

Chapter 2

Gaussian Processes

The theory of Gaussian processes and fields is rich and varied, and many excellent books have been written on the subject, among them Bogachev [15], Dudley [33], Fernique [39], Hida and Hitsuda [49], Janson [52], Ledoux and Talagrand [60], Lifshits [61] and Piterbarg [68], not to mention *RFG* and another old favourite of ours, another set of lecture notes, [3]. In particular, a new book [11] by Jean-Marc Azaïs and Mario Wschebor has recently appeared that has a lot of material similar, but generally complementary, to what interests us.

We have no intention to go into any detail in the current notes, however, and so will take a quick route towards defining Gaussian processes on general parameter spaces that will get us where we need to go with the minimum of fuss. All you will need to know to follow this is some rather basic graduate level probability, and the definition of the multivariate Gaussian distribution.¹ We shall start, however, with a very simple example which requires nothing beyond undergraduate probability and some innovative calculus, but which is already extremely instructive.

¹ Recall that a \mathbb{R}^d valued random variable $X = (X_1, \dots, X_d)$ is said to be *multivariate Gaussian* if, for every $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{R}^d$, the real valued variable $\langle \alpha, X \rangle = \sum_{i=1}^d \alpha_i X_i$ is univariate Gaussian. In this case there exists a mean vector $m \in \mathbb{R}^d$ with $m_j = \mathbb{E}\{X_j\}$ and a non-negative definite $d \times d$ covariance matrix C , with elements $c_{ij} = \mathbb{E}\{(X_i - m_i)(X_j - m_j)\}$, such that the probability density of X is given by

$$\phi_d(x) = \frac{1}{(2\pi)^{d/2} |C|^{1/2}} e^{-(x-m)C^{-1}(x-m)'/2}, \quad x \in \mathbb{R}^d, \quad (2.0.1)$$

where $|C| = \det(C)$. We write this as $X \sim N(m, C)$, or $X \sim N_d(m, C)$ if we need to emphasise the dimension, and also adopt the standard but heavily overworked symbol ϕ to denote the density ϕ_1 of a $N(0, 1)$ random variable.

2.1 The Cosine Process

Perhaps the grandfather of all smooth stochastic processes is the *cosine random process* on \mathbb{R} . It is defined as

$$f(t) \triangleq \xi \cos \lambda t + \xi' \sin \lambda t, \quad (2.1.1)$$

where ξ and ξ' are uncorrelated, equidistributed, random variables and λ is a positive constant.

It is elementary trigonometry to see that the cosine process can also be written as

$$f(t) = R \cos(\lambda(t - \theta)), \quad (2.1.2)$$

where $R^2 = \xi^2 + (\xi')^2 \geq 0$ and $\theta = \arctan(\xi/\xi') \in (-\pi, \pi]$, from whence the name ‘cosine process’. Assuming, for convenience, that $\mathbb{E}\{\xi\} = 0$, we have that the covariance function of f is given by

$$\begin{aligned} C(s, t) &= \mathbb{E}\{f(s)f(t)\} \\ &= \mathbb{E}\{(\xi \cos \lambda s + \xi' \sin \lambda s)(\xi \cos \lambda t + \xi' \sin \lambda t)\} \\ &= \mathbb{E}\{\xi^2\} \cos(\lambda(t - s)), \end{aligned}$$

on using the fact that ξ and ξ' are uncorrelated and equidistributed. Consequently, regardless of the distribution of ξ , the cosine process is stationary. (See Sect. 2.6 below for definitions of stationarity and isotropy.)

One of the nice aspects of the cosine process is that many things that are either difficult or impossible to compute for more general processes can be computed exactly, and from first principles, once some assumptions are made on the distribution of ξ . We shall therefore now assume that ξ and ξ' are independent, Gaussian variables, with zero mean and common variance σ^2 . As an example of what can be computed, consider, for $u > 0$, the exceedence probability

$$\mathbb{P} \left\{ \sup_{0 \leq t \leq T} f(t) \geq u \right\}, \quad (2.1.3)$$

which we met in the Introduction.

Under the Gaussian assumption, R^2 has an exponential distribution with mean $1/(2\sigma^2)$, θ has a uniform distribution on $(-\pi, \pi]$, and R and θ are independent. We can use this information to compute some exceedence probabilities directly, and start by defining the *number of upcrossings* by f of the level u in time $[0, T]$,

$$N_u = N_u(f, T) = \#\{t \in [0, T] : f(t) = u \text{ and } df(t)/d(t) > 0\}.$$

It is trivial to see that the exceedence probability that we are after can now be written as

$$\begin{aligned}\mathbb{P}\left\{\sup_{0 \leq t \leq T} f(t) \geq u\right\} &= \mathbb{P}\{f(0) \geq u\} + \mathbb{P}\{f(0) < u, N_u \geq 1\} \\ &= \Psi\left(\frac{u}{\sigma}\right) + \mathbb{P}\{f(0) < u, N_u \geq 1\}.\end{aligned}\quad (2.1.4)$$

where

$$\Psi(x) \triangleq 1 - \Phi(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du. \quad (2.1.5)$$

is the tail probability function for a standard Gaussian variable.

We now restrict attention to the case $T \leq \pi/\lambda$, in which case, since f has period $2\pi/\lambda$, the event $\{f(0) \geq u, N_u \geq 1\}$ is empty, implying that

$$\mathbb{P}\{f(0) < u, N_u \geq 1\} = \mathbb{P}\{N_u \geq 1\}.$$

Again using the fact that $T < \pi/\lambda$, note that N_u is either 0 or 1. In order that it be 1, two independent events must occur. Firstly, we must have $R > u$, with probability $e^{-u^2/2\sigma^2}$. Secondly (draw a picture) θ must fall in an interval of length λT , so that the final result is

$$\mathbb{P}\left\{\sup_{0 \leq t \leq T} f(t) \geq u\right\} = \Psi\left(\frac{u}{\sigma}\right) + \frac{\lambda T}{2\pi\sigma} e^{-u^2/2\sigma^2}, \quad (2.1.6)$$

and the probability density of the supremum is given by

$$\frac{1}{\sigma} \phi\left(\frac{u}{\sigma}\right) + \frac{\lambda T u}{2\pi\sigma^2} e^{-u^2/2\sigma^2}. \quad (2.1.7)$$

This computation was so simple, that one is tempted to believe that it must be easy to extend to many other processes. In fact, this is not the case, and the cosine process and field, which we shall meet in a moment, are the *only* differentiable, stationary, Gaussian processes for which the exceedence probabilities are explicitly known.

However, before we leave it, we can use the cosine process to motivate a more general approach. Note first that since, as noted above, N_u is either 0 or 1 when $T < \pi/\lambda$, we can rewrite (2.1.4) as

$$\mathbb{P}\left\{\sup_{0 \leq t \leq T} f(t) \geq u\right\} = \Psi\left(\frac{u}{\sigma}\right) + \mathbb{E}\{N_u\}. \quad (2.1.8)$$

Thus, rather than arguing as above, we could concentrate on finding an expression for the mean number of upcrossings.

