Ito Integrals

10

10.1 Summary

Kiyoshi Ito (1915–2008) was awarded the inaugural Gauss Prize by the International Mathematical Union in 2006. Stochastic integration in the narrow sense can be traced back to his early work published in Japanese in the forties of the last century. We precede the general definition of the Ito integral with a special case. Concluding, we discuss the (quadratic) variation of a process without which a sound understanding of Ito's lemma will not be possible.

10.2 A Special Case

We start with a special case of Ito integration, so to speak the mother of all stochastic integrals. Thereby we will understand that, besides the Ito integral, infinitely many related integrals of a similar structure exist.

Problems with the Definition

The starting point is again a partition

$$P_n\left([0,t]\right): \quad 0 = s_0 < s_1 < \ldots < s_n = t\,,$$

that gets finer for n growing since we continue to maintain (8.2). Given this decomposition of $[0,t]$, we define analogously to the Riemann-Stieltjes sum for $s_i^* \in [s_{i-1}, s_i)$:

$$S_n(W) = \sum_{i=1}^{n} W(s_i^*)\left(W(s_i) - W(s_{i-1})\right). \tag{10.1}$$

U. Hassler, *Stochastic Processes and Calculus*, Springer Texts in Business and Economics, DOI 10.1007/978-3-319-23428-1_10

For $n \to \infty$ we would like to denote the limit as $\int_0^t W(s)\,dW(s)$, which looks like a Stieltjes integral of a WP with respect to a WP. However, we will realize that:

1. The limit of $S_n(W)$ is not unique, but depends on the choice of s_i^*;
2. the limit of $S_n(W)$ is not defined as a Stieltjes integral.

As the Stieltjes integral has a unique limit independently of s_i^*, the second claim follows from the first one. If one chooses in particular the lower endpoint of the interval as support, $s_i^* = s_{i-1}$, then this leads to the Ito integral. Hence, this special case is called the **Ito sum**:

$$I_n(W) = \sum_{i=1}^{n} W(s_{i-1})\,(W(s_i) - W(s_{i-1}))\,. \tag{10.2}$$

The following proposition specifies the dependence on s_i^*. The special case $\gamma = 0$ leading to the Ito integral will be proved in Problem 10.1; the general result is established e.g. in Tanaka (1996, eq. (2.40)). The convergence is again in mean square.

Proposition 10.1 (Stochastic Integrals in Mean Square) *Let $s_i^* = (1-\gamma)\,s_{i-1} + \gamma\,s_i$ with $0 \leq \gamma < 1$. Then it holds for the sum from (10.1) with $n \to \infty$ under (8.2):*

$$S_n(W) \stackrel{2}{\to} \frac{1}{2}\left(W^2(t) - t\right) + \gamma\,t\,.$$

Before we discuss two special cases of Proposition 10.1, this striking result is to be somewhat better understood. We call it striking because it is counter-intuitive at first glance that the choice of s_i^* should matter with the intervals $[s_{i-1}, s_i)$ getting narrower and narrower for $n \to \infty$. To better understand this, we temporarily denote the limit of $S_n(W)$ as $S(\gamma)$:

$$S_n(W) \stackrel{2}{\to} S(\gamma)\,.$$

Then one observes immediately:

$$S(\gamma) = S(0) + \gamma\,t\,.$$

This means that the variance of all these stochastic integrals $S(\gamma)$ is identical, i.e. equal to the variance of $S(0)$. Hence, the choice of different support points s_i^* is only reflected in the expected value:

$$\begin{aligned}\mathrm{E}\left(S(\gamma)\right) &= \frac{1}{2}\left(\mathrm{E}(W^2(t)) - t\right) + \gamma\, t \\ &= \gamma\, t\,.\end{aligned}$$

However, this expected value can be well understood as for finite sums it can be shown that (see Problem 10.4):

$$\mathrm{E}\left(S_n(W)\right) = \gamma\, t\,.$$

This simply follows from the fact that $W(s_i^*)$ is not independent of $W(s_i) - W(s_{i-1})$ for $\gamma > 0$. Next, we turn towards the case $\gamma = 0$.

Ito Integral

For $\gamma = 0$, $S_n(W)$ from Proposition 10.1 merges into $I_n(W)$ from (10.2). The proposition guarantees two different things: First, that the limit of $I_n(W)$ converges in mean square. We call this limit the **Ito integral** and write instead of $S(0)$ the following integral:

$$I_n(W) \xrightarrow{2} \int_0^t W(s)\, dW(s)\,.$$

Secondly, the proposition yields an expression for this Ito integral:

$$\int_0^t W(s)\, dW(s) = \frac{1}{2}\, W^2(t) - \frac{1}{2}\, t\,. \tag{10.3}$$

By the way, (10.3) is just the "stochastified chain rule" for Wiener processes from (1.14).[1] Note the analogy and the contrast to the deterministic case (with $f(0) = 0$):

$$\int_0^t f(s)\, df(s) = \frac{1}{2} f^2(t) \quad \text{for } f(0) = 0\,. \tag{10.4}$$

In Eq. (10.3) we find, so to speak, the archetype of Ito calculus, i.e. of stochastic calculus using Ito's lemma. The latter will be covered in the next chapter.

[1] In particular for $t = 1$, (10.3) accomplishes the transition from (1.11) to (1.12) for the Dickey-Fuller distribution.

