ANNALES
DE LA FACULTÉ
DES SCIENCES
DE
TOULOUSE
Mathématiques

MARIO WSCHEBOR

Smoothing and occupation measures of stochastic processes

Tome XV, nº 1 (2006), p. 125-156.

<http://afst.cedram.org/item?id=AFST_2006_6_15_1_125_0>

Annales de la Faculté des Sciences de Toulouse Vol. XV, n° 1, 2006
pp. 125–156

Smoothing and occupation measures of stochastic processes[(*)]

Mario Wschebor [(1)]

Abstract. — This is a review paper about some problems of statistical inference for one-parameter stochastic processes, mainly based upon the observation of a convolution of the path with a non-random kernel. Most of the results are known and presented without proofs. The tools are first and second order approximation theorems of the occupation measure of the path, by means of functionals defined on the smoothed paths. Various classes of stochastic processes are considered starting with the Wiener process, Gaussian processes, continuous semi-martingales and Lévy processes. Some statistical applications are also included in the text.

Résumé. — Cet article est une révision d'un certain nombre de problèmes statistiques concernant les processus aléatoires à un paramètre continu. En général, on suppose que l'observable est une régularisation de la trajectoire du processus, obtenue par convolution avec un noyau détérministe. La plupart des résultats ici exposés est connue et presentée sans démonstration. Les énoncés des théorèmes contiennent des approximations de la mesure d'occupation, au premier et deuxième ordre, basées sur des fonctionnelles définies sur les régularisées des trajectoires. On considère diverses classes de processus, à savoir, le processus de Wiener, les processus gaussiens, les semi-martingales continues et les processus de Lévy. Nous avons inclus les détails de certaines applications statistiques.

1. Introduction

The content of this paper is motivated by the purpose of making statistical inference on continuous parameter stochastic processes, on the basis of the observation of a smooth approximation of a trajectory. Our interest lies in the situation in which this random path is a non-smooth function. We intend to make inference on those parameters affecting the regularity,

(*) Reçu le 22 octobre 2004, accepté le 23 février 2005
(1) Centro de Matemática, Facultad de Ciencias, Universidad de la República, Calle Iguá 4225. 11400, Montevideo (Uruguay).
E-mail: wschebor@cmat.edu.uy

so that understanding the irregularity of the path should allow to obtain information on them.

Most of the results in this paper are known and published. The intention is a unified presentation including some applications to statistical problems. There are also some new results for which we will provide proofs or at least some indications of how they might be proved. However, several important questions remain unanswered.

Let us consider as a typical example the stochastic differential equation

$$dX(t) = b(t, X(t)\ dt + \sigma(t, X(t)\ dZ(t)\ (t > 0) \quad X(0^+) = x_0 \tag{1.1}$$

where b, σ are regular functions and Z is some noise.

In a certain number of relevant cases, statistics on the drift function b is classical and well-established since a long time. (See for example the books by Lipster and Shiryaev [L-S], Ch. 7, 17 or Prakasa Rao [PR]).

On the contrary, if one wants to make inference on the noise part, i.e. on the function σ, the situation becomes more difficult due to the singularity of the measures induced on the space of trajectories by different parameter values, which is an obstacle to apply likelihood methods. One can try to overcome these difficulties by looking at the behaviour of the likelihood quotient for different parameter values of the finite-dimensional projections in the path space, and obtain asymptotic expansions when the time-grid is refined. This corresponds to observing the solution of (1.1) on a finite number of parameter values or, equivalently, its associated polygonal approximation, and the natural problems turn into understanding the behaviour of the level sets of random polygonals, when one refines the grid.

Results on first order approximations of the number of level crossings of random polygonals and some other related functionals have been considered in the context of the theories of random walks and convergence of empirical measures. A typical result is the following (see [R],[C-R1],[C-R2] and references therein): in an appropriate probability space, if $\{S_n : n \geqslant 0\}$ is a random walk with centered *i.i.d.* jumps $\{X_n : n \geqslant 0\}$ and the common law of the X_n's satisfies certain regularity and boundedness conditions, then, for any $\delta > 0$, almost surely:

$$\sup_{u \in \mathcal{R}} \left| \rho N_{u,n} - L^W(u, [0, n]) \right| = \ o(n^{\frac{1}{4}+\delta}) \tag{1.2}$$

where

- $N_{u,n}$ denotes the number of roots of the equation $S_k(t) = u$ in the interval $[0, n]$, $\{S_k(t) : t \geqslant 0\}$ denoting the function with polygonal graph and vertices $\{(k, S_k) : k \geqslant 1\}$,
- $L^W(u, J)$ is the local time of the Wiener process on interval J at the level u,
- $\rho = E(|X_1|)$.

A related problem consists in approximating the local time by the normalized number of crossings of polygonal approximations of the paths of a stochastic process. Using rescaling, a simple consequence of (1.2) is the following: Denote by $X^{(n)}$ the polygonal approximation of the path W corresponding to the grid $\{\frac{k}{n} : 0 \leqslant k \leqslant n\}$, that is, $X^{(n)}(t) = (1-nt+k)W(\frac{k}{n}) + (nt-k)W(\frac{k+1}{n})$ if $\frac{k}{n} \leqslant t \leqslant \frac{k+1}{n}$. Also, $N_u(g, I)$ denotes the number of roots of equation $g(t) = u$ such that $t \in I$.

Then, (1.2) implies that

$$\sqrt{\frac{\pi}{2n}} N_0(X^{(n)}, [0,1]) \text{ converges to } L^W(0, [0,1]) \text{ as } n \to +\infty. \tag{1.3}$$

in the sense of weak convergence of probability distributions on the line. In fact one can prove by using some other methods that (1.3) holds true in the L^p of the probability space for every $p > 0$ (see for example [A1],[A2] where L^p-convergence of normalized crossings of polygonal approximations is studied for various classes of random processes).

Almost sure approximations of the local time, seem to have started with Paul Lévy's work. Classical well-known results are the following:

1. almost surely,

$$\sqrt{\frac{\pi\varepsilon}{2}} \nu_\varepsilon^W([0,t]) \to 2L^W(0,[0,t]) \quad \text{as } \varepsilon \to 0 \quad \text{for all } t \geqslant 0 \tag{1.4}$$

where $\nu_\varepsilon^W([0,t])$ denotes the number of excursions of the path W with respect to level $u = 0$, having length greater than ε.

2. almost surely,

$$\varepsilon D_t^W(\varepsilon) \to 2L^W(0,[0,t]) \quad \text{as } \varepsilon \to 0 \quad \text{for all } t \geqslant 0 \tag{1.5}$$

where $D_t^W(\varepsilon)$ denotes the number of downcrossings of the strip $[0, \varepsilon]$ performed by the path W during the time interval $[0, t]$.

General treatments of this subject can be found, for example in the books [I-M],[I-W],[K-S]. Both results (1.4) and (1.5) have been extended to a general class of real-valued Markov processes (see [F-T]).

If one is willing to use these kind of results for statistical purposes, a certain number of difficulties arise.

First, the use of polygonal approximations or Lévy-type results, requires the observation of functionals of the actual path, which may be a diffusion or a diffusion-like process, which is non-differentiable, and in principle can be observed only after smoothing. Then, these results can't be applied directly since one must know what happens when replacing the values of the process at the grid times by the values of an approximating observable process at the same instants. This is by no means trivial.

The majority of the results we are going to consider in this paper try to overcome this difficulty by observing a regularization of the path obtained with a convolution device. So that, instead of polygonal approximations of the path we will deal with smooth functions. This will not happen all the time: in some cases we will go back to polygonal approximations and the corresponding results.

Second, for statistical purposes, theorems on almost sure convergence or convergence in probability – which we call here "first order approximations" – are not enough, one also needs speeds. We have included some speed results, both for polygonal approximations and smoothing.

A *third* problem is that in some of the first order approximation or speed theorems, the local time appears in the statements. Of course this is a serious difficulty, since generally speaking the local time of a path can't be observed and the approximations that we know are too slow, so that we are unable to put them instead of the local time to obtain asymptotic results. To face this problem we integrate in the state space of the process. Once this is done, the approximations we get are not for each level (as local time approximations) but integrated results and one can handle them to obtain theorems which are satisfactory from the standpoint of the statistical applications.

The account we give below of this type of results is far from complete. We have not included multiparameter processes, for which this author is aware of first order results only for Gaussian stationary fields defined on $\mathcal{R}^d$ (in which some speed results are also known, see [B-W]) and for the $d-$parameter Wiener sheet ([W1],[W2]). For one-parameter continuous semi-martingales, we give a quite general picture, excepting for technical generalizations. On the contrary, there are many gaps in the understanding of processes with jumps, Lévy processes having jump part with locally unbounded variation, or solutions of SDE with noise part having jumps.

We have included without proofs some few specific statistical applications, and discuss some of them for the first time here. The simplest example, which already contains problems that show some general difficulties, is making inference on the variance in a simple regression model in continuous time.

2. First order approximations

2.1. Wiener process

Let $\{W(t) : t \geqslant 0\}$ be a standard Wiener process. Our starting point is the following property of the paths:

$$a.s. \quad \lambda\left(\left\{t \in [0,1] : \frac{W(t+\varepsilon) - W(t)}{\sqrt{\varepsilon}} \leqslant x\right\}\right) \to P(\xi \leqslant x) \quad \text{as } \varepsilon \downarrow 0 \tag{2.1}$$

for every real x, where λ denotes Lebesgue measure on the line and ξ is a standard normal random variable.

(2.1) is easily proved by using the Borel-Cantelli Lemma and the Hölder condition for the Wiener paths (see [W4]). In fact one can prove a somewhat stronger result, i.e. that almost surely, as $\varepsilon \to 0$, one has moment convergence of the distribution of the functions $t \rightsquigarrow \frac{W(t+\varepsilon)-W(t)}{\sqrt{\varepsilon}}$ defined on the probability space $([0,1], \lambda)$.

One can consider (2.1) as a positive result as opposite to the law of the iterated logarithm. We mean the following:

Consider the normalized increments $Z_\varepsilon(t) = \frac{W(t+\varepsilon)-W(t)}{a(\varepsilon)}$ where the normalizing function a is non-random and satisfies the mild natural property:

$$a \text{ is non-decreasing on some interval of the form } (0, \varepsilon_0) \text{ and } a(0^+) = 0. \tag{2.2}$$

A consequence of the law of the iterated logarithm is that there is no normalizing function a such that almost surely Z_ε converges to a non-trivial limit for almost every t.

A similar conclusion is obtained if instead of almost everywhere convergence with respect to the t-variable one looks at convergence in $L^p([0,1], \lambda)$.

