Introduction to the theory of random processes

N.V. Krylov

Author address:

School of Mathematics, University of Minnesota, Minneapolis, MN, 55455

E-mail address: krylov@math.umn.edu

2000 *Mathematics Subject Classification.* Primary 60-01;
Secondary 60G99.

The author was supported in part by NSF Grant DMS–9876586.

ABSTRACT. These lecture notes concentrate on some general facts and ideas of the theory of stochastic processes. The main objects of study are the Wiener process, the stationary processes, the infinitely divisible processes, and the Itô stochastic equations.

Although it is not possible to cover even a noticeable portion of the topics listed above in a short course, the author sincerely hopes that after having followed the material presented here the reader acquires a good understanding of what kind of results are available and what kind of techniques are used to obtain them.

These notes are intended for graduate students and scientists in mathematics, physics and engineering interested in the theory of Random Processes and its applications.

Contents

Chapter 2

The Wiener Process

1. Brownian motion and the Wiener process

Robert Brown, an English botanist, observed (1828) that pollen grains suspended in water perform an unending chaotic motion. L. Bachelier (1900) derived the law governing the position w_t at time t of a single grain performing a one-dimensional Brownian motion starting at $a \in \mathbb{R}$ at time $t = 0$:

$$P_a\{w_t \in dx\} = p(t, a, x)\, dx, \tag{1}$$

where

$$p(t, a, x) = \frac{1}{\sqrt{2\pi t}} e^{-(x-a)^2/(2t)}$$

is the fundamental solution of the heat equation

$$\frac{\partial u}{\partial t} = \frac{1}{2} \frac{\partial^2 u}{\partial a^2}.$$

Bachelier (1900) also pointed out the Markovian nature of the Brownian path and used it to establish the law of maximum displacement

$$P_a\{\max_{s \le t} w_s \le b\} = \frac{2}{\sqrt{2\pi t}} \int_0^b e^{-x^2/(2t)}\, dx, \quad t > 0, b \ge 0.$$

Einstein (1905) also derived (1) from statistical mechanics considerations and applied it to the determination of molecular diameters. Bachelier was unable to obtain a clear picture of the Brownian motion, and his ideas were unappreciated at the time. This is not surprising, because the precise mathematical definition of the Brownian motion involves a measure on the path space, and even after the ideas of Borel, Lebesgue, and Daniell appeared, N. Wiener (1923) only constructed a Daniell integral on the path space which later was revealed to be the Lebesgue integral against a measure, the so-called Wiener measure.

The simplest model describing movement of a particle subject to hits by much smaller particles is the following. Let η_k, $k = 1, 2, ...$, be independent identically distributed random variables with $E\eta_k = 0$ and $E\eta_k^2 = 1$. Fix an integer n, and at times $1/n, 2/n, ...$ let our particle experience instant displacements by $\eta_1 n^{-1/2}$, $\eta_2 n^{-1/2}, ...$ At moment zero let our particle be at zero. If

$$S_k := \eta_1 + ... + \eta_k,$$

then at moment k/n our particle will be at the point $S_k/\sqrt{n}$ and will stay there during the time interval $[k/n, (k+1)/n)$. Since real Brownian motion has continuous paths, we replace our piecewise constant trajectory by a continuous piecewise linear one preserving its positions at times k/n. Thus we come to the process

$$\xi_t^n := S_{[nt]}/\sqrt{n} + (nt - [nt])\eta_{[nt]+1}/\sqrt{n}. \qquad (2)$$

This process gives a very rough caricature of Brownian motion. Clearly, to get a better model we have to let $n \to \infty$. By the way, precisely this necessity dictates the intervals of time between collisions to be $1/n$ and the displacements due to collisions to be $\eta_k/\sqrt{n}$, since then ξ_t^n is asymptotically normal with parameters $(0, 1)$.

It turns out that under a very special organization of randomness, which generates different $\{\eta_k; k \geq 1\}$ for different n, one can get the situation where the ξ_t^n converge for each ω uniformly on each finite interval of time. This is a consequence of a very general result due to Skorokhod. We do not use this result, confining ourselves to the weak convergence of the distributions of $\xi_\cdot^n$.

1. Lemma. *The sequence of distributions of $\xi_\cdot^n$ in C is relatively compact.*

Proof. For simplicity we assume that $m_4 := E\eta_k^4 < \infty$, referring the reader to [**Bi**] for the proof in the general situation. Since $\xi_0^n = 0$, by Theorem 1.4.7 it suffices to prove that

$$E|\xi_t^n - \xi_s^n|^4 \leq N|t - s|^2 \quad \forall s, t \in [0, 1], \qquad (3)$$

where N is independent of n, t, s.

Without loss of generality, assume that $s < t$. Denote $a_n = E(S_n)^4$. By virtue of the independence of the η_k and the conditions $E\eta_k = 0$ and $E\eta_k^2 = 1$, we have

$$a_{n+1} = E(S_n + \eta_{n+1})^4 = a_n + 4ES_n^3\eta_{n+1} + 6ES_n^2\eta_{n+1}^2 + 4ES_n\eta_{n+1}^3 + m_4 = a_n + 6n + m_4.$$

Hence (for instance, by induction),

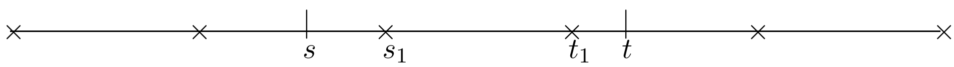
$$a_n = 3n(n-1) + nm_4 \leq 3n^2 + nm_4.$$

Furthermore, if s and t belong to the same interval $[k/n, (k+1)/n]$, then

$$|\xi_t^n - \xi_s^n| = \sqrt{n}|\eta_{k+1}|\,|t-s|,$$

$$E|\xi_t^n - \xi_s^n|^4 = n^2 m_4 |t-s|^4 \leq m_4|t-s|^2. \tag{4}$$

Now, consider the following picture, where s and t belong to different intervals of type $[k/n, (k+1)/n)$ and by crosses we denote points of type k/n:



Clearly

$$s_1 - s \leq t - s, \quad t - t_1 \leq t - s, \quad t_1 - s_1 \leq t - s, \quad (t_1 - s_1)/n \leq (t_1 - s_1)^2,$$

$$s_1 = ([ns]+1)/n, \quad t_1 = [nt]/n, \quad [nt] - ([ns]+1) = n(t_1 - s_1).$$

Hence and from (4) and the inequality $(a+b+c)^4 \leq 81(a^4 + b^4 + c^4)$ we conclude that

$$E|\xi_t^n - \xi_s^n|^4 \leq 81E(|\xi_t^n - \xi_{t_1}^n|^4 + |\xi_{t_1}^n - \xi_{s_1}^n|^4 + |\xi_{s_1}^n - \xi_s^n|^4)$$

$$\leq 162(t-s)^2 m_4 + 81E|S_{[nt]}/\sqrt{n} - S_{[ns]+1}/\sqrt{n}|^4$$

$$= 162(t-s)^2 m_4 + 81n^{-2}a_{[nt]-([ns]+1)}$$

$$\leq 162(t-s)^2 m_4 + 243(t-s)^2 + 81(t_1-s_1)m_4/n \leq 243(m_4+1)|t-s|^2.$$

Thus for all positions of s and t we have (3) with $N = 243(m_4+1)$. The lemma is proved.

Remember yet another definition from probability theory. We say that a sequence $\xi^n, n \geq 1$, of $\mathbb{R}^k$-valued random variables is *asymptotically normal* with parameters (m, R) if F_{ξ^n} converges weakly to the Gaussian distribution with parameters (m, R) (by F_ξ we denote the distribution of a random variable ξ). Below we use the fact that the weak convergence of distributions is equivalent to the pointwise convergence of their characteristic functions.

2. Lemma. *For every* $0 \leq t_1 < t_2 < ... < t_k \leq 1$ *the vectors* $(\xi^n_{t_1}, \xi^n_{t_2}, ..., \xi^n_{t_k})$ *are asymptotically normal with parameters* $(0, (t_i \wedge t_j))$.

Proof. We only consider the case $k = 2$. Other k's are treated similarly. We have

$$\lambda_1 \xi^n_{t_1} + \lambda_2 \xi^n_{t_2} = (\lambda_1+\lambda_2) S_{[nt_1]}/\sqrt{n} + \lambda_2 (S_{[nt_2]} - S_{[nt_1]+1})/\sqrt{n}$$

$$+\eta_{[nt_1]+1}\{(nt_1 - [nt_1])\lambda_1/\sqrt{n} + \lambda_2/\sqrt{n}\} + \eta_{[nt_2]+1}(nt_2 - [nt_2])\lambda_2/\sqrt{n}.$$

On the right, we have a sum of independent terms. In addition, the coefficients of $\eta_{[nt_1]+1}$ and $\eta_{[nt_2]+1}$ go to zero and

$$E\exp(ia_n \eta_{[nt]+1}) = E\exp(ia_n\eta_1) \to 1 \quad \text{as} \quad a_n \to 0.$$

Finally, by the central limit theorem, for $\varphi(\lambda) = E\exp(i\lambda\eta_1)$,

$$\lim_{n\to\infty} \varphi^n(\lambda/\sqrt{n}) = e^{-\lambda^2/2}.$$

Hence,

$$\lim_{n\to\infty} Ee^{i(\lambda_1\xi^n_{t_1}+\lambda_2\xi^n_{t_2})} = \lim_{n\to\infty} \big(\varphi(\lambda_1/\sqrt{n}+\lambda_2/\sqrt{n})\big)^{[nt_1]}\big(\varphi(\lambda_2/\sqrt{n})\big)^{[nt_2]-[nt_1]-1}$$

$$= \exp\{-((\lambda_1+\lambda_2)^2 t_1 + \lambda_2^2(t_2-t_1))/2\}$$

$$= \exp\{-(\lambda_1^2(t_1\wedge t_1) + 2\lambda_1\lambda_2(t_1\wedge t_2) + \lambda_2^2(t_2\wedge t_2))/2\}.$$

The lemma is proved.

3. Theorem (Donsker)**.** *The sequence of distributions* $F_{\xi^n_\cdot}$ *weakly converges on* C *to a measure. This measure is called the Wiener measure.*

Proof. Owing to Lemma 1, there is a sequence $n_i \to \infty$ such that $F_{\xi^{n_i}_\cdot}$ converges weakly to a measure μ. By Exercise 1.2.10 it only remains to prove that the limit is independent of the choice of subsequences.

Let $F_{\xi^{m_i}_\cdot}$ be another weakly convergent subsequence and ν its limit. Fix $0 \le t_1 < t_2 < ... < t_k \le 1$ and define a continuous function on C by the formula $\pi(x_\cdot) = (x_{t_1}, ..., x_{t_k})$. By Lemma 2, considering π as a random element on $(C, \mathfrak{B}(C), \mu)$, for every bounded continuous $f(x^1, ..., x^k)$, we get

$$\int_{\mathbb{R}^k} f(x^1, ..., x^k)\, \mu\pi^{-1}(dx) = \int_C f(x_{t_1}, ..., x_{t_k})\, \mu(dx_\cdot)$$

$$= \lim_{i\to\infty} \int_C f(x_{t_1}, ..., x_{t_k})\, F_{\xi^{n_i}_\cdot}(dx_\cdot) = \lim_{i\to\infty} Ef(\xi^{n_i}_{t_1}, ..., \xi^{n_i}_{t_k}) = Ef(\zeta_1, ..., \zeta_k),$$

where $(\zeta_1, ..., \zeta_k)$ is a random vector normally distributed with parameters $(0, t_i \wedge t_j)$. One gets the same result considering m_i instead of n_i. By Theorem 1.2.4, we conclude that $\mu\pi^{-1} = \nu\pi^{-1}$. This means that for every Borel $B^{(k)} \subset \mathbb{R}^k$ the measures μ and ν coincide on the set $\{x_\cdot : (x_{t_1}, ..., x_{t_k}) \in B^{(k)}\}$. The collection of all such sets (with varying k, $t_1, ..., t_k$) is an algebra. By a result from measure theory, a measure on a σ-field is uniquely determined by its values on an algebra generating the σ-field. Thus $\mu = \nu$ on $\mathfrak{B}(C)$, and the theorem is proved.

Below we will need the conclusion of the last argument from the above proof, showing that there can be only one measure on $\mathfrak{B}(C)$ with given values on finite dimensional cylinder subsets of C.

4. Remark. Since Gaussian distributions are uniquely determined by their means and covariances, finite-dimensional distributions of Gaussian processes are uniquely determined by mean value and covariance functions. Hence, given a continuous Gaussian process ξ_t, its distribution on $(C, \mathfrak{B}(C))$ is uniquely determined by the functions m_t and $R(s,t)$.

5. Definition. By a Wiener process we mean a continuous Gaussian process on $[0,1]$ with $m_t = 0$ and $R(s,t) = s \wedge t$.

As follows from above, the distributions of all Wiener processes on $(C, \mathfrak{B}(C))$ coincide if the processes exist at all.

6. Exercise*. Prove that if w_t is a Wiener process on $[0,1]$ and c is a constant with $c \ge 1$, then cw_{t/c^2} is also a Wiener process on $[0,1]$. This property is called *self-similarity* of the Wiener process.

7. Theorem. *There exists a Wiener process, and its distribution on* $(C, \mathfrak{B}(C))$ *is the Wiener measure.*

Proof. Let μ be the Wiener measure. On the probability space $(C, \mathfrak{B}(C), \mu)$ define the process $w_t(x.) = x_t$. Then, for every $0 \leq t_1 < ... < t_k \leq 1$ and continuous bounded $f(x^1, ..., x^k)$, as in the proof of Donsker's theorem, we have

$$Ef(w_{t_1}, ..., w_{t_k}) = \int_C f(x_{t_1}, ..., x_{t_k})\, \mu(dx.)$$

$$= \lim_{n\to\infty} Ef(\xi^n_{t_1}, ..., \xi^n_{t_k}) = Ef(\zeta^1, ..., \zeta^k),$$

where ζ is a Gaussian vector with parameters $(0, (t_i \wedge t_j))$. Since f is arbitrary, we see that the distribution of $(w_{t_1}, ..., w_{t_k})$ and $(\zeta^1, ..., \zeta^k)$ coincide, and hence $(w_{t_1}, ..., w_{t_k})$ is Gaussian with parameters $(0, (t_i \wedge t_j))$. Thus, w_t is a Gaussian process, $Ew_{t_i} = 0$, and $R(t_i, t_j) = Ew_{t_i} w_{t_j} = E\zeta_i\zeta_j = t_i \wedge t_j$. The theorem is proved.

This theorem and the remark before it show that the limit in Donsker's theorem is independent of the distributions of the η_k as long as $E\eta_k = 0$ and $E\eta_k^2 = 1$. In this framework Donsker's theorem is called *the invariance principle* (although there is no more "invariance" in this theorem than in the central limit theorem).

2. Some properties of the Wiener process

First we prove two criteria for a process to be a Wiener process.

1. Theorem. *A continuous process on* $[0, 1]$ *is a Wiener process if and only if*

(i) $w_0 = 0$ *(a.s.),*

(ii) $w_t - w_s$ *is normal with parameters* $(0, |t - s|)$ *for every* $s, t \in [0, 1]$,

(iii) $w_{t_1}, w_{t_2} - w_{t_1}, ...w_{t_n} - w_{t_{n-1}}$ *are independent for every* $n \geq 2$ *and* $0 \leq t_1 \leq t_2 \leq ... \leq t_n \leq 1$.

Proof. First assume that w_t is a Wiener process. We have $w_0 \sim N(0, 0)$, hence $w_0 = 0$ (a.s.). Next take $0 \leq t_1 \leq t_2 \leq ... \leq t_n \leq 1$ and let

$$\xi_1 = w_{t_1}, \quad \xi_2 = w_{t_2} - w_{t_1}, ..., \xi_n = w_{t_n} - w_{t_{n-1}}.$$

The vector $\xi = (\xi_1, ..., \xi_n)$ is a linear transform of $(w_{t_1}, ..., w_{t_n})$. Therefore ξ is Gaussian. In particular ξ_i and, generally, $w_t - w_s$ are Gaussian. Obviously, $E\xi_i = 0$ and, for $i > j$,

$$E\xi_i\xi_j = E(w_{t_i} - w_{t_{i-1}})(w_{t_j} - w_{t_{j-1}}) = Ew_{t_i}w_{t_j} - Ew_{t_{i-1}}w_{t_j} - Ew_{t_i}w_{t_{j-1}}$$

$$+Ew_{t_{i-1}}w_{t_{j-1}} = t_j - t_j - t_{j-1} + t_{j-1} = 0.$$

Similarly, the equality $Ew_tw_s = s \wedge t$ implies that $E|w_t - w_s|^2 = |t-s|$. Thus $w_t - w_s \sim N(0, |t-s|)$, and we have proved (ii). In addition $\xi_i \sim N(0, t_i - t_{i-1})$, $E\xi_i^2 = t_i - t_{i-1}$, and

$$E\exp\{i\sum_k \lambda_k\xi_k\} = \exp\{-\frac{1}{2}\sum_{k,r}\lambda_k\lambda_r \mathrm{cov}\,(\xi_k, \xi_r)\}$$

$$= \exp\{-\frac{1}{2}\sum_k \lambda_k^2(t_k - t_{k-1})\} = \prod_k E\exp\{i\lambda_k\xi_k\}.$$

This proves (iii).

Conversely, let w_t be a continuous process satisfying (i) through (iii). Again take $0 \le t_1 \le t_2 \le ... \le t_n \le 1$ and the same ξ_i's. From (i) through (iii), it follows that $(\xi_1, ..., \xi_n)$ is a Gaussian vector. Since $(w_{t_1}, ..., w_{t_n})$ is a linear function of $(\xi_1, ..., \xi_n)$, $(w_{t_1}, ..., w_{t_n})$ is also a Gaussian vector; hence w_t is a Gaussian process. Finally, for every $t_1, t_2 \in [0,1]$ satisfying $t_1 \le t_2$, we have

$$m_{t_1} = E\xi_1 = 0, \quad R(t_1, t_2) = R(t_2, t_1) = Ew_{t_1}w_{t_2} = E\xi_1(\xi_1 + \xi_2)$$

$$= E\xi_1^2 = t_1 = t_1 \wedge t_2.$$

The theorem is proved.

2. Theorem. *A continuous process on* $[0,1]$ *is a Wiener process if and only if*

(i) $w_0 = 0$ *(a.s.),*

(ii) $w_t - w_s$ *is normal with parameters* $(0, |t-s|)$ *for every* $s, t \in [0,1]$*,*

(iii) *for every* $n \ge 2$ *and* $0 \le t_1 \le t_2 \le ... \le t_n \le 1$*, the random variable* $w_{t_n} - w_{t_{n-1}}$ *is independent of* $w_{t_1}, w_{t_2}, ... w_{t_{n-1}}$*.*

Proof. It suffices to prove that properties (iii) of this and the previous theorems are equivalent under the condition that (i) and (ii) hold. We are going to use the notation from the previous proof. If (iii) of the present theorem holds, then

$$E\exp\{i\sum_{k=1}^n \lambda_k\xi_k\} = E\exp\{i\lambda_n\xi_n\}E\exp\{i\sum_{k=1}^{n-1}\lambda_k\xi_k\},$$

since $(\xi_1, ..., \xi_{n-1})$ is a function of $(w_{t_1}, ..., w_{t_{n-1}})$. By induction,

$$E\exp\{i\sum_{k=1}^n \lambda_k\xi_k\} = \prod_k E\exp\{i\lambda_k\xi_k\}.$$

This proves property (iii) of the previous theorem. Conversely if (iii) of the previous theorem holds, then one can carry out the same computation in the opposite direction and get that ξ_n is independent of $(\xi_1, ..., \xi_{n-1})$ and of $(w_{t_1}, ..., w_{t_{n-1}})$, since the latter is a function of the former. The theorem is proved.

3. Theorem (Bachelier)**.** *For every* $t \in (0,1]$ *we have* $\max_{s\le t} w_s \sim |w_t|$, *which is to say that for every* $x \ge 0$

$$P\{\max_{s\le t} w_s \le x\} = \frac{2}{\sqrt{2\pi t}} \int_0^x e^{-y^2/(2t)}\,dy.$$

Proof. Take independent identically distributed random variables η_k so that $P(\eta_k = 1) = P(\eta_k = -1) = 1/2$, and define ξ_t^n by (1.2). First we want to find the distribution of

$$\zeta^n = \max_{[0,1]} \xi_t^n = n^{-1/2} \max_{k\le n} S_k.$$

Observe that, for each n, the sequence $(S_1, ..., S_n)$ takes its every particular value with the same probability 2^{-n}. In addition, for each integer $i > 0$, the number of sequences favorable for the events

$$\{\max_{k\le n} S_k \ge i, S_n < i\} \quad \text{and} \quad \{\max_{k\le n} S_k \ge i, S_n > i\} \tag{1}$$

is the same. One proves this by using the reflection principle; that is, one takes each sequence favorable for the first event, keeps it until the moment when it reaches the level i and then *reflects* its remaining part about this level. This implies equality of the probabilities of the events in (1). Furthermore, due to the fact that i is an integer, we have

$$\{\zeta^n \ge in^{-1/2},\ \xi_1^n < in^{-1/2}\} = \{\max_{k\le n} S_k \ge i, S_n < i\}$$

and

$$\{\zeta^n \ge in^{-1/2},\ \xi_1^n > in^{-1/2}\} = \{\max_{k\le n} S_k \ge i, S_n > i\}.$$

Hence,

$$P\{\zeta^n \ge in^{-1/2},\ \xi_1^n < in^{-1/2}\} = P\{\zeta^n \ge in^{-1/2},\ \xi_1^n > in^{-1/2}\}.$$

Moreover, obviously,

$$P\{\zeta^n \ge in^{-1/2},\ \xi_1^n > in^{-1/2}\} = P\{\xi_1^n > in^{-1/2}\},$$

$$P\{\zeta^n \geq in^{-1/2}\} = P\{\zeta^n \geq in^{-1/2},\ \xi_1^n > in^{-1/2}\}$$
$$+ P\{\zeta^n \geq in^{-1/2},\ \xi_1^n < in^{-1/2}\} + P\{\xi_1^n = in^{-1/2}\}.$$

It follows that

$$P\{\zeta^n \geq in^{-1/2}\} = 2P\{\xi_1^n > in^{-1/2}\} + P\{\xi_1^n = in^{-1/2}\} \tag{2}$$

for every integer $i > 0$. The last equality also obviously holds for $i = 0$. We see that for numbers a of type $in^{-1/2}$, where i is a nonnegative integer, we have

$$P\{\zeta^n \geq a\} = 2P\{\xi_1^n > a\} + P\{\xi_1^n = a\}. \tag{3}$$

Certainly, the last probability goes to zero as $n \to \infty$ since ξ_1^n is asymptotically normal with parameters $(0, 1)$. Also, keeping in mind Donsker's theorem, it is natural to think that

$$P\{\max_{s \leq 1} \xi_s^n \geq a\} \to P\{\max_{s \leq 1} w_s \geq a\}, \quad 2P\{\xi_1^n > a\} \to 2P\{w_1 > a\}.$$

Therefore, (3) naturally leads to the conclusion that

$$P\{\max_{s \leq 1} w_s \geq a\} = 2P\{w_1 > a\} = P\{|w_1| > a\} \quad \forall a \geq 0,$$

and this is our statement for $t = 1$.

To justify the above argument, notice that (2) implies that

$$P\{\zeta^n = in^{-1/2}\} = P\{\zeta^n \geq in^{-1/2}\} - P\{\zeta^n \geq (i+1)n^{-1/2}\}$$
$$= 2P\{\xi_1^n = (i+1)n^{-1/2}\} + P\{\xi_1^n = in^{-1/2}\} - P\{\xi_1^n = (i+1)n^{-1/2}\}$$
$$= P\{\xi_1^n = (i+1)n^{-1/2}\} + P\{\xi_1^n = in^{-1/2}\}, \quad i \geq 0.$$

Now for every bounded continuous function $f(x)$ which vanishes for $x < 0$ we get

$$Ef(\zeta^n) = \sum_{i=0}^{\infty} f(in^{-1/2}) P\{\zeta^n = in^{-1/2}\} = Ef(\xi_1^n - n^{-1/2}) + Ef(\xi_1^n).$$

By Donsker's theorem and by the continuity of the function $x_\cdot \to \max_{[0,1]} x_t$ we have

$$Ef(\max_{[0,1]} w_t) = 2Ef(w_1) = Ef(|w_1|).$$

We have proved our statement for $t = 1$. For smaller t one uses Exercise 1.6, saying that cw_{s/c^2} is a Wiener process for $s \in [0,1]$ if $c \geq 1$. The theorem is proved.

4. Theorem (on the modulus of continuity)**.** *Let w_t be a Wiener process on $[0,1]$, $1/2 > \varepsilon > 0$. Then for almost every ω there exists $n \geq 0$ such that for each $s, t \in [0,1]$ satisfying $|t-s| \leq 2^{-n}$, we have*

$$|w_t - w_s| \leq N|t-s|^{1/2-\varepsilon},$$

where N depends only on ε. In particular, $|w_t| = |w_t - w_0| \leq Nt^{1/2-\varepsilon}$ for $t \leq 2^{-n}$.

Proof. Take a number $\alpha > 2$ and denote $\beta = \alpha/2 - 1$. Let $\xi \sim N(0,1)$. Since $w_t - w_s \sim N(0, |t-s|)$, we have $w_t - w_s \sim \xi|t-s|^{1/2}$. Hence

$$E|w_t - w_s|^{\alpha} = |t-s|^{\alpha/2} E|\xi|^{\alpha} = N_1(\alpha)|t-s|^{1+\beta}.$$

Next, let

$$K_n(a) = \{x_\cdot \in C : |x_0| \leq 2^n, |x_t - x_s| \leq N(a)|t-s|^a \quad \forall |t-s| \leq 2^{-n}\}.$$

By Theorem 1.4.6, for $0 < a < \beta\alpha^{-1}$, we have

$$P\{w_\cdot \in \bigcup_{n=1}^{\infty} K_n(a)\} = 1.$$

Therefore, for almost every ω there exists $n \geq 0$ such that for all $s, t \in [0,1]$ satisfying $|t-s| \leq 2^{-n}$, we have $|w_t(\omega) - w_s(\omega)| \leq N(a)|t-s|^a$. It only remains to observe that we can take $a = 1/2 - \varepsilon$ if from the very beginning we take $\alpha > 1/\varepsilon$ (for instance $\alpha = 2/\varepsilon$). The theorem is proved.