More importantly, note that for any T , and, indeed, for *any differentiable random process*, the above argument always gives

$$\mathbb{P} \left\{ \sup_{0 \leq t \leq T} f(t) \geq u \right\} \leq \mathbb{P} \{f(0) \geq u\} + \mathbb{E}\{N_u\}. \quad (2.1.9)$$

Thus there would seem to be a close relationship between exceedence probabilities and level crossing rates, that actually becomes exact for the cosine process over certain intervals. In fact, since, for a one dimensional set, its Euler characteristic is given by the number of its connected components, the expectation in the right hand sides of both (2.1.8) and (2.1.9) could be written as $\mathbb{E}\{\varphi(A_u(f, T))\}$, where φ is the Euler characteristic.

2.2 The Cosine Field

The cosine field is a straightforward extension to \mathbb{R}^N of the cosine process, and has the representation

$$f(t) = f(t_1, \dots, t_N) \triangleq \frac{1}{\sqrt{N}} \sum_{k=1}^N f_k(\lambda_k t_k), \quad (2.2.1)$$

where each f_k is the process on \mathbb{R} given by

$$f_k(t) = \xi_k \cos t + \xi'_k \sin t.$$

The λ_k are fixed, and the ξ_k and ξ'_k are taken to be identically distributed and uncorrelated.

Again, it is a simple exercise to check that the cosine field is both stationary and isotropic but it is somewhat harder to compute its exceedence probabilities. To see what can be done, we restrict attention to the cosine process on a rectangle of the form $T = \prod_{k=1}^N [0, T_k]$. Then, given the structure of the cosine field as a sum, it is immediate that

$$\sup_{t \in T} f(t) = \frac{1}{\sqrt{N}} \sum_{k=1}^N \sup_{0 \leq t_k \leq T_k} f_k(t).$$

If we assume that the ξ_k and ξ'_k are all independent $N(0, \sigma^2)$, then the suprema of the individual f_k are also independent. Further assuming that each $T_k \in (0, \pi/\lambda_k]$, (2.1.6) and (2.1.7) give their individual distributions. The distribution of the supremum of the cosine field is then the convolution of these. The computations involved in actually doing the convolution are not easy, but Piterbarg [68] showed that, if $p_N(u)$ is the density function of the

supremum, ϕ the standard Gaussian density and $\phi^{(k)}$ its k -derivative, then there are simple constants, C_{nk} , depending only on n and k , such that

$$p_N(u) = \phi\left(\frac{u}{\sigma}\right) + \sum_{k=1}^N (-1)^k C_{nk} \phi^{(k)}\left(\frac{u}{\sigma}\right) \sum_{j_1 \dots j_k} \prod_{i=1}^k \frac{\lambda_{j_i} T_{j_i}}{\sigma}. \quad (2.2.2)$$

The inner sum here is over the $\binom{N}{k}$ subsets of size k of $\{1, \dots, N\}$.

Now assume that all the λ_j are identical. Then, appropriately rewritten, this result will recall the Gaussian kinematic formula. Setting $\sigma^2 = 1$ for convenience, and recalling the definition of the Lipschitz-Killing curvatures of rectangles at (1.2.12), we can write

$$p_N(u) = \sum_{k=0}^N (-1)^k C'_{nk} \phi^{(k)}(u) \lambda^k \mathcal{L}_j(T). \quad (2.2.3)$$

Going a little further, integrating over u , and applying some non-trivial asymptotics (cf. Sect. 2.5 of [68]) one finds that

$$\mathbb{P}\left\{\sup_{t \in T} f(t) \geq u\right\} = e^{-u^2/2} \sum_{k=0}^N C''_{nk} H_{k-1}(u) \lambda^k \mathcal{L}_k(T) + o\left(e^{-(1+\eta)u^2/2}\right), \quad (2.2.4)$$

for some $\eta > 0$, The Hermite polynomials H_n are defined by

$$H_n(x) = n! \sum_{j=0}^{\lfloor n/2 \rfloor} \frac{(-1)^j x^{n-2j}}{j! (n-2j)! 2^j}, \quad n \geq 0, \quad x \in \mathbb{R}. \quad (2.2.5)$$

where $\lfloor a \rfloor$ is the largest integer less than or equal to a and

$$H_{-1}(x) \triangleq \sqrt{2\pi} e^{x^2/2} \Psi(x), \quad (2.2.6)$$

where Ψ is the tail probability (2.1.5) of a standard normal.

The easily checked fact that

$$\frac{d^j}{dx^j} e^{-x^2/2} = (-1)^j H_j(x) e^{-x^2/2}, \quad (2.2.7)$$

along with (2.2.3) explains why Hermite polynomials arise in the exceedence probabilities of the cosine field.

In fact, it turns out Hermite polynomials will arise in expressions for exceedence probabilities of all real valued C^2 Gaussian fields. Furthermore, along with the factor $e^{-u^2/2}$, they can be written in terms of Gaussian Minkowski functionals, a fact that we shall prove in Sect. 3.5. Thus the main term in the

right hand side of (2.2.4) is now very reminiscent of the right hand side of the Gaussian kinematic formula.

However, even as a stand-alone result, it is already fascinating in that it links exceedence probabilities to the geometry of the parameter space.

2.3 Constructing Gaussian Processes

Since the construction of cosine processes and fields as a sum of deterministic functions with random amplitudes worked so well, we now try something similar in general. Thus, with M a potential parameter space, choose a finite or infinite set of functions $\varphi_1, \varphi_2, \dots, \varphi_j : M \rightarrow \mathbb{R}$ satisfying only

$$\sum_j \varphi_j^2(t) < \infty, \quad \text{for all } t \in M. \quad (2.3.1)$$

Let ξ_1, ξ_2, \dots be a sequence of independent, mean zero, variance 1, Gaussian random variables, and define the random field $f : M \rightarrow \mathbb{R}$ by

$$f(t) = \sum_j \xi_j \varphi_j(t). \quad (2.3.2)$$

That the sum converges in L^2 , for each fixed $t \in M$, is a consequence of (2.3.1). How f behaves, as a function of t , is another issue, that we shall turn to later. Clearly, though, the smoother the φ_j are, the better behaved f will be.