In fact, if almost surely Z_ε converges in $L^p([0,1],\lambda)$ to a non-trivial random function Z for some $p>0$, one can show that $\frac{a(\varepsilon)}{\sqrt{\varepsilon}}$ must have a finite non-zero limit as $\varepsilon \downarrow 0$. It follows that Z is independent of the $\sigma-$algebra generated by $\{W(t) : 0 \leqslant t \leqslant 1\}$. Hence, for each $t \in [0,1]$ $Z(t)$ is almost surely non-random and one can check that convergence holds in $L^2([0,1],\lambda)$. The moment computation

$$E\left(\int_0^1 |Z_\varepsilon(t) - Z(t)|^2 \, dt\right) = \frac{h}{a^2(h)} + \int_0^1 Z^2(t)dt$$

shows that this is not possible.

However, (2.1) says that $a(\varepsilon) = \sqrt{\varepsilon}$ is a good normalizing function, in the sense that a positive convergence result is obtained if we replace the topology of almost everywhere (or L^p) convergence by weak convergence of measures.

A natural extension of (2.1) is as follows.

Let $\psi : \mathcal{R} \to \mathcal{R}^+$ be a C^1-function having compact support contained in $[-1,1]$, $\int_{-1}^1 \psi(x)dx = 1$. Put, for $\varepsilon > 0$, $\psi_\varepsilon(t) = \varepsilon^{-1}\psi(t/\varepsilon)$.

For any locally bounded real-valued measurable function g defined on the real line, denote

$$g^\varepsilon(t) = (g * \psi_\varepsilon)(t) = \int_{-\infty}^{+\infty} \psi_\varepsilon(t-s)g(s)ds.$$

the convolution of g with the approximation of unity ψ_ε . With these hypotheses, g^ε is of class C^1. We will use the same notation W^ε for the convolution of ψ_ε with the extension of $W(.)$ to the whole line, putting the value $W(0)$ on the negative half-axis.

Then, almost surely, for every bounded continuous function $f : \mathcal{R} \to \mathcal{R}$ and every bounded interval I contained in $\mathcal{R}^+$:

$$\sqrt{\frac{\pi\varepsilon}{2}} \frac{1}{\|\psi\|_2} \int_{-\infty}^{+\infty} f(u)\, N_u(W^\varepsilon, I)\, du \to \int_{-\infty}^{+\infty} f(u)\, L^W(u,I)\, du = \int_I f\left[W(t)\right]\, dt \tag{2.3}$$

as $\varepsilon \to 0$. In (2.3) $\|\psi\|_2$ is the L^2-norm of ψ and $L^W(u,I)$ the local time of the Wiener path at the value u, corresponding to interval I, that is, the continuous version of the Radon-Nikodym derivative with respect to Lebesgue measure of the occupation measure $\mu_I(B) = \lambda(\{t \in I, W(t) \in B\})$, B a Borel set in the line.

(2.3) can be proved by showing that

- (2.1) holds true if one replaces the normalized increment $\frac{W(t+\varepsilon)-W(t)}{\sqrt{\varepsilon}}$ by $\sqrt{\varepsilon}\,(W_\varepsilon)'\,(t)$ (in fact, the first corresponds to putting $\psi = 1_{[-1,0]}$ in the second) and the variance of ξ by $\|\psi\|_2^2$. This implies that for every continuous function $h : I \to \mathcal{R}$:

$$\int_I h(t)\left|\sqrt{\varepsilon}\,(W^\varepsilon)'\,(t)\right| dt \to \sqrt{\frac{2}{\pi}}\frac{1}{\|\psi\|_2}\int_I h(t)dt \quad \text{as } \varepsilon \to 0 \tag{2.4}$$

- Check that formula

$$\int_{-\infty}^{+\infty} f(u)\ N_u(g, I)\ du = \int_I f\left[g(t)\right]|g'(t)|\ dt \tag{2.5}$$

  is valid for continuous f and g of class C^1.

- Use (2.4), (2.5) and the fact that W^ε converges to W uniformly on I as $\varepsilon \to 0$.

The limit result (2.3) can be considered as an almost-sure-weak approximation of the local time of the Wiener process. It differs in two ways from classical almost sure approximations of the local time.

1. First, we are integrating in the state space instead of putting a Dirac δ-function in the place where function f is standing.

2. Second, we approximate the local time by means of a functional defined on the smoothed path W^ε instead of the underlying path W.

2.2. Semi-martingales with continuous paths

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\{\mathcal{F}_t\}_{t\geqslant 0}$ a filtration in it. $\{M(t) : t \geqslant 0\}$ is a real-valued local martingale adapted to the filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ having continuous paths. We denote $\{A_t : t \geqslant 0\}$ its quadratic variation process. For almost every ω, A_t' is the almost everywhere defined derivative of A_t. Notice that A_t' is non-negative whenever it exists.

For convinience in what follows we assume that $M(t)$ is defined for negative values of the argument, $M(t) = M(0)$ if $t < 0$.

The next two statements extend (2.1) and (2.3) to the almost sure weak convergence of the oscillations of $M(.)$ and the approximation of the occupation measure by the normalized crossings of the regularized path.

Theorem 2.1 [AW2]. — *Almost surely, for any real $x \neq 0$ one has*

$$\lambda(\{t \in [0,1] : \frac{M(t+\varepsilon)-M(t)}{\sqrt{\varepsilon}} \leqslant x\}) \to \int_0^1 P\left(A_t'^{1/2}\xi \leqslant x\right)\, dt. \qquad (2.6)$$

where ξ is standard normal and independent of the process M.

Notice that the - random - limit in (2.6) is a continuous function of x at $x \neq 0$ and has a discontinuity at $x = 0$ if and only if the set $\{t \in [0,1] : A_t' = 0\}$ has positive Lebesgue measure.

Theorem 2.2 [AW2]. — *Let I be a bounded interval in the real line. Then, almost surely,*

$$\sqrt{\frac{\pi\varepsilon}{2}}\frac{1}{\|\psi\|_2}\int_{-\infty}^{+\infty} f(u)\; N_u(M^\varepsilon, I)\; du \to \int_I f(M_t)\; A_t'^{1/2}\; dt \quad \text{as } \varepsilon \to 0 \quad (2.7)$$

One can check that exactly the same results hold true if instead of a martingale we have a semi-martingale with continuous paths, that is, a process $\{X(t) : t \geqslant 0\}$ such that $X(t) = M(t)+V(t)$, M is a local martingale as above and $\{V(t) : t \geqslant 0\}$ is an adapted process with continous paths and locally bounded variation. That is, the process V does not appear in this type of first order approximation of the normalized number of crossings of X^ε.

2.3. Lévy processes

One can extend (2.1) to Lévy processes. We will use the canonical representation for Lévy processes $\{X(t) : t \geqslant 0\}$ which will be assumed to have cadlag paths, as a sum of independent processes (see [G-S]):

$$X(t) = mt + \sigma W(t) + \int_{|x|\geqslant 1} x\nu_t(dx) + \int_{|x|<1} x\nu_t^*(dx) \qquad (2.8)$$

In formula (2.8) the ingredients are the following:

- The Lévy-Khinchin representation of the Fourier transform of the distribution of the random variable $X(t)$, which can be written, for $t \geqslant 0$, $s \in \mathcal{R}$, as:

$$E\left[\exp(izX(t))\right] = \exp\left[itmz - \frac{\sigma^2 z^2}{2}t + t\int_{-\infty}^{+\infty}(e^{izx} - 1 - izg(x))N(dx)\right]$$

where m and σ are real constants, $\sigma \geqslant 0$, $g(x) = x\ 1_{(-1,1)}(x)$ and $N(dx)$ is a Borel measure on the real line, $N(\{0\}) = 0$, $N(\{x : |x| \geqslant a\}) < \infty$ for every $a > 0$ and $\int_{|x|<1} x^2 N(dx) < \infty$.

For $x > 0$, we put:

$$|N|(x) = N(x,1) + N(-1,-x), \quad U(x) = \int_{|y|\leqslant x} y^2 N(dy), \quad V(x) = \frac{U(x)}{x^2}$$

- $\{W(t) : t \geqslant 0\}$ is a standard Wiener process,
- $\{\nu_t(dx) : t \geqslant 0\}$ is the Poisson measure of discontinuities, that is, for every Borel set B in the real line, $0 \notin B$, $\nu_t(B) = card\{s : 0 \leqslant s \leqslant t, X(s) - X(s^-) \in B\}$, $E(\nu_t(B)) = tN(B)$. If $B \subset \{x : |x| \geqslant a\}$ for some $a > 0$, then $\nu_t(B)$ has a Poisson distribution.
- $\nu_t^*(B) = \nu_t(B) - tN(B) \quad (t \geqslant 0)$.
- The families of random variables $\{W(t) : t \geqslant 0\}, \{\nu_t(B) : t \geqslant 0, B$ a Borel set in the line$\}$ are independent.
- The first integral in (2.8) is an ordinary Lebesgue integral, the second one is a Wiener integral with respect to the additive set function ν_t^*.

For simplicity we consider here only Lévy processes with symmetric one-dimensional distributions (for the general case see [W5]).

We have the following statement:

THEOREM 2.3. — *([W5] Assume that the law of X_1 is symmetric, that is $m = 0$ and $N(x,+\infty) = N(-\infty,-x)$ for every positive x. Then,*

(a) There exists a normalizing function a (i.e. satisfying (2.2)) such that

$$\textit{almost surely} \quad \lambda(\{t \in [0,1] : \frac{X(t+\varepsilon) - X(t)}{a(\varepsilon)} \leqslant x\}) \to \lambda^*(x) \quad \textit{as } \varepsilon \downarrow 0 \tag{2.9}$$

for some Borel probability distribution λ^* on the line, $\lambda^* \neq \delta_0$, and every continuity point x of λ^*, if and only if, one of the conditions in the first column of the table below holds true.

Necessary and sufficient conditions	$a(\varepsilon)$	*Log. Fourier of* λ^*
1. $\sigma > 0$	$\varepsilon^{1/2}$	$\frac{-\sigma^2 z^2}{2}$
2. $\sigma = 0$, *U slowly varying at* 0	$V^{-1}(\frac{1}{\varepsilon})$	$-\frac{z^2}{2}$
3. $\sigma = 0$, $N(x) = x^{-\rho} L(x)$, $0 < \rho < 2$	$\|N\|^{-1}(\frac{1}{\varepsilon})$	$-K_\rho \|z\|^\rho$

In the table, L is a slowly varying function at $x = 0$*, the inverse functions are defined by* $F^{-1}(x) = \inf\{t : F(t) \leqslant x\}$ *and* K_ρ *is the positive constant such that* $-K_\rho |z|^\rho$ *is the log-Fourier transform of the stable symmetric distribution with Lévy-Khinchin measure* $N_\rho(dx) = (\rho/2)\,|x|^{-\rho-1}\,dx$*, so that* $K_\rho = \frac{\rho}{2}\int_{-\infty}^{+\infty} \frac{1-\cos y}{|y|^{\rho+1}} dy$

(b) For each condition in the first column of the table, the second and third columns exhibit respectively a normalizing function and the logarithm of the Fourier transform of the corresponding limiting distribution.

