entries.

To finish, we only need to compute $E(|\det(G)|)$. One way to do it, is to observe that $|\det(G)|$ is the volume (in \mathbb{R}^m) of the set

$$\{v \in \mathbb{R}^m : v = \sum_{k=1}^m \lambda_k g_k, 0 \leq \lambda_k \leq 1, k = 1, \dots, m\}$$

where $\{g_1, \dots, g_m\}$ are the columns of G . Then using the invariance of the standard normal law in \mathbb{R}^m with respect to isometries, we get:

$$E(|\det(G)|) = \prod_{k=1}^m E(\|\eta_k\|),$$

where η_k is standard normal in \mathbb{R}^k . An elementary computation gives:

$$E(\|\eta_k\|) = \sqrt{2} \frac{\Gamma((k+1)/2)}{\Gamma(k/2)}$$

which implies:

$$E(|\det(G)|) = \frac{1}{\sqrt{2\pi}} 2^{(m+1)/2} \Gamma((m+1)/2).$$

Using (12.9), (12.8) and (12.5), we get the result. \square

Remark

When the hypotheses of Theorem 12.1 are verified, and moreover all the degrees d_i ($i = 1, \dots, m$) are equal, formula (12.4) was first proved by Kostlan. In what follows, we will call such a model the KSS (Kostlan-Shub-Smale) model.

1.2. Variance of the number of roots. We restrict this subsection to the KSS model. In this case, a few asymptotic results have been proved on variances, when the number m of unknowns tends to ∞ . More precisely, consider the normalized random variable

$$n^X = \frac{N^X}{\sqrt{D}}.$$

It is an obvious consequence of Theorem 12.1 that $E(n^X) = 1$. Let us denote $\sigma_{m,d}^2 = \text{Var}(n^X)$. We have:

THEOREM 12.2. *Assume that the random polynomial system (12.1) is a KSS system with common degree equal to d , $d \geq 2$, and assume that $d \leq d_0 < \infty$, where d_0 is some constant independent of m .*

Then, as $m \rightarrow +\infty$:

- *If $d = 2$, $\text{Var}(n^X) \approx \frac{1}{2} \frac{\log m}{m}$*
- *If $d = 3$, $\text{Var}(n^X) \approx \frac{3}{2} \frac{\log m}{m^2}$*
- *If $d \geq 4$, $\text{Var}(n^X) \approx \frac{K_d}{m^{3 \wedge (d-2)}}$ where $K_4 = \frac{15}{2}$, $K_d = \frac{3465}{64}$ if $d \geq 5$.*

Remark

A simple but interesting corollary of the fact that $\text{Var}(n^X)$ tends to zero as $m \rightarrow +\infty$ is that

$$n^X = \frac{N^X}{d^{m/2}}$$

tends to 1 in probability. Using a similar method, it is also possible to obtain the same type of result if we allow d tend to infinity, slowly enough. For $d \geq 3$ one can find a proof of this weaker result in Wschebor (2005). Notice that the theorem above is more precise, it gives the equivalent of the normalized variance as $m \rightarrow +\infty$.

PROOF OF THEOREM 12.2. We use the same notations as in the beginning of this section. We have:

$$(12.10) \quad \text{Var}(n^X) = \frac{\mathbb{E}((N^X)^2)}{d^m} - 1 = \frac{1}{4} \frac{\mathbb{E}(N^Y(N^Y - 1))}{d^m} + \frac{1}{2d^{m/2}} - 1.$$

We will not perform the detailed computations of the proof, which turn out to be somewhat heavy, but only sketch the main steps and give some more details in the cases $d = 2$ and $d > 5$. The remaining ones are similar and the detailed computations are contained in Wschebor (2007). The general scheme is the following: we show that the first term in the right-hand side of (12.10) has the form $1 + \alpha_m$ where α_m has the speed in the statement. This will be sufficient.

We use Rice formula:

$$(12.11) \quad \mathbb{E}(N^Y(N^Y - 1)) = \int_{S^m \times S^m} \mathbb{E}(|\det(Y'(s))| |\det(Y'(t))| | Y(s) = Y(t) = 0) p_{Y(s), Y(t)}(0, 0) \sigma_{m-1}(ds) \sigma_{m-1}(dt),$$

where:

$$p_{Y(s), Y(t)}(0, 0) = \frac{1}{(2\pi)^m (1 - \langle s, t \rangle)^{2d}}.$$

Conditional expectation.

- Let $s, t \in S^m$ be linearly independent.
- v_2, \dots, v_m pairwise orthogonal, $v_k \perp s, t$ for $k = 2, \dots, m$.
- $B_s = \{v'_1, v_2, \dots, v_m\}$ orthonormal basis of the tangent space $T_s(S^m) = s^\perp$ (in \mathbb{R}^{m+1})
- $B_t = \{v'_1, v_2, \dots, v_m\}$ orthonormal basis of $T_t(S^m) = t^\perp$ (in \mathbb{R}^{m+1})

We express the derivatives $Y'(s)$ and $Y'(t)$ in the basis B_s and B_t respectively and compute the covariances of the pairs of coordinates. This is standard calculation. Once this has been done, we can perform the Gaussian regression of the matrices $Y'(s)$ and $Y'(t)$ on the condition $Y(s) = Y(t) = 0$ and replace the conditional expectation in 12.11 by:

$$d^m \mathbb{E}(|\det(M^s)| |\det(M^t)|),$$

where the matrices M^s, M^t have the following joint law:

- (M_{ik}^s, M_{ik}^t) ($i, k = 1, \dots, m$) are independent bivariate Gaussian centered random vectors,
- for $i = 1, \dots, m; k = 2, \dots, m$,

$$\begin{aligned} \mathbb{E}((M_{ik}^s)^2) &= \mathbb{E}((M_{ik}^t)^2) = 1, \\ \mathbb{E}(M_{ik}^s M_{ik}^t) &= \langle s, t \rangle^{d-1} \end{aligned}$$

•

$$\begin{aligned} \sigma^2 &= \mathbb{E}((M_{i1}^s)^2) = \mathbb{E}((M_{i1}^t)^2) = 1 - \frac{d \langle s, t \rangle^{2d-2}}{1 + \langle s, t \rangle^2 + \dots + \langle s, t \rangle^{2d-2}}, \\ \tau &= \mathbb{E}((M_{i1}^s M_{i1}^t)) \langle s, t \rangle^{d-2} \left[1 - \frac{d}{1 + \langle s, t \rangle^2 + \dots + \langle s, t \rangle^{2d-2}} \right] \end{aligned}$$

Gaussian regression of M_{ik}^t on M_{ik}^s . For $i = 1, \dots, m; k = 2, \dots, m$:

$$M_{ik}^t = M_{ik}^t - \langle s, t \rangle^{d-1} M_{ik}^s + \langle s, t \rangle^{d-1} M_{ik}^s = \zeta_{ik} + \langle s, t \rangle^{d-1} M_{ik}^s,$$

where $\mathbb{E}(\zeta_{ik}^2) = 1 - \langle s, t \rangle^{2d-2}$ and ζ_{ik} is independent of all the rest.

For $i = 1, \dots, m; k = 1$ the regression has the form:

$$M_{i1}^t = M_{i1}^t - \frac{\tau}{\sigma^2} M_{i1}^s + \frac{\tau}{\sigma^2} M_{i1}^s = \zeta_{i1} + \frac{\tau}{\sigma^2} M_{i1}^s,$$

with $\mathbb{E}(\zeta_{i1}^2) = \sigma^2 - \frac{\tau^2}{\sigma^2}$ and again ζ_{i1} is independent of all the rest.

Notice that if $d = 2$, one has $\sigma^2 = \frac{1-\langle s,t \rangle^2}{1+\langle s,t \rangle^2} = -\tau$ which obviously implies that $M_{i1}^t = -M_{i1}^s$ almost surely, for $i = 1, \dots, m$. So, we distinguish in the computation between $d = 2$ and $d > 2$. In the last case, $|\tau| < \sigma^2$.

Case $d = 2$

Replace the above results in the right-hand side of (12.11). We get:

$$(12.12) \quad \mathbb{E}(N^Y(N^Y - 1)) = \frac{2^m}{(2\pi)^m} \int \int_{S^m \times S^m} \frac{1}{(1 - \langle s, t \rangle^4)^{m/2}} \Delta \sigma_m(ds) \sigma_m(dt),$$

where

- $\Delta = (1 - \langle s, t \rangle^2)^{(m-1)/2} \mathbb{E}(|\det(A^s)| |\det(A^t)|)$,
- $(a_{ik}^s, a_{ik}^t)_{i,k=1,\dots,m}$ are centered Gaussian independent pairs,
- $\mathbb{E}((a_{ik}^s)^2) = \mathbb{E}((a_{ik}^t)^2) = 1$ and $\mathbb{E}(a_{ik}^s a_{ik}^t) = \langle s, t \rangle$ for $i = 1, \dots, m; k = 2, \dots, m$,
- $\sigma^2 = \mathbb{E}((a_{i1}^s)^2) = \frac{1-\langle s,t \rangle^2}{1+\langle s,t \rangle^2}$ and $a_{i1}^t = -a_{i1}^s$ for $i = 1, \dots, m$.

We divide the integral in (12.12) into two parts: I_1 is the integral over the pairs $(s, t) \in S^m \times S^m$ such that $|\langle s, t \rangle| \geq \delta_m = \frac{1}{m^\delta}$ where $\delta > 0$ will be chosen afterwards in such a way that this part is negligible, and I_2 over the pairs such that $|\langle s, t \rangle| \leq \delta_m$. The second part will be the relevant one.

Bound for I_1 . We take common factor $\sqrt{\frac{1-\langle s,t \rangle^2}{1+\langle s,t \rangle^2}}$ in the first column of both matrices A^s, A^t . Since each one of the resulting matrices is standard normal, we apply the Cauchy-Schwarz inequality to bound $\mathbb{E}(|\det(A^s)| |\det(A^t)|)$ and use that if G is a GOE $m \times m$ matrix, then

$$\mathbb{E}((\det(G))^2) = m!.$$

So,

$$I_1 \leq \frac{2^m}{(2\pi)^m} \int \int_{|\langle s,t \rangle| \geq \delta_m} \frac{(1 - \langle s, t \rangle^2)^{(m-1)/2}}{(1 - \langle s, t \rangle^4)^{m/2}} m! \frac{1 - \langle s, t \rangle^2}{1 + \langle s, t \rangle^2} \sigma_m(ds) \sigma_m(dt).$$

Now we use the invariance of the integrand and the measure under the orthogonal group and the form of the volume element on S^m . So, rewriting the right-hand side:

$$\begin{aligned} I_1 &\leq \frac{2^m}{(2\pi)^m} \sigma_m(S^m) \sigma_{m-1}(S^{m-1}) m! \int_{|t_0| \geq \delta_m} \frac{(1 - t_0^2)^{(m+1)/2}}{(1 - t_0^4)^{m/2}} (1 - t_0^2)^{\frac{m}{2}-1} dt_0 \\ &\leq \frac{2^m}{(2\pi)^m} \frac{2\pi^{(m+1)/2}}{\Gamma((m+1)/2)} \frac{2\pi^{m/2}}{\Gamma(m/2)} m! \int_{|t_0| \geq \delta_m} (1 - t_0^2)^{(m-1)/2} dt_0 \\ &\leq (\text{const}) 2^m m^{1-\delta} \exp\left[-\frac{1}{2} m^{1-2\delta}\right], \end{aligned}$$

using the usual Stirling's formula and the choice $\delta_m = \frac{1}{m^\delta}$. If $0 < \delta < 1/2$, this shows that $\frac{I_1}{2^m}$ goes to zero faster than any power of m .

Equivalent for I_2 . This is finer than the bound for I_1 . For the conditional expectation, we use the computation of the absolute value of the determinant as the volume of the parallelotope generated by the columns. On account of the invariance of the standard normal distribution under the isometries of the underlying Euclidean space and using a similar expression to the one to obtain the bound for I_1 , we see that Δ in (12.12) can be written as:

$$\Delta = (1 - \langle s, t \rangle^2)^{(m-1)/2} \mathbb{E}(\|\xi\|_m^2) \frac{1 - \langle s, t \rangle^2}{1 + \langle s, t \rangle^2} \left[\prod_{k=1}^{m-1} \mathbb{E}(\|\xi\|_k \|\eta + \frac{\langle s, t \rangle}{(1 - \langle s, t \rangle^2)^{1/2}} \xi\|_k) \right]$$

and

$$(12.13) \quad I_2 = \frac{2^m}{(2\pi)^m} \sigma_m(S^m) \sigma_{m-1}(S^{m-1}) m \prod_{k=1}^{m-1} \mathbb{E}(\|\xi\|_k)^2 \int_{-\delta_m}^{\delta_m} \frac{(1-t_0^2)^{(m-1)/2}}{(1+t_0^2)^{(m+2)/2}} \\ \cdot \prod_{k=1}^{m-1} \frac{\mathbb{E}(\|\xi\|_k \|\eta + \frac{t_0}{(1-t_0^2)^{1/2}} \xi\|_k)}{[\mathbb{E}(\|\xi\|_k)]^2} dt_0.$$

On replacing the various terms here, and making the change of variables $t_0 = \frac{\theta}{\sqrt{m}}$:

$$(12.14) \quad \frac{I_2}{4 \cdot 2^m} = \frac{1}{\sqrt{\pi}} \frac{\Gamma((m+1)/2)}{\Gamma(m/2)} m \int_{-\delta_m}^{\delta_m} \frac{(1-t_0^2)^{(m-1)/2}}{(1+t_0^2)^{(m+2)/2}} \prod_{k=1}^{m-1} \frac{\mathbb{E}(\|\xi\|_k \|\eta + \frac{t_0}{(1-t_0^2)^{1/2}} \xi\|_k)}{[\mathbb{E}(\|\xi\|_k)]^2} dt_0 \\ = C_m \int_{-m^{\frac{1}{2}-\delta}}^{m^{\frac{1}{2}-\delta}} \frac{(1-\theta^2/m)^{(m-1)/2}}{(1+\theta^2/m)^{(m+2)/2}} \prod_{k=1}^{m-1} \frac{\mathbb{E}(\|\xi\|_k \|\eta + \frac{\theta/\sqrt{m}}{(1-\theta^2/m)^{1/2}} \xi\|_k)}{[\mathbb{E}(\|\xi\|_k)]^2} d\theta,$$

where

$$C_m = \frac{1}{\sqrt{\pi m}} \frac{\Gamma((m+1)/2)}{\Gamma(m/2)}.$$

We want to write this expression for $\frac{I_2}{4 \cdot 2^m}$ in the form $1 + \alpha_m$. This is based upon the following lemmas, some of which are well-known and the remaining ones are proved by standard computations.

LEMMA 12.3. *We have the expansion, valid for real z , $z \rightarrow +\infty$ (see Erdelyi et al. (1953), p 57).*

$$\Gamma(z) = e^{-z} z^{z-\frac{1}{2}} (2\pi)^{\frac{1}{2}} \left[1 + \frac{1}{12z} + \frac{1}{288z^2} - \frac{139}{51840z^3} - \dots \right].$$

LEMMA 12.4. *For $c \in \mathbb{R}$ and $k = 1, 2, \dots$ let:*

$$G_k(c) = \mathbb{E}([\eta_1 + c]^2 + \eta_2^2 \dots + \eta_k^2)^{1/2},$$

where the random variables η_1, \dots, η_k are i.i.d. standard normal.

Then:

- (1) $G_k(0) = \sqrt{2} \frac{\Gamma((k+1)/2)}{\Gamma(k/2)}$.
- (2) $G'_k(0) = 0$.
- (3) $0 \leq G'''(c) \leq G''(0) = \frac{1}{k} G_k(0)$ for all c .
- (4) $|G'''_k(c)| \leq 3(\sqrt{2/\pi} + |c|) \mathbb{E}(\|\eta\|_{k-1}^{-3})$

where η is standard normal in \mathbb{R}^{k-1} and the notation (that we use without further explanation) is that $\|\eta\|_h$ denotes Euclidean norm of a vector in \mathbb{R}^h .

This inequality has some interest if $k \geq 5$ since otherwise the right-hand side is infinite.

The proof of parts (1),(2),(3) are given in Lemma 12.13 below.

LEMMA 12.5. $C_m = \frac{1}{\sqrt{2\pi}} \left[1 - \frac{1}{4m} + \frac{1}{32m^2} + \frac{5}{128m^3} + O\left(\frac{1}{m^4}\right) \right]$

LEMMA 12.6. *For $k = 1, 2, \dots$; j an integer, set $m_{kj} = \mathbb{E}(\|\xi\|_k^j)$. Then,*

$$m_{kj} = 2^{\frac{j}{2}} \frac{\Gamma((j+k)/2)}{\Gamma(k/2)}$$

LEMMA 12.7. *For fixed integer j , we have:*

$$m_{kj} = k^{\frac{j}{2}} \left[1 + \frac{1}{k} \left(\frac{j^2}{4} - \frac{j}{2} \right) + \frac{1}{k^2} \left(-\frac{j^3}{12} + \frac{j^2}{4} - \frac{j}{6} \right) \right. \\ \left. + \frac{1}{k^3} \left(\frac{j^4}{24} - \frac{5}{24} j^3 - \frac{j^2}{12} + \frac{j}{6} \right) + O(1/k^4) \right]$$

where the bound in "O" depends on j .

With these ingredients and some additional effort, one can prove that:

$$\begin{aligned}
 \frac{I_2}{4 \cdot 2^m} &= \left(1 - \frac{1}{4m} + O\left(\frac{1}{m^2}\right)\right) \left[\int_{-m^{\frac{1}{2}-\delta}}^{m^{\frac{1}{2}-\delta}} \frac{1}{\sqrt{2\pi}} e^{-\frac{\theta^2}{2}} \left(1 + \frac{\theta^2 \log m}{2m}\right) d\theta + O(1/m) \right] \\
 (12.15) \quad &= \left(1 - \frac{1}{4m} + O\left(\frac{1}{m^2}\right)\right) \left[\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{\theta^2}{2}} \left(1 + \frac{\theta^2 \log m}{2m}\right) d\theta + O(1/m) \right] \\
 &= 1 + \frac{1}{2} \frac{\log m}{m} + O(1/m)
 \end{aligned}$$

This shows that

$$\text{Var}(n^X) \approx \frac{1}{2} \frac{\log m}{m}$$

and finishes the computation when $d = 2$.

Case $d \geq 3$. Instead of the general formula in the case $d = 2$, we have:

$$(12.16) \quad \mathbb{E}(N^Y(N^Y - 1)) = \frac{d^m}{(2\pi)^m} \int \int_{S^m \times S^m} \frac{1}{(1 - \langle s, t \rangle^{2d})^{m/2}} \Delta \sigma_m(ds) \sigma_m(dt),$$

where

$$\begin{aligned}
 \Delta &= (1 - \langle s, t \rangle^{2d-2})^{(m-1)/2} (\sigma^4 - \tau^2)^{1/2} \\
 &\quad \cdot \left[\prod_{k=1}^{m-1} \mathbb{E}(\|\xi\|_k \|\eta + \frac{\langle s, t \rangle^{d-1}}{(1 - \langle s, t \rangle^{2d-2})^{1/2}} \xi\|_k) \mathbb{E}(\|\xi\|_m \|\eta + \frac{\tau}{(\sigma^4 - \tau^2)^{1/2}} \xi\|_m) \right].
 \end{aligned}$$

In each factor, ξ, η are independent standard normal vectors in \mathbb{R}^k . The proof that the part of the integral corresponding to the pairs $(s, t) \in S^m \times S^m$ such that $|\langle s, t \rangle| \geq \delta_m$, with $\delta_m = 1/m^\delta$, $0 < \delta < 1/2$ is negligible is similar to the case $d = 2$ (take into account that $\sigma^2 \leq (\text{const})(1 - \langle s, t \rangle^2)$) and the question is again the equivalent of the integral over the set $|\langle s, t \rangle| < \delta_m$. The overall computation is similar to the case $d = 2$, with some minor differences.

Using as above the invariance under isometries, we have to consider the new integral:

$$\frac{I_2}{4d^m} = \frac{\Gamma((m+1)/2)}{\sqrt{\pi}\Gamma(m/2)} \int_{|t_0| < \delta_m} \frac{(1 - t_0^{2d-2})^{(m-1)/2}}{(1 - t_0^{2d})^{m/2}} (\sigma^4 - \tau^2)^{1/2} (1 - t_0^2)^{\frac{m}{2}-1} H_m dt_0,$$

where

- $H_m = \prod_{k=1}^m \left[\frac{1}{(G_k(0))^2} \mathbb{E}(\|\xi\|_k \|\eta + \alpha_k \xi\|_k) \right]$
- $\alpha_k = \frac{t_0^{d-1}}{(1 - t_0^{2d-2})^{1/2}}$ if $k = 1, \dots, m-1$ and $\alpha_m = \frac{\tau}{(\sigma^4 - \tau^2)^{1/2}}$.

Again we perform the change of variables $t_0 = \theta/\sqrt{m}$:

$$\frac{I_2}{4d^m} = C_m \int_{-m^{1/2-\delta}}^{m^{1/2-\delta}} \frac{(1 - (\frac{\theta^2}{m})^{d-1})^{(m-1)/2}}{(1 - (\frac{\theta^2}{m})^d)^{m/2}} (\sigma^4 - \tau^2)^{1/2} (1 - \frac{\theta^2}{m})^{\frac{m}{2}-1} H_m d\theta,$$

with

$$H_m = \left(1 + \frac{1}{2} \frac{\tau^2}{\sigma^4 - \tau^2} c_m\right) \prod_{k=1}^{m-1} \left[1 + \frac{1}{2} \frac{(\theta^2/m)^{d-1}}{1 - (\theta^2/m)^{d-1}} c_k\right],$$

where

$$c_k = \frac{m k_3}{k m_{k1}}.$$

LEMMA 12.8.

$$c_k = 1 + \frac{1}{k} - \frac{1}{4k^2} - \frac{55}{48k^3} + O(1/k^4).$$

LEMMA 12.9. *Let*

$$A_m = \frac{(1 - (\frac{\theta^2}{m})^{d-1})^{(m-1)/2}}{(1 - (\frac{\theta^2}{m})^d)^{m/2}} (1 - \frac{\theta^2}{m})^{\frac{m}{2}-1}.$$

Then,

$$A_m = \exp \left[-\frac{\theta^2}{2} + \frac{1}{m}(\theta^2 - \frac{1}{4}\theta^4) + \frac{1}{m^2}(\frac{1}{2}\theta^4 - \frac{1}{6}\theta^6) + \frac{1}{m^3}(\frac{1}{3}\theta^6 - \frac{1}{8}\theta^8) \right. \\ \left. - \frac{1}{2} \frac{\theta^{2d-2}}{m^{d-2}} + \frac{1}{m^{d-1}}(\frac{1}{2}\theta^{2d-2} + \frac{1}{2}\theta^{2d}) - \frac{\theta^{4d-4}}{4m^{2d-3}} + O(1/m^{8\delta}) \right]$$

LEMMA 12.10. *Let* $x = t_0^2 \leq \frac{1}{m^{2\delta}} \rightarrow 0$.

Then,

$$\tau^2 = x^{d-2} [(d-1)^2 + d^2(x^2 + x^{2d} - 2x^{d+1} + 2x^{d+2}) \\ - 2d(1-d)(-x + x^d - x^{d+1} + x^{2d})] + O(x^8) \\ \sigma^4 - \tau^2 = 1 - (d-1)^2 x^{d-2} + \\ d(d-2)x^{d-1} [2 - x - x^{d-1} + 2x^d - x^{d+1} - x^{2d-1}] + O(x^8).$$

Case $d > 5$. We have

- $\tau^2 = O_u(\frac{1}{m^{8\delta}})$
- $\sigma^4 - \tau^2 = 1 + O_u(\frac{1}{m^{8\delta}})$
- $H_m = 1 + O_u(1/m^{10\delta-1})$

We choose δ so that $10\delta - 1 > 3$.

$$\frac{I_2}{4d^m} = \frac{1}{\sqrt{2\pi}} \left[1 - \frac{1}{4m} + \frac{1}{32m^2} + \frac{5}{128m^3} + O(\frac{1}{m^4}) \right] [1 + O_u(1/m^{10\delta-1})] \\ \cdot \int_{-m^{\frac{1}{2}-\delta}}^{m^{\frac{1}{2}-\delta}} \exp \left[-\frac{\theta^2}{2} + \frac{1}{m}(\theta^2 - \frac{1}{4}\theta^4) + \frac{1}{m^2}(\frac{1}{2}\theta^4 - \frac{1}{6}\theta^6) + \frac{1}{m^3}(\frac{1}{3}\theta^6 - \frac{1}{8}\theta^8) + O_u(\frac{1}{m^{10\delta-1}}) \right] d\theta.$$

Notice (key point) that excluding the first term, all other terms in the exponent are $O_u(1)$ as $m \rightarrow +\infty$, that is, are uniformly small if m is large enough.

Expanding and using the moments of the standard normal distribution (up to order 12), we get, for $d > 5$:

$$\frac{I_2}{4d^m} = 1 + \frac{3465}{64} \frac{1}{m^3} + O(\frac{1}{m^{10\delta-1}}),$$

so that $\text{Var}(n^X) \approx \frac{3465}{64} \frac{1}{m^3}$.

The cases $d = 3, 4, 5$ can be treated in a similar way. □

2. More general models

The probability law of the Shub-Smale model defined in section 1 has the simplifying property of being invariant under the orthogonal group of the underlying Euclidean space \mathbb{R}^m . In this section we present the extension of formula (12.4) to general systems which share the same invariance property. This paragraph follows Azaïs and Wschebor (2005 b).

We require the polynomial random fields X_i ($i = 1, \dots, m$) to be centered, Gaussian, independent and their covariances

$$r^{X_i}(s, t) = \mathbb{E}(X_i(s)X_i(t))$$

to be invariant under orthogonal linear transformation of \mathbb{R}^m , i.e. $r^{X_i}(Us, Ut) = r^{X_i}(s, t)$ for any orthogonal transformation U and any pair $s, t \in \mathbb{R}^m$. This implies in particular that the

coefficients $a_j^{(i)}$ remain independent for different i 's but can now be correlated from one j to another for the same value of i . It is easy to check that this implies that for each $i = 1, \dots, m$, the covariance $r^{X_i}(s, t)$ is a function of the triple $(\langle s, t \rangle, \|s\|^2, \|t\|^2)$. It is somewhat harder but can also be proved (see Spivak (1979)) that this function is in fact a polynomial with real coefficients, say $Q^{(i)}$

$$(12.17) \quad r^{X_i}(s, t) = Q^{(i)}(\langle s, t \rangle, \|s\|^2, \|t\|^2),$$

satisfying the symmetry condition

$$(12.18) \quad Q^{(i)}(u, v, w) = Q^{(i)}(u, w, v)$$

A natural question is which are the polynomials $Q^{(i)}$ such that the function in the right-hand side of (12.17) is a covariance, that is, non-negative definite. A simple way to construct a class of covariances of this type is to take

$$(12.19) \quad Q^{(i)}(u, v, w) = P(u, vw),$$

where P is a polynomial in two variables with non-negative coefficients. In fact, the functions $(s, t) \rightsquigarrow \langle s, t \rangle$ and $(s, t) \rightsquigarrow \|s\|^2 \|t\|^2$ are covariances and the set of covariances is closed under linear combinations with non-negative coefficients as well as under multiplication, so that $P(\langle s, t \rangle, \|s\|^2 \|t\|^2)$ is also the covariance of some random field.

The situation becomes simpler if one considers only functions of the scalar product, i.e.

$$Q(u, v, w) = \sum_{k=0}^d c_k u^k.$$

The necessary and sufficient condition for $\sum_{k=0}^d c_k \langle s, t \rangle^k$ to be a covariance is that $c_k \geq 0 \forall k = 0, 1, \dots, d$. In that case, it is the covariance of the random field $X(t) := \sum_{\|j\| \leq d} a_j t^j$ where the a_j 's are centered, Gaussian, independent random variables, $\text{Var}(a_j) = c_{\|j\|} \frac{\|j\|!}{j!}$ (The proof of this is left to the reader). The Shub-Smale model is the special case corresponding to the choice $c_k = \binom{d}{k}$.

The general description of the polynomial covariances which are invariant under the action of the orthogonal group, is in Kostlan (2002), part II.

We now state the extension of the Shub-Smale formula to the general case.

THEOREM 12.11. *Assume that the X_i are independent centered Gaussian polynomial random fields with covariances $r^{X_i}(s, t) = Q^{(i)}(\langle s, t \rangle, \|s\|^2, \|t\|^2)$ ($i = 1, \dots, m$).*

Let us denote by $Q_u^{(i)}, Q_w^{(i)}, Q_{uv}^{(i)}, \dots$ the partial derivatives of $Q^{(i)}$. We assume that $Q^{(i)}(x, x, x)$ and $Q_u^{(i)}(x, x, x)$ do not vanish for $x \geq 0$. Set

$$q_i(x) := \frac{Q_u^{(i)}}{Q^{(i)}}$$

$$r_i(x) := \frac{Q^{(i)}(Q_{uu}^{(i)} + 2Q_{uv}^{(i)} + 2Q_{uw}^{(i)} + 4Q_{vw}^{(i)}) - (Q_u^{(i)} + Q_v^{(i)} + Q_w^{(i)})^2}{(Q^{(i)})^2}$$

where the functions in the right-hand sides are always computed at the triple (x, x, x) . Put

$$h_i(x) := 1 + x \frac{r_i(x)}{q_i(x)}.$$

Then, for all Borel sets V we have

$$(12.20) \quad \mathbb{E}(N^X(V)) = \kappa_m \int_V \left(\prod_{i=1}^m q_i(\|t\|^2) \right)^{1/2} E_h(\|t\|^2) dt.$$

In this formula,

$$E_h(x) := \mathbb{E} \left(\left(\sum_{i=1}^m h_i(x) \xi_i^2 \right)^{1/2} \right),$$

where ξ_1, \dots, ξ_m are i.i.d. standard normal in \mathbb{R} and

$$\kappa_m = \frac{1}{\sqrt{2\pi}} \frac{\Gamma(m/2)}{\pi^{m/2}}.$$

PROOF. .