The moments of the Ito integral can be determined by (10.3). According to this equation it holds for the expected value:

$$\mathrm{E}\left(\int_0^t W(s)\,dW(s)\right) = 0\,.$$

The variance of the integral as well can be calculated elementarily[2]:

$$\begin{aligned}
\mathrm{Var}\left(\int_0^t W(s)\,dW(s)\right) &= \mathrm{Var}\left(\frac{W^2(t)}{2} - \frac{t}{2}\right)\\
&= \mathrm{E}\left[\left(\frac{W^2(t)}{2} - \frac{t}{2}\right)^2\right]\\
&= \frac{1}{4}\,\mathrm{E}\left(W^4(t) - 2\,t\,W^2(t) + t^2\right)\\
&= \frac{1}{4}\,\left(3\,t^2 - 2\,t^2 + t^2\right)\\
&= \frac{t^2}{2}\,,
\end{aligned}$$

where the kurtosis of 3 for Gaussian random variables was used. Hence, we have the first two results of the following proposition

Proposition 10.2 (Moments of $\int_0^t W(s)\,dW(s)$) *For $I(t) = \int_0^t W(s)\,dW(s)$ it holds that*

$$E(I(t)) = 0 \quad \textit{and} \quad Var\,(I(t)) = \frac{t^2}{2}\,,$$

and

$$E(I(t)\,I(t+h)) = \frac{t^2}{2} \quad \textit{for } h \geq 0\,.$$

---

[2] An alternative, interesting method uses the fact that the variance of a chi-squared distributed random variable equals twice its degrees of freedom:

$$\mathrm{Var}\left(\frac{W^2(t)}{2} - \frac{t}{2}\right) = \frac{1}{4}\,\mathrm{Var}(W^2(t)) = \frac{t^2}{4}\,\mathrm{Var}\left(\left(\frac{W(t)}{\sqrt{t}}\right)^2\right) = \frac{t^2}{4}\cdot 2 = \frac{t^2}{2}\,,$$

as it holds that $W(t)/\sqrt{t} \sim \mathcal{N}(0,1)$ and therefore

$$\left(\frac{W(t)}{\sqrt{t}}\right)^2 \sim \chi^2(1)\,.$$

As a third result it holds in Proposition 10.2, just as for the Stieltjes integral, that the autocovariance coincides with the variance, i.e.

$$\mathrm{E}(I(t)\, I(t+h)) = \mathrm{Var}\,(I(t)) \ , \quad h \geq 0.$$

We will prove this in Problem 10.2.

Stratonovich Integral

For a reason that will soon become evident, sometimes a considered competitor of the Ito integral is the **Stratonovich integral**. It is defined as the limit of $S_n(W)$ from (10.1) with the midpoints of the intervals as s_i^*:

$$s_i^* = \frac{s_{i-1} + s_i}{2} \, .$$

This corresponds to the choice of $\gamma = 0.5$ in Proposition 10.1. Let the limit in mean square be denoted as follows:

$$\sum_{i=1}^{n} W\left(\frac{s_{i-1} + s_i}{2}\right) (W(s_i) - W(s_{i-1})) \stackrel{2}{\rightarrow} \int_0^t W(s)\, \partial W(s) \, ,$$

where "∂" does not stand for the partial derivative but denotes the Stratonovich integral in contrast to the Ito integral. By the way, with $\gamma = 0.5$ Proposition 10.1 yields:

$$\int_0^t W(s)\, \partial W(s) = \frac{W^2(t)}{2} \, .$$

Hence, the Stratonovich integral stands out due to the fact that the familiar integration rule known from ordinary calculus holds true. In differential notation this rule can be formulated symbolically as follows:

$$\frac{\partial W^2(t)}{2} = W(t)\, \partial W(t) \, .$$

This just corresponds to the ordinary chain rule, cf. (10.4). Although the Ito and the Stratonovich integral are distinguished from each other only by the choice of s_i^* with intervals getting shorter and shorter, they still have drastically different properties. Obviously, it holds for the expected value

$$\mathrm{E}\left(\int_0^t W(s)\, \partial W(s)\right) = \frac{t}{2} \, ,$$

while the Ito integral is zero on average. However, as aforementioned, the variances of $\int W(s)dW(s)$ and $\int W(s)\partial W(s)$ coincide, cf. Problem 10.3 as well.

Example 10.1 (Alternative Stratonovich Sum) Sometimes the Stratonovich integral is defined as the limit of the following sum, see e.g. Klebaner (2005, eq. (5.65)):

$$\sum_{i=1}^{n} \frac{W(s_{i-1}) + W(s_i)}{2} \, (W(s_i) - W(s_{i-1})) \ .$$

The intuition behind this is, due to the continuity of the WP, that

$$W(s_{i-1}) \approx W\left(\frac{s_{i-1} + s_i}{2}\right) \approx W(s_i)\,.$$

In fact, it can be shown more explicitly that the following difference becomes negligible in mean square:

$$\Gamma := W\left(\frac{s_{i-1} + s_i}{2}\right) - \frac{W(s_{i-1}) + W(s_i)}{2} \overset{2}{\rightarrow} 0\,.$$

For this purpose we consider as the mean square deviation with $s_i^* = (s_{i-1} + s_i)/2$:

$$\begin{aligned}
\mathrm{MSE}(\Gamma, 0) &= \mathrm{E}\left[(\Gamma - 0)^2\right] \\
&= \mathrm{E}\left[W^2(s_i^*) - W(s_i^*)\,(W(s_{i-1}) + W(s_i))\right] \\
&\quad + \mathrm{E}\left[\frac{W^2(s_{i-1}) + 2\,W(s_{i-1})W(s_i) + W^2(s_i)}{4}\right].
\end{aligned}$$

Due to $s_{i-1} < s_i^* < s_i$ the familiar variance and covariance formulas yield:

$$\begin{aligned}
\mathrm{MSE}(\Gamma, 0) &= s_i^* - s_{i-1} - s_i^* + \frac{s_{i-1} + 2\,s_{i-1} + s_i}{4} \\
&= \frac{s_i - s_{i-1}}{4}.
\end{aligned}$$

As for $n \to \infty$ the partition gets finer and finer, $s_i - s_{i-1} \to 0$, the replacement of $W(s_i^*)$ by $(W(s_{i-1}) + W(s_i))\,/2$ is asymptotically well justified. ■

10.3 General Ito Integrals

After covering general Ito integrals, we define so-called diffusions that we will be concerned with in the following chapters.

Definition and Moments

In order to define general Ito integrals, we consider for a stochastic process X as a generalization of the sum $I_n(W)$:

$$I_n(X) = \sum_{i=1}^{n} X(s_{i-1})\ (W(s_i) - W(s_{i-1}))\ . \tag{10.5}$$

For the Ito integral two special things apply: First, the lower endpoint of the interval is chosen as $s_i^* = s_{i-1}$, i.e. $X(s_{i-1})$; and secondly, we integrate with respect to the WP, $(W(s_i) - W(s_{i-1}))$. If X was integrated with respect to another stochastic process, then one would obtain even more general stochastic integrals, which we are not interested in here.