(c) If a is a normalizing function satisfying (2.9), then any other normalizing function $\widetilde{a}$ *has the same property if and only if* $\widetilde{a}(\varepsilon)/a(\varepsilon)$ *has a finite non-zero limit as* $\varepsilon \downarrow 0$*. The corresponding limiting measure* $\widetilde{\lambda}^*$ *is obtained from* λ^* *by means of a change of scale. If there is a normalizing function satisfying (2.9) there is also a continuous one with the same property.*

(d) Let $g : [0,1] \times \Omega \to \mathcal{R}$ *be a measurable function and assume that for each* ω, $g(.,\omega) \in L^1([0,1],\lambda)$*. Put* $G(dt) = g(t,.)dt$*. Assume that the process verifies one of the conditions in the first column of the table.*

Then,

$$\text{almost surely} \quad G(\{t \in [0,1] : \frac{X(t+\varepsilon) - X(t)}{a(\varepsilon)} \leqslant x\}) \longrightarrow \lambda^*(x) \int_0^1 g(t,.)dt \quad \text{as } \varepsilon \downarrow 0 \tag{2.10}$$

where a stands for the corresponding function in the second column and λ^* *for the probability measure having log-Fourier transform in the third column.*

With respect to the first order approximation of the occupation measure by normalized crossings of the smoothed path, let us only mention the case of stable symmetric processes, for which a result of this type is contained in [A-W1], theorem 5.1.

2.4. Gaussian processes

First order approximation of the local time of stationary Gaussian processes using normalized crossings of smoothed paths has been considered in [A-F]. More general results are contained in [A-W1], from which we take as an example the special case of Gaussian processes with stationary increments, in which the statement is somewhat simpler than in the general case, even though the proofs are quite similar.

Our setting is the following:

- $\{X(t) : t \in \mathcal{R}\}$ is a Gaussian centered process with continuous paths, and covariance function $r(s,t) = E[X(s)X(t)]$. We assume that the process has stationary increments, i.e. $Var[X(t) - X(s)]$ depends only on $|t-s|$. Let $a(\varepsilon) = [Var(X(\varepsilon) - X(0))]^{1/2}$.
- We assume that:

  (i) the incremental standard deviation $a : (0,1) \to \mathcal{R}$ is regularly varying at $\varepsilon = 0$ with exponent α, $0 < \alpha < 1$, i.e. $a(\varepsilon) = \varepsilon^{\alpha} L(\varepsilon)$ where L is slowly varying at zero, i.e., $L(\varepsilon x)/L(x) \to 1$ as $x \downarrow 0$ for each fixed $\varepsilon > 0$.

  (ii) The covariance r is twice continuously differentiable outside the diagonal $t = s$ and satisfies the following regularly varying condition: there exists some $\eta > 0$ and a non-increasing slowly varying function at zero L_1 so that if $-\eta \leqslant s < t \leqslant 1 + \eta$, then

  $$\left| \frac{\partial^2 r(s,t)}{\partial s \partial t} \right| \leqslant (t-s)^{2\alpha - 2} L_1(t-s)$$

  where α is the same as in (i).

THEOREM 2.4. — *Let $\{X(t) : t \in \mathcal{R}\}$ satisfy the above conditions. Then,*

(i) almost surely, for every bounded interval I, as $\varepsilon \downarrow 0$,

$$\lambda\left(\left\{t \in I : \frac{\varepsilon.\, (X^{\varepsilon})'(t)}{a(\varepsilon)} \leqslant x\right\}\right) \to \lambda(I)\ \Phi\left(K_{\psi}^{-1} x\right) \quad \textit{for every } x \in \mathcal{R}$$

where Φ denotes the standard normal distribution function and

$$K_{\psi} = \left[-\frac{1}{2} \iint_{-1}^{1} |u - v|\ \ \psi(du)\psi(dv)\right]^{\frac{1}{2}} .$$

(ii) almost surely, for every continuous real-valued function f and every bounded interval I, as $\varepsilon \downarrow 0$,

$$\sqrt{\frac{\pi}{2}} \frac{\varepsilon}{K_{\psi} a(\varepsilon)} \int_{-\infty}^{+\infty} f(u)\ N_u(X^{\varepsilon}, I)\ du \to \int_I f[X(t)]\ dt.$$

3. Second order approximations. Speeds

3.1. Wiener process

We look at the speed of convergence in (2.3).

Let us put

$$\mathfrak{E}_\varepsilon(t) = \sqrt{\frac{\pi\varepsilon}{2}}\frac{1}{\|\psi\|_2}\int_{-\infty}^{+\infty} f(u)\; N_u(W^\varepsilon,[0,t])\; du - \int_{-\infty}^{+\infty} f(u)\; L^W(u,[0,t])\; du \tag{3.1}$$

where the notations are the same as above. Then,

Theorem 3.1. — *Assume that the function $f:\mathcal{R}\to\mathcal{R}$ is twice continuously differentiable and its second derivative is bounded.*

Then, as $\varepsilon\to 0$, the law of the stochastic process $\left\{\frac{1}{\sqrt{\varepsilon}}\mathfrak{E}_\varepsilon(t) : t\geqslant 0\right\}$ converges weakly in the space $C([0,+\infty),\mathcal{R})$ to the law of the process

$$D\int_0^t f(W(s))\; dB(s), \quad t\geqslant 0, \tag{3.2}$$

where

- *$\{B(s) : s\geqslant 0\}$ is a new Wiener process, independent of W,*
- *Conditionally on W the integral in (3.2) is an ordinary Wiener integral,*
- *the constant D depends only on the kernel ψ and is given by the explicit formula*

$$D^2 = 2\int_0^2 \left[r(t)\; Ar\sin[r(t)] + \sqrt{1-r^2(t)} - 1\right]\; dt \tag{3.3}$$

where $r(t)$ is the covariance function

$$r(t) = \frac{1}{\|\psi\|_2^2}\int_{-\infty}^{+\infty}\psi(t+u)\;\psi(u)\; du.$$

The first published proof ot Theorem 3.1 is based on Wiener chaos expansions and was given by Berzin and Leon in [B-L1]. See also [B-L-O] for extensions to stationary Gaussian processes.

A similar result to Theorem 3.1 holds true – mutatis mutandis – if instead of smoothing by convolution one uses polygonal approximation. In the latter case, the statement is valid if one replaces

- W^ε by $W^{(n)}$ the polygonal approximation of W with vertices $(\frac{k}{n}, W(\frac{k}{n}))$ $k = 0, 1, 2, ...$
- ε by $1/n$.
- the normalizing constant $\sqrt{\frac{\pi\varepsilon}{2}}\frac{1}{\|\psi\|_2}$ by $\sqrt{\frac{\pi}{2n}}$
- the constant D by $\sqrt{\frac{\pi}{2} - 1}$.

Jacod [8] has given a speed theorem using normalized crossings of polygonals to approximate the local time of the Wiener process at a fixed level u, that is, when one replaces the regular function f in the statement of the last Theorem by a Dirac δ. It is as follows:

Let the function $h(x, y)$ satisfy certain boundedness conditions and $\theta(h)$ be a suitable centering constant depending on h, then the stochastic process

$$n^{\frac{1}{4}}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]} h\left(\sqrt{n}\left(W(\frac{i-1}{n} - u\right), \sqrt{n}\left(W(\frac{i}{n}) - W(\frac{i-1}{n})\right)\right)\right.$$
$$\left. -\theta(h)\ L^W(u, [0, t])\right] \tag{3.4}$$

converges weakly to a Gaussian martingale conditionally on the given process W.

3.2. Semimartingales with continuous paths

Extensions of Theorem 3.1 to continuous semi-martingales have been given in [PW1] and [PW2]. Let us state the main result in [PW2] with some detail. Consider an Itô semi-martingale with values in $\mathcal{R}^d$, d a positive integer having the form:

$$X(t) = x_0 + \int_0^t a(s)\ dW(s) + V(t) \tag{3.5}$$

where

- $\{W(s) : s \geqslant 0\}$ is a standard Wiener process in $\mathcal{R}^d$. We denote $\mathcal{F} = \{\mathcal{F}_s : s \geqslant 0\}$ the filtration it generates.

- $a = \{a(s) : s \geqslant 0\}, V = \{V(s) : s \geqslant 0\}$ are stochastic processes adapted to $\mathcal{F}$, having continuous paths and values in the space of real $d \times d$ matrices and $\mathcal{R}^d$ respectively.

- $V(0) = 0$, so that $x_0 \in \mathcal{R}^d$ is a given – non-random – initial value.

We assume that a and V satisfy the following conditions:

1. Almost surely, the coordinates of the vector-valued random function $s \rightsquigarrow V(s)$ have locally bounded variation.

   a is strongly elliptic in each bounded interval I, which means that

   $$\inf_{s \in I} \inf_{\|x\|=1} \left\|a^T(s).x\right\| \geqslant C_I > 0$$

   where $\|.\|$ is Euclidean norm in $\mathcal{R}^d$, A^T denotes the transposed of matrix A and C_I is a non-random constant.

2. Let $a(s) = ((a^{jk}(s)))_{j,k=1,\ldots,d}$. We assume that for $s \geqslant 0$ and $0 < \varepsilon \leqslant \varepsilon_0$:

   $$\frac{a^{jk}(s+\varepsilon) - a^{jk}(s)}{\sqrt{\varepsilon}} = (\overline{a}^{jk}(s))^T \; Z^{jk}_{s,\varepsilon} + r^{jk}_{s,\varepsilon} \tag{3.6}$$

   where $\overline{a}^{jk}(s)$ and $Z^{jk}_{s,\varepsilon}$ are random vectors in $\mathcal{R}^d$ and $r^{jk}_{s,\varepsilon}$ is real-valued, such that:

- $\overline{a}^{jk}(s)$ is $\mathcal{F}_s$-mesurable, $Z^{jk}_{s,\varepsilon}$ and $r^{jk}_{s,\varepsilon}$ are $\mathcal{F}_{s+\varepsilon}$-mesurable and for each $p > 0$ the coordinates of $a(s)$, $\overline{a}^{jk}(s)$, $Z^{jk}_{s,\varepsilon}$ are uniformly bounded in $L^p(\Omega)$ for $0 < \varepsilon \leqslant \varepsilon_0$ and s in a bounded set.

- For each $p > 0$, $E(\left|r^{jk}_{s,\varepsilon}\right|^p) \to 0$ as $\varepsilon \downarrow 0$ uniformly for s in a bounded set.