Let us put $K_j = E(\|\eta_j\|)$ with η_j standard normal in \mathbb{R}^j . An elementary computations gives:

$$K_m = \sqrt{2} \frac{\Gamma((m+1)/2)}{\Gamma(m/2)}.$$

We define the integral

$$J_m := \int_0^{+\infty} \frac{\rho^{m-1}}{(1+\rho^2)^{(m+1)/2}} d\rho = \sqrt{\pi/2} \frac{1}{K_m}$$

that will appear later on. Consider the normalized Gaussian fields

$$Z_i(t) := \frac{X_i(t)}{(Q^{(i)}(\|t\|^2, \|t\|^2, \|t\|^2))^{1/2}}$$

which have variance 1. Denote $Z(t) = (Z_1(t), \dots, Z_m(t))^T$. Applying Rice Formula for the expectation of the number of zeros of Z :

$$E(N^X(V)) = E(N^Z(V)) = \int_V E(|\det(Z'(t))| | Z(t) = 0) \frac{1}{(2\pi)^{\frac{m}{2}}} dt,$$

where $Z'(t) := [Z'_1(t) \ \vdots \ \dots \ \vdots \ Z'_m(t)]$ is the matrix obtained by concatenation of the vectors $Z'_1(t), \dots, Z'_m(t)$. Note that since $E(Z_i^2(t))$ is constant, it follows that $E(Z_i(t) \frac{\partial Z_i}{\partial t_j}(t)) = 0$ for all $i, j = 1, \dots, m$. Since the field is Gaussian this implies that $Z_i(t)$ and $Z'_i(t)$ are independent and given that the coordinate fields Z_1, \dots, Z_m are independent, one can conclude that for each t , $Z(t)$ and $Z'(t)$ are independent. So

$$(12.21) \quad E(N^X(V)) = E(N^Z(V)) = \frac{1}{(2\pi)^{\frac{m}{2}}} \int_V E(|\det(Z'(t))|) dt.$$

A straightforward computation shows that the (α, β) - entry, $\alpha, \beta = 1, \dots, m$, in the covariance matrix of $Z'_i(t)$ is

$$E\left(\frac{\partial Z_i}{\partial t_\alpha}(t) \frac{\partial Z_i}{\partial t_\beta}(t)\right) = \frac{\partial^2}{\partial s_\alpha \partial t_\beta} r^{Z_i}(s, t) |_{s=t} = r_i(\|t\|^2) t_\alpha t_\beta + q_i(\|t\|^2) \delta_{\alpha\beta},$$

where $\delta_{\alpha,\beta}$ denotes the Kronecker symbol. This can be rewritten as

$$\text{Var}(Z'_i(t)) = q_i I_m + r_i t t^T,$$

where the functions in the right-hand side are to be computed at the point $\|t\|^2$. Let U be the orthogonal transformation of \mathbb{R}^m that gives the coordinates in a basis with first vector $\frac{t}{\|t\|}$, we get

$$\text{Var}(UZ'_i(t)) = \text{Diag}((r_i \cdot \|t\|^2 + q_i), q_i, \dots, q_i),$$

so that

$$\text{Var}\left(\frac{UZ'_i(t)}{\sqrt{q_i}}\right) = \text{Diag}(h_i, 1, \dots, 1).$$

Put now

$$T_i := \frac{UZ'_i(t)}{\sqrt{q_i}}$$

and set

$$T := [T_1 \ \vdots \ \dots \ \vdots \ T_m]$$

We have

$$(12.22) \quad |\det(Z'(t))| = |\det(T)| \prod_{i=1}^m q_i^{1/2}.$$

Now, we write

$$T = \begin{bmatrix} W_1 \\ \cdots \\ \cdots \\ \cdots \\ W_m \end{bmatrix},$$

where the W_i are random row vectors. Because of the properties of independence of all the entries of T , we know that :

- W_2, \dots, W_m are independent standard normal vectors in \mathbb{R}^m
- W_1 is independent from the other $W_i, i \geq 2$, and has a centered Gaussian distribution with variance matrix $Diag(h_1, \dots, h_m)$.

Now $E(|\det(T)|)$ is calculated as the expectation of the volume of the parallelotope generated by W_1, \dots, W_m in \mathbb{R}^m . That is,

$$|\det(T)| = \|W_1\| \prod_{j=2}^m d(W_j, S_{j-1}),$$

where S_{j-1} denotes the subspace of \mathbb{R}^m generated by W_1, \dots, W_{j-1} and d denotes the Euclidean distance. Using the invariance under isometries of the standard normal distribution of \mathbb{R}^m we know that, conditioning on W_1, \dots, W_{j-1} , the projection $P_{S_{j-1}^\perp}(W_j)$ of W_j on the orthogonal S_{j-1}^\perp of S_{j-1} has a distribution which is standard normal on the space S_{j-1}^\perp which is of dimension $m - j + 1$ with probability 1. Thus $E(d(W_j, S_{j-1})|W_1, \dots, W_{j-1}) = K_{m-j+1}$. By successive conditionings on W_1, W_1, W_2 etc... , we get:

$$E(|\det(T)|) = E\left(\left(\sum_{i=1}^m h_i(x)\xi_i^2\right)^{1/2}\right) \times \prod_{j=1}^{m-1} K_j,$$

where ξ_1, \dots, ξ_m are i.i.d. standard normal in \mathbb{R} . Using (12.22) and (12.21) we obtain (12.20) . \square

2.1. Examples. 1.- Let $Q^{(i)}(u, v, w) = Q^{i}(u)$ for some polynomial Q . We get:

$$q_i(x) = l_i q(x) = l_i \frac{Q'(x)}{Q(x)}, \quad h_i(x) = h(x) = 1 - x \frac{Q'^2(x) - Q(x)Q''(x)}{Q(x)Q'(x)}.$$

Applying formula (12.20) with $V = \mathbb{R}^m$, and using polar coordinates:

$$(12.23) \quad E(N^X) = \frac{2}{\sqrt{\pi}} \frac{\Gamma((m+1)/2)}{\Gamma(m/2)} \sqrt{l_1 \dots l_m} \int_0^\infty \rho^{m-1} q(\rho^2)^{m/2} \sqrt{h(\rho^2)} d\rho.$$

2.- If in Example 1.- we put $Q(u) = 1 + u$ we get the Shub-Smale model. Replacing in (12.23) an elementary computation reproduces (12.4).

3.- A simple variant of Shub & Smale theorem corresponds to taking $Q^{(i)}(u) = 1 + u^d$ for all $i = 1, \dots, m$ (here all the X_i 's have the same law). Even though in this case the derivative $Q^{(i)}(u)$ vanishes at zero, the reader can easily check that the conclusion of Theorem 12.11 remains valid, and

$$q(x) = q_i(x) = \frac{du^{d-1}}{1 + u^d} ; \quad h(x) = h_i(x) = \frac{d}{1 + u^d},$$

$$E(N^X) = \sqrt{\frac{2}{\pi}} K_m \int_0^{+\infty} \frac{\rho^{md-1}}{(1 + \rho^{2d})^{(m+1)/2}} d\rho = d^{(m-1)/2},$$

which differs by a constant factor from the analogous Shub & Smale result for $(1 + u)^d$ which is $d^{m/2}$.

4.- Linear systems with a quadratic perturbation

Consider linear systems with a quadratic perturbation

$$X_i(s) = \xi_i + \langle \eta_i, s \rangle + \zeta_i \|s\|^2,$$

where the $\xi_i, \zeta_i, \eta_i, i = 1, \dots, m$ are independent and standard normal in \mathbb{R}, \mathbb{R} and \mathbb{R}^m respectively. This corresponds to the covariance $r^{X_i}(s, t) = 1 + \langle s, t \rangle + \|s\|^2 \|t\|^2$.

If there is no quadratic perturbation, it is obvious that the number of roots is almost surely equal to 1.

For the perturbed system, applying Theorem 12.11 and performing the computations required in this case, we obtain:

$$q(x) = \frac{1}{1+x+x^2}; \quad r(x) = \frac{4}{1+x+x^2} - \frac{(1+2x)^2}{(1+x+x^2)^2}; \quad h(x) = \frac{1+4x+x^2}{1+x+x^2}$$

and

$$E(N^X) = \frac{H_m}{J_m} \quad \text{with} \quad H_m = \int_0^{+\infty} \frac{\rho^{m-1}(1+4\rho^2+\rho^4)^{\frac{1}{2}}}{(1+\rho^2+\rho^4)^{\frac{m}{2}+1}} d\rho.$$

An elementary computation shows that $E(N^X) = o(1)$ as $m \rightarrow +\infty$ (see the next example for a more precise behavior). In other words, the probability that the perturbed system has no solution tends to 1 as $m \rightarrow +\infty$.

5.- More general perturbed systems

Let us consider the covariances given by the polynomials

$$Q^i(u, v, w) = Q(u, v, w) = 1 + 2u^d + (vw)^d.$$

This corresponds to adding a perturbation depending on the product of the norms of s, t to the modified Shub-Smale systems considered in our first example. We know that for the unperturbed system, one has $E(N^X) = d^{\frac{m-1}{2}}$. Notice that the factor 2 in Q has only been added for computational convenience and does not modify the random variable N^X of the unperturbed system. For the perturbed system, we get

$$q(x) = \frac{2dx^{d-1}}{(1+x^d)^2}; \quad r(x) = \frac{2d(d-1)x^{d-2}}{(1+x^d)^2}; \quad h(x) = d.$$

Therefore,

$$\begin{aligned} (12.24) \quad E(N^X) &= \sqrt{\frac{2}{\pi}} K_m \int_0^{+\infty} \rho^{m-1} \left(\frac{2d\rho^{2(d-1)}}{(1+\rho^{2d})^2} \right)^{\frac{m}{2}} \sqrt{d} d\rho \\ &= \sqrt{\frac{2}{\pi}} K_m 2^{m/2} d^{\frac{m+1}{2}} \int_0^{+\infty} \frac{\rho^{md-1}}{(1+\rho^{2d})^m} d\rho. \end{aligned}$$

The integral can be evaluated by an elementary computation and we obtain

$$E(N^X) = 2^{-\frac{m-2}{2}} d^{\frac{m-1}{2}},$$

which shows that the mean number of zeros is reduced by the perturbation at a geometrical rate as m grows.

6.-Polynomial in the scalar product, real roots

Consider again the case in which the polynomials $Q^{(i)}$ are all equal and the covariances depend only on the scalar product, i.e. $Q^{(i)}(u, v, w) = Q(u)$. We assume further that the roots of Q , that we denote $-\alpha_1, \dots, -\alpha_d$, are real ($0 < \alpha_1 \leq \dots \leq \alpha_d$). We get

$$q(x) = \sum_{h=1}^d \frac{1}{x + \alpha_h}; \quad r(x) = \sum_{h=1}^d \frac{1}{(x + \alpha_h)^2}; \quad h(x) = \frac{1}{q_i(x)} \sum_{h=1}^d \frac{\alpha_h}{(x + \alpha_h)^2}.$$

It is easy now to write an upper bound for the integrand in (12.20) and compute the remaining integral, thus obtaining the inequality

$$\mathbb{E}(N^X) \leq \sqrt{\frac{\alpha_d}{\alpha_1}} d^{m/2},$$

which is sharp if $\alpha_1 = \dots = \alpha_d$.

If we further assume that $d = 2$, with no loss of generality $Q(u)$ has the form $Q(u) = (u+1)(u+\alpha)$ with $\alpha \in [0, 1]$. Replacing q by $\frac{1}{x+1} + \frac{1}{x+\alpha}$ in formula (12.23) we get:

$$(12.25) \quad \mathbb{E}(N^X) = \sqrt{2/\pi} K_m \int_0^\infty \rho^{m-1} \left(\frac{1}{1+\rho^2} + \frac{1}{\alpha+\rho^2} \right)^{(m-1)/2} \left(\frac{1}{(1+\rho^2)^2} + \frac{\alpha}{(\alpha+\rho^2)^2} \right)^{1/2} d\rho.$$

One can compute the limit of the right-hand side as $\alpha \rightarrow 0$. For this purpose, notice that the function $\alpha \rightarrow \frac{\alpha}{(\alpha+\rho^2)^2}$ attains its maximum at $\alpha = \rho^2$ and is dominated by $\frac{1}{4\rho^2}$. We divide the integral in the right-hand side of (12.25) into two parts, setting for some $\delta > 0$

$$I_{\delta,\alpha} := \int_0^\delta \rho^{m-1} \left(\frac{1}{1+\rho^2} + \frac{1}{\alpha+\rho^2} \right)^{(m-1)/2} \left(\frac{1}{(1+\rho^2)^2} + \frac{\alpha}{(\alpha+\rho^2)^2} \right)^{1/2} d\rho,$$

and

$$J_{\delta,\alpha} := \int_\delta^{+\infty} \rho^{m-1} \left(\frac{1}{1+\rho^2} + \frac{1}{\alpha+\rho^2} \right)^{(m-1)/2} \left(\frac{1}{(1+\rho^2)^2} + \frac{\alpha}{(\alpha+\rho^2)^2} \right)^{1/2} d\rho.$$

By dominated convergence,

$$J_{\delta,\alpha} \rightarrow \int_\delta^{+\infty} \left(\frac{2\rho^2+1}{\rho^2+1} \right)^{(m-1)/2} \frac{d\rho}{1+\rho^2},$$

as $\alpha \rightarrow 0$. On the other hand

$$I_{\delta,\alpha}^- \leq I_{\delta,\alpha} \leq I_{\delta,\alpha}^+$$

where

$$(12.26) \quad I_{\delta,\alpha}^- := \int_0^\delta \left(\frac{\rho^2}{1+\rho^2} + \frac{\rho^2}{\alpha+\rho^2} \right)^{(m-1)/2} \frac{\sqrt{\alpha}}{\rho^2+\alpha} d\rho \\ = \int_0^{\delta/\alpha} \left(\frac{\alpha z^2}{1+\alpha z^2} + \frac{\alpha z^2}{\alpha(z^2+1)} \right)^{(m-1)/2} \frac{dz}{z^2+1} \rightarrow J_m,$$

as $\alpha \rightarrow 0$, and

$$(12.27) \quad I_{\delta,\alpha}^+ := \int_0^\delta \left(\frac{\rho^2}{1+\rho^2} + \frac{\rho^2}{\alpha+\rho^2} \right)^{(m-1)/2} \left(\frac{1}{1+\rho^2} + \frac{\sqrt{\alpha}}{\rho^2+\alpha} \right) d\rho \\ \rightarrow \int_0^\delta \left(\frac{2\rho^2+1}{\rho^2+1} \right)^{(m-1)/2} \frac{d\rho}{1+\rho^2} + J_m,$$

as $\alpha \rightarrow 0$. Since δ is arbitrary, the integral in the right-hand side of (12.27) can be chosen arbitrarily small. Using the identity $K_m J_m = \sqrt{\pi}/2$, we get

$$\mathbb{E}(N^X) \rightarrow v := 1 + \frac{1}{J_m} \int_0^{+\infty} \left(\frac{2\rho^2+1}{\rho^2+1} \right)^{(m-1)/2} \frac{d\rho}{1+\rho^2},$$

as $\alpha \rightarrow 0$. Since $\frac{2\rho^2}{\rho^2+1} < \frac{2\rho^2+1}{\rho^2+1} < 2$:

$$1 + 2^{(m-1)/2} < v < 1 + \frac{2^{(m-1)/2} \pi}{J_m}.$$

3. Non-centered systems (smoothed analysis)

The aim of this section is to remove the hypothesis that the coefficients have zero expectation. Let us start with a non-random system

$$(12.28) \quad P_i(t) = 0 \quad (i = 1, \dots, m),$$

and perturb it with a polynomial noise $\{X_i(t) : i = 1, \dots, m\}$, that is, we consider the new system:

$$P_i(t) + X_i(t) = 0 \quad (i = 1, \dots, m)$$

What can one say about the number of roots of the new system? Of course, to obtain results on $E(N^{P+X})$ we need a certain number of hypotheses both on the noise X and the polynomial “signal” P , especially the relation between the size of P and the probability distribution of X .

Some of these hypotheses are of technical nature, allowing to perform the computations. Beyond this, roughly speaking, Theorem 12.12 below says that if the relation signal over noise is neither too big nor too small, in a sense that we make precise later on, then there exist positive constants $C, \theta, 0 < \theta < 1$ such that

$$(12.29) \quad E(N^{P+X}) \leq C \theta^m E(N^X).$$

Inequality (12.29) becomes of interest if the starting non-random system (12.28) has a large number of roots, possibly infinite, and m is large. In this situation, the effect of adding polynomial noise is a reduction at a geometric rate of the expected number of roots, as compared to the centered case. In formula (12.29), $E(N^X)$ can be computed or estimated using the results in sections 1 and 2 and bounds for the constants C, θ can be explicitly deduced from the hypotheses.

Before the statement we need to introduce some additional notations and hypotheses: H_1 and H_2 concern only the noise, H_3 and H_4 include relations between noise and signal.

The noise will correspond to polynomials $Q^{(i)}(u, v, w) = \sum_{k=0}^{d_i} c_k^{(i)} u^k$, $c_k^{(i)} \geq 0$, considered in section 2, i.e. the covariances are only function of the scalar product. Also, each polynomial $Q^{(i)}$ has effective degree d_i , i.e.

$$c_{d_i}^{(i)} > 0 \quad (i = 1, \dots, m).$$

and does not vanish for $u \geq 0$, which amounts to saying that for each t the distribution of $X_i(t)$ does not degenerate.

An elementary calculation then shows that for each polynomial $Q^{(i)}$, as $u \rightarrow +\infty$:

$$(12.30) \quad q_i(u) \sim \frac{d_i}{1+u},$$

$$(12.31) \quad h_i(u) \sim \frac{c_{d_i-1}^{(i)}}{d_i c_{d_i}^{(i)}} \cdot \frac{1}{1+u}.$$

Since we are interested in the large m asymptotics, the polynomials P, Q can vary with m and we will require somewhat more than relations (12.30) and (12.31), as specified in the following hypotheses:

H_1) h_i is independent of i ($i = 1, \dots, m$) (but may vary with m). We put $h = h_i$.

H_2) There exist positive constants D_i, E_i ($i = 1, \dots, m$) and \underline{q} such that

$$(12.32) \quad 0 \leq D_i - (1+u)q_i(u) \leq \frac{E_i}{1+u}, \quad \text{and} \quad (1+u)q_i(u) \geq \underline{q}$$

for all $u \geq 0$, and moreover

$$\max_{1 \leq i \leq m} D_i, \quad \max_{1 \leq i \leq m} E_i$$

are bounded by constants \bar{D}, \bar{E} respectively, which are independent of m . \underline{q} is also independent of m .

Also, there exist positive constants \underline{h}, \bar{h} such that

$$(12.33) \quad \underline{h} \leq (1+u)h(u) \leq \bar{h}$$

for $u \geq 0$.

Notice that the auxiliary functions q_i, r_i, h ($i = 1, \dots, m$) will also vary with m . To simplify somewhat the notations we are dropping the parameter m in P, Q, q_i, r_i, h . However, in H_2) the constants \underline{h}, \bar{h} do not depend on m . One can check that these conditions imply that $h(u) \geq 0$ when $u \geq 0$.

Let us now describe the second set of hypotheses. Let P be a polynomial in m real variables with real coefficients having degree d and Q a polynomial in one variable with non-negative coefficients, having also degree d , $Q(u) = \sum_{k=0}^d c_k u^k$. We assume that Q does not vanish on $u \geq 0$ and $c_d > 0$. Define

$$H(P, Q) := \sup_{t \in \mathbb{R}^m} \left\{ (1 + \|t\|) \cdot \left\| \nabla \left(\frac{P}{\sqrt{Q(\|t\|^2)}} \right) (t) \right\| \right\}$$

$$K(P, Q) := \sup_{t \in \mathbb{R}^m \setminus \{0\}} \left\{ (1 + \|t\|^2) \cdot \left| \frac{\partial}{\partial \rho} \left(\frac{P}{\sqrt{Q(\|t\|^2)}} \right) (t) \right| \right\}$$

where $\frac{\partial}{\partial \rho}$ denotes the derivative in the direction defined by $\frac{t}{\|t\|}$, at each point $t \neq 0$.

For $r > 0$, put:

$$L(P, Q, r) := \inf_{\|t\| \geq r} \frac{P(t)^2}{Q(\|t\|^2)}.$$

One can check by means of elementary computations, that for each pair P, Q as above, one has

$$H(P, Q) < \infty, \quad K(P, Q) < \infty.$$

With these notations, we introduce the following hypotheses on the systems P, Q , as m grows:

H_3)

$$(12.34) \quad A_m = \frac{1}{m} \cdot \sum_{i=1}^m \frac{H^2(P_i, Q^{(i)})}{i} = o(1) \quad \text{as } m \rightarrow +\infty$$

$$(12.35) \quad B_m = \frac{1}{m} \cdot \sum_{i=1}^m \frac{K^2(P_i, Q^{(i)})}{i} = o(1) \quad \text{as } m \rightarrow +\infty.$$

H_4) There exist positive constants r_0, l such that if $r \geq r_0$:

$$L(P_i, Q^{(i)}, r) \geq l \quad \text{for all } i = 1, \dots, m.$$

THEOREM 12.12 (Armentano and Wschebor (2007)). *Under the hypotheses H_1, H_2, H_3, H_4 , one has*

$$(12.36) \quad \mathbb{E}(N^{P+X}) \leq C \theta^m \mathbb{E}(N^X)$$

where C, θ are positive constants, $0 < \theta < 1$.

3.1. Remarks on the statement of Theorem 12.12.

1.- It is obvious that our problem does not depend on the order in which the equations

$$P_i(t) + X_i(t) = 0 \quad (i = 1, \dots, m)$$

appear. However, conditions (12.34) and (12.35) in hypothesis H_3) do depend on the order. One can restate them saying that there exists an order $i = 1, \dots, m$ on the equations, such that (12.34) and (12.35) hold true.

2.- Condition H_3) can be interpreted as a bound on the quotient signal over noise. In fact, it concerns the gradient of this quotient. In (12.35) appears the radial derivative, which happens to decrease faster as $\|t\| \rightarrow \infty$ than the other components of the gradient.

Clearly, if $H(P_i, Q^{(i)}), K(P_i, Q^{(i)})$ are bounded by fixed constants, (12.34) and (12.35) hold true. Also, some of them may grow as $m \rightarrow +\infty$ provided (12.34) and (12.35) remain satisfied.

- 3.- Hypothesis H_4) goes in the opposite direction: for large values of $\|t\|$ we need a lower bound of the relation signal over noise.
- 4.- A result of the type of Theorem 12.12 can not be obtained without putting some restrictions on the relation signal over noise. In fact consider the system

$$(12.37) \quad P_i(t) + \sigma X_i(t) = 0 \quad (i = 1, \dots, m)$$

where σ is a positive real parameter. As $\sigma \downarrow 0$ the expected value of the number of roots of (12.37) tends to the number of roots of $P_i(t) = 0$, ($i = 1, \dots, m$), for which no a priori bound is available. In this case, the relation signal over noise tends to infinity. On the other hand, if we let $\sigma \rightarrow +\infty$, the relation signal over noise tends to zero and the expected number of roots will tend to $E(N^X)$.

PROOF OF THEOREM 12.12 We follow the same lines of the proof of Theorem 12.11

Let

$$Z_j(t) = \frac{P_j(t) + X_j(t)}{\sqrt{Q^{(j)}(\|t\|^2)}} \quad (j = 1, \dots, m)$$

and

$$Z = (Z_1, \dots, Z_m)^T.$$

Clearly,

$$N^{P+X}(V) = N^Z(V)$$

for any subset V of \mathbb{R}^m .

$\{Z_j(t) : t \in \mathbb{R}^m\}$ ($j = 1, \dots, m$) are independent centered Gaussian process,

$$E(Z_j^2(t)) = 1$$

for all $j = 1, \dots, m$ and all $t \in \mathbb{R}^m$. This implies that $Z_j(t)$ and $\nabla Z_j(t)$ are independent for each $t \in \mathbb{R}^m$. We apply Rice formula to compute $E(N^Z(V))$, that is:

$$E(N^Z(V)) = \int_V E(|\det(Z'(t))| | Z(t) = 0) \cdot p_{Z(t)}(0) dt,$$

Using the independence between $Z'(t)$ and $Z(t)$, one gets:

$$(12.38) \quad E(N^Z(V)) = \int_V E(|\det(Z'(t))|) \cdot \frac{1}{(2\pi)^{m/2}} \cdot \exp\left[-\frac{1}{2} \left(\frac{P_1(t)^2}{Q^{(1)}(\|t\|^2)} + \dots + \frac{P_m(t)^2}{Q^{(m)}(\|t\|^2)} \right)\right] dt$$

and our main problem consists in the evaluation of $E(|\det(Z'(t))|)$.

As in the centered case, we have:

$$\text{Cov}\left(\frac{\partial Z_i}{\partial t_\alpha}(t), \frac{\partial Z_j}{\partial t_\beta}(t)\right) = \delta_{ij} [r_i(\|t\|^2)t_\alpha t_\beta + q_i(\|t\|^2)\delta_{\alpha\beta}]$$

for $i, j, \alpha, \beta = 1, \dots, m$.

For each $t \neq 0$, let U_t be an orthogonal transformation of \mathbb{R}^m that takes the first element of the canonical basis into the unit vector $\frac{t}{\|t\|}$. Then

$$(12.39) \quad \text{Var}\left(\frac{U_t \nabla Z_j(t)}{\sqrt{q_j(\|t\|^2)}}\right) = \text{Diag}(h(\|t\|^2), 1, \dots, 1)$$

where we denote the gradient $\nabla Z_j(t)$ as a column vector.

$\text{Diag}(\lambda_1, \dots, \lambda_m)$ denotes the $m \times m$ diagonal matrix with elements $\lambda_1, \dots, \lambda_m$ in the diagonal. So we can write

$$\frac{U_t \nabla Z_j(t)}{\sqrt{q_j(\|t\|^2)}} = \zeta_j + \alpha_j \quad (j = 1, \dots, m)$$

where ζ_j is a Gaussian, centered random vector in \mathbb{R}^m having covariance given by (12.39), ζ_1, \dots, ζ_m are independent and α_j is the non-random vector

$$(12.40) \quad \alpha_j = \frac{U_t \nabla \left(\frac{P_j(t)}{\sqrt{Q^{(j)}(\|t\|^2)}} \right)}{\sqrt{q_j(\|t\|^2)}} (\alpha_{1j}, \dots, \alpha_{mj}) \quad (j = 1, \dots, m).$$

We denote by T the $m \times m$ random matrix having columns $\zeta_j + \alpha_j$, ($j = 1, \dots, m$). We have

$$|\det(Z'(t))| = |\det(T)| \cdot \prod_{i=1}^m (q_i(\|t\|^2))^{1/2}$$

so that

$$(12.41) \quad E(|\det(Z'(t))|) = E(|\det(T)|) \cdot \prod_{i=1}^m (q_i(\|t\|^2))^{1/2}.$$

Denote by η_1, \dots, η_m the columns of T , i.e.

$$\eta_j = \zeta_j + \alpha_j \quad (j = 1, \dots, m)$$

where the ζ_{ij} are Gaussian centered independent and

$$\begin{aligned} \text{Var}(\zeta_{ij}) &= 1 \quad \text{for } i = 2, \dots, m; j = 1, \dots, m \\ \text{Var}(\zeta_{1j}) &= h(\|t\|^2) \quad \text{for } j = 1, \dots, m. \end{aligned}$$

Proceeding as in the centered case to compute the volume of the associated parallelepiped, we obtain the bound:

$$(12.42) \quad E(|\det T|) \leq \sqrt{h(\|t\|^2)} \cdot \prod_{j=1}^m E(\|\xi_j + c_j(t)\|_j)$$

where $\|\cdot\|_j$ denotes Euclidean norm in \mathbb{R}^j , ($\|\cdot\| = \|\cdot\|_m$), ξ_j is a random vector with normal standard distribution in \mathbb{R}^j and $c_j(t)$ is a non-random vector in \mathbb{R}^j having norm

$$\|\tilde{\alpha}_j\|, \quad j = 1, \dots, m.$$

where $\tilde{\alpha}_j = (\alpha_{1j}/\sqrt{h(\|t\|^2)}, \alpha_{2j}, \dots, \alpha_{mj})^T$ and the α_{ij} are given in (12.40). We denote

$$\gamma_j(c) = E(\|\xi_j + c\|_j)$$

where $c \in \mathbb{R}^j$ is non-random. We have (see the auxiliary Lemma 12.13 after this proof):

$$\gamma_j(c) \leq \left(1 + \|c\|_j^2 \frac{1}{2j}\right) \gamma_j(0).$$

Replacing in (12.42) and using (12.38), (12.41) we get:

$$(12.43) \quad E(N^Z) \leq \frac{1}{(2\pi)^{m/2}} L_m \cdot \int_{\mathbb{R}^m} \left\{ \sqrt{h(\|t\|^2)} \left(\prod_{i=1}^m q_i(\|t\|^2) \right)^{1/2} \cdot \exp \left[-\frac{1}{2} \sum_{i=1}^m \frac{P_i(t)^2}{Q^{(i)}(\|t\|^2)} + \frac{1}{2} \sum_{j=1}^m \|c_j(t)\|_j^2 \frac{1}{j} \right] \right\},$$

where

$$L_m = \prod_{i=1}^m E(\|\xi_i\|_i) = \frac{1}{\sqrt{2\pi}} 2^{(m+1)/2} \Gamma\left(\frac{m+1}{2}\right).$$

Our final task is to obtain an adequate bound for the integral in (12.43). For $j = 1, \dots, m$ (use H_1):

$$|\tilde{\alpha}_{1j}| = \frac{1}{\sqrt{h(\|t\|^2) q_j(\|t\|^2)}} \cdot \left| \frac{\partial}{\partial \rho} \frac{P_j(\|t\|^2)}{\sqrt{Q^{(j)}(\|t\|^2)}} \right| \leq \frac{1}{\sqrt{h} \underline{q}} K(P_j, Q^{(j)})$$

and

$$\|\alpha_j\| = \frac{\left\| \nabla \left(\frac{P_j(t)}{\sqrt{Q^{(j)}(\|t\|^2)}} \right) \right\|}{\sqrt{q_j(\|t\|^2)}} \leq \frac{1}{\sqrt{q}} H(P_j, Q^{(j)}).$$

Then, if we bound $\|\tilde{\alpha}_j\|^2$ by:

$$\|\tilde{\alpha}_j\|^2 \leq |\tilde{\alpha}_{1j}|^2 + \|\alpha_j\|^2$$

we obtain

$$\|\tilde{\alpha}_j\|^2 \leq \frac{1}{\underline{h}\underline{q}} K^2(P_j, Q^{(j)}) + \frac{1}{\underline{q}} H^2(P_j, Q^{(j)})$$

which implies

$$\sum_{j=1}^m \|c_j\|_j^2 \cdot \frac{1}{j} \leq \frac{1}{\underline{q}} m A_m + \frac{1}{\underline{h}\underline{q}} m B_m.$$

Replacing in (12.43) we get the bound:

$$\mathbb{E}(N^Z) \leq s_m H_m$$

where

$$s_m = \left(\frac{\bar{h}}{\underline{h}} \right)^{1/2} \cdot \exp \left(\frac{1}{2} \left(\frac{1}{\underline{q}} m A_m + \frac{1}{\underline{h}\underline{q}} m B_m \right) \right) = e^{o(m)} \quad (\text{as } m \rightarrow +\infty)$$

and

$$(12.44) \quad H_m = \frac{1}{\pi^{(m+1)/2}} \Gamma \left(\frac{m}{2} \right) \cdot \int_{\mathbb{R}^m} \left(\prod_{i=1}^m q_i(\|t\|^2) \right)^{1/2} \sqrt{h(\|t\|^2)} \mathbb{E}(\|\xi_m\|) e^{-\frac{1}{2} \sum_{i=1}^m \frac{P_i(t)^2}{Q^{(i)}(\|t\|^2)}} dt.$$

The integrand in (12.44) is the same as in the expectation in the centered case, except for the exponential, which will help for large values of $\|t\|$.