5. Exercise. Prove that there exists a constant N such that for almost every ω there exists $n \geq 0$ such that for each $s, t \in [0,1]$ satisfying $|t-s| \leq 2^{-n}$, we have

$$|w_t - w_s| \leq N\sqrt{|t-s|(-\ln|t-s|)},$$

The result of Exercise 5 is not far from the best possible. P. Lévy proved that

$$\overline{\lim_{\substack{0 \leq s < t \leq 1 \\ u = t-s \to 0}}} \frac{|w_t - w_s|}{\sqrt{2u(-\ln u)}} = 1 \quad \text{(a.s.)}.$$

6. Theorem (on quadratic variation)**.** *Let* $0 = t_{0n} \le t_{1n} \le ... \le t_{k_n n} = 1$ *be a sequence of partitions of* $[0,1]$ *such that* $\max_i(t_{i+1,n} - t_{in}) \to 0$ *as* $n \to \infty$*. Also let* $0 \le s \le t \le 1$*. Then, in probability as* $n \to \infty$*,*

$$\sum_{s \le t_{in} \le t_{i+1,n} \le t} (w_{t_{i+1,n}} - w_{t_{in}})^2 \to t - s. \tag{4}$$

Proof. Let

$$\xi_n := \sum_{s \le t_{in} \le t_{i+1,n} \le t} (w_{t_{i+1,n}} - w_{t_{in}})^2$$

and observe that ξ_n is a sum of independent random variables. Also use that if $\eta \sim N(0, \sigma^2)$, then $\eta = \sigma\zeta$, where $\zeta \sim N(0,1)$, and $\operatorname{Var} \eta^2 = \sigma^4 \operatorname{Var} \zeta^2$. Then, for $N := \operatorname{Var} \zeta$, we obtain

$$\operatorname{Var} \xi_n = \sum_{s \le t_{in} \le t_{i+1,n} \le t} \operatorname{Var} [(w_{t_{i+1,n}} - w_{t_{in}})^2] = N \sum_{s \le t_{in} \le t_{i+1,n} \le t} (t_{i+1,n} - t_{in})^2$$

$$\le N \max_i (t_{i+1,n} - t_{in}) \sum_{0 \le t_{in} \le t_{i+1,n} \le 1} (t_{i+1,n} - t_{in}) = N \max_i (t_{i+1,n} - t_{in}) \to 0.$$

In particular, $\xi_n - E\xi_n \to 0$ in probability. In addition,

$$E\xi_n = \sum_{s \le t_{in} \le t_{i+1,n} \le t} (t_{i+1,n} - t_{in}) \to t - s.$$

Hence $\xi_n - (t-s) = \xi_n - E\xi_n + E\xi_n - (t-s) \to 0$ in probability, and the theorem is proved.

7. Exercise. Prove that if $t_{in} = i/2^n$, then the convergence in (4) holds almost surely.

8. Corollary. *It is not true that there exist functions* $\varepsilon(\omega)$ *and* $N(\omega)$ *such that with positive probability* $\varepsilon(\omega) > 0$*,* $N(\omega) < \infty$*, and*

$$|w_t(\omega) - w_s(\omega)| \le N(\omega)|t - s|^{1/2 + \varepsilon(\omega)}$$

whenever $t, s \in [0,1]$ *and* $|t - s| \le \varepsilon(\omega)$*.*

Indeed, if $|w_t(\omega) - w_s(\omega)| \le N(\omega)|t - s|^{1/2+\varepsilon(\omega)}$ for $|t - s|$ sufficiently small, then

$$\sum_i (w_{t_{i+1,n}}(\omega) - w_{t_{in}}(\omega))^2 \le N^2 \sum_i (t_{i+1,n} - t_{in})^{1+2\varepsilon} \to 0.$$

9. Corollary. $P\{\operatorname{Var}_{[0,1]} w_t = \infty\} = 1$.

This follows from the fact that, owing to the continuity of w_t,

$$\sum_i (w_{t_{i+1,n}}(\omega) - w_{t_{in}}(\omega))^2 \leq \max_i |w_{t_{i+1,n}}(\omega) - w_{t_{in}}(\omega)| \operatorname{Var}_{[0,1]} w_t(\omega) \to 0$$

if $\operatorname{Var}_{[0,1]} w_t(\omega) < \infty$.

10. Exercise. Let w_t be a one-dimensional Wiener process. Find

$$P\{\max_{s\leq 1} w_s \geq b, w_1 \leq a\}.$$

The following exercise is a particular case of the Cameron-Martin theorem regarding the process $w_t - \int_0^t f_s\,ds$ with nonrandom f. Its extremely powerful generalization for random f is known as Girsanov's Theorem 6.8.8.

11. Exercise. Let w_t be a one-dimensional Wiener process on a probability space $(\Omega, \mathcal{F}, P)$. Prove that

$$Ee^{w_t - t/2} = 1.$$

Introduce a new measure by $Q(d\omega) = e^{w_1 - 1/2} P(d\omega)$. Prove that $(\Omega, \mathcal{F}, Q)$ is a probability space, and that $w_t - t$, $t \in [0,1]$, is a Wiener process on $(\Omega, \mathcal{F}, Q)$.

12. Exercise. By using the results in Exercise 11 and the fact that the distributions on $(C, \mathfrak{B}(C))$ of Wiener processes coincide, show that

$$P\{\max_{s\leq 1}[w_s + s] \leq a\} = Ee^{w_1 - 1/2} I_{\max_{s\leq 1} w_s \leq a}.$$

Then by using the result in Exercise 10, compute the last expectation.

Unboundedness of the variation of Wiener trajectories makes it hard to justify the following argument. In real situations the variance of Brownian motion of pollen grains should depend on the water temperature. If the temperature is piecewise constant taking constant value on each interval of a partition $0 \leq t_1 < t_2 < ... < t_n = 1$, then the trajectory can be modeled by

$$\sum_{t_{i+1}\leq t} (w_{t_{i+1}} - w_{t_i}) f_i + (w_t - w_{t_k}) f_k,$$

where $k = \max\{i : t_i \leq t\}$ and the factor f_i reflects the dependence of the variance on temperature for $t \in [t_i, t_{i+1})$. The difficulty comes when one tries to pass from piecewise constant temperatures to continuously changing ones, because the sum should converge to an integral against w_t as we make partitions finer and finer. On the other hand, the integral against w_t is not defined since the variation of w_t is infinite for almost each ω. Yet there is a rather narrow class of functions f, namely functions of bounded variation, for which one can define the Riemann integral against w_t pathwise (see Theorem 3.22). For more general functions one defines the integral against w_t in the mean-square sense.

3. Integration against random orthogonal measures

The reader certainly knows the basics of the theory of L_p spaces, which can be found, for instance, in [**Du**] and which we only need for $p = 1$ and $p = 2$. Our approach to integration against random orthogonal measures requires a version of this theory which starts with introducing step functions using not all measurable sets but rather some collection of them. Actually, the version is quite parallel to the usual theory, and what follows below should be considered as just a reminder of the general scheme of the theory of L_p spaces.

Let X be a set, Π *some* family of subsets of X, $\mathfrak{A}$ a σ-algebra of subsets of X, and μ a measure on $(X, \mathfrak{A})$. Suppose that $\Pi \subset \mathfrak{A}$ and $\Pi_0 := \{\Delta \in \Pi : \mu(\Delta) < \infty\} \neq \emptyset$. Let $S(\Pi) = S(\Pi, \mu)$ denote the set of all *step functions*, that is, functions

$$\sum_{i=1}^{n} c_i I_{\Delta(i)}(x),$$

where c_i are complex numbers, $\Delta(i) \in \Pi_0$ (not Π!), $n < \infty$ is an integer. For $p \in [1, \infty)$, let $L_p(\Pi, \mu)$ denote the set of all $\mathfrak{A}^\mu$-measurable complex-valued functions f on X for each of which there exists a sequence $f_n \in S(\Pi)$ such that

$$\int_X |f - f_n|^p \, \mu(dx) \longrightarrow 0 \quad \text{as} \quad n \to \infty. \tag{1}$$

A sequence $f_n \in S(\Pi)$ that satisfies (1) will be called a *defining sequence* for f. From the convexity of $|t|^p$, we infer that $|a+b|^p \leq 2^{p-1}|a|^p + 2^{p-1}|b|^p$, $|f|^p \leq 2^{p-1}|f_n|^p + 2^{p-1}|f - f_n|^p$ and therefore, if $f \in L_p(\Pi, \mu)$, then

$$\|f\|_p := \Big(\int_X |f|^p \, \mu(dx) \Big)^{1/p} < \infty. \tag{2}$$

The expression $\|f\|_p$ is called the L_p *norm* of f. For $p = 2$ it is also useful to define the *scalar product* (f, g) of elements $f, g \in L_2(\Pi, \mu)$:

$$(f,g) := \int_X f\bar{g}\,\mu(dx). \tag{3}$$

This integral exists and is finite, since $|f\bar{g}| \le |f|^2 + |g|^2$. The expression $\|f-g\|_p$ defines a distance in $L_p(\Pi,\mu)$ between the elements $f, g \in L_p(\Pi,\mu)$. It is "almost" a metric on $L_p(\Pi,\mu)$, in the sense that, although the equality $\|f-g\|_p = 0$ implies that $f = g$ only almost everywhere with respect to μ, nevertheless $\|f-g\|_p = \|g-f\|_p$ and the *triangle inequality* holds:

$$\big\|f+g\big\|_p \le \big\|f\big\|_p + \big\|g\big\|_p.$$

If $f_n, f \in L_p(\Pi,\mu)$ and $\|f_n - f\|_p \to 0$ as $n \to \infty$, we will naturally say that f_n *converges to* f *in* $L_p(\Pi,\mu)$. If $\|f_n - f_m\|_p \to 0$ as $n, m \to \infty$, we will call f_n a *Cauchy sequence* in $L_p(\Pi,\mu)$. The following results are useful. For their proofs we refer the reader to [**Du**].

1. Theorem. (i) *If f_n is a Cauchy sequence in $L_p(\Pi,\mu)$, then there exists a subsequence $f_{n(k)}$ such that $f_{n(k)}$ has a limit μ-a.e. as $k \to \infty$.*

(ii) *$L_p(\Pi,\mu)$ is a linear space, that is, if a, b are complex numbers and $f, g \in L_p(\Pi,\mu)$, then $af + bg \in L_p(\Pi,\mu)$.*

(iii) *$L_p(\Pi,\mu)$ is a complete space, that is, for every Cauchy sequence $f_n \in L_p(\Pi,\mu)$, there exists an $\mathfrak{A}$-measurable function f for which* (1) *is true; in addition, every $\mathfrak{A}^\mu$-measurable function f that satisfies* (1) *for some sequence $f_n \in L_p(\Pi,\mu)$ is an element of $L_p(\Pi,\mu)$.*

2. Exercise*. Prove that if Π is a *σ-field*, then $L_p(\Pi,\mu)$ is simply the set of all Π^μ-measurable functions f that satisfy (2).

3. Exercise. Prove that if Π_0 consists of only one set Δ, then $L_p(\Pi,\mu)$ is the set of all functions μ-almost everywhere equal to a constant times the indicator of Δ.

4. Exercise. Prove that if $(X,\mathfrak{A},\mu) = ([0,1],\mathfrak{B}[0,1],\ell)$ and $\Pi = \{(0,t] : t \in (0,1)\}$, then $L_p(\Pi,\mu)$ is the space of all Lebesgue measurable functions summable to the pth power on $[0,1]$.

We now proceed to the main contents of this section. Let $(\Omega,\mathcal{F},P)$ be a probability space and suppose that to every $\Delta \in \Pi_0$ there is assigned a random variable $\zeta(\Delta) = \zeta(\omega,\Delta)$.

5. Definition. We say that ζ is a *random orthogonal measure with reference measure* μ if (a) $E\,|\zeta(\Delta)|^2 < \infty$ for every $\Delta \in \Pi_0$, (b) $E\,\zeta(\Delta_1)\bar{\zeta}(\Delta_2) = \mu(\Delta_1 \cap \Delta_2)$ for all $\Delta_1, \Delta_2 \in \Pi_0$.

6. Example. If $(X, \mathfrak{A}, \mu) = (\Omega, \mathcal{F}, P)$ and $\Pi = \mathfrak{A}$, then $\zeta(\Delta) := I_\Delta$ is a random orthogonal measure with reference measure μ. In this case, for each ω, ζ is just the Dirac measure concentrated at ω.

Generally, random orthogonal measures are not measures for each ω, because they need not even be defined on a σ-field. Actually, the situation is even more interesting, as the reader will see from Exercise 21.

7. Example. Let w_t be a Wiener process on $[0,1]$ and

$$(X, \mathfrak{A}, \mu) = ([0,1], \mathfrak{B}([0,1]), \ell).$$

Let $\Pi = \{[0,t] : t \in (0,1]\}$ and, for each $\Delta = [0,t] \in \Pi$, let $\zeta(\Delta) = w_t$. Then, for $\Delta_i = [0,t_i] \in \Pi$, we have

$$E\zeta(\Delta_1)\overline{\zeta(\Delta_2)} = Ew_{t_1}w_{t_2} = t_1 \wedge t_2 = \ell(\Delta_1 \cap \Delta_2),$$

which shows that ζ is a random orthogonal measure with reference measure ℓ.

8. Exercise*. Let τ_n be a sequence of independent random variables exponentially distributed with parameter 1. Define a sequence of random variables $\sigma_n = \tau_1 + ... + \tau_n$ and the corresponding counting process

$$\pi_t = \sum_{n=1}^{\infty} I_{[\sigma_n,\infty)}(t).$$

Observe that π_t is a function of locally bounded variation (at least for almost all ω), so that the usual integral against $d\pi_t$ is well defined: if f vanishes outside a finite interval, then

$$\int_0^\infty f(t)\,d\pi_t = \sum_{n=1}^{\infty} f(\sigma_n).$$

Prove that, for every bounded continuous real-valued function f given on $\mathbb{R}$ and having compact support and every $s \in \mathbb{R}$,

$$\varphi(s) := E\exp\{i\int_0^\infty f(s+t)\,d\pi_t\} = \exp(\int_0^\infty (e^{if(s+t)} - 1)\,dt).$$

Conclude from here that $\pi_t - \pi_s$ has Poisson distribution with parameter $|t-s|$. In particular, prove $E\pi_t = t$ and $E(\pi_t - t)^2 = t$. Also prove that π_t is a process with independent increments, that is, $\pi_{t_2} - \pi_{t_1}, ..., \pi_{t_{k+1}} - \pi_{t_k}$ are independent as long as the intervals $(t_j, t_{j+1}]$ are disjoint. The process π_t is called *a Poisson process with parameter 1*.

9. Example. Take the Poisson process π_t from Exercise 8. Denote $m_t = \pi_t - t$. If $0 \leq s \leq t$, then

$$Em_s m_t = Em_s^2 + Em_s(m_t - m_s) = Em_s^2 = s = s \wedge t.$$

Therefore, if in Example 7 we replace w_t with π_t, we again have a random orthogonal measure with reference measure ℓ.

We will always assume that ζ satisfies the assumptions of Definition 5. Note that by Exercise 2 we have $\zeta(\Delta) \in L_2(\mathcal{F}, P)$ for every $\Delta \in \Pi_0$. The word "orthogonal" in Definition 5 comes from the fact that if $\Delta_1 \cap \Delta_2 = \emptyset$, then $\zeta(\Delta_1) \perp \zeta(\Delta_2)$ in the Hilbert space $L_2(\mathcal{F}, P)$. The word "measure" is explained by the property that if $\Delta, \Delta_i \in \Pi_0$, the Δ_i's are pairwise disjoint, and $\Delta = \bigcup_i \Delta_i$, then $\zeta(\Delta) = \sum_i \zeta(\Delta_i)$, where the series converges in the mean-square sense. Indeed,

$$\begin{aligned}
&\lim_{n\to\infty} E|\zeta(\Delta) - \sum_{i\leq n} \zeta(\Delta_i)|^2 \\
&= \lim_{n\to\infty} [E|\zeta(\Delta)|^2 + \sum_{i\leq n} E|\zeta(\Delta_i)|^2 - 2\text{Re} \sum_{i\leq n} E\zeta(\Delta)\bar{\zeta}(\Delta_i)] \\
&= \lim_{n\to\infty} [\mu(\Delta) + \sum_{i\leq n} \mu(\Delta_i) - 2\sum_{i\leq n} \mu(\Delta_i)] = 0.
\end{aligned}$$

Interestingly enough, our explanation of the word "measure" is void in Examples 7 and 9, since there is no $\Delta \in \Pi$ which is representable as a countable union of disjoint members of Π.

10. Lemma. *Let $\Delta_i, \Gamma_j \in \Pi_0$, and let c_i, d_j be complex numbers, $i = 1, ..., n$, $j = 1, ..., m$. Assume $\sum_{i\leq n} c_i I_{\Delta_i} = \sum_{j\leq m} d_j I_{\Gamma_j}$ (μ-a.e.). Then*

$$\sum_{i\leq n} c_i\zeta(\Delta_i) = \sum_{j\leq m} d_j\zeta(\Gamma_j) \quad \textit{(a.s.)}, \tag{4}$$

$$E|\sum_{i\leq n} c_i\zeta(\Delta_i)|^2 = \int_X |\sum_{i\leq n} c_i I_{\Delta_i}|^2 \, \mu(dx). \tag{5}$$

Proof. First we prove (5). We have

$$E|\sum_{i\leq n} c_i\zeta(\Delta_i)|^2 = \sum_{i,j\leq n} c_i\bar{c}_j E\zeta(\Delta_i)\bar{\zeta}(\Delta_j) = \sum_{i,j\leq n} c_i\bar{c}_j \mu(\Delta_i \cap \Delta_j)$$

$$= \int_X \sum_{i,j\leq n} c_i\bar{c}_j I_{\Delta_i} I_{\Delta_j} \, \mu(dx) = \int_X |\sum_{i\leq n} c_i I_{\Delta_i}|^2 \, \mu(dx).$$

Hence,

$$E|\sum_{i\leq n} c_i\zeta(\Delta_i) - \sum_{j\leq m} d_j\zeta(\Gamma_j)|^2 = \int_X |\sum_{i\leq n} c_i I_{\Delta_i} - \sum_{j\leq m} d_j I_{\Gamma_j}|^2 \, \mu(dx) = 0.$$

The lemma is proved.

11. Remark. The first statement of the lemma looks quite surprising in the situation when μ is concentrated at only one point x_0. Then the equality $\sum_{i\leq n} c_i I_{\Delta_i} = \sum_{j\leq m} d_j I_{\Gamma_j}$ holds μ-almost everywhere if and only if

$$\sum_{i\leq n} c_i I_{\Delta_i}(x_0) = \sum_{j\leq m} d_j I_{\Gamma_j}(x_0),$$

and this may hold for very different $c_i, \Delta_i, d_j, \Gamma_j$. Yet each time (4) holds true.

Next, on $S(\Pi)$ define an operator I by the formula

$$I : \sum_{i\leq n} c_i I_{\Delta_i} \to \sum_{i\leq n} c_i\zeta(\Delta_i).$$

In the future we will always identify two elements of an L_p space which coincide almost everywhere. Under this stipulation, Lemma 10 shows that I is a well defined linear unitary operator from a subset $S(\Pi)$ of $L_2(\Pi,\mu)$ into $L_2(\mathcal{F}, P)$. In addition, by definition $S(\Pi)$ is dense in $L_2(\Pi,\mu)$ and every isometric operator is uniquely extendible from a dense subspace to the whole space. By this we mean the following result, which we suggest as an exercise.

12. Lemma. *Let B_1 and B_2 be Banach spaces and B_0 a linear subset of B_1. Let a linear isometric operator I be defined on B_0 with values in B_2 ($|Ib|_{B_2} = |b|_{B_1}$ for every $b \in B_0$). Then there exists a unique linear isometric operator $\tilde{I} : \bar{B}_0 \to B_2$ ($\bar{B}_0$ is the closure of B_0 in B_1) such that $\tilde{I}b = Ib$ for every $b \in B_0$.*

Combining the above arguments, we arrive at the following.

13. Theorem. *There exists a unique linear operator $I : L_2(\Pi,\mu) \to L_2(\mathcal{F}, P)$ such that*

(i) *$I(\sum_{i\leq n} c_i I_{\Delta_i}) = \sum_{i\leq n} c_i\zeta(\Delta_i)$ (a.s.) for all finite n, $\Delta_i \in \Pi_0$ and complex c_i;*

(ii) *$E|If|^2 = \int_X |f|^2 \, \mu(dx)$ for all $f \in L_2(\Pi,\mu)$.*

For $f \in L_2(\Pi,\mu)$ we write

$$If = \int_X f(x)\,\zeta(dx)$$

and we call If *the stochastic integral of* f with respect to ζ. Observe that, by continuity of I, to find If it suffices to construct step functions f_n converging to f in the $L_2(\Pi,\mu)$ sense, and then

$$\int_X f(x)\,\zeta(dx) = \underset{n\to\infty}{\text{l.i.m.}} \int_X f_n(x)\,\zeta(dx).$$

The operator I preserves not only the norm but also the scalar product:

$$E\int_X f(x)\,\zeta(dx)\overline{\int_X g(x)\,\zeta(dx)} = \int_X f\bar{g}\,\mu(dx), \quad f,g\in L_2(\Pi,\mu). \tag{6}$$

This follows after comparing the coefficients of the complex parameter λ in the equal (by Theorem 13) polynomials $E|I(f+\lambda g)|^2$ and $\int |f+\lambda g|^2\,\mu(dx)$.

14. Exercise. Take π_t from Example 9. Prove that for every Borel $f\in L_2(0,1)$ the stochastic integral of f against $\pi_t - t$ equals the usual integral; that is,

$$-\int_0^1 f(s)\,ds + \sum_{\sigma_n\le 1} f(\sigma_n).$$

15. Remark. If $E\zeta(\Delta)=0$ for every $\Delta\in\Pi_0$, then for every $f\in L_2(\Pi,\mu)$, we have

$$E\int_X f\,\zeta(dx) = 0.$$

Indeed, for $f\in S(\Pi)$, this equality is verified directly; for arbitrary $f\in L_2(\Pi,\mu)$ it follows from the fact that, by Cauchy's inequality for $f_n\in S(\Pi)$,

$$|E\int_X f\,\zeta(dx)|^2 = |E\int_X (f-f_n)\,\zeta(dx)|^2$$

$$\le E|\int_X (f-f_n)\,\zeta(dx)|^2 = \int_X |f-f_n|^2\,\mu(dx).$$

We now proceed to the question as to when $L_p(\Pi,\mu)$ and $L_p(\mathfrak{A},\mu)$ coincide, which is important in applications. Remember the following definitions.

16. Definition. Let X be a set, $\mathfrak{B}$ a family of subsets of X. Then $\mathfrak{B}$ is called a *π-system* if $A_1 \cap A_2 \in \mathfrak{B}$ for every $A_1,\ A_2 \in \mathfrak{B}$. It is called a *$\lambda$-system* if

(i) $X \in \mathfrak{B}$ and $A_2 \setminus A_1 \in \mathfrak{B}$ for every $A_1, A_2 \in \mathfrak{B}$ such that $A_1 \subset A_2$;

(ii) for every $A_1, A_2, ... \in \mathfrak{B}$ such that $A_i \cap A_j = \emptyset$ when $i \neq j$, $\bigcup\limits_{n=1}^{\infty} A_n \in \mathfrak{B}$.

A typical example of λ-system is given by the collection of all subsets on which two given probability measures coincide.

17. Exercise*. Prove that if $\mathfrak{B}$ is both a λ-system and a π-system, then it is a σ-field.

A very important property of π- and λ-systems is stated as follows.

18. Lemma. *If Λ is a λ-system and Π is a π-system and $\Pi \subset \Lambda$, then $\sigma(\Pi) \subset \Lambda$.*

Proof. Let Λ_1 denote the smallest λ-system containing Π (Λ_1 is the intersection of all λ-systems containing Π). It suffices to prove that $\Lambda_1 \supset \sigma(\Pi)$. To do this, it suffices to prove, by Exercise 17, that Λ_1 is a π-system, that is, it contains the intersection of every two of its sets. For $B \in \Lambda_1$ let $\Lambda(B)$ denote the family of *all* $A \in \Lambda_1$ such that $A \cap B \in \Lambda_1$. Obviously, $\Lambda(B)$ is a λ-system. In addition, if $B \in \Pi$, then $\Lambda(B) \supset \Pi$ (since Π is a π-system). Consequently, if $B \in \Pi$, then by the definition of Λ_1, we have $\Lambda(B) \supset \Lambda_1$. But this means that $\Lambda(A) \supset \Pi$ for each $A \in \Lambda_1$, so that as before, $\Lambda(A) \supset \Lambda_1$ for each $A \in \Lambda_1$, that is, Λ_1 is a π-system. The lemma is proved.

19. Theorem. *Let $\mathfrak{A}_1 = \sigma(\Pi)$. Assume that Π is a π-system and that there exists a sequence $\Delta(1), \Delta(2), ... \in \Pi_0$ such that $\Delta(n) \subset \Delta(n+1)$, $X = \bigcup_n \Delta(n)$. Then $L_p(\Pi, \mu) = L_p(\mathfrak{A}_1, \mu)$.*

Proof. Let Σ denote the family of all subsets A of X such that

$$I_A I_{\Delta(n)} \in L_p(\Pi, \mu)$$

for every n. Observe that Σ is a λ-system. Indeed for instance, if $A_1, A_2, ... \in \Sigma$ are pairwise disjoint and $A = \bigcup_k A_k$, then

$$I_A I_{\Delta(n)} = \sum_k I_{A_k} I_{\Delta(n)},$$

where the series converges in $L_p(\Pi, \mu)$ since $\bigcup_{k \geq m} A_k \downarrow \emptyset$ as $m \to \infty$, $\mu(\Delta(n)) < \infty$, and

$$\int_X |\sum_{k\geq m} I_{A_k} I_{\Delta(n)}|^p \,\mu(dx) = \int_X \sum_{k\geq m} I_{A_k} I_{\Delta(n)} \,\mu(dx) = \mu\big(\Delta(n)\cap \bigcup_{k\geq m} A_k\big) \to 0$$

as $m \to \infty$.