If $X(t)$ is a process with finite variance where the variance varies continuously in the course of time, and if $X(t)$ only depends on the past of the WP, $W(s)$ with $s \leq t$, but not on its future, then the Ito sum converges uniquely and independently of the partition. The limit is called Ito integral and is denoted as follows:

$$\int_0^t X(s)\, dW(s)\,.$$

The assumptions about $X(t)$ are stronger than necessary, however, they guarantee the existence of the moments of an Ito integral, too. Similar assumptions can be found in Klebaner (2005, Theorem 4.3) or Øksendal (2003, Corollary 3.1.7).

Proposition 10.3 (General Ito Integral) *Let $X(s)$ be a stochastic process on $[0, t]$ with two properties:*

(i) $\mu_2(s) = E\left(X^2(s)\right) < \infty$ is a continuous function,
(ii) $X(s)$ is independent of $W(s_j) - W(s_i)$ with $s \leq s_i < s_j$.

Then it holds that

(a) the sum from (10.5) converges in mean square:

$$\sum_{i=1}^{n} X(s_{i-1})\,(W(s_i) - W(s_{i-1})) \xrightarrow{2} \int_0^t X(s)\, dW(s)\,;$$

(b) the moments of the Ito integral are determined as:

$$E\left(\int_0^t X(s)\, dW(s)\right) = 0\,, \quad Var\left(\int_0^t X(s)\, dW(s)\right) = \int_0^t E\left(X^2(s)\right)\, ds\,.$$

Naturally, for $X(s) = W(s)$ the extensively discussed example from the previous section is obtained. In particular, in (b) the moments from Proposition 10.2 are reproduced, and it holds for the variance that:

$$\operatorname{Var}\left(\int_0^t W(s)\,dW(s)\right) = \int_0^t \mathrm{E}(W^2(s))\,ds = \int_0^t s\,ds = \frac{t^2}{2}.$$

Example 10.2 (Stieltjes Integral) Consider the special case where $X(s)$ is not stochastic but deterministic,

$$X(s) = f(s),$$

where $f(s)$ is continuous. Then, the conditions of existence are fulfilled which can easily be verified: The square, $\mu_2(s) = f^2(s)$, is continuous as well, and the deterministic function is independent of $W(s)$. Hence, it holds that

$$\sum_{i=1}^{n} f(s_{i-1})\ (W(s_i) - W(s_{i-1})) \xrightarrow{2} \int_0^t f(s)\,dW(s).$$

In other words: For deterministic processes, $X(s) = f(s)$, the Stieltjes and the Ito integral coincide; the former is a special case of the latter. Due to $\mathrm{E}\left(f^2(s)\right) = f^2(s)$ the already familiar formulas for expectation and variance from Proposition 9.2 are embedded in the general Proposition 10.3. ■

Distribution and Further Properties

As is well known, the special case of the Stieltjes integral is Gaussian. For the Ito integral this does not hold in general. This can clearly be seen in (10.3):

$$\int_0^t W(s)\,dW(s) = \frac{W^2(t) - t}{2} \geq \frac{-t}{2},$$

i.e. the support of the distribution is bounded. Then again, the integral of a WP with respect to a thereof stochastically independent WP amounts to a Gaussian distribution. The following result is by Phillips and Park (1988).

Proposition 10.4 (Ito Integral of an Independent WP) *Let $W(t)$ and $V(s)$ be stochastically independent Wiener processes. Then it holds that*

$$\left(\int_0^1 V^2(s)\,ds\right)^{-0.5} \int_0^1 V(s)\,dW(s) \sim \mathcal{N}(0,1).$$

By showing the conditional distribution of the left-hand side given $V(t)$ to just follow a $\mathcal{N}(0, 1)$ distribution and therefore not to depend on this condition, one proves the result claimed in Proposition 10.4; see Phillips and Park (1988, Appendix) for details.

Note that the Ito integral again defines a stochastic process in the bounds from 0 to t whose properties could be discussed, which we will not do at this point. In the literature, however, it can be looked up that the Ito integral and the Wiener process share the continuity and the martingale properties. What is more, the integration rules outlined at the end of Sect. 8.2 hold true for Ito integrals as well, see e.g. Klebaner (2005, Thm. 4.3).

Diffusions

For economic modeling the Ito integral is an important ingredient. However, it gains its true importance only when combined with Riemann integrals. In the following chapters, the sum of both integrals constitutes so-called **diffusions**[3] (diffusion processes). Hence, we now define processes $X(t)$ (with starting value $X(0)$) as follows:

$$X(t) = X(0) + \int_0^t \mu(s)\, ds + \int_0^t \sigma(s)\, dW(s)\,.$$

Frequently, we will write this integral equation in differential form as follows:

$$dX(t) = \mu(t)\, dt + \sigma(t)\, dW(t)\,.$$

The conditions set to $\mu(s)$ and $\sigma(s)$ that guarantee the existence of such processes can be adopted from Propositions 8.1 and 10.3. In general, $\mu(s)$ and $\sigma(s)$ are stochastic; particularly, they are allowed to be dependent on $X(s)$ itself. Therefore, we write $\mu(s)$ and $\sigma(s)$ as abbreviations for functions which firstly explicitly depend on time and secondly depend on X simultaneously:

$$\mu(s) = \mu\left(s, X(s)\right)\,, \quad \sigma(s) = \sigma\left(s, X(s)\right)\,.$$

Processes μ and σ satisfying these conditions are used to define diffusions $X(t)$:

$$dX(t) = \mu\left(t, X(t)\right)\, dt + \sigma\left(t, X(t)\right)\, dW(t)\,, \quad t \in [0, T]\,. \tag{10.6}$$

[3]The name stems from molecular physics, where diffusions are used to model the change of location of a molecule due to a deterministic component (drift) and an erratic (stochastic) component. Physically, the influence of temperature on the motion hides behind the stochastics: The higher the temperature of the matter in which the particles move, the more erratic is their behavior.