Let us write H_m as

$$H_m = H_m^{(1)}(r) + H_m^{(2)}(r)$$

where $H_m^{(1)}(r)$ corresponds to integrating on $\|t\| \leq r$ and $H_m^{(2)}(r)$ on $\|t\| > r$ instead of the whole \mathbb{R}^m in formula (12.44). We first choose r large enough so that the condition in hypothesis H_4) is satisfied. Then

$$(12.45) \quad H_m^{(2)}(r) \leq e^{-lm/2} \mathbb{E}(N^X).$$

We now turn to $H_m^{(1)}(r)$. We have, bounding the exponential in the integrand by 1 and using hypothesis H_2):

$$(12.46) \quad H_m^{(1)}(r) \leq \frac{1}{\pi^{(m+1)/2}} \Gamma \left(\frac{m}{2} \right) \bar{h}^{1/2} \mathbb{E}(\|\xi_m\|) \left(\prod_{i=1}^m D_i^{1/2} \right) \sigma_{m-1} \int_0^r \frac{\rho^{m-1}}{(1+\rho^2)^{(m+1)/2}} d\rho,$$

where σ_{m-1} is the $(m-1)$ -dimensional area measure of S^{m-1} . The integral in the right hand side is bounded by

$$\frac{\pi}{2} \left(\frac{r^2}{1+r^2} \right)^{\frac{m-1}{2}}.$$

Again using H_2), we have the lower bound:

$$\begin{aligned} \mathbb{E}(N^X) &\geq \\ &\geq \frac{1}{\pi^{(m+1)/2}} \Gamma \left(\frac{m}{2} \right) \underline{h}^{1/2} \mathbb{E}(\|\xi_m\|) \int_0^{+\infty} \left[\prod_{i=1}^m \left(\frac{D_i}{1+\|t\|^2} - \frac{E_i}{(1+\|t\|^2)^2} \right)^{1/2} \right] \frac{1}{(1+\|t\|^2)^{1/2}} dt \\ &= \frac{1}{\pi^{(m+1)/2}} \Gamma \left(\frac{m}{2} \right) \underline{h}^{1/2} \mathbb{E}(\|\xi_m\|) \left(\prod_{i=1}^m D_i^{1/2} \right) \sigma_{m-1} \int_0^{+\infty} \frac{\rho^{m-1}}{(1+\rho^2)^{(m+1)/2}} \prod_{i=1}^m \left(1 - \frac{F_i}{1+\rho^2} \right)^{1/2} d\rho \end{aligned}$$

where we have denoted $F_i = E_i/D_i$, ($i = 1, \dots, m$), which implies

$$F_i \leq \frac{\max_k E_k}{\underline{q}} \quad (i = 1, \dots, m).$$

Choose now $\tau > 1$ large enough, to have

$$\lambda^2 = \frac{r^2}{1+r^2} < 1 - \frac{\max_{1 \leq i \leq m} F_i}{1+\tau^2 r^2} = \nu^2$$

and we get for $E(N^X)$ the lower bound:
(12.47)

$$E(N^X) \geq \frac{1}{\pi^{(m+1)/2}} \Gamma\left(\frac{m}{2}\right) \underline{h}^{1/2} E(\|\xi_m\|) \left(\prod_{i=1}^m D_i^{1/2}\right) \sigma_{m-1} \nu^m \int_{\tau r}^{+\infty} \frac{\rho^{m-1}}{(1+\rho^2)^{(m+1)/2}} d\rho.$$

Now, compare (12.46) and (12.47) and use the elementary equivalence, for each $a > 0$

$$\int_a^{+\infty} \frac{\rho^{m-1}}{(1+\rho^2)^{(m+1)/2}} d\rho \simeq \sqrt{\frac{\pi}{2m}} \quad \text{as } m \rightarrow +\infty.$$

We get:

$$H_m^1(r) \leq C_1 \lambda_1^m E(N^X)$$

where C_1 is a positive constant and $\lambda/\nu < \lambda_1 < 1$. This implies

$$E(N^{P+X}) \leq s_m \left[C_1 \lambda_1^m + e^{-\ell m/2} \right] E(N^X) \leq C \theta^m E(N^X)$$

for positive constants C, θ , $0 < \theta < 1$.

More precisely, we can obtain first θ and then, m_0 and the constant C , in such a way that whenever $m \geq m_0$, inequality (12.36) holds true. The reader can verify, following step by step the proof, that a possible choice is the following:

Choose r_0 and ℓ from H_4),

$$\theta_1 = \max \left\{ \frac{r_0}{\sqrt{r_0^2 + \frac{1}{2}}}, e^{-\ell/2} \right\}, \quad \theta = \frac{1 + \theta_1}{2}$$

Put $F_i = E_i/D_i$, ($i = 1, \dots, m$) and $\bar{F} = \max\{F_1, \dots, F_m\}$. From the hypotheses, one has $\bar{F} \leq \bar{E}/\underline{q}$. Let $\tau > 0$ such that:

$$\frac{\bar{F}}{1 + \tau^2 r_0^2} < \frac{1}{2} \frac{1}{1 + r_0^2}.$$

Choose m_0 (using H_3) so that if $m \geq m_0$ one has:

$$e^{\frac{1}{2} \left[\frac{m A_m}{\underline{q}} + \frac{m B_m}{\underline{h} \underline{q}} \right]} \theta_1^m \sqrt{m} \leq \theta^m$$

$$\pi \left(\frac{\tau^2 r_0^2}{1 + \tau^2 r_0^2} \right)^{(m-1)/2} < \frac{e^{-2}}{\sqrt{m}}.$$

Then, (12.36) is satisfied for $m \geq m_0$, with

$$C = 30. \frac{\bar{h}}{\underline{h}} \frac{\sqrt{1+r_0^2}}{r_0}.$$

□

3.2. Auxiliary Lemma.

LEMMA 12.13. Let $\gamma : \mathbb{R}^k \rightarrow \mathbb{R}$, $k \geq 1$ be defined as

$$\gamma(c) = E(\|\xi + c\|)$$

where ξ is a standard normal random vector in \mathbb{R}^k , and $c \in \mathbb{R}^k$. Then

$$(i) \quad \gamma(0) = \sqrt{2} \frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})}.$$

(ii) γ is a function of $\|c\|$ and verifies:

$$(12.48) \quad \gamma(c) \leq \gamma(0) \left(1 + \frac{1}{2k} \|c\|^2\right).$$

PROOF. (i) follows on integrating in polar coordinates.

(ii) That γ is a function of $\|c\|$ is a consequence of the invariance of the distribution of ξ under the orthogonal group of \mathbb{R}^k . For $k = 1$, (12.48) follows from the exact computation

$$\gamma(c) = \sqrt{2/\pi} e^{-\frac{1}{2}c^2} + c \int_{-c}^c \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

and a Taylor expansion at $c = 0$, which gives

$$\gamma(c) \leq \sqrt{2/\pi} \left(1 + \frac{1}{2} c^2\right).$$

For $k \geq 2$, we write

$$\gamma(c) = E \left([(\xi_1 + a)^2 + \xi_2^2 + \dots + \xi_k^2]^{1/2} \right) = G(a)$$

where $a = \|c\|$ and ξ_1, \dots, ξ_k are independent standard normal variables. Differentiating under the expectation sign, we get:

$$G'(a) = E \left(\frac{\xi_1 + a}{[(\xi_1 + a)^2 + \xi_2^2 + \dots + \xi_k^2]^{1/2}} \right),$$

so that $G'(0) = 0$ due to the symmetry of the distribution of ξ .

One can differentiate formally once more, obtaining:

$$(12.49) \quad G''(a) = E \left(\frac{\xi_2^2 + \dots + \xi_k^2}{[(\xi_1 + a)^2 + \xi_2^2 + \dots + \xi_k^2]^{3/2}} \right).$$

For the validity of equality (12.49) for $k \geq 3$ one can use that if $d \geq 2$, $\frac{1}{\|x\|}$ is integrable in \mathbb{R}^d with respect to the Gaussian standard measure. For $k = 2$ one must be more careful but it holds true and left to the reader. The other ingredient of the proof is that one can verify that G'' has a maximum at $a = 0$. Hence, on applying Taylor's formula, we get

$$G(a) \leq G(0) + \frac{1}{2} a^2 G''(0).$$

Check that $G''(0) = \frac{\sqrt{2}}{k} \frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})}$ which, together with (i) gives:

$$\frac{G''(0)}{G(0)} = \frac{1}{k}$$

which implies (ii). □

3.3. Examples.

• *Shub-Smale noise* Assume that the noise follows the Shub-Smale model. If the degrees d_i are uniformly bounded, one can easily check that H_1) and H_2) are satisfied.

For the signal, we give two simple examples. Let

$$P_i(t) = \|t\|^{d_i} - r^{d_i},$$

where d_i is even and $r > 0$ remains bounded as m varies. One has:

$$\begin{aligned} \frac{\partial}{\partial \rho} \left(\frac{P_i}{\sqrt{Q^{(i)}}} \right) (t) &= \frac{d_i \|t\|^{d_i-1} + d_i r^{d_i} \|t\|}{(1 + \|t\|^2)^{\frac{d_i}{2}+1}} \leq \frac{d_i(1 + r^{d_i})}{(1 + \|t\|^2)^{3/2}} \\ \nabla \left(\frac{P_i}{\sqrt{Q^{(i)}}} \right) (t) &= \frac{d_i \|t\|^{d_i-2} + d_i r^{d_i}}{(1 + \|t\|^2)^{\frac{d_i}{2}+1}} \cdot t \end{aligned}$$

which implies

$$\left\| \nabla \left(\frac{P_i}{\sqrt{Q^{(i)}}} \right) (t) \right\| \leq \frac{d_i(1 + r^{d_i})}{(1 + \|t\|^2)^{3/2}}.$$

So, since the degrees d_1, \dots, d_m are uniformly bounded, H_3) follows. H_4) also holds under the same hypothesis.

Notice that an interest in this choice of the P_i 's lies in the fact that obviously the system $P_i(t) = 0$ ($i = 1, \dots, m$) has infinite roots (all points in the sphere of radius r centered at the origin are solutions), but the expected number of roots of the perturbed system is geometrically smaller than the Shub-Smale expectation \sqrt{D} , when m is large.

Our second example of signal is as follows. Let T be a polynomial of degree d in one variable that has d distinct real roots. Define:

$$P_i(t_1, \dots, t_m) = T(t_i) \quad (i = 1, \dots, m).$$

One can easily check that the system verifies our hypotheses, so that there exist C, θ positive constants, $0 < \theta < 1$ such that

$$E(N^{P+X}) \leq C \theta^m d^{m/2}$$

where we have used the Kostlan-Shub-Smale formula. On the other hand, it is clear that $N^P = d^m$ so that the diminishing effect of the noise on the number of roots can be observed.

• $Q^{(i)} = Q$, *Only real roots* Assume that all the $Q^{(i)}$'s are equal, $Q^{(i)} = Q$, and Q has only real roots. Since Q does not vanish on $u \geq 0$, all the roots should be strictly negative, say $-\alpha_1, \dots, -\alpha_d$ where $0 < \alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_d$. With no loss of generality, we may assume that $\alpha_1 \geq 1$.

We will assume again that the degree d (which can vary with m) is bounded by a fixed constant \bar{d} as well as the roots $\alpha_k \leq \bar{\alpha}$ ($k = 1, \dots, d$) for some constant $\bar{\alpha}$. One verifies (12.32), choosing $D_i = d$, $E_i = d \cdot \max_{1 \leq k \leq d} (\alpha_k - 1)$. Similary, a direct computation gives (12.33).

Again let us consider the particular example of signals:

$$P_i(t) = \|t\|^{d_i} - r^{d_i}$$

where d_i is even and r is positive and remains bounded as m varies.

$$\left| \frac{\partial}{\partial \rho} \left(\frac{P_i}{\sqrt{Q^{(i)}}} \right) \right| \leq d_i (\bar{\alpha} + r^{d_i}) \frac{1}{(1 + \|t\|^2)^{3/2}}$$

so that $K(P_i, Q^{(i)})$ is uniformly bounded. A similar computation shows that $H(P_i, Q^{(i)})$ is uniformly bounded. Finally, it is obvious that

$$L(P_i, Q^{(i)}, r) \geq \left(\frac{1}{1 + \bar{\alpha}} \right)^{\bar{d}}$$

for $i = 1, \dots, m$ and any $r \geq 1$. So the conclusion of Theorem 12.12 can be applied.

One can check that the second signal in the previous example also works with respect to this noise.

•*Some other examples.* Assume that the covariance of the noise has the form of Example 4.1.1. Q is a polynomial in one variable having degree ν and positive coefficients, $Q(u) = \sum_{k=0}^{\nu} b_k u^k$. Q may depend on m , as well as the exponents l_1, \dots, l_m . Notice that $d_i = \nu \cdot l_i$ ($i = 1, \dots, m$). One easily verifies that H_1) is satisfied.

We assume that the coefficients b_0, \dots, b_{ν} of the polynomial Q verify the conditions

$$b_k \leq \frac{\nu - k + 1}{k} b_{k-1} \quad (k = 1, 2, \dots, \nu)$$

Moreover, l_1, \dots, l_m, ν are bounded by a constant independent of m and there exist positive constants \underline{b}, \bar{b} such that

$$\underline{b} \leq b_0, b_1, \dots, b_{\nu} \leq \bar{b}.$$

Under these conditions, one can check that H_2) holds true, with $D_i = d_i$ ($i = 1, \dots, m$).

For the relation signal over noise, conditions are similar to the previous example.

Notice that already if $\nu = 2$, and we choose for Q the fixed polynomial:

$$Q(u) = 1 + 2a u + b u^2$$

with $0 < a \leq 1$, $\sqrt{b} > a \geq b > 0$, then the conditions in this example are satisfied, but the polynomial Q (hence Q^{d_i}) does not have real roots, so that it is not included in Example 3.3.

4. Systems having a law invariant under orthogonal transformations and translations

In this section we assume that $X_i : \mathbb{R}^m \rightarrow \mathbb{R}$, $i = 1, \dots, m$ are independent centered Gaussian random fields with covariances having the form

$$(12.50) \quad r^{X_i}(s, t) = \gamma_i(\|t - s\|^2), \quad (i = 1, \dots, m).$$

We will assume that γ_i is of class \mathcal{C}^2 and, with no loss of generality, that $\gamma_i(0) = 1$.

The computation of the expectation of the number of roots belonging to a Borel set V can be done using Rice formula (12.21), obtaining:

$$(12.51) \quad \mathbb{E}(N^X(V)) = (2\pi)^{-m/2} \mathbb{E}(|\det(X'(0))|) \lambda_m(V)$$

To prove (12.51) we take into account that the law of the random field is invariant under translations and for each t , $X(t)$ and $X'(t)$ are independent. Compute, for $i, \alpha, \beta = 1, \dots, m$

$$\mathbb{E} \left(\frac{\partial X_i}{\partial t_{\alpha}}(0) \frac{\partial X_i}{\partial t_{\beta}}(0) \right) = \left. \frac{\partial^2 r^{X_i}}{\partial s_{\alpha} \partial t_{\beta}} \right|_{t=s} = -2\gamma'_i(0) \delta_{\alpha\beta},$$

which implies, using a similar method to the one in the proof of Theorem 12.1 :

$$\mathbb{E}(|\det(X'(0))|) = \frac{1}{\sqrt{\pi}} 2^m \Gamma((m+1)/2) \prod_{i=1}^m |\gamma'_i(0)|^{1/2}$$

and replacing in (12.51)

$$(12.52) \quad \mathbb{E}(N^X(V)) = \frac{1}{\sqrt{\pi}} \left(\frac{2}{\pi}\right)^{m/2} \Gamma((m+1)/2) \left[\prod_{i=1}^m |\gamma'_i(0)|^{1/2} \right] \lambda_m(V).$$

Next, let us consider the variance. One can prove that under certain additional technical conditions, the variance of the normalized number of roots:

$$n^X(V) = \frac{N^X(V)}{\mathbb{E}(N^X(V))}$$

- which has obviously mean value equal to 1- grows exponentially when the dimension m tends to infinity. This establishes a striking difference with respect to the results in section 1. In other words, one should expect to have large fluctuations of $n^X(V)$ around its mean for systems having large m .

Our additional requirements are the following:

- 1) All the γ_i coincide, $\gamma_i = \gamma$, $i = 1, \dots, m$,

2) the function γ is such that $(s, t) \rightsquigarrow \gamma(\|t - s\|^2)$ is a covariance for all dimensions m .

It is well known (Schoenberg, 1938) that γ satisfies 2) and $\gamma(0) = 1$ if and only if there exists a probability measure G on $[0, +\infty)$ such that

$$(12.53) \quad \gamma(x) = \int_0^{+\infty} e^{-xw} G(dw) \quad \text{for all } x \geq 0.$$

THEOREM 12.14 (Azaïs-Wschebor, 2005b). *Let $r^{X_i}(s, t) = \gamma(\|t - s\|^2)$ for $i = 1, \dots, m$ where γ is of the form (12.53). We assume further that*

(1) *G is not concentrated at a single point and*

$$\int_0^{+\infty} x^2 G(dx) < \infty.$$

(2) *$\{V_m\}_{m=1,2,\dots}$ is a sequence of Borel sets, $V_m \subset \mathbb{R}^m$, $\lambda_m(\partial V_m) = 0$ and there exist two positive constants δ, Δ such that for each m , V_m contains a ball with radius δ and is contained in a ball with radius Δ .*

Then,

$$(12.54) \quad \text{Var}(n^X(V_m)) \rightarrow +\infty,$$

exponentially fast as $m \rightarrow +\infty$.

PROOF. To compute the variance of $N^X(V)$ we start as in the case of the KSS model:

$$(12.55) \quad \text{Var}(N^X(V)) = \mathbb{E}(N^X(V)(N^X(V) - 1)) + \mathbb{E}(N^X(V)) - [\mathbb{E}(N^X(V))]^2,$$

so that to prove (12.54), it suffices to show that

$$(12.56) \quad \frac{\mathbb{E}(N^X(V)(N^X(V) - 1))}{[\mathbb{E}(N^X(V))]^2} \rightarrow +\infty$$

exponentially fast as $m \rightarrow +\infty$. The denominator in (12.56) is given by formula (12.52). For the numerator, we can apply Rice formula for the second order factorial moment:

$$(12.57) \quad \mathbb{E}(N^X(V)(N^X(V) - 1)) \\ = \iint_{V \times V} \mathbb{E}(|\det(X'(s)) \det(X'(t))| | X(s) = X(t) = 0) p_{X(s), X(t)}(0, 0) ds dt,$$

Next, we compute the ingredients of the integrand in (12.57). Because of invariance under translations, the integrand is a function of $\tau = t - s$. We denote with τ_1, \dots, τ_m the coordinates of τ .

The Gaussian density is immediate:

$$(12.58) \quad p_{X(s), X(t)}(0, 0) = \frac{1}{(2\pi)^m} \frac{1}{[1 - \gamma^2(\|\tau\|^2)]^{m/2}}.$$

Let us turn to the conditional expectation in (12.57). We put

$$\mathbb{E}(|\det(X'(s)) \det(X'(t))| | X(s) = X(t) = 0) = \mathbb{E}(|\det(A^s) \det(A^t)|),$$

where $A^s = ((A_{i\alpha}^s))$, $A^t = ((A_{i\alpha}^t))$ are $m \times m$ random matrices having as joint - Gaussian - distribution the conditional distribution of the pair $X'(s), X'(t)$ given that $X(s) = X(t) = 0$. So, to describe this joint distribution we must compute the conditional covariances of the elements of the matrices $X'(s)$ and $X'(t)$ given the condition $\mathcal{C} : \{X(s) = X(t) = 0\}$. This is easily done using standard regression formulas:

$$\mathbb{E}\left(\frac{\partial X_i}{\partial s_\alpha}(s) \frac{\partial X_i}{\partial s_\beta}(s) | \mathcal{C}\right) = \frac{\partial^2 r}{\partial s_\alpha \partial t_\beta} \Big|_{t=s} - \frac{1}{1 - (r(s, t))^2} \frac{\partial r}{\partial s_\alpha}(s, t) \frac{\partial r}{\partial s_\beta}(s, t) \\ \mathbb{E}\left(\frac{\partial X_i}{\partial s_\alpha}(s) \frac{\partial X_i}{\partial t_\beta}(t) | \mathcal{C}\right) = \frac{\partial^2 r}{\partial s_\alpha \partial t_\beta}(s, t) + \frac{1}{1 - (r(s, t))^2} \frac{\partial r}{\partial s_\alpha}(s, t) \frac{\partial r}{\partial t_\beta}(s, t) r(s, t).$$

Replacing in our case, we obtain

$$(12.59) \quad \mathbb{E}(A_{i\alpha}^s A_{i\beta}^s) = \mathbb{E}(A_{i\alpha}^t A_{i\beta}^t) = -2\gamma'(0)\delta_{\alpha\beta} - 4\frac{\gamma'^2 \tau_\alpha \tau_\beta}{1 - \gamma^2},$$

$$(12.60) \quad \mathbb{E}(A_{i\alpha}^s A_{i\beta}^t) = -4\gamma'' \tau_\alpha \tau_\beta - 2\gamma' \delta_{\alpha\beta} - 4\frac{\gamma\gamma'^2 \tau_\alpha \tau_\beta}{1 - \gamma^2},$$

and for every $i \neq j$:

$$\mathbb{E}(A_{i\alpha}^s A_{j\beta}^s) = \mathbb{E}(A_{i\alpha}^t A_{j\beta}^t) = \mathbb{E}(A_{i\alpha}^s A_{j\beta}^t) = 0,$$

where $\gamma = \gamma(\|\tau\|^2)$, $\gamma' = \gamma'(\|\tau\|^2)$, $\gamma'' = \gamma''(\|\tau\|^2)$.

Take now an orthonormal basis of \mathbb{R}^m having the unit vector $\frac{\tau}{\|\tau\|}$ as first element. Then the variance $(2m) \times (2m)$ matrix of the pair A_i^s, A_i^t - the i -th rows of A^s and A^t respectively - takes the following form:

$$T = \left[\begin{array}{cccc|cccc} U_0 & \cdots & \cdot & \cdot & U_1 & \cdots & \cdot & \cdot \\ \cdot & V_0 & \cdots & \cdot & \cdot & V_1 & \cdots & \cdot \\ \cdot & \cdot & \ddots & \cdot & \cdot & \cdot & \ddots & \cdot \\ \cdot & \cdot & \cdots & V_0 & \cdot & \cdot & \cdots & V_1 \\ \hline U_1 & \cdots & \cdot & \cdot & U_0 & \cdots & \cdot & \cdot \\ \cdot & V_1 & \cdots & \cdot & \cdot & V_0 & \cdots & \cdot \\ \cdot & \cdot & \ddots & \cdot & \cdot & \cdot & \ddots & \cdot \\ \cdot & \cdot & \cdots & V_1 & \cdot & \cdot & \cdots & V_0 \end{array} \right],$$

where

$$\begin{aligned} U_0 &= U_0(\|\tau\|^2) = -2\gamma'(0) - 4\frac{\gamma'^2 \|\tau\|^2}{1 - \gamma^2}; \\ V_0 &= -2\gamma'(0); \\ U_1 &= U_1(\|\tau\|^2) = -4\gamma'' \|\tau\|^2 - 2\gamma' - 4\frac{\gamma\gamma'^2 \|\tau\|^2}{1 - \gamma^2}; \\ V_1 &= V_1(\|\tau\|^2) = -2\gamma'; \end{aligned}$$

and there are zeros outside the diagonals of each one of the four blocks. Let us perform a second regression of $A_{i\alpha}^t$ on $A_{i\alpha}^s$, that is, write the orthogonal decompositions

$$A_{i\alpha}^t = B_{i\alpha}^{t,s} + C_\alpha A_{i\alpha}^s \quad (i, \alpha = 1, m),$$

where $B_{i\alpha}^{t,s}$ is centered Gaussian independent of the matrix A^s , and

$$\begin{aligned} \text{For } \alpha = 1, \quad C_1 &= \frac{U_1}{U_0}, \quad \text{Var}(B_{i1}^{t,s}) = U_0 \left(1 - \frac{U_1^2}{U_0^2}\right); \\ \text{For } \alpha > 1, \quad C_\alpha &= \frac{V_1}{V_0}, \quad \text{Var}(B_{i\alpha}^{t,s}) = V_0 \left(1 - \frac{V_1^2}{V_0^2}\right). \end{aligned}$$

Conditioning we have :

$$\mathbb{E}(|\det(A^s)| |\det(A^t)|) = \mathbb{E}[|\det(A^s)| \mathbb{E}(|\det((B_{i\alpha}^{t,s} + C_\alpha A_{i\alpha}^s)_{i,\alpha=1,\dots,m})| |A^s)]$$

with obvious notations. For the inner conditional expectation, we can proceed in the same way as we did in the proof of Theorem 12.11 to compute the determinant, obtaining a product of expectations of Euclidean norms of non-centered Gaussian vectors in \mathbb{R}^k for $k = 1, \dots, m$. Now we use the well-known inequality

$$\mathbb{E}(\|\xi + v\|) \geq \mathbb{E}(\|\xi\|)$$

valid for ξ standard normal in \mathbb{R}^k and v any vector in \mathbb{R}^k , and it follows that

$$\mathbb{E}(|\det(A^s)| |\det(A^t)|) \geq \mathbb{E}(|\det(A^s)|) \mathbb{E}(|\det(B^{t,s})|).$$

Since the elements of A^s (resp. $B^{t,s}$) are independent, centered Gaussian with known variance, we obtain:

$$\mathbb{E} |\det(A^s) \det(A^t)| \geq U_0 V_0^{m-1} \left(1 - \frac{U_1^2}{U_0^2}\right)^{1/2} \left(1 - \frac{V_1^2}{V_0^2}\right)^{(m-1)/2} L_m^2.$$

Going back to (12.56) and on account of (12.52) and (12.57) we have

$$(12.61) \quad \frac{\mathbb{E}(N^X(V) (N^X(V) - 1))}{\mathbb{E}(N^X(V))^2} \geq (\lambda_m(V))^{-2} \iint_{V \times V} ds dt \left[\frac{1 - V_1^2 V_0^{-2}}{1 - \gamma^2} \right]^{m/2} H(\|\tau\|^2).$$

Let us put $V = V_m$ in (12.61) and study the integrand in the right hand side. The function

$$H(x) = \left(\frac{U_0^2(x) - U_1^2(x)}{V_0^2 - V_1^2(x)} \right)^{1/2}$$

is continuous for $x > 0$. Let us show that it does not vanish if $x > 0$.

It is clear that $U_1^2 \leq U_0^2$ on applying the Cauchy-Schwarz inequality to the pair of variables A_{i1}^s, A_{i1}^t . The equality holds if and only if the variables A_{i1}^s, A_{i1}^t are linearly dependent. This would imply that the distribution - in \mathbb{R}^4 - of the random vector

$$\zeta := (X(s), X(t), \partial_1 X(s), \partial_1 X(t))$$

would degenerate for $s \neq t$ (we have denoted ∂_1 differentiation with respect to the first coordinate). We will show that this is not possible. Notice first that for each $w > 0$, the function

$$(s, t) \rightsquigarrow e^{-\|t-s\|^2 w}$$

is positive definite, hence the covariance of a centered Gaussian stationary field defined on \mathbb{R}^m , say $\{Z^w(t) : t \in \mathbb{R}^m\}$ whose spectral measure has the non-vanishing density:

$$f^w(x) = (2\pi)^{-m/2} (2w)^{-m/2} \exp\left(-\frac{\|x\|^2}{4w}\right) \quad (x \in \mathbb{R}^m).$$

The field $\{Z^w(t) : t \in \mathbb{R}^m\}$ satisfies the conditions of Proposition 3.1 of Azais & Wschebor (2004) so that the distribution of the 4-tuple

$$\zeta^w := (Z^w(s), Z^w(t), \partial_1 Z^w(s), \partial_1 Z^w(t))$$

does not degenerate for $s \neq t$. On account of (12.53) we have,

$$\text{Var}(\zeta) = \int_0^{+\infty} \text{Var}(\zeta^w) G(dw),$$

where integration of the matrix is integration term by term. This implies that the distribution of ζ does not degenerate for $s \neq t$ and that $H(x) > 0$ for $x > 0$.