- Denote for $t \geqslant 0, \varepsilon > 0$, $W^{\varepsilon,t}(u) = \frac{W(t+\varepsilon u) - W(t)}{\sqrt{\varepsilon}}$ $(u \geqslant 0)$ which is a new Wiener process. Then, we assume that almost every pair $(s,t), s \neq t$, the set $\left((Z^{jk}_{s,\varepsilon}, Z^{jk}_{t,\varepsilon})_{j,k=1,..,d}, W^{\varepsilon,s}_{.}, W^{\varepsilon,t}_{.}\right)$ converges weakly to the law of a random variable $\zeta(s,t)$ taking values in the appropriate space. This law is symmetric, independent of $\mathcal{F}_\infty$ and if $\{s,t\}$ and $\{s',t'\}$ are disjoint, then $\zeta(s,t)$ and $\zeta(s',t')$ are independent.

Some of these technical conditions appear to be complicated at first sight. However, they are satisfied in certain relevant cases:

1. Our first example is solutions of stochastic differential equations (SDE). Let the process $\{X(t) : t \geqslant 0\}$ be the strong solution of the SDE in $\mathcal{R}^d$:

$$dX(t) = \sigma(t, X(t))\ dW(t) + b(t, X(t))\ dt, \quad X(0) = x_0 \tag{3.7}$$

We assume that $\sigma(t, x) = ((\sigma^{jk}(t, x)))_{j,k=1,\ldots,d}$ and $b(t, x) = ((b^j(t, x)))_{j=1,\ldots,d}$ satisfy usual hypotheses such as Lipshitz local behaviour and degree one polynomial bound at ∞, and moreover, that σ is twice continuously differentiable with bounded second derivatives and $\left\|\sigma^T(s, x)v\right| \geqslant C_{s_0} \|v\|$ for some $C_{s_0} > 0$ and all $s \in [0, s_0]$, $x, v \in \mathcal{R}^d$. Then, one can show that the representation (3.6) holds true and verifies the conditions in the list above, with

$$\begin{aligned} (\overline{a}^{jk}(s))^T &= (D_x\sigma^{jk})(s, X(s))\ \sigma(s, X(s)) \\ Z^{jk}_{s,\varepsilon} &= \frac{W(s+\varepsilon) - W(s)}{\sqrt{\varepsilon}} = W^{\varepsilon,s}(1) \end{aligned}$$

2. Our second example is given by smoother integrands a. If for each $t \geqslant 0$ and $j, k = 1, ..., d$ one has for $\varepsilon > 0$:

$$\sup_{0 \leqslant s \leqslant t} \left|a^{jk}(s+\varepsilon) - a^{jk}(s)\right| \leqslant C\varepsilon^{\alpha(t)} \tag{3.8}$$

where C is a positive random variable having finite moments of all orders and $\alpha(t) > 1/2$, then the conditions are satisfied with $\overline{a}^{jk}(s) = 0$.

3. In dimension $d = 1$, put in (3.5) $a(s) = f(W(s))$, $V(s) = 0$ and choose $f : \mathcal{R} \to \mathcal{R}$, $f(x) = 1 + \beta g(x)$, g of class C^3, non-negative with compact support, $g''(0) \neq 0$ and $\beta > 0$ small enough so that $\beta\ \sup_{x \in \mathcal{R}} |g''(x)| < 2$.

One can check that our conditions are satisfied with $\overline{a}(s) = f'(W(s))$, $Z_{s,\varepsilon} = W^{\varepsilon,s}(1)$. However, the process $\{X(t) : t \geqslant 0\}$ is a non-Markovian continuous martingale (for a proof, see [N-W], 4.2.) with respect to any natural filtration and hence, it can not be the solution of a Markovian SDE. Obviously it does not satisfy (3.8) either.

Besides the continuous semimartingale (3.5) the other ingredient in the next Theorem is smoothing of paths. Since we are considering vector-valued processes, we need to introduce some slight changes in notation and some additional requirements with respect to one-dimensional smoothing.

In what follows, $\psi(x) = ((\psi^{jk}(x)))_{j,k=1,\ldots,d}$ is a deterministic matrix kernel, each function $\psi^{jk}(x)$ being C^∞ real-valued of one real variable, support contained in the interval $[-1,1]$, $\int_{-\infty}^{+\infty}\psi(x)\,dx=((\int_{-\infty}^{+\infty}\psi^{jk}(x)\,dx))_{j,k=1,\dots,d}$ $=I_d=$ identity matrix $d\times d$. We put $\psi_\varepsilon(x)\;=\varepsilon^{-1}\psi(\varepsilon^{-1}x)$.

We also add the following technical condition on the smoothing kernel. Denote $\underline{\lambda}(x)$ (resp. $\overline{\lambda}(x)$) the minimal (resp. maximal) eigenvalue of $\psi(x)\psi^T(x)$. We assume that there exists a positive constant L such that $\overline{\lambda}(x)\leqslant L\;\underline{\lambda}(x)$ for all $x\in\mathcal{R}$. This condition plays some role only if $d>1$ and limits the anisotropy that is allowed for the regularization mechanism.

Define the smoothed path by

$$X^\varepsilon(t)=\int_{-\infty}^{+\infty}\psi_\varepsilon(t-s)\;X(s)\;ds \tag{3.9}$$

where in (3.9) we put $X(s)=X(0)$ if $s<0$.

Finally, let $f:\mathcal{R}^d\to\mathcal{R}$ be of class C^2 with bounded second derivatives and $g:\mathcal{R}^+\to\mathcal{R}$ be of class C^2, $|g'(r)|\leqslant C_g(1+r^m)$ for some $m\geqslant 1$, positive constant C_g and all $r\in\mathcal{R}^+$.

We consider the observable functional defined on the smoothed path:

$$\theta_{\varepsilon,\tau}=\theta_{\varepsilon,\tau}(f,g)=\int_0^\tau f\left[X^\varepsilon(t)\right]\;g\left(\left\|\sqrt{\varepsilon}(X^\varepsilon)'(t)\right\|\right)\;dt \tag{3.10}$$

Now, we are prepared to state the following Central Limit Theorem for $\theta_{\varepsilon,\tau}$:

THEOREM 3.2. — *With the hypotheses above, as $\varepsilon\downarrow 0$:*

$$\frac{1}{\sqrt{\varepsilon}}\left[\theta_{\varepsilon,\tau}-\int_0^\tau f\left(X(t)\right)\;E\left(g(\left\|\Sigma_t^{1/2}\xi\right\|/\mathcal{F}_\infty)\right)\,dt\right]\Longrightarrow B_{\overline{\sigma}^2(\tau)} \tag{3.11}$$

where

- $\Longrightarrow$ *denotes weak convergence of probability measures in* $C\left([0,+\infty),\mathcal{R}\right)$,
- B *denotes a new one-dimensional Wiener process, independent of* $\mathcal{F}_\infty$.
- $\Sigma_t=\int_{-1}^1\psi(x)\;a(t)a^T(t)\;\psi^T(x)\;dx.$
- ξ *is a standard normal random variable in* $\mathcal{R}^d$, *independent of* $\mathcal{F}_\infty$.

- $$\overline{\sigma}^2(\tau) = \int_0^\tau f^2\left(X(t)\right) dt \iint_{-1}^{1} F(t,v,\overline{v})\ dv d\overline{v} \tag{3.12}$$

where

$F(t,v,\overline{v}) =$

$E\left[g'(\|\eta_{t,v}\|)g'(\left\|\overline{\eta}_{t,\overline{v}}\right\|)sg(\eta_{t,v}^T)\psi(v)\ a(t)a^T(t)\ \psi^T(\overline{v})sg(\overline{\eta}_{t,\overline{v}}^T)/\mathcal{F}_\infty\right]$

and the conditional distribution of the pair of $\mathcal{R}^d$*-valued random variables* $\eta_{t,v}, \overline{\eta}_{t,\overline{v}}$ *given* $\mathcal{F}_\infty$ *is centered Gaussian with covariance structure:*

$$\begin{aligned} E\left(\eta_{t,v}\eta_{t,v}^T/\mathcal{F}_\infty\right) &= E\left(\overline{\eta}_{t,\overline{v}}\overline{\eta}_{t,\overline{v}}^T/\mathcal{F}_\infty\right) = \Sigma_t, \\ E\left(\eta_{t,v}\overline{\eta}_{t,\overline{v}}^T/\mathcal{F}_\infty\right) &= \int_{-1}^{v\wedge\overline{v}} \psi(-x)\ a(t)a^T(t)\ \psi^T(-x+|\overline{v}-v|)\ dx. \end{aligned}$$

3.2.1. Remarks on the Theorem. Related results

1. To make more clear the statement of Theorem 3.2 let us consider two examples of interesting functionals.
    - (Normalized curve length). Let $g(r) = r$ and assume that f is a C^2-approximation of the indicator function of a set $B \subset \mathcal{R}^d$ having a sectionally smooth boundary. Then, $\theta_{\varepsilon,\tau}$ is an approximation of $\sqrt{\varepsilon}\mathcal{L}_\varepsilon(\tau,B)$, $\mathcal{L}_\varepsilon(\tau,B)$ denoting the length of the part of the curve $t \rightsquigarrow X^\varepsilon(t)$ contained in the observation window B. Generally speaking, in the situations of interest, $\mathcal{L}_\varepsilon(\tau,B) \to +\infty$ as $\varepsilon \downarrow 0$, $\sqrt{\varepsilon}$ is the normalization for first order approximation and $\sqrt{\varepsilon}$ is the order of the speed.
    - (Normalized kinetic energy). Let $g(r) = r^2$ and f as in the previous example.Then, $\theta_{\varepsilon,\tau}$ becomes an approximation of $\varepsilon K_\varepsilon(\tau,B)$, $K_\varepsilon(\tau,B)$ denoting the kinetic energy of the same part of the smoothed path.

2. The above statement can be compared with results based upon polygonal approximations. Estimation methods of the diffusion coefficient in a SDE seem to start with the pionner work by Dacunha-Castelle and Florens [D-F]. See also [B]. In this direction, the following statement is proved in [F].

    Let $\{X(t) : t \geqslant 0\}$ be a strong solution of the one-dimensional SDE

    $$dX(t) = \sigma(X(t))dW(t) + b(X(t))dt, \quad X(0) = x_0 \tag{3.13}$$

where σ and b satisfy certain regularity conditions, and $k \leqslant \sigma(x) \leqslant K$ for all $x \in \mathcal{R}$, k, K some positive constants. Take a sequence of positive numbers $\{h_n\}$ – the window size in space – such that $nh_n \to +\infty$, $nh_n^3 \to 0$ and consider the estimator of $\sigma^2(x)$:

$$S_n(x) = \frac{\sum_{i=1}^{n-1} 1_{\{|X(\frac{i}{n})-x|<h_n\}} n\left(X(\frac{i+1}{n}) - X(\frac{i}{n})\right)^2}{\sum_{i=1}^{n-1} 1_{\{|X(\frac{i}{n})-x|<h_n\}}}$$

whenever the denominator does not vanish. Denote T_x the waiting time defined by $T_x = \inf\{t \in [0,1] : X(t) = x\}$ if $\{.\}$ is non empty and $T_x = 1$ if $\{.\}$ is empty. Then conditionally on $T_x < 1$,

$$\sqrt{nh_n}\left[\frac{S_n(x)}{\sigma^2(x)} - 1\right]$$

converges in distribution to

$$\left[L^X(x,[0,1])\right]^{-1/2}.\xi$$

where L^X stands for local time, ξ is standard normal and both factors are independent random variables.