Recall that this differential equation actually means the following:

$$X(t) = X(0) + \int_0^t \mu\,(s, X(s))\; ds + \int_0^t \sigma\,(s, X(s))\; dW(s)\,.$$

Example 10.3 (Brownian Motion with Drift) We consider the Brownian motion with drift and a starting value 0:

$$\begin{aligned} X(t) &= \mu t + \sigma\, W(t) \\ &= \mu \int_0^t ds + \sigma \int_0^t dW(s)\,. \end{aligned}$$

Therefore, the differential notation reads

$$dX(t) = \mu\, dt + \sigma\, dW(t)\,.$$

Hence, this is a diffusion whose drift and volatility are constant:

$$\mu\,(t, X(t)) = \mu \quad \text{and} \quad \sigma\,(t, X(t)) = \sigma\,. \quad \blacksquare$$

10.4 (Quadratic) Variation

From (10.3) we know that

$$\int_0^t W(s)\, dW(s) = \frac{W^2(t) - t}{2}\,.$$

Now, we want to understand where the expression t comes from that is subtracted from $W^2(t)$. It will be made clear that this is the so-called quadratic variation.

(Absolute) Variation

Again, the considerations are based on an adequate partition of the interval $[0, t]$,

$$P_n\,([0, t])\,:\; 0 = s_0 < s_1 < \ldots < s_n = t\,.$$

For a function g the **variation** over this partition is defined as[4]:

$$V_n(g, t) = \sum_{i=1}^{n} |g(s_i) - g(s_{i-1})|\,.$$

[4]Sometimes we speak of absolute variation in order to avoid confusion with e.g. quadratic variation.

If the limit exists independently of the decomposition for $n \to \infty$ under (8.2), then one says that g is of finite variation and writes[5]:

$$V_n(g,t) \;\to\; V(g,t)\,, \quad n \to \infty\,.$$

The finite sum $V_n(g,t)$ measures for a certain partition the absolute increments of the function g on the interval $[0,t]$. If the function evolves sufficiently smooth, then $V(g,t)$ takes on a finite value for $(n \to \infty)$. For very jagged functions, however, it may be that for an increasing refinement $(n \to \infty)$ the increments of the graph of g become larger and larger even for fixed t, such that g is not of finite variation.

Example 10.4 (Monotonic Functions) For monotonic finite functions the variation can be calculated very easily and intuitively. In this case, $V(g,t)$ is simply the absolute value of the difference of the function at endpoints of the interval, $|g(t)-g(0)|$. First, let us assume that g grows monotonically on $[0,t]$,

$$g(s_i) \geq g(s_{i-1}) \quad \text{for } s_i > s_{i-1}\,.$$

Obviously, it then holds by (8.3) that

$$V_n(g,t) = \sum_{i=1}^{n} (g(s_i)-g(s_{i-1})) = g(t)-g(0) = V(g,t)\,.$$

For a monotonically decreasing function, it results quite analogously:

$$\begin{aligned} V_n(g,t) &= \sum_{i=1}^{n} |g(s_i)-g(s_{i-1})| \\ &= -\sum_{i=1}^{n} (g(s_i)-g(s_{i-1})) \\ &= g(0)-g(t) \\ &= V(g,t)\,. \end{aligned}$$

Monotonic functions are hence of finite variation. ■

Without the requirement of monotonicity, an intuitive sufficient condition exists for the function to be smooth enough to be of finite variation where this variation then has a familiar form as well.

---

[5]If g is a deterministic function, then "$\to$" means the usual convergence of analysis. If we allow for $g(t)$ to be a stochastic process, then we mean the convergence in mean square: "$\xrightarrow{2}$".

Proposition 10.5 (Variation of Continuously Differentiable Functions) *Let g be a continuously differentiable function with derivative g' on $[0, t]$. Then g is of finite variation and it holds that*

$$V(g, t) = \int_0^t |g'(s)| \, \mathrm{d}s.$$

The proof is given in Problem 10.6.

Example 10.5 (Sine Wave) Let us consider a sine cycle of the frequency k on the interval $[0, 2\pi]$:

$$g_k(s) = \sin(ks), \quad k = 1, 2, \dots .$$

The derivative reads

$$g_k'(s) = k \cos(ks).$$

Accounting for the sign one obtains as the variation:

$$\begin{aligned} V(g_1, 2\pi) &= \int_0^{2\pi} |\cos(s)| \, ds = 4 \int_0^{\pi/2} \cos(s) \, ds \\ &= 4 \left(\sin\left(\frac{\pi}{2}\right) - \sin(0) \right) \\ &= 4 , \end{aligned}$$

$$\begin{aligned} V(g_2, 2\pi) &= \int_0^{2\pi} 2 \, |\cos(2s)| \, ds = 8 \int_0^{\pi/4} 2 \cos(2s) \, ds \\ &= 8 \left(\sin\left(\frac{\pi}{2}\right) - \sin(0) \right) \\ &= 8 , \end{aligned}$$

$$\begin{aligned} V(g_k, 2\pi) &= \int_0^{2\pi} k \, |\cos(ks)| \, ds = 4k \int_0^{\pi/2k} k \cos(ks) \, ds \\ &= 4k \left(\sin\left(\frac{\pi}{2}\right) - \sin(0) \right) \\ &= 4k . \end{aligned}$$

In Fig. 10.1 it can be observed, how the sum of (absolute) differences in amplitude grows with k growing. Accordingly, the absolute variation of $g_k(s) = \sin(ks)$ multiplies with k. For $k \to \infty$, g_k' tends to infinity such that this derivative is not