We now show that for $\tau \neq 0$:

$$\frac{1 - V_1^2(\|\tau\|^2) V_0^{-2}}{1 - \gamma^2(\|\tau\|^2)} > 1$$

which is equivalent to

$$(12.62) \quad -\gamma'(x) < -\gamma'(0)\gamma(x), \quad \forall x > 0.$$

The left-hand side of (12.62) can be written as

$$-\gamma'(x) = \frac{1}{2} \iint_0^{+\infty} (w_1 \exp(-xw_1) + w_2 \exp(-xw_2)) G(dw_1) G(dw_2)$$

and the right-hand side

$$-\gamma'(0)\gamma(x) = \frac{1}{2} \iint_0^{+\infty} (w_1 \exp(-xw_2) + w_2 \exp(-xw_1)) G(dw_1) G(dw_2),$$

so that

$$-\gamma'(0)\gamma(x) + \gamma'(x) = \frac{1}{2} \iint_0^{+\infty} (w_2 - w_1) (\exp(-xw_1) - \exp(-xw_2)) G(dw_1) G(dw_2),$$

which is ≥ 0 and is equal to zero only if G is concentrated at a point, which is not the case. This proves (12.62). Now, using the hypotheses on the inner and outer diameter of V_m , the result follows by a compactness argument. \square

Random fields and condition numbers of random matrices

Let A be an invertible $n \times n$ real matrix and $b \in \mathbb{R}^n$. We are interested in understanding how the solution $x \in \mathbb{R}^n$ of the linear system of equations

$$(13.1) \quad Ax = b$$

is affected by perturbations in the input (A, b) .

Early work by Turing (1948) and von Neumann and Goldstine (1947) identified that the key quantity is:

$$\kappa(A) = \|A\| \cdot \|A^{-1}\|$$

where $\|A\|$ denotes the operator norm of A defined in the usual way:

$$\|A\| = \max_{\|x\|=1} \|Ax\|.$$

Here $\|v\|$ denotes the Euclidean norm of $v \in \mathbb{R}^n$. Of course, other norms can be considered, but in this chapter we will restrain ourselves to the Euclidean norm. If A is singular, we put $\kappa(A) = +\infty$. Turing called $\kappa(A)$ the *condition number* of A .

The first meaning of $\kappa(A)$ is a consequence of the following property that the reader can easily check.

Let $x + \Delta x$ be the solution of system (13.1) when the input is $(A + \Delta A, b + \Delta b)$ instead of (A, b) . If $\kappa(A) \frac{\|\Delta A\|}{\|A\|} < 1$, then

$$(13.2) \quad \frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa(A)}{1 - \kappa(A) \frac{\|\Delta A\|}{\|A\|}} \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|} \right).$$

In fact,

$$(A + \Delta A)\Delta x = \Delta b - (\Delta A)x$$

which implies, whenever $\|A^{-1}(\Delta A)\| < 1$:

$$\Delta x = (A + \Delta A)^{-1}(\Delta b - (\Delta A)x) = (I + A^{-1}(\Delta A))^{-1}A^{-1}(\Delta b - (\Delta A)x)$$

and taking norms:

$$\begin{aligned} \|\Delta x\| &\leq \|(I + A^{-1}(\Delta A))^{-1}\| \|A^{-1}\| (\|\Delta b\| + \|\Delta A\| \|x\|) \\ &\leq \|(I + A^{-1}(\Delta A))^{-1}\| \|A^{-1}\| \left(\frac{\|\Delta b\|}{\|b\|} + \frac{\|\Delta A\|}{\|A\|} \right) \|A\| \|x\|, \end{aligned}$$

from which (13.2) follows.

Notice that the factor $\frac{\kappa(A)}{1 - \kappa(A) \frac{\|\Delta A\|}{\|A\|}}$ tends to $\kappa(A)$ when $\|\Delta A\| \rightarrow 0$. Thus, $\kappa(A)$ is a bound for the amplification of the relative error between output and input in the system (13.1), when the last one is small. The reader may also easily check that, in addition, $\kappa(A)$ is sharp in the sense that no smaller number will satisfy a similar inequality for all increments ΔA and Δb . In other words, if one thinks on binary floating point arithmetic, $\log_2 \kappa(A)$ measures the loss of precision due to error in the input. So, for numerical analysis purposes, it is usual that the relevant function of the matrix A appears to be $\log_2 \kappa(A)$.

Matrices A with $\kappa(A)$ small are said to be *well-conditioned*, those with $\kappa(A)$ large are said to be *ill-conditioned*. The set $\Sigma = \{A : \kappa(A) = +\infty\}$ is called the set of *ill-posed* matrices.

The distance of a matrix A to the set Σ is closely related to $\kappa(A)$ as shown in the next theorem. For a proof, see Blum et al. (1998):

THEOREM 13.1 (Eckart-Young, 1936). *For any $n \times n$ real matrix A one has*

$$\kappa(A) = \frac{\|A\|}{d_F(A, \Sigma)}$$

Here d_F means distance in \mathbb{R}^{n^2} with respect to the Frobenius norm $\|A\| = \sqrt{\sum a_{ij}^2}$.

The relationship between conditioning and distance to ill-posedness is a recurrent theme in numerical analysis (see for example Demmel, 1987). $\kappa(A)$ appears in more elaborate round-off analysis of algorithms, in which errors may occur in all the operations. As an example, let us mention such an analysis for Cholesky's method (see Wilkinson, 1963). If A is symmetric and positive definite we may solve the linear system $Ax = b$ by using Cholesky's factorization. Assume the length of the mantissa in the binary representation used in the computation is equal to ℓ . Then, if ℓ is large enough, one can prove that

$$\frac{\|\Delta x\|}{\|x\|} \leq 3n^3 2^{-\ell} \kappa(A).$$

The interested reader can find a variety of related subjects in the books by Higham (1996) and Trefethen and Bau (1997).

Next, we introduce some notation. Given A , an $n \times n$ real matrix, we denote by ν_1, \dots, ν_n , $0 \leq \nu_1 \leq \dots \leq \nu_n$ the squares of the singular values of A , that is, the eigenvalues of $A^T A$. If $X : S^{n-1} \rightarrow \mathbb{R}$ is the quadratic polynomial $X(x) = x^T A^T A x$, then:

- $\nu_n = \|A\|^2 = \max_{x \in S^{n-1}} X(x)$
- in case A is non-singular, it follows that $\nu_1 = \frac{1}{\|A^{-1}\|^2} = \min_{x \in S^{n-1}} X(x)$.

Hence,

$$\kappa(A) = \left(\frac{\nu_n}{\nu_1} \right)^{\frac{1}{2}}$$

when $\nu_1 > 0$, and $\kappa(A) = +\infty$ if $\nu_1 = 0$. Notice also that $\kappa(A) \geq 1$ and $\kappa(rA) = \kappa(A)$ for any real r , $r \neq 0$.

That is, the computation of the condition number of an $n \times n$ matrix A is a problem about the spectrum of the non-negative definite matrix $A^T A$ or, more precisely, about the largest and the smallest singular values of A .

Suppose one is interested in the analysis of a certain algorithm in which the condition number $\kappa(A)$ plays a role. Typically, the condition number will be a component appearing in some bound for the cost of the algorithm, in which - as in the example above - the size of the problem and the length of the mantissa in the floating point computation will also be present. A natural setting consists in imagining that the algorithm is applied to a problem which is drawn at random from a certain family of problems, which in our case amount to choose at random the coefficients of the system of equations. The cost of the algorithm and $\kappa(A)$ become random variables. One can try to compute, or give bounds, for the expectation or the higher moments of the cost, or estimate its distribution function and this will depend on the moments or the probability distribution of $\kappa(A)$ itself. Thus, in our case, the question of probabilistic analysis of algorithms, becomes a problem on the spectrum of random matrices.

There is large body of knowledge on random matrices and specifically, on their singular values or eigenvalues. Our general reference is Mehta's book (3d edition, 2004), to which we have already referred in Chapter 12. See also Girko's book (1996).

Since the 1930's motivation and methods have come from different sources, may be starting with Fisher in 1939 in multivariate statistics (see Muirhead's book, 1982, or Kendall, Stuart and

Ord, 1983). By the mid-1950's, Wigner's work (see for example his papers of 1958 and 1967) was followed by a big interest on random matrices in some areas of mathematical physics, which continues until the present times (see for example Soshnikov, 1998, Davidson and Szarek, 2001, Tracy and Widom, 1999). A third source of interest has been numerical analysis, motivated by the above mentioned condition number problems, may be starting with a paper by Steve Smale (1985). A very interesting survey of analytical methods, originally inspired by numerical analysis problems but also including a diversity of applications, is Edelman and Rao (2005).

In this chapter we are only considering a very small part of the subject, which concerns the condition number $\kappa(A)$. One reason to include this topic here is that the methods we will present are based upon random fields and Rice formulas, which is the core of this book. In some cases, these methods do not give optimal results, as it happens in the canonical case of matrices having i.i.d. standard normal entries. However, for non-centered Gaussian matrices our methods still provide the best bounds for the tails of the probability distribution of $\kappa(A)$ of which the authors are aware.

We start in Section 1 with some elementary upper-bounds for $E(\log \kappa(A))$ when A is a random matrix with i.i.d. - not necessarily Gaussian - entries. The methods are ad-hoc, but we are not aware of the existence of better general bounds.

The remainder of the chapter concerns Gaussian matrices. Section 2 is on centered matrices and Section 3 on non-centered ones.

1. Condition numbers of non-Gaussian matrices

The results of this section are extracted from Cuesta-Albertos and Wschebor (2003).

Throughout this section, we assume that $A = ((a_{ij})), i, j = 1, 2, \dots, n$ is an $n \times n$ matrix, where the a_{ij} 's are independent identically distributed real-valued random variables defined on some probability space (Ω, A, P) . We denote with μ the common distribution measure of the a_{ij} 's.

1.1. A general bound for $E(\log \kappa(A))$ for symmetric entries.

THEOREM 13.2. *We assume that the distribution μ satisfies the following conditions:*

- (1) *For any pair α, β of real numbers, $\alpha < \beta$, one has*

$$(13.3) \quad \mu\left([\alpha, \beta]\right) \leq \mu\left(\left[-\frac{\beta - \alpha}{2}, \frac{\beta - \alpha}{2}\right]\right).$$

- (2) $E[|a_{1,1}|^r] = 1$, for some $r > 0$.
 (3) *There exist positive numbers C, γ such that*

$$\mu([- \alpha, \alpha]) \leq C\alpha^\gamma, \text{ for all } \alpha > 0.$$

Then,

$$(13.4) \quad E[\log \kappa(A)] \leq \left(1 + \frac{2}{r}\right) \log n + \frac{1}{r} + \frac{1}{\gamma} \left\{[(2 + \gamma) \log n + \log C]^+ + 1\right\},$$

where $x^+ = \max(x, 0)$ for real x .

PROOF. Notice that $\|A\| \leq \left(\sum_{i,j=1}^n a_{i,j}^2\right)^{1/2}$. So, with the only assumption that the random entries are identically distributed, one has, for $t > 0$:

$$(13.5) \quad P[\|A\| > t] \leq P\left[n^2 \sup_{i,j=1,\dots,n} a_{i,j}^2 > t^2\right] \leq P\left[\bigcup_{i,j=1}^n \left\{|a_{i,j}| > \frac{t}{n}\right\}\right] \leq n^2 P\left[|a_{1,1}| > \frac{t}{n}\right].$$

Hence, for $\alpha_n \geq 0$

$$\begin{aligned}
 (13.6) \quad \mathbb{E}[\log \|A\|] &\leq \alpha_n + \int_{\alpha_n}^{\infty} \mathbb{P}[\log \|A\| > x] dx \\
 &= \alpha_n + \int_{\alpha_n}^{\infty} \mathbb{P}[\|A\| > e^x] dx \\
 &\leq \alpha_n + \int_{\alpha_n}^{\infty} n^2 \mathbb{P}\left[|a_{1,1}| > \frac{e^x}{n}\right] dx \\
 &\leq \alpha_n + \int_{\alpha_n}^{\infty} n^2 \left(\frac{n}{e^x}\right)^r dx = \alpha_n + n^{2+r} \frac{1}{r} e^{-r\alpha_n},
 \end{aligned}$$

where the last inequality follows from Markov inequality and Assumption (2).

Choose $\alpha_n \geq 0$ to minimize the right-hand side of (13.6), i.e. $\alpha_n = (1 + \frac{2}{r}) \log n$ and it follows that

$$(13.7) \quad \mathbb{E}[\log \|A\|] \leq \left(1 + \frac{2}{r}\right) \log n + \frac{1}{r}.$$

We now consider the factor $\|A^{-1}\|$. Denote $A^{-1} = ((b_{i,j}))_{i,j=1,\dots,n}$, that is:

$$b_{i,j} = \frac{a^{i,j}}{\det(A)}, \quad i, j = 1, \dots, n,$$

where $a^{i,j}$ is the cofactor of the position (i, j) in the matrix A .

Clearly the r.v.'s $|b_{i,j}|$, $i, j = 1, \dots, n$ are identically distributed so that we may apply (13.5) to the matrix A^{-1} instead of A :

$$\begin{aligned}
 \mathbb{P}[\|A^{-1}\| > t] &\leq n^2 \mathbb{P}\left[|b_{1,1}| > \frac{t}{n}\right] = n^2 \mathbb{P}\left[\left|\frac{a^{1,1}}{\sum_{j=1}^n a_{1,j} a^{1,j}}\right| > \frac{t}{n}\right] \\
 &= n^2 \mathbb{P}\left[\left|a_{1,1} + \sum_{j=2}^n a_{1,j} \frac{a^{1,j}}{a^{1,1}}\right| < \frac{n}{t}\right].
 \end{aligned}$$

The random variables

$$a_{1,1} \quad \text{and} \quad \eta = \sum_{j=2}^n a_{1,j} \frac{a^{1,j}}{a^{1,1}}$$

are independent, so that, for each $\alpha > 0$, denoting by \mathbb{P}_η the probability distribution of η , and using Fubini's theorem and Assumption (1), we have:

$$\begin{aligned}
 \mathbb{P}[|a_{1,1} + \eta| < \alpha] &= \int_{-\infty}^{\infty} \mu[(-\alpha - y, \alpha - y)] \mathbb{P}_\eta(dy) \\
 &\leq \int_{-\infty}^{\infty} \mu[(-\alpha, \alpha)] \mathbb{P}_\eta(dy) = \mu[(-\alpha, \alpha)].
 \end{aligned}$$

Hence, by Assumption (3),

$$(13.8) \quad \mathbb{P}[\|A^{-1}\| > t] \leq n^2 \mu\left[\left[-\frac{n}{t}, \frac{n}{t}\right]\right] \leq n^2 C \left(\frac{n}{t}\right)^\gamma,$$

and, with $\beta_n \geq 0$:

$$\begin{aligned}
 \mathbb{E}[\log \|A^{-1}\|] &\leq \beta_n + \int_{\beta_n}^{\infty} \mathbb{P}[\|A^{-1}\| > e^x] dx \\
 &\leq \beta_n + \int_{\beta_n}^{\infty} C n^{2+\gamma} e^{-\gamma x} dx = \beta_n + C \frac{n^{2+\gamma}}{\gamma} e^{-\gamma \beta_n}.
 \end{aligned}$$

Choosing $\beta_n = \frac{1}{\gamma} [(2 + \gamma) \log n + \log C]^+$, one obtains:

$$(13.9) \quad \mathbb{E}[\log \|A^{-1}\|] \leq \frac{1}{\gamma} \{[(2 + \gamma) \log n + \log C]^+ + 1\},$$

and putting together (13.7) and (13.9), (13.4) follows. \square

Remarks on the statement of Theorem 13.2.

It is not hard to see that Assumption (1) implies that the measure μ is symmetric around 0. In particular, this implies that in case the random variables $a_{i,j}, i, j = 1, \dots, n$ are integrable, their common expectation must be 0.

Since $\kappa(\lambda A) = \kappa(A)$ for any non-zero real number λ and any nonsingular matrix A , in case $m_r = \int_{-\infty}^{\infty} |x|^r \mu(dx) < \infty$ it is possible to replace A by $m_r^{-1/r} A$ so that assumption 2 holds true, without modifying the condition number. Of course, in this case one must change accordingly the constant C in assumption 3. In this sense, assumption 2 is not more restrictive than the finiteness of the r -th moment of the probability measure μ .

1.2. Some examples.

Density Assume that μ has a density function f , that f is even and non-increasing on $[0, \infty)$ and that $m_r = \int_{-\infty}^{\infty} |x|^r f(x) dx < \infty$ for some $r > 0$.

We replace the original density f by $m_r^{1/r} f(m_r^{1/r} x)$ so that assumption 2 is satisfied without changing $\kappa(A)$; assumption 3 is verified with $\gamma = 1$ and $C = 2m_r^{1/r} f(0)$. Inequality (13.4) becomes

$$(13.10) \quad \mathbb{E}[\log \kappa(A)] \leq \left(1 + \frac{2}{r}\right) \log n + \frac{1}{r} + \left[3 \log n + \frac{1}{r} \log m_r + \log(2f(0))\right]^+ + 1.$$

Uniform distribution. Let μ be the uniform distribution on $[-H, H]$, $H > 0$. In this case, $m_r = H^r (r+1)^{-1}$ and (13.10) holds true for any $r > 0$. Letting $r \rightarrow +\infty$, we obtain

$$(13.11) \quad \mathbb{E}[\log \kappa(A)] \leq 4 \log n + 1.$$

Strong concentration near the mean. Assume that the density of μ has the form

$$\frac{1}{2} \frac{\gamma}{|x|^{1-\gamma}} \mathbf{1}_{[-1,1]}(x),$$

for some γ , $0 < \gamma < 1$.

One has $m_r = \frac{\gamma}{r+\gamma}$ for each $r > 0$ and easily checks that introducing the modification suggested above, assumptions 1, 2 and 3 are satisfied with $C = m_r^{\gamma/r}$. Hence, Theorem 13.2 implies that for any $r > 0$:

$$\mathbb{E}[\log \kappa(A)] \leq \left(1 + \frac{2}{r}\right) \log n + \frac{1}{r} + \frac{1}{\gamma} \left\{ \left[(2 + \gamma) \log n + \frac{\gamma}{r} \log \frac{\gamma}{r + \gamma} \right]^+ + 1 \right\},$$

and, letting $r \rightarrow +\infty$ it follows that

$$\mathbb{E}[\log \kappa(A)] \leq \left(2 + \frac{2}{\gamma}\right) \log n + \frac{1}{\gamma}.$$

Particular distributions. The bound in Theorem 13.2 can be improved by using the actual distribution μ instead of the Markov inequality in (13.6) or the bound in (13.8). This is, for example, the case for symmetric exponential or standard normal distributions but. However, in the later case, this method is not fine enough: the precise behavior of $\mathbb{E}[\log \kappa(A)]$ as $n \rightarrow +\infty$ was given by Edelman (1988), using analytic methods. It is the following:

$$\mathbb{E}[\log \kappa(A)] = \log n + C_0 + \varepsilon_n,$$

where C_0 is a known constant ($C_0 \simeq 1, 537$) and $\varepsilon_n \rightarrow 0$.

1.3. Smoothed analysis. We consider now the condition number when the r.v.'s in the matrix $A = (a_{i,j})_{i,j=1,\dots,n}$ have the form

$$a_{i,j} = m_{i,j} + \psi_{i,j}, \quad i, j = 1, \dots, n,$$

where $M = (m_{i,j})_{i,j=1,\dots,n}$ is non-random and $(\psi_{i,j})_{i,j=1,\dots,n}$ are i.i.d. r.v.'s with common distribution μ satisfying Assumptions (1), (2) and (3) in Theorem 13.2. This has been called “smoothed analysis” and corresponds to the idea of exploring what happens to the condition number when a non-random matrix is perturbed with a noise having a law which verifies a certain number of requirements. We will be only looking to the effect on the moments of the loss of precision $\log \kappa(A)$, when performing this operation. (See Spielman and Teng (2002) and Tao and Vu (2007), where similar questions are considered, in the last case, allowing the measure μ to be purely atomic).

THEOREM 13.3. *Under the conditions stated above, if we put*

$$m_n = \sup_{i,j=1,\dots,n} |m_{i,j}| \leq n^{2/r},$$

then

$$(13.12) \quad \mathbb{E}[\log \kappa(A)] \leq \left(1 + \frac{2}{r}\right) \log n + \log 2 + \frac{1}{r} + \frac{1}{\gamma} \{[(2 + \gamma) \log n + \log C]^+ + 1\}$$

PROOF. The proof (as well as the result) is very similar to that of Theorem 13.2. For $t > 0$ one has:

$$\begin{aligned} \mathbb{P}[\|A\| > t] &\leq \mathbb{P}\left[\sum_{i,j=1}^n a_{i,j}^2 > t^2\right] \leq \sum_{i,j=1}^n \mathbb{P}\left[a_{i,j}^2 > \frac{t^2}{n^2}\right] \\ &= \sum_{i,j=1}^n \mathbb{P}\left[|m_{i,j} + \psi_{i,j}| > \frac{t}{n}\right] \leq n^2 \mathbb{P}\left[|\psi_{1,1}| > \frac{t}{n} - m_n\right]. \end{aligned}$$

Now choose $\alpha_n = \left(1 + \frac{2}{r}\right) \log n + \log 2$. If $x > \alpha_n$, then

$$\frac{e^x}{n} - m_n > \frac{1}{2n} e^x.$$

Thus,

$$\begin{aligned} \mathbb{E}[\log \|A\|] &\leq \alpha_n + \int_{\alpha_n}^{\infty} \mathbb{P}[\|A\| > e^x] dx \\ &\leq \alpha_n + n^2 \int_{\alpha_n}^{\infty} \mathbb{P}\left[|\psi_{1,1}| > \frac{1}{2n} e^x\right] dx \\ &\leq \alpha_n + n^2 \int_{\alpha_n}^{\infty} \frac{1}{\left(\frac{1}{2n} e^x\right)^r} dx \\ &= \left(1 + \frac{2}{r}\right) \log n + \log 2 + \frac{1}{r}. \end{aligned}$$

On the other hand, with the same notation as in the proof of Theorem 13.2,

$A^{-1} = (b_{i,j})_{i,j=1,\dots,n}$ and

$$\mathbb{P}[\|A^{-1}\| > t] \leq \sum_{i,j=1}^n \mathbb{P}\left[|b_{i,j}| > \frac{t}{n}\right].$$

For each term in this sum it is possible to repeat exactly the same computations as in the proof of Theorem 13.2 to bound $\mathbb{P}\left[|b_{1,1}| > \frac{t}{n}\right]$ and obtain the same bound as there for $\mathbb{E}[\log \|A^{-1}\|]$. This finishes the proof. \square

For higher order moments, one can obtain upper bounds for $\mathbb{E}[(\log \kappa(A))^k]$, $k = 2, 3, \dots$ much in the same way as we did for $k = 1$. We consider here the centered case, for smoothed analysis,

the situation is similar.

Since $\log \kappa(A) \geq 0$ we have that

$$\mathbb{E} [(\log \kappa(A))^k] \leq 2^k \left[\mathbb{E} \left\{ (\log^+ \|A\|)^k \right\} + \mathbb{E} \left\{ (\log^+ \|A^{-1}\|)^k \right\} \right].$$

Using the same estimates as in the case $k = 1$ for the tails of the probability distributions of $\|A\|$ and $\|A^{-1}\|$, after an elementary computation, it is possible to obtain that if k satisfies the inequalities $2 \leq k \leq 1 + (2 + \gamma \wedge r) \log n$, then

$$\mathbb{E} [(\log \kappa(A))^k] \leq (2 \log n)^k \left[\left(1 + \frac{2}{r}\right)^k (1+k) + \left(1 + \frac{2}{\gamma}\right)^k (1+Ck) \right]$$

2. Condition numbers of centered Gaussian matrices

The purpose of the present section is to prove the following

THEOREM 13.4 (Azaïs-Wschebor, 2005c). *Assume that $A = ((a_{ij}))_{i,j=1,\dots,n}$, $n \geq 3$, and that the a_{ij} 's are i.i.d. standard normal random variables.*

Then, there exist universal positive constants c, C such that for $x > 1$:

$$(13.13) \quad \frac{c}{x} < \mathbb{P}(\kappa(A) > n.x) < \frac{C}{x}$$

Remarks

(1) The limiting distribution of $\kappa(A)/n$ as $n \rightarrow \infty$, has been computed in Edelman's thesis (1989). The interest of this theorem lies in the uniformity of the statement and in the relationship that the proof below establishes with Rice formulas.

(2) This Theorem, and related ones, can be considered as results on the Wishart matrix $A^T A$. Introducing some minor changes, it is possible to use the same methods to study the condition number of $A^T A$ for rectangular $n \times m$ matrices A having i.i.d. standard normal entries, $n > m$.

(3) We will see below that $c = 0.13$, $C = 5.60$ satisfy (13.13) for every $n = 3, 4, \dots$. Using the same methods one can obtain more precise upper and lower bounds for each n . Improved values for the constants, as well as extensions to rectangular matrices and to other canonical non-Gaussian distributions can be found in Edelman and Sutton's paper of (2005), where the proofs are based upon the analytic theory of random matrices. In particular, for the constant C these authors show evidence for the value $C = 2$. See the numerical application in the next section.

PROOF OF THEOREM 13.4. Recall the notation in the introduction of this chapter. It is easy to see that, almost surely, the eigenvalues of $A^T A$ are pairwise different. We introduce the following additional notations:

- $\{e_1, \dots, e_n\}$ is the canonical basis of \mathbb{R}^n .
- $B = A^T A = ((b_{ij}))_{i,j=1,\dots,n}$
- For $s \neq 0$ in \mathbb{R}^n $\pi_s : \mathbb{R}^n \rightarrow \mathbb{R}^n$ denotes the orthogonal projection onto $\{s\}^\perp$, the orthogonal complement of s in \mathbb{R}^n
- For a differentiable function F defined on a smooth manifold M embedded in some Euclidean space, $F'(s)$ and $F''(s)$ are the first and the second derivative of F that we will represent, in each case, with respect to an appropriate orthonormal basis of the tangent space.

Instead of (13.13) we prove the equivalent statement: for $x > n$:

$$(13.14) \quad \frac{c n}{x} < \mathbb{P}(\kappa(A) > x) < \frac{C n}{x}$$

We break the proof into several steps. Our main task is to estimate the joint density of the pair (ν_n, ν_1) ; this will be done in Step 4.

Step 1 For $a, b \in \mathbb{R}$, $a > b$, one has almost surely:

$$(13.15) \quad \{\nu_n \in (a, a + da), \nu_1 \in (b, b + db)\} \\ = \left\{ \begin{array}{l} \exists s, t \in S^{n-1}, \langle s, t \rangle = 0, X(s) \in (a, a + da), X(t) \in (b, b + db), \\ \pi_s(Bs) = 0, \pi_t(Bt) = 0, X''(s) \prec 0, X''(t) \succ 0 \end{array} \right\}$$

The random integer $N_{a,b,da,db}$ of pairs (s, t) belonging to the right-hand side of (13.15) is equal to 0 or to 4, so that:

$$(13.16) \quad P(\nu_n \in (a, a + da), \nu_1 \in (b, b + db)) = \frac{1}{4} E(N_{a,b,da,db})$$

Step 2. In this step we will give a bound for $E(N_{a,b,da,db})$ using a Rice-type formula. Let

$$V = \{(s, t) : s, t \in S^{n-1}, \langle s, t \rangle = 0\}.$$

V is a C^∞ -differentiable manifold without boundary, embedded in \mathbb{R}^{2n} , $\dim(V) = 2n - 3$. We will denote by $\tau = (s, t)$ a generic point in V and by $\sigma_V(d\tau)$ the geometric measure on V . We will need the total measure $\sigma_V(V)$, which is a particular case of the following lemma (we will use the full statement in the next section).

LEMMA 13.5. *Let $a, b > 0$. We define:*

$$V_{a,b} = \{(s, t) \in \mathbb{R}^n \times \mathbb{R}^n : \|s\|^2 = a, \|t\|^2 = b, \langle s, t \rangle = 0\},$$

Denote by $\mu_{a,b}$ the geometric measure of the compact C^∞ -manifold $V_{a,b}$ embedded in \mathbb{R}^{2n} . Then:

$$\mu_{a,b} = (a + b)^{1/2} (ab)^{\frac{m-2}{2}} \sigma_{m-1} \sigma_{m-2}$$

where σ_{n-1} denotes the surface area of $S^{n-1} \subset \mathbb{R}^n$, that is $\sigma_{n-1} = \frac{2\pi^{n/2}}{\Gamma(n/2)}$.

PROOF. Notice that, for each point $(s, t) \in V_{a,b}$, the triplet

$$\left(\frac{s}{\|s\|}, 0 \right), \left(0, \frac{t}{\|t\|} \right), \frac{1}{\sqrt{\|s\|^2 + \|t\|^2}}(t, s)$$

is an orthonormal basis of the normal space to $V_{a,b}$ at (s, t) , which correspond respectively to the unit vectors orthogonal to each one of the $(2n - 1)$ -dimensional manifolds in \mathbb{R}^{2n} given by equations:

$$(13.17) \quad \begin{aligned} \|s\|^2 &= a \\ \|t\|^2 &= b \\ \langle s, t \rangle &= 0. \end{aligned}$$

So, as $\delta \downarrow 0$, the $2n$ -dimensional Lebesgue measure of the set

$$E_\delta = \{(s, t) \in \mathbb{R}^{2n} : \sqrt{a} - \delta < \|s\| < \sqrt{a} + \delta, \sqrt{b} - \delta < \|t\| < \sqrt{b} + \delta, |\langle s, t \rangle| < \delta\sqrt{a+b}\}$$

is equivalent to:

$$(2\delta)^3 \mu_{a,b}.$$

On the other hand:

$$(13.18) \quad \lambda_{2n}(E_\delta) = \int_{\{\sqrt{a}-\delta < \|s\| < \sqrt{a}+\delta\}} ds \int_{\{\sqrt{b}-\delta < \|t\| < \sqrt{b}+\delta, |\langle s, t \rangle| < \delta\sqrt{a+b}\}} dt$$

Using polar coordinates in each iterate of the double integral in 13.18, the result follows. \square

WE GO BACK TO THE PROOF OF THE THEOREM. Since $V = V_{1,1}$, Lemma 13.5 implies that $\sigma_V(V) = \sqrt{2}\sigma_{n-1} \cdot \sigma_{n-2}$.