The mean of f is zero, and its covariance function is given by

$$C(s, t) = \mathbb{E}\{f(s)f(t)\} = \sum_j \varphi_j(s)\varphi_j(t). \quad (2.3.3)$$

So we have seen how to go from a sum like (2.3.2) to a covariance function. Usually, however, Gaussian processes are defined by their covariance functions, rather than vice versa, so let's make a couple of calculations and then try to work backwards. Firstly, define a class of functions

$$S = \left\{ u : M \rightarrow \mathbb{R} : u(\cdot) = \sum_{i=1}^n a_i C(s_i, \cdot), \text{ } a_i \text{ real, } s_i \in M, n \geq 1 \right\}. \quad (2.3.4)$$

Define an inner product on S by

$$(u, v)_H = \left(\sum_{i=1}^n a_i C(s_i, \cdot), \sum_{j=1}^m b_j C(t_j, \cdot) \right)_H$$

$$= \sum_{i=1}^n \sum_{j=1}^m a_i b_j C(s_i, t_j). \quad (2.3.5)$$

It is easy to check that if $u \in S$, then the following unusual property holds:

$$(u(\cdot), C(t, \cdot))_H = u(t). \quad (2.3.6)$$

This is known as the *reproducing kernel* property. The completion of S under this above inner product is known as the *reproducing kernel Hilbert space* (RKHS) of f , and all its elements also satisfy the reproducing property.

What is most interesting in this construction is that it also works in the other direction. That is, given a positive definite function C on a space M , one can define the completion of the space S of (2.3.4) under the inner product of (2.3.5), find an orthonormal basis $\{\varphi_k\}$ for $H(C)$ and define the Gaussian process (2.3.2). This will have C as its covariance function. The RKHS is now associated with C rather than f , and is denoted by $H(C)$.

For further details see *RFG* (or virtually any of the other texts mentioned at the beginning of this chapter) where you will also find a proof of the following, harder and much deeper, result, which holds under the implicit assumption, assumed throughout these notes, that we are dealing only with separable random processes.²

Theorem 2.3.1. *Suppose that C is a bounded, positive definite function, continuous on $M \times M$, and that*

$$\sup_{s, t \in M} |C(s, s) + C(t, t) - 2C(s, t)| < \infty. \quad (2.3.7)$$

Let f be defined from C as above. Then f is a.s. continuous, if, and only if, the sum (2.3.2) converges uniformly on M , with probability one.

For the French among you, here is an (almost familiar) example. The *Brownian sheet* is the zero mean, Gaussian, random field on the positive orthant $[0, \infty)^N$ with covariance function

$$\mathbb{E}\{W(s)W(t)\} = (s_1 \wedge t_1) \times \cdots \times (s_N \wedge t_N). \quad (2.3.8)$$

Replacing each j in the above sums by a multi-index $j = (j_1, \dots, j_N)$, it is then not too hard to check that the φ_j for W are given, for W restricted

² Recall that a real valued random process is called separable if there exists a countable dense subset D of M and a fixed event N with $\mathbb{P}\{N\} = 0$ such that, for any closed $B \subset \mathbb{R}$ and open $I \subset T$,

$$\{\omega : f(t, \omega) \in B, \forall t \in I\} \Delta \{\omega : f(t, \omega) \in B, \forall t \in I \cap D\} \subset N,$$

where Δ denotes the usual symmetric difference operator.

to $[0, 1]^N$, by

$$\varphi_j(t) = 2^{N/2} \prod_{i=1}^N \frac{2}{(2j_i + 1)\pi} \sin \left(\frac{1}{2} (2j_i + 1) \pi t_i \right).$$

When $N = 1$, W is the completely familiar Brownian motion. The corresponding expansion is due to Lévy, and the corresponding RKHS is known as Cameron–Martin space.

The message of this section should, by now, be clear. When dealing with continuous Gaussian processes, we lose no generality whatsoever by treating them as sums of deterministic functions with independent Gaussian coefficients. This will be important throughout these notes.

2.4 The Canonical Process on $S(\mathbb{R}^l)$

There is a school of thought that takes the basic ideas of the previous section even further. Note that, for any $t \in M$, the sequence $\tilde{\varphi}(t) = \{\varphi_1(t), \varphi_2(t), \dots\}$ belongs to ℓ^2 . (cf. (2.3.1).) Consider the image of M in ℓ^2 under the mapping $t \rightarrow x = \tilde{\varphi}(t)$, denote it by B , and define a new Gaussian process \tilde{f} by setting

$$\tilde{f}(x) = f(\tilde{\varphi}^{-1}(x)), \quad (2.4.1)$$

assuming always that φ is one to one.³ Note that

$$\begin{aligned} \mathbb{E} \{ \tilde{f}(x) \tilde{f}(y) \} &= \mathbb{E} \{ f(\tilde{\varphi}^{-1}(x)) f(\tilde{\varphi}^{-1}(y)) \} \\ &= \sum_j \varphi_j(\tilde{\varphi}^{-1}(x)) \varphi_j(\tilde{\varphi}^{-1}(y)) \\ &= \sum_j x_j y_j \\ &= \langle x, y \rangle_{\ell^2}. \end{aligned} \quad (2.4.2)$$

³ This is actually a perfectly reasonable assumption. If there are two different points $s, t \in M$ mapping to the same point in $S(\mathbb{R}^l)$, then we must have

$$\mathbb{E} \{ [f(t) - f(s)]^2 \} = \sum_1^\ell [\varphi_j(t) - \varphi_j(s)]^2 = 0,$$

which implies that $f(t)$ and $f(s)$ are, almost surely, identical, and so one of the points s, t can be dropped from the parameter set.

In other words, there is really only *one* Gaussian process. It is defined on a subset of ℓ^2 and its covariance function is the natural inner product on ℓ^2 . It is known as the *isonormal process*, and all of its properties must be properties only of the parameter set B , and so accessible via the techniques of Banach spaces.

While we shall not exactly adopt this approach, and, to some extent, it would fail us if we did, it will be particularly helpful in certain special cases.

In particular, suppose that f has constant variance, which for notational simplicity we take to be one, and, somewhat more restrictively, that the expansion (2.3.2) is finite. Consequently,

$$f(t) = \sum_{j=1}^l \xi_j \varphi_j(t), \quad (2.4.3)$$

for some $1 \leq l < \infty$ and

$$\sum_{j=1}^k \varphi_j^2(t) = \mathbb{E} \{f^2(t)\} = 1. \quad (2.4.4)$$

Thus, the set $B = \tilde{\varphi}(M)$ of the previous section is now embedded in $S(\mathbb{R}^l)$, the unit sphere of \mathbb{R}^l , and the random field defined on it can be easily extended to the entire sphere. The corresponding field is known as *canonical (isotropic) process on $S(\mathbb{R}^l)$* . It has covariance $C(s, t) = \langle s, t \rangle$, and can be realised as

$$\tilde{f}(t) = \sum_{j=1}^l t_j \xi_j. \quad (2.4.5)$$

The isotropy comes from the fact that $C(s, t)$ is function of only the (geodesic) distance between s and t . (cf. Sect. 2.6 for a definition and discussion of isotropy.)

The diagram of Fig. 1.1.1 can now be modified somewhat. In fact, if we take d independent copies of f and \tilde{f} so that now $f = (f_1, \dots, f_d)$ and $\tilde{f} = (\tilde{f}_1, \dots, \tilde{f}_d)$, we can write

$$f(t) = \tilde{f}(\tilde{\varphi}(t)) = (\tilde{f} \circ \tilde{\varphi})(t).$$

The picture is now as in Fig. 2.4.1, where we have neglected the final mapping F in Fig. 1.1.1.