Since $\Sigma \supset \Pi$, because Π is a π-system, it follows by Lemma 18 that $\Sigma \supset \mathfrak{A}_1$. Consequently, it follows from the definition of $L_p(\mathfrak{A}_1, \mu)$ that $I_{\Delta(n)} f \in L_p(\Pi, \mu)$ for $f \in L_p(\mathfrak{A}_1, \mu)$ and $n \geq 1$. Finally, a straightforward application of the dominated convergence theorem shows that $||I_{\Delta(n)} f - f||_p \to 0$ as $n \to \infty$. Hence $f \in L_p(\Pi, \mu)$ if $f \in L_p(\mathfrak{A}_1, \mu)$ and $L_p(\mathfrak{A}_1, \mu) \subset L_p(\Pi, \mu)$. Since the reverse inclusion is obvious, the theorem is proved.

It turns out that, under the conditions of Theorem 19, one can extend ζ from Π_0 to the larger set $\mathfrak{A}_0 := \sigma(\Pi) \cap \{\Gamma : \mu(\Gamma) < \infty\}$. Indeed, for $\Gamma \in \mathfrak{A}_0$ we have $I_\Gamma \in L_2(\Pi, \mu)$, so that the definition

$$\tilde{\zeta}(\Gamma) = \int_X I_\Gamma \,\zeta(dx)$$

makes sense. In addition, if $\Gamma_1, \Gamma_2 \in \mathfrak{A}_0$, then by (6)

$$E\tilde{\zeta}(\Gamma_1)\overline{\tilde{\zeta}(\Gamma_2)} = E\int_X I_{\Gamma_1}\,\zeta(dx)\overline{\int_X I_{\Gamma_2}\,\zeta(dx)}$$

$$= \int_X I_{\Gamma_1} I_{\Gamma_2}\,\mu(dx) = \mu(\Gamma_1 \cap \Gamma_2).$$

Since obviously $\zeta(\Delta) = \tilde{\zeta}(\Delta)$ (a.s.) for every $\Delta \in \Pi_0$, we have an extension indeed. In Sec. 7 we will see that sometimes one can extend ζ even to a larger set than $\mathfrak{A}_0$.

20. Exercise. Let $X \in \Pi_0$, and let Π be a π-system. Show that if $\tilde{\zeta}_1$ and $\tilde{\zeta}_2$ are two extensions of ζ to $\sigma(\Pi)$, then

$$\int_X f(x)\,\tilde{\zeta}_1(dx) = \int_X f(x)\,\tilde{\zeta}_2(dx)$$

(a.s.) for every $f \in L_2(\sigma(\Pi), \mu)$. In particular, $\tilde{\zeta}_1(\Gamma) = \tilde{\zeta}_2(\Gamma)$ (a.s.) for any $\Gamma \in \sigma(\Pi)$.

21. Exercise. Come back to Example 7. By what is said above there is an extension of ζ to $\mathfrak{B}([0,1])$. By using the independence of increments of w_t, prove that

$$E\exp(-\sum_n |\zeta((a_{n+1}, a_n])|) = 0,$$

where $a_n = 1/n$. Derive from here that for almost every ω the function $\zeta(\Gamma), \Gamma \in \mathfrak{B}([0,1])$, has unbounded variation and hence cannot be a measure.

Let us apply the above theory of stochastic integration to modeling Brownian motion when the temperature varies in time.

Take the objects introduced in Example 7. By Theorem 19 (and by Exercise 2), for every $f \in L_2(0,1)$ (where $L_2(0,1)$ is the usual L_2 space of square integrable functions on $(0,1)$) the stochastic integral $\int_X f(t)\,\zeta(dt)$ is well defined. Usually, one writes this integral as

$$\int_0^1 f(t)\,dw_t.$$

Observe that (by the continuity of the integral) if $f^n \to f$ in $L_2(0,1)$, then

$$\int_0^1 f^n(t)\,dw_t \to \int_0^1 f(t)\,dw_t$$

in the mean-square sense. In addition, if

$$f^n(t) = \sum_i f(t_{i+1,n}) I_{t_{in}<t\le t_{i+1,n}} = \sum_i f(t_{i+1,n})[I_{t\le t_{i+1,n}} - I_{t\le t_{in}}]$$

with $0 \le t_{in} \le t_{i+1,n} \le 1$, then (by definition and linearity)

$$\int_0^1 f(t)\,dw_t = \operatorname*{l.i.m.}_{n\to\infty} \int_0^1 f^n(t)\,dw_t = \operatorname*{l.i.m.}_{n\to\infty} \sum_i f(t_{i+1,n})(w_{t_{i+1,n}} - w_{t_{in}}). \tag{7}$$

Naturally, the integral

$$\int_0^t f(s)\,dw_s := \int_0^1 I_{s\le t} f(s)\,dw_s$$

gives us a representation of Brownian motion in the environment with changing temperature. However, for each individual t this integral is an element of $L_2(\mathcal{F}, P)$ and thus is uniquely defined only up to sets of probability zero. For describing individual trajectories of Brownian motion we should take an appropriate representative of $\int_0^t f(s)\,dw_s$ for each $t \in [0,1]$. At this moment it is absolutely not clear whether this choice can be performed so that we will have continuous trajectories, which is crucial from the practical point of view. Much later (see Theorem 6.1.10) we will prove that one can indeed make the right choice even when f is a random function. The good news is that this issue can be easily settled at least for some functions f.

22. Theorem. *Let $t \in [0,1]$, and let f be absolutely continuous on $[0,t]$. Then*

$$\int_0^t f(s)\,dw_s = f(t)w_t - \int_0^t w_s f'(s)\,ds \quad (a.s.).$$

Proof. Define $t_{in} = ti/n$. Then the functions $f^n(s) := f(t_{in})$ for $s \in (t_{in}, t_{i+1,n}]$ converge to $f(s)$ uniformly on $[0,t]$ so that (cf. (7)) we have

$$\int_0^t f(s)\,dw_s = \int_0^1 I_{s\le t} f(s)\,dw_s = \operatorname*{l.i.m.}_{n\to\infty} \sum_{i\le n-1} f(t_{in})(w_{t_{i+1,n}} - w_{t_{in}})$$

$$= f(t)w_t - \operatorname*{l.i.m.}_{n\to\infty} \sum_{i\le n-1} w_{t_{i+1,n}} \big(f(t_{i+1,n}) - f(t_{in})\big)$$

(summation by parts), where the last sum is written as

$$\int_0^t w_{\kappa(s,n)} f'(s)\,ds \tag{8}$$

with $\kappa(n,s) = t_{i+1,n}$ for $s \in (t_{in}, t_{i+1,n}]$. By the continuity of w_s we have $w_{\kappa(s,n)} \to w_s$ uniformly on $[0,t]$, and by the dominated convergence theorem (f' is integrable) we see that (8) converges to $\int_0^t w_s f'(s)\,ds$ for every ω. It only remains to remember that the mean-square limit coincides (a.s.) with the pointwise limit if both exist. The theorem is proved.

23. Exercise*. Prove that if a real-valued $f \in L_2(0,1)$, then $\int_0^t f(s)\,dw_s$, $t \in [0,1]$, is a Gaussian process with zero mean and covariance

$$R(s,t) = \int_0^{s\wedge t} f^2(u)\,du = \Big(\int_0^s f^2(u)\,du\Big) \wedge \Big(\int_0^t f^2(u)\,du\Big).$$

The construction of the stochastic integral with respect to a random orthogonal measure is not specific to probability theory. We have considered the case in which $\zeta(\Delta) \in L_2(\mathcal{F}, P)$, where P is a *probability* measure. Our arguments could be repeated almost word for word for the case of an arbitrary measure. It would then turn out that the Fourier integral of L_2 functions is a particular case of integrals with respect to random orthogonal measure. In this connection we offer the reader the following exercise.

24. Exercise. Let Π be the set of all intervals $(a,b]$, where $a,b \in (-\infty,\infty)$, $a < b$. For $\Delta = (a,b] \in \Pi$, define a function $\zeta(\Delta) = \zeta(\omega,\Delta)$ on $(-\infty,\infty)$ by

$$\zeta(\Delta) = \frac{1}{i\omega}\big(e^{i\omega b} - e^{i\omega a}\big) = \int_\Delta e^{i\omega x}\,dx.$$

Define $L_p = L_p(\Pi, \ell) = L_p(\mathfrak{B}(\mathbb{R}), \ell)$. Prove, using a change of variable, that the number $\big(\zeta(\Delta_1), \zeta(\Delta_2)\big)$ equals its complex conjugate, that is, it is real, and that $\|\zeta(\Delta)\|_2^2 = c\,\ell(\Delta)$ for $\Delta_1, \Delta_2, \Delta \in \Pi$, where c is a constant independent of Δ. Use this and the observation that $\zeta(\Delta_1 \cup \Delta_2) = \zeta(\Delta_1) + \zeta(\Delta_2)$ if $\Delta_1, \Delta_2, \Delta_1 \cup \Delta_2 \in \Pi$, $\Delta_1 \cap \Delta_2 = \emptyset$, to deduce that in that case $\big(\zeta(\Delta_1), \zeta(\Delta_2)\big) = 0$. Using the fact that $\Delta_1 = (\Delta_1 \setminus \Delta_2) \cup (\Delta_1 \cap \Delta_2)$ and adding an interval between Δ_1, Δ_2 if they do not intersect, prove that $\big(\zeta(\Delta_1), \zeta(\Delta_2)\big) = c\,\ell(\Delta_1 \cap \Delta_2)$ for every $\Delta_1, \Delta_2 \in \Pi$ and, consequently, that we can construct an integral with respect to ζ, such that *Parseval's equality* holds for every $f \in L_2$:

$$c\|f\|_2^2 = \Big\| \int f\zeta(dx) \Big\|_2^2.$$

Keeping in mind that for $f \in S(\Pi)$, obviously,

$$\int f\zeta(dx) = \int_{-\infty}^{\infty} f(x)e^{i\omega x}\,dx \quad \text{(a.e.)},$$

generalize this equality to all $f \in L_2 \cap L_1$. Putting $f = \exp(-x^2)$ and using the characteristic function of the normal distribution, prove that $c = 2\pi$. Finally, use Fubini's theorem to prove that for $f \in L_1$ and $-\infty < a < b < \infty$, we have

$$\int_a^b \Big(\int_{-\infty}^{\infty} \bar f(\omega) e^{i\omega x}\,d\omega \Big) dx = \int_{-\infty}^{\infty} \frac{1}{i\omega}\big(e^{i\omega b} - e^{i\omega a}\big) \bar f(\omega)\,d\omega.$$

In other words, if $f \in L_1 \cap L_2$, then $\big(\zeta(\Delta), f\big) = c(I_\Delta, g)$, where

$$\bar g(x) = c^{-1} \int \bar f(\omega)\zeta(x, d\omega),$$

and (by definition) this leads to the *inversion formula for the Fourier transform*:

$$f(\omega) = \int g(x)\zeta(\omega, dx).$$

Generalize this formula from the case $f \in L_1 \cap L_2$ to all $f \in L_2$.

4. The Wiener process on $[0,\infty)$

The definition of the Wiener process on $[0,\infty)$ is the same as on $[0,1]$ (cf. Definition 1.5). Clearly for the Wiener process on $[0,\infty)$ one has the corresponding counterparts of Theorems 2.1 and 2.2 about the independence of increments and the independence of increments of previous values of the process. Also as in Exercise 1.6, if w_t is a Wiener process on $[0,\infty)$ and c is a strictly positive constant, then cw_{t/c^2} is also a Wiener process on $[0,\infty)$. This property is called *self-similarity* of the Wiener process.

1. Theorem. *There exists a Wiener process defined on* $[0,\infty)$.

Proof. Take any smooth function $f(t)>0$ on $[0,1)$ such that

$$\int_0^1 f^2(t)\,dt=\infty.$$

Let $\varphi(r)$ be the inverse function to $\int_0^t f^2(s)\,ds$. For $t<1$ define

$$y(t)=f(t)w_t-\int_0^t w_s f'(s)\,ds.$$

Obviously $y(t)$ is a continuous process. By Theorem 3.22 we have

$$y(t)=\int_0^t f(s)\,dw_s=\int_0^1 I_{s\le t}f(s)\,dw_s \quad \text{(a.s.)}.$$

By Exercise 3.23, y_t is a Gaussian process with zero mean and covariance

$$\int_0^{s\wedge t} f^2(u)\,du=\Big(\int_0^s f^2(u)\,du\Big)\wedge\Big(\int_0^t f^2(u)\,du\Big),\quad s,t<1.$$

Now, as is easy to see, $x(r):=y(\varphi(r))$ is a continuous Gaussian process defined for $r\in[0,\infty)$ with zero mean and covariance $r_1\wedge r_2$. The theorem is proved.

Apart from the properties of the Wiener process on $[0,\infty)$ stated in the beginning of this section, which are similar to the properties on $[0,1]$, there are some new ones, of which we will state and prove only two.

2. Theorem. *Let w_t be a Wiener process for $t\in[0,\infty)$ defined on a probability space $(\Omega,\mathcal{F},P)$. Then there exists a set $\Omega'\in\mathcal{F}$ such that $P(\Omega')=1$ and, for each $\omega\in\Omega'$, we have*

$$\lim_{t\downarrow 0} t w_{1/t}(\omega)=0.$$

Furthermore, for $t>0$ define

$$\xi_t(\omega)=\begin{cases} tw_{1/t}(\omega) & \text{if} \quad \omega\in\Omega', \\ 0 & \text{if} \quad \omega\notin\Omega', \end{cases}$$

and let $\xi_0(\omega)\equiv 0$. *Then* ξ_t *is a Wiener process.*

Proof. Define $\tilde{\xi}_t = tw_{1/t}$ for $t>0$ and $\tilde{\xi}_0\equiv 0$. As is easy to see, $\tilde{\xi}_t$ is a Gaussian process with zero mean and covariance $s\wedge t$. It is also continuous on $(0,\infty)$. It follows, in particular, that $\sup_{s\in(0,t]}|\tilde{\xi}_s(\omega)|$ equals the sup over rational numbers on $(0,t]$. Since this sup is an increasing function of t, its limit as $t\downarrow 0$ can also be calculated along rational numbers. Thus,

$$\Omega' := \{\omega : \lim_{t\downarrow 0}\sup_{s\in(0,t]}|\tilde{\xi}_s(\omega)|=0\}\in\mathcal{F}.$$

Next, let C' be the set of all (maybe unbounded) continuous functions on $(0,1]$, and $\Sigma(C')$ the cylinder σ-field of subsets of C', that is, the smallest σ-field containing all sets $\{x_\cdot\in C' : x_t\in\Gamma\}$ for all $t\in(0,1]$ and $\Gamma\in\mathfrak{B}(\mathbb{R})$. Then the distributions of $\tilde{\xi}_\cdot$ and $w_\cdot$ on $(C',\Sigma(C'))$ coincide (cf. Remark 1.4).

Define

$$A=\{x_\cdot\in C' : \lim_{t\downarrow 0}\sup_{s\in(0,t]}|x_s|=0\}.$$

Since $x_\cdot\in C'$ are continuous in $(0,1]$, it is easy to see that $A\in\Sigma(C')$. Therefore,

$$P(\tilde{\xi}_\cdot\in A)=P(w_\cdot\in A),$$

which is to say,

$$P(\lim_{t\downarrow 0}\sup_{s\in(0,t]}|\tilde{\xi}_s|=0)=P(\lim_{t\downarrow 0}\sup_{s\in(0,t]}|w_s|=0).$$

The last probability being 1, we conclude that $P(\Omega')=1$, and it only remains to observe that ξ_t is a continuous process and $\xi_t=\tilde{\xi}_t$ on Ω' or almost surely, so that ξ_t is a Gaussian process with zero mean and covariance $s\wedge t$. The theorem is proved.

3. Corollary. *Let* $1/2>\varepsilon>0$. *By Theorem* 2.4 *for almost every* ω *there exists* $n(\omega)<\infty$ *such that* $|\xi_t(\omega)|\le Nt^{1/2-\varepsilon}$ *for* $t\le 2^{-n(\omega)}$, *where* N *depends only on* ε. *Hence, for* w_t, *for almost every* ω *we have* $|w_t|\le Nt^{1/2+\varepsilon}$ *if* $t\ge 2^{n(\omega)}$.

4. Remark. Having the Wiener process on $[0,\infty)$, we can repeat the construction of the stochastic integral and define $\int_0^\infty f(t)\,dw_t$ for every $f\in L_2([0,\infty))$ starting with the random orthogonal measure $\zeta(0,a]=w_a$ defined for all $a\geq 0$. Of course, this integral has properties similar to those of $\int_0^1 f(t)\,dw_t$. In particular, the results of Theorem 3.22 on integrating by parts and of Exercise 3.23 still hold.

5. Markov and strong Markov properties of the Wiener process

Let $(\Omega,\mathcal{F},P)$ be a probability space carrying a Wiener process w_t, $t\in[0,\infty)$. Also assume that for every $t\in[0,\infty)$ we are given a σ-field $\mathcal{F}_t\subset\mathcal{F}$ such that $\mathcal{F}_s\subset\mathcal{F}_t$ for $t\geq s$. We call such a collection of σ-fields an (increasing) *filtration of σ-fields*.

A trivial example of filtration is given by $\mathcal{F}_t\equiv\mathcal{F}$.

1. Definition. Let Σ be a σ-field, $\Sigma\subset\mathcal{F}$ and ξ a random variable taking values in a measurable space $(X,\mathfrak{B})$. We say that ξ and Σ are *independent* if $P(A,\xi\in B)=P(A)P(\xi\in B)$ for every $A\in\Sigma$ and $B\in\mathfrak{B}$.

2. Exercise*. Prove that if ξ and Σ are independent, $f(x)$ is a measurable function, and η is Σ-measurable, then $f(\xi)$ and η are independent as well.

3. Definition. We say that w_t is a Wiener process relative to the filtration $\mathcal{F}_t$ if w_t is $\mathcal{F}_t$-measurable for every t and $w_{t+h}-w_t$ is independent of $\mathcal{F}_t$ for every $t,h\geq 0$. In that case the couple $(w_t,\mathcal{F}_t)$ is called *a Wiener process*.

Below we assume that $(w_t,\mathcal{F}_t)$ is a Wiener process, explaining first that there always exists a filtration with respect to which w_t is a Wiener process.

4. Lemma. *Let*

$$\mathcal{F}_t^w:=\sigma\{\{\omega: w_s(\omega)\in B\}, s\leq t, B\in\mathfrak{B}(\mathbb{R})\}.$$

Then $(w_t,\mathcal{F}_t^w)$ is a Wiener process.

Proof. By definition $\mathcal{F}_t^w$ is the smallest σ-field containing all sets $\{\omega: w_s(\omega)\in B\}$ for $s\leq t$ and Borel B. Since each of them is (as an element) in $\mathcal{F}$, $\mathcal{F}_t^w\subset\mathcal{F}$. The inclusion $\mathcal{F}_s^w\subset\mathcal{F}_t^w$ for $t\geq s$ is obvious, since $\{\omega: w_r(\omega)\in B\}$ belong to $\mathcal{F}_t^w$ for $r\leq s$ and $\mathcal{F}_s^w$ is the smallest σ-field containing them. Therefore $\mathcal{F}_t^w$ is a filtration.

Next, $\{\omega: w_t(\omega)\in B\}\in\mathcal{F}_t^w$ for $B\in\mathfrak{B}(\mathbb{R})$; hence w_t is $\mathcal{F}_t^w$-measurable. To prove the independence of $w_{t+h}-w_t$ and $\mathcal{F}_t^w$, fix a $B\in\mathfrak{B}(\mathbb{R})$, $t,h\geq 0$, and define

$$\mu(A)=P(A,w_{t+h}-w_t\in B),\quad \nu(A)=P(A)P(w_{t+h}-w_t\in B).$$

One knows that μ and ν are measures on $(\Omega, \mathcal{F})$. By Theorem 2.2 these measures coincide on every A of type $\{\omega : (w_{t_1}(\omega), ..., w_{t_n}(\omega)) \in B^{(n)}\}$ provided that $t_i \leq t$ and $B^{(n)} \in \mathfrak{B}(\mathbb{R}^n)$. The collection of these sets is an algebra (Exercise 1.3.3). Therefore μ and ν coincide on the smallest σ-field, say Σ, containing these sets. Observe that $\mathcal{F}_t^w \subset \Sigma$, since the collection generating Σ contains $\{\omega : w_s(\omega) \in D\}$ for $s \leq t$ and $D \in \mathfrak{B}(\mathbb{R})$. Hence μ and ν coincide on $\mathcal{F}_t^w$. It only remains to remember that B is an arbitrary element of $\mathfrak{B}(\mathbb{R})$. The lemma is proved.

We see that one can always take $\mathcal{F}_t^w$ as $\mathcal{F}_t$. However, it turns out that sometimes it is very inconvenient to restrict our choice of $\mathcal{F}_t$ to $\mathcal{F}_t^w$. For instance, we can be given a multi-dimensional Wiener process $(w_t^1, ..., w_t^d)$ (see Definition 6.4.1) and study only its first coordinate. In particular, while introducing stochastic integrals of random processes against dw_t^1 we may be interested in integrating functions depending not only on w_t^1 but on all other components as well.

5. Exercise*. Let $\bar{\mathcal{F}}_t^w$ be the completion of $\mathcal{F}_t^w$. Prove that $(w_t, \bar{\mathcal{F}}_t^w)$ is a Wiener process.

6. Theorem (Markov property)**.** *Let $(w_t, \mathcal{F}_t)$ be a Wiener process. Fix t, $h_1, .., h_n \geq 0$. Then the vector $(w_{t+h_1} - w_t, ..., w_{t+h_n} - w_t)$ and the σ-field $\mathcal{F}_t$ are independent. Furthermore, $w_{t+s} - w_t$, $s \geq 0$, is a Wiener process.*

Proof. The last statement follows directly from the definitions. To prove the first one, without losing generality we assume that $h_1 \leq ... \leq h_n$ and notice that, since $(w_{t+h_1} - w_t, ..., w_{t+h_n} - w_t)$ is obtained by a linear transformation from η_n, where $\eta_k = (w_{t+h_1} - w_{t+h_0}, ..., w_{t+h_k} - w_{t+h_{k-1}})$ and $h_0 = 0$, we need only show that η_n and $\mathcal{F}_t$ are independent. We are going to use the theory of characteristic functions. Take $A \in \mathcal{F}_t$ and a vector $\lambda \in \mathbb{R}^n$. Notice that

$$EI_A \exp(i\lambda \cdot \eta_n) = EI_A \exp(i\mu \cdot \eta_{n-1}) \exp(i\lambda^n (w_{t+h_n} - w_{t+h_{n-1}})),$$

where $\mu = (\lambda^1, ..., \lambda^{n-1})$. Here I_A is $\mathcal{F}_t$-measurable and, since $\mathcal{F}_t \subset \mathcal{F}_{t+h_{n-1}}$, it is $\mathcal{F}_{t+h_{n-1}}$-measurable as well. It follows that $I_A \exp(i\mu \cdot \eta_{n-1})$ is $\mathcal{F}_{t+h_{n-1}}$-measurable. Furthermore, $w_{t+h_n} - w_{t+h_{n-1}}$ is independent of $\mathcal{F}_{t+h_{n-1}}$. Hence, by Exercise 2

$$EI_A \exp(i\lambda \cdot \eta_n) = EI_A \exp(i\mu \cdot \eta_{n-1}) E \exp(i\lambda^n (w_{t+h_n} - w_{t+h_{n-1}})),$$

and by induction and independence of increments of w_t

$$EI_A \exp(i\lambda \cdot \eta_n) = EI_A \prod_{j=1}^{n} E \exp(i\lambda^n (w_{t+h_j} - w_{t+h_{j-1}})) = P(A) E \exp(i\lambda \cdot \eta_n).$$

It follows from the theory of characteristic functions that for every Borel bounded g

$$EI_A g(\eta_n) = P(A)Eg(\eta_n).$$

It only remains to substitute here the indicator of a Borel set in place of g. The theorem is proved.

Theorem 6 says that, for every fixed $t \geq 0$, the process $w_{t+s} - w_t$, $s \geq 0$, starts afresh as a Wiener process forgetting everything that happened to w_r before time t. This property is quite natural for Brownian motion. It also has a natural extension when t is replaced with a random time τ, provided that τ does not depend on the future in a certain sense. To describe exactly what we mean by this, we need the following.

7. Definition. Let τ be a random variable taking values in $[0, \infty]$ (including ∞). We say that τ is a *stopping time* (relative to $\mathcal{F}_t$) if $\{\omega : \tau(\omega) > t\} \in \mathcal{F}_t$ for every $t \in [0, \infty)$.

The term "stopping time" is discussed after Exercise 3.3.3. Trivial examples of stopping times are given by nonrandom positive constants. A much more useful example is the following.

8. Example. Fix $a \geq 0$ and define

$$\tau = \tau_a = \inf\{t \geq 0 : w_t \geq a\} \quad (\inf \emptyset := \infty)$$

as the first hitting time of the point a by w_t. It turns out that τ is a stopping time.

Indeed, one can easily see that

$$\{\omega : \tau(\omega) > t\} = \{\omega : \max_{s \leq t} w_s(\omega) < a\}, \tag{1}$$

where, for ρ defined as the set of all rational points on $[0, \infty)$,

$$\max_{s \leq t} w_s = \sup_{r \in \rho, r \leq t} w_r,$$

which shows that $\max_{s \leq t} w_s$ is an $\mathcal{F}_t$-measurable random variable.

9. Exercise*. Let $a < 0 < b$ and let τ be the first exit time of w_t from (a, b):

$$\tau = \inf\{t \geq 0 : w_t \notin (a, b)\}.$$

Prove that τ is a stopping time.