On V we define the random field

$$Y : V \rightarrow \mathbb{R}^{2n}$$

by means of

$$Y(s, t) = \begin{pmatrix} \pi_s(Bs) \\ \pi_t(Bt) \end{pmatrix}.$$

For $\tau = (s, t)$ a given point in V , we have that

$$Y(\tau) \in \{(t, -s)\}^\perp \cap \{\{s\}^\perp \times \{t\}^\perp\} = W_\tau$$

for any value of the matrix B , where $\{(t, -s)\}^\perp$ is the orthogonal complement of the point $(t, -s)$ in \mathbb{R}^{2n} . In fact, $(t, -s) \in \{s\}^\perp \times \{t\}^\perp$ and

$$\begin{aligned} \langle Y(s, t), (t, -s) \rangle_{\mathbb{R}^{2n}} &= \langle \pi_s(Bs), t \rangle - \langle \pi_t(Bt), s \rangle \\ &= \langle Bs - \langle s, Bs \rangle s, t \rangle - \langle Bt - \langle t, Bt \rangle t, s \rangle = 0 \end{aligned}$$

since $\langle s, t \rangle = 0$ and B is symmetric.

Notice that $\dim(W_\tau) = 2n - 3$.

We also set

$$\Delta(\tau) = \left[\det \left[(Y'(\tau))^T Y'(\tau) \right] \right]^{\frac{1}{2}}$$

For $\tau = (s, t) \in V$, F_τ denotes the event

$$F_\tau = \{X(s) \in (a, a + da), X(t) \in (b, b + db), X''(s) < 0, X''(t) > 0\}$$

and $p_{Y(\tau)}(\cdot)$ is the density of the random vector $Y(\tau)$ in the $(2n - 3)$ -dimensional subspace W_τ of \mathbb{R}^{2n} .

Applying Rice formula:

$$\begin{aligned} (13.19) \quad \mathbb{E}(N_{a,b,da,db}) &= \int_a^{a+da} dx \int_b^{b+db} dy \int_V \mathbb{E}(\Delta(s, t) \mathbf{1}_{\{X''(s) < 0, X''(t) > 0\}} | X(s) = x, X(t) = y, Y(s, t) = 0) \\ &\quad \cdot p_{X(s), X(t), Y(s, t)}(x, y, 0) \sigma_V(d(s, t)) \end{aligned}$$

The invariance of the law of A with respect to the orthogonal group of \mathbb{R}^n implies that the integrand in (13.19) does not depend on $(s, t) \in V$. Hence, we have proved that the joint law of λ_n and λ_1 has a density $g(a, b)$, $a > b$, and

$$\begin{aligned} (13.20) \quad g(a, b) &= \frac{\sqrt{2}}{4} \sigma_{n-1} \cdot \sigma_{n-2} \mathbb{E}(\Delta(e_1, e_2) \mathbf{1}_{\{X''(e_1) < 0, X''(e_2) > 0\}} | X(e_1) = a, X(e_2) = b, Y(e_1, e_2) = 0) \\ &\quad \cdot p_{X(e_1), X(e_2), Y(e_1, e_2)}(a, b, 0). \end{aligned}$$

Step 3 Next, we compute the ingredients in the right-hand side of (13.20). We take as orthonormal basis for the subspace $W_{(e_1, e_2)}$:

$$\{(e_3, 0), \dots, (e_n, 0), (0, e_3), \dots, (0, e_n), \frac{1}{\sqrt{2}}(e_2, e_1)\} = L_1.$$

We have:

$$\begin{aligned} X(e_1) &= b_{11} \\ X(e_2) &= b_{22} \\ X''(e_1) &= B_1 - b_{11}I_{n-1} \\ X''(e_2) &= B_2 - b_{22}I_{n-1} \end{aligned}$$

where B_1 (resp. B_2) is the $(n - 1) \times (n - 1)$ matrix obtained by suppressing the first (resp. the second) row and column in B .

$$Y(e_1, e_2) = (0, b_{21}, b_{31}, \dots, b_{n1}, b_{12}, 0, b_{32}, \dots, b_{n2}, b_{12})^T$$

so that, it has the expression in the orthonormal basis L_1 :

$$Y(e_1, e_2) = \sum_{i=3}^n (b_{i1}(e_i, 0) + b_{i2}(0, e_i)) + \sqrt{2}b_{12}\left(\frac{1}{\sqrt{2}}(e_2, e_1)\right)$$

So, the joint density of $X(e_1), X(e_2), Y(e_1, e_2)$ appearing in (13.20) in the space $\mathbb{R} \times \mathbb{R} \times W_{(e_1, e_2)}$ is the joint density of the r.v.'s

$$b_{11}, b_{22}, \sqrt{2}b_{12}, b_{31}, \dots, b_{n1}, b_{32}, \dots, b_{n2}$$

at the point $(a, b, 0, \dots, 0)$. To compute this density, we first compute the joint density q of

$$b_{31}, \dots, b_{n1}, b_{32}, \dots, b_{n2},$$

given a_1, a_2 , where a_j denotes the j -th column of A which is standard normal in \mathbb{R}^n .

q is the normal density in $\mathbb{R}^{2(n-2)}$, centered with variance matrix

$$\begin{pmatrix} \|a_1\|^2 I_{n-2} & \langle a_1, a_2 \rangle I_{n-2} \\ \langle a_1, a_2 \rangle I_{n-2} & \|a_2\|^2 I_{n-2} \end{pmatrix}.$$

Set

$$a'_j = \frac{a_j}{\|a_j\|}, \quad j = 1, 2.$$

Now we compute the density of the triplet

$$(b_{11}, b_{22}, b_{12}) = (\|a_1\|^2, \|a_2\|^2, \|a_1\| \|a_2\| \langle a'_1, a'_2 \rangle)$$

at the point $(a, b, 0)$.

Since $\langle a'_1, a'_2 \rangle$ and $(\|a_1\|, \|a_2\|)$ are independent, the density of the triplet at $(a, b, 0)$ is equal to:

$$\chi_n^2(a)\chi_n^2(b)(ab)^{-1/2}p_{\langle a'_1, a'_2 \rangle}(0)$$

where $\chi_n^2(\cdot)$ denotes the χ^2 density with n degrees of freedom.

Let $\xi = (\xi_1, \dots, \xi_n)^T$ be standard normal in \mathbb{R}^n . Clearly, $\langle a'_1, a'_2 \rangle$ has the same distribution as $\frac{\xi_1}{\|\xi\|}$, because of the invariance under the orthogonal group.

$$\begin{aligned} \frac{1}{2t}P\{|\langle a'_1, a'_2 \rangle| \leq t\} &= \frac{1}{2t}P\left\{\frac{\xi_1^2}{\chi_{n-1}^2} \leq \frac{t^2}{1-t^2}\right\} = \frac{1}{2t}P\{F_{1, n-1} \leq \frac{t^2(n-1)}{1-t^2}\} \\ &= \frac{1}{2t} \int_0^{\frac{t^2(n-1)}{1-t^2}} f_{1, n-1}(x)dx, \end{aligned}$$

where $\chi_{n-1}^2 = \xi_2^2 + \dots + \xi_n^2$ and $F_{1, n-1}$ has the Fisher distribution with $(1, n-1)$ degrees of freedom and density $f_{1, n-1}$. Letting $t \rightarrow 0$, we obtain

$$p_{\langle a'_1, a'_2 \rangle}(0) = \frac{1}{\sqrt{\pi}} \frac{\Gamma(n/2)}{\Gamma((n-1)/2)}$$

Summing up, the density in (13.20) is equal to :

$$(13.21) \quad \frac{1}{\sqrt{2}}(2\pi)^{2-n}\pi^{-\frac{1}{2}} \frac{1}{\Gamma(n/2)\Gamma((n-1)/2)} 2^{-n} \frac{1}{\sqrt{ab}} \exp\left(-\frac{a+b}{2}\right).$$

We now consider the conditional expectation in (13.20). The tangent space to V at the point (s, t) is parallel to the orthogonal complement in $\mathbb{R}^n \times \mathbb{R}^n$ of the triplet of vectors $(s, 0); (0, t); (t, s)$. This is immediate from the definition of V .

To compute the associated matrix for $Y'(e_1, e_2)$ take the set

$$\{(e_3, 0), \dots, (e_n, 0), (0, e_3), \dots, (0, e_n), \frac{1}{\sqrt{2}}(e_2, -e_1)\} = L_2.$$

as orthonormal basis in the tangent space and the canonical basis in \mathbb{R}^{2n} . A direct calculation gives :

$$Y'(e_1, e_2) = \begin{pmatrix} -v^T & 0_{1,n-2} & -\frac{1}{\sqrt{2}}b_{21} \\ w^T & 0_{1,n-2} & \frac{1}{\sqrt{2}}(-b_{11} + b_{22}) \\ B_{12} - b_{11}I_{n-2} & 0_{n-2,n-2} & \frac{1}{\sqrt{2}}w \\ 0_{1,n-2} & -w^T & \frac{1}{\sqrt{2}}(-b_{11} + b_{22}) \\ 0_{1,n-2} & v^T & \frac{1}{\sqrt{2}}b_{21} \\ 0_{n-2,n-2} & B_{12} - b_{22}I_{n-2} & -\frac{1}{\sqrt{2}}v \end{pmatrix}$$

where $v^T = (b_{31}, \dots, b_{n1})$, $w^T = (b_{32}, \dots, b_{n2})$, $0_{i,j}$ is a null matrix with i rows and j columns and B_{12} is obtained from B by suppressing the first and second rows and columns. The columns represent the derivatives in the directions of L_2 at the point (e_1, e_2) . The first n rows correspond to the components of $\pi_s(Bs)$, the last n ones to those of $\pi_t(Bt)$.

Thus, under the conditioning in (13.20),

$$Y'(e_1, e_2) = \begin{pmatrix} 0_{1,n-2} & 0_{1,n-2} & 0 \\ 0_{1,n-2} & 0_{1,n-2} & \frac{1}{\sqrt{2}}(b-a) \\ B_{12} - aI_{n-2} & 0_{n-2,n-2} & 0_{n-2,1} \\ 0_{1,n-2} & 0_{1,n-2} & \frac{1}{\sqrt{2}}(b-a) \\ 0_{1,n-2} & 0_{1,n-2} & 0 \\ 0_{n-2,n-2} & B_{12} - bI_{n-2} & 0_{n-2,1} \end{pmatrix}$$

and

$$\left[\det \left[(Y'(e_1, e_2))^T Y'(e_1, e_2) \right] \right]^{\frac{1}{2}} = |\det(B_{12} - aI_{n-2})| |\det(B_{12} - bI_{n-2})| (a-b)$$

Step 4 Notice that $B_1 - aI_{n-1} < 0 \Rightarrow B_{12} - aI_{n-2} < 0$, and similarly, $B_2 - bI_{n-1} > 0 \Rightarrow B_{12} - bI_{n-2} > 0$, and that for $a > b$, under the conditioning in (13.20), there is equivalence in these relations.

It is also clear that, since $B_{12} > 0$ one has

$$|\det(B_{12} - aI_{n-2})| \mathbf{1}_{B_{12} - aI_{n-2} < 0} \leq a^{n-2}$$

and it follows that the conditional expectation in (13.20) is bounded by:

$$(13.22) \quad a^{n-1} \mathbf{E}(|\det(B_{12} - bI_{n-2})| \mathbf{1}_{B_{12} - bI_{n-2} > 0} | b_{11} = a, b_{22} = b, b_{12} = 0, b_{i1} = b_{i2} = 0 \\ (i = 3, \dots, n)).$$

We further condition on a_1 and a_2 . Since unconditionally a_3, \dots, a_n are i.i.d. standard normal vectors in \mathbb{R}^n , under the conditioning, their joint law becomes the law of i.i.d. standard normal vectors in \mathbb{R}^{n-2} and independent of the condition. That is, (13.22) is equal to

$$(13.23) \quad a^{n-1} \mathbf{E}(|\det(M - bI_{n-2})| \mathbf{1}_{M - bI_{n-2} > 0}),$$

where M is an $(n-2) \times (n-2)$ random matrix with entries $M_{ij} = \langle v_i, v_j \rangle$, $(i, j = 1, \dots, n-2)$ and the vectors v_1, \dots, v_{n-2} are i.i.d. standard normal in \mathbb{R}^{n-2} . The expression in (13.23) is bounded by

$$a^{n-1} \mathbf{E}(\det(M)) = a^{n-1} (n-2)!,$$

The last equality is contained in the following lemma (see, for example, Mehta, 2004). We include a proof for completeness.

LEMMA 13.6. *Let ξ_1, \dots, ξ_m be i.i.d. random vectors in \mathbb{R}^p , $p \geq m$, their common distribution being Gaussian centered with variance I_p .*

Denote $W_{m,p}$ the matrix

$$W_{m,p} = ((\langle \xi_i, \xi_j \rangle))_{i,j=1,\dots,m}.$$

and

$$D(\lambda) = \det(W_{m,p} - \lambda I_m)$$

its characteristic polynomial.

Then,

$$(13.24) \quad \mathbb{E}(\det(W_{m,p})) = p(p-1)\dots(p-m+1)$$

$$(13.25) \quad \mathbb{E}(D(\lambda)) = \sum_{k=0}^m (-1)^k \binom{m}{k} \frac{p!}{(p-m+k)!} \lambda^k$$

PROOF. Fix the values of ξ_1, \dots, ξ_{m-1} , which are linearly independent with probability 1. Let $\{w_1, \dots, w_p\}$ be an orthonormal basis of \mathbb{R}^p such that $w_1, \dots, w_{m-1} \in V_{m-1}$, where V_j denotes the subspace generated by ξ_1, \dots, ξ_j .

Observe that $\det(W_{m,p})$ is the square of the volume - in $V_m \subset \mathbb{R}^p$ - of the parallelotope $\{\sum_{i=1}^m c_i \xi_i, 0 \leq c_i \leq 1, i = 1, \dots, m\}$. So,

$$\det(W_{m,p}) = d^2(\xi_m, V_{m-1}) \cdot \det(W_{m-1,p})$$

where $d(\xi_m, V_{m-1})$ is the - Euclidean - distance from ξ_m to V_{m-1} . Because of rotational invariance of the standard normal distribution in \mathbb{R}^p , the conditional distribution of $d^2(\xi_m, V_{m-1})$ given ξ_1, \dots, ξ_{m-1} is independent of the condition and χ_{p-m+1}^2 . Hence,

$$\begin{aligned} \mathbb{E}(\det(W_{m,p})) &= \mathbb{E}(\mathbb{E}(d^2(\xi_m, V_{m-1}) \cdot \det(W_{m-1,p}) \mid \xi_1, \dots, \xi_{m-1})) \\ &= (p-m+1)\mathbb{E}(\det(W_{m-1,p})) \end{aligned}$$

Iterating the procedure we get (13.24).

Let us prove (13.25). Clearly

$$(13.26) \quad D(\lambda) = \sum_{k=0}^m \frac{D^{(k)}(0)}{k!} \lambda^k$$

Standard differentiation of the determinant with respect to λ shows that for $k = 1, \dots, m-1$ one has

$$D^{(k)}(\lambda) = (-1)^k \sum \det(W_{m,p}^{i_1, \dots, i_k} - \lambda I_{m-k})$$

where the sum is over all k -tuples i_1, \dots, i_k of pairwise different non-negative integers that are smaller or equal than m , and the $(m-k) \times (m-k)$ matrix $W_{m,p}^{i_1, \dots, i_k}$ is obtained by suppressing in $W_{m,p}$ the rows and columns numbered i_1, \dots, i_k . Hence, applying (13.24) to each term in this sum, and on account of the number of terms, we get

$$\mathbb{E}(D^{(k)}(0)) = (-1)^k m(m-1)\dots(m-k+1)p(p-1)\dots(p-(m-k)+1).$$

To finish, take expectations in (13.26), and notice that

$$D^{(m)}(\lambda) = (-1)^m m! \quad ; \quad \mathbb{E}(D^{(0)}(0)) = p(p-1)\dots(p-m+1).$$

□

RETURNING TO THE PROOF OF THE THEOREM and summing up this part, after replacing in (13.20), we get

$$(13.27) \quad g(a, b) \leq C_n \frac{\exp(-(a+b)/2)}{\sqrt{ab}} a^{n-1},$$

where

$$C_n = \frac{1}{4(n-2)!}.$$

Step 5 Now we prove the upper-bound part in (13.14). One has, for $x > 1$

$$(13.28) \quad \mathbb{P}\{\kappa(A) > x\} = \mathbb{P}\left\{\frac{\nu_n}{\nu_1} > x^2\right\} \leq \mathbb{P}\left\{\nu_1 < \frac{L^2 n}{x^2}\right\} + \mathbb{P}\left\{\frac{\nu_n}{\lambda_1} > x^2, \nu_1 \geq \frac{L^2 n}{x^2}\right\}$$

where L is a positive number to be chosen later on.

For the first term in (13.28), we need some auxiliary lemmas that we take from Sankar et al (2002) in a modified form included in Cuesta-Albertos and Wschebor (2004):

LEMMA 13.7. Assume that $A = (a_{i,j})_{i,j=1,\dots,n}$, $a_{i,j} = m_{i,j} + g_{i,j}$ ($i, j = 1, \dots, n$), where the $m_{i,j}$'s are non-random and the $g_{i,j}$'s are i.i.d. standard normal r.v.'s. Let $v \in S^{n-1}$. Then, for $x > 0$:

$$\mathbb{P} [\|A^{-1}v\| > x] = \mathbb{P} (|\xi| < \frac{1}{x}) < \left(\frac{2}{\pi}\right)^{1/2} \frac{1}{x},$$

where ξ is a standard normal random variable.

PROOF. Because of the orthogonal invariance of the standard normal distribution, it suffices to prove the result for $v = e_1$. Denote by a_1, \dots, a_n the rows of A and α_1 the first column of A^{-1} . Clearly, α_1 is orthogonal to a_2, \dots, a_n and $\langle \alpha_1, a_1 \rangle = 1$, so that $\alpha_1 = \gamma w$, where $|\gamma| = \|\alpha_1\|$ and the unit vector w is measurable with respect to a_2, \dots, a_n and orthogonal to each one of these vectors. Also:

$$\|\alpha_1\| |\langle w, a_1 \rangle| = 1.$$

Observe now that the conditional distribution of the real-valued random variable $\langle w, a_1 \rangle$ given the random vectors a_2, \dots, a_n is standard normal, so that, under this condition, $\|\alpha_1\|$ can be written as

$$\|\alpha_1\| = \frac{1}{|\xi|}$$

where ξ is standard normal. The equalities:

$$\mathbb{P} [\|A^{-1}v\| > x] = \mathbb{P} [\|\alpha_1\| > x] = \mathbb{P} (|\xi| < \frac{1}{x}),$$

finish the proof of the Lemma. \square

LEMMA 13.8. Let $U = (U_1, \dots, U_n)$ be an n -dimensional vector chosen uniformly on S^{n-1} and let t_{n-1} be a real valued r.v. with a Student distribution with $n - 1$ degrees of freedom.

Then, if $c \in (0, n)$, we have that

$$\mathbb{P} \left[U_1^2 > \frac{c}{n} \right] = \mathbb{P} \left[t_{n-1}^2 > \frac{n-1}{n-c} c \right].$$

PROOF. U can be written as

$$U = \frac{V}{\|V\|}.$$

where $V = (V_1, \dots, V_n)$ is an n -dimensional random vector with standard normal distribution. Let us denote, to simplify the notation $K = V_2^2 + \dots + V_n^2$. Then the statement

$$\frac{V_1^2}{V_1^2 + K} > \frac{c}{n}$$

is equivalent to

$$\frac{V_1^2}{K} > \frac{c}{n-c},$$

and we have

$$\mathbb{P} \left[U_1^2 > \frac{c}{n} \right] = \mathbb{P} \left[\frac{(n-1)V_1^2}{K} > \frac{n-1}{n-c} c \right] = \mathbb{P} \left[t_{n-1}^2 > \frac{n-1}{n-c} c \right],$$

where t_{n-1} is a real valued r.v. having Student's distribution with $n - 1$ degrees of freedom. \square

LEMMA 13.9. Let C be an $n \times n$ real non-singular matrix. Then, there exists $w \in \mathbb{R}^n$, $\|w\| = 1$ such that, for every $u \in \mathbb{R}^n$, $\|u\| = 1$, one has:

$$\|Cu\| \geq \|C\| |\langle w, u \rangle|.$$

PROOF. Since $\|C\|^2$ is the maximum eigenvalue of the symmetric matrix $C^T C$, take for w an eigenvector, $\|w\| = 1$, so that $\|Cw\| = \|C\|$. Then, if $\|u\| = 1$,

$$\|C\|^2 \langle w, u \rangle = \langle Cw, Cu \rangle,$$

which implies

$$\|C\|^2 \langle w, u \rangle \leq \|C\| \|Cu\|.$$

\square

LEMMA 13.10. Assume that $A = (a_{i,j})_{i,j=1,\dots,n}$, $a_{i,j} = m_{i,j} + g_{i,j}$ ($i, j = 1, \dots, n$), where the $g_{i,j}$'s are i.i.d. standard normal r.v.'s and $M = (m_{i,j})_{i,j=1,\dots,n}$ is non random.

Then, for each $x > 0$:

$$(13.29) \quad \mathbb{P}[\|A^{-1}\| \geq x] \leq C_2(n) \frac{n^{1/2}}{x},$$

where

$$C_2(n) = \left(\frac{2}{\pi}\right)^{1/2} \left(\sup_{c \in (0,n)} \sqrt{c} \mathbb{P} \left[t_{n-1}^2 > \frac{n-1}{n-c} c \right] \right)^{-1} \leq C_2(\infty) = C_2 \simeq 2.34737\dots$$

PROOF. Let U be a random vector, independent of A with uniform distribution on S^{n-1} . Applying Lemma 13.7 we have that

$$(13.30) \quad \mathbb{P}[\|A^{-1}U\| > x] = \mathbb{E} \{ \mathbb{P}[\|A^{-1}U\| > x | U] \} \leq \left(\frac{2}{\pi}\right)^{1/2} \frac{1}{x}.$$

Now, since if w_A , $\|w_A\| = 1$ satisfies $\|A^{-1}w_A\| = \|A^{-1}\|$, and $\|u\| = 1$, then,

$$\|A^{-1}u\| \geq \|A^{-1}\| \times |\langle w_A, u \rangle|,$$

we have that, if $c \in (0, n)$, then

$$\begin{aligned} \mathbb{P}[\|A^{-1}U\| \geq x \left(\frac{c}{n}\right)^{1/2}] &\geq \mathbb{P}[\{\|A^{-1}\| \geq x\} \text{ and } \{|\langle w_A, U \rangle| \geq \left(\frac{c}{n}\right)^{1/2}\}] \\ &= \mathbb{E} \left\{ \mathbb{P} \left[\{\|A^{-1}\| \geq x\} \text{ and } \left\{ |\langle w_A, U \rangle| \geq \left(\frac{c}{n}\right)^{1/2} \right\} \middle| A \right] \right\} \\ &= \mathbb{E} \left\{ \mathbf{1}_{\{\|A^{-1}\| \geq x\}} \mathbb{P} \left[|\langle w_A, U \rangle| \geq \left(\frac{c}{n}\right)^{1/2} \middle| A \right] \right\} \\ &= \mathbb{E} \left\{ \mathbf{1}_{\{\|A^{-1}\| \geq x\}} \mathbb{P} \left[t_{n-1}^2 > \frac{n-1}{n-c} c \right] \right\} \\ &= \mathbb{P} \left[t_{n-1}^2 > \frac{n-1}{n-c} c \right] \mathbb{P}[\|A^{-1}\| \geq x]. \end{aligned}$$

where we have applied Lemma 13.8. From here and (13.30) we have that

$$\mathbb{P}[\|A^{-1}\| \geq x] \leq \frac{1}{\mathbb{P} \left[t_{n-1}^2 > \frac{n-1}{n-c} c \right]} \left(\frac{2}{\pi}\right)^{1/2} \frac{1}{x} \left(\frac{n}{c}\right)^{1/2}.$$

To end the proof, notice that, if g is a standard normal random variable, then

$$(13.31) \quad \begin{aligned} \sup_{c \in (0,n)} c^{1/2} \mathbb{P} \left[t_{n-1}^2 > \frac{n-1}{n-c} c \right] &\geq \sup_{c \in (0,1)} c^{1/2} \mathbb{P} \left[t_{n-1}^2 > \frac{n-1}{n-c} c \right] \\ &\geq \sup_{c \in (0,1)} c^{1/2} \mathbb{P} [t_{n-1}^2 > c] \\ &\geq \sup_{c \in (0,1)} c^{1/2} \mathbb{P} [g^2 > c] \\ &\geq 0.565^{1/2} \mathbb{P} [g^2 > 0.565] \simeq 0.3399. \end{aligned}$$

□

WE RETURN AGAIN TO THE PROOF OF THEOREM 13.4. For the first term in the right-hand side of (13.28), using the auxiliary lemmas we obtain:

$$\mathbb{P} \left\{ \nu_1 < \frac{L^2 n}{x^2} \right\} = \mathbb{P} \left\{ \|A^{-1}\| > \frac{x}{L\sqrt{n}} \right\} \leq C_2(n) \frac{Ln}{x}$$

Here,

$$C_2(n) = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \left[\sup_{0 < c < n} \sqrt{c} \mathbb{P} \left(t_{n-1}^2 > \frac{(n-1)c}{n-c} \right) \right]^{-1} \leq C_2(+\infty) \simeq 2.3473$$

where t_{n-1} is a random variable having Student's distribution with $n-1$ degrees of freedom.

Let us now turn to the second term in the right-hand side of (13.28),

$$P\left\{\frac{\nu_n}{\lambda_1} > x^2, \nu_1 \geq \frac{L^2 n}{x^2}\right\} = \int_{L^2 n x^{-2}}^{+\infty} db \int_{bx^2}^{+\infty} g(a, b) da \leq G_n(x^2)$$

with

$$G_n(y) = C_n \int_{L^2 n y^{-1}}^{+\infty} db \int_{by}^{+\infty} \frac{\exp(-(a+b)/2)}{\sqrt{ab}} a^{n-1} da,$$

using (13.27). We have:

$$(13.32) \quad G'_n(y) = C_n \left[- \int_{L^2 n y^{-1}}^{+\infty} \exp(-b/2) \sqrt{b} \exp(-(by)/2) (by)^{n-3/2} db + \right. \\ \left. L^2 n y^{-2} \int_{L^2 n}^{+\infty} \exp\left(-\frac{1}{2}\left(a + \frac{L^2 n}{y}\right)\right) a^{n-3/2} L^{-1} n^{-\frac{1}{2}} y^{\frac{1}{2}} da \right]$$

which implies

$$\begin{aligned} -G'_n(y) &\leq C_n y^{n-3/2} \int_{L^2 n y^{-1}}^{+\infty} \exp\left(-\frac{b(1+y)}{2}\right) b^{n-1} db \\ &= \frac{y^{-3/2}}{4(n-2)!} \left(\frac{y}{1+y}\right)^n 2^n \int_{\frac{L^2 n}{2y}(1+y)}^{+\infty} e^{-z} z^{n-1} dz \\ &\leq \frac{y^{-3/2}}{4(n-2)!} 2^n \int_{\frac{L^2 n}{2}}^{+\infty} e^{-z} z^{n-1} dz \end{aligned}$$

Put $I_n(a) = \int_a^{+\infty} e^{-z} z^{n-1} dz$. On integrating by parts we get:

$$I_n(a) = e^{-a} [a^{n-1} + (n-1)a^{n-2} + (n-1)(n-2)a^{n-3} + \dots + (n-1)!]$$

so that for $a > 2.5m$

$$I_n(a) \leq \frac{5}{3} e^{-a} a^{n-1}.$$

If $L^2 > 5$ we obtain the bound

$$-G'_n(y) \leq D_n y^{-3/2} \quad \text{with} \quad D_n = \frac{5}{6} \frac{n^{n-1}}{(n-2)!} L^{2(n-1)} \exp\left(-\frac{L^2 n}{2}\right)$$

We now apply Stirling's formula (Abramovitz and Stegun, 1968) i.e. for all $x > 0$

$$\Gamma(x+1) \exp\left(-\frac{1}{12x}\right) \leq \left(\frac{x}{e}\right)^x \sqrt{2\pi x} \leq \Gamma(x+1),$$

to get

$$D_n \leq \frac{5\sqrt{2}}{12\sqrt{\pi}L^2} \frac{n}{\sqrt{n-2}} \exp\left(-n \frac{L^2 - 4\log(L) - 2}{2}\right) \leq \frac{5\sqrt{2}}{12\sqrt{\pi}L^2} n.$$

if we choose for L the only root larger than 1 of the equation $L^2 - 4\log(L) - 2 = 0$ (check that $L \approx 2.3145$). To finish,

$$0 \leq G_n(y) = \int_y^{+\infty} -G'_n(t) dt < D_n \int_y^{+\infty} \frac{dt}{t^{3/2}} = 2D_n y^{-\frac{1}{2}}.$$

Replacing y by x^2 and performing the numerical evaluations, the upper bound in (13.14) follows and we get for the constant C the value 5.60.

Step 6 We consider now the lower bound in (13.14). For $\gamma > 0$ and $x > 1$, we have :

$$(13.33) \quad P\{\kappa(A) > x\} = P\left\{\frac{\nu_n}{\nu_1} > x^2\right\} \geq P\left\{\frac{\nu_n}{\nu_1} > x^2, \nu_1 < \frac{\gamma^2 n}{x^2}\right\} \\ = P\left\{\nu_1 < \frac{\gamma^2 n}{x^2}\right\} - P\left\{\frac{\nu_n}{\nu_1} \leq x^2, \nu_1 < \frac{\gamma^2 n}{x^2}\right\}$$

A lower bound for the first term in the right-hand side of (13.33) is obtained using the following inequality, that we state as a new auxiliary lemma. The reader can find similar statements in the papers by Szarek (1991) and Edelman (1988).