On the other hand, Jacod has extended (3.4) to diffusions satisfying certain conditions and used this extension to make non-parametric inference on σ [J2]. See also [H] and references therein and [G-J], [G-J-L].

Theorem 3.2 points towards two kinds of different problems, that did not seem to be solved for diffusions before: a) Testing hypotheses on the diffusion function σ beyond considering a fixed argument x; b) Instead of discrete sampling, using functionals defined on regularizations of the underlying path.

At the same time, integrating in the state space without compressing the window to a single point, permits to improve the speed of convergence, passing from $n^{-1/4}$ for fixed level to $n^{-1/2}$ for the integrated result. These is somewhat similar to what occurs in other statistical problems, in which integrating a function instead of putting a Dirac δ gives as a by-product a better speed. Notice that the presence of the local time, as we pointed out above implies serious inconvenients if one is willing to make statistical inference.

3. A nice property of Theorem 3.2 is that the drift part V does not appear explicity, either in the centering term or in the asymptotic probability law. Of course, it is hidden in the process X. This becomes useful to make inference on the martingale part, i.e. on a.

We will see later on that this is not the case if V has jumps, a possibility that we have excluded for the moment.

4. In fact, the actual situation is more complicated. Consider the special important case in which X is a solution of a SDE with $a(t) = \sigma(t, X(t))$. Theorem 3.2 as it has been stated, can't be used for statistical applications when the observation is the smoothed path X^ε, since the unobserved underlying path X appears in the statement, both in the centering term and in the asymptotic variance $\overline{\sigma}^2(\tau)$.

   In the variance $\overline{\sigma}^2(\tau)$ this is not a serious problem, since it is clear from (3.12) that if one replaces $X(t)$ by $X^\varepsilon(t)$, this expression is continuous as a function of ε and the asymptotic law is close to its limit. This is less obvious for the centering term in (3.11) since one must divide by $\sqrt{\varepsilon}$, but it holds true under certain additional conditions given in the next Theorem.

THEOREM 3.3. — *Let us consider the process $\{X(t) : t \geqslant 0\}$ solution of the SDE (3.7) satisfying the hypotheses of Example 1. We assume further that σ and b are only functions of space i.e., $\sigma(t, x) = \sigma(x)$, $b(t, x) = b(x)$ and that the smoothing kernel is isotropic, that is, $\psi(x) = \psi^*(x)\ I_d$, where ψ^* is real-valued.*

Then, as $\varepsilon \downarrow 0$:

(i)

$$\frac{1}{\sqrt{\varepsilon}}\left[\theta_{\varepsilon,\tau} - m_\varepsilon(\tau)\right] \Longrightarrow B_{\overline{\sigma}^2(\tau)} \tag{3.14}$$

where:

- $m_\varepsilon(\tau) = \int_0^\tau f\left(X^\varepsilon(t)\right)\ E\left(g\left[\|\psi^*\|_2 \left\|\sigma^T\left(X^\varepsilon(t)\right)\xi\right\|\right]/\mathcal{F}_\infty\right)\ dt,$
- *the symbols $\theta_{\varepsilon,\tau}$, $\Longrightarrow$, ξ , B, $\overline{\sigma}^2(\tau)$ are as in Theorem 3.2.*

(ii) a.s. $\overline{\sigma}^2_\varepsilon(\tau) \to \overline{\sigma}^2(\tau)$ where

$$\overline{\sigma}^2_\varepsilon(\tau) = \int_0^\tau f^2\left(X^\varepsilon(t)\right) dt \iint_{-1}^{1} \psi^*(-v)\psi^*(-\overline{v}).$$
$$.E\left[g'(\|\eta_{t,v}\|)g'(\left\|\overline{\eta}_{t,\overline{v}}\right\|)sg(\eta^T_{t,v})\sigma(X^\varepsilon(t))\sigma^T(X^\varepsilon(t))sg(\overline{\eta}^T_{t,\overline{v}})/\mathcal{F}_\infty\right] dvd\overline{v}$$

and for the conditional distribution of the pair of $\mathcal{R}^d$-valued random variables $\eta_{t,v}, \overline{\eta}_{t,\overline{v}}$ given $\mathcal{F}_\infty$, one must change in the statement of Theorem 3.2

Σ_t by $\|\psi^*\|^2\,\sigma(X^\varepsilon(t))\sigma^T(X^\varepsilon(t))$ and $E\left(\eta_{t,v}\overline{\eta}^T_{t,\overline{v}}/\mathcal{F}_\infty\right)$ by $K(v,\overline{v})\sigma(X^\varepsilon(t))\sigma^T(X^\varepsilon(t))$ with

$$K(v,\overline{v}) = \int_{-1}^{v\wedge\overline{v}} \psi^*(-x)\ \psi^T(-x+|\overline{v}-v|)\ dx. \tag{3.15}$$

3.2.2. Some statistical examples

1. **Hypothesis testing for σ.**

   We consider the case $d=1$.

   Let X be the solution of the SDE (3.13) and assume the conditions of Theorem 3.3 are satisfied. We put $g(r)=r$ and assume $\inf_{x\in\mathcal{R}}\sigma(x) > 0$.

   Suppose we want to test the null hypothesis

   $H_0 : \sigma(x) = \sigma_0(x)$ for all $x \in \mathcal{R}$,

   against the local alternative

   $H_\varepsilon : \sigma_\varepsilon(x) = \sigma_0(x) + \sqrt{\varepsilon}\sigma_1(x) + \gamma_\varepsilon(x)$ for all $x\in\mathcal{R}$,

   where $\gamma_\varepsilon(x) = o(\sqrt{\varepsilon})$ and $\gamma'_\varepsilon(x) = o(\sqrt{\varepsilon})$ as $\varepsilon\downarrow 0$ uniformly on $x\in\mathcal{R}$. $\sigma_0(.)$, $\sigma_1(.)$ and $\gamma_\varepsilon(.)$ are functions with continuous and bounded second derivatives and at most degree one polynomial growth at ∞.

   The application of Theorem 3.3 is not straightforward under these conditions, since under the alternative, the solution process depends also on ε. If we denote it by $X(.;\varepsilon)$, then:

   $$dX(t;\varepsilon) = \sigma_\varepsilon(X(t;\varepsilon))\ dW(t) + b(X(t;\varepsilon))\ dt, \quad X(0;\varepsilon) = x_0.$$

   Let us put $X^\varepsilon(t) = (\psi_\varepsilon * X(.,\varepsilon))(t)$.

   One can prove that under H_ε, as $\varepsilon\downarrow 0$, one has:

$$\frac{1}{\sqrt{\varepsilon}}\left[\sqrt{\frac{\pi\varepsilon}{2}}\frac{1}{\|\psi\|_2}\int_{-\infty}^{+\infty} f(u)\ N_u(X^\varepsilon,[0,\tau])\ du - \int_0^\tau f(X^\varepsilon(t))\ \sigma_0(X^\varepsilon(t))\ dt\right]$$
$$\approx \int_0^\tau f(X^\varepsilon(t))\ \sigma_1(X^\varepsilon(t))\ dt + \sqrt{\frac{\pi}{2}}\frac{1}{\|\psi\|_2}B_{\overline{\sigma}^2_\varepsilon(\tau)}$$

   One should interpret this result as follows: As $\varepsilon\downarrow 0$ the law of the left-hand side converges in $C([0,+\infty),\mathcal{R})$ to the law of the random process $\int_0^\tau f(X(t))\ \sigma_1(X(t))\ dt + \sqrt{\frac{\pi}{2}}\frac{1}{\|\psi\|_2}B_{\overline{\sigma}^2_0(\tau)}$ and furthermore, the right-hand side converges to this process almost surely as $\varepsilon\downarrow 0$.

Notice that in case our test is based upon the observation of X^ε, this is well-adapted to the statistical purpose, since both the centering term in the left-hand member and the right-hand member can be computed from the hypotheses and from functionals defined on $X^\varepsilon(t), 0 \leqslant t \leqslant \tau$.

2. $d > 1$. **Testing isotropy of the noise part.**

   Assume now that $d > 1$ and the SDE (3.13) satisfies the hypotheses of Theorem 3.3. We put again $g(r) = r$.

   We want to test the null hypothesis of isotropy of the noise, that is

   $H_0 : \Sigma(x) = \sigma(x)\sigma^T(x) = I_d$ for all $x \in \mathcal{R}^d$

   against the alternative

   $H_\varepsilon : \Sigma(x) = I_d + \sqrt{\varepsilon}\Sigma_1(x) + \Gamma_\varepsilon(x)$ for all $x \in \mathcal{R}$

   where $\|\Gamma_\varepsilon(x)\| = o(\sqrt{\varepsilon})$ and $\|\Gamma'_\varepsilon(x)\| = o(\sqrt{\varepsilon})$ as $\varepsilon \downarrow 0$ uniformly on $x \in \mathcal{R}$.

   Here, $\Sigma_1(x)$ and $\Gamma_\varepsilon(x)$ are positive semi-definite $d \times d$ real matrix with elements that are twice continuously differentiable functions with bounded second derivatives and at most degree two polynomial growth at ∞. The norm is any norm on $d \times d$ matrices.

   In this case, the statistical result takes the form

$$\frac{1}{\sqrt{\varepsilon}}\left[\int_0^\tau f(X^\varepsilon(t)) \; \left[\left\|\sqrt{\varepsilon}\,(X^\varepsilon)'(t))\right\| - \|\psi^*\|_2\, J_d\right] \; dt\right] \qquad (3.16)$$
$$\approx \frac{J_d}{2d}\,\|\psi^*\|_2 \int_0^\tau f(X^\varepsilon(t))\; tr(\Sigma_1(X^\varepsilon(t)))\; dt + B_{\overline{\sigma}^2_\varepsilon(\tau)}$$

   where the interpretation of the sign $\approx$ is the same as in the previous example and:

   - $J_d = \frac{(2p)!}{(2^p p!)^2} p\sqrt{8\pi}$ if $d = 2p$, $J_d = \frac{(2^p p!)^2}{(2p)!}\frac{1}{\sqrt{2\pi}}$ if $d = 2p+1$.
   - $\overline{\sigma}^2_\varepsilon(\tau) = \int_0^\tau f^2(X^\varepsilon(t))\; dt \iint_{-1}^1 \psi^*(-v)\psi^*(-\overline{v})\; A_d(\widetilde{K}(v,\overline{v}))\; dv d\overline{v}$

   where $\widetilde{K}(v,\overline{v}) = \frac{K(v,\overline{v})}{\|\psi^*\|_2^2}$, $K(v,\overline{v})$ the same function (3.15) as in the previous example and the function $A_d : [-1,1] \to \mathcal{R}$ is defined by

$$A_d(\rho) = E\left[\frac{\langle \xi, \eta\rangle}{\|\xi\|\,\|\eta\|}\right]$$

   where the pair ξ, η of random vectors has a joint Gaussian, centered distribution in $\mathcal{R}^d \times \mathcal{R}^d$, each one of them is standard normal in $\mathcal{R}^d$ and $E\left(\xi\eta^T\right) = \rho\, I_d$.