It turns out that for many purposes it suffices to work with the second half of the figure, from $\tilde{\varphi} \rightarrow \mathbb{R}^d$. In the following two subsections we shall see two examples of this.

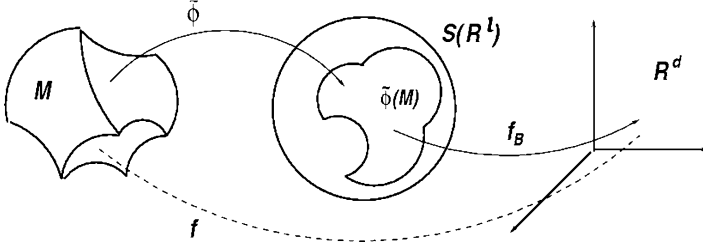


Fig. 2.4.1 The new setting with the canonical process on $S(\mathbb{R}^l)$ intervening

2.4.1 The Canonical Processes and Exceedence Probabilities

We are now going to look at the exceedence probabilities and continue the process of connecting them to geometry. The underlying technique is known as the *tube method* and has its roots in a pair of papers by Hotelling [50] and Weyl [86] in 1939. In the setting in which we shall apply it, it was developed primarily in [53, 58, 77].

Retaining the notation of the previous section, it is trivial that

$$\sup_{t \in M} f(t) \equiv \sup_{x \in \tilde{\phi}(M)} \tilde{f}(x), \quad (2.4.6)$$

so that in computing exceedence probabilities for unit variance, finite expansion Gaussian fields, we can concentrate first on treating only the canonical process over subsets of $S(\mathbb{R}^l)$. Thus, for the moment, let f be the canonical process on $S(\mathbb{R}^l)$, and let $M \in S(\mathbb{R}^l)$ be a nice set. Adopting the representation (2.4.5), we write $f(t)$ as $\langle \xi, t \rangle$, for $\xi \sim N(0, I_{\ell \times \ell})$ and $t \in S(\mathbb{R}^l)$.

Then we can argue as follows, writing $\mathbb{P}_{|\xi|}$ for the distribution of $|\xi|$:

$$\begin{aligned} \mathbb{P} \left\{ \sup_{t \in M} f_t \geq u \right\} &= \int_0^\infty \mathbb{P} \left\{ \sup_{t \in M} f_t \geq u \mid |\xi| = r \right\} \mathbb{P}_{|\xi|}(dr) \\ &= \int_0^\infty \mathbb{P} \left\{ \sup_{t \in M} \langle \xi, t \rangle \geq u \mid |\xi| = r \right\} \mathbb{P}_{|\xi|}(dr) \\ &= \int_u^\infty \mathbb{P} \left\{ \sup_{t \in M} \langle \xi, t \rangle \geq u \mid |\xi| = r \right\} \mathbb{P}_{|\xi|}(dr) \\ &= \int_u^\infty \mathbb{P} \left\{ \sup_{t \in M} \langle \xi/r, t \rangle \geq u/r \mid |\xi| = r \right\} \mathbb{P}_{|\xi|}(dr). \end{aligned} \quad (2.4.7)$$

Consider the integrand here. Since ξ is multivariate Gaussian, it is standard fare that the vector $\xi/|\xi|$ is uniformly distributed on $S(\mathbb{R}^l)$, independently of $|\xi|$, which is distributed as the square root of a χ_l^2 random variable. If

we now write η_l to denote the uniform measure over $S(\mathbb{R}^l)$, we can rewrite the integrand as a simple volume computation, once we take a moment to consider tubes on spheres.

Our definition (1.2.6) of tubes extends from the simple Euclidean setting to subsets of spheres by adopting the standard geodesic metric on $S(\mathbb{R}^l)$ given by

$$\tau(s, t) = \cos^{-1}(\langle s, t \rangle).$$

Thus the tube of radius ρ around a closed set $M \in S(\mathbb{R}^l)$ is given by

$$\begin{aligned} \text{Tube}(M, \rho) &= \{t \in S(\mathbb{R}^l) : \tau(t, M) \leq \rho\} \\ &= \{t \in S(\mathbb{R}^l) : \exists s \in M \text{ such that } \langle s, t \rangle \geq \cos(\rho)\} \\ &= \left\{t \in S(\mathbb{R}^l) : \sup_{s \in M} \langle s, t \rangle \geq \cos(\rho)\right\}. \end{aligned} \quad (2.4.8)$$

With this behind us, we can now continue the development of (2.4.7) to obtain

$$\mathbb{P} \left\{ \sup_{t \in M} f_t \geq u \right\} = \int_u^\infty \eta_l(\text{Tube}(M, \cos^{-1}(u/r))) \mathbb{P}_{|\xi|}(dr) \quad (2.4.9)$$

Thus, the exceedence probability that we seek is weighted average of the volume of tubes around M of varying radii, and if we could compute

$$\eta_l(\text{Tube}(M, \rho))$$

for all $\rho \leq 1$ we would, basically, be done, since the averaging, over the square root of a χ_l^2 random variable is, in principle, straightforward.

This approach – almost – works.

Firstly, not surprisingly, there are analogues of Steiner's formula (1.2.7), now called a tube formula, for subsets of spheres, with the Lipschitz-Killing curvatures appearing in the Euclidean case replaced by their spherical counterparts. We shall treat these in some detail in Chap. 3.4.2, but, for the moment, let us write them as $\mathcal{L}_j^1(M)$ so that, assuming the existence of a tube formula, (2.4.9) becomes

$$\begin{aligned} \mathbb{P} \left\{ \sup_{t \in M} f_t \geq u \right\} &= \sum_{j=0}^{\dim M} C_{\ell j} \mathcal{L}_j^1(M) \int_u^\infty (\cos^{-1}(u/r))^{\ell-j} \mathbb{P}_{|\xi|}(dr), \\ &= \sum_{j=0}^{\dim M} C_{\ell j} \mathcal{L}_j^1(M) G_{\ell j}(u) \end{aligned} \quad (2.4.10)$$

for some identifiable constants $C_{\ell j}$ and functions $G_{\ell j}$. Note that the final expression here is starting to take on the form of the right hand side of the Gaussian kinematic formula.

Where this argument breaks down is that the tube formula only works for small enough ρ or, in our case, small enough r . If r is large in the integrand of (2.4.7) then the tube around M has radius close to $\pi/2$, and it becomes easy, and, indeed, typical, for the tube to intersect itself ‘on the other side’ of the sphere in which it is embedded. Once a self-intersection of this kind occurs, tube formulae are no longer valid.