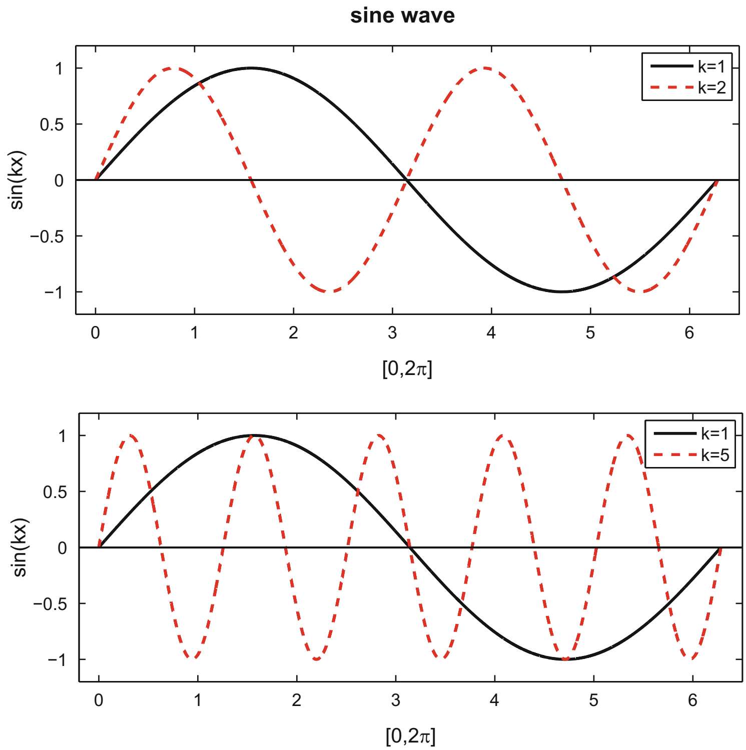

Fig. 10.1 Sine cycles of different frequencies (Example 10.5)

continuous anymore. Consequently, the absolute variation is not finite in the limiting case $k \to \infty$. ■

Quadratic Variation

In the same way as $V_n(g,t)$ a q-variation can be defined where we are only interested in the case $q = 2$, – the quadratic variation:

$$Q_n(g,t) = \sum_{i=1}^{n} |g(s_i) - g(s_{i-1})|^2 = \sum_{i=1}^{n} (g(s_i) - g(s_{i-1}))^2 \ .$$

As would seem natural, g is called of finite quadratic variation if it holds that

$$Q_n(g,t) \ \to \ Q(g,t)\,, \quad n \to \infty\,.$$

If g is a stochastic function, i.e. a stochastic process, then $Q(g,t)$ and $V(g,t)$ are defined as limits in mean square. Between the absolute variation $V(g,t)$ and the quadratic variation $Q(g,t)$ there are connections which we want to deal with now. If a continuous function is of finite variation, then it is of finite quadratic variation as well, where the latter is in fact zero. This is the statement of the following proposition. As it seems counterintuitive at first sight that Q, as the limit of a positive sum of squares Q_n, can become zero, we start with an example.

Example 10.6 (Identity Function) Let id be the identity function on $[0,t]$:

$$id(s) = s.$$

As the functions increases monotonically, it is of finite variation with

$$V(id,t) = id(t) - id(0) = t\,.$$

For finite n it holds that:

$$\begin{aligned} Q_n(id,t) &= \sum_{i=1}^{n} (id(s_i) - id(s_{i-1}))^2 \\ &= \sum_{i=1}^{n} (s_i - s_{i-1})^2 \\ &> 0\,. \end{aligned}$$

Q_n consists of n terms, where the lengths $s_i - s_{i-1} > 0$ are of the magnitude $\frac{1}{n}$. Due to the squaring, the n terms are of the magnitude $\frac{1}{n^2}$. Hence, the sum converges to zero for $n \to \infty$. This intuition can be formalized as follows:

$$\begin{aligned} Q_n(id,t) &= \sum_{i=1}^{n} (s_i - s_{i-1})^2 \\ &\le \max_{1\le i\le n} (s_i - s_{i-1}) \sum_{i=1}^{n} (s_i - s_{i-1}) \\ &= \max_{1\le i\le n} (s_i - s_{i-1})\, V_n(id,t) \\ &= \max_{1\le i\le n} (s_i - s_{i-1})\ t \\ &\to 0, \end{aligned}$$

as $\max(s_i - s_{i-1}) \to 0$ for $n \to \infty$. ■

The next proposition gives a sufficient condition for the quadratic variation to vanish.

Proposition 10.6 (Absolute and Quadratic Variation) *Let g be a continuous function on* $[0, t]$. *It then holds under (8.2) for* $n \to \infty$*:*

$$V_n(g,t) \;\to\; V(g,t) < \infty$$

implies

$$Q_n(g,t) \;\to\; 0\,.$$

If g is a stochastic process, then "→" is to be understood as convergence in mean square.

The proof is given in Problem 10.7. From the proposition it follows by contraposition that: If we have a positive (finite) quadratic variation, then the process does not have a finite variation. Formally, we write: From

$$Q_n(g,t) \to Q(g,t) < \infty \quad \text{with} \quad Q(g,t) > 0$$

it follows that there is no finite variation:

$$V_n(g,t) \to \infty\,.$$

If a function g is so smooth that it has a continuous derivative, then $Q(g,t) = 0$ by Propositions 10.5 and 10.6; the other way round, values of $Q(g,t) > 0$ characterize how little smooth or jagged the function is.

Wiener Processes

As we know, the WP is nowhere differentiable, therefore it is everywhere so jagged that there is no valid tangent line approximation. Due to this extreme jaggedness the WP is of infinite variation as well, as we will show in a moment. More explicitly, we prove that the WP is of positive quadratic variation and does not have a finite absolute variation due to Proposition 10.6. We save the proof for an exercise (Problem 10.8).

Proposition 10.7 (Quadratic Variation of the WP) *For the Wiener process with $n \to \infty$ it holds under (8.2):*

$$Q_n(W,t) \xrightarrow{2} t = Q(W,t)\,.$$

The expression $Q(W,t) = t$ characterizes the level of jaggedness or irregularity of the Wiener process on the interval $[0,t]$. This non-vanishing quadratic variation causes the problems and specifics of the Ito integral. Let us recapitulate: If the Wiener process was continuously differentiable, then it would be of finite variation due to Proposition 10.5 and it would have a vanishing quadratic variation due to Proposition 10.6. However, this is just not the case.

Symbolic Notation

In finance textbooks one frequently finds a notation for time that is strange at first sight:

$$(dW(t))^2 = dt\,. \tag{10.7}$$

How is this to be understood? Formal integration yields

$$\int_0^t (dW(s))^2 = t\,.$$

As would seem natural, the "integral" on the left-hand side here stands for $Q(W,t)$:

$$Q_n(W,t) = \sum_{i=1}^{n} (W(s_i) - W(s_{i-1}))^2 \xrightarrow{2} \int_0^t (dW(s))^2 := Q(W,t)\,.$$

Therefore, the integral equation and hence (10.7) is justified by Proposition 10.7: $Q(W,t) = t$. We adopt the result into the following proposition. The expressions

$$dW(t)\,dt = 0 \quad \text{and} \quad (dt)^2 = 0 \tag{10.8}$$

are to be understood similarly, namely in the sense of Proposition 10.8.