10. Definition. Random processes $\eta^1_t, ..., \eta^n_t$ defined for $t \geq 0$ are called *independent* if for every $t_1, ..., t_k \geq 0$ the vectors $(\eta^1_{t_1}, ..., \eta^1_{t_k}), ..., (\eta^n_{t_1}, ..., \eta^n_{t_k})$ are independent.

In what follows we consider some processes at random times, and these times occasionally can be infinite even though this happens with probability zero. In such situations we use the notation

$$x_\tau = x_\tau(\omega) = \begin{cases} x_{\tau(\omega)}(\omega) & \text{if} \quad \tau(\omega) < \infty, \\ 0 & \text{if} \quad \tau(\omega) = \infty. \end{cases}$$

11. Lemma. *Let $(w_t, \mathcal{F}_t)$ be a Wiener process and let τ be an $\mathcal{F}_t$-stopping time. Assume $P(\tau < \infty) = 1$. Then the processes $w_{t\wedge\tau}$ and $B_t := w_{\tau+t} - w_\tau$ are independent and the latter one is a Wiener process.*

Proof. Take $0 \leq t_1 \leq ... \leq t_k$. As is easy to see, we need only prove that for any Borel nonnegative functions $f(x_1, ..., x_k)$ and $g(x_1, ..., x_k)$

$$I_\tau := Ef(w_{t_1\wedge\tau}, ..., w_{t_k\wedge\tau})g(B_{t_1}, ..., B_{t_k})$$

$$= Ef(w_{t_1\wedge\tau}, ..., w_{t_k\wedge\tau})Eg(w_{t_1}, ..., w_{t_k}). \tag{2}$$

Assume for a moment that the set of values of τ is countable, say $r_1 < r_2 < ....$ By noticing that $\{\tau = r_n\} = \{\tau > r_{n-1}\} \setminus \{\tau > r_n\} \in \mathcal{F}_{r_n}$ and

$$F_n := f(w_{t_1\wedge\tau}, ..., w_{t_k\wedge\tau})I_{\tau=r_n} = f(w_{t_1\wedge r_n}, ..., w_{t_k\wedge r_n})I_{\tau=r_n},$$

we see that the first term is $\mathcal{F}_{r_n}$-measurable. Furthermore,

$$I_{\tau=r_n}g(B_{t_1}, ..., B_{t_k}) = I_{\tau=r_n}g(w_{r_n+t_1} - w_{r_n}, ..., w_{r_n+t_k} - w_{r_n}),$$

where, by Theorem 6, the last factor is independent of $\mathcal{F}_{r_n}$, and

$$Eg(w_{r_n+t_1} - w_{r_n}, ..., w_{r_n+t_k} - w_{r_n}) = Eg(w_{t_1}, ..., w_{t_k}).$$

Therefore,

$$I_\tau = \sum_{r_n} EF_n g(w_{r_n+t_1} - w_{r_n}, ..., w_{r_n+t_k} - w_{r_n})$$

$$= Eg(w_{t_1}, ..., w_{t_k}) \sum_{r_n} EF_n.$$

The last sum equals the first term on the right in (2). This proves the theorem for our particular τ.

In the general case we approximate τ and first notice (see, for instance, Theorem 1.2.4) that equation (2) holds for all Borel nonnegative f, g if and only if it holds for all bounded continuous f, g. Therefore, we assume f, g to be bounded and continuous.

Now, for $n = 1, 2, ...$, define

$$\tau_n(\omega) = (k+1)2^{-n} \quad \text{for } \omega \text{ such that} \quad k2^{-n} < \tau(\omega) \leq (k+1)2^{-n}, \qquad (3)$$

$k = -1, 0, 1, ...$ It is easily seen that $\tau \leq \tau_n \leq \tau + 2^{-n}$, $\tau_n \downarrow \tau$, and for $t \geq 0$

$$\{\omega : \tau_n > t\} = \{\omega : \tau(\omega) > 2^{-n}[2^n t]\} \in \mathcal{F}_{2^{-n}[2^n t]} \subset \mathcal{F}_t,$$

so that τ_n are stopping times. Hence, by the above result,

$$I_\tau = \lim_{n\to\infty} I_{\tau_n} = Eg(w_{t_1}, ..., w_{t_k}) \lim_{n\to\infty} Ef(w_{t_1\wedge\tau_n}, ..., w_{t_k\wedge\tau_n}),$$

and this leads to (2). The lemma is proved.

The following theorem states that the Wiener process has the strong Markov property.

12. Theorem. *Let $(w_t, \mathcal{F}_t)$ be a Wiener process and τ an $\mathcal{F}_t$-stopping time. Assume that $P(\tau < \infty) = 1$. Let*

$$\mathcal{F}^w_{\leq\tau} = \sigma\{\{\omega : w_{s\wedge\tau} \in B\}, s \geq 0, B \in \mathfrak{B}(\mathbb{R})\},$$

$$\mathcal{F}^w_{\geq\tau} = \sigma\{\{\omega : w_{\tau+s} - w_\tau \in B\}, s \geq 0, B \in \mathfrak{B}(\mathbb{R})\}.$$

Then the σ-fields $\mathcal{F}^w_{\leq\tau}$ and $\mathcal{F}^w_{\geq\tau}$ are independent in the sense that for every $A \in \mathcal{F}^w_{\leq\tau}$ and $B \in \mathcal{F}^w_{\geq\tau}$ we have $P(AB) = P(A)P(B)$. Furthermore, $w_{\tau+t} - w_\tau$ is a Wiener process.

Proof. The last assertion is proved in Lemma 11. To prove the first one we follow the proof of Lemma 4 and first let $B = \{\omega : (w_{\tau+s_1} - w_\tau, ..., w_{\tau+s_k} - w_\tau) \in \Gamma\}$, where $\Gamma \in \mathfrak{B}(\mathbb{R}^k)$. Consider two measures $\mu(A) = P(AB)$ and $\nu(A) = P(A)P(B)$ as measures on sets A. By Lemma 11 these measures coincide on every A of type $\{\omega : (w_{t_1\wedge\tau}, ..., w_{t_n\wedge\tau}) \in B^{(n)}\}$ provided that $B^{(n)} \in \mathfrak{B}(\mathbb{R}^n)$. The collection of these sets is an algebra (Exercise 1.3.3). Therefore μ and ν coincide on the smallest σ-field, which is $\mathcal{F}^w_{\leq\tau}$, containing these sets. Hence $P(AB) = P(A)P(B)$ for all $A \in \mathcal{F}^w_{\leq\tau}$ and our particular B. It only remains to repeat this argument relative to B upon fixing A. The theorem is proved.

6. Examples of applying the strong Markov property

First, we want to apply Theorem 5.12 to τ_a from Example 5.8. Notice that Bachelier's Theorem 2.3 holds not only for $t \in (0,1]$ but for $t \geq 1$ as well. One proves this by using the self-similarity of the Wiener process (cw_{t/c^2} is a Wiener process for every constant $c \neq 0$). Then, owing to (5.1), for $t > 0$ we find that $P(\tau_a > t) = P(|w_t| < a) = P(|w_1|\sqrt{t} < a)$, which tends to zero as $t \to \infty$, showing that $P(\tau_a < \infty) = 1$. Now Theorem 5.12 allows us to conclude that $w_{\tau+t} - w_\tau = w_{\tau+t} - a$ is a Wiener process independent of the trajectory on $[0, \tau]$. This makes rigorous what is quite clear intuitively. Namely, after reaching a, the Wiener process starts "afresh", forgetting everything which happened to it before. The same happens when it reaches a higher level $b > a$ after reaching a, and moreover, $\tau_b - \tau_a$ has the same distribution as τ_{b-a}. This is part of the following theorem, in which, as well as above, we allow ourselves to consider random variables like $\tau_b - \tau_a$ which may not be defined on a set of probability zero. We set $\tau_b(\omega) - \tau_a(\omega) = 0$ if $b > a > 0$ and $\tau_b(\omega) = \tau_a(\omega) = \infty$.

1. Theorem. (i) *For every* $0 < a_1 < a_2 < ... < a_n < \infty$ *the random variables* $\tau_{a_1}, \tau_{a_2} - \tau_{a_1}, ..., \tau_{a_n} - \tau_{a_{n-1}}$ *are independent.*

(ii) *For* $0 < a < b$, *the law of* $\tau_b - \tau_a$ *coincides with that of* τ_{b-a}, *and* τ_a *has* Wald's distribution *with density*

$$p(t) = (2\pi)^{-1/2} a t^{-3/2} \exp(-a^2/(2t)), \quad t > 0.$$

Proof. (i) It suffices to prove that $\tau_{a_n} - \tau_{a_{n-1}}$ is independent of $\tau_{a_1}, ...,$ $\tau_{a_{n-1}}$ (cf. the proof of Theorem 2.2). To simplify notation, put $\tau(a) = \tau_a$. Since $a_i \leq a_{n-1}$ for $i \leq n-1$, we can rewrite (5.1) as

$$\{\omega : \tau(a_i) > t\} = \{\omega : \sup_{s \in \rho, s \leq t} w_{s \wedge \tau(a_{n-1})} < a_i\},$$

which implies that the $\tau(a_i)$ are $\mathcal{F}_{\leq \tau(a_{n-1})}$-measurable. On the other hand, for $t \geq 0$,

$$\{\omega : \tau(a_n) - \tau(a_{n-1}) > t\}$$
$$= \{\omega : \tau(a_n) - \tau(a_{n-1}) > t, \tau(a_{n-1}) < \infty\}$$
$$= \{\omega : \sup_{s \in \rho, s \leq t} (w_{\tau(a_{n-1})+s} - w_{\tau(a_{n-1})}) < a_n - a_{n-1}, \tau(a_{n-1}) < \infty\}$$
$$= \{\omega : 0 < \sup_{s \in \rho, s \leq t} (w_{\tau(a_{n-1})+s} - w_{\tau(a_{n-1})}) < a_n - a_{n-1}\}, \qquad (1)$$

which shows that $\tau(a_n) - \tau(a_{n-1})$ is $\mathcal{F}_{\geq \tau(a_{n-1})}$-measurable. Referring to Theorem 5.12 finishes the proof of (i).

(ii) Let $n=2$, $a_1=a$, and $a_2=b$. Then in the above notation $\tau(a_n)=\tau_b$ and $\tau(a_{n-1})=\tau_a$. Since $w_{\tau(a_{n-1})+t}-w_{\tau(a_{n-1})}=w_{\tau_a+t}-w_{\tau_a}$ is a Wiener process and the distributions of Wiener processes coincide, the probability of the event on the right in (1) equals

$$P(\sup_{s\in\rho,s\leq t} w_s<a_n-a_{n-1}=b-a)=P(\tau_{b-a}>t).$$

This proves the first assertion in (ii). To find the distribution of τ_a, remember that

$$P(\tau_a>t)=P(\max_{s\leq t} w_s<a)=P(|w_1|\sqrt{t}<a)=\frac{2}{\sqrt{2\pi}}\int_0^{a/\sqrt{t}} e^{-y^2/2}\,dy.$$

By differentiating this formula we immediately get our density. The theorem is proved.

2. Exercise. We know that the Wiener process is self-similar in the sense that cw_{t/c^2} is a Wiener process for every constant $c\neq 0$. The process τ_a, $a\geq 0$, also has this kind of property. Prove that, for every $c>0$, the process $c\tau_{a/\sqrt{c}}$, $a\geq 0$, has the same finite-dimensional distributions as τ_a, $a\geq 0$. Such processes are called *stable*. The Wiener process is a stable process of order 2, and the process τ_a is a stable process of order 1/2.

Our second application exhibits the importance of the operator $u\to u''$ in computing various expectations related to the Wiener process. The following results can be obtained quite easily on the basis of Itô's formula from Chapter 6. However, the reader might find it instructive to see that there is a different approach using the strong Markov property.

3. Lemma. *Let u be a twice continuously differentiable function defined on $\mathbb{R}$ such that u, u', and u'' are bounded. Then, for every $\lambda>0$,*

$$u(0)=E\int_0^\infty e^{-\lambda t}(\lambda u(w_t)-(1/2)u''(w_t))\,dt. \tag{2}$$

Proof. Since w_t is a normal $(0,t)$ variable, the right-hand side of (2) equals

$$I:=\int_0^\infty e^{-\lambda t}E(\lambda u(w_t)-(1/2)u''(w_t))\,dt$$
$$=\int_0^\infty e^{-\lambda t}\Big(\int_{\mathbb{R}}(\lambda u(x)-(1/2)u''(x))p(t,x)\,dx\Big)\,dt,$$

where

$$p(t,x):=\frac{1}{\sqrt{2\pi t}}e^{-x^2/(2t)},\quad t>0.$$

We continue our computation, integrating by parts. One can easily check that

$$\frac{1}{2}\frac{\partial^2 p}{(\partial x)^2} = \frac{\partial p}{\partial t}, \quad e^{-\lambda t}\lambda p - \frac{e^{-\lambda t}}{2}\frac{\partial^2 p}{(\partial x)^2} = -\frac{\partial}{\partial t}(e^{-\lambda t}p).$$

Hence

$$I = \lim_{\varepsilon \downarrow 0} \int_\varepsilon^\infty e^{-\lambda t}\Big(\int_{\mathbb{R}} (\lambda u(x) - (1/2)u''(x))p(t,x)\,dx\Big)\,dt$$

$$= -\lim_{\varepsilon \downarrow 0} \int_\varepsilon^\infty \frac{\partial}{\partial t}\Big(e^{-\lambda t}\int_{\mathbb{R}} u(x)p(t,x)\,dx\Big)\,dt = \lim_{\varepsilon \downarrow 0} e^{-\lambda\varepsilon}\int_{\mathbb{R}} u(x)p(\varepsilon,x)\,dx$$

$$= \lim_{\varepsilon \downarrow 0} Eu(w_\varepsilon) = u(0).$$

The lemma is proved.

4. Theorem. *Let* $-\infty < a < 0 < b < \infty$, *and let* u *be a twice continuously differentiable function given on* $[a,b]$. *Let* τ *be the first exit time of* w_t *from the interval* (a,b) (*see Exercise* 5.9). *Then, for every* $\lambda \geq 0$,

$$u(0) = E\int_0^\tau e^{-\lambda t}(\lambda u(w_t) - (1/2)u''(w_t))\,dt + Ee^{-\lambda\tau}u(w_\tau). \tag{3}$$

Proof. If needed, one can continue u outside $[a,b]$ and have a function, for which we keep the same notation, satisfying the assumptions of Lemma 3. Denote $f = \lambda u - u''$. Notice that obviously $\tau \leq \tau_b$, and, as we have seen above, $P(\tau_b < \infty) = 1$. Therefore by Lemma 3 we find that, for $\lambda > 0$,

$$u(0) = E\int_0^\infty ... = E\int_0^\tau ... + E\int_\tau^\infty ...$$

$$= E\int_0^\tau e^{-\lambda t}f(w_t)\,dt + \int_0^\infty e^{-\lambda t}Ee^{-\lambda\tau}f(w_\tau + B_t)\,dt =: I + J,$$

where $B_t = w_{\tau+t} - w_\tau$. Now we want to use Theorem 5.12. The reader who did Exercise 5.9 understands that τ is $\mathcal{F}^w_{\leq\tau}$-measurable. Furthermore, $w_{t\wedge\tau}I_{\tau<\infty} \to w_\tau$ as $t \to \infty$, so that w_τ is also $\mathcal{F}^w_{\leq\tau}$-measurable. Hence (τ, w_τ) and B_t are independent, and

$$J = \int_0^\infty e^{-\lambda t}Ee^{-\lambda\tau}f(w_\tau + B_t)\,dt = Ee^{-\lambda\tau}v(w_\tau),$$

where

$$v(y) := E\int_0^\infty e^{-\lambda t}f(y + B_t)\,dt = E\int_0^\infty e^{-\lambda t}f(y + w_t)\,dt.$$

Upon applying Lemma 3 to $u(x+y)$ in place of $u(x)$, we immediately get that $v = u$, and this proves the theorem if $\lambda > 0$.

To prove (3) for $\lambda = 0$ it suffices to pass to the limit, which is possible due to the dominated convergence theorem if we know that $E\tau < \infty$. However, for the function $u_0(x) = (x-a)(b-x)$ and the result for $\lambda > 0$, we get

$$|a|b = u_0(0) = E\int_0^\tau e^{-\lambda t}(\lambda u(w_t) + 1)\,dt \geq E\int_0^\tau e^{-\lambda t}\,dt,$$

$$E\int_0^\tau e^{-\lambda t}\,dt \leq |a|b$$

and it only remains to apply the monotone convergence theorem to get $E\tau \leq |a|b < \infty$. The theorem is proved.

In the following exercises we suggest the reader use Theorem 4.

5. Exercise. (i) Prove that $E\tau = |a|b$.

(ii) By noticing that

$$Eu(w_\tau) = u(b)P(\tau = \tau_b) + u(a)P(\tau < \tau_b)$$

and taking an appropriate function u, show that the probability that the Wiener process hits b before hitting a is $|a|/(|a|+b)$.

6. Exercise. Sometimes one is interested in knowing how much time the Wiener process spends in a subinterval $[c,d] \subset (a,b)$ before exiting from (a,b). Of course, by this time we mean Lebesgue measure of the set $\{t < \tau : w_t \in [c,d]\}$.

(i) Prove that this time equals

$$\gamma := \int_0^\tau I_{[c,d]}(w_t)\,dt.$$

(ii) Prove that for any Borel nonnegative f we have

$$E\int_0^\tau f(w_t)\,dt = \frac{2}{b-a}\big(b\int_a^0 f(y)(y-a)\,dy - a\int_0^b f(y)(b-y)\,dy\big),$$

and find $E\gamma$.

7. Exercise. Define $x_t = w_t + t$, and find the probability that x_t hits b before hitting a.

7. Itô stochastic integral

In Sec. 3 we introduced the stochastic integral of nonrandom functions on $[0,1]$ against dw_t. It turns out that a slight modification of this procedure allows one to define stochastic integrals of random functions as well. The way we proceed is somewhat different from the traditional one, which will be presented in Sec. 6.1. We decided to give this definition just in case the reader decides to study stochastic integration with respect to arbitrary square integrable martingales.

Let $(w_t, \mathcal{F}_t)$ be a Wiener process in the sense of Definition 5.3, given on a probability space $(\Omega, \mathcal{F}, P)$. To proceed with defining Itô stochastic integral in the framework of Sec. 3 we take

$$X = \Omega \times (0, \infty), \quad \mathfrak{A} = \mathcal{F} \otimes \mathfrak{B}((0, \infty)), \quad \mu = P \times \ell \tag{1}$$

and define Π as the collection of all sets $A \times (s, t]$ where $0 \le s \le t < \infty$ and $A \in \mathcal{F}_s$. Notice that, for $A \times (s, t] \in \Pi$,

$$\mu(A \times (s, t]) = P(A)(t - s) < \infty,$$

so that $\Pi_0 = \Pi$. For $A \times (s, t] \in \Pi$ let

$$\zeta(A \times (s, t]) = (w_t - w_s) I_A.$$

1. Definition. Denote $\mathcal{P} = \sigma(\Pi)$ and call $\mathcal{P}$ the σ-field of *predictable sets*. The functions on $\Omega \times (0, \infty)$ which are $\mathcal{P}$-measurable are called *predictable* (relative to $\mathcal{F}_t$).

By the way, the name "predictable" comes from the observation that the simplest $\mathcal{P}$-measurable functions are indicators of elements of Π which have the form $I_A I_{(s,t]}$ and are left-continuous, thus predictable on the basis of past observations, functions of time.

2. Exercise*. Prove that Π is a π-system, and by relying on Theorem 3.19 conclude that $L_2(\Pi, \mu) = L_2(\mathcal{P}, \mu)$.

3. Theorem. *The function ζ on Π is a random orthogonal measure with reference measure μ, and $E\zeta(\Delta) = 0$ for every $\Delta \in \Pi$.*

Proof. We have to check the conditions of Definition 3.5. Let $\Delta_1 = A_1 \times (t_1, t_2]$, $\Delta_2 = A_2 \times (s_1, s_2] \in \Pi$. Define

$$f_t(\omega) = I_{\Delta_1}(\omega, t) + I_{\Delta_2}(\omega, t)$$

and introduce the points $r_1 \leq ... \leq r_4$ by ordering t_1, t_2, s_1, and s_2. Obviously, for every $t \geq 0$, the functions $I_{\Delta_i}(\omega, t+)$ are $\mathcal{F}_t$-measurable and the same holds for $f_{t+}(\omega)$. Furthermore, for each ω, $f_t(\omega)$ is piecewise constant and left continuous in t. Therefore,

$$f_t(\omega) = \sum_{i=1}^{3} g_i(\omega) I_{(r_i, r_{i+1}]}(t), \tag{2}$$

where the $g_i = f_{r_i+}$ are $\mathcal{F}_{r_i}$-measurable.

It turns out that for every ω

$$\zeta(\Delta_1) + \zeta(\Delta_2) = \sum_{i=1}^{3} g_i(\omega)(w_{r_{i+1}} - w_{r_i}). \tag{3}$$

One can prove (3) in the following way. Fix an ω and define a continuous function A_t, $t \in [r_1, r_4]$, so that A_t is piecewise linear and equals w_{r_i} at all r_i's. Then by integrating through (2) against dA_t, remembering the definition of f_t and the fact that the integral of a sum equals the sum of integrals, we come to (3).

It follows from (3) that

$$E(\zeta(\Delta_1) + \zeta(\Delta_2))^2 = \sum_{i=1}^{3} E g_i^2 (w_{r_{i+1}} - w_{r_i})^2$$

$$+2 \sum_{i<j} E g_i g_j (w_{r_{i+1}} - w_{r_i})(w_{r_{j+1}} - w_{r_j}),$$

where all expectations make sense because $0 \leq f \leq 2$ and $Ew_t^2 = t < \infty$. Remember that $E(w_{r_{j+1}} - w_{r_j}) = 0$ and $E(w_{r_{i+1}} - w_{r_i})^2 = r_{i+1} - r_i$. Also notice that $(w_{r_{i+1}} - w_{r_i})^2$ and g_i^2 are independent by Exercise 5.2 and, for $i < j$, the g_i are $\mathcal{F}_{r_i}$-measurable and $\mathcal{F}_{r_j}$-measurable, owing to $\mathcal{F}_{r_i} \subset \mathcal{F}_{r_j}$, so that $g_i g_j (w_{r_{i+1}} - w_{r_i})$ is $\mathcal{F}_{r_j}$-measurable and hence independent of $w_{r_{j+1}} - w_{r_j}$. Then we see that

$$E(\zeta(\Delta_1) + \zeta(\Delta_2))^2 = \sum_{i=1}^{3} E f_{r_i+}^2 (r_{i+1} - r_i) = E \int_{r_1}^{r_4} f_t^2 \, dt$$

$$= E \int_{r_1}^{r_4} (I_{\Delta_1} + I_{\Delta_2})^2 \, dt = E \int_{r_1}^{r_4} I_{\Delta_1} \, dt + 2E \int_{r_1}^{r_4} I_{\Delta_1 \cap \Delta_2} \, dt + E \int_{r_1}^{r_4} I_{\Delta_2} \, dt$$

$$= \mu(\Delta_1) + 2\mu(\Delta_1 \cap \Delta_2) + \mu(\Delta_2). \tag{4}$$

By plugging in $\Delta_1 = \Delta_2 = \Delta$, we find that $E\zeta^2(\Delta) = \mu(\Delta)$. Then, developing $E(\zeta(\Delta_1) + \zeta(\Delta_2))^2$ and coming back to (4), we get $E\zeta(\Delta_1)\zeta(\Delta_2) = \mu(\Delta_1 \cap \Delta_2)$. Thus by Definition 3.5 the function ζ is a random orthogonal measure with reference measure μ.

The fact that $E\zeta = 0$ follows at once from the independence of $\mathcal{F}_s$ and $w_t - w_s$ for $t \geq s$. The theorem is proved.

Theorem 3 allows us to apply Theorem 3.13. By combining it with Exercise 2 and Remark 3.15 we come to the following result.

4. Theorem. *In notation* (1) *there exists a unique linear isometric operator* $I : L_2(\mathcal{P}, \mu) \to L_2(\mathcal{F}, P)$ *such that, for every* $n = 1, 2, ...,$ *constants* c_i, $s_i \leq t_i$, *and* $A_i \in \mathcal{F}_{s_i}$ *given for* $i = 1, ..., n$, *we have*

$$I(\sum_{i=1}^{n} c_i I_{A_i} I_{(s_i,t_i]}) = \sum_{i=1}^{n} c_i I_{A_i}(w_{t_i} - w_{s_i}) \quad (a.s.). \tag{5}$$

In addition, $EIf = 0$ *for every* $f \in L_2(\mathcal{P}, \mu)$.

5. Exercise*. Formula (5) admits the following generalization. Prove that for every $n = 1, 2, ...,$ constants $s_i \leq t_i$, and $\mathcal{F}_{s_i}$-measurable functions g_i given for $i = 1, ..., n$ and satisfying $Eg_i^2 < \infty$, we have

$$I(\sum_{i=1}^{n} g_i I_{(s_i,t_i]}) = \sum_{i=1}^{n} g_i(w_{t_i} - w_{s_i}) \quad (a.s.).$$

6. Definition. We call If, introduced in Theorem 4, *the Itô stochastic integral of* f, and write

$$If =: \int_0^\infty f(\omega, t)\, dw_t.$$

The Itô integral between nonrandom a and b such that $0 \leq a \leq b \leq \infty$ is naturally defined by

$$\int_a^b f(\omega, t)\, dw_t = \int_0^\infty f(\omega, t) I_{(a,b]}(t)\, dw_t.$$

The comments in Sec. 3 before Theorem 3.22 are valid for Itô stochastic integrals as well as for integrals of nonrandom functions against dw_t. It is natural to notice that for nonrandom functions both integrals introduced in this section and in Sec. 3 coincide (a.s.). This follows from formula (3.7), valid for both integrals (and from the possibility of finding appropriate f^n, a possibility which is either known to the reader or will be seen from Remark 8.6).

Generally it is safe to say that the properties of the Itô integral are absolutely different from those of the integral of nonrandom functions. For instance Exercise 3.23 implies that for nonrandom integrands the integral is either zero or its distribution has density. About 1981 M. Safonov constructed an example of random f_t satisfying $1 \leq f_t \leq 2$ and such that the distribution of $\int_0^1 f_t\,dw_t$ is singular with respect to Lebesgue measure.