One way around this, which we shall not adopt in these notes, is to note that since the problems arise only for large r , and these have small probability under $\mathbb{P}_{|\xi|}$, one can ignore the tail of the integral, in a u -dependent fashion, and estimate the error involved in doing so. Then, however, (2.4.10) becomes an approximation rather than an exact result. We prefer to use the Euler characteristic approximation of (1.3.7) and will justify it later. In most cases, the two approaches yield identical approximations (cf. [78].)

The second problem with approaching everything via the canonical process on the sphere is that most random fields do not live in the sphere, and although the mapping from $M \rightarrow \tilde{\varphi}(M)$ is natural one, in the final analysis one would like to have answers that depend not on the structure of $\tilde{\varphi}(M)$, but rather on the structure of M and the covariance structure of f . This is, in fact, not too hard to do, and we shall see later how to relate the $\mathcal{L}_j^1(\tilde{\varphi}(M))$ to the $\mathcal{L}_j(M)$.

The final problem with this approach, however, is highly non-trivial: Not all random fields have orthogonal expansions with only a finite number of terms. In fact, this is the exception rather than the rule. For example, *no* isotropic random field on \mathbb{R}^N has a finite expansion! Nevertheless, the isonormal process on the sphere turns out to be the key example for generating results for general processes, as we shall see later.

2.4.2 The Canonical Process and Geometry

Returning now to the original random field f on M , consider how the excursion sets of f relate to those of \tilde{f} . That is, what is the relation between

$$A_D = \{t \in M : f(t) \in D\} \quad \text{and} \quad \tilde{A}_D = \{x \in \tilde{\varphi}(M) : \tilde{f}(x) \in D\}?$$

The first thing to notice is that since $\tilde{\varphi}$ is one-one (already assumed) and if we assume that it is C^2 or smoother (in fact, it will always be at least C^4 for us) then the fact that $\tilde{\varphi}$ is a diffeomorphism implies that the Euler characteristics of A_D and \tilde{A}_D will be identical. Consequently,

$$\mathbb{E}\{\mathcal{L}_0(A_D)\} = \mathbb{E}\{\mathcal{L}_0(\tilde{A}_D)\}, \quad (2.4.11)$$

so that if we can compute the expected Euler characteristics of excursion sets for the canonical process on spheres, then we can, at least in principle, compute them for all Gaussian random fields with finite expansions.

Of course, we shall still face the same two problems that we faced above. The answers will depend on the structure of $\tilde{\varphi}(M)$, rather than on the structure of M and the covariance structure of f , and they will only hold for random fields with finite expansions.

Furthermore, it is not at all clear if one can extend (2.4.11) to Lipschitz-Killing curvatures other than the Euler characteristic. For example, it is certainly not true in general that

$$\mathcal{L}_N(A_D) \equiv \mathcal{H}_N(A_D) = \mathcal{H}_N(\tilde{A}_D) \equiv \mathcal{L}_N(\tilde{A}_D),$$

where $\mathcal{H}_N(A_D)$ is the Euclidean volume of A_D but $\mathcal{H}_N(\tilde{A}_D)$ is the surface area of \tilde{A}_D as a subset of the sphere. That there is nevertheless a way to obtain the general Gaussian kinematic formula, which gives an expression for means of all the $\mathcal{L}_j(A_D)$ from a parallel result for the canonical process, is one of the mysteries that will be unravelled as you proceed through these notes.

2.5 The Basic Theory of Gaussian Fields

To make the lecture course for which these notes were prepared complete and self-contained, we would have needed another 24 h or so to give a mini-course on Gaussian processes. Thus, for example, if you look at *RFG* (and by now you should have ordered a personal copy from Springer) you will see that the first third of the book is devoted to this material.

Clearly this was not possible. On the other hand, we do need some results from the general theory, and some specific moment results, for later use, and so they are collected in the following sections, with no attempt to prove anything. Everything is proven in *RFG* in full detail.

In fact, if you are reading through these notes by yourself, and have an impatient nature, you can actually skip these sections for now, go directly to the geometry of Chap. 3, and return later, as needed.

We should really begin by actually defining real valued *Gaussian (random) fields* or *Gaussian (random) processes*, something which have not actually done yet, as being a random fields for which the (finite dimensional) distributions of $(f_{t_1}, \dots, f_{t_n})$ are multivariate Gaussian for each $1 \leq n < \infty$ and each collection $(t_1, \dots, t_n) \in M^n$.

Since multivariate Gaussian distributions are determined by means and covariances, it is immediate that Gaussian random fields are determined by their mean and covariance functions defined, respectively, by

$$m(t) = \mathbb{E}\{f(t)\} \tag{2.5.1}$$

and

$$C(s, t) = \mathbb{E} \{ (f(s) - m(s)) (f(t) - m(t)) \}. \quad (2.5.2)$$

In fact, this is one of the main reasons, beyond ubiquitous but not always justified appeals to the central limit theorem, that Gaussian processes are such popular and useful choices for models for random processes on general spaces.

2.5.1 Regularity for Gaussian Process

We have already spoken about continuous and differentiable fields, but have said nothing about conditions that ensure this. In the Gaussian case, everything is dependent on the size of the parameter space, which we shall measure via the canonical metric.⁴

The *canonical metric*, d , of a zero mean Gaussian field on a topological space M , is defined by setting

$$d(s, t) \triangleq [\mathbb{E} \{ (f(s) - f(t))^2 \}]^{\frac{1}{2}}, \quad (2.5.3)$$

in a notation that will henceforth remain fixed.⁵ A ball in this metric, of radius ε and centered at a point $t \in M$ is denoted by

$$B_d(t, \varepsilon) \triangleq \{s \in M : d(s, t) \leq \varepsilon\}. \quad (2.5.4)$$

Assume that M is d -compact, in the sense that

$$\text{diam}(M) \triangleq \sup_{s, t \in M} d(s, t) < \infty. \quad (2.5.5)$$

Fix $\varepsilon > 0$ and let $N(M, d, \varepsilon) \equiv N(\varepsilon)$ denote the smallest number of d -balls of radius ε whose union covers M . Set

$$H(M, d, \varepsilon) \equiv H(\varepsilon) = \ln(N(\varepsilon)). \quad (2.5.6)$$

Then N and H are called the (metric) *entropy* and *log-entropy* functions for M (or f).

⁴ There is also a more powerful approach based on the notion of *majorising measures* which we shall not adopt. For information on this approach see *RFG* and the far more serious treatment in [60].

⁵ Actually, d is only a pseudo-metric, since although it satisfies all the other demands of a metric, $d(s, t) = 0$ does not necessarily imply $s = t$.

Here then is the main result about Gaussian continuity and boundedness, due originally, more or less in the form given below, to Richard Dudley [31, 32]. It is not the latest word in the topic, but it will more than suffice for our purposes. Note how the topological and geometric structure of M blend together with the covariance structure of f to give a measure, the metric entropy, which determines everything in this result.

This blending of the geometry of the parameter space together with a metric derived from the random field will also lie at the heart of the Gaussian kinematic formula, although it will be different geometry and a different metric.