LEMMA 13.11. *If $0 < a < 1/n$, then*

$$P\{\nu_1 < a\} \geq \beta\sqrt{an},$$

where we can choose for β the value $\beta = \left(\frac{2}{3}\right)^{3/2}e^{-1/3}$

PROOF. Define the index $i_X(t)$ of a critical point $t \in S^{n-1}$ of the function X as the number of negative eigenvalues of $X''(t)$. For each $a > 0$ put

$$N_i(a) = \#\{t \in S^{n-1} : X(t) = t^T B t < a, X'(t) = 0, i_X(t) = i\},$$

for $i = 0, 1, \dots, n-1$. One easily checks that if the eigenvalues of B are ν_1, \dots, ν_n , $0 < \nu_1 < \dots < \nu_n$, then:

- if $a \leq \nu_1$ then $N_i(a) = 0$
for $i = 0, 1, \dots, n-1$
- if $\nu_i < a \leq \nu_{i+1}$ then $N_k(a) = 2$
for some $i = 0, 1, \dots, n-1$ for $k = 0, \dots, i-1$
 $N_k(a) = 0$
for $k = i, \dots, n-1$
- if $\nu_n < a$ then $N_i(a) = 2$
for $i = 0, 1, \dots, n-1$

Now consider

$$M(a) = \sum_{i=0}^{n-1} (-1)^i N_i(a)$$

$M(a)$ is the Euler-Poincaré characteristic of the set $S = \{t \in S^{n-1} : X(t) < a\}$ (see for example Adler, 1981). Since:

- if $N_0(a) = 0$, then $N_i(a) = 0$ for $i = 1, \dots, n-1$, hence $M(a) = 0$ and
- if $N_0(a) = 2$, then $M(a) = 0$ or 2 ,

in any case we have the inequality:

$$M(a) \leq N_0(a).$$

Hence,

$$(13.34) \quad P\{\nu_1 < a\} = P\{N_0(a) = 2\} = \frac{1}{2}E(N_0(a)) \geq \frac{1}{2}E(M(a))$$

Given the definition of $M(a)$, its expectation can be written using the following Rice formula:

$$\begin{aligned} E(M(a)) &= \int_0^a dy \int_{S^{n-1}} E[\det(X''(t)) | X(t) = y, X'(t) = 0] p_{X(t), X'(t)}(y, 0) \sigma_{n-1}(dt) \\ &= \int_0^a \sigma_{n-1}(S^{n-1}) E[\det(X''(e_1)) | X(e_1) = y, X'(e_1) = 0] p_{X(e_1), X'(e_1)}(y, 0) dy. \end{aligned}$$

where we have used again invariance under isometries. Applying a similar Gaussian regression - similar to what we did in Step 4 - we obtain:

$$(13.35) \quad E(M(a)) = \int_0^a E[\det(Q - yI_{n-1})] \frac{\sqrt{2\pi}}{2^{n-1}} \Gamma^{-2}(n/2) \frac{\exp(-y/2)}{\sqrt{y}} dy$$

where Q is an $(n-1) \times (n-1)$ random matrix with entry i, j equal to $\langle v_i, v_j \rangle$ and v_1, \dots, v_{n-1} are i.i.d. standard normal in \mathbb{R}^{n-1} . We now use part (ii) of Lemma 13.6 :

$$(13.36) \quad E[\det(Q - yI_{n-1})] = (n-1)! \sum_{k=0}^{n-1} \binom{n-1}{k} \frac{(-y)^k}{k!}$$

Under condition $0 < a < n^{-1}$, since $0 < y < a$, as k increases the terms of the sum in the right-hand side of (13.36) have decreasing absolute value, so that:

$$E[\det(Q - yI_{n-1})] \geq (n-1)! [1 - (n-1)y].$$

Substituting into the right-hand side of (13.35), we get :

$$E[M(a)] \geq \frac{\sqrt{2\pi} (n-1)!}{2^{n-1} \Gamma^2(n/2)} J_n(a),$$

where, using again $0 < a < n^{-1}$:

$$J_n(a) = \int_0^a (1 - (n-1)y) \frac{\exp(-y/2)}{\sqrt{y}} dy \geq \int_0^a \frac{(1 - (n-1)y)}{\sqrt{y}} (1 - y/2) dy \geq \frac{4}{3} \sqrt{a},$$

by an elementary computation. Going back to (13.35), applying Stirling's formula and remarking that $(1 + 1/n)^{n+1} \geq e$ we get

$$P\{\nu_1 < a\} \geq \left(\frac{2}{3}\right)^{3/2} e^{-1/3} \sqrt{an}$$

This proves the lemma. □

ENDING THE PROOF OF THEOREM 13.4 Using Lemma 13.11, the first term in the right-hand side of (13.33) is bounded below by

$$\beta \gamma \frac{n}{x}.$$

To obtain a bound for the second term, we use again our upper bound (13.27) on the joint density $g(a, b)$, as follows:

$$\begin{aligned} (13.37) \quad P\left\{\frac{\nu_n}{\nu_1} \leq x^2, \nu_1 < \frac{\gamma^2 n}{x^2}\right\} &= \int_0^{\gamma^2 n x^{-2}} db \int_b^{bx^2} g(a, b) da \\ &\leq C_m \int_0^{\gamma^2 n x^{-2}} db \int_b^{bx^2} \frac{\exp(-(a+b)/2)}{\sqrt{ab}} a^{n-1} da \leq C_n \int_0^{\gamma^2 n x^{-2}} b(x^2 - 1) b^{-\frac{1}{2}} (bx^2)^{n-3/2} db \\ &\leq \frac{1}{4(n-2)!} \frac{x^2 - 1}{x^3} \gamma^{2n} n^{n-1} \leq \frac{\sqrt{2}}{8\sqrt{\pi}} e^n \gamma^{2n} \frac{n}{x}, \end{aligned}$$

on applying Stirling's formula. Choosing now $\gamma = 1/e$, we see that the hypothesis of Lemma 13.11 is satisfied and

$$P\left\{\frac{\nu_n}{\nu_1} \leq x^2, \nu_1 < \frac{\gamma^2 n}{x^2}\right\} \leq \frac{\sqrt{2}}{8\sqrt{\pi}} e^{-3} \frac{n}{x}$$

Replacing into (13.33), we obtain the lower bound in (13.13) with

$$c = \left(\frac{2}{3}\right)^{3/2} e^{-4/3} - \frac{\sqrt{2}}{8\sqrt{\pi}} e^{-3} \approx 0.138.$$

□

2.1. Monte-Carlo experiment. To study the tail of the distribution of the condition number of Gaussian matrices of various size, we used the following Matlab functions

- `normrnd` to simulate normal variables
- `cond` to compute the Condition number of matrix A .

Results over 40,000 simulations using Matlab are given in Table 13.1 and in Figure 13.1.

This table suggests, taking into account the simulation variability, that the constants c and C should take values smaller than 0.88 and bigger than 2.00 respectively.

TABLE 13.1. Values of the estimations $P\{\kappa(A) > mx\}$ for $x = 1, 2, 3, 5, 10, 15, 30, 50, 100$ and $m = 3, 5, 10, 30, 100, 300, 500$ by Monte-Carlo method over 40,000 simulations.

Probability	Value of x									
	1	2	3	5	10	20	30	50	100	
Lower b.: $.13/x$.13	.065	.043	.026	.013	.007	.004	.003	.001	
Upper b.: $5.6/x$	1	1	1	1	.56	.28	.187	.112	.056	
m = 3	.881	.57	.41	.26	.13	.067	.044	.027	.013	
m=5	.931	.66	.48	.30	.16	.079	.053	.033	.016	
m=10	.959	.71	.52	.34	.17	.088	.059	.035	.017	
m=30	.974	.75	.56	.36	.19	.096	.063	.038	.019	
m=100	.978	.77	.58	.38	.20	.098	.066	.040	.019	
m=300	.982	.77	.58	.38	.20	.101	.069	.041	.022	
m =500	.980	.77	.59	.38	.20	.100	.066	.039	.020	

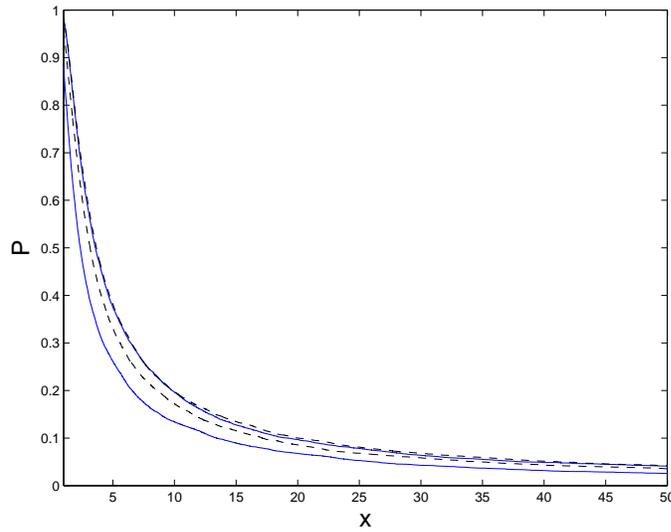


FIGURE 13.1. Values of $P\{\kappa(A) > mx\}$ as a function of x for $m = 3$, (down) 10, 100 and 500 (up)

3. Non-centered Gaussian matrices

Let $A = ((a_{ij}))_{i,j=1,\dots,n}$ be a random matrix. Throughout this section, we will assume that the a'_{ij} s are independent random variables with Gaussian distribution having expectations $m_{ij} = E(a_{ij})$ and common variance σ^2 . We denote $M = ((m_{ij}))$.

The aim of the section is to prove Theorem 13.12 below, which gives a bound for the tails of probability distribution of $\kappa(A)$. One way of looking to this result, is as follows: we start with a non-random matrix M and add noise by putting independent centered Gaussian random variables with variance σ^2 at each location. We ask for the condition number of the perturbed matrix. Notice that the starting matrix M can have arbitrarily large (or infinite) condition number, but for the new one, it turns out that we are able to give an upper-bound for $P(\kappa(A) > x)$ which has a similar form to the one in the centered case.

THEOREM 13.12. *Under the above hypotheses on A , one has, for $x > 0$:*

$$(13.38) \quad P(\kappa(A) > n.x) < \frac{1}{x} \left(\frac{1}{4\sqrt{2\pi n}} + C(M, \sigma, n) \right)$$

where

$$C(M, \sigma, n) = 7 \left(5 + \frac{4 \|M\|^2 (1 + \log n)}{\sigma^2 n} \right)^{\frac{1}{2}}$$

Remarks.

(1) Theorem 13.12 implies that if $0 < \sigma \leq 1$ and $\|M\| \leq 1$ then, for $x > 0$:

$$(13.39) \quad P(\kappa(A) > n \cdot x) < \frac{20}{\sigma x}$$

This is an immediate consequence of the statement in the theorem.

(2) With similar calculations than the ones we will perform for the proof of Theorem 13.12, one can improve somewhat the constants in (13.38) and (13.39).

PROOF. Due to the homogeneity of $\kappa(A)$, with no loss of generality we may assume $\sigma = 1$, changing the expected matrix M by $\frac{1}{\sigma}M$ in the final result.

We follow closely the proof of Theorem 13.4 in the previous section, with some changes to adapt it to the present conditions. In exactly the same way, we apply Rice formula and prove that the joint density $g(a, b)$, $a > b$ of the random variables ν_n, ν_1 is given by:

$$(13.40) \quad g(a, b) = \int_V \mathbb{E} (\Delta(s, t) \mathbb{1}_{\{X''(s) < 0, X''(t) > 0\}} | X(s) = a, X(t) = b, Y(s, t) = 0) \cdot p_{X(s), X(t), Y(s, t)}(a, b, 0) \sigma_V(d(s, t)).$$

where the notations are also borrowed from the previous section.

Next, we compute the ingredients in the right-hand side of (13.40). This has some differences with the centered case. Put $a_{ij} = m_{ij} + g_{ij}$ with the g_{ij} 's i.i.d. standard normal and $G = ((g_{ij}))$.

For each $(s, t) \in V$, we take an orthonormal basis of \mathbb{R}^n so that its first two elements are respectively s and t , say $\{s, t, w_3, \dots, w_m\}$. When expressing the linear transformation $x \rightsquigarrow Ax$ ($x \in \mathbb{R}^n$) in this new basis, we denote $A^{s,t}$ the associated matrix and by $a_{ij}^{s,t}$ its i, j entry. In a similar way we get $G^{s,t}, M^{s,t}, B^{s,t}$. Notice that $G^{s,t}$ has the same law as G , but the non-random part $M^{s,t}$ can vary with the point (s, t) .

We denote by $B_1^{s,t}$ (respectively $B_2^{s,t}$) the $(n - 1) \times (n - 1)$ matrix obtained from $B^{s,t}$ by suppressing the first (respectively the second) row and column. $B_{1,2}^{s,t}$ denotes the $(n - 2) \times (n - 2)$ matrix obtained from $B^{s,t}$ by suppressing the first and second row and column.

To get an estimate for the right-hand side in (13.40) we start with the density $p_{X(s), X(t), Y(s, t)}(a, b, 0)$. We denote $B^{s,t} = ((b_{ij}^{s,t}))$ (and similarly for the other matrices).

We have:

$$\begin{aligned} X(s) &= b_{11}^{s,t} \\ X(t) &= b_{22}^{s,t} \\ X''(s) &= B_1^{s,t} - b_{11}^{s,t} I_{n-1} \\ X''(t) &= B_2^{s,t} - b_{22}^{s,t} I_{n-1}. \end{aligned}$$

Take the following orthonormal basis of the subspace $W_{(s,t)}$:

$$\{(w_3, 0), \dots, (w_n, 0), (0, w_3), \dots, (0, w_n), \frac{1}{\sqrt{2}}(t, s)\} = L_{s,t}$$

Since the expression of $Y(s, t)$ in the canonical basis of \mathbb{R}^{2n} is:

$$Y(s, t) = (0, b_{21}^{s,t}, b_{31}^{s,t}, \dots, b_{n1}^{s,t}, b_{12}^{s,t}, 0, b_{32}^{s,t}, \dots, b_{n2}^{s,t}, b_{12}^{s,t})^T,$$

it is written in the orthonormal basis $L_{s,t}$ as the linear combination:

$$Y(s, t) = \sum_{i=3}^n [b_{i1}^{s,t} \cdot (w_i, 0) + b_{i2}^{s,t} \cdot (0, w_i)] + \sqrt{2} b_{12}^{s,t} \cdot \left[\frac{1}{\sqrt{2}}(t, s) \right]$$

It follows that the joint density of $X(s), X(t), Y(s, t)$ appearing in (13.40) in the space $\mathbb{R} \times \mathbb{R} \times W_{(s,t)}$ is the joint density of the r.v.'s

$$b_{11}^{s,t}, b_{22}^{s,t}, \sqrt{2}b_{12}^{s,t}, b_{31}^{s,t}, \dots, b_{n1}^{s,t}, b_{32}^{s,t}, \dots, b_{n2}^{s,t}$$

at the point $(a, b, 0)$. To compute this density, first compute the joint density q of

$$b_{31}^{s,t}, \dots, b_{n1}^{s,t}, b_{32}^{s,t}, \dots, b_{n2}^{s,t},$$

given $a_1^{s,t}, a_2^{s,t}$, where $a_j^{s,t}$ denotes the j -th column of $A^{s,t}$, with the additional conditions that

$$\|a_1^{s,t}\| = b_{11}^{s,t} = a, \quad \|a_2^{s,t}\| = b_{22}^{s,t} = b, \quad \langle a_1^{s,t}, a_2^{s,t} \rangle = b_{12}^{s,t} = 0.$$

q is the normal density in $\mathbb{R}^{2(n-2)}$, with the same variance matrix as in the centered case, that is

$$\begin{pmatrix} a.I_{n-2} & 0 \\ 0 & b.I_{n-2} \end{pmatrix}.$$

but not necessarily centered.

So, the conditional density q is bounded above by

$$(13.41) \quad \frac{1}{(2\pi)^{n-2}} \frac{1}{(ab)^{(n-2)/2}}.$$

Our next task is to obtain an upper bound useful for our purposes for the density of the triplet

$$(b_{11}^{s,t}, b_{22}^{s,t}, b_{12}^{s,t}) = (\|a_1^{s,t}\|^2, \|a_2^{s,t}\|^2, \langle a_1^{s,t}, a_2^{s,t} \rangle)$$

at the point $(a, b, 0)$ which together with (13.41) will provide an upper bound for $p_{X(s), X(t), Y(s,t)}(a, b, 0)$.

We do this in the next Lemma, which we will apply afterwards with $\xi = a_1^{s,t}, \eta = a_2^{s,t}$.

LEMMA 13.13. *Let ξ, η be to independent Gaussian vectors in \mathbb{R}^n ($n \geq 2$), $E(\xi) = \mu, E(\eta) = \nu, \text{Var}(\xi) = \text{Var}(\eta) = I_n$.*

Then, the density p of the random triplet $(\|\xi\|^2, \|\eta\|^2, \langle \xi, \eta \rangle)$ satisfies the following inequality, for $a \geq 4 \|\mu\|^2$:

$$(13.42) \quad p(a, b, 0) \leq \frac{1}{4(2\pi)^n} \sigma_{n-1} \sigma_{n-2} (ab)^{(n-3)/2} \exp\left(-\frac{a}{8}\right) \quad (a, b > 0)$$

PROOF. . Let $F : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^3$ be the function

$$F(x, y) = \left(\|x\|^2, \|y\|^2, \langle x, y \rangle \right)^T$$

According to the co-area formula Prop 6.30, the density p at the point $(a, b, 0)$ can be written as

$$p(a, b, 0) = \int_{F^{-1}(a,b,0)} \left(\det \left[F'(x, y) \cdot (F'(x, y))^T \right] \right)^{-\frac{1}{2}} \frac{1}{(2\pi)^n} e^{-\frac{1}{2}[\|x-\mu\|^2 + \|y-\nu\|^2]} d\gamma(x, y)$$

where γ denotes the geometric measure on $F^{-1}(a, b, 0)$.

Recall the manifold $V_{a,b}$ given by the set of equations

$$\|x\|^2 = a, \quad \|y\|^2 = b, \quad \langle x, y \rangle = 0$$

Using Lemma 13.5,

$$\gamma(V_{a,b}) = (a + b)^{\frac{1}{2}} \sigma_{n-1} \sigma_{n-2} (ab)^{\frac{n-2}{2}}.$$

On the other hand,

$$F'(x, y) = \begin{pmatrix} 2.x^T & 0 \\ 0 & 2.y^T \\ y^T & x^T \end{pmatrix}$$

so that if $(x, y) \in F^{-1}(a, b, 0)$, one gets:

$$\det \left[F'(x, y) \cdot (F'(x, y))^T \right] = 16.ab(a + b).$$

Replacing into (13.43) and taking into account condition $a \geq 4 \|\mu\|^2$, the result in the lemma follows. \square

Summing up this part, (13.41) plus (13.42) imply that

$$(13.43) \quad p_{X(s), X(t), Y(s,t)}(a, b, 0) \leq \frac{1}{2^{2n-\frac{3}{2}} \pi^{n-2}} \frac{1}{\Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{n-1}{2}\right)} \frac{\exp\left(-\frac{a}{8}\right)}{\sqrt{ab}}.$$

We now consider the conditional expectation in (13.40).

First, observe that the $(2n-3)$ -dimensional tangent space to V at the point (s, t) is parallel to the orthogonal complement in $\mathbb{R}^n \times \mathbb{R}^n$ of the triplet of vectors $(s, 0); (0, t); (t, s)$.

To compute the associated matrix for $Y'(s, t)$ take the set

$$\{(w_3, 0), \dots, (w_n, 0), (0, w_3), \dots, (0, w_n), \frac{1}{\sqrt{2}}(t, -s)\} = K_{s,t}.$$

as orthonormal basis in the tangent space. As for the codomain of Y , we take the canonical basis in \mathbb{R}^{2n} . A direct calculation gives :

$$Y'(s, t) = \begin{pmatrix} -v^T & 0_{1,n-2} & -\frac{1}{\sqrt{2}}b_{21}^{s,t} \\ w^T & 0_{1,n-2} & \frac{1}{\sqrt{2}}(-b_{11}^{s,t} + b_{22}^{s,t}) \\ B_{12}^{s,t} - b_{11}^{s,t}I_{n-2} & 0_{n-2,n-2} & \frac{1}{\sqrt{2}}w \\ 0_{1,n-2} & -w^T & \frac{1}{\sqrt{2}}(-b_{11}^{s,t} + b_{22}^{s,t}) \\ 0_{1,n-2} & v^T & \frac{1}{\sqrt{2}}b_{21}^{s,t} \\ 0_{n-2,n-2} & B_{12}^{s,t} - b_{22}^{s,t}I_{n-2} & -\frac{1}{\sqrt{2}}v \end{pmatrix}$$

where $v^T = (b_{31}^{s,t}, \dots, b_{n1}^{s,t})$, $w^T = (b_{32}^{s,t}, \dots, b_{n2}^{s,t})$, $0_{i,j}$ is a null matrix with i rows and j columns. The columns represent the derivatives in the directions of $K_{s,t}$ at the point (s, t) . The first n rows correspond to the components of $\pi_s(Bs)$, the last n ones to those of $\pi_t(Bt)$.

Thus, under the conditioning in (13.40),

$$Y'(s, t) = \begin{pmatrix} 0_{1,n-2} & 0_{1,n-2} & 0 \\ 0_{1,n-2} & 0_{1,n-2} & \frac{1}{\sqrt{2}}(b-a) \\ B_{12}^{s,t} - aI_{n-2} & 0_{n-2,n-2} & 0_{m-2,1} \\ 0_{1,n-2} & 0_{1,n-2} & \frac{1}{\sqrt{2}}(b-a) \\ 0_{1,n-2} & 0_{1,n-2} & 0 \\ 0_{n-2,n-2} & B_{12}^{s,t} - bI_{n-2} & 0_{n-2,1} \end{pmatrix}$$

and

$$\left[\det \left[(Y'(s, t))^T Y'(s, t) \right] \right]^{\frac{1}{2}} = |\det(B_{12}^{s,t} - aI_{n-2})| |\det(B_{12}^{s,t} - bI_{n-2})| (a-b)$$

Since $B_{12}^{s,t} \succ 0$ one has

$$|\det(B_{12}^{s,t} - aI_{n-2})| \mathbf{1}_{B_{12}^{s,t} - aI_{n-2} < 0} \leq a^{n-2}$$

and the conditional expectation in (13.40) is bounded by:

(13.44)

$$a^{n-1} \mathbf{E} \left[|\det(B_{12}^{s,t} - bI_{n-2})| \mathbf{1}_{B_{12}^{s,t} - bI_{n-2} > 0} | b_{11}^{s,t} = a, b_{22}^{s,t} = b, b_{12}^{s,t} = 0, b_{i1}^{s,t} = b_{i2}^{s,t} = 0 \ (i = 3, \dots, n) \right].$$

We further condition on $a_1^{s,t}$ and $a_2^{s,t}$, with the additional requirement that $\|a_1^{s,t}\|^2 = a$, $\|a_2^{s,t}\|^2 = b$, $\langle a_1^{s,t}, a_2^{s,t} \rangle = 0$. Since unconditionally, a_3, \dots, a_n are independent Gaussian vectors in \mathbb{R}^n each having variance equal to 1 and mean smaller or equal to $\|M\|$, under the conditioning, their joint law becomes the law of $(n-2)$ Gaussian vectors in \mathbb{R}^{n-2} , independent of the condition and also having variance equal to 1 and mean with Euclidean norm smaller than or equal to $\|M\|$.

As a consequence, the conditional expectation in (13.44) is bounded by

$$\mathbf{E}(\det(C^{s,t}))$$

where $C^{s,t}$ is an $(n-2) \times (n-2)$ random matrix, $C^{s,t} = ((c_{ij}^{s,t}))$, $c_{ij}^{s,t} = \langle u_i^{s,t}, u_j^{s,t} \rangle$, $(i, j = 3, \dots, n)$,

$$u_i^{s,t} = \zeta_i + \mu_i^{s,t} \quad i = 3, \dots, n,$$

ζ_3, \dots, ζ_n are i.i.d. standard normal in \mathbb{R}^{n-2} and $\|\mu_i^{s,t}\| \leq \|M\|$ for $i = 3, \dots, n$.

The usual argument to compute $\det(C^{s,t})$ as the square of the volume in \mathbb{R}^{n-2} of the set of linear combinations of the form $\sum_{i=3}^{i=n} \nu_i u_i^{s,t}$ with $0 \leq \nu_i \leq 1$ ($i = 3, \dots, n$), shows that

$$\begin{aligned} \mathbb{E}(\det(C^{s,t})) &\leq \left(1 + \|M\|^2\right) \left(2 + \|M\|^2\right) \dots \left(n - 2 + \|M\|^2\right) \\ &= (n-2)! \prod_{i=1}^{i=n-2} \left(1 + \frac{\|M\|^2}{i}\right) \\ &\leq (n-2)! \left[\left(1 + \|M\|^2 \frac{1 + \log n}{n}\right) \right]^n \end{aligned}$$

where we have bounded the geometric mean by the arithmetic mean.

Replacing in (13.44) and on account of the bound (13.43) we get from (13.40) the following bound for the joint density, valid for $a \geq 4\|M\|^2$:

$$(13.45) \quad g(a, b) \leq \bar{C}_n \frac{e^{-\frac{a}{8}}}{\sqrt{ab}} a^{n-1}$$

where

$$\bar{C}_n = \frac{1}{4(n-2)!} \left[1 + \|M\|^2 \frac{1 + \log n}{n} \right]^n.$$

We now turn to the proof of (13.38). One has, for $x > 1$:

$$(13.46) \quad \mathbb{P}(\kappa(A) > x) = \mathbb{P}\left(\frac{\nu_n}{\nu_1} > x^2\right) \leq \mathbb{P}\left(\nu_1 < \frac{L^2 n}{x^2}\right) + \mathbb{P}\left(\frac{\nu_n}{\nu_1} > x^2, \nu_1 \geq \frac{L^2 n}{x^2}\right)$$

where L is a positive number to be chosen later on.

For the first term in (13.46), we use Lemma 13.10.

$$(13.47) \quad \mathbb{P}\left(\nu_1 < \frac{L^2 n}{x^2}\right) = \mathbb{P}\left(\|A^{-1}\| > \frac{x}{L\sqrt{n}}\right) \leq C_2 \frac{Ln}{x}$$

Impose first on L the condition

$$L^2 n \geq 4\|M\|^2$$

so that for the second term in (13.46) we can make use of the bound (13.45) on the joint density $g(a, b)$:

$$(13.48) \quad \mathbb{P}\left(\frac{\nu_n}{\nu_1} > x^2, \nu_1 \geq \frac{L^2 n}{x^2}\right) = \int_{L^2 n x^{-2}}^{+\infty} db \int_{bx^2}^{+\infty} g(a, b) da \leq H_n(x^2)$$

with

$$H_n(y) = \bar{C}_n \int_{L^2 n y^{-1}}^{+\infty} db \int_{by}^{+\infty} \frac{\exp(-\frac{a}{8})}{\sqrt{ab}} a^{n-1} da,$$

We have:

$$H'_n(y) = \bar{C}_n \left[- \int_{L^2 n y^{-1}}^{+\infty} \exp(-\frac{by}{8})(by)^{n-1} \frac{db}{\sqrt{y}} + \frac{L n^{\frac{1}{2}}}{y^{\frac{3}{2}}} \int_{L^2 n}^{+\infty} \exp(-\frac{a}{4}) a^{n-3/2} da \right]$$

which implies

$$\begin{aligned} -H'_n(y) &\leq \bar{C}_n y^{n-3/2} \int_{L^2 n y^{-1}}^{+\infty} \exp\left(-\frac{by}{8}\right) b^{n-1} db \\ &\leq \frac{\bar{C}_n}{y^{3/2}} 8^n \int_{\frac{L^2 n}{8}}^{+\infty} e^{-z} z^{n-1} dz \leq \frac{\bar{C}_n}{y^{3/2}} 8^n \frac{5}{3} e^{-\frac{L^2 n}{8}} \left(\frac{L^2 n}{8}\right)^{n-1} = \bar{D}_n \frac{1}{y^{3/2}} \end{aligned}$$

if we choose $L^2 > 20$.

So,

$$(13.49) \quad H_n(y) = - \int_y^{+\infty} H'_n(s) ds \leq \bar{D}_n \int_y^{+\infty} \frac{ds}{s^{\frac{3}{2}}} \leq 2\bar{D}_n \frac{1}{y^{1/2}}$$

where

$$\bar{D}_n \leq \frac{10}{3\sqrt{2\pi}L^2} \frac{n}{\sqrt{n-2}} \exp \left[\left(1 - \frac{L^2}{8} + \log L^2 + \log \theta \right) n \right]$$

where $\theta = 1 + \|M\|^2 \frac{1+\log n}{n}$.

Choosing

$$L = 2\sqrt{2}(1 + 4\theta)^{\frac{1}{2}}$$

conditions $L^2 > 20$ and $L^2 n \geq 4\|M\|^2$ are verified and $1 - \frac{L^2}{8} + \log L^2 + \log \theta < 0$.