One may wonder why we took sets like $A \times (s,t]$ and not $A \times [s,t)$ as a starting point for stochastic integration. Actually, for the Itô stochastic integral against the Wiener process this is irrelevant, and the second approach even has some advantages, since then (cf. Exercise 5) almost by definition we would have a very natural formula:

$$\int_0^\infty f(t)\,dw_t = \sum_{i=1}^n f(t_i)(w_{t_{i+1}} - w_{t_i})$$

provided that $f(t)$ is $\mathcal{F}_t$-measurable and $E|f(t)|^2 < \infty$ for every t, and $0 \leq t_1 \leq ... \leq t_{n+1} < \infty$ are nonrandom and such that $f(t) = f(t_i)$ for $t \in [t_i, t_{i+1})$ and $f(t) = 0$ for $t \geq t_{n+1}$. We show that this formula is indeed true in Theorem 8.8.

However, there is a significant difference between the two approaches if one tries to integrate with respect to discontinuous processes. Several unusual things may happen, and we offer the reader the following exercises showing one of them.

7. Exercise. In completely the same way as above one introduces a stochastic integral against $\bar{\pi}_t := \pi_t - t$, where π_t is the Poisson process with parameter 1. Of course, one needs an appropriate filtration of σ-fields $\mathcal{F}_t$ such that π_t is $\mathcal{F}_t$-measurable and $\pi_{t+h} - \pi_t$ is independent of $\mathcal{F}_t$ for all $t, h \geq 0$. On the other hand, one can integrate against $\bar{\pi}_t$ as usual, since this function has bounded variation on each interval $[0,T]$. In connection with this, prove that

$$E(\text{usual})\int_0^1 \pi_t\,d\bar{\pi}_t \neq 0,$$

so that either π_t is not stochastically integrable or the usual integral is different from the stochastic one. (As follows from Theorem 8.2, the latter is true.)

8. Exercise. In the situation of Exercise 7, prove that for every predictable nonnegative f_t we have

$$E(\text{usual})\int_0^1 f_t\,d\pi_t = E\int_0^1 f_t\,dt.$$

Conclude that π_t is *not* predictable, and is not $\mathcal{P}^\mu$-measurable either.

8. The structure of Itô integrable functions

Dealing with Itô stochastic integrals quite often requires much attention to tiny details, since often what seems true turns out to be absolutely wrong. For instance, we will see below that the function $I_{(0,\infty)}(w_t)I_{(0,1)}(t)$ is Itô integrable and consequently its Itô integral has zero mean. This may look strange due to the following.

Represent the open set $\{t : w_t > 0\}$ as the countable union of disjoint intervals (α_i, β_i). Clearly $w_{\alpha_i} = w_{\beta_i} = 0$, and

$$I_{(0,\infty)}(w_t)I_{(0,1)}(t) = \sum_i I_{(0,1)\cap(\alpha_i,\beta_i)}(t). \tag{1}$$

In addition it looks natural that

$$\int_0^\infty I_{(0,1)\cap(\alpha_i,\beta_i)}(t)\,dw_t = w_{1\wedge\alpha_i} - w_{1\wedge\beta_i}, \tag{2}$$

where the right-hand side is different from zero only if $\alpha_i < 1$, $\beta_i > 1$, and $w_1 > 0$, i.e. if $1 \in (\alpha_i, \beta_i)$. In that case the right-hand side of (2) equals $(w_1)_+$, and since the integral of a sum should be equal to the sum of integrals, formula (1) shows that the Itô integral of $I_{(0,\infty)}(w_t)I_{(0,1)}(t)$ should equal $(w_1)_+$. However, this is impossible since $E(w_1)_+ > 0$.

The contradiction here comes from the fact that the terms in (1) are not Itô integrable and (2) just does not make sense.

One more example of an integral with no sense gives $\int_0^1 w_1\,dw_t$. Again its mean value should be zero, but under every reasonable way of defining this integral it should equal $w_1\int_0^1 dw_t = w_1^2$.

All this leads us to the necessity of investigating the set of Itô integrable functions. Due to Theorem 3.19 and Exercise 3.2 this is equivalent to investigating which functions are $\mathcal{P}^\mu$-measurable.

1. Definition. A function $f_t(\omega)$ given on $\Omega \times (0,\infty)$ is called $\mathcal{F}_t$*-adapted* if it is $\mathcal{F}_t$-measurable for each $t > 0$. By H we denote the set of all real-valued $\mathcal{F}_t$-adapted functions $f_t(\omega)$ which are $\mathcal{F} \otimes \mathfrak{B}(0,\infty)$-measurable and satisfy

$$E\int_0^\infty f_t^2\,dt < \infty.$$

The following theorem says that all elements of H are Itô integrable. The reader is sent to Sec. 7 for necessary notation.

2. Theorem. *We have* $H \subset L_2(\mathcal{P},\mu)$.

Proof (Doob). It suffices only to prove that $f \in L_2(\mathcal{P},\mu)$ for $f \in H$ such that $f_t(\omega) = 0$ for $t \geq T$, where T is a constant. Indeed, by the dominated convergence theorem

$$\int_X |f_t - f_t I_{t\leq n}|^2\,dPdt = E\int_n^\infty f_t^2\,dt \to 0$$

as $n \to \infty$, so that, if $f_t I_{t\leq n} \in L_2(\mathcal{P},\mu)$, then $f_t \in L_2(\mathcal{P},\mu)$ due to the completeness of $L_2(\mathcal{P},\mu)$.

Therefore we fix an $f \in H$ and $T < \infty$ and assume that $f_t = 0$ for $t \geq T$. It is convenient to assume that f_t is defined for negative t as well, and $f_t = 0$ for $t \leq 0$. Now we recall that it is known from integration theory that every L_2-function is continuous in L_2. More precisely, if $h \in L_2([0,T])$ and $h(t) = 0$ outside $[0,T]$, then

$$\lim_{a\to 0}\int_{-T}^T |h(t+a) - h(t)|^2\,dt = 0.$$

This and the inequality

$$\int_{-T}^T |f_{t+a} - f_t|^2\,dt \leq 2\Big(\int_{-T}^T f_{t+a}^2\,dt + \int_{-T}^T f_t^2\,dt\Big) \leq 4\int_0^T f_t^2\,dt$$

along with the dominated convergence theorem imply that

$$\lim_{a\to 0} E\int_{-T}^T |f_{t+a} - f_t|^2\,dt = 0. \tag{3}$$

Now let

$$\rho_n(t) = k2^{-n} \quad \text{for} \quad t \in (k2^{-n}, (k+1)2^{-n}].$$

Changing variables $t + s = u, t = v$ shows that

$$\int_0^1 E\int_0^T |f_{\rho_n(t+s)-s}-f_t|^2\,dtds=\int_0^{T+1}\big(E\int_{u-1}^{u\wedge T}|f_{\rho_n(u)-u+v}-f_v|^2\,dv\big)\,du.$$

The last expectation tends to zero owing to (3) uniformly with respect to u, since $0\le u-\rho_n(u)\le 2^{-n}$. It follows that there is a sequence $n(k)\to\infty$ such that for almost every $s\in[0,1]$

$$\lim_{k\to\infty}E\int_0^T|f_{\rho_{n(k)}(t+s)-s}-f_t|^2\,dt=0. \tag{4}$$

Fix any s for which (4) holds, and denote $f_t^k=f_{\rho_{n(k)}(t+s)-s}$. Then (4) and the inequality $|a|^2\le 2|b|^2+2|a-b|^2$ show that $|f_t^k|^2$ is μ-integrable at least for all large k.

Furthermore, it turns out that the f_t^k are predictable. Indeed,

$$f_{\rho_n(t+s)-s}=\sum_i f_{i2^{-n}-s}I_{(i2^{-n}-s,(i+1)2^{-n}-s]}(t)=\sum_{i:i2^{-n}-s>0}\quad. \tag{5}$$

In addition, $f_{t_1}I_{(t_1,t_2]}$ is predictable if $0\le t_1\le t_2$, since for any Borel B

$$\{(\omega,t):f_{t_1}(\omega)I_{(t_1,t_2]}(t)\in B\}$$
$$=(\{\omega:f_{t_1}(\omega)\in B\}\times(t_1,t_2])\cup\{(\omega,t):I_{(t_1,t_2]}(t)=0\in B\}\in\mathcal{P}.$$

Therefore (5) yields the predictability of f_t^k, and the integrability of $|f_t^k|^2$ now implies that $f_t^k\in L_2(\mathcal{P},\mu)$. The latter space is complete, and owing to (4) we have $f_t\in L_2(\mathcal{P},\mu)$. The theorem is proved.

3. Exercise*. By following the above proof, show that left continuous $\mathcal{F}_t$-adapted processes are predictable.

4. Exercise. Go back to Exercise 7.7 and prove that if f_t is left continuous, $\mathcal{F}_t$-adapted, and $E\int_0^1 f_t^2\,dt<\infty$, then the usual integral $\int_0^1 f_t\,d\bar\pi_t$ coincides with the stochastic one (a.s.). In particular, prove that the usual integral $\int_0^1 \pi_{t-}\,d\bar\pi_t$ coincides with the stochastic integral $\int_0^1\pi_t\,d\bar\pi_t$ (a.s.).

5. Exercise. Prove that if $f\in L_2(\mathcal{P},\mu)$, then there exists $h\in H$ such that $f=h$ μ-a.e. and in this sense $H=L_2(\mathcal{P},\mu)$.

6. Remark. If f_t is independent of ω, (4) implies that for almost any $s\in[0,1]$

$$\lim_{k\to\infty}\int_0^T|f_{\rho_{n(k)}(t+s)-s}-f_t|^2\,dt=0,\qquad \int_0^T f_t\,dt=\lim_{k\to\infty}\int_0^T f_{\rho_{n(k)}(t+s)-s}\,dt.$$

This means that appropriate Riemann sums converge to the Lebesgue integral of f.

7. Remark. It is seen from the proof of Theorem 2 that, if $f \in H$, then for any integer $n \geq 1$ one can find a partition $0 = t_{n0} < t_{n1} < ... < t_{nk(n)} = n$ such that $\max_i(t_{n,i+1} - t_{ni}) \leq 1/n$ and

$$\lim_{n\to\infty} E \int_0^\infty |f_t - f_t^n|^2 \, dt = 0,$$

where $f^n \in H$ are defined by $f_t^n = f_{t_{ni}}$ for $t \in (t_{ni}, t_{n,i+1}]$, $i \leq k(n) - 1$, and $f_t^n = 0$ for $t > n$. Furthermore, the f_t^n are predictable, and by Theorem 7.4

$$\int_0^\infty f_t \, dw_t = \operatorname*{l.i.m.}_{n\to\infty} \int_0^\infty f_t^n \, dw_t. \tag{6}$$

One can apply the same construction to vector-valued functions f, and then one sees that the above partitions can be taken the same for any finite number of f's.

Next we prove two properties of the Itô integral. The first one justifies the notation $\int_0^\infty f_t \, dw_t$, and the second one shows a kind of local property of this integral.

8. Theorem. (i) *If $f \in H$, $0 = t_0 < t_1 < ... < t_n < ...$, $f_t = f_{t_i}$ for $t \in [t_i, t_{i+1})$ and $i \geq 0$, then in the mean square sense*

$$\int_0^\infty f_t \, dw_t = \sum_{i=0}^\infty f_{t_i}(w_{t_{i+1}} - w_{t_i}).$$

(ii) *If $g, h \in H$, $A \in \mathcal{F}$, and $h_t(\omega) = g_t(\omega)$ for $t \geq 0$ and $\omega \in A$, then $\int_0^\infty g_t \, dw_t = \int_0^\infty h_t \, dw_t$ on A (a.s.).*

Proof. (i) Define $f_t^i = f_{t_i} I_{(t_i, t_{i+1}]}$ and observe the simple fact that $f = \sum_i f^i$ μ-a.e. Then the linearity and continuity of the Itô integral show that to prove (i) it suffices to prove that

$$\int_0^\infty g I_{(r,s]}(t) \, dw_t = (w_s - w_r) g \tag{7}$$

(a.s.) if g is $\mathcal{F}_r$-measurable, $Eg^2 < \infty$, and $0 \leq r < s < \infty$.

If g is a step function (having the form $\sum_{i=1}^n c_i I_{A_i}$ with constant c_i and $A_i \in \mathcal{F}_r$), then (7) follows from Theorem 7.4. The general case is suggested as Exercise 7.5.

To prove (ii), take common partitions for g and h from Remark 7 and on their basis construct the sequences g_t^n and h_t^n. Then by (i) the left-hand sides of (6) for $f_t^n = g_t^n$ and $f_t^n = h_t^n$ coincide on A (a.s.). Formula (6) then says that the same is true for the integrals of g and h. The theorem is proved.

Much later (see Sec. 6.1) we will come back to Itô stochastic integrals with variable upper limit. We want these integrals to be continuous. For this purpose we need some properties of martingales which we present in the following chapter. The reader can skip it if he/she is only interested in stationary processes.

9. Hints to exercises

2.5 Use Exercise 1.4.14, with $R(x) = x$, and estimate $\int_0^x \sqrt{(-\ln y)/y}\,dy$ through $\sqrt{x(-\ln x)}$ by using l'Hospital's rule.

2.10 The cases $a \le b$ and $a > b$ are different. At some moment you may like to consult the proof of Theorem 2.3 taking there 2^{2n} in place of n.

2.12 If $P(\xi \le a, \eta \le b) = \int_{-\infty}^b f(x)\,dx$ for every b, then $Eg(\eta)I_{\xi \le a} = \int_{\mathbb{R}} g(x)f(x)\,dx$. The result of these computations is given in Sec. 6.8.

3.4 It suffices to prove that the indicators of sets $(s,t]$ are in $L_p(\Pi,\mu)$.

3.8 Observe that

$$\varphi(s) = E\exp(i\sum_{n=1}^\infty f(s+\sigma_n)),$$

and by using the independence of the τ_n and the fact that $EF(\tau_1,\tau_2,...) = E\Phi(\tau_1)$, where $\Phi(t) = EF(t,\tau_2,...)$, show that

$$\varphi(s) = \int_0^\infty e^{if(s+t)-t}\varphi(s+t)\,dt = e^s\int_s^\infty e^{if(t)}(e^{-t}\varphi(t))\,dt.$$

Conclude first that φ is continuous, then that $\varphi(s)e^{-s}$ is differentiable, and solve the above equation. After that, approximate by continuous functions the function which is constant on each interval $(t_j,t_{j+1}]$ and vanishes outside of the union of these intervals.

3.14 Prove that, for every Borel nonnegative f, we have

$$E\sum_{\sigma_n\le 1} f(\sigma_n) = \int_0^1 f(s)\,ds,$$

and use it to pass to the limit from step functions to arbitrary ones.

3.21 For $b_n > 0$ with $b_n \to 1$, we have $\prod b_n = 0$ if and only if $\sum_n(1-b_n) = \infty$.

3.23 Use Remark 1.4.10 and (3.6).

5.9 Take any continuous function $u(x)$ defined on $[a,b]$ such that $u<0$ in (a,b) and $u(a)=u(b)=0$, and use it to write a formula similar to (5.1).

6.7 Define τ as the first exit time of x_t from (a,b) and, similarly to (6.3), prove that

$$u(0)=E\int_0^{\tau} e^{-\lambda t}(\lambda u(x_t)-u'(x_t)-(1/2)u''(x_t))\,dt+Ee^{-\lambda\tau}u(x_\tau).$$

7.7 Observe that $\int_{(0,t]}\pi_s\,d\pi_s=\pi_t(\pi_t+1)/2$.

7.8 First take $f_t=I_\Delta$.

8.4 Keep in mind the proof of Theorem 8.2, and redo Exercise 7.5 for π_t in place of w_t.

8.5 Take a sequence of step functions converging to f μ-a.e., and observe that step functions are $\mathcal{F}_t$-adapted.

Chapter 5

Infinitely Divisible Processes

The Wiener process has independent increments and the distribution of each increment depends only on the length of the time interval over which the increment is taken. There are many other processes possessing this property; for instance, the Poisson process or the process τ_a, $a \geq 0$, from Example 2.5.8 are examples of those (see Theorem 2.6.1).

In this chapter we study what can be said about general processes of that kind. They are supposed to be given on a *complete* probability space $(\Omega, \mathcal{F}, P)$ usually behind the scene. The assumption that this space is complete will turn out to be convenient to use starting with Exercise 5.5. One more stipulation is that unless explicitely stated otherwise, all the processes under consideration are assumed to be *real valued*. Finally, after Theorem 1.5 we tacitly assume that all processes under consideration are stochastically continuous without specifying this each time.

1. Stochastically continuous processes with independent increments

We start with processes having independent increments. The main goal of this section is to show that these processes, or at least their modifications, have rather regular trajectories (see Theorem 11).

1. Definition. A real- or vector-valued random process ξ_t given on $[0, \infty)$ is said to be a process *with independent increments* if $\xi_0 = 0$ (a.s.) and $\xi_{t_1}, \xi_{t_2} - \xi_{t_1}, ..., \xi_{t_n} - \xi_{t_{n-1}}$ are independent provided $0 \leq t_1 \leq ... \leq t_n < \infty$.

We will be only dealing with stochastically continuous processes.

2. Definition. A real- or vector-valued random process ξ_t given on $[0,\infty)$ is said to be *stochastically continuous* at a point $t_0 \in [0,\infty)$ if $\xi_t \xrightarrow{P} \xi_{t_0}$ as $t \to t_0$. We say that ξ_t is stochastically continuous on a set if it is stochastically continuous at each point of the set.

Clearly, ξ_t is stochastically continuous at t_0 if $E|\xi_t - \xi_{t_0}| \to 0$ as $t \to t_0$. Stochastic continuity is very weakly related to the continuity of trajectories. For instance, for the Poisson process with parameter 1 (see Exercise 2.3.8) we have $E|\xi_t - \xi_{t_0}| = |t - t_0|$. However, all trajectories of ξ_t are discontinuous. By the way, this example shows also that the requirement $\beta > 0$ in Kolmogorov's Theorem 1.4.8 is essential. The trajectories of τ_a, $a \geq 0$, are also discontinuous, but this process is stochastically continuous too since (see Theorem 2.6.1 and (2.5.1))

$$P(|\tau_b - \tau_a| > \varepsilon) = P(\tau_{|b-a|} > \varepsilon) = P(\max_{t \leq \varepsilon} w_s < |b-a|) \to 0 \quad \text{as} \quad b \to a.$$

3. Exercise. Prove that, for any ω, the function τ_a, $a > 0$, is *left continuous* in a.

4. Definition. A (real-valued) random process ξ_t given on $[0,\infty)$ is said to be *bounded in probability* on a set $I \subset [0,\infty)$ if

$$\lim_{c \to \infty} \sup_{t \in I} P(|\xi_t| > c) = 0.$$

As in usual analysis, one proves the following.

5. Theorem. *If the process ξ_t is stochastically continuous on $[0,T]$ ($T < \infty$), then*

(i) *it is uniformly stochastically continuous on $[0,T]$, that is, for any $\gamma, \varepsilon > 0$ there exists $\delta > 0$ such that*

$$P(|\xi_{t_1} - \xi_{t_2}| > \varepsilon) < \gamma,$$

whenever $t_1, t_2 \in [0,T]$ and $|t_1 - t_2| \leq \delta$;

(ii) *it is bounded in probability on $[0,T]$.*

The proof of this theorem is left to the reader as an exercise.

From this point on we will only consider stochastically continuous processes on $[0,\infty)$, without specifying this each time.

To prove that processes with independent increments admit modifications without second-type discontinuities, we need the following lemma.

6. Lemma (Ottaviani's inequality). *Let η_k, $k = 1, ..., n$, be independent random variables, $S_k = \eta_1 + ... + \eta_k$, $a \geq 0$, $0 \leq \alpha < 1$, and*

$$P\{|S_n - S_k| \geq a\} \leq \alpha \quad \forall k.$$

Then for all $c \geq 0$

$$P\{\max_{k \leq n} |S_k| \geq a + c\} \leq \frac{1}{1-\alpha} P\{|S_n| \geq c\}. \tag{1}$$

Proof. The probability on the left in (1) equals

$$\sum_{k=1}^{n} P\{|S_i| < a + c, i < k, |S_k| \geq a + c\}$$

$$\leq \frac{1}{1-\alpha} \sum_{k=1}^{n} P\{|S_i| < a + c, i < k, |S_k| \geq a + c, |S_n - S_k| < a\}$$

$$\leq \frac{1}{1-\alpha} \sum_{k=1}^{n} P\{|S_i| < a+c, i < k, |S_k| \geq a+c, |S_n| \geq c\} \leq \frac{1}{1-\alpha} P\{|S_n| \geq c\}.$$

The lemma is proved.

7. Theorem. *Let ξ_t be a process with independent increments on $[0, \infty)$, $T \in [0, \infty)$, and let ρ be the set of all rational points on $[0, T]$. Then*

$$P\{\sup_{r \in \rho} |\xi_r| < \infty\} = 1.$$

Proof. Obviously it suffices to prove that for some $h > 0$ and all $t \in [0, T]$ we have

$$P\{\sup_{r \in [t, t+h] \cap \rho} |\xi_r| < \infty\} = 1. \tag{2}$$

Take $h > 0$ so that $P\{|\xi_u - \xi_{u+s}| \geq 1\} \leq 1/2$ for all s, u such that $0 \leq s \leq h$ and $s + u \leq T$. Such a choice is possible owing to the uniform stochastic continuity of ξ_t on $[0, T]$. Fix $t \in [0, T]$ and let

$$r_1, ..., r_n \in [t, t + h] \cap \rho, \quad r_1 \leq ... \leq r_n.$$

Observe that $\xi_{r_k} = \xi_{r_1} + (\xi_{r_2} - \xi_{r_1}) + ... + (\xi_{r_k} - \xi_{r_{k-1}})$, where the summands are independent. In addition, $P\{|\xi_{r_n} - \xi_{r_k}| \geq 1\} \leq 1/2$. Hence by Lemma 6

$$P\{\sup_{k\le n}|\xi_{r_k}|\ge 1+c\}\le 2\sup_{t\in[0,T]}P\{|\xi_t|\ge c\}. \tag{3}$$

The last inequality is true for any arrangement of the points $r_k\in[t,t+h]\cap\rho$ which may not be necessarily ordered increasingly. Therefore, now we can think of the set $\{r_1,r_2,...\}$ as being the whole $\rho\cap[t,t+h]$. Then, passing to the limit in (3) as $n\to\infty$ and noticing that

$$\sup\{|\xi_{r_k}|:k=1,2,...\}\uparrow\sup\{|\xi_r|:r\in\rho\cap[t,t+h]\},$$

we find that

$$P\{\sup_{r\in[t,t+h]\cap\rho}|\xi_r|>1+c\}\le 2\sup_{t\in[0,T]}P\{|\xi_t|\ge c\}.$$

Finally, by letting $c\to\infty$ and using the uniform boundedness of ξ_r in probability, we come to (2). The theorem is proved.

Define $D[0,\infty)$ to be the set of all complex-valued right-continuous functions on $[0,\infty)$ which have finite left limits at each point $t\in(0,\infty)$. Similarly one defines $D[0,T]$. We say that a function $x_\cdot$ is a *cadlag* function on $[0,T]$ if $x_\cdot\in D[0,T]$, and just cadlag if $x_\cdot\in D[0,\infty)$.

8. Exercise*. Prove that if $x^n_\cdot\in D[0,\infty)$, $n=1,2,...$, and the x^n_t converge to x_t as $n\to\infty$ uniformly on each finite time interval, then $x_\cdot\in D[0,\infty)$.

9. Lemma. *Let $\rho=\{r_1,r_2,...\}$ be the set of all rational points on $[0,1]$, x_t a real-valued (nonrandom) function given on ρ. For $a<b$ define $\beta_n(x_\cdot,a,b)$ to be the number of upcrossings of the interval (a,b) by the function x_t restricted to the set $r_1,r_2,...,r_n$. Assume that*

$$\lim_{n\to\infty}\beta_n(x_\cdot,a,b)<\infty$$

for any rational a and b. Then the function

$$\tilde{x}_t:=\lim_{\rho\ni r\downarrow t}x_r$$

is well defined for any $t\in[0,1)$, is right continuous on $[0,T)$, and has (perhaps infinite) left limits on $(0,T]$.

This lemma is set as an exercise on properties of $\overline{\lim}$ and $\underline{\lim}$.

10. Lemma. *Let $\psi(t,\lambda)$ be a complex-valued function defined for $\lambda \in \mathbb{R}$ and $t \in [0,1]$. Assume that $\psi(t,\lambda)$ is continuous in t and never takes the zero value. Let ξ_t be a stochastically continuous process such that*

(i) $\sup_{r\in\rho} |\xi_r| < \infty$ *(a.s.);*

(ii) $\lim_{n\to\infty} E\beta_n(\eta^i_\cdot(\lambda), a, b) < \infty$ *for any* $-\infty < a < b < \infty$, $\lambda \in \mathbb{R}$, $i = 1,2$, *where*

$$\eta^1_t(\lambda) = \mathrm{Re}\,[\psi(t,\lambda)e^{i\lambda\xi_t}], \quad \eta^2_t(\lambda) = \mathrm{Im}\,[\psi(t,\lambda)e^{i\lambda\xi_t}].$$

Then the process ξ_t admits a modification, all trajectories of which belong to $D[0,1]$.