Theorem 2.5.1. *Let f be a centered Gaussian field on a d -compact M , d the canonical metric, and H the corresponding log-entropy. Then there exists a universal constant K such that*

$$\mathbb{E} \left\{ \sup_{t \in M} f_t \right\} \leq K \int_0^{\text{diam}(M)} H^{1/2}(\varepsilon) d\varepsilon, \quad (2.5.7)$$

and

$$\mathbb{E} \{ \omega_{f,d}(\delta) \} \leq K \int_0^\delta H^{1/2}(\varepsilon) d\varepsilon, \quad (2.5.8)$$

where

$$\omega_{f,d}(\delta) \triangleq \sup_{d(s,t) \leq \delta} |f(t) - f(s)|, \quad \delta > 0, \quad (2.5.9)$$

Furthermore, there exists a random $\eta \in (0, \infty)$ and a universal constant K such that

$$\omega_{f,d}(\delta) \leq K \int_0^\delta H^{1/2}(\varepsilon) d\varepsilon,$$

for all $\delta < \eta$.

A complement to this result states that f is also stationary, then

$$\begin{aligned} f \text{ is a.s. continuous on } M &\iff f \text{ is a.s. bounded on } M \\ &\iff \int_0^\delta H^{1/2}(\varepsilon) d\varepsilon < \infty, \quad \forall \delta > 0. \end{aligned} \quad (2.5.10)$$

For necessary and sufficient conditions in the general case one needs to turn to the notion of majorising measures mentioned above.

2.5.2 Gaussian Fields on \mathbb{R}^N

The entropy conditions above yield very simple sufficient conditions for continuity of centered Gaussian fields on compact sets M of \mathbb{R}^N . In fact, it is easy to check that, defining

$$p^2(u) \triangleq \sup_{|s-t| \leq u} \mathbb{E} \{ |f_s - f_t|^2 \}, \quad (2.5.11)$$

a.s. continuity and boundedness follow if, for some $\delta > 0$, either

$$\int_0^\delta (-\ln u)^{\frac{1}{2}} dp(u) < \infty \quad \text{or} \quad \int_\delta^\infty p(e^{-u^2}) du < \infty. \quad (2.5.12)$$

Furthermore, there exists a constant K' , dependent only on the dimension N , and a random $\delta_o > 0$, such that, for all $\delta < \delta_o$,

$$\omega_f(\delta) \leq K' \int_0^{p(\delta)} (-\ln u)^{\frac{1}{2}} dp(u), \quad (2.5.13)$$

where the modulus of continuity ω_f is as in (2.5.9), but taken with respect to the usual Euclidean metric rather than the canonical one. A similar bound, in the spirit of (2.5.8), holds for $\mathbb{E}\{\omega_f(\delta)\}$.

A sufficient condition for either integral in (2.5.12) to be finite is that, for some $0 < K < \infty$ and $\alpha, \eta > 0$,

$$\mathbb{E} \{ |f_s - f_t|^2 \} \leq \frac{K}{|\log |s - t||^{1+\alpha}},$$

for all s, t with $|s - t| < \eta$. Related conditions hold on the spectral density in the stationary case. See *RFG* for details.

In practical situations, it is rare indeed that one even gets close to the logarithmic behavior of (2.5.14). The more common situation is that the covariance function has a power series representation of the form

$$C(s, t) = C(t, t) + (t - s)\Lambda_t(t - s)' + o(|t - s|^{2+\delta}), \quad (2.5.14)$$

for $|t - s|$ small and some $\delta > 0$, or, in the stationary case

$$C(t) = C(0) + t\Lambda t' + o(|t|^{2+\delta}), \quad (2.5.15)$$

for t in the neighborhood of the origin. The matrices Λ_t and Λ are $N \times N$ and positive definite.

2.5.3 Differentiability

Since we shall also be requiring that our random functions are a.s. C^2 , a few words on this condition are also in order. Firstly, unlike continuity, which requires nothing of the parameter space M other than it have a topology (so one can talk about continuity) differentiability requires that M itself has a differentiable structure. For the moment, we limit ourselves to \mathbb{R}^N with its usual structure.

It then turns out that, at least in the Gaussian scenario, differentiability can be handled within the framework of continuity since derivatives, if they exist, must still be Gaussian. Since this is an important observation, that has been missed by many authors in the past, we shall deviate from the policy of this section and actually give details of how to do things.

To start, we need to define L^2 derivatives. Choose a point $t \in \mathbb{R}^N$ and a sequence of k 'directions' t'_1, \dots, t'_k in \mathbb{R}^N , and write these as $t' = (t'_1, \dots, t'_k)$. We say that f has a k -th order L^2 partial derivative at t , in the direction t' , if the limit

$$D_{L^2}^k f(t, t') \triangleq \lim_{h_1, \dots, h_k \rightarrow 0} \frac{1}{\prod_{i=1}^k h_i} \Delta^k f(t, t', h) \quad (2.5.16)$$

exists in mean square, where $h = (h_1, \dots, h_k)$. Here $\Delta^k f(t, t', h)$ is the symmetrized difference

$$\Delta^k f(t, t', h) = \sum_{s \in \{0,1\}^k} (-1)^{k - \sum_{i=1}^k s_i} f\left(t + \sum_{i=1}^k s_i h_i t'_i\right)$$

and the limit in (2.5.16) is interpreted sequentially, i.e. first send h_1 to 0, then h_2 , etc. Note that if f is Gaussian then so are its L^2 derivatives, when they exist.

By choosing $t' = (e_{i_1}, \dots, e_{i_k})$, where e_i is the vector with i -th element 1 and all others zero, we can talk of the mean square partial derivatives

$$\frac{\partial^k}{\partial t_{i_1} \dots \partial t_{i_k}} f(t) \triangleq D_{L^2}^k f(t, (e_{i_1}, \dots, e_{i_k})) \quad (2.5.17)$$

of f of various orders.

Moving now to almost sure differentiability, first endow the space $\mathbb{R}^N \times \otimes^k \mathbb{R}^N$ with the norm

$$\|(s, s')\|_{N,k} \triangleq |s| + \|s'\|_{\otimes^k \mathbb{R}^N} = |s| + \left(\sum_{i=1}^k |s'_i|^2 \right)^{1/2},$$

and write $B_{N,k}(y, h)$ for the ball centered at $y = (t, t')$ and of radius h in the metric induced by $\|\cdot\|_{N,k}$. Furthermore, write

$$M_{k,\rho} \triangleq M \times \{t' : \|t'\|_{\otimes^k \mathbb{R}^N} \in (1 - \rho, 1 + \rho)\}$$

for the product of M with the ρ -tube around the unit sphere in $\otimes^k \mathbb{R}^N$. This is enough to allow us to formulate

Theorem 2.5.1. *Suppose f is a centered Gaussian random field on an open $M \in \mathbb{R}^N$, possessing k -th order partial derivatives in the L^2 sense in all directions everywhere inside M . Suppose, furthermore, that there exists $0 < K < \infty$, and $\rho, \delta, h_0 > 0$ such that for $0 < \eta_1, \eta_2, h < h_0$,*

$$\begin{aligned} & \mathbb{E} \left\{ [\eta_1^{-k} \Delta^k f(t, t', \eta_1 1) - \eta_2^{-k} \Delta^k f(s, s', \eta_2 1)]^2 \right\} \\ & < K |\ln(\|(t, t') - (s, s')\|_{N,k} + |\eta_1 - \eta_2|)|^{-(1+\delta)}, \end{aligned} \quad (2.5.18)$$

for all

$$((t, t'), (s, s')) \in M_{k,\rho} \times M_{k,\rho} : (s, s') \in B_{N,k}((t, t'), h),$$

where $\eta_j 1$ denotes the k -vector all of whose elements are η_j . Then, with probability one, f is k times continuously differentiable.