Proposition 10.8 (Symbolic Notation) *It holds for $n \to \infty$ under (8.2):*

$$(a)\ \sum_{i=1}^{n} (W(s_i) - W(s_{i-1}))^2 \xrightarrow{2} \int_0^t (dW(s))^2 = t\,,$$

$$(b)\ \sum_{i=1}^{n} (W(s_i) - W(s_{i-1}))\ (s_i - s_{i-1}) \xrightarrow{2} \int_0^t dW(s)\,ds = 0\,,$$

$$(c)\ \sum_{i=1}^{n} (s_i - s_{i-1})^2 \to \int_0^t (ds)^2 = 0\,.$$

In symbols, these facts are frequently formulated as in (10.7) and (10.8).

Note that the expression in (c) in Proposition 10.8 is the quadratic variation of the identity function $id(s) = s$:

$$Q_n(id, t) \to \int_0^t (ds)^2 := Q(id, t) = 0\,.$$

Hence, the third claim is already established by Example 10.6. The expression from (b) in Proposition 10.8 is sometimes also called covariation (of $W(s)$ and $id(s) = s$). The claimed convergence to zero in mean square is shown in Problem 10.9.

10.5 Problems and Solutions

Problems

10.1 Prove Proposition 10.1 for $\gamma = 0$ (Ito integral).
Hint: Use Proposition 10.7.

10.2 Prove the autocovariance from Proposition 10.2.

10.3 Derive that the Ito integral from (10.3) and the corresponding Stratonovich integral have the same variance.

10.4 Show for $S_n(W)$ from (10.1) with s_i^* from Proposition 10.1,

$$s_i^* = (1 - \gamma)\, s_{i-1} + \gamma\, s_i\,, \quad 0 \le \gamma < 1\,,$$

that it holds:

$$\mathrm{E}\left(S_n(W)\right) = \gamma\, t\,.$$

10.5 Show: $\Gamma_n \xrightarrow{2} 0$ with

$$\Gamma_n = W\left((1-\gamma)\, s_{i-1} + \gamma\, s_i\right) - \left[(1-\gamma)\, W(s_{i-1}) + \gamma\, W(s_i)\right]$$

for $\gamma \in [0, 1]$ for an adequate partition, i.e. for $s_i - s_{i-1} \to 0$.

10.6 Prove Proposition 10.5.

10.7 Prove Proposition 10.6.

10.8 Determine the quadratic variation of the Wiener process, i.e. verify Proposition 10.7.

10.9 Show (b) from Proposition 10.8.

10.10 Determine the covariance of $W(s)$ and $\int_0^t W(r)\, dW(r)$ for $s \le t$.

Solutions

10.1 In order to prove Proposition 10.1 for $\gamma = 0$, it has to be shown that $I_n(W)$ from (10.2) converges in mean square, namely to the expression given in (10.3). For this purpose, we write $I_n(W)$ as follows:

$$\begin{aligned}
I_n(W) &= \sum_{i=1}^{n} W(s_{i-1})\ \left(W(s_i) - W(s_{i-1})\right)\\
&= \frac{1}{2}\left[2\sum_{i=1}^{n} W(s_i)\, W(s_{i-1}) - 2\sum_{i=1}^{n} W^2(s_{i-1})\right]\\
&= \frac{1}{2}\left[\sum_{i=1}^{n}\left(W^2(s_i) - W^2(s_{i-1})\right)\right.\\
&\qquad \left. - \sum_{i=1}^{n}\left(W^2(s_i) - 2\, W(s_i)\, W(s_{i-1}) + W^2(s_{i-1})\right)\right],
\end{aligned}$$

where for the last equation we added zero, $\left(W^2(s_i) - W^2(s_i)\right)$. Hence, it furthermore follows by means of the quadratic variation:

$$\begin{aligned} I_n(W) &= \frac{1}{2}\left[W^2(s_n) - W^2(s_0) - \sum_{i=1}^{n}(W(s_i) - W(s_{i-1}))^2\right] \\ &= \frac{1}{2}\left(W^2(t) - W^2(0)\right) - \frac{1}{2}Q_n(W,t)\,. \end{aligned}$$

The Wiener process is of finite quadratic variation and we know from Proposition 10.7: $Q_n(W,t) \stackrel{2}{\to} t$. This verifies the claim (as it holds that $W(0) = 0$ with probability 1).

10.2 Based on (10.3) we consider the process

$$I(t) = \int_0^t W(s)dW(s) = \frac{W^2(t) - t}{2}.$$

Due to the vanishing expected value, it holds for the autocovariance that:

$$\begin{aligned} \mathrm{E}(I(t)\,I(s)) &= \frac{1}{4}\mathrm{E}\left[W^2(t)W^2(s) - tW^2(s) - sW^2(t) + st\right] \\ &= \frac{1}{4}\left[\mathrm{E}\left(W^2(t)W^2(s)\right) - ts - st + st\right]. \end{aligned}$$

By adding zero one obtains:

$$\begin{aligned} \mathrm{E}[W^2(t)W^2(s)] &= \mathrm{E}\left[(W(t) - W(s) + W(s))^2\,W^2(s)\right] \\ &= \mathrm{E}\left[(W(t) - W(s))^2\,W^2(s)\right] \\ &\quad + 2\,\mathrm{E}\left[(W(t) - W(s))W^3(s)\right] + \mathrm{E}\left[W^4(s)\right]. \end{aligned}$$

If we assume w.l.o.g. that $s \le t$, then due to the independence of non-overlapping increments of W it holds that:

$$\begin{aligned} \mathrm{E}\left[(W(t) - W(s))^2\,W^2(s)\right] &= \mathrm{E}\left[(W(t) - W(s))^2\right]\mathrm{E}\left[W^2(s)\right] \\ &= \mathrm{Var}\,(W(t) - W(s))\,\mathrm{Var}(W(s)) \\ &= (t - s)s \end{aligned}$$

and

$$\mathrm{E}\left[(W(t)-W(s))W^3(s)\right]=\mathrm{E}(W(t)-W(s))\mathrm{E}(W^3(s))=0.$$