Proof. Denote $\eta_t(\lambda) = \psi(t,\lambda)e^{i\lambda\xi_t}$ and

$$\Omega' = \bigcap_{m=1}^{\infty} \bigcap_{\substack{a<b\\ a,b \text{ rational}}} \{\lim_{n\to\infty} \beta_n(\eta^i_\cdot(\tfrac{1}{m}), a, b) < \infty, i = 1,2\} \cap \{\sup_{r\in\rho} |\xi_r| < \infty\}.$$

Obviously, $P(\Omega') = 1$. For $\omega \in \Omega'$ Lemma 9 allows us to let

$$\tilde{\eta}_t(\tfrac{1}{m}) = \lim_{\rho\ni r\downarrow t} \eta_r(\tfrac{1}{m}), \quad t < 1, \quad \tilde{\eta}_1(\tfrac{1}{m}) = \eta_1(\tfrac{1}{m}).$$

For $\omega \notin \Omega'$ let $\tilde{\eta}_t(\frac{1}{m}) \equiv 0$. Observe that, since ψ is continuous in t and $P(\Omega') = 1$ and ξ_t is stochastically continuous, we have that

$$\begin{aligned} &\tilde{\eta}_t(\tfrac{1}{m}) = P\text{-}\lim_{\rho\ni r\downarrow t} \eta_r(\tfrac{1}{m}) = \eta_t(\tfrac{1}{m}) \quad \text{(a.s.)} \quad \forall t < 1, \\ &\tilde{\eta}_1(\tfrac{1}{m}) = \eta_1(\tfrac{1}{m}). \end{aligned} \tag{4}$$

Furthermore, $|\tilde{\eta}_t(\frac{1}{m})\psi^{-1}(t, \frac{1}{m})| \le 1$ for all ω and t.

Now define $\mu = \mu(\omega) = [\sup_{r\in\rho} |\xi_r|] + 1$ and

$$\tilde{\xi}_t = \mu \arcsin \mathrm{Im}\, \tilde{\eta}_t(1/\mu)\psi^{-1}(t, 1/\mu)I_{\Omega'}.$$

By Lemma 9, $\tilde{\eta}_\cdot(\frac{1}{m}) \in D[0,1]$ for any ω. Hence, $\tilde{\xi}_t \in D[0,1]$ for any ω.

It only remains to prove that $P\{\tilde{\xi}_t = \xi_t\} = 1$ for any $t \in [0,1]$. For $t \in \rho$ we have this equality from (4) and from the formula

$$\xi_t = \mu \arcsin \mathrm{Im}\, \eta_t(1/\mu)\psi^{-1}(t, 1/\mu),$$

which holds for $\omega \in \Omega'$. For other t, owing to the stochastic continuity of ξ_t and the right continuity of $\tilde{\xi}_t$, we have

$$\xi_t = P\text{-}\lim_{\rho \ni r \downarrow t} \xi_r = P\text{-}\lim_{\rho \ni r \downarrow t} \tilde{\xi}_r = \tilde{\xi}_t$$

(a.s.). The lemma is proved.

11. Theorem. *Stochastically continuous processes with independent increments admit modifications which are right continuous and have finite left limits for any* ω.

Proof. Let ξ_t be a process in question. It suffices to construct a modification with the described properties on each interval $[n, n+1]$, $n = 0, 1, 2, ...$ The reader can easily combine these modifications to get what we want on $[0, \infty)$. We will confine ourselves to the case $n = 0$. Let ρ be the set of all rational points on $[0, 1]$, and let

$$\varphi(t, \lambda) = Ee^{i\xi_t \lambda}, \quad \varphi(t_1, t_2, \lambda) = Ee^{i\lambda(\xi_{t_2} - \xi_{t_1})}.$$

Since the process ξ_t is stochastically continuous, the function $\varphi(t_1, t_2, \lambda)$ is continuous in $(t_1, t_2) \in [0, 1] \times [0, 1]$ for any λ. Therefore, this function is uniformly continuous on $[0, 1] \times [0, 1]$, and, because $\varphi(t, t, \lambda) = 1$, there exists $\delta(\lambda) > 0$ such that $|\varphi(t_1, t_2, \lambda)| \geq 1/2$ whenever $|t_1 - t_2| < \delta(\lambda)$ and $t_1, t_2 \in [0, 1]$. Furthermore, for any $t \in [0, 1]$ and $\lambda \in \mathbb{R}$ one can find $n \geq 1$ and $0 = t_1 \leq t_2 \leq ... \leq t_n = t$ such that $|t_k - t_{k-1}| < \delta(\lambda)$. Then, using the independence of increments, we find that

$$\varphi(t, \lambda) = \varphi(t_1, t_2, \lambda) \cdot ... \cdot \varphi(t_{n-1}, t_n, \lambda),$$

which implies that $\varphi(t, \lambda) \neq 0$. In addition, $\varphi(t, \lambda)$ is continuous in t.

For fixed λ consider the process

$$\eta_t = \eta_t(\lambda) = \varphi^{-1}(t, \lambda) e^{i\lambda\xi_t}.$$

Let $s_1, s_2, ..., s_n$ be rational numbers in $[0, 1]$ such that $s_1 \leq ... \leq s_n$. Define

$$\mathcal{F}_k = \sigma\{\xi_{s_1}, \xi_{s_2} - \xi_{s_1}, ..., \xi_{s_k} - \xi_{s_{k-1}}\}.$$

Notice that $(\operatorname{Re}\eta_{s_k}, \mathcal{F}_k)$ and $(\operatorname{Im}\eta_{s_k}, \mathcal{F}_k)$ are martingales. Indeed, by virtue of the independence of $\xi_{s_{k+1}} - \xi_{s_k}$ and $\mathcal{F}_k$, we have

$$E\{\operatorname{Re}\eta_{s_{k+1}} | \mathcal{F}_k\} = \operatorname{Re} E\{e^{i\lambda\xi_{s_k}} \varphi^{-1}(s_{k+1}, \lambda) e^{i\lambda(\xi_{s_{k+1}} - \xi_{s_k})} | \mathcal{F}_k\}$$

$$= \operatorname{Re} e^{i\lambda\xi_{s_k}} \varphi^{-1}(s_{k+1}, \lambda) \varphi(s_k, s_{k+1}, \lambda) = \operatorname{Re}\eta_{s_k} \quad \text{(a.s.)}.$$

Hence by Doob's upcrossing theorem, if $r_i \in \rho$, $\{r_1, ..., r_n\} = \{s_1, ..., s_n\}$, and $0 \leq s_1 \leq ... \leq s_n$, then

$$E\beta_n(\operatorname{Re}\eta_\cdot, a, b) \leq (E|\operatorname{Re}\eta_{s_n}| + |a|)/(b - a)$$

$$\leq (\sup_{t\in[0,1]} \varphi^{-1}(t,\lambda) + |a|)/(b-a) < \infty,$$

$$\sup_n E\beta_n(\operatorname{Im}\eta_\cdot, a, b) < \infty.$$

It only remains to apply Lemma 10. The theorem is proved.

12. Exercise* (cf. Exercise 3)**.** Take the stable process τ_a, $a \geq 0$, from Theorem 2.6.1. Observe that τ_a increases in a and prove that its cadlag modification, the existence of which is asserted in Theorem 11, is given by τ_{a+}, $a \geq 0$.

2. Lévy-Khinchin theorem

In this section we prove a remarkable Lévy-Khinchin theorem. It is worth noting that this theorem was originally proved for so-called infinitely divisible laws and not for infinitely divisible processes. As usual we are only dealing with one-dimensional processes (the multidimensional case is treated, for instance, in [**GS**]).

1. Definition. A process ξ_t with independent increments is called *time homogeneous* if, for every $h > 0$, the distribution of $\xi_{t+h} - \xi_t$ is independent of t.

2. Definition. A stochastically continuous time-homogeneous process ξ_t with independent increments is called an *infinitely divisible process*.

3. Theorem (Lévy-Khinchin)**.** *Let ξ_t be an infinitely divisible process on $[0,\infty)$. Then there exist a finite nonnegative measure on $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ and a number $b \in \mathbb{R}$ such that, for any $t \in [0,\infty)$ and $\lambda \in \mathbb{R}$, we have*

$$Ee^{i\lambda\xi_t} = \exp\{t\int_{\mathbb{R}} f(\lambda,x)\,\mu(dx) + itb\lambda\}, \tag{1}$$

where

$$f(\lambda,x) = (e^{i\lambda x} - 1 - i\lambda\sin x)\frac{1+x^2}{x^2}, \quad x \neq 0, \quad f(\lambda,0) := -\frac{\lambda^2}{2}.$$

Proof. Denote $\varphi(t,\lambda) = Ee^{i\lambda\xi_t}$. In the proof of Theorem 1.11 we saw that $\varphi(t,\lambda)$ is continuous in t and $\varphi(t,\lambda) \neq 0$. In addition $\varphi(t,\lambda)$ is continuous with respect to the pair (t,λ). Define

$$a(t,\lambda) = \arg\varphi(t,\lambda), \quad l(t,\lambda) = \ln|\varphi(t,\lambda)|.$$

By using the continuity of φ and the fact that $\varphi \neq 0$, one can uniquely define $a(t,\lambda)$ to be continuous in t and in λ and satisfy $a(0,\lambda) = a(t,0) = 0$.

Clearly, $l(t,\lambda)$ is a finite function which is also continuous in t and in λ. Furthermore,

$$\varphi(t,\lambda) = \exp\{l(t,\lambda) + ia(t,\lambda)\}.$$

Next, it follows from the homogeneity and independence of increments of ξ_t that

$$\varphi(t+s,\lambda) = \varphi(t,\lambda)\varphi(s,\lambda).$$

Hence, by definition of a, we get that, for each λ, it satisfies the equation

$$f(t+s) = f(t) + f(s) + 2\pi k(s,t),$$

where $k(s,t)$ is a continuous integer-valued function. Since $k(t,0) = 0$, in fact, $k \equiv 0$, and a satisfies $f(t+s) = f(t) + f(s)$. The same equation is also valid for l. Any continuous solution of this equation has the form ct, where c is a constant. Thus,

$$a(t,\lambda) = ta(\lambda), \quad l(t,\lambda) = tl(\lambda),$$

where $a(\lambda) = a(1,\lambda)$ and $l(\lambda) = l(1,\lambda)$. By defining $g(\lambda) := l(\lambda) + ia(\lambda)$, we write

$$\varphi(t,\lambda) = e^{tg(\lambda)},$$

where g is a continuous function of λ and $g(0) = 0$. We have reduced our problem to finding g.

Observe that

$$g(\lambda) = \lim_{t\downarrow 0} \frac{e^{tg(\lambda)} - 1}{t} = \lim_{t\downarrow 0} \frac{\varphi(t,\lambda) - 1}{t}. \tag{2}$$

Moreover, from Taylor's expansion of $\exp(tg(\lambda))$ with respect to t one easily sees that the convergence in (2) is uniform on each set of values of λ on which $g(\lambda)$ is bounded. In particular, this is true on each set $[-h,h]$ with $0 \le h < \infty$.

By taking t of type $1/n$ and denoting F_t the distribution of ξ_t, we conclude that

$$n\int_{\mathbb{R}} (e^{i\lambda x} - 1)\, F_{1/n}(dx) \to g(\lambda) \tag{3}$$

as $n \to \infty$ uniformly in λ on any finite interval. Integrate this against $d\lambda$ to get

$$\lim_{n\to\infty} n \int_{\mathbb{R}} \left(1 - \frac{\sin xh}{xh}\right) F_{1/n}(dx) = -\frac{1}{2h}\int_{-h}^{h} g(\lambda)\, d\lambda. \tag{4}$$

Notice that the right-hand side of (4) can be made arbitrarily small by choosing h small, since g is continuous and vanishes at zero. Furthermore, as is easy to see, $1 - \sin xh/xh \geq 1/2$ for $|xh| \geq 2$. It follows that, for any $\varepsilon > 0$, there exists $h > 0$ such that

$$\overline{\lim_{n\to\infty}} (n/2) \int_{|x|\geq 2/h} F_{1/n}(dx) \leq \varepsilon.$$

In turn, it follows that, for all large n,

$$n \int_{|x|\geq 2/h} F_{1/n}(dx) \leq 4\varepsilon. \tag{5}$$

By reducing h one can accommodate any finite set of values of n and find an h such that (5) holds for all $n \geq 1$ rather than only for large ones.

To derive yet another consequence of (4), notice that there exists a constant $\gamma > 0$ such that

$$1 - \frac{\sin x}{x} \geq \gamma \frac{x^2}{1+x^2} \quad \forall x \in \mathbb{R}.$$

Therefore, from (4) with $h = 1$, we obtain that there exists a finite constant c such that for all n

$$n \int_{\mathbb{R}} \frac{x^2}{1+x^2} F_{1/n}(dx) \leq c. \tag{6}$$

Finally, upon introducing measures μ_n by the formula

$$\mu_n(dx) = n \frac{x^2}{1+x^2} F_{1/n}(dx),$$

and noticing that $\mu_n \leq nF_{1/n}$, from (5) and (6), we see that the family $\{\mu_n, n = 1, 2, ...\}$ is weakly compact. Therefore, there exist a subsequence $n' \to \infty$ and a finite measure μ such that

$$\int_{\mathbb{R}} f(x)\, \mu_{n'}(dx) \to \int_{\mathbb{R}} f(x)\, \mu(dx)$$

for every bounded and continuous f. As is easy to check $f(\lambda, x)$ is bounded and continuous in x. Hence,

$$g(\lambda) = \lim_{n\to\infty} n \int_{\mathbb{R}} (e^{i\lambda x} - 1)\, F_{1/n}(dx)$$

$$= \lim_{n\to\infty} [\int_{\mathbb{R}} f(\lambda, x)\mu_n(dx) + i\lambda n \int_{\mathbb{R}} \sin x\, F_{1/n}(dx)]$$

$$= \lim_{n'\to\infty} [\int_{\mathbb{R}} f(\lambda, x)\mu_{n'}(dx) + i\lambda n' \int_{\mathbb{R}} \sin x\, F_{1/n'}(dx)]$$

$$= \int_{\mathbb{R}} f(\lambda, x)\mu(dx) + i\lambda b,$$

where

$$b := \lim_{n'\to\infty} n' \int_{\mathbb{R}} \sin x\, F_{1/n'}(dx),$$

and the existence and finiteness of this limit follows from above computations in which all limits exist and are finite. The theorem is proved.

Formula (1) is called *Khinchin's formula.* The following *Lévy's formula* sheds more light on the structure of the process x_t:

$$\varphi(t, \lambda) = \exp t\{\int_{\mathbb{R}} (e^{i\lambda x} - 1 - i\lambda \sin x)\, \Lambda(dx) + ib\lambda - \sigma^2\lambda^2/2\},$$

where Λ is called *the Lévy measure of* ξ_t. This is a nonnegative, generally speaking, infinite measure on $\mathfrak{B}(\mathbb{R})$ such that

$$\int_{\mathbb{R}} \frac{x^2}{1+x^2}\, \Lambda(dx) < \infty, \quad \Lambda(\{0\}) = 0. \tag{7}$$

Any such measure is called *a Lévy measure.* One obtains one formula from the other by introducing the following relations between μ and the pair (Λ, σ^2):

$$\mu(\{0\}) = \sigma^2, \quad \Lambda(\Gamma) = \int_{\Gamma\setminus\{0\}} \frac{1+x^2}{x^2}\, \mu(dx).$$

4. Exercise*. Prove that if one introduces (Λ, σ^2) by the above formulas, then one gets Lévy's formula from Khinchin's formula, and, in addition, Λ satisfies (7).

5. Exercise*. Let a measure Λ satisfy (7). Define

$$\mu(\Gamma) = \int_\Gamma \frac{x^2}{1+x^2}\,\Lambda(dx) + I_\Gamma(0)\sigma^2.$$

Show that μ is a finite measure for which Lévy's formula transforms into Khinchin's formula.

6. Theorem (uniqueness). *There can exist only one finite measure μ and one number b for which $\varphi(t,\lambda)$ is representable by Khinchin's formula. There can exist only one measure Λ satisfying (7) and unique numbers b and σ^2 for which $\varphi(t,\lambda)$ is representable by Lévy's formula.*

Proof. Exercises 4 and 5 show that we may concentrate only on the first part of the theorem. The exponent in Khinchin's formula is continuous in λ and vanishes at $\lambda = 0$. Therefore it is uniquely determined by $\varphi(t,\lambda)$, and we only need prove that μ and b are uniquely determined by the function

$$g(\lambda) := \int_{\mathbb{R}} f(\lambda, x)\,\mu(dx) + ib\lambda.$$

Clearly, it suffices only to show that μ is uniquely determined by g.

For $h > 0$, we have

$$g(\lambda) - \frac{g(\lambda+h)+g(\lambda-h)}{2} = \int_{\mathbb{R}} e^{i\lambda x}\,\frac{1-\cos xh}{x^2}\,(1+x^2)\,\mu(dx) \qquad (8)$$

with the agreement that $(1-\cos xh)/x^2 = h^2/2$ if $x = 0$. Define a new measure

$$\nu_h(\Gamma) = \int_\Gamma \rho(x,h)\,\mu(dx), \quad \rho(x,h) = \frac{1-\cos xh}{x^2}\,(1+x^2)$$

and use

$$\int_{\mathbb{R}} f(x)\,\nu_h(dx) = \int_{\mathbb{R}} f(x)\rho(x,h)\,\mu(dx)$$

for all bounded Borel f. Then we see from (8) that the characteristic function of ν_h is uniquely determined by g. Therefore, ν_h is uniquely determined by g for any $h > 0$.

Now let Γ be a bounded Borel set and h be such that $\Gamma \subset [-1/h, 1/h]$. Take $f(x) = \rho^{-1}(x,h)$ for $x \in \Gamma$ and $f(x) = 0$ elsewhere. By the way, observe that f is a bounded Borel function. For this f

$$\int_{\mathbb{R}} f(x)\,\nu_h(dx) = \int_{\mathbb{R}} f(x)\rho(x,h)\,\mu(dx) = \mu(\Gamma),$$

where the left-hand side is uniquely determined by g. The theorem is proved.

7. Corollary. *Define*

$$\mu_t(dx) = \frac{x^2}{t(1+x^2)}\,F_t(dx), \quad b_t = \frac{1}{t}\int_{\mathbb{R}} \sin x\,F_t(dx).$$

Then $\mu_t \to \mu$ weakly and $b_t \to b$ as $t \downarrow 0$.

Indeed, similarly to (3) we have

$$\frac{1}{t}\int_{\mathbb{R}} (e^{i\lambda x} - 1)\,F_t(dx) \to g(\lambda),$$

which as in the proof of the Lévy-Khinchin theorem shows that the family $\{\mu_t; t \leq 1\}$ is weakly compact. Next, if $\mu_{t_{n'}} \xrightarrow{w} \nu$, then, again as in the proof of the Lévy-Khinchin theorem, $b_{t_{n'}}$ converges, and if we denote its limit by c, then Khinchin's formula holds with $\mu = \nu$ and $b = c$. Finally, the uniqueness implies that all weak limit points of μ_t, $t \downarrow 0$, coincide with μ and hence (cf. Exercise 1.2.10) $\mu(t) \xrightarrow{w} \mu$ as $t \downarrow 0$. This obviously implies that b_t also converges and its limit is b.

8. Corollary. *In Lévy's formula*

$$\sigma^2 = \lim_{n\to\infty} \lim_{t\downarrow 0} \frac{1}{t}\,E\xi_t^2 I_{|\xi_t| \leq \varepsilon_n},$$

where ε_n is a sequence such that $\varepsilon_n > 0$ and $\varepsilon_n \downarrow 0$. Moreover, F_t/t converges weakly on $\mathbb{R} \setminus \{0\}$ as $t \downarrow 0$ to Λ, that is,

$$\lim_{t\downarrow 0} \frac{1}{t}\int_{\mathbb{R}} f(x)\,F_t(dx) = \lim_{t\downarrow 0} \frac{1}{t}\,Ef(\xi_t) = \int_{\mathbb{R}} f(x)\,\Lambda(dx) \tag{9}$$

for each bounded continuous function f which vanishes in a neighborhood of 0.

Proof. By the definition of Λ and Corollary 7, for each bounded continuous function f which vanishes in a neighborhood of 0, we have

$$\int_{\mathbb{R}} f(x)\,\Lambda(dx) = \int_{\mathbb{R}} f(x)\,\frac{1+x^2}{x^2}\,\mu(dx)$$

$$= \lim_{t\downarrow 0} \int_{\mathbb{R}} f(x)\,\frac{1+x^2}{x^2}\,\mu_t(dx) = \lim_{t\downarrow 0} \frac{1}{t}\int_{\mathbb{R}} f(x)\,F_t(dx).$$

This proves (9).

Let us prove the first assertion. By the dominated convergence theorem, for every sequence of nonnegative $\varepsilon_n \to 0$ we have

$$\sigma^2 = \mu(\{0\}) = \int_{\mathbb{R}} I_{[0,0]}(x)\,\mu(dx)$$

$$= \int_{\mathbb{R}} I_{[0,0]}(x)(1+x^2)\,\mu(dx) = \lim_{n\to\infty} \int_{\mathbb{R}} I_{[-\varepsilon_n,\varepsilon_n]}(x)(1+x^2)\,\mu(dx).$$

By Theorem 1.2.11 (v), if $\mu(\{\varepsilon_n\}) = \mu(\{-\varepsilon_n\}) = 0$, then

$$\int_{\mathbb{R}} I_{[-\varepsilon_n,\varepsilon_n]}(x)(1+x^2)\,\mu(dx)$$

$$= \lim_{t\downarrow 0} \frac{1}{t} \int_{\mathbb{R}} I_{[-\varepsilon_n,\varepsilon_n]}(x)x^2\,F_t(dx) = \lim_{t\downarrow 0} \frac{1}{t}\,E\xi_t^2 I_{|\xi_t|\le\varepsilon_n}.$$

It only remains to notice that the set of x such that $\mu(\{x\}) > 0$ is countable, so that there exists a sequence ε_n such that $\varepsilon_n \downarrow 0$ and $\mu(\{\varepsilon_n\}) = \mu(\{-\varepsilon_n\}) = 0$. The corollary is proved.

9. Exercise. Prove that, if $\xi_t \ge 0$ for all $t \ge 0$ and ω, then $\Lambda((-\infty,0]) = 0$. One can say more in that case, as we will see in Exercise 3.15.

We know that the Wiener process has independent increments, and also that it is homogeneous and stochastically continuous (even just continuous). In Lévy's formula, to get $E\exp(i\lambda w_t)$ one takes $\Lambda = 0$, $b = 0$, and $\sigma = 1$.

If in Lévy's formula we take $\sigma = 0$, $\Lambda(\Gamma) = I_\Gamma(1)\mu$, and $b = \mu\sin 1$, where μ is a nonnegative number, then the corresponding process is called *the Poisson process with parameter* μ.

If $\sigma = b = 0$ and $\Lambda(dx) = ax^{-2}\,dx$ with a constant $a > 0$, the corresponding process is called *the Cauchy process*.

Clearly, for the Poisson process π_t with parameter μ we have

$$Ee^{i\lambda\pi_t} = e^{t\mu(e^{i\lambda}-1)},$$

so that π_t has Poisson distribution with parameter $t\mu$. In particular,

$$E|\pi_{t+h} - \pi_t| = E\pi_h = h\mu$$

for $t,h \ge 0$. The values of π_t are integers and π_t is not identically constant (the expectation grows). Therefore π_t does not have continuous modification, which shows, in particular, that the requirement $\beta > 0$ in Theorem 1.4.8 is essential. For $\mu = 1$ we come to the Poisson process introduced in Exercise 2.3.8.

10. Exercise. Prove that for the Cauchy process we have $\varphi(t,\lambda) = \exp(-ct|\lambda|)$, with a constant $c > 0$.

11. Exercise*. Prove that the Lévy measure of the process τ_{a+}, $a \geq 0$ (see Theorem 2.6.1, and Exercise 1.12) is concentrated on the positive half line and is given by $I_{x>0}(2\pi)^{-1/2}x^{-3/2}\,dx$. This result will be used in Sec. 6.

You may also like to show that

$$\varphi(t,\lambda) = \exp(-t|\lambda|^{1/2}(a - ib\,\mathrm{sign}\,\lambda)),$$

where

$$a = (2\pi)^{-1/2}\int_0^\infty x^{-3/2}(1-\cos x)\,dx, \quad b = (2\pi)^{-1/2}\int_0^\infty x^{-3/2}\sin x\,dx,$$

and, furthermore, that $a = b = 1$.

12. Exercise. Prove that if in Lévy's formula we have $\Lambda = 0$ and $\sigma = 0$, then $\xi_t = bt$ (a.s.) for all t, where b is a constant.

3. Jump measures and their relation to Lévy measures

Let ξ_t be an infinitely divisible cadlag process on $[0,\infty)$. Define

$$\Delta\xi_t = \xi_t - \xi_{t-}.$$

For any set $\Gamma \subset \mathbb{R}_+ \times \mathbb{R} := [0,\infty) \times \mathbb{R}$ let $p(\Gamma)$ be the number of points $(t,\Delta\xi_t) \in \Gamma$. It may happen that $p(\Gamma) = \infty$. Obviously $p(\Gamma)$ is a σ-additive measure on the family of all subsets of $\mathbb{R}_+ \times \mathbb{R}$. The measure $p(\Gamma)$ is called *the jump measure of* ξ_t.

For $T, \varepsilon \in (0,\infty)$ define

$$R_{T,\varepsilon} = [0,T] \times \{x : |x| \geq \varepsilon\}.$$

1. Remark. Notice that $p(R_{T,\varepsilon}) < \infty$ for any ω, which is to say that on $[0,T]$ there may be only finitely many t such that $|\Delta\xi_t| \geq \varepsilon$. This property follows immediately from the fact that the trajectories of ξ_t do not have discontinuities of the second kind. It is also worth noticing that $p(\Gamma)$ is concentrated at points $(t,\Delta\xi_t)$ and each point of this type receives a unit mass.

We will need yet another measure defined on subsets of $\mathbb{R}$. For any $B \subset \mathbb{R}$ define

$$p_t(B) = p((0,t] \times B).$$

2. Remark. By Remark 1, if B is separated from zero, then $p_t(B)$ is finite. Moreover, let $f(x)$ be a Borel function (perhaps unbounded) vanishing for $|x| < \varepsilon$, where $\varepsilon > 0$. Then, the process

$$\eta_t := \eta_t(f) := \int_{\mathbb{R}} f(x)\, p_t(dx)$$

is well defined and is just equal to the (finite) sum of $f(\Delta\xi_s)$ for all $s \le t$ such that $|\Delta\xi_s| \ge \varepsilon$.