Proof. Recalling that we have assumed the existence of L^2 derivatives, we can define the Gaussian field

$$\widehat{f}(t, t', \eta) = \begin{cases} \Delta^k f(t, \eta t') & \eta \neq 0, \\ D_{L^2}^k f(t, t') & \eta = 0, \end{cases}$$

where $D_{L^2}^k f$ is the mean square derivative (2.5.16). This process is defined on the parameter space $\widehat{M} \triangleq M_{k,\rho} \times (-h, h)$, an open subset of the finite dimensional vector space $\mathbb{R}^N \times \otimes^k \mathbb{R}^N \times \mathbb{R}$, with norm

$$\|(t, t', \eta)\|_{N,k,1} = \|(t, t')\|_{N,k} + |\eta|.$$

Whether or not \widehat{f} is k times differentiable on M is clearly the same issue as whether or not \widehat{f} is continuous in \widehat{M} , with the issue of the continuity of \widehat{f} really being only on the hyperplane where $\eta = 0$. But this puts us back into the setting of the previous subsection, and it is easy to check that condition (2.5.14) there translates to (2.5.18) in the current scenario. \square

As for continuity, it is rare in practice to get close to the upper bound in (2.5.18), and this condition will easily be satisfied if, in analogy to (2.5.14) and (2.5.15), the covariance function has a Taylor series expansion of up to order $2k$ with a remainder of $o(|h|^{2k+\eta})$ for some $\eta > 0$.

2.6 Stationarity, Isotropy, and Constant Variance

Although we have already met both stationarity and isotropy more than once, the time has now come to define them properly and list some of their basic properties.

We start by noting that a random field on a general parameter space M is called (second order) *stationary*, or *homogeneous*, if it has constant means and the covariance function $C(s, t)$ is a function of the difference $s - t$ only. With some abuse of notation, we shall write $C(s, t) = C(s - t)$.

Of course, if M is general, there is no reason why $s, t \in M$ implies that $s - t$ is also in M , and so it is necessary to assume that M has a group structure. In these notes, when discussing stationarity, we shall be concerned only with the cases $M = \mathbb{R}^N$ or $M = S_\lambda(\mathbb{R}^N)$. Note that for Gaussian processes this definition of stationarity also implies what is known as *strong stationarity*, which is that the finite dimensional distributions of the field are invariant under translations.

A stationary field is called *isotropic* if the covariance function is direction independent, in the sense that $C(t) = C(|t|)$.

We now restrict attention to random fields on \mathbb{R}^N . There are two basic results in the theory of stationary processes. One is known as the *spectral distribution theorem* and one as the *spectral representation theorem*. The first, which is the only one that we shall need in these notes, is due originally to Bochner in a non-probabilistic setting. For fields on \mathbb{R}^N it states that if a continuous function $C : \mathbb{R}^N \rightarrow \mathbb{R}$ is non-negative definite, and so the covariance function of a stationary random field, if and only if there exists a finite measure ν on the Borel σ -field \mathbb{B}^N of \mathbb{R}^N such that

$$C(t) = \int_{\mathbb{R}^N} e^{i\langle t, \lambda \rangle} \nu(d\lambda), \quad (2.6.1)$$

for all $t \in \mathbb{R}^N$.

The measure ν is called the *spectral measure* and, since C is real, must be symmetric, in the sense that $\nu(A) = \nu(-A)$ for all $A \in \mathbb{B}^N$. Similarly, if C is isotropic then ν must be spherically symmetric, in the sense that $\nu(A) = \nu(\Theta A)$ for all $A \in \mathbb{B}^N$ and any rotation Θ .

2.6.1 Spectral Moments and Derivatives of Random Fields

Given the spectral representation (2.6.1) we define the *spectral moments*

$$\lambda_{i_1 \dots i_N} \triangleq \int_{\mathbb{R}^N} \lambda_1^{i_1} \cdots \lambda_N^{i_N} \nu(d\lambda), \quad (2.6.2)$$

for all (i_1, \dots, i_N) with $i_j \geq 0$. Note that, since ν is symmetric, the odd ordered spectral moments, when they exist, are zero; i.e.

$$\lambda_{i_1 \dots i_N} = 0, \quad \text{if } \sum_{j=1}^N i_j \text{ is odd.} \quad (2.6.3)$$

Spectral moments turn out to be closely related to the variances and covariances of derivatives of random fields.

Recalling the notion of mean square partial derivatives from (2.5.17) it is a straightforward exercise to check that, in general, their covariance functions are given by

$$\mathbb{E} \left\{ \frac{\partial^k f(s)}{\partial s_{i_1} \partial s_{i_1} \dots \partial s_{i_k}} \frac{\partial^k f(t)}{\partial t_{i_1} \partial t_{i_1} \dots \partial t_{i_k}} \right\} = \frac{\partial^{2k} C(s, t)}{\partial s_{i_1} \partial t_{i_1} \dots \partial s_{i_k} \partial t_{i_k}}. \quad (2.6.4)$$

When f is stationary, the corresponding variances and covariances also have a nice representation in terms of spectral moments. For example, if f has mean square partial derivatives of orders $\alpha + \beta$ and $\gamma + \delta$ for $\alpha, \beta, \gamma, \delta \in \{0, 1, 2, \dots\}$, then

$$\begin{aligned} \mathbb{E} \left\{ \frac{\partial^{\alpha+\beta} f(t)}{\partial s_{t_i} \partial s_{t_j}} \frac{\partial^{\gamma+\delta} f(t)}{\partial \gamma t_k \partial \delta t_l} \right\} &= (-1)^{\alpha+\beta} \frac{\partial^{\alpha+\beta+\gamma+\delta} C(t)}{\partial s_{t_i} \partial s_{t_j} \partial \gamma t_k \partial \delta t_l} \Big|_{t=0} \quad (2.6.5) \\ &= (-1)^{\alpha+\beta} i^{\alpha+\beta+\gamma+\delta} \int_{\mathbb{R}^N} \lambda_i^\alpha \lambda_j^\beta \lambda_k^\gamma \lambda_l^\delta \nu(d\lambda). \end{aligned}$$

Note that although this equation seems to have some asymmetries in the powers, these disappear due to the fact that all odd ordered spectral moments, like all odd ordered derivatives of C , are identically zero.