Hence,

$$2\bar{D}_n \leq \frac{1}{4} \sqrt{\frac{n}{2\pi}}.$$

On account of (13.47), (13.48) and (13.49), replacing in the right-hand side of (13.46), inequality (13.38) in the statement of the theorem follows. \square

Notations

Notation	Meaning
$(const)$	a positive constant, its value can change from one occurrence to an other.
A^C	Complement of the set A
ϕ	Density of the standard normal distribution in \mathbb{R}
Φ	Cumulative distribution function of the standard normal distribution in \mathbb{R}
χ_d^2	Chi-square distribution with d degrees of freedom
λ, λ_d	depending on the context : Lebesgue measure, in \mathbb{R}^d
	Spectral moment of order d
$(const)$	A non important constant, it value can change from one occurrence to another
$p_\xi(x)$	Density of the random variable or vector ξ at point x
$B(u; r)$	Open ball with center u and radius r
$\overline{B}(u; r)$	Closed ball with center u and radius r
$N_u(f, T)$ or N_u	number of roots of $f(t) = u$ that belong to the set T
$U_u(f, T)$ or U_u	number of up-crossings of the level u by the function f on the set T
$D_u(f, T)$ or D_u	number of down-crossings of the level u by the function f on the set T
M_T^X or M or M_T	$\sup_{t \in T} X_t$
$F_M(u)$	$P\{M \leq u\}$
$\mu(X)$	a median for the real random variable X
$r(s, t)$	the covariance of a (often Gaussian) stochastic process
$r_{ij}(s, t)$	$\frac{\partial^{i+j} r}{\partial s^i \partial t^j}$
$\Gamma(t)$	the covariance function of a stationary process : $\text{Cov}(X(z), X(z+t))$
càd-làg	French acronym for “continue à droite et limité à gauche”: right continuous with left limits
z^+	$\sup(0, z)$
z^-	$-\inf(0, z)$
$z^{[k]}$	$z(z-1) \dots (z-k+1)$ z and k positive integers, $l \leq z$
$\nu_k(u, T)$	$E(U_u^{[k]})$.
$\tilde{\nu}_k(u, T)$	$E(U_u^{[k]} \mathbf{1}_{X(0) < u})$.
\mathcal{C}_u	level set : $\{t \in S : X(t) = u\}$
$\ \cdot\ _p$	L^p -norm, $p > 0$
$\ \cdot\ _\infty$	the sup norm
\approx	Equivalence of two functions
\simeq	Numerical approximative equality
$\stackrel{D}{=}$	equality in distribution
$\{W(t) : t \geq 0\}$	the Wiener process
$\{B_H(t) : t \geq 0\}$	fractional Brownian motion
(Ω, \mathcal{A}, P)	a probability space
λ_k	k -th moment of the spectral measure
$M \succ 0$	the symmetric square matrix M is positive definite
$M \prec 0$	the symmetric square matrix M is negative definite
$M_{u,1}(X, S),$	number of local maxima of the random field X , having value bigger than u and belonging to the set S .
$M_{u,2}(X, S),$	number of critical points of the random field X , having value bigger than u and belonging to the set S .
$H_n(x),$	Hermite polynomials see p. 156
$\bar{H}_n(x),$	modified Hermite polynomials see p. 156
$GOE,$	Gaussian Orthogonal Ensemble
$I_n,$	identity $n \times n$ real matrix
σ_d	geometrical measure of size d
$\{e_1, \dots, e_n\}$	the canonical basis.

References

- Abrahams, J. (1982). The zero-crossing problem for some non-stationary Gaussian processes, *IEEE Trans. Inform. Theory*, Vol IT-28, 677-678.
- Abrahams, J. (1986). A Survey of Recent Progress on Level-Crossing Problems for Random Processes, in *Communications and Networks, A Survey of Recent Advances*, Ed. I.F. Blake and H.V. Poor, Springer-Verlag, pp 6-25.
- Abramovitz, M and Stegun, I. (1968). *Handbook of Mathematical Functions with formulas, graphs and mathematical tables*. National Bureau of Standards.
- Adler, R.J. (1976). On generalizing the notion of up-crossings to random fields. *Adv. Appl. Prob.*, 8, 789-805.
- Adler, R.J. (1976). Excursions Above a Fixed Level by n-dimensional Random Fields, *J. Appl. Probab.*, 13, 276-289.
- Adler, R.J. (1977). Excursions above high levels by Gaussian Random Fields, *Stoch. Proc. and Applic.* 5, 21-25.
- Adler, R.J. (1981). *The Geometry of Random Fields*. Wiley, New York.
- Adler, R.J. (1984). The supremum of a particular Gaussian field, *Ann. Probab.*, Vol 12, 436-444.
- Adler, R.J. (1990). *An Introduction to Continuity, Extrema and Related Topics for General Gaussian Processes*, IMS, Hayward, Ca.
- Adler, R.A. (2000). On excursion sets, tube formulae and maxima of Gaussian fields. *Ann. Appl. Probab.*, 10, 1-74.
- Adler, R.A. (2004). Gaussian random fields on manifolds. *Stochastic Analysis, Random Fields and Applications*, 4, 3-20.
- Adler, R.J. and Firman, D. (1981). A non-Gaussian model for random surfaces. *Trans. Royal Soc. London, A*, 303, 433-462.
- Adler, R. and Hasofer, A.M. (1976). Level Crossings of Random Fields *Ann. of Probab.* , N1, 1-12.
- Adler, R.J. and Samorodnisky, G. (1987). Tail behaviour for the suprema of Gaussian processes with applications to empirical processes. *Ann. Probab.*, 15, 1339-1351.
- Adler, R.A. and Samorodnisky, G. (1997). Level crossings of absolutely continuous stationary symmetric α -stable processes. *Ann. Appl. Probab.*, 6, 460-493.
- Adler, R.A. , Samorodnisky, G. and Gadrich T.(1993). The expected number of level crossings for stationary harmonisable symmetric stable processes, *Ann. Appl. Probab*, 2, 553-575.
- Adler, R.J. and Wilson, R.J. (1982). The structure of Gaussian fields near a level crossing. *Adv. Appl. Prob.*, 14, 543-565.
- Adler, R.J. and Taylor, J. (2007). *Random fields and geometry* Springer.
- Andrews, D.W.K. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica*, 62, 1383-1414.
- Armentano, D. and Wschebor, M. (2008). Random systems of polynomial equations. The expected number of roots under smooth analysis. *Bernoulli*, to appear.

- Aronowich, M. and Adler, R.J. (1985). Behaviour of χ^2 -processes at Extrema, *Adv. Appl. Prob.* 17, 280-297.
- Azaïs, J-M (1989). Approximation des trajectoires et temps local des diffusions. *Ann. Inst. H. Poincaré, Probab. Statist.*, 25, No.2, 175-194.
- Azaïs, J-M, Bardet, J-M and Wschebor, M. (2002). On the Tails of the Distribution of the Maximum of a Smooth Stationary Gaussian Process. *ESAIM: Probability and Statistics*, Vol. 6, 177-184.
- Azaïs, J-M., Cierco-Ayrolles, C. and Croquette, A. (1999). Bounds and asymptotic expansions for the distribution of the maximum of a smooth stationary Gaussian process. *ESAIM: Probability and Statistics*, 3, 107-129.
- Azaïs, J-M. and Cierco-Ayrolles, (2002). An asymptotic test for quantitative gene detection. *Ann. Inst. H. Poincaré, Probab. Statist.*, 38, 1087-1092.
- Azaïs, J-M., Déjean S., León J. and Zwolska F. (2008). Transformed Gaussian stationary models for ocean waves. *Submitted*.
- Azaïs, J-M. and Delmas, C. (2002). Asymptotic expansions for the distribution of the maximum of a Gaussian random fields. *Extremes* (2002)5(2), 181-212.
- Azaïs, J-M., Gassiat, E. and Mercadier, C. (2006). Asymptotic distribution and power of the likelihood ratio test for mixtures: bounded and unbounded case. *Bernoulli*, vol 12(5), 2006, 775-799.
- Azaïs, J-M., Gassiat, E. and Mercadier, C. (2007). The likelihood ratio test for general mixture models with or without structural parameter. *To appear in ESAIM: Probability and Statistics*
- Azaïs, J-M., León J. and Ortega J. (2005). Geometrical Characteristic of Gaussian sea Waves. *Journal of Applied Probability*, 42, 1-19.
- Azaïs, J-M. and Mercadier C. (2003). Asymptotic Poisson character of extremes in non-stationary Gaussian models. *Extremes*, 6, 4, 301-318.
- Azaïs, J-M. and Wschebor M. (1996). Almost sure oscillation of certain random processes. *Bernoulli*, Vol. 2, Fasc. 3, 257-270.
- Azaïs, J-M. and Wschebor M. (1997). Oscillation presque sûre de martingales continues. *Sém. Probab. XXXI, Lecture Notes in Mathematics N 1655*, Springer Verlag, 69-76.
- Azaïs, J-M. and Wschebor M. (1997). Une formule pour calculer la distribution du maximum d'un processus stochastique. *C.R. Acad. Sci. Paris*, t. 324, série I, 225-230.
- Azaïs, J-M. and Wschebor M. (1999). Régularité de la loi du maximum de processus gaussiens réguliers. *C.R. Acad. Sci. Paris*, t. 328, Sér. I, 333-336.
- Azaïs, J-M and Wschebor, M. (2001). On the Regularity of the Distribution of the Maximum of one-parameter Gaussian Processes. *Probability Theory and Related Fields* 119, 70-98.
- Azaïs, J-M and Wschebor, M. (2002). The Distribution of the Maximum of a Gaussian Process: Rice Method Revisited. *In and out of equilibrium: probability with a physical flavour*, Progress in Probability, Birkhäuser, 321-348.
- Azaïs, J-M. and Wschebor, M. (2005a). On the Distribution of the Maximum of a Gaussian Field with d Parameters, *Annals of Applied Probab.*, Vol. 15, No. 1A, 254-278.
- Azaïs J-M. and Wschebor M. (2005b). On the roots of a random system of equations. The theorem of Shub and Smale and some extensions. *Found. Comp. Math.*, 125-144.
- Azaïs, J-M. and Wschebor, M. (2005c). Upper and lower bounds for the tails of the distribution of the condition number of a Gaussian matrix, *SIAM J. Matrix Anal. Appl.*, Vol. 26, No.2, 426-440.
- Azaïs J-M. and Wschebor, M. (2006). A self contained proof of the Rice formula for random fields. Preprint available at <http://www.lsp.ups-tlse.fr/Azais/publi/completestproof.pdf>.

- Azaïs, J-M. and Wschebor, M (2008). A general formula for the distribution of the maximum of a Gaussian field and the approximation of the tail. *Stochastic Processes and their Applications*, Vol 118/7, pp. 1190-1218.
- Bar David, I.; Nemirovsky, A. (1972). Level crossings of non-differentiable shot processes. *IEEE Trans. Inform. Theory*, Vol. IT-18, 27-34.
- Baxevani A., Podgórski K. and Rychlik I, (2003). Velocities for Random Surfaces. *Probabilistic Engineering Mechanics* ,**18**, 251-271.
- Bharucha-Reid, A.T. and Sambandham, M (1986). *Random polynomials*. Probability and Mathematical Statistics. Academic Press Inc., Orlando, FL.
- Belayev, Y. (1961). Continuity and Hölder conditions for sample functions of stationary Gaussian processes. *Proc. Fourth Berk. Symp. on Math. Stat. and Probability*, 2, 23-33.
- Belayev, Y. (1966). On the number of intersections of a level by a Gaussian Stochastic process. *Theory Prob. Appl.*, 11, 106-113.
- Belayev, Y. (1967). Bursts and Shines of Random Fields, *Soviet Math. Doklady*, Vol. 8, N 5, 1107-1109.
- Belayev, Y. (1968). On the Number of Exits Across the Boundary of a Region by a Vector Stochastic Process, *Th. Prob. Appl.*, Vol. XIII, 320-324.
- Belayev, Y. (1970). Distribution of the Maximum of a Random Field and its Application to Reliability Problems. *Engineering Cybernetics*, N 2, 269-276.
- Belayev, Y. (1972a). Point processes and first passage problems. Proceedings of the 6th Berkeley Symp. on Math. Statist. and Prob., Berkeley. 1-17.
- Belayev, Y. (1972b). On the mean number of crossings of a level by a stationary Gaussian stochastic process. *Proc. Sixth Berkeley Symp. Math. Stat. Prob.*, 3,1-17.
- Belayev, Y. and Malevich, T.L. (1968). Asymptotic Normality of the Number of Intersections of a Fixed Line by a Stationary Gaussian Process, *Th. Prob. Appl.* XIII, 543-544.
- Belayev, Y. and Piterbarg, V.I. (1972). Asymptotics of the average number of overshoot of a Gaussian Field beyond a high level, *Soviet Math. Doklady*, Vol 13, N 2, 309-313.
- Benzaquen, S. and Cabaña, E.M. (1982). The expected measure of the level set of a regular stationary Gaussian process. *Pacific J. Math.*, 103, 1, 9-16.
- Berman, S.M. (1971a). Asymptotic Independence of the Numbers of High and Low Level Crossings of Stationary Gaussian Processes, *The Ann. Math. Stat.*, Vol. 42, N 3, 927-945.
- Berman, S.M. (1971b). Excursions above high levels for stationary Gaussian processes. *Pacific J. Math.*, 3, 663-79.
- Berman, S.M. (1974). Sojourns and Extremes of Gaussian Processes, *Ann. Probab.* 2, 999-1026.
- Berman, S.M. (1979). High Level Sojourns for Strongly Dependent Gaussian Processes, *Zeit. Wahr.*, 50, 223-236.
- Berman, S.M. (1985a). An asymptotic formula for the distribution of the maximum of a Gaussian process with stationary increments. *J. Appl. Prob.*, 22, 454-460.
- Berman, S.M. (1985b). The maximum of a Gaussian process with nonconstant variance. *Ann. Inst. H. Poincaré Probab. Statist.*, 21, 383-391.
- Berman, S.M. (1985c). Sojourns above a high level for a Gaussian process with a point of maximum variance Comm. Pure Appl. Math. 38, N 5, 519-528.
- Berman, S.M. (1986). The supremum of a process with stationary independent and symmetric increments Stoch. Proc. Appl., 23, N 2, 281-290.
- Berman, S.M. (1987). Extreme sojourns of a Gaussian process with a point of maximum variance Prob. Th. Rel. Fields, 74, N 1, 113-124.

- Berman, S.M. (1992a). *Sojourns and extremes of stochastic processes*, The Wadworth and Brooks, Probability Series.
- Berman, S.M. (1992b). A Central Limit Theorem for the renormalized self-intersection local time of a stationary vector Gaussian process. *Ann. Probab.* 20, 61-81 .
- Berman, S.M. and Kôno, N. (1989). The maximum of a gaussian process with non-constant variance: a sharp bound for the distribution of the tail. *Ann. Probab.*, 17, 632-650.
- Berzin, C. and León, J.R. (2005). Convergence in fractional models and applications. *Electronic Journal of Probability*, 10, 326-370.
- Besson, J-L and Wschebor, M. (1983). Sur la finitude des moments du nombre des passages par un niveau d'une fonction aléatoire, *C.R. Acad. Sci. Paris*, t. 297, Sér I, 361-363.
- Billingsley P. (1999). *Convergence of probability measures*. Second Edition. Wiley Series in Probability and Statistics.
- Bloch, A. and Polya, G. (1932). On the number of real roots of a random algebraic equation. *Proc. Cambridge Phil. Soc.*, 33, 102-114.
- Blum, L., Cucker, F., Shub, M. and Smale, S. (1998). *Complexity and real computation*. Springer-Verlag, New York.
- Borell, C. (1975). The Brunn-Minkowski inequality in Gauss space. *Invent. Math.*, 30, 207-216.
- Borell, C. (2003). The Ehrhard inequality. *C.R. Acad. Sci. Paris*, Sér. I, 337, 663-666.
- Breuer J. Major P. Central limit theorems for non-linear functionals of Gaussian fields *J. Mult. Ann.* 13, 425-444 (1983).
- Brillinger D. R., (1972). On the number of solutions of systems of random equations. *The Annals of Math. Statistics*, 43, 534-540.
- Brown, J.R. (1976). *Ergodic Theory and Topological Dynamics*. Academic Press.
- Bulinskaya, E.V. (1961). On the mean number of crossings of a level by a stationary Gaussian stochastic process, *Theory Prob. Appl.*, 6, 435-438.
- Bürgisser, P. (2007) Average Euler Characteristic of Random Real Algebraic Varieties. *Comptes Rendus Academie des Sciences, Paris, Mathématique*, Vol 345/9, pp. 507-512.
- Bürgisser, P., Cucker, F. and Lotz, M. (2006). Smoothed analysis of complex conic condition numbers. *J. Math. Pures et Appl.*, 86, 293-309.
- Cabaña, E. M., (1985). Esperanzas de integrales sobre conjuntos de de nivel aleatorios (spanish). *Actas del segundo Congreso Latinoamericano de Probabilidad y Estadística Matemática*. Caracas, 65-81.
- Cabaña E. M. (1987). Affine process a test of isotropy based in level sets. *SIAM J. App. Math.* 47 886-891.
- Cabaña, E.M. (1991). A Gaussian process with parabolic covariance *J. Appl. Prob.* 28,898-902.
- Cabaña, E.M and Wschebor, M. (1981). An estimate for the tails of the distribution of the supremum for a class of stationary multiparameter Gaussian processes. *J. Appl. Prob.*, Vol 18, 536-541.
- Cartwright, D.E. and Longuet-Higgins, M.S. (1956). The Statistical Distribution of the Maxima of a Random Function, *Proc. Royal Soc., Ser. A*, 237, (1209), 212-232.
- Caussinus H. and Ruiz, A. (1995). Metrics for finding typical structures by means of Principal component analysis. *Data science and its application*, 117-192.
- Cierco, C. (1996). *Problèmes statistiques liés à la détection et à la localisation d'un gène à effet quantitatif*. PHD dissertation. University of Toulouse, France.

- Cierco-Ayrolle, C. Croquette, A. and Delmas (2003). Computing the distribution of the maximum of Gaussian random processes. *Methodology and Computing in Applied Probability*, Vol 5, 4,427-438.
- Corrsin S. (1955). A measure of the area of a homogeneous random surface in space. *Quart. Appl. Math.* **12**, 404 - 408.
- Cramér, H. (1965). A limit theorem for the maximum values of certain stochastic processes. *Theory Prob. Appl.*, 10, 126-128.
- Cramér, H. and Leadbetter, M.R. (1965). The moments of the number of crossings of a level by a stationary normal process. *Ann. Math. Statist.*, 36, 1656-1663.
- Cramér, H. and Leadbetter, M.R. (1967). *Stationary and Related Stochastic Processes*, J. Wiley & Sons, New-York.
- Cressie, N (1980). The asymptotic distribution of the scan statistic under uniformity. *Ann. Probab*, 8, 828-840.
- Cressie, N. and Davies, R.W. (1981). The supremum distribution of another Gaussian process. *J. Appl. Prob.*, Vol 18, 131-138.
- Cucker, F., Diao, H. and Wei, Y. (2005). Smoothed analysis of some condition numbers. *Num. Linear Alg. Appl.* 13, 71-84.
- Cucker, F.; Krick, T.; Malajovich, G. and Wschebor, M. Real polynomial systems II. Randomization and condition number, *preprint*, 2008.
- Cucker, F. and Wschebor, M. (2003). On the Expected Condition Number of Linear Programming Problems, *Numerische Mathematik*, 94, 3, 419-478.
- Cuesta-Albertos, J. and Wschebor M. (2003). Some Remarks on the Condition Number of a Real Random Square Matrix, *Journal of Complexity*, Vol. 19, 4, 548-554.
- Cuesta-Albertos, J. and Wschebor, M. (2004). Condition Numbers and Extrema of Random Fields, *IV Ascona Seminar on Stochastic Analysis, Random Fields and Applications*, Progress in Probability, Birkhäuser, pp 69-82.
- Cuzick, J. (1975). Conditions for finite moments of the number of zero crossings for Gaussian processes. *Ann. Probab*, 3, 849-858.
- Cuzick, J. (1981). Boundary crossing probabilities for stationary Gaussian processes and Brownian motion, *Trans. Am. Math. Soc.*, 263, N 2, 469-492.
- Dacunha-Castelle, D. and Gassiat, E. (1997). Testing in locally conic models, and application to mixture models. *ESAIM: Probability and Statistics*, 1, 285-317.
- Dacunha-Castelle, D. and Gassiat, E. (1999). Testing the order of a model using locally conic parameterization: population mixtures and stationary ARMA processes. *Ann. Statist.*, 27(4):1178-1209.
- Darling, D. A. (1983). On the supremum of certain Gaussian processes. *Ann. Probab.* 11, 803-806.
- Davidson, K.R. and Szarek, S.J. (2001). Local operator theory, random matrices and Banach spaces. In *Handbook of the geometry of Banach spaces, Vol. I*, 317-366, North-Holland.
- Davies, R.B. (1977). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika* 64, 247-254.
- Davies, R.B. (1987). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika* 74, 33-43.
- Davies, R.B. (2002) Hypothesis testing when a nuisance parameter is present only under the alternative: Linear model case. *Biometrika* 89, 484-489.
- Davis, J. D. (1975). *Interpolation and approximation*. Dover, New York.
- Delmas, C. (2001). Distribution du maximum d'un champ aléatoire et applications statistiques. PHD Thesis, *Université P. Sabatier*, Toulouse.

- Delmas, C. (2003 a). On likelihood ratio test in Gaussian mixture models *Sankya* 65,3,513-531.
- Delmas, C. (2003 b). Projection on spherical cones, maximum of Gaussian fields and Rice's method. *Statistics and Probability Letters*, 64,3,263-270.
- DeLong, D. M. (1981). Crossing probabilities for a square root boundary by a Bessel Process. *Communication in Statistics-Theory and methods*.A10, 2197-2213.
- Dembo, A., Poonen, B., Shao, Q.-M. and Zeitouni, O. (2002), Random polynomials having few or no real zeros *Journal of the american mathematical society*, Volume 15, Number 4, Pages 857-892
- Demmel, J. (1987). On condition numbers and the distance to the nearest ill-posed problem, *Num. Math.* 51, 251-289.
- Demmel, J. (1988). The probability that a numerical analysis problem is difficult. *Math. Comp.* 50, 449-480.
- Diebolt, J. (1979). Sur la loi du maximum de certains processus stochastiques. *Pub. Inst. Stat. Univ. Paris*, XXIV, fasc.3-4, 31-67.
- Diebolt, J. (1981). Sur la loi du maximum de certain processus stochastiques sur le tore. *Ann. Inst. Henri Poincaré*, B, Vol. XVII, N2, 165-179.
- Diebolt, J. and Posse, C. (1996). On the Density of the Maximum of Smooth Gaussian Processes. *Ann. Probab.*, 24, 1104-1129.
- Dietrich, C. R. and Newsam G. N. (1997). Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix. *SIAM J. Sci. Comput.* 18, 4, 1088-1107.
- Dobric, V.; Marcus, M.B. and Weber, M. (1988). The distribution of large values of the supremum of a Gaussian process *Astérisque*, 157-8, 95-128.
- Dudley, R.M. (1967). The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *J. Funct. Analysis* 1, 290-330.
- Dudley, R.M. (1973). Sample functions of the Gaussian process *Ann. Probab.*, 1, 66-103.
- Dudley, R.M. (1989). *Real Analysis and Probability*. Chapman and Hall.
- Durbin, J. (1985), The first-passage density of a continuous Gaussian process to a general boundary, *J. Appl. Probab.*, 22, 99-122.
- Eckart, C. and Young, G. (1936). The approximation of one matrix by another of lower rank, *Psychometrika* 1, 211-218.
- Edelman, A. (1988). Eigenvalues and condition numbers of random matrices, *SIAM J. of Matrix Anal. and Appl.*, 9, 543-556.
- Edelman, A. (1989). Eigenvalues and condition numbers of random matrices, PHD Thesis, MIT, Cambridge, Ma.
- Edelman A. and Kostlan, E. (1995). How many zeros of a random polynomial are real? *Bull. Amer. Math. Soc. (N.S.)*, 32 (1), 1-37.
- Edelman, A. and Rao, N. R. (2005). *Random matrix theory*, *Acta Numerica*, 233-297.
- Edelman, A. and Sutton, B.D. (2005). Tails of condition number distributions, *SIAM J. Matrix Anal. Appl.*, Vol. 27, No. 2, 547-560.
- Ehm, W. (1981). Sample properties of multiparameter stable processes. *Zeit. für Wahr. und v.g.*, 56, 195-228.
- Ehrhard, A. (1981). Une démonstration de l'inégalité de Borell *Ann. Scient. Univ. Clermont-Ferrand*, 69, 165-184.
- Ehrhard, A. (1983). Symétrisation dans l'espace de Gauss. *Math. Scand.* 53, 281-301.
- Ehrhard, A. (1986). Éléments extrémaux pour les inégalités de Brunn-Minkowski gaussiennes, *Ann. Inst. H. Poincaré*, 22, 149-168.

- Ein-Gal, M. and Bar-Davis, I. (1975). Passages and maxima for a particular Gaussian process *Ann. Probab.*, Vol 3, 549-556.
- Erdelyi A., Magnus W. , Oberhettinger F. , Tricomi F.G. (1953). Higher Transcendental Functions, *The Bateman Project*, Vol. 1, McGraw-Hill.
- Federer, H. (1969). *Geometric measure theory*. Springer-Verlag.
- Feller, W. (1966). *An introduction to Probability Theory and its Applications Vol. II*. Wiley, New-York.
- Fernique, X. (1964). Continuité des processus Gaussiens *Comptes R. Ac. Sci. Paris*, 6058-6060.
- Fernique, X. (1971). Régularité des processus gaussiens *Invent. Math.*, 12, 304-321.
- Fernique, X.(1974). Régularité des trajectoires des fonctions aléatoires gaussiennes. *Ecole d'Eté de Probabilités de Saint Flour. Lecture Notes in Mathematics*, 480, Springer-Verlag, New-York.
- Fernique, X. (1983). Régularité de fonctions aléatoires non-gaussiennes *Lecture Notes in Math. N 976*, 1-74, Springer-Verlag.
- Fyodorov, Y. (2006). Complexity of Random Energy Landscapes, Glass Transition and Absolute Value of Spectral Determinant of Random Matrices. *Physical Review Letters*, Vol 92,(2004),240601 (4 pages). Erratum, *ibid*, Vol 93, 149901 (1 page).
- Gassiat, E. (2002). Likelihood ratio inequalities with application to various mixtures. *Ann. Inst. H. Poincaré, Probab. Statist.*, 38, 897-906.
- Geman, D. (1972). On the variance of the number of zeros of stationary Gaussian process, *Ann. Math. Statist.*, Vol 43, N 3, 977-982.
- Ghosh, J. and Sen, P. (1985). On the asymptotic performance of the log-likelihood ratio statistic for the mixture model and related results. *Proceedings of the Berkeley conference in honor of Jerzy Neyman and Jack Kiefer*. Le Cam, L.M. and Olshen, R.A. eds.
- Girko, V. L. (1996). *Random matrices*. Handbook of algebra, Vol. 1, 27–78, North-Holland.
- Goldman, M. (1971). On the first passage of the integrated Wiener process. *The Annals Math. Statist.*, 42, 6, 2150-2155.
- Gordon, Y. (1985). Some inequalities for Gaussian processes and applications *Israel J. Math.*, 50, 265-289.
- Gordon, Y. (1988). Gaussian processes and almost spherical sections of convex bodies *Ann. Probab.*, 16,180-188.
- Haldane J.B.S. (1919). The combination of linkage values, and the calculation of distances between the loci of linked factors. *J. Genet.* 8, 299-309.
- Hansen, B.E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica*, 64, 2, 413-430.
- Hasofer, A.M. (1974). The up-crossing rate of a class of stochastic processes, *Studies in Probability and Statistics*, E.J. Williams, ed., New York, NY, North Holland, 153-159.
- Hasofer, A.M. (1976). The Mean Number of Maxima Above a High Level in Gaussian Random Fields, *J. Appl. Probab.*, 13, 377-379.
- Hasofer, A.M. (1978). Up-crossings of random fields *Adv. Appl. Prob.*, 10, 14-21.
- Higham, N. (1996). *Accuracy and Stability of Numerical Algorithms*, SIAM.
- Hoeffding W. and Robbins H. (1948). The Central Limit Theorem for dependent random variables, *Duke Math- J.* 15, 773-730 .
- Ibragimov A. , Sudakov V. N and Tsirelson B. S.(1976). Norms of Gaussian sample functions. *Proceedings of the third Japan-USSR Symposium on Probability Theory. Lecture Notes in Math.* 550, 20–41 . Springer-Verlag.

- Ikeda N. and Watanabe S.(1981), Stochastic Differential Equations and Diffusion Processes. North-Holland .
- Iribarren I (1989). Asymptotic behaviour of the integral of a function on the level set of a mixing random field, *Probab. Math. Statistics*, Vol. 10, 1, 45 - 56.
- Ito, K. (1964). The expected number of zeros of continuous stationary Gaussian processes, *J. Math. Kyoto Univ.* 3-2, 207-221.
- Jacod, J. (1998). Rates of convergence to the local time of a diffusion. *Ann. Inst. H. Poincaré, Sér. B*, 34,505-544.
- Jacod, J. (2000). Non-parametric kernel estimation of the diffusion coefficient of a diffusion. *Scand. J. Statistics* 27,1,83-96.
- 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 (1985). Some random series of functions 2a. ed. Cambridge Univ. Press.
- Kahane, J-P. (1986). Une inégalité du type de Slepian et Gordon sur les processus gaussiens Israel *J. Math.*, 55, 109-110.
- Karatzas, I. and Shreve (1998). Brownian motion and stochastic calculus. Springer-Verlag , 1998.
- Kendall, M., Stuart, A. and Ord, J.K. (1983). *The advanced theory of statistics, Vol. 3. Design and analysis and time series.* 4th. edition. Ed. Macmillan.
- Kinsman, B. (1965). Wind waves: Their generation and propagation on the ocean surface, Prentice-Hall, Englewood Cliffs. Reprinted by Dover Phoenix Editions, New York, 2002.
- Kobelkov, G.S. (2005). The ruin problem for the stationary Gaussian process *Th. Probab and Applications*, 155-163.
- Kôno, N. (1980). Sample path properties of stochastic processes. *J. Math. Kyoto Univ.* 20, 295-313.
- Kostlan, E. (2002). On the expected number of real roots of a system of random polynomial equations. In *Foundations of computational mathematics (Hong Kong, 2000)*, pp 149-188. World Sci. Publishing, River Edge, NJ.
- Kratz, M.F.(2006). Level crossings and other level functionals of stationary Gaussian processes. *Probability Surveys* Vol. 3, 230-288
- Kratz, M.F. and León, J.R. (1997). Hermite polynomial expansion for non-smooth functionals of stationary Gaussian processes: Crossings and extremes *Stoch. Proc. Appl.*, 66, 237-252.
- Kratz,M. and León, J.R. (2001). Central Limit Theorems for Level Functionals of Stationary Gaussian Processes and Fields, *Journal of Theoretical Probability*, 14 639- 671.
- Kratz, M. and León, J. (2006). On the second moment of the number of crossings by a stationary Gaussian process. *Ann. Probab.* Vol.34, No.4.
- Krée, P. and Soize, C. (1983). *Mécanique Aléatoire.* Dunod.
- Lachal, A. (1991). Sur le premier instant de passage de l'intégrale du mouvement Brownien. *Ann. Inst. H. Poincaré Probab. Statist.* 27, 385-405.
- Landau, H.J. and Shepp, L.A. (1970). On the supremum of a Gaussian process. *Sankya Ser. A* 32, 369-378.
- Lander, E. and Bolstein, D. (1989). Mapping mendelian factors underlying quantitative traits using rflp linkage maps. *Genetics society of America.*
- Latala, R. (1996). A note on the Ehrhard inequality. *Studia Math.* 118, 169-174.
- Latala, R. (2002). On some inequalities for Gaussian measures. *ICM 2002*, Vol. II, 813-822.