   One can see from (3.16) that a good choice for the function f in a test against alternative H_ε is $f(x) = tr(\Sigma_1(x))$ in which case the asymptotic bias under H_ε is $\frac{J_d}{2d}\,\|\psi^*\|_2 \int_0^\tau \left[tr(\Sigma_1(X^\varepsilon(t)))\right]^2\; dt$.

3. **Simple regression**.

Let us consider the simple regression model in continuous time

$$X(t) = m(t) + \sigma\ W(t), \quad t \geqslant 0, \tag{3.17}$$

where $m : \mathcal{R}^+ \to \mathcal{R}$ is a continuous non-random function having locally bounded variation, σ is a positive constant and W stands for Wiener process. Notice that there are no jumps.

If one wants to make inference on the value of σ, the usual approach is to assume that one can measure X on a grid, say at the points $\left\{\frac{k}{n} : k=0,1,...,n\right\}$ of the unit inteval $[0,1]$ and use the consistency of the estimator

$$\hat{\sigma}_n^2 = \sum_{i=1}^{n} \left[X(\frac{i}{n}) - X(\frac{i-1}{n})\right]^2$$

which converges in probability to the true value σ^2.

Of course, the next step is to compute the speed of convergence, that is to find a sequence $\{a_n\}$ of positive numbers tending to $+\infty$ such that

$$a_n \left[\hat{\sigma}_n^2 - \sigma^2\right] \tag{3.18}$$

converges in distribution to some non-trivial limit.

The problem here is that the existence of such a sequence $\{a_n\}$ depends on the regularity of the drift function m. For example, if m is absolutely continuous with respect to Lebesgue measure and has a bounded Radon-Nikodym derivative m', then it is an elementary fact that $a_n = \sqrt{n}$ will do the job and the limit distribution of (3.18) is centered normal with variance equal to $2\sigma^4$. However, this is not true in general.

For example, if $m(t) = t^\alpha$ with $0 < \alpha < \frac{1}{4}$, one can easily check that there is no normalizing sequence $\{a_n\}$ such that (3.18) converges in distribution, and this example can be used as a basis to construct a set of other simple ones, in which the same phenomenon takes place. Notice that in this example, the function m is not only of bounded variation, but also absolutely continous with an unbounded derivative.

In other words, using the observation of X on a grid to make inference on σ in this way, requires m to have certain regularity properties that may not be verified. In a given situation one may not be able to decide if it is reasonable to assume that the unknown function m verifies these regularity conditions. This appears to be a difficulty to use this kind of statistics.

On the other hand, let us assume that we observe a regularization of X of the type described before the statement of Theorem 3.2. In our case, the statement of this Theorem becomes

$$\frac{1}{\sqrt{\varepsilon}}\left[\theta_{\varepsilon,\tau}-E\ \left[g(|\sigma\xi|)\right]\int_0^\tau f\left(X(t)\right)\ dt\right]\Longrightarrow B_{\overline{\sigma}^2(\tau)} \tag{3.19}$$

with

$$\overline{\sigma}^2(\tau)=\sigma^2\int_0^\tau f^2\left(X(t)\right)dt\iint_{-1}^{1}E\left[g'(|\eta_{t,v}|)g'(|\overline{\eta}_{t,\overline{v}}|)sg(\eta^T_{t,v})sg(\overline{\eta}^T_{t,\overline{v}})\right]\psi(v)\psi(\overline{v})dvd\overline{v} \tag{3.20}$$

where the joint distribution of $\eta_{t,v},\overline{\eta}_{t,\overline{v}}$ is centered Gaussian, $Var(\eta_{t,v})=Var(\overline{\eta}_{t,\overline{v}})=\sigma^2\left\|\psi\right\|_2^2$, $Cov(\eta_{t,v},\overline{\eta}_{t,\overline{v}})=\sigma^2K(v,\overline{v})$.

This is not yet adequate for statistical use, one still needs to replace $X(t)$ by $X^\varepsilon(t)$ in (3.19) and in (3.20). In what follows we prove that this can be done, so that it is possible to use the method based on the observation of the smoothed path without restrictions on the function m for estimation or hypothesis testing on σ in the model (3.17).

Again it is clear that this is not a problem for the variance (3.20) and that replacing $X(t)$ by $X^\varepsilon(t)$ we obtain $\overline{\sigma}^2_\varepsilon(\tau)\approx\overline{\sigma}^2(\tau)$.

Let us turn to (3.19). We need to prove that

PROPOSITION 3.4. — *Let* $0\leqslant\tau\leqslant1$. *Then*:

$$\frac{1}{\sqrt{\varepsilon}}\int_0^\tau\left[f\left(X^\varepsilon(t)\right)-f\left(X(t)\right)\right]\ dt \tag{3.21}$$

tends to zero in probability, as $\varepsilon\downarrow0$.

Proof. — It is clear that replacing $f(X(t))$ by $f(X(t-\varepsilon))$ in (3.21) for almost every ω the error is $O(\sqrt{\varepsilon})$. Using a Taylor expansion for f :

$$\frac{1}{\sqrt{\varepsilon}}\int_0^\tau\left[f\left(X^\varepsilon(t)\right)-f\left(X(t-\varepsilon)\right)\right]\ dt=T_\varepsilon(\tau)+R_\varepsilon(\tau) \tag{3.22}$$

where

$$T_\varepsilon(\tau)=\frac{1}{\sqrt{\varepsilon}}\int_0^\tau f'\left(X(t-\varepsilon)\right)\left[X^\varepsilon(t)-X(t-\varepsilon)\right]\ dt=\frac{1}{\sqrt{\varepsilon}}\int_0^\tau f_\varepsilon(t)\ dt$$

$$|R_\varepsilon(\tau)|\leqslant\frac{\|f''\|_\infty}{2\sqrt{\varepsilon}}\int_0^\tau\left[X^\varepsilon(t)-X(t-\varepsilon)\right]^2dt$$

We prove that $T_\varepsilon(\tau)$ and $R_\varepsilon(\tau)$ both tend to zero in probability as $\varepsilon \downarrow 0$.

For $R_\varepsilon(\tau)$ one has:

$$E\left(|R_\varepsilon(\tau)|\right) \leqslant \frac{(const)}{\sqrt{\varepsilon}} \left[\int_0^\tau \left[m^\varepsilon(t) - m(t-\varepsilon)\right]^2 dt \right.$$
$$\left. + \int_0^\tau E\left(\left[W^\varepsilon(t) - W(t-\varepsilon)\right]^2\right) dt\right] \tag{3.23}$$

Let us show that the first term in brackets in the right-hand member of (3.23) is $O(\varepsilon)$. Since m has bounded variation in any bounded interval, it suffices to prove this when m is non-decreasing. In this case:

$$\begin{aligned} \int_0^\tau \left[m^\varepsilon(t) - m(t-\varepsilon)\right]^2 dt &\leqslant m(1+\varepsilon) \int_0^\tau \left[m^\varepsilon(t) - m(t-\varepsilon)\right] \ dt \\ &= m(1+\varepsilon) \int_0^\tau dt \int_{-1}^1 \psi(u) \left[m(t-\varepsilon u) - m(t-\varepsilon)\right] \ du \\ &\leqslant m(1+\varepsilon) \int_{-1}^1 \psi(u) \ du \int_0^\tau \left[m(t+\varepsilon) - m(t-\varepsilon)\right] \ dt = O(\varepsilon) \end{aligned}$$

For the second term in (3.23), one easily checks that $E([W^\varepsilon(t) - W(t-\varepsilon)]^2) \leqslant (const)\varepsilon$. So, $E\left(|R_\varepsilon(\tau)|\right) = O(\sqrt{\varepsilon})$.

Let us consider now the first term in (3.22). We have:

$$\begin{aligned} E\left(T_\varepsilon^2(\tau)\right) &= \frac{1}{\varepsilon} \iint_0^\tau E\left(f_\varepsilon(s) f_\varepsilon(t)\right) \ dsdt \\ &= \frac{2}{\varepsilon} \left[\iint_{s \leqslant t < s+2\varepsilon} E\left(f_\varepsilon(s) f_\varepsilon(t)\right) \ dsdt + \iint_{t \geqslant s+2\varepsilon} E\left(f_\varepsilon(s) f_\varepsilon(t)\right) \ dsdt\right] \\ &= \frac{2}{\varepsilon} \left[I_\varepsilon + J_\varepsilon\right] \end{aligned}$$

For I_ε:

$$|I_\varepsilon| \leqslant \iint_{s \leqslant t < s+2\varepsilon} \left[E\left(f_\varepsilon^2(s)\right) E(f_\varepsilon^2(t))\right]^{\frac{1}{2}} dsdt = o(\varepsilon) \text{ as } \varepsilon \downarrow 0.$$

For J_ε, if $t \geqslant s + 2\varepsilon$ denoting with $\{\mathcal{F}_t\}_{t \geqslant 0}$ the filtration generated by the process, we have:

$$\begin{aligned} E\left(f_\varepsilon(s) f_\varepsilon(t)\right) &= E\left(E\left(f_\varepsilon(s) f_\varepsilon(t) / \mathcal{F}_{t-\varepsilon}\right)\right) \\ &= E\left(f'\left(X(s-\varepsilon)\right)\right. \\ &\quad \left.\left[X^\varepsilon(s) - X(s-\varepsilon)\right] f'\left(X(t-\varepsilon)\right) E\left(X^\varepsilon(t) - X(t-\varepsilon) / \mathcal{F}_{t-\varepsilon}\right)\right) \\ &= \left[m^\varepsilon(t) - m(t-\varepsilon)\right] \\ &\quad E\left(f'\left(X(s-\varepsilon)\right) \left[X^\varepsilon(s) - X(s-\varepsilon)\right] f'\left(X(t-\varepsilon)\right)\right) \end{aligned}$$

which implies that

$$|J_\varepsilon| \leqslant (const) \int_0^\tau E\left(\left[X^\varepsilon(s) - X(s-\varepsilon)\right]^2\right)^{\frac{1}{2}} ds \int_0^\tau |m^\varepsilon(t) - m(t-\varepsilon)| \; dt$$
$$= o(\varepsilon).$$

This ends the proof. □

3.3. Lévy processes with non-vanishing Gaussian part

The results we present on speeds of convergence for Lévy processes are partial, they concern only some special cases. We use without additional reference, the notation in the previous sections.