The structure of η_t is pretty simple. Indeed, fix an ω and let $0 \le s_1 < ... < s_n < ...$ be all s for which $|\Delta\xi_s| \ge \varepsilon$ (if there are only $m < \infty$ such s, we let $s_n = \infty$ for $n \ge m+1$). Then, of course, $s_n \to \infty$ as $n \to \infty$. Also $s_1 > 0$, because ξ_t is right continuous and $\xi_0 = 0$. With this notation

$$\eta_t = \sum_{s_n \le t} f(\Delta\xi_{s_n}). \tag{1}$$

We see that η_t starts from zero, is constant on each interval $[s_{n-1}, s_n)$, $n = 1, 2, ...$ (with $s_0 := 0$), and

$$\Delta\eta_{s_n} = f(\Delta\xi_{s_n}). \tag{2}$$

3. Lemma. *Let $f(x)$ be a function as in Remark 2. Assume that f is continuous. Let $0 \le t < \infty$ and t_i^n be such that*

$$s = t_1^n < ... < t_{k(n)+1}^n = t, \qquad \max_{j=1,...,k(n)} (t_{j+1}^n - t_j^n) \to 0$$

as $n \to \infty$. Then for any ω

$$\eta_t(f) - \eta_s(f) = \int_{\mathbb{R}_+ \times \mathbb{R}} I_{(s,t]}(u) f(x)\, p(dudx) = \lim_{n\to\infty} \sum_{j=1}^{k(n)} f(\xi_{t_{j+1}^n} - \xi_{t_j^n}). \tag{3}$$

Proof. We have noticed above that the set of all $u \in (s,t]$ for which $|\Delta\xi_u| \ge \varepsilon$ is finite. Let $\{u_1, ..., u_N\}$ be this set. Single out those intervals $(t_j^h, t_{j+1}^n]$ which contain at least one of the u_i's. For large n we will have exactly N such intervals. First we prove that, for large n,

$$|\xi_{t_{j+1}^n} - \xi_{t_j^n}| < \varepsilon, \quad f(\xi_{t_{j+1}^n} - \xi_{t_j^n}) = 0$$

if the interval $(t_j^n, t_{j+1}^n]$ does not contain any of the u_i's. Indeed, if this were not true, then one could find a sequence s_k, t_k such that $|\xi_{t_k} - \xi_{s_k}| \ge \varepsilon$, $s_k, t_k \in (s,t]$, $s_k < t_k$, $t_k - s_k \to 0$, and on $(s_k, t_k]$ there are no points u_i.

Without loss of generality, we may assume that $s_k, t_k \to u \in (s,t]$ (actually, one can obviously assume that $u \in [s,t]$, but since the trajectories are right continuous, $\xi_{s_k}, \xi_{t_k} \to \xi_s$ if $s_k, t_k \to s$, so that $u \neq s$).

Furthermore, there are infinitely many s_k's either to the right or to the left of u. Therefore, using subsequences if needed, we may assume that the sequence s_k is monotone and then that t_k is monotone as well. Then, since ξ_t has finite right and left limits, we have that $s_k \uparrow u$, $s_k < u$, and $t_k \downarrow u$, which implies that $|\Delta\xi_u| \geq \varepsilon$. But then we would have a point $u \in \{u_1, ..., u_N\}$ which belongs to $(s_k, t_k]$ for all k (after passing to subsequences). This is a contradiction, which proves that for all large n the sum on the right in (3) contains at most N nonzero terms. These terms correspond to the intervals $(t_j, t_{j+1}]$ containing u_i's, and they converge to $f(\Delta\xi_{u_i})$.

It only remains to observe that the first equality in (3) is obvious and, by Remark 2,

$$\int_{\mathbb{R}_+\times\mathbb{R}} I_{(s,t]}(u) f(x)\, p(dudx) = \sum_{i=1}^{N} f(\Delta\xi_{u_i}).$$

The lemma is proved.

4. Definition. For $0 \leq s < t < \infty$ define $\mathcal{F}^{\xi}_{s,t}$ as the completion of the σ-field generated by $\xi_r - \xi_s$, $r \in [s,t]$. Also set $\mathcal{F}^{\xi}_t = \mathcal{F}^{\xi}_{0,t}$.

5. Remark. Since the increments of ξ_t are independent, the σ-fields $\mathcal{F}^{\xi}_{0,t_1}$, $\mathcal{F}^{\xi}_{t_1,t_2}$,..., $\mathcal{F}^{\xi}_{t_{n-1},t_n}$ are independent for any $0 < t_1 < ... < t_n$.

Next remember Definition 2.5.10.

6. Definition. Random processes $\eta^1_t, ..., \eta^n_t$ defined for $t \geq 0$ are called *independent* if for any $t_1, ..., t_k \geq 0$ the vectors $(\eta^1_{t_1}, ..., \eta^1_{t_k}), ..., (\eta^n_{t_1}, ..., \eta^n_{t_k})$ are independent.

7. Lemma. *Let ζ_t be an $\mathbb{R}^d$-valued process starting from zero and such that $\zeta_t - \zeta_s$ is $\mathcal{F}^{\xi}_{s,t}$-measurable whenever $0 \leq s < t < \infty$. Also assume that, for all $0 \leq s < t < \infty$, the random variables $\zeta^1_t - \zeta^1_s$,..., $\zeta^d_t - \zeta^d_s$ are independent. Then the process ζ_t has independent increments and the processes ζ^1_t,..., ζ^d_t are independent.*

Proof. That ζ_t has independent increments follows from Remark 5. To prove that the vectors

$$(\zeta^1_{t_1}, ..., \zeta^1_{t_n}), ..., (\zeta^d_{t_1}, ..., \zeta^d_{t_n}) \tag{4}$$

are independent if $0 = t_0 < t_1 < ..., < t_n$, it suffices to prove that

$$(\zeta^1_{t_1} - \zeta^1_{t_0}, \zeta^1_{t_2} - \zeta^1_{t_1}, ..., \zeta^1_{t_n} - \zeta^1_{t_{n-1}}), ..., (\zeta^d_{t_1} - \zeta^d_{t_0}, \zeta^d_{t_2} - \zeta^d_{t_1}, ..., \zeta^d_{t_n} - \zeta^1_{t_{n-1}}) \tag{5}$$

are independent. Indeed, the vectors in (4) can be obtained after applying a linear transformation to the vectors in (5). Now take $\lambda^k_j \in \mathbb{R}$ for $k = 1, ..., d$ and $j = 1, .., n$, and write

$$E \exp\big(i \sum_{k,j} \lambda^k_j (\zeta^k_{t_j} - \zeta^k_{t_{j-1}})\big)$$

$$= E \exp\big(i \sum_{j \le n-1,k} \lambda^k_j (\zeta^k_{t_j} - \zeta^k_{t_{j-1}})\big) E\{\exp\big(i \sum_k \lambda^k_n (\zeta^k_{t_n} - \zeta^k_{t_{n-1}})\big) | \mathcal{F}^\xi_{0,t_{n-1}}\}$$

$$= E \exp\big(i \sum_{j \le n-1,k} \lambda^k_j (\zeta^k_{t_j} - \zeta^k_{t_{j-1}})\big) E \exp\big(i \sum_k \lambda^k_n (\zeta^k_{t_n} - \zeta^k_{t_{n-1}})\big)$$

$$= E \exp\big(i \sum_{j \le n-1,k} \lambda^k_j (\zeta^k_{t_j} - \zeta^k_{t_{j-1}})\big) \prod_k E \exp\big(i \lambda^k_n (\zeta^k_{t_n} - \zeta^k_{t_{n-1}})\big).$$

An obvious induction allows us to represent the characteristic function of the family $\{\zeta^k_{t_j} - \zeta^k_{t_{j-1}}, k = 1, ..., d, j = 1, ..., n\}$ as the product of the characteristic functions of its members, thus proving the independence of all $\zeta^k_{t_j} - \zeta^k_{t_{j-1}}$ and, in particular, of the vectors (5). The lemma is proved.

8. Lemma. *Let f be as in Remark 2 and let f be continuous. Take $\alpha \in \mathbb{R}$ and denote $\zeta_t = \eta_t(f) + \alpha \xi_t$. Then*

(i) *for every $0 \le s < t < \infty$, the random variable $\zeta_t - \zeta_s$ is $\mathcal{F}^\xi_{s,t}$-measurable;*

(ii) *the process ζ_t is an infinitely divisible cadlag process and*

$$E e^{i\zeta_t} = \exp t\Big\{\int_{\mathbb{R}} (e^{i(f(x)+\alpha x)} - 1 - i\alpha \sin x)\, \Lambda(dx) + i\alpha b - \alpha^2\sigma^2/2\Big\}. \tag{6}$$

Proof. Assertion (i) is a trivial consequence of (3). In addition, Remark 5 shows that ζ_t has independent increments.

(ii) The homogeneity of ζ_t follows immediately from (3) and the similar property of ξ_t. Furthermore, Remark 2 shows that ζ_t is cadlag. From the homogeneity and right continuity of ζ_t we get

$$\lim_{s \uparrow t} E e^{i\lambda(\zeta_t - \zeta_s)} = \lim_{s \uparrow t} E e^{i\lambda \zeta_{t-s}} = E e^{i\lambda \zeta_0} = 1, \quad t > 0.$$

Similar equations hold for $s \downarrow t$ with $t \geq 0$. Therefore, $\zeta_s \xrightarrow{P} \zeta_t$ as $s \to t$, and ζ_t is stochastically continuous.

To prove (6), take Khinchin's measure μ and take μ_t and b_t from Corollary 2.7. Also observe that

$$\lim_{n\to\infty} a_n^n = \lim_{n\to\infty} e^{n\log a_n} = \lim_{n\to\infty} e^{n(a_n-1)}$$

provided $a_n \to 1$ and one of the limits exists. Then we have

$$Ee^{i(\eta_t+\alpha\xi_t)} = \lim_{n\to\infty}\left(Ee^{i(f(\xi_{t/n})+\alpha\xi_{t/n})}\right)^n$$

$$= \lim_{n\to\infty} \exp n\int_{\mathbb{R}} (e^{i(f(x)+\alpha x)} - 1)\, F_{t/n}(dx),$$

with

$$\lim_{n\to\infty} n\int_{\mathbb{R}} (e^{i(f(x)+\alpha x)} - 1)\, F_{t/n}(dx)$$

$$= \lim_{n\to\infty} t\int_{\mathbb{R}} (e^{i(f(x)+\alpha x)} - 1 - i\alpha\sin x)(1+x^2)/x^2\,\mu_{t/n}(dx) + i\alpha t b$$

$$= t\int_{\mathbb{R}} (e^{i(f(x)+\alpha x)} - 1 - i\alpha\sin x)(1+x^2)/x^2\,\mu(dx) + i\alpha t b.$$

Now to get (6) one only has to refer to Exercise 2.4. The lemma is proved.

9. Theorem. (i) *For $ab > 0$ the process $p_t(a,b]$ is a Poisson process with parameter $\Lambda((a,b])$, and, in particular,*

$$Ep_t(a,b] = t\Lambda((a,b]); \tag{7}$$

(ii) *if $a_m < b_m$, $a_m b_m > 0$, $m = 1, ..., n$, and the intervals $(a_m, b_m]$ are pairwise disjoint, then the processes $p_t(a_1,b_1], ..., p_t(a_n,b_n]$ are independent.*

Proof. To prove (i), take a sequence of bounded continuous functions $f_k(x)$ such that $f_k(x) \to \lambda I_{(a,b]}(x)$ as $k \to \infty$ and $f_k(x) = 0$ for $|x| < \varepsilon := (|a| \wedge |b|)/2$. Then, for each ω,

$$\int_{\mathbb{R}} f_k(x)\, p_t(dx) \to \lambda p_t(a,b]. \tag{8}$$

Moreover, $|\exp\{if_k(x)\} - 1| \leq 2I_{|x|\geq\varepsilon}$ and

$$\int_{\mathbb{R}} I_{|x|\geq\varepsilon}\,\Lambda(dx) \leq \frac{1+\varepsilon^2}{\varepsilon^2}\int_{\mathbb{R}} \frac{x^2}{1+x^2}\,\Lambda(dx) < \infty. \tag{9}$$

Hence, by Lemma 8 and by the dominated convergence theorem,

$$Ee^{i\lambda p_t(a,b]} = \exp t\int_{\mathbb{R}} (e^{i\lambda I_{(a,b]}(x)} - 1)\,\Lambda(dx) = \exp\{t\Lambda((a,b])(e^{i\lambda}-1)\}.$$

The homogeneity of $p_t(a,b]$ and independence of its increments follow from (8) and Lemma 8. Remark 2 shows that $p_t(a,b]$ is a cadlag process. As in Lemma 8, this leads to the conclusion that $p_t(a,b]$ is stochastically continuous. This proves (i).

(ii) Formula (8) and Lemma 8 imply that $p_t(a,b] - p_s(a,b]$ is $\mathcal{F}^{\xi}_{s,t}$-measurable if $s<t$. By Lemma 7, to prove that the processes $p_t(a_1,b_1]$,..., $p_t(a_n,b_n]$ are independent, it suffices to prove that, for any $s<t$, the random variables

$$p_t(a_1,b_1] - p_s(a_1,b_1], ..., p_t(a_n,b_n] - p_s(a_n,b_n] \tag{10}$$

are independent.

Take $\lambda_1,...,\lambda_n \in \mathbb{R}$ and define $f(x) = \lambda_m$ for $x \in (a_m,b_m]$ and $f = 0$ outside the union of the $(a_m,b_m]$. Also take a sequence of bounded continuous functions f_n vanishing in a neighborhood of zero such that $f_n(x) \to f(x)$ for all $x \in \mathbb{R}$. Then

$$\eta_t(f_n) - \eta_s(f_n) \to \eta_t(f) - \eta_s(f) = \sum_{m=1}^{n} \lambda_m\{p_t(a_m,b_m] - p_s(a_m,b_m]\}.$$

Hence and from Lemma 8 we get

$$E\exp(i\sum_{m=1}^{n} \lambda_m\{p_t(a_m,b_m] - p_s(a_m,b_m]\}) = \lim_{n\to\infty} Ee^{i(\eta_t(f_n)-\eta_s(f_n))}$$

$$= \lim_{n\to\infty} Ee^{i\eta_{t-s}(f_n)} = \lim_{n\to\infty} \exp\{(t-s)\int_{\mathbb{R}} (e^{if_n(x)}-1)\,\Lambda(dx)\}$$

$$= \exp\{(t-s)\int_{\mathbb{R}} (e^{if(x)}-1)\,\Lambda(dx)\} = \prod_{m=1}^{n} \exp\{(t-s)\Lambda((a_m,b_m])(e^{i\lambda_m}-1)\}.$$

This and assertion (i) prove that the random variables in (5) are independent. The theorem is proved.

10. Corollary. *Let f be a Borel nonnegative function. Then, for each $t \geq 0$,*

$$\int_{\mathbb{R}\setminus\{0\}} f(x)\, p_t(dx)$$

is a random variable and

$$E\int_{\mathbb{R}\setminus\{0\}} f(x)\, p_t(dx) = t\int_{\mathbb{R}} f(x)\, \Lambda(dx). \tag{11}$$

Notice that on the right in (11) we write the integral over $\mathbb{R}$ instead of $\mathbb{R}\setminus\{0\}$ because $\Lambda(\{0\}) = 0$ by definition. To prove the assertion, take $\varepsilon > 0$ and let Σ be the collection of all Borel Γ such that $p_t(\Gamma\setminus(-\varepsilon,\varepsilon))$ is a random variable and

$$\nu_\varepsilon(\Gamma) := Ep_t(\Gamma\setminus(-\varepsilon,\varepsilon)) = t\Lambda_\varepsilon(\Gamma) := t\Lambda(\Gamma\setminus(-\varepsilon,\varepsilon)).$$

It follows from (7) and from the finiteness of $\Lambda(\mathbb{R}\setminus(-\varepsilon,\varepsilon))$ that $\mathbb{R} \in \Sigma$. By adding an obvious argument we conclude that Σ is a λ-system. Furthermore, from Theorem 9 (i) we know that Σ contains $\Pi := \{(a,b] : ab > 0\}$, which is a π-system. Therefore, $\Sigma = \mathfrak{B}(\mathbb{R})$. Now a standard measure-theoretic argument shows that, for every Borel nonnegative f, we have

$$E\int_{\mathbb{R}\setminus(-\varepsilon,\varepsilon)} f(x)\, p_t(dx) = \int_{\mathbb{R}} f(x)\, \nu_\varepsilon(dx)$$

$$= t\int_{\mathbb{R}} f(x)\, \Lambda_\varepsilon(dx) = t\int_{\mathbb{R}\setminus(-\varepsilon,\varepsilon)} f(x)\, \Lambda(dx).$$

It only remains to let $\varepsilon \downarrow 0$ and use the monotone convergence theorem.

11. Corollary. *Every continuous infinitely divisible process has the form $bt + \sigma w_t$, where σ and b are the constants from Lévy's formula and w_t is a Wiener process if $\sigma \neq 0$ and $w_t \equiv 0$ if $\sigma = 0$.*

Indeed, for a continuous ξ_t we have $p_t(\alpha,\beta] = 0$ if $\alpha\beta > 0$. Hence $\Lambda((\alpha,\beta]) = 0$ and $\varphi(t,\lambda) = \exp\{ibt\lambda - \sigma^2\lambda^2 t/2\}$. For $\sigma \neq 0$, it follows that $\eta_t := (\xi_t - bt)/\sigma$ is a continuous process with independent increments, $\eta_0 = 0$, and $\eta_t - \eta_s \sim N(0, |t-s|)$. As we know, η_t is a Wiener process. If $\sigma = 0$, then $\xi_t - bt = 0$ (a.s.) for any t and, actually, $\xi_t - bt = 0$ for all t at once (a.s.) since $\xi_t - bt$ is continuous.

12. Corollary. *Let an open set $G \subset \mathbb{R}\setminus\{0\}$ be such that $\Lambda(G) = 0$. Then there exists $\Omega' \in \mathcal{F}$ such that $P(\Omega') = 1$ and, for each $t \geq 0$ and $\omega \in \Omega'$, $\Delta\xi_t(\omega) \notin G$.*

Indeed, represent G as a countable union (perhaps with intersections) of intervals $(a_m, b_m]$. Since $\Lambda((a_m, b_m]) = 0$, we have $Ep_t(a_m, b_m] = 0$ and $p_t(a_m, b_m] = 0$ (a.s.). Adding to this that $p_t(a_m, b_m]$ increases in t, we conclude that $p_t(a_m, b_m] = 0$ for all t (a.s.). Now let

$$\Omega' = \bigcap_m \{\omega : p_t(a_m, b_m] = 0 \quad \forall t \geq 0\}.$$

Then $P(\Omega') = 1$ and

$$p((0,t] \times G) \leq \sum_m p_t(a_m, b_m] = 0$$

for each $\omega \in \Omega'$ and $t \geq 0$, as asserted.

The following corollary will be used for deriving an integral representation of ξ_t through jump measures.

13. Corollary. *Denote $q_t(a,b] = p_t(a,b] - t\Lambda((a,b])$. Let some numbers satisfying $a_i \leq b_i$ and $a_i b_i > 0$ be given for $i = 1, 2$. Then, for all $t, s \geq 0$,*

$$Eq_t(a_1, b_1]q_s(a_2, b_2] = (s \wedge t)\Lambda((a_1, b_1] \cap (a_2, b_2]). \tag{12}$$

Indeed, without loss of generality assume $t \geq s$. Notice that both parts of (12) are additive in the sense that if, say, $(a_1, b_1] = (a_3, b_3] \cup (a_4, b_4]$ and $(a_3, b_3] \cap (a_4, b_4] = \emptyset$, then

$$q_t(a_1, b_1] = q_t(a_3, b_3] + q_t(a_4, b_4],$$

$$\Lambda((a_1, b_1] \cap (a_2, b_2]) = \Lambda((a_3, b_3] \cap (a_2, b_2]) + \Lambda((a_4, b_4] \cap (a_2, b_2]).$$

It follows easily that to prove (12) it suffices to prove it only for two cases: (i) $(a_1, b_1] \cap (a_2, b_2] = \emptyset$ and (ii) $a_1 = a_2$, $b_1 = b_2$.

In the first case (12) follows from the independence of the processes $p_\cdot(a_1, b_1]$ and $p_\cdot(a_2, b_2]$ and from (7). In the second case, it suffices to remember that the variance of a random variable having the Poisson distribution with parameter Λ is Λ and use the fact that

$$q_t(a,b] = q_s(a,b] + (q_t(a,b] - q_s(a,b]),$$

where the summands are independent.

We will also use the following theorem, which is closely related to Theorem 9.

14. Theorem. *Take $a>0$ and define*

$$\eta_t = \int_{[a,\infty)} x\, p_t(dx) + \int_{(-\infty,-a]} x\, p_t(dx). \tag{13}$$

Then:

(i) *the process η_t is infinitely divisible, cadlag, with $\sigma = b = 0$ and Lévy measure $\Lambda(\Gamma \setminus (-a,a))$;*

(ii) *the process $\xi_t - \eta_t$ is infinitely divisible, cadlag, and does not have jumps larger in magnitude than a;*

(iii) *the processes η_t and $\xi_t - \eta_t$ are independent.*

Proof. Assertion (i) is proved like the similar assertion in Theorem 9 on the basis of Lemma 8. Indeed, take a sequence of continuous functions $f_k(x) \to x(1 - I_{(-a,a)}(x))$ such that $f_k(x) = 0$ for $|x| \le a/2$. Then, for any ω,

$$\int_{\mathbb{R}} f_k(x)\, p_t(dx) \to \eta_t. \tag{14}$$

This and Lemma 8 imply that η_t is a homogeneous process with independent increments. That it is cadlag follows from Remark 2. The stochastic continuity of η_t follows from its right continuity and homogeneity as in Lemma 8. To find the Lévy measure of η_t, observe that $|\exp\{i\lambda f_k(x)\} - 1| \le 2I_{|x|\ge a/2}$. By using (9), Lemma 8, and the dominated convergence theorem, we conclude that

$$Ee^{i\lambda\eta_t} = \exp t \int_{\mathbb{R}} (e^{i\lambda x(1-I_{(-a,a)}(x))} - 1)\, \Lambda(dx) = \exp t \int_{\mathbb{R}\setminus(-a,a)} (e^{i\lambda x} - 1)\, \Lambda(dx).$$

In assertion (ii) the fact that $\xi_t - \eta_t$ is an infinitely divisible cadlag process is proved as above on the basis of Lemma 8. The assertion about its jumps is obvious because of Remark 2. Another explanation of the same fact can be obtained from Lemma 8, which implies that

$$Ee^{i(\lambda\eta_t+\alpha\xi_t)}$$

$$= \exp t\Big\{ \int_{\mathbb{R}} (e^{i\lambda x(1-I_{(-a,a)}(x))+i\alpha x} - 1 - i\alpha \sin x)\,\Lambda(dx) + i\alpha b - \alpha^2\sigma^2/2\Big\}$$

$$= \exp t\Big\{ \int_{(-a,a)} (e^{i\alpha x} - 1 - i\alpha \sin x)\,\Lambda(dx)$$

$$+ \int_{\mathbb{R}\setminus(-a,a)} (e^{i(\lambda+\alpha)x} - 1 - i\alpha \sin x)\,\Lambda(dx) + i\alpha b - \alpha^2\sigma^2/2\Big\}, \quad (15)$$

where, for $\lambda = -\alpha$, the expression in the last braces is

$$\int_{(-a,a)} (e^{i\alpha x} - 1 - i\alpha \sin x)\,\Lambda(dx) + i\alpha(b - \int_{\mathbb{R}\setminus(-a,a)} \sin x\,\Lambda(dx)) - \alpha^2\sigma^2/2,$$

which shows that the Lévy measure of $\xi_t - \eta_t$ is concentrated on $(-a, a)$.

To prove (iii), first take $\lambda = \beta - \alpha$ in (15). Then we see that

$$Ee^{i\beta\eta_t+i\alpha(\xi_t-\eta_t)} = e^{tg},$$

where

$$g = \int_{\mathbb{R}\setminus(-a,a)} (e^{i\beta x} - 1)\,\Lambda(dx)$$

$$+ \int_{(-a,a)} (e^{i\alpha x} - 1 - i\alpha \sin x)\,\Lambda(dx) + i\alpha(b - \int_{\mathbb{R}\setminus(-a,a)} \sin x\,\Lambda(dx)) - \alpha^2\sigma^2/2,$$

so that $Ee^{i\beta\eta_t+i\alpha(\xi_t-\eta_t)} = Ee^{i\beta\eta_t}Ee^{i\alpha(\xi_t-\eta_t)}$. Hence, for any t, η_t and $\xi_t - \eta_t$ are independent.

Furthermore, for any constants $\lambda, \alpha \in \mathbb{R}$, the process $\lambda\eta_t + \alpha(\xi_t - \eta_t) = (\lambda - \alpha)\eta_t + \alpha\xi_t$ is a homogeneous process, which is proved as above by using Lemma 8. It follows that the two-dimensional process $(\eta_t, \xi_t - \eta_t)$ has homogeneous increments. In particular, if $s < t$, the distributions of $(\eta_{t-s}, \xi_{t-s} - \eta_{t-s})$ and $(\eta_t - \eta_s, \xi_t - \eta_t - (\xi_s - \eta_s))$ coincide, and since the first pair is independent, so is the second. Now the independence of the processes η_t and $\xi_t - \eta_t$ follows from Lemma 7 and from the fact that $\eta_t - \eta_s$, $\xi_t - \xi_s$, and $(\eta_t - \eta_s, \xi_t - \eta_t - (\xi_s - \eta_s))$ are $\mathcal{F}^{\xi}_{s,t}$-measurable (see (14) and Lemma 8). The theorem is proved.

The following exercise describes all nonnegative infinitely divisible cadlag processes.

15. Exercise. Let ξ_t be an infinitely divisible cadlag process satisfying $\xi_t \geq 0$ for all $t \geq 0$ and ω. Take $\eta_t = \eta_t(a)$ from Theorem 14.

(i) By using Exercise 2.9, show that all jumps of ξ_t are nonnegative.

(ii) Prove that for every $t \geq 0$, we have $P(\eta_t(a) = 0) = \exp(-t\Lambda([a,\infty)))$.

(iii) From Theorem 14 and (ii), derive that $\xi_t - \eta_t(a) \geq 0$ (a.s.) for each $t \geq 0$.