As the kurtosis of a Gaussian random variable is 3, it follows that

$$\mathrm{E}[W^4(s)]=3s^2.$$

Therefore, these results jointly yield the claimed outcome:

$$\begin{aligned}\mathrm{E}(I(t)\,I(s)) &= \frac{(t-s)s}{4}+\frac{3}{4}s^2-\frac{st}{4}\\ &= \frac{s^2}{2}, \quad s\le t\,.\end{aligned}$$

10.3 We know about the aforementioned Ito integral from Proposition 10.2 that its variance is $t^2/2$. Hence, it is to be shown that:

$$\mathrm{Var}\left(\int_0^t W(s)\,\partial W(s)\right)=\frac{t^2}{2}\,.$$

We use Proposition 10.1,

$$\int_0^t W(s)\,\partial W(s)=\frac{W^2(t)}{2}\,,$$

from which it follows immediately that

$$\mathrm{E}\left(\int_0^t W(s)\,\partial W(s)\right)=\frac{t}{2}\,.$$

The usual variance decomposition, see (2.1), hence yields

$$\mathrm{Var}\left(\frac{W^2(t)}{2}\right)=\frac{1}{4}\left(\mathrm{E}\left(W^4(t)\right)-t^2\right)\,.$$

Due to a kurtosis of 3, the fourth moment of a $\mathcal{N}(0,t)$-distribution just amounts to $3\,t^2$. Thus, one obtains

$$\mathrm{Var}\left(\frac{W^2(t)}{2}\right)=\frac{1}{4}\left(3\,t^2-t^2\right)=\frac{t^2}{2}\,,$$

which proves the claim.

10.4 The expectation of the sum $S_n(W)$ is equal to the sum of the expectations. Therefore, we consider an individual expectation,

$$\begin{aligned} \mathrm{E}\left[W(s_i^*)\ (W(s_i) - W(s_{i-1})\right] &= \mathrm{E}\left[W(s_i^*)W(s_i) - W(s_i^*)W(s_{i-1})\right] \\ &= \min(s_i^*, s_i) - \min(s_i^*, s_{i-1}) \\ &= (1-\gamma)\, s_{i-1} + \gamma\, s_i - s_{i-1} \\ &= \gamma\, (s_i - s_{i-1}), \end{aligned}$$

where simply the well-known covariance formula was used. Hence, summation yields as desired

$$\begin{aligned} \mathrm{E}\,(S_n(W)) &= \gamma \sum_{i=1}^{n} (s_i - s_{i-1}) \\ &= \gamma(\,s_n - s_0) \\ &= \gamma\,(t-0)\,. \end{aligned}$$

10.5 Convergence in mean square implies that the mean squared error tends to zero. The MSE with the limit zero reads $\mathrm{MSE}(\varGamma_n, 0) = \mathrm{E}(\varGamma_n^2)$. Therefore, it remains to be shown that: $\mathrm{E}(\varGamma_n^2) \to 0$.

For this purpose one considers with $s_i^* = (1-\gamma)s_{i-1} + \gamma\, s_i$:

$$\begin{aligned} \varGamma_n^2 = W^2(s_i^*)\ &-\ 2W(s_i^*)\,[(1-\gamma)W(s_{i-1}) + \gamma\, W(s_i)] \\ &+\ (1-\gamma)^2\, W^2(s_{i-1}) + 2\gamma\,(1-\gamma)\, W(s_{i-1})W(s_i) \\ &+\ \gamma^2 W^2(s_i). \end{aligned}$$

Forming expectation yields:

$$\begin{aligned} \mathrm{E}(\varGamma_n^2) &= s_i^* - 2\left[(1-\gamma)s_{i-1} + \gamma\, s_i^*\right] + (1-\gamma)^2 s_{i-1} \\ &\quad +2\gamma(1-\gamma)\, s_{i-1} + \gamma^2\, s_i \\ &= (1-\gamma)s_{i-1} + \gamma\, s_i - 2(1-\gamma)\, s_{i-1} - 2\gamma(1-\gamma)s_{i-1} - 2\,\gamma^2\, s_i \\ &\quad +(1-\gamma)^2\, s_{i-1} + 2\,\gamma(1-\gamma)\, s_{i-1} + \gamma^2\, s_i \\ &= s_i(\gamma-\gamma^2) + s_{i-1}\left((1-\gamma)^2 - (1-\gamma)\right) \\ &= s_i(\gamma-\gamma^2) + s_{i-1}(\gamma^2-\gamma) \\ &= (s_i - s_{i-1})\,\gamma\,(1-\gamma). \end{aligned}$$

Hence, for $n \to \infty$ the required result is obtained as $s_i - s_{i-1}$ tends to zero.

10.6 For a given partition we write

$$V_n(g,t) = \sum_{i=1}^{n} |g(s_i) - g(s_{i-1})| = \sum_{i=1}^{n} \left| \int_{s_{i-1}}^{s_i} g'(s)\, ds \right| .$$

As the derivative is continuous, $|g'(s)|$ is continuous as well and hence integrable. According to the mean value theorem an $s_i^* \in [s_{i-1}, s_i]$ exists with

$$\int_{s_{i-1}}^{s_i} g'(s)\, ds = g'(s_i^*)\,(s_i - s_{i-1}) .$$

Thus it follows that

$$\begin{aligned} V_n(g,t) &= \sum_{i=1}^{n} |g'(s_i^*)|\,(s_i - s_{i-1}) \\ &\to \int_0^t |g'(s)|\, ds . \end{aligned}$$

Quod erat demonstrandum.

10.7 The claim is based on the bound

$$\begin{aligned} Q_n(g,t) &\le \max_{1\le i\le n} (|g(s_i) - g(s_{i-1})|) \sum_{i=1}^{n} |g(s_i) - g(s_{i-1})| \\ &= \max_{1\le i\le n} (|\, g(s_i) - g(s_{i-1})\, |)\; V_n(g,t) . \end{aligned}$$

Due to continuity it holds that

$$\max_{1\le i\le n} (|\, g(s_i) - g(s_{i-1})\, |) \;\to\; 0 .$$

Hence, the claim immediately follows from the bound.