Here are some important special cases of the above, for which we adopt the shorthand $f_j = \partial f / \partial t_j$ and $f_{ij} = \partial^2 f / \partial t_i \partial t_j$ along with a corresponding shorthand for the partial derivatives of C .

- (a) f_j has covariance function $-C_{jj}$ and thus variance $\lambda_{2e_j} = -C_{jj}(0)$, where e_j is the vector with a 1 in the j -th position and zero elsewhere.
- (b) In view of (2.6.3), and taking $\alpha = \gamma = \delta = 0$, $\beta = 1$ in (2.6.5)

$$f(t) \text{ and } f_j(t) \text{ are uncorrelated,} \quad (2.6.6)$$

for all j and all t . If f is Gaussian, this is equivalent to independence. Note that (2.6.6) does *not* imply that f and f_j are uncorrelated *as processes*.

In general, for $s \neq t$, we will have that $\mathbb{E}\{f(s)f_j(t)\} = -C_j(s-t) \neq 0$.

- (c) Taking $\alpha = \gamma = \delta = 1$, $\beta = 0$ in (2.6.5) gives that

$$f_i(t) \text{ and } f_{kl}(t) \text{ are uncorrelated} \quad (2.6.7)$$

for all i, k, l and all t . Again, if f is Gaussian, this is equivalent to independence.

Under the additional condition of isotropy, with its implication of spherical symmetry for the spectral measure, the structure of the spectral moments simplifies significantly, as do the correlations between various derivatives of f . In particular, it follows immediately from (2.6.5) that

$$\mathbb{E}\{f_i(t)f_j(t)\} = -\mathbb{E}\{f(t)f_{ij}(t)\} = \lambda_2\delta_{ij} \quad (2.6.8)$$

where δ_{ij} is the Kronecker delta and λ_2 is the *second spectral moment*

$$\lambda_2 \triangleq \int_{\mathbb{R}^N} \lambda_i^2 \nu(d\lambda), \quad (2.6.9)$$

which, because of isotropy, is independent of the value of i . Consequently, if f is Gaussian, then the first order derivatives of f are independent of one another, in addition to being independent of f itself.

Finally, we note that a similar argument shows that even if f is neither stationary nor isotropic, but does have constant variance, then it is still true that f and its first order derivatives are uncorrelated.

2.6.2 Local Isotropy and the Induced Metric

Of all the relationships between spectral moments in the previous subsection, the most important is probably (2.6.8), which describes the lack of correlation between first order derivatives of random fields under isotropy. It turns out that, in the case of Gaussian fields, this makes many computations that are, a priori, quite forbidding actually quite easy. Thus it is not surprising that the theory of Gaussian fields began with the isotropic case.

It is not in general possible to transform non-isotropic fields to isotropic ones, but there are a number of ways to ensure that first order derivatives are uncorrelated. This property is important enough that we shall give it a name, defining random fields with constant mean and variance, and uncorrelated first order derivatives, to be *locally isotropic*.

It turns out that it is easy to transform non-isotropic but stationary random fields f on \mathbb{R}^N to locally isotropic ones. If Λ is the $N \times N$ matrix of second spectral moments λ_{ij} , then it is trivial to check that the field defined by $\tilde{f}(t) = f(\Lambda^{-1/2}t)$ is locally isotropic. (cf. (2.5.14).)

In the non-stationary case there is no such simple transformation available. However, there is a trick, based on Riemannian geometry, that allows one to change the Riemannian structure of the parameter space by introducing a Riemannian metric related to the covariance function that makes all first order *Riemannian* derivatives uncorrelated. It was this trick that, in many

ways, was one of the most important themes of *RF*G, and is what allows one to move from a theory of stationary random fields on subsets of \mathbb{R}^N to non-stationary fields on stratified manifolds. We shall see how this works later, in Sect. 4.5 when we introduce this special (induced) metric at (4.5.1).

2.7 Three Gaussian Facts

We close this chapter with three facts about multivariate Gaussian random variables that we shall need later. All are well known and easy to check, and we include them now only so that they will be easy to refer back to later.

It follows from the form (2.0.1) of the multivariate Gaussian density that if $X \sim N_d(m, C)$ then its characteristic function is given by

$$\phi(\theta) = \mathbb{E}\{e^{i\langle\theta, X\rangle}\} = e^{i\langle\theta, m\rangle - \theta C \theta' / 2}, \quad (2.7.1)$$

where $\theta \in \mathbb{R}^d$. From this follows the fact that, if A is a $d \times d$ matrix, then

$$XA \sim N(mA, A'CA). \quad (2.7.2)$$

Next, if $n < d$, make the partitions

$$\begin{aligned} X &= (X^1, X^2) = ((X_1, \dots, X_n), (X_{n+1}, \dots, X_d)), \\ m &= (m^1, m^2) = ((m_1, \dots, m_n), (m_{n+1}, \dots, m_d)), \\ C &= \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}, \end{aligned}$$

where C_{11} is an $n \times n$ matrix. Then each X^i is $N(m^i, C_{ii})$ and the conditional distribution of X^i given X^j is also Gaussian, with mean vector

$$m_{i|j} = m^i + (X^j - m^j)C_{jj}^{-1}C_{ji} \quad (2.7.3)$$

and covariance matrix

$$C_{i|j} = C_{ii} - C_{ij}C_{jj}^{-1}C_{ji}. \quad (2.7.4)$$

Finally, we quote a fundamental moment result known as *Wick's formula*. This states that if $X = (X_1, X_2, \dots, X_d) \sim N(0, C)$ then, for any non-negative integer m ,

$$\mathbb{E}\{X_1 X_2 \cdots X_{2m+1}\} = 0, \quad (2.7.5)$$

$$\begin{aligned} \mathbb{E}\{X_1 X_2 \cdots X_{2m}\} &= \sum \mathbb{E}\{X_{i_1} X_{i_2}\} \cdots \mathbb{E}\{X_{i_{2m-1}} X_{i_{2m}}\} \\ &= \sum C(i_1, i_2) \cdots C(i_{2m-1}, i_{2m}), \end{aligned} \quad (2.7.6)$$

where the sum is taken over the $(2m)!/m!2^m$ different ways of grouping X_1, \dots, X_{2m} into m pairs. Wick's formula can be proven by successive differentiation of the characteristic function (2.7.1).

Topological Complexity of Smooth Random Functions

École d'Été de Probabilités de Saint-Flour XXXIX-2009

Adler, R.; Taylor, J.E.

2011, VIII, 122 p. 15 illus., 9 illus. in color., Softcover

ISBN: 978-3-642-19579-2