- Leadbetter, M.R. (1966). On Crossings of Levels and Curves by a Wide Class of Stochastic Processes, *Ann. Math. Stat.*, 37, 260-267.
- Leadbetter, M.R. (1967). A Note on the Moments of the Number of Axis Crossings by a Stochastic Process, *Bull. AMS*, 73, 129-132.
- Leadbetter, M.R. (1968). On three basic results in the theory of stationary point processes *Proc. Am. Math. Soc.*, 115-117.
- Leadbetter, M.R. (1972). Point processes generated by level crossings, *Stochastic Point Processes*, PAW Lewis, ed., New York NY, Wiley, 436-467.
- Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983). *Extremes and related properties of random sequences and processes*. Springer-Verlag, New-York.
- Leadbetter, M.R. and Rootzén, H. (1988). Extremal theory for stochastic processes *Ann. Probab.*, 16, 431-478.
- Ledoux, M. (1996). Isoperimetry and Gaussian Analysis. *Ecole d'Eté de Probabilités de Saint-Flour 1994*. Lecture Notes in Math. 1648, 165-264. Springer-Verlag. New York.
- Ledoux, M. (2001). *The Concentration of Measure Phenomenon*. American Math. Soc., Providence, RI.
- Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces*, Springer-Verlag, New-York.
- León J. and Ortega J. (1989). Weak convergence of different types of variation for biparametric Gaussian processes, *Colloquia Math. Soc. J. Bolyai*, n 57, 1989, Limit theorems in Proba. and Stat. Pecs.
- León J. (2006). A note on Breuer-Major CLT for non-linear functionals of continuous time stationary Gaussian processes. Preprint
- Li, W. V. and Shao, Q.-M. (2002). A normal comparison inequality and its applications. *Probab. Theory Related Fields*, 122(4):494-508, (2002).
- Lifshits, M.A. (1987). On the distribution of the maximum of a Gaussian process *Th. Prob. Appl.*, 125-132.
- Lifshits, M.A.(1995). *Gaussian random functions* . Kluwer, The Netherlands.
- Lindgren, G. (1970). Some properties of a normal process near a local maximum *Ann. Math. Statist.*, Vol 41, 1870-1883.
- Lindgren, G. (1972). Wave-length and Amplitude in Gaussian Noise . *Adv. Appl. Prob.*, 4, 81-108.
- Lindgren, G. (1980). Extreme values and crossings for the c2-process and other functions of multidimensional Gaussian processes with reliability applications *Adv..Appl. Prob.*, Vol 12, 746-774.
- Lindgren, G. (1983). Use and structure of Slepian model processes for prediction and detection in crossing and extreme value theory, University of Lund and Lund Institute of Technology, Dept. of Math Stat. Report N 1983:4.
- Lindgren, G. (2006). Slepian models for the stochastic shape of individual Lagrange sea waves. *Adv. Appl. Prob.*, 38, 430-450.
- Lindgren, G. and Rychlik, I. (1991). Slepian models and regression approximations in crossings and extreme value theory *International Statistical Review*, 59, 2, 195-225.
- Lindgren G. and Rychlik I. (1995). Wave analysis toolbox a tutorial. Preprint University of Lund.
- Littlewood J.E. and Offord A.C.(1938). On the number of real roots of a random algebraic equation. *J. London Math. Soc.*, 13, 288-295.
- Littlewood, J.E. and Offord, A.C. (1939). On the roots of certain algebraic equation. *Proc. London Math. Soc.*, 35, 133-148.

- Longuet-Higgins, M.S. (1957). The statistical analysis of a random moving surface, *Phil. Trans. Royal Soc. London*, Ser. A, 249, 321-387.
- Longuet-Higgins, M.S. (1962a). The statistical geometry of random surfaces. Hydraulic Instability, Symp. in Appl. Math., Vol 13, 105-143.
- Longuet-Higgins, M.S. (1962b). The distribution of intervals between zeros of a stationary random function, *Phil. Trans. Royal Soc. London*, 254A, 557-599.
- Malevich, T. L. (1979). On Conditions for the Moments of the Number of Zeros of Gaussian Stationary Processes to be Finite, *Th. Prob. Appl.*, N 4, Vol. XXIV, 741-754.
- Malevich, T. L. (1984). On conditions for finiteness of the factorial moments of the number of zeros of Gaussian stationary processes *Th. Prob. Appl.*, Vol XXIX, N 3, 534-545.
- Malevich, T. L. and Volodina, L.N. (1980). Upper and lower bounds for factorial moments of the number of zeros of a Gaussian process (Russian). *Izv Akad Nauk Uz USSR*, No. 4, 26-33.
- Malmquist, S. (1954). On certain confidence contours for distribution functions. *Ann. Math. Statist.*, 25, 523-533.
- Marcus, M. (1970). A bound for the distribution of the maximum of continuous Gaussian processes *Ann. Math. Statist.*, Vol. 41, N 1, 305-309.
- Marcus, M. (1972). Upper bounds for the asymptotic maxima of continuous Gaussian processes *Ann. Math. Statist.*, Vol 43, N 2, 522-533.
- Marcus, M. (1973). Continuity of Gaussian Processes and Random Fourier Series, *The Ann. Probab.*, 1, 968-981.
- Marcus, M. (1977). Level Crossings of a Stochastic Process with Absolutely Continuous Sample Paths, *Ann. Probab.*, 5, 52-71.
- Marcus, M. (1989). Some bounds for the expected number of level crossings of symmetric harmonizable p -stable processes *Stoch. Proc. Appl.*, 33, 217-231.
- Marcus, M.B. and Shepp, L.A. (1972). Sample behaviour of Gaussian processes. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.*, 2, 423-442.
- de Maré, J. (1977). Reconstruction of a stationary Gaussian process from its sign changes *J. Appl. Prob.*, Vol 14, 38-57.
- Maruyama G. (1949). The harmonic analysis of stationary stochastic processes. *Mem. Fac. Sci. Kyusyu Univ.*, 4, 45-106.
- McKean, H.P. (1963). A winding problem for a resonant driven by a white noise. *J. Math. Kyoto Univ.*, 2, 227-235.
- McKean, H.P. (1969). *Stochastic Integrals*. Academic Press.
- Mehta, M.L. (2004). *Random Matrices*. 3d. edition. Academic Press.
- Mercadier, C. (2002). Caractère asymptotiquement Gaussien des extrêmes en modèles non-stationnaires. Rapport de DEA. Université Paul Sabatier. Toulouse.
- Mercadier, C. (2005). MAGP toolbox, <http://math.univ-lyon1.fr/~mercadier/>
- Mercadier, C. (2006). Numerical bounds for the distribution of the maximum of one- and two-dimensional processes, *Adv. in Appl. Probab.* 38, no. 1, 149-170.
- Milnor, J. W. (1965). *Topology from the differentiable viewpoint*. The University Press of Virginia, Charlottesville.
- Miroshin, R. N. (1974). Rice series in the theory of random functions. *Vestnik Leningrad Univ. Math.*, 1, 143-155.
- Miroshin, R. N. (1977). Condition for finiteness of moments of the number of zeros of stationary Gaussian processes. *Th. Prob. Appl.*, 22, 615-624.
- Miroshin, R. N. (1983). The use of Rice series. *Th. Prob. Appl.* 28, 714-726.

- Miroshin, R.N. (1985). A simple criterion for finiteness of the moments of the number of zeros of a Gaussian stationary process *Th. Prob. Appl.*, 566-569.
- Muirhead, R.J. (1982). *Aspects of multivariate statistical theory*. Wiley Series in Probability and Mathematical Statistics.
- Neveu, J. (1977). *Processus ponctuels*. Lecture Notes in Math., 598, Springer.
- Nisio, M. (1967). On the extreme values of Gaussian processes *Osaka J. Math.*, 4, 313-326.
- Nosko, V.P. (1969). Local Structure of Gaussian Random Fields in the Vicinity of High Level Light Sources. *Soviet Math. Doklady*, Vol. 10, N 6, 1481-1484.
- Nosko, V.P. (1973). On the Possibility of Using the Morse Inequalities for the Estimation of the Number of Excursions of Random Field in a Domain, *Th. Prob. Appl.*, 18, 821-822.
- Nosko, V.P. (1979a). On the definition of the number of excursions of a random field above a fixed level *Th. Prob. Appl.*, 24, N 3, 598-602.
- Nosko, V.P. (1979b). On the average number of distinct level curve points of a random field on the plane. *Th. Prob. Appl.*, 24, 181-183.
- Nosko, V.P. (1986). The local structure of a homogeneous Gaussian random field in a neighbourhood of high level points. *Th. Prob. Appl.*, 30, 767-782.
- Nosko, V.P. (1987). On the distribution of the arc length of a high-level excursion of a stationary Gaussian process *Th. Prob. Appl.*, 521-523.
- Nualart, D. and Vives, J. (1988). Continuité absolue de la loi du maximum d'un processus continu, *C.R. Acad. Sci. Paris*, t. 307, Série I, 349-354.
- Nualart, D. and Wschebor, M. (1991). Intégration par parties dans l'espace de Wiener et approximation du temps local. *Prob. Th. Rel. Fields*, 90, 83-109.
- Ochi, M. (1998). *Ocean Waves*. Cambridge University press.
- Orsingher, E. (1979). Level crossings of transformations of stationary Gaussian processes *Metron*, Vol 37, 81-100.
- Orsingher, E. (1980). Planar Gaussian motions generated by crossings events *Pub. Inst. Stat. Univ. Paris*, Vol 25, 61-86.
- Orsingher, E. (1981). Upcrossings and conditional behaviour of the lognormal process *Boll. Un. Mat. Ital.*, Vol 18-B, 1017-1034.
- Palm C. (1943). Intensitätsschwankungen in Fernsprechverkehr. *Ericson Technics*, 1-189.
- Paranjape S.R; Park, C. (1973). Distribution of the supremum of the two-parameter Yeh-Wiener process on the boundary, *J. Appl. Probab*, 10, 4, 875-880. 2.
- Perera, G. and Wschebor, M. (1998). Crossings and occupation measures for a class of semi-martingales. *Ann. Probab.*, Vol. 26, N 1, 253-266.
- Perera, G. and Wschebor, M. (2002). Inference on the Variance and Smoothing of the Paths of Diffusions. *Ann. Inst. H. Poincaré*, Vol 38, 6, 1009-1022.
- Pickands, J. III (1969 a). Upcrossing probabilities for stationary Gaussian processes *Trans. Am. Math. Soc.* Vol 145, 51-73.
- Pickands, J. III (1969b). Asymptotic properties of the maximum in a stationary Gaussian process *Trans. Am. Math. Soc.*, Vol 145, 75-86.
- Piterbarg, V.I. (1978). The Central Limit Theorem for the Number of Level Crossings for Gaussian Processes, *Th. Prob. Appl.*, 23, N 1, 178-181.
- Piterbarg, V; I. (1981). Comparison of distribution functions of maxima of Gaussian processes. *Th, Proba. Appl.*, 26, 687-705.
- Piterbarg, V; I. (1996). *Asymptotic Methods in the Theory of Gaussian Processes and Fields*. American Mathematical Society. Providence, Rhode Island.

- Piterbarg V. I., (1996b). Rice's Method for Large Excursions of Gaussian Random Fields. Technical Report No. 478, University of North Carolina. Translation of Rice's method for Gaussian random fields.
- Plackett, R. I. (1954). *A reduction formula for normal multivariate integrals*. *Biometrika* 41:351–360, (1954).
- Podgórski K., Rychlik I. and Sjö E. (2000). Statistics for velocities of Gaussian waves. *Inter. J. of Offshore and Polar Engineering* **10**, 91-98.
- Qualls, C. (1968). On a limit distribution of high level crossings of a stationary Gaussian process *Ann. Math. Statist.*, 39, 2108-2113.
- Qualls, C. (1969). On the joint distribution of crossings of high multiple levels by a stationary Gaussian process *Arkiv för Matematik*, 8, 15, 129-137.
- Qualls, C. and Watanabe, H. (1973a). Asymptotic properties of Gaussian random fields *Trans. Am. Math. Soc.*, 155-171.
- Qualls, C. and Watanabe, H. (1973b). Asymptotic properties of Gaussian process. *Ann. Math. Statistics*, 43, 580-596
- Radchenko, A.N. (1982). On geometric measures of the level set of a random field *Th. Prob. Math. Statist.*, 25, 125-135.
- Rényi, A. (1967). Remarks on the Poisson process. *Studia Sci. Math. Hungar*, 2:119-123.
- Revuz D. and Yor M. (1998). *Continuous martingales and Brownian motion*. 3d ED. Springer
- Rice, S.O. (1939). The distribution of the maxima of a random curve. *American Journal of Mathematics*, 61, 409-416
- Rice, S.O. (1944-1945). Mathematical Analysis of Random Noise. *Bell System Tech. J.*, 23, 282-332; 24, 45-156.
- Rice, S.O. (1958). Distribution of the duration of fades in radio transmission *Bell Sys. Tech. J.*, Vol 37, 620-623, 630-631.
- Rickard, J.T. (1977). The zero crossing interval statistic of the smoothed random telegraph signal *Inf. Sci.*, Vol 13, 253-268.
- Rickard, J.T. and Pawula, R.F. (1975). The zero crossing variance of the smoothed random telegraph signal *Int. J. Control*, Vol 21, 743-752.
- Rychlik, I. (1987). A note on Durbin's formula for the first-passage density, *Statistics and Probability Letters*, vol. 5, pp. 425-428.
- Rychlik, I. (1990). New bounds for the first passage, wavelength and amplitude density. *Stoch. Proc. Appl.*, 34, 313-339.
- Rychlik, I., Johannesson, P. and Leadbetter, M. R. (1997). Modelling and statistical analysis of Ocean-wave Data using transformed Gaussian processes. *Marine structure* 10, 13-47.
- Sard, A. (1942). The measure of the critical points of differentiable maps. *Bull. Amer. Math. Soc.*, 48, 883-890.
- Schoenberg, I.J. (1938). Metric spaces and completely monotone functions. *Ann. of Math. (2)*, 39(4), 811-841.
- Shepp, L. A. (1971). First passage time for a particular Gaussian process. *The Ann. of Math. Stat.*, 42, 946-951.
- Shepp, L. A. (1979). The joint density of the maximum and its location for a Wiener process with drift. *J. Appl. Prob.* 16, 423-427.
- Shepp, L. A. and Slepian, D. (1976). First-passage time for a particular stationary periodic Gaussian process. *J. Appl. Prob.*, 13, 27-38.

- Shub, M. and Smale, S. (1993). Complexity of Bezout's theorem, II. Volumes and probabilities. In *Computational algebraic geometry (Nice, 1992)*, volume 109 of *Progr. Math.*, pages 267-285. Birkhäuser Boston, Boston, MA.
- 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.* 42, 463-501.
- Slepian, D. (1963). On the zeros of Gaussian noise Time series analysis, Wiley, New York.
- Slepian, D. (1983). Estimation of the Gauss-Markov process from observation of its signs *Stoch. Proc. Appl.*, Vol 14, 249-265.
- Slud, E. (1991). Multiple Wiener-Itô integral expansions for level-crossing-count functionals *Prob. Th. Rel. Fields*, 87, 349-364.
- Smale, S. (1985). On the efficiency of algorithms of analysis. *Bull. Amer. Math. Soc. (N.S.)*. 13 (2), 87-121.
- Smale, S. (1997). Complexity theory and numerical analysis, In A. Iserles (Ed.), *Acta Numerica*, 523-551, Cambridge University Press.
- Soshnikov, A. (1998). Level spacings distribution for large random matrices: Gaussian fluctuations. *Ann. of Math. (2)*. 148, no. 2, 573-617.
- Szarek, J. (1991). *Condition numbers of random matrices*. *J. Complexity*, 7 (2): 131-149.
- Spielman, D.A. and Teng, S.H. (2002). Smoothed analysis of algorithms, *Proc. of ICM, Vol. I*, 597-606.
- Spivak, M. (1979). *A comprehensive introduction to differential geometry. Vol.V*. Publish or Perish Inc., Wilmington, Del., 2d. edition.
- Sudakov, V.N. and Tsirelson, B.S. (1974). Extremal properties of half spaces for spherically invariant measures (in Russian). *Zap. Nauchn. Sem. LOMI*, 45, 75-82.
- Sun, J. (1993). Tail Probabilities of the Maxima of Gaussian Random Fields, *Ann. Probab.*, 21, 34-71.
- Switzer, P. (1976). Geometrical Measures of the Smoothness of Random Functions, *J. Appl. Probab.*, 13, 86-95.
- Talagrand, M. (1987). Regularity of Gaussian processes. *Acta Math.*, 159, 99-149.
- Talagrand, M. (1988). Small tails for the supremum of a Gaussian process. *Ann. Inst. H. Poincaré, Sér. B*, 24, 2, 307-315.
- Talagrand, M. (1990). Sample boundedness of stochastic processes under increment conditions *Ann. Probab.*, 18, 1-49.
- Talagrand, M. (1994). The supremum of some canonical processes *Amer. J. Math.*, 116, 284-325.
- Talagrand, M. (1995). Concentration of measure and isoperimetric inequalities in product spaces, *IHES*, 81, 73-205.
- Talagrand, M. (1996). Majorising measures: the general chaining. *Ann. Probab.* 24, 1049-1103.
- Talagrand, M. (2001). Majorizing measures without measures. *Ann. Probab.* 29, No.1, 411-417.
- Talagrand, M. (2005). The generic chaining. *Upper and Lower Bounds of Stochastic Processes*. Springer-Verlag.
- Tao, T. and Vu, V. (2007). The condition number of a randomly perturbed matrix, *J. ACM*, 248-255.
- Taylor, J.E. and Adler, R. J. (2003). Euler characteristics for Gaussian fields on manifolds. *Ann. Probab.*, 31, 533-563.

- Taylor J.E., Takemura A. and Adler R.J. (2005). Validity of the expected Euler Characteristic heuristic. *Ann. Probab.*, 33, 4, 1362-1396.
- Tracy, C.A. and Widom, H. (1999). Universality of the Distribution Functions of Random Matrix Theory. In *Integrable systems from classical to quantum: Proceedings of the 38th session of the Séminaire de mathématiques supérieures. Université de Montréal*
- Trefethen, L. and Bau III, D. (1997). *Numerical Linear Algebra*, SIAM.
- Tsirelson, V.S. (1975). The Density of the Maximum of a Gaussian Process. *Th. Probab. Appl.*, 20, 817-856.
- Turing, A. (1948). Rounding-off errors in matrix processes, *Quart. J. Mech. Appl. Math.*, 1, 287-308.
- Volkonskii V.A. and Rozanov Y.A (1959,1961). Somme limit theorems for random functions I and II . *Teoria Veroyatnostei i ee Primeneniya* 4,186-207 and 6,202-215.
- Von Neumann, J.and Goldstine, H. (1947). Numerical inverting matrices of high order, *Bull. AMS*, 53, 1021-1099.
- WAFO-group (2000). WAFO - A Matlab Toolbox for Analysis of Random Waves and Loads - A Tutorial. *Math. Stat., Center for Math. Sci., Lund Univ., Lund, Sweden.* ISBN XXXX, URL <http://www.maths.lth.se/matstat/wafo>.
- Weber, M. (1985). Sur la densité du maximum d'un processus gaussien. *J. Math. Kyoto Univ.*, 25, 515-521.
- Wilkinson, J.H. (1963). *Rounding Errors in Algebraic Processes*. Prentice-Hall.
- Wigner, E. (1958). On the distribution of the roots of certain symmetric matrices, *Ann. of Math.*, 67, 325-327.
- Wigner, E. (1967). Random matrices in physics, *SIAM Review*, 9, 1-23.
- Wong, E. (1966). Some results concerning the zero-crossings of Gaussian noise *Siam J. Appl. Math.*, 14, 1246-1254.
- Wschebor, M. (1982). Formule de Rice en dimension d. *Zeit. Wahr. verw. geb.*, Vol. 60, 393-401.
- Wschebor, M. (1983). On crossings of Gaussian fields. *Stoch. Proc. Appl.*, Vol. 14, N 2, 147-155.
- Wschebor, M. (1985). *Surfaces aléatoires. Mesure géométrique des ensembles de niveau.* Lecture Notes in Math., 1147, Springer-Verlag.
- Wschebor, M. (1986). Random lengths for high levels *Acta Cient. Venezolana*, Vol. 37, N 6, 633-635.
- Wschebor, M. (1992). Sur les accroissements du Processus de Wiener. *C.R. Acad. Sci., Paris*, T 315, Série I, pp. 1293-1296.
- Wschebor, M. (2000). Sur la loi du sup de certains processus Gaussiens non bornés. *C.R. Acad. Sci. Paris*, t. 331, Série I, 823-826.
- Wschebor, M.(2004). Smoothed analysis of K(A), *Journal of Complexity*, Vol. 20, 1, 97-107.
- Wschebor, M. (2005). On the Kostlan-Shub-Smale model for random polynomial systems. Variance of the number of roots. *Journal of Complexity*, 21, 773-789.
- Wschebor, M. (2006). Smoothing and the approximation of the occupation measure of random processes. *Ann. Fac. Sci. Toulouse Math.*, (6), 15, no. 1, 125-156.
- Wschebor, M. (2008). Random systems of equations, a review of some recent results, *In and out of equilibrium 2, Progress in Probability*, Vol. 60, Birkhäuser, 559-574.
- Ylvisaker, D. (1965). The expected number of zeros of a stationary Gaussian process *Ann. Math. Statist.*, 1043-1046.
- Ylvisaker, D. (1966). On a theorem of Cramér and Leadbetter *Ann. Math. Statist.*, 37, 682-685.

Ylvisaker, D. (1968). A Note on the Absence of Tangencies in Gaussian Sample Paths. *The Ann. of Math. Stat.*, 39, 261-262.

Index

- Airy relation, 209
- Area formula, 119

- Bézout number, 223
- Back-cross, 82
- Banach's formula, 68, 201
- Berman's condition, 189
- Bulinskaya's lemma, 132

- càd-làg, 84
- Camilla hurricane, 218
- Canonical distance, 30
- Central limit theorem for crossings, 203
- Co-area formula, 132, 268
- Comparison of means, 83
- Computation of moments, 106
- Condition number, 249, 251
- Contiguity, 94
- Continuity of the paths, 22
- convex cone, 160
- Crest, 213
- Critical points, 126
- Crossings, 55
 - computation of moments, 78
 - down-crossing, 55
 - finiteness of moments, 64
 - Hermite expansion, 200
 - up-crossing, 55
- Cylinder set, 16

- Davies bound, 74, 101, 113
- Density of the maximum, 137, 146
 - asymptotics, 154, 166
 - bound, 160
 - constant variance process, 170
 - derivative, 146
 - implicit formula, 138
 - non constant variance process, 166
 - one parameter, 140
 - random fields, 148
 - stationary and isotropic Gaussian fields, 150, 161
- Differentiability in quadratic mean, 28
- Differentiable paths, 27
- Direct method, 154
- Directional power spectrum, 210
- Discretization of time, 192
- Distribution
 - of the maximum (bound), 74
 - asymptotic, 94, 97
 - Gaussian, 17
 - Gumbel, 190
 - non degenerated , 68
 - normal , 17
 - Palm, 211
 - Poisson, 86
- Double sum method, 153
- Dual cone, 160

- Entropy, 30, 51
- EPC method, 154
- Euler-Poincaré characteristic, 153, 160, 188, 264
- Excursion set, 160
- Exploding paths, 157
- Extended outwards derivative, 160

- Factorial moments, 62
- Finiteness of moments, 64, 253
- Formula
 - Banach, 182
 - Durbin, 184
 - Rychlik, 181
- Fractional Brownian motion, 36

- Gaussian matrix, 166, 179, 251
- Gaussian Orthogonal Ensemble, 163
- Gaussian regression, 18
- Geman's condition, 68, 77, 189, 190
- Gene detection, 82
- Genetic marker, 83
- GOE, 163, 178, 226

- Hölder Conditions, 27
- Haldane distance, 82
- Hermite expansion of crossings, 200
- Hermite polynomials, 156, 200, 218
 - expansion, 165
 - modified, 156
- Hermite rank, 196, 197

- Inequalities for the tails, 45, 48, 50, 74, 77
- Inequality
 - Borell, 51
 - Borell, Sudakov, Tsirelson , 46
 - Dudley, 51
 - Ehrhard, 42
 - isoperimetric, 45
 - Li and Shao, 39
 - Slepian, 39
- Invariant systems, 229
- Ito integrals, 24
- Ito's formula, 46

- Kac's formula, 55

- Karhunen-Loève representation, 20
- KSS model, 224, 244
- Law
 - zero one, 20
- Law of iterated logarithm, 184
- Level curve, 212
- Level set, 131
- Linear model, 82
- Lipschitz condition, 120
- Local asymptotic alternative, 94
- Local extrema, 15, 31, 57, 211
- Local time, 6, 70, 71
- MAGP toolbox, 183, 185, 188
- Manifold, 138, 268
 - embedded, 129
 - Riemannian, 128
- Matrix
 - condition number, 249, 251
 - Frobenius norm, 250
 - ill-conditioned, 249
 - non-Gaussian, 251
 - random, 250
 - well-conditioned, 249
- Maxima on manifolds , 126
- Maximum of the absolute value, 81
- Mehler's formula, 197
- Mixtures of Distributions , 92
- Monte-Carlo experiment, 265
- Monte-Carlo method, 90, 101, 111
- Morgan, 82
- Multivariate statistics, 250
- Narrow spectrum, 71
- Non-centered systems, 235
- Non-linear functionals, 199
- Non-parametric transformation, 218
- Normal comparison lemma, 39, 54
- Numerical analysis, 251
- ordinary moments, 62
- Point processes, 190
- Polish space, 84
- Polyhedron, 161
- Polynomial in the scalar product, 233
- Polynomial interpolation, 102
- Polynomials, 221
- Polynomials systems, 222
- Process
 - χ^2 , 70
 - diffusion, 118
 - ergodic, 195, 211
 - m-dependent, 196
 - non smooth, 184
 - Ornstein-Uhlenbeck, 41, 86, 89, 100, 112, 185
 - periodic, 100
 - Poisson, 82, 87, 189, 194
 - sine-cosine, 99
 - Slepian, 41
 - stationary, 19, 23, 62, 64, 77, 79, 105
 - stationary and isotropic, 243
 - unbounded, 157
 - Wiener, 7, 24, 27, 51, 70, 118, 184, 198
 - with Gaussian covariance, 41
- process
 - two-parameter, 185
- QTL, 82
- Quadratic perturbation, 232
- Random fields, 29, 194, 206
- Random matrices, 250
- Rayleigh density, 71
- Record method, 181
- Record times
 - one dimension, 181
 - two dimensions, 185
- Rice formula, 55, 257
 - χ^2 process, 70
 - random field, 121
 - counterexample, 67
 - factorial moments, 62
 - first moments, 62
 - Gaussian, 57, 63
 - geometric measure, 131
 - higher moments, 121
 - integral on the level set, 131
 - local extrema, 67, 70
 - non-Gaussian, 60, 129
 - upper-bound, 69
 - variations, 66
 - weighted , 122
- Rice series, 101, 183
 - numerical aspects, 111
- Round-off errors, 250
- Second fundamental form, 129
- Separability, 26
- Shub-Smale model, 222
- Singular values, 250
- Skorohod
 - embedding, 85
 - space, 190
 - topology, 84
- Smoothed analysis, 235, 254
- Smoothing, 89
- Spectral measure, 19
- Spectral moment, 20
- Speed of crossings, 214
- Standard normal distribution, 18
- Stationarity, 195
- Stationary and isotropic Gaussian field, 150, 154, 161, 243
- Statistical test, 73, 82
- Stirling formula, 263
- Student distribution, 262
- Systems of random equations, 221
- Tail of the distribution of the maximum
 - asymptotics, 182
 - two dimension, 187
- Tangencies, 31
- Taylor expansion of Rice functions, 106
- Test
 - likelihood, 83, 93
 - moment, 93
- Theorem
 - Birkhoff-Khintchine, 195
 - Bochner, 17, 134
 - central limit, 196, 202

ergodic, 196
Kolmogorov's Extension, 16
Maruyama, 195
Prohorov, 84
Rolle, 65
Shub-Smale, 221
Sudakov-Fernique, 42
Tsirelson, 137
Volkonskii and Rozanov, 189
Ylvisaker, 32, 188
Tightness, 84
Transformed Gaussian models, 218
Tube method, 153

Unit speed, 183

Variance of the number of roots, 224
Velocity of crests, 216
Velocity of level curves, 215
Version, 17

Wave-height, 211
Wave-length, 211
Waves
 Euler equations, 209
 Lagrange model, 219
Weak convergence, 83–85
Wiener chaos, 189