10.8 It is to be shown that the mean squared error,

$$\mathrm{MSE}\,(Q_n(W,t),\ t) = \mathrm{E}\left[(Q_n(W,t) - t)^2\right] ,$$

tends to zero. For this purpose we proceed in two steps. In the first one we show that the MSE coincides with the variance of $Q_n(W,t)$. In the second step it will be shown that the variance converges to zero.

(1) For the first step we only need to derive $\mathrm{E}(Q_n(W,t)) = t$. With

$$Q_n(W,t) = \sum_{i=1}^{n} (W(s_i) - W(s_{i-1}))^2$$

the required expectation can be easily determined:

$$\begin{aligned}\mathrm{E}\,(Q_n(W,t)) &= \sum_{i=1}^{n} \mathrm{Var}\,(W(s_i) - W(s_{i-1})) \\ &= \sum_{i=1}^{n} (s_i - s_{i-1}) = s_n - s_0 = t - 0 \\ &= t\,.\end{aligned}$$

(2) Due to the independence of the increments of the WP one has

$$\mathrm{Var}\,(Q_n(W,t)) = \sum_{i=1}^{n} \mathrm{Var}\left[(W(s_i) - W(s_{i-1}))^2\right]\,.$$

Due to $W(s_i) - W(s_{i-1}) \sim \mathcal{N}(0, s_i - s_{i-1})$ and with a kurtosis of 3 for Gaussian random variables, it furthermore holds that:

$$\begin{aligned}\mathrm{Var}\left[(W(s_i) - W(s_{i-1}))^2\right] &= \mathrm{E}\left[(W(s_i) - W(s_{i-1}))^4\right] - \left(\mathrm{E}\left[(W(s_i) - W(s_{i-1}))^2\right]\right)^2 \\ &= 3\,\left[\mathrm{Var}(W(s_i) - W(s_{i-1}))\right]^2 - (s_i - s_{i-1})^2 \\ &= 2\,(s_i - s_{i-1})^2\,.\end{aligned}$$

Hence, plugging in yields

$$\begin{aligned}\mathrm{Var}(Q_n(W,t)) &= 2\sum_{i=1}^{n} (s_i - s_{i-1})^2 \\ &\le 2 \max_{1\le i\le n} (s_i - s_{i-1}) \sum_{i=1}^{n} (s_i - s_{i-1}) \\ &= 2 \max_{1\le i\le n} (s_i - s_{i-1})\,(s_n - s_0) \\ &\to 0\,, \quad n \to \infty\,,\end{aligned}$$

which completes the proof.

10.9 Let us call the aforementioned covariation CV_n:

$$CV_n = \sum_{i=1}^{n} (W(s_i) - W(s_{i-1}))\ (s_i - s_{i-1}) .$$

The claim reads: $\text{MSE}(CV_n,\, 0) \to 0$. As it obviously holds that $\text{E}(CV_n) = 0$, we obtain

$$\text{MSE}(CV_n,\, 0) = \text{Var}(CV_n) .$$

Hence, it remains to be shown that this variance tends to zero: Due to the independence of the increments of the WP, one determines

$$\text{Var}(CV_n) = \sum_{i=1}^{n} \text{Var}\,(W(s_i) - W(s_{i-1}))\ (s_i - s_{i-1})^2 ,$$

and hence

$$\begin{aligned}
\text{Var}(CV_n) &= \sum_{i=1}^{n} (s_i - s_{i-1})^3 \\
&\le \max_{1\le i\le n} (s_i - s_{i-1}) \sum_{i=1}^{n} (s_i - s_{i-1})^2 \\
&= \max_{1\le i\le n} (s_i - s_{i-1})\, Q_n(id, t) \\
&\to 0 ,
\end{aligned}$$

where $Q_n(id, t)$ is the quadratic variation of the identity function, see Example 10.6. Hence, the claim is established.

10.10 We want to obtain the expected value of $Y(s, t)$ with

$$Y(s, t) := W(s) \int_0^t W(r)dW(r) , \quad s \le t.$$

Due to (10.3) it again holds that:

$$\begin{aligned}
\text{E}\,(Y(s, t)) &= \text{E}\left[W(s) \frac{W^2(t) - t}{2} \right] \\
&= \frac{1}{2} \text{E}\left[W(s) W^2(t) \right] .
\end{aligned}$$

Therefore, we study

$$\begin{aligned}\mathrm{E}\left[W(s)W^2(t)\right] &= \mathrm{E}\left[W(s)(W(t)-W(s)+W(s))^2\right]\\ &= \mathrm{E}\left[W(s)(W(t)-W(s))^2\right]\\ &\quad +2\mathrm{E}\left[W^2(s)(W(t)-W(s))\right]+\mathrm{E}[W^3(s)].\end{aligned}$$

Let us consider the last three terms one by one. Due to the independence of the increments, one obtains:

$$\mathrm{E}\left[W(s)(W(t)-W(s))^2\right] = \mathrm{E}(W(s))\mathrm{E}\left((W(t)-W(s))^2\right) = 0.$$

Moreover, it is obvious that the second term is also zero. For the third term the symmetry of the Gaussian distribution yields $\mathrm{E}\left(W^3(s)\right) = 0$. Summing up, we have shown that

$$\mathrm{E}\left(W(s)W^2(t)\right) = 0\,, \quad s \leq t,$$

and hence

$$\mathrm{E}\left[W(s)\int_0^t W(r)dW(r)\right] = 0\,, \quad s \leq t.$$

References

Klebaner, F. C. (2005). *Introduction to stochastic calculus with applications* (2nd ed.). London: Imperical College Press.

Øksendal, B. (2003). *Stochastic differential equations: An introduction with applications* (6th ed.). Berlin/New York: Springer.

Phillips, P. C. B., & Park, J. Y. (1988). Asymptotic equivalence of ordinary least squares and generalized least squares in regressions with integrated regressors. *Journal of the American Statistical Association, 83*, 111–115.

Tanaka, K. (1996). *Time series analysis: Nonstationary and noninvertible distribution theory*. New York: Wiley.