We say that the Lévy process satisfies condition (V) if

$$\int_{|x|<1} |x| \, N(dx) < \infty \tag{3.24}$$

which implies that the jump part of the process has locally finite variation. Our first theorem is on the speed of convergence in (2.9).

THEOREM 3.5. — *Let $\{X(t) : t \geqslant 0\}$ be a Lévy process with parameters m, σ and N. Assume that condition (V) is satisfied and that $\sigma > 0$. Let g be any real-valued continuously differentiable function, bounded and with a bounded derivative and*

$$Y_\varepsilon(t) = \frac{1}{\sqrt{\varepsilon}} \left[\int_0^t g\left(\frac{X(s+\varepsilon) - X(s)}{\sqrt{\varepsilon}}\right) ds - tE\left(g\left(\sigma\xi\right)\right)\right] \tag{3.25}$$

Then,

$$(Y_\varepsilon, X) \Longrightarrow (D_g B, X) \quad as \quad \varepsilon \downarrow 0$$

where

- $\Longrightarrow$ *denotes weak convergence in the product space $C([0,+\infty), \mathcal{R}) \times D([0,+\infty), \mathcal{R})$ of the space of continuous functions times the Skorokhod space,*
- *B is a new standard Wiener process, independent of X,*
- *ξ is a standard normal random variable,*

- *the constant D_g is given by the formula*

$$D_g^2 = 2\int_0^1 \left[E\left(g\left(\sigma\xi_1\right)g\left(\sigma\xi_2\right)\right) - \left[E\left(g\left(\sigma\xi\right)\right)\right]^2\right]\, d\rho$$

in which the pair of random variables $\xi_1,\ \xi_2$ has joint normal distribution, centered and $Var(\xi_1) = Var(\xi_2) = 1, Cov(\xi_1,\ \xi_2) = \rho$.

Corollary 3.6. — *With the same hypotheses of the last theorem and the previous notations for smoothing, if f is twice continously differentiable with bounded second derivative, one has*

$$\frac{1}{\sqrt{\varepsilon}}\left[\int_0^t f(X^\varepsilon(s))g\left(\sqrt{\varepsilon}\,|(X^\varepsilon)'(s)|\right)ds - E\left(g(\sigma\,\|\psi\|_2^{-1}\,|\xi|)\right)\int_0^t f(X(s))ds\right]$$

$$\Longrightarrow D_{g,\psi}\int_0^t f(X(s))dB(s) \tag{3.26}$$

as $\varepsilon \downarrow 0$.

In (3.26) $D_{g,\psi}^2 = \int_{-\infty}^{+\infty}\overline{g}\left(|U_0|\right)\overline{g}\left(|U_u|\right)du$, where $\overline{g}(x) = g(\sigma x) - E(g\left(\sigma\xi\right))$ and U_u is a stationary centered Gaussian process with covariance

$$r(u) = \|\psi\|_2^{-2}\int_{-\infty}^{+\infty}\psi(u-y)\psi(-y)\ dy. \tag{3.27}$$

In the corollary we may replace again $X(s)$ by $X^\varepsilon(s)$ and obtain approximate results based on the observation of the smoothed path. However, both the theorem and its corollary do not include the special case $g(r) = r$ which in one dimension corresponds to observing crossings of the smoothed path. In the next theorem we consider this case, for which the proof is somewhat more complicated (see [MW2]) than the one of Theorem 3.5.

Theorem 3.7. — *Let $\{X(t) : t \geqslant 0\}$ be a Lévy process with parameters m, σ and N. Assume that conditon (V) is satisfied and $\sigma > 0$.*

Then, for each function f, twice continuously differentiable with bounded second derivatives, as $\varepsilon \downarrow 0$,

$$\frac{1}{\sqrt{\varepsilon}}\left[\sqrt{\frac{\pi\varepsilon}{2}}\frac{1}{\|\psi\|_2}\int_{-\infty}^{+\infty} f(u)\ N_u(X^\varepsilon, [0,t])\ du - \sigma\int_0^t f(X(s))ds\right]$$

$$-\sum_{0<s\leqslant t} L(f,s)\,|\Delta X(s)|$$

$$\overset{(c)}{\Longrightarrow} D_\psi\int_0^t f(X(s))dB(s) \tag{3.28}$$

where:

- $\overset{(c)}{\Longrightarrow}$ *means cylindric convergence,*
- $L(f,s) =$
$$\sqrt{\frac{\pi}{2}}\frac{1}{\|\psi\|_2}\int_{-1}^{1}\psi(z)\ f\left(X(s^-)\int_z^1\psi(w)dw + X(s)\int_{-1}^{z}\psi(w)dw\right)dz,$$
- $\Delta X(s) = X(s) - X(s^-)$,
- B *is a Wiener process, independent of* X,
-
$$D_\psi^2 = 2\sigma^2\int_0^2\left[r(u)Ar\sin r(u) + \sqrt{1-r^2(u)} - 1\right]\ du$$
where $r(u)$ *is given by (3.27).*

Remark 3.8. — One can not expect weak convergence in the Skorokhod space in the statement of Theorem 3.7 if the jump part of the process does not vanish, excepting for trivial functions f. If this convergence would hold true, the limit should be the right-hand member of (3.28). But if the jump part of X does not vanish, for a generic f the bias term $\sum_{0<s\leqslant t}L(f,s)|\Delta X(s)|$ in the left hand member of (3.28) has non-vanishing jumps with positive probability, so that there can not be weak convergence to a process with continuous paths. Notice also that $\sum_{0<s\leqslant t} L(f,s)\,|\Delta X(s)|$ term can be passed to the right-hand member and look at the result as a speed result for the approximation of the occupation measure by normalized crossings. The bias in the limit measure is due to the presence of jumps.

On the other hand, one can obtain weak convergence replacing the bias term by $\sqrt{\frac{\pi}{2}}\frac{1}{\|\psi\|_2}\int_0^t f(X^\varepsilon(s))\left|(S^\varepsilon)'(s)\right|\ ds$ where S is the jump part of the process X and S^ε is its smooth approximation by convolution. Even though this holds in a stronger topology, it looks less interesting from the point of view of the interpretation of the bias term.

Theorem 3.7 has the following counterpart in polygonal approximation.

Theorem 3.9. — *Assume the same hypotheses of Theorem 3.7 are satisfied, and let us consider the polygonal approximation* $X^{(n)}$ *of the paths of the process* X, *as defined above. Then, as* $n \to +\infty$,

$$\sqrt{n}\left[\sqrt{\frac{\pi}{2n}}\int_{-\infty}^{+\infty} f(u)\ N_u(X^{(n)},[0,t])\ du - \sigma\int_0^t f(X(s))ds\right] - Z(t)$$
$$\overset{(c)}{\Longrightarrow} D\int_0^t f(X(s))dB(s) \qquad (3.29)$$

where

- $D^2 = \sigma^2(\frac{\pi}{2} - 1), (c)$
- $Z(t) = \sqrt{\frac{\pi}{2}}\frac{1}{2}\sum_{0<s\leqslant t}\left[f(X(s)) + f(X(s^-))\right] \ \ |\Delta X(s)|.$

3.4. Pure jump Lévy processes

Let X be a Lévy process with vanishing Gaussian component, i.e. $\sigma = 0$. We assume again, for simplicity, that it has symmetric one-dimensional distributions, so that $m = 0$ and $N(dx)$ is even.

In this section we sketch the calculations that lead to certain second order results, only in some special cases. One can get some more general approximations using similar methods. However, as far as this author knows, no general result is available.

We assume that the process satisfies condition 3. in the table included in the statement of Theorem 3.5, i.e. regular variation at zero of the function N, so that the corresponding conclusions hold true. We use the same notations as in Theorem 3.5.

The general scheme is as follows: let $g : \mathcal{R} \to \mathcal{R}$ be continuous and bounded, and ξ a random variable having the limit distribution λ^*. Denote $m_g(\varepsilon) = E\left[g\left(\frac{X_\varepsilon}{a(\varepsilon)}\right)\right]$. One can show the following CLT.

As $\varepsilon \downarrow 0$,

$$\frac{1}{\sqrt{\varepsilon}}\left[\int_0^t g\left[\frac{X(s+\varepsilon) - X(s)}{a(\varepsilon)}\right] ds - tE\left(g\left(\xi\right)\right)\right] - t\ H(\varepsilon) \Longrightarrow D_{g,\rho}B_t \quad (3.30)$$

where

- $\Longrightarrow$ denotes weak convergence in the space $C([0,+\infty), \mathcal{R})$,
- the random variable ξ has distribution λ^*,
- B is a Wiener process independent of X,
- $H(\varepsilon) = \frac{1}{\sqrt{\varepsilon}}\left[m_g(\varepsilon) - E\left(g\left(\xi\right)\right)\right],$
- the constant $D_{g,\rho}$ is given by formula

$$D^2_{g,\rho} = 2\int_0^1 \left[G_\rho(\theta, g) - \left[E\left(g\left(\xi\right)\right)\right]^2\right]\ d\theta, \qquad (3.31)$$

$$G_\rho(\theta, g) = E\left[g\left(\theta^{\frac{1}{\rho}}\xi_1 + (1-\theta)^{\frac{1}{\rho}}\xi_2\right) g\left((1-\theta)^{\frac{1}{\rho}}\xi_2 + \theta^{\frac{1}{\rho}}\xi_3\right)\right]$$

where ξ_1, ξ_2, ξ_3 are independent random variables, each one of them with distribution λ^*.

The proof of (3.30) is standard. One proves tightness using fourth moments and the independence of increments of X. A variance computation gives (3.31) and the remainder is plain.

The only point that remains in order to have a useful result, is the behaviour of $H(\varepsilon)$ as $\varepsilon \downarrow 0$. Denote by φ_ε the Fourier transform of the probability distribution of $\frac{X_\varepsilon}{a(\varepsilon)}$. Then,

$$\begin{aligned} H(\varepsilon) &= \frac{1}{\sqrt{\varepsilon}}\int_{-\infty}^{+\infty} g(x)dx \frac{1}{2\pi}\int_{-\infty}^{+\infty}\exp(-ixz)\left(\varphi_\varepsilon(z) - e^{-K_\rho|z|^\rho}\right) dz \quad (3.32)\\ &= \int_{-\infty}^{+\infty} g(x)dx \frac{1}{2\pi}\int_{-\infty}^{+\infty} \exp(-ixz - K_\rho\,|z|^\rho)\ \gamma_\rho(\varepsilon, z)\ dz + o(1) \end{aligned}$$

where

$$\gamma_\rho(\varepsilon, z) = \frac{1}{\sqrt{\varepsilon}}\left\{\exp\left[K_\rho\,|z|^\rho - \frac{\varepsilon}{a^2(\varepsilon)}z^2\int_{|x|<1} G(x)\ \exp(\frac{izx}{a(\varepsilon)})\ dx\right] - 1\right\} \quad (3.33)$$

with $G(x) = \int_x^1 N(y,1)dy$.

In case the process X is ρ-stable, and $X(1)$ has probability distribution λ^*, one can show that

$$\gamma_\rho(\varepsilon, z) = \frac{1}{\sqrt{\varepsilon}}\left\{\exp\left[-\frac{\rho}{2}\,|z|^\rho\int_{|y|\geqslant\frac{|z|}{\varepsilon^{1/\rho}}}\frac{1-\cos y}{|y|^{\rho+1}}dy\right] - 1\right\}$$

which implies $|\gamma_\rho(\varepsilon, z)| \leqslant \varepsilon^{1/2}$ if ε is small enough. So, in this case, if also $g \in L^1(\mathcal{R}, dx)$, it follows from (3.32) that $H(\varepsilon) \to 0$ as $\varepsilon \downarrow 0$, and we have proved that (3.30) holds true on replacing $H(\varepsilon)$ by zero.

The general case in which the function $N(x,1)$ is ρ-regularly varying at zero is more complicated. One can prove, with the same hypotheses on the function g and using similar calculations, the following result.

Assume that $N(x,1) = x^{-\rho}L(x)$ where $L(x) = L(0) + C(x)\ x^\alpha$ for $0 < x \leqslant 1$, $L(0) > 0$, C, C' are bounded functions, $C(0^+) \neq 0$, $\alpha \geqslant 0$.

Then, the behaviour of $H(\varepsilon)$ is given by:

- If $\alpha > \frac{1}{2}\rho$, then $H(\varepsilon) \to 0$.

- If $\alpha < \frac{1}{2}\rho$, then $H(\varepsilon) \to \infty$, and more precisely, $H(\varepsilon) \approx (const)\, \varepsilon^{\frac{\alpha}{\rho}-\frac{1}{2}}$ as $\varepsilon \downarrow 0$, the constant factor depending on the function g.

- If $\alpha = \frac{1}{2}\rho$, then $H(\varepsilon) \to K_{\rho,g,C(0^+)}$. This limit can be computed by the formula:

$$K_{\rho,g,C(0^+)} = C(0^+) \int_0^{+\infty} \frac{\sin y}{y^{\frac{\rho}{2}}} dy \int_{-\infty}^{+\infty} g(x) dx \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(-ixz - K_\rho \left|z\right|^\rho) z^\rho dz.$$

Remarks. —

1. In the last section, we have considered second order approximation for Lévy processes with no Gaussian component only for the increments of the Lévy process, in other words, only in the case that the convolution kernel is $\psi(x) = 1_{[-1,0]}(x)$. With some extrawork, it is possible to obtain similar theorems for a general convolution kernel, under the same hypotheses, when instead of $\frac{X(s+\varepsilon)-X(s)}{a(\varepsilon)}$ one considers $\frac{\varepsilon}{a(\varepsilon)}(X^\varepsilon)'(s)$. The results have the same form, excepting that one has to multiply the stable random variables ξ, ξ_1, ξ_2, ξ_3 appearing in the above formulae by the constant $\left\|\psi\right\|_\rho$.

2. We will not pursue here the subject of statistical applications of the above results in the case of Lévy processes. It is clear from the statements that they may be used for this purpose.

As examples, theorems (3.7) and (3.9) fit well to certain inference problems, such as testing the hypothesis that are no jumps or estimation of σ in the presence of jumps. Also, notice that if one knows that the Lévy process has no Gaussian part and the Lévy measure is regularly varying at zero with exponent ρ, $0 < \rho < 2$, the results in the last section show, even in a restricted framework, that one can make inference from the smoothed path on parameters appearing in the slowly varying function L.

Bibliography

[A1] Azaïs (J.-M.). — Conditions for convergence of number of crossings to the local time, Applications to stable processes with independent increments and to Gaussian processes, Probab. Math. Stat., 11, 1 (1990).

[A2] Azaïs (J.-M.). — Approximation des trajectoires et temps local des diffusions, Ann. Inst. H. Poincaré, B, 25, (2) p. 175-194 (1989).

[A-F] Azaïs (J.-M.), Florens-Szmirou (D.). — Approximation du temps local des processus gaussiens stationnaires par régularisation des trajectoires, Probab. Th. Rel. Fields, 76, p. 121-132 (1987).

[A-W1] Azaïs (J.-M.), Wschebor (M.). — Almost sure oscillation of certain random processes, Bernoulli, Vol. 2, Fasc. 3, p. 257-270 (1996).

[A-W2] Azaïs (J.-M.), Wschebor (M.). — Oscillation presque sûre de martingales continues, Séminaires de Probabilités XXXI, Eds. J. Azéma, M. Emery, M. Yor, Lecture Notes Math., Vol. 1655, Springer-Verlag, p. 69-76 (1997).

[B-L1] Berzin (C.), Leon (J.R.). — Weak convergence of the integrated number of level crossings to the local time of the Wiener process, Comptes R. Acad. Sc. Paris, Sér. I, 319, p. 1311-1316 (1994).

[B-L-O] Berzin (C.), Leon (J.R.), Ortega (J.). — Level crossings and local time for regularized Gaussian processes, Probab. Math. Statist. 18, No. 1, p. 39-81 (1998).

[B-W] Berzin (C.), Wschebor (M.). — Approximation du temps local des surfaces gaussiennes, Probab. Th. Rel. Fields, 96, p. 1-32 (1993).

[B-I] Borodin (A.N.), Ibragimov (I.A.). — Limit theorems for functionals of random walks, Proc. Steklov Institute Math., 195, AMS, Providence, RI (1995).

[B] Brugière (P.). — Estimation de la variance d'un processus de diffusion dans le cas multidimensionel, Comptes R. Acad. Sc. Paris, Sér. I, 312, p. 999-1004 (1991).

[C-R1] Csörgö (M.), Révész (P.). — Three strong approximations of the local time of a Wiener process and their applications to invariance. In: Coll. Math. Soc. J. Bolyai, 36 (1984). Limit Theorems in Probability and Statistics, Veszprém, p. 223-254 (1982).

[C-R2] Csörgö (M.), Révészv (P.). — On strong invariance for local time of partial sums, Stoch. Proc. Appl., 20, p. 59-84 (1985).

[D-F] Dacunha-Castelle (D.), Florens-Zmirou (D.). — Estimation of the coefficient of a diffusion from discrete observations, Stochastics, 19, p. 263-284 (1986).

[F-T] Fristedt (B.), Taylor (S.J.). — Constructions of local time for a Markov process. Z. Wahr.verw. gebiete, 62, p. 73-112 (1983).

[F] Florens-Zmirou (D.). — On estimating the diffusion coefficient from discrete observations, J. Appl. Prob., 30, p. 790-804 (1993).

[G-J] Génon-Catalot (V.), Jacod (J.). — On the estimation of the diffusion coefficient for multidimensional diffusion processes, Ann. Inst. H. Poincaré, Prob. Stat. 29, p. 119-151 (1993).

[G-S] Guikhman (I.), Skorokhod (A.). — Introduction à la théorie des processus aléatoires (MIR, Moscow) (1980).

[G-J] Génon-Catalot (V.), Jacod (J.). — On the estimation of the diffusion coefficient for multidimensional diffusion processes, Ann. Inst. H. Poincaré, Prob. Stat., 29, p. 119-151 (1993).

[G-J-L] Génon-Catalot (V.), Jeantheau (T.), Laredo (C.). — Limit theorems for discretely observed stochastic volatility models, Bernoulli, 4, No. 3, p. 283-304 (1998).

[H] Hoffmann (M.). — L^p estimation of the diffusion coefficient, Bernoulli, 5, No. 3, p. 447-481 (1999).

[I-W] Ikeda (N.), Watanabe (S.). — Stochastic Differential Equations and Diffusion Processes, North Holland (1982).

[I-M] Itô (K.), Mc Kean (H.P.). — Diffusion processes and their sample paths, Academic Press (1965).

[J] Jacod (J.). — Rates of convergence to the local time of a diffusion, Ann. Inst. H. Poincaré, Prob. Stat, 34, 505-544 (1998).

[J] JACOD (J.). — Non-parametric kernel estimation of the diffusion coefficient of a diffusion, Scand. J. Statist., 27, 1, p. 83-96 (2000).

[K-S] KARATZAS (I.). — Shreeve, Brownian motion and stochastic calculus, Springer-Verlag (1998).

[L-S] LIPSTER (R.S.), Shiryaev (A.N.). — Statistics of Random Processes, Vol. I, II. 2d ed. Springer-Verlag (2001).

[M-W1] MORDECKI (E.), WSCHEBOR (M.). — Smoothing of paths and weak approximation of the occupation measure of Lévy processes, Pub. Mat. Uruguay, to appear (2005).

[M-W2] MORDECKI (E.), WSCHEBOR (M.). — Approximation of the occupation measure of Lévy processes, Comptes Rendus de l'Académie des Sciences, Paris, Sér. I, 340, 605-610 (2005).

[N-W] NUALART (D.), WSCHEBOR (M.). — Intégration par parties dans l'espace de Wiener et approximation du temps local, Probab. Th. Rel. Fields, 90, p. 83-109 (1991).

[P-W1] PERERA (G.), WSCHEBOR (M.). — Crossings and occupation measures for a class of semimartingales, Ann. Probab., Vol. 26, N 1, p. 253-266 (1998).

[P-W2] PERERA (G.), WSCHEBOR (M.). — Inference on the Variance and Smoothing of the Paths of Diffusions, Ann. Inst. H. Poincaré, Prob. Stat. 38, 6, p. 1009-1022 (2002).

[PR] PRAKASA RAO (B.L.S). — Semimartingales and their Statistical Inference. Chapman & Hall (1999).

[R] RÉVÉSZ (P.). — Local time and invariance, Lecture Notes in Math. No. 861, Springer-Verlag, p. 128-145 (1981).

[W1] WSCHEBOR (M.). — Régularisation des trajectoires et approximation du temps local, C.R. Acad. Sci. Paris, Sér. I, p. 209-212 (1984).

[W2] WSCHEBOR (M.). — Surfaces aléatoires. Mesure géometrique des ensembles de niveau. Lecture Notes Math., Vol. 1147, Springer-Verlag (1985).

[W3] WSCHEBOR (M.). — Crossings and local times of one-dimensional diffusions, Pub. Mat. Uruguay, N 3, p. 69-100 (1990).

[W4] WSCHEBOR (M.). — Sur les accroissements du processus de Wiener, C. R. Acad. Sci. Paris, t. 315, Sér. I, p. 1293-1296 (1992).

[W5] WSCHEBOR (M.). — Almost sure weak convergence of the increments of Lévy processes, Stoch. Proc. Appl., 55, p. 253-270 (1995).