(iv) Since obviously $\eta_t(a)$ increases as a decreases, conclude that $\eta_t(0+) \leq \xi_t < \infty$ (a.s.) for each $t \geq 0$. From (15) with $\alpha = 0$ find the characteristic function of $\eta_t(0+)$ and prove that $\xi_t - \eta_t(0+)$ has normal distribution. By using that $\xi_t - \eta_t(0+) \geq 0$ (a.s.), prove that $\xi_t = \eta_t(0+)$ (a.s.).

(v) Prove that

$$\int_0^1 x\Lambda(dx) < \infty, \quad \xi_t = \int_{(0,\infty)} x\, p(t,dx) \quad \text{(a.s.)},$$

and, in particular, ξ_t is a pure jump process with nonnegative jumps.

4. Further comments on jump measures

1. Exercise. Let $f(t,x)$ be a Borel nonnegative function such that $f(t,0) = 0$. Prove that $\int_{\mathbb{R}_+\times\mathbb{R}} f(s,x)\, p(dsdx)$ is a random variable and

$$E\int_{\mathbb{R}_+\times\mathbb{R}} f(s,x)\, p(dsdx) = \int_{\mathbb{R}_+\times\mathbb{R}} f(s,x)\, ds\Lambda(dx). \tag{1}$$

2. Exercise. Let $f(t,x) = f(\omega,t,x)$ be a bounded function such that $f = 0$ for $|x| < \varepsilon$ and for $t \geq T$, where the constants $\varepsilon, T \in (0,\infty)$. Also assume that $f(\omega,t,x)$ is left continuous in t for any (ω,x) and $\mathcal{F}_t^\xi \otimes \mathfrak{B}(\mathbb{R})$-measurable for any t. Prove that the following version of (1) holds:

$$E\int_{\mathbb{R}_+\times\mathbb{R}} f(s,x)\, p(dsdx) = \int_{\mathbb{R}_+\times\mathbb{R}} Ef(s,x)\, ds\Lambda(dx).$$

The following two exercises are aimed at generalizing Theorem 3.9.

3. Exercise. Let $f(t,x)$ be a bounded Borel function such that $f = 0$ for $|x| < \varepsilon$, where the constant $\varepsilon > 0$. Prove that, for $t \in [0,\infty)$,

$$\varphi(t) := E\exp\{i\int_{(0,t]\times\mathbb{R}} f(s,x)\, p(dsdx)\} = \exp\{\int_{(0,t]\times\mathbb{R}} (e^{if(s,x)}-1)\, ds\Lambda(dx)\}.$$

4. Exercise. By taking f in Exercise 3 as linear combinations of the indicators of Borel subsets $\Gamma_1, ..., \Gamma_n$ of $\mathbb{R}_+ \times \mathbb{R}$, prove that, if the sets are disjoint, then $p(\Gamma_1), ..., p(\Gamma_n)$ are independent. Also prove that, if $\Gamma_1 \subset R_{T,\varepsilon}$, then $p(\Gamma_1)$ is Poisson with parameter $(\ell \times \Lambda)(\Gamma_1)$.

The following exercise shows that Poisson processes without common jumps are independent.

5. Exercise. Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $\mathcal{F}_t$ be σ-fields defined for $t \geq 0$ such that $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for $s \leq t$. Assume that ξ_t and η_t are two Poisson processes with parameters μ and ν respectively defined on Ω, and such that ξ_t and η_t are $\mathcal{F}$-measurable for each t and $\xi_{t+h} - \xi_t$ and $\eta_{t+h} - \eta_t$ are independent of $\mathcal{F}_t$ for all $t, h \geq 0$. Finally, assume that ξ_t and η_t do not have common jumps, that is, $(\Delta\xi_t)\Delta\eta_t = 0$ for all t and ω. Prove that the processes ξ_t and η_t are independent.

5. Representing infinitely divisible processes through jump measures

We start with a simple result.

1. Theorem. *Let ξ_t be an infinitely divisible cadlag process with parameters σ, b, and Lévy measure concentrated at points $x_1, ..., x_n$.*

(i) *If $\sigma \neq 0$, then there exist a Wiener process w_t and Poisson processes p_t^1,..., p_t^n with parameters $\Lambda(\{x_1\}), ..., \Lambda(\{x_n\})$, respectively, such that w_t, p_t^1,..., p_t^n are mutually independent and*

$$\xi_t = x_1 p_t^1 + ... + x_n p_t^n + bt + \sigma w_t \quad \forall t \geq 0 \quad (a.s.). \tag{1}$$

(ii) *If $\sigma = 0$, assertion* (i) *still holds if one does not mention w_t and drops the term σw_t in* (1).

Proof. (i) Of course, we assume that $x_i \neq x_j$ for $i \neq j$. Notice that $\Lambda(\{0\}) = 0$. Therefore, $x_m \neq 0$. Also

$$\Lambda(\mathbb{R} \setminus \{x_1, ..., x_n\}) = 0.$$

Hence, by Corollary 3.12, we may assume that all jumps of ξ_t belong to the set $\{x_1, ..., x_n\}$.

Now take $a > 0$ such that $a < |x_i|$ for all i, and define η_t by (3.13). By Theorem 3.14 the process $\xi_t - \eta_t$ does not have jumps and is infinitely divisible. By Corollary 3.11 we conclude that

$$\xi_t - \eta_t = bt + \sigma w_t.$$

In addition, formula (3.1) shows also that

$$\eta_t = x_1 p_t(\{x_1\}) + ... + x_n p_t(\{x_n\}) = x_1 p_t(a_1, b_1] + ... + x_n p_t(a_n, b_n],$$

where a_m, b_m are any numbers satisfying $a_m b_m > 0$, $a_m < x_m \leq b_m$, and such that $(a_m, b_m]$ are mutually disjoint. This proves (1) with $p_t^m = p_t(a_m, b_m]$, which are Poisson processes with parameters $\Lambda(\{x_m\})$.

To prove that $w_t, p_t^1, ..., p_t^n$ are mutually independent, introduce p^η as the jump measure of η_t and observe that by Theorem 3.14 the processes $\xi_t - \eta_t = bt + \sigma w_t$ and η_t (that is, w_t and η_t) are independent. It follows from Lemma 3.3 that, if we take any continuous functions $f_1, ..., f_n$ vanishing in the neighborhood of the origin, then the process w_t and the vector-valued process

$$\Big(\int_{\mathbb{R}} f_1(x)\, p_t^\eta(dx), ..., \int_{\mathbb{R}} f_n(x)\, p_t^\eta(dx) \Big)$$

are independent. By taking appropriate approximations we conclude that the process w_t and the vector-valued process

$$(p_t^\eta(a_1, b_1], ..., p_t^\eta(a_n, b_n])$$

are independent. Finally, by observing that, by Theorem 3.9, the processes $p_t^\eta(a_1, b_1], ..., p_t^\eta(a_n, b_n]$ are independent and, obviously (cf. (3.2)), $p^\eta = p$, we get that $w_t, p_t^1, ..., p_t^n$ are mutually independent. The theorem is proved.

The above proof is based on the formula

$$\xi_t = \zeta_t^a + \eta_t^a, \tag{2}$$

where

$$\eta_t^a = \int_{\mathbb{R}\setminus(-a,a)} x\, p_t(dx), \quad \zeta_t^a = \xi_t - \eta_t^a, \quad a > 0,$$

and the fact that for small a all processes η_t^a are the same. In the general case we want to let $a \downarrow 0$ in (2). The only trouble is that generally there is no limit of η_t^a as $a \downarrow 0$. On the other hand, the left-hand side of (2) does have a limit, just because it is independent of a. So there is a hope that if we subtract an appropriate quantity from ζ_t^a and add it to η_t^a, the results will converge. This appropriate quantity turns out to be the stochastic integral against *the centered Poisson measure* q introduced by

$$q_t(a, b] = p_t(a, b] - t\Lambda((a, b]) \quad \text{if} \quad ab > 0.$$

2. Lemma. *Let* $\Pi = \{(0,t] \times (a,b] : t > 0, a < b, ab > 0\}$ *and for* $A = (0,t] \times (a,b] \in \Pi$ *let* $q(A) = q_t(a,b]$. *Then* Π *is a* π*-system and* q *is a random orthogonal measure on* Π *with reference measure* $\ell \times \Lambda$.

Proof. Let $A = (0,t_1] \times (a_1,b_1], B = (0,t_2] \times (a_2,b_2] \in \Pi$. Then

$$AB = (0, t_1 \wedge t_2] \times (c,d], \quad (c,d] := (a_1,b_1] \cap (a_2,b_2],$$

which shows that Π is a π-system. That q is a random orthogonal measure on Π with reference measure $\ell \times \Lambda$ is stated in Corollary 3.13. The lemma is proved.

3. Remark. We may consider Π as a system of subsets of $\mathbb{R}_+ \times \mathbb{R} \setminus \{0\}$. Then as is easy to see, $\sigma(\Pi) = \mathfrak{B}(\mathbb{R}_+) \otimes \mathfrak{B}(\mathbb{R} \setminus \{0\})$. By Theorem 2.3.19, $L_2(\Pi, \Lambda) = L_2(\sigma(\Pi), \ell \times \Lambda)$. Therefore, Lemma 2 and Theorem 2.3.13 allow us to define the stochastic integral

$$\int_{\mathbb{R}_+ \times (\mathbb{R} \setminus \{0\})} f(t,x)\, q(dtdx)$$

for every Borel f satisfying

$$\int_{\mathbb{R}_+ \times \mathbb{R}} |f(t,x)|^2 \, dt\Lambda(dx) < \infty$$

(we write this integral over $\mathbb{R}_+ \times \mathbb{R}$ instead of $\mathbb{R}_+ \times (\mathbb{R} \setminus \{0\})$ because $\Lambda(\{0\}) = 0$ by definition). Furthermore,

$$\begin{aligned} E\Big|\int_{\mathbb{R}_+ \times (\mathbb{R} \setminus \{0\})} f(t,x)\, q(dtdx)\Big|^2 &= \int_{\mathbb{R}_+ \times \mathbb{R}} |f(t,x)|^2 \, dt\Lambda(dx), \\ E\int_{\mathbb{R}_+ \times (\mathbb{R} \setminus \{0\})} f(t,x)\, q(dtdx) &= 0, \end{aligned} \tag{3}$$

the latter following from the fact that $Eq(A) = 0$ if $A \in \Pi$ (see Remark 2.3.15).

4. Remark. Denote

$$\int_{\mathbb{R} \setminus \{0\}} f(x)\, q_t(dx) = \int_{\mathbb{R}_+ \times (\mathbb{R} \setminus \{0\})} I_{(0,t]}(u) f(x)\, q(dudx). \tag{4}$$

Then (3) shows that, for each Borel f satisfying $\int_{\mathbb{R}} |f(x)|^2 \, \Lambda(dx) < \infty$ and every $t, s \in [0, \infty)$,

$$E\Big|\int_{\mathbb{R}\setminus\{0\}} f(x)\,q_t(dx)-\int_{\mathbb{R}\setminus\{0\}} f(x)\,q_s(dx)\Big|^2=|t-s|\int_{\mathbb{R}}|f(x)|^2\,\Lambda(dx),$$

$$E\Big|\int_{\mathbb{R}\setminus\{0\}} f(x)\,q_t(dx)\Big|^2=t\int_{\mathbb{R}}|f(x)|^2\,\Lambda(dx).$$

In the following exercise we use for the first time our assumption that $(\Omega,\mathcal{F},P)$ is a complete probability space. This assumption allowed us to complete $\sigma(\xi_s : s\le t)$ and have this completion, denoted $\mathcal{F}^\xi_t$, to be part of $\mathcal{F}$. This assumption implies that, if we are given two random variables satisfying $\zeta=\eta$ (a.s) and ζ is $\mathcal{F}^\xi_t$-measurable, so is η.

5. Exercise*. Prove that if f is a bounded Borel function vanishing in a neighborhood of zero, then $\int_{\mathbb{R}}|f(x)|^2\,\Lambda(dx)<\infty$ and

$$\int_{\mathbb{R}\setminus\{0\}} f(x)\,q_t(dx)=\int_{\mathbb{R}} f(x)\,p_t(dx)-t\int_{\mathbb{R}} f(x)\,\Lambda(dx)\quad\text{(a.s.).}\qquad(5)$$

By using Lemma 3.8, conclude that the left-hand side of (5) is $\mathcal{F}^\xi_t$-measurable for every $f\in L_2(\mathfrak{B}(\mathbb{R}),\Lambda)$.

6. Exercise*. As a continuation of Exercise 5, prove that (5) holds for every Borel f satisfying $f(0)=0$ and $\int_{\mathbb{R}}(|f|+|f|^2)\,\Lambda(dx)<\infty$.

7. Lemma. *For every Borel $f\in L_2(\mathfrak{B}(\mathbb{R}),\Lambda)$ the stochastic integral*

$$\eta_t:=\int_{\mathbb{R}\setminus\{0\}} f(x)\,q_t(dx)$$

is an infinitely divisible $\mathcal{F}^\xi_t$-adapted process such that, if $0\le s\le t<\infty$, then $\eta_t-\eta_s$ and $\mathcal{F}^\xi_s$ are independent. By Theorem 1.11 the process η_t admits a modification with trajectories in $D[0,\infty)$. If we keep the same notation for the modification, then for every $T\in[0,\infty)$

$$E\sup_{t\le T}\eta_t^2\le 4T\int_{\mathbb{R}}|f(x)|^2\,\Lambda(dx).\qquad(6)$$

Proof. If f is a bounded continuous function vanishing in a neighborhood of zero, the first statement follows from Exercise 5 and Lemma 3.8. An obvious approximation argument and Remark 4 allow us to extend the result to arbitrary f in question.

To prove (6) take $0\le t_1\le ...\le t_n\le T$ and observe that, owing to the independence of $\eta_{t_{k+1}}-\eta_{t_k}$ and $\mathcal{F}^\xi_{t_k}$, we have

$$E(\eta_{t_{k+1}} - \eta_{t_k} | \mathcal{F}^{\xi}_{t_k}) = E(\eta_{t_{k+1}} - \eta_{t_k}) = 0.$$

Therefore, $(\eta_{t_k}, \mathcal{F}^{\xi}_{t_k})$ is a martingale. By Doob's inequality

$$E \sup_k \eta^2_{t_k} \le 4E\eta^2_T = 4T \int_{\mathbb{R}} |f(x)|^2 \,\Lambda(dx).$$

Clearly the inequality between the extreme terms has nothing to do with ordering t_k. Therefore by ordering the set ρ of all rationals on $[0,T]$ and taking the first n rationals as t_k, $k = 1, ..., n$, and then sending n to infinity, by Fatou's theorem we find that

$$E \sup_{r \in \rho, r < T} \eta^2_r \le 4T \int_{\mathbb{R}} |f(x)|^2 \,\Lambda(dx).$$

Now equation (6) immediately follows from the right continuity and the stochastic continuity (at point T) of $\eta_\cdot$, since (a.s.)

$$\sup_{t \le T} \eta^2_t = \sup_{t < T} \eta^2_t = \sup_{r \in \rho, r < T} \eta^2_r.$$

The lemma is proved.

8. Theorem. *Let ξ_t be an infinitely divisible cadlag process with parameters σ, b, and Lévy measure Λ.*

(i) *If $\sigma \neq 0$, then there exist a constant $\bar{b}$ and a Wiener process w_t, which is independent of all processes $p_t(c, d]$, such that, for each $t \ge 0$,*

$$\xi_t = \bar{b}t + \sigma w_t + \int_{(-1,1)} x\, q_t(dx) + \int_{\mathbb{R}\setminus(-1,1)} x\, p_t(dx) \quad (a.s.). \tag{7}$$

(ii) *If $\sigma = 0$, assertion (i) still holds if one does not mention w_t and drops the term σw_t in (7).*

Proof. For $a \in (0,1)$ write (2) as

$$\xi_t = \zeta^a_t + \int_{(-1,1)\setminus(-a,a)} x\, p_t(dx) + \int_{\mathbb{R}\setminus(-1,1)} x\, p_t(dx).$$

Here, by Exercise 5,

$$\int_{(-1,1)\setminus(-a,a)} x\, p_t(dx) = \int_{(-1,1)\setminus(-a,a)} x\, q_t(dx) + t \int_{(-1,1)\setminus(-a,a)} x\, \Lambda(dx),$$

so that

$$\xi_t = \kappa_t^a + \int_{(-1,1)\setminus(-a,a)} x\, q_t(dx) + \int_{\mathbb{R}\setminus(-1,1)} x\, p_t(dx), \tag{8}$$

where

$$\kappa_t^a = \zeta_t^a + t\int_{(-1,1)\setminus(-a,a)} x\,\Lambda(dx).$$

By Lemma 7, for any $T \in (0,\infty)$,

$$E\sup_{t\leq T}\Big|\int_{(-1,1)\setminus(-a,a)} x\, q_t(dx) - \int_{(-1,1)} x\, q_t(dx)\Big|^2 \to 0$$

as $a \downarrow 0$. Therefore, there exists a sequence $a_n \downarrow 0$, along which with probability one the first integral on the right in (8) converges uniformly on each finite time interval to the first integral on the right in (7). It follows from (8) that almost surely $\kappa_t^{a_n}$ also converges uniformly on each finite time interval to a process, say κ_t. Bearing in mind that the κ_t^a are cadlag and using Exercise 1.8, we see that κ_t is cadlag too. By Theorem 3.14, the process ζ_t^a is infinitely divisible cadlag. It follows that κ_t^a and κ_t are infinitely divisible cadlag as well.

Furthermore, since ζ_t^a does not have jumps larger in magnitude than a, the process κ_t does not have jumps at all and hence is continuous (the last conclusion is easily proved by contradiction). Again by Theorem 3.14, the process ζ_t^a is independent of η_t^a and, in particular, is independent of the jump measure of η_t^a (cf. Lemma 3.3). The latter being $p_t((c,d]\setminus(-a,a))$ (cf. (3.2)) shows that ζ_t^a as well as κ_t^a are independent of all processes $p_t((c,d]\setminus(-a,a))$. By letting $a \downarrow 0$, we conclude that κ_t is independent of all processes $p_t(c,d]$.

To conclude the proof it only remains to use Corollary 3.11. The theorem is proved.

9. Exercise. It may look as though assertion (i) of Theorem 8 holds even if $\sigma = 0$. Indeed, in this case $\sigma w_t \equiv 0$ anyway. However, generally this assertion is false if $\sigma = 0$. The reader is asked to give an example in which this happens.

6. Constructing infinitely divisible processes

Here we want to show that for an arbitrary Lévy measure and constants b and σ there exists an infinitely divisible process ξ_t, defined on an appropriate probability space, such that

$$Ee^{i\lambda\xi_t} = \exp t\{\int_{\mathbb{R}} (e^{i\lambda x} - 1 - i\lambda \sin x)\,\Lambda(dx) + ib\lambda - \sigma^2\lambda^2/2\}. \tag{1}$$

By the way, this will show that generally there are no *additional* properties of Λ apart from those listed in (2.7).

The idea is that if we have at least one process with "arbitrarily" small jumps, then by "redirecting" the jumps we can get jump measures corresponding to arbitrary infinitely divisible process. We know that at least one such "test" process exists, the increasing 1/2-stable process τ_{a+}, $a \geq 0$ (see Theorem 2.6.1 and Exercises 1.12).

The following lemma shows how to redirect the jumps of τ_{a+}.

1. Lemma. *Let Λ be a positive measure on $\mathfrak{B}(\mathbb{R})$ such that $\Lambda(\mathbb{R}\setminus(-a,a)) < \infty$ for any $a > 0$ and $\Lambda(\{0\}) = 0$. Then there exists a finite Borel function $f(x)$ on $\mathbb{R}$ such that $f(0) = 0$ and for any Borel Γ*

$$\Lambda(\Gamma) = \int_{f^{-1}(\Gamma)} |x|^{-3/2}\,dx.$$

Proof. For $x > 0$, define $2F(x) = \Lambda\{(x,\infty)\}$. Notice that $F(x)$ is right continuous on $(0,\infty)$ and $F(\infty) = 0$. For $x > 0$ let

$$f(x) = \inf\{y > 0 : 1 \geq xF^2(y)\}.$$

Since $F(\infty) = 0$, f is a finite function.

Next notice that, if $t > 0$ and $f(x) > t$, then for any $y > 0$ satisfying $1 \geq xF^2(y)$, we have $y > t$, which implies that $1 < xF^2(t)$. Hence,

$$\{x > 0 : f(x) > t\} \subset \{x > 0 : xF^2(t) > 1\}. \tag{2}$$

On the other hand, if $t > 0$ and $xF^2(t) > 1$, then due to the right continuity of F also $xF^2(t+\varepsilon) > 1$, where $\varepsilon > 0$. In that case, $f(x) \geq t + \varepsilon > t$. Thus the sets in (2) coincide if $t > 0$, and hence

$$\Lambda\{(t,\infty)\} = 2F(t) = \int_{1/F^2(t)}^{\infty} x^{-3/2}\,dx = \int_{x:xF^2(t)>1} x^{-3/2}\,dx = \nu\{(t,\infty)\},$$

where

$$\nu(\Gamma) = \int_{x>0:f(x)\in\Gamma} x^{-3/2}\,dx.$$

A standard measure-theoretic argument allows us to conclude that

$$\Lambda(\Gamma \cap (0,\infty)) = \nu(\Gamma)$$

not only for $\Gamma = (t,\infty)$, $t > 0$, but for all Borel $\Gamma \subset (0,\infty)$.

Similarly, one constructs a negative function $g(x)$ on $(-\infty, 0)$ such that

$$\Lambda(\Gamma \cap (-\infty,0)) = \int_{x<0:g(x)\in\Gamma} |x|^{-3/2}\,dx.$$

Finally, the function we need is given by $f(x)I_{x>0} + g(x)I_{x<0}$. The lemma is proved.

We also need the following version of Lemma 3.8.

2. Lemma. *Let p_t be the jump measure of an infinitely divisible cadlag process with Lévy measure Λ, and let f be a finite Borel function such that $f(0) = 0$ and $\Lambda(\{x : f(x) \neq 0\}) < \infty$. Then*

(i) *we have*

$$\int_{\mathbb{R}\setminus\{0\}} |f(x)|\,p_t(dx) < \infty$$

(a.s.), and

$$\xi_t := \int_{\mathbb{R}\setminus\{0\}} f(x)\,p_t(dx)$$

is well defined and is cadlag;

(ii) *ξ_t is an infinitely divisible process, and*

$$Ee^{i\xi_t} = \exp t\int_{\mathbb{R}} (e^{if(x)} - 1)\,\Lambda(dx). \tag{3}$$

Proof. (i) By Corollary 3.10

$$Ep_t(\{x : f(x) \neq 0\}) = t\Lambda(\{x : f(x) \neq 0\}) < \infty.$$

Since the measure p_t is integer valued, it follows that (a.s.) there are only finitely many points in $\{x : f(x) \neq 0\}$ to which p_t assigns a nonzero mass. This proves (i).

To prove (ii) we use approximations. The inequality $|e^{if} - 1| \leq 2I_{f\neq 0}$ and the dominated convergence theorem show that, if assertion (ii) holds for some functions $f_n(x)$ such that $f_n \xrightarrow{\Lambda} f$, $\Lambda(\{x : \sup_n |f_n(x)| > 0\}) < \infty$, and

$$\int_{\mathbb{R}\setminus\{0\}} |f - f_n|\, p_t(dx) \xrightarrow{P} 0, \tag{4}$$

then (ii) is also true for f. By taking $f_n = (-n) \vee f \wedge n$, we see that it suffices to prove (ii) for bounded f. Then considering $f_n = fI_{1/n<|x|<n}$ reduces the general case further to bounded functions vanishing for small and large $|x|$. Any such function can be approximated in $L_1(\mathfrak{B}(\mathbb{R}), \Lambda)$ by continuous functions f_n, for which (4) holds automatically due to Corollary 3.10 and (3) holds due to Lemma 3.8 (ii) with $\alpha = 0$. The lemma is proved.

Now let Λ be a Lévy measure and b and σ some constants. Take a probability space carrying two independent copies $\eta_t^{\pm}$ of the process τ_{t+}, $t \geq 0$, and a Wiener process w_t independent of (η_t^+, η_t^-). By Exercise 2.11, the Lévy measure of $\eta_t^{\pm}$ is given by $c_0 x^{-3/2} I_{x>0}\, dx$, where c_0 is a constant. Define

$$\Lambda_0(dx) = c_0 |x|^{-3/2}\, dx$$

and take the function f from Lemma 1 constructed from Λ/c_0 in place of Λ, so that, for any $\Gamma \in \mathfrak{B}(\mathbb{R})$,

$$\Lambda(\Gamma) = \Lambda_0 f^{-1}(\Gamma) = \Lambda_0(\{x : f(x) \in \Gamma\}). \tag{5}$$

3. Remark. Equation (5) means that, for any $\Gamma \in \mathfrak{B}(\mathbb{R})$ and $h = I_\Gamma$,

$$\int_{\mathbb{R}} h(x)\, \Lambda(dx) = \int_{\mathbb{R}} h(f(x))\, \Lambda_0(dx). \tag{6}$$

A standard measure-theoretic argument shows that (6) is true for each Borel nonnegative h and also for each Borel h for which at least one of

$$\int_{\mathbb{R}} |h(x)|\, \Lambda(dx) \quad \text{and} \quad \int_{\mathbb{R}} |h(f(x))|\, \Lambda_0(dx)$$

is finite. In particular, if h is a Borel function, then $h \in L_2(\mathfrak{B}(\mathbb{R}), \Lambda)$ if and only if $h(f) \in L_2(\mathfrak{B}(\mathbb{R}), \Lambda_0)$.

4. Theorem. *Let $p^{\pm}$ be the jump measures of $\eta_t^{\pm}$ and $q^{\pm}$ the centered Poisson measures of $\eta_t^{\pm}$. Define*

$$\xi_t^{\pm} = \int_{\mathbb{R}\setminus\{0\}} f(\pm x) I_{|f(\pm x)|<1}\, q_t^{\pm}(dx) + \int_{\mathbb{R}\setminus\{0\}} f(\pm x) I_{|f(\pm x)|\geq 1}\, p_t^{\pm}(dx)$$
$$=: \alpha_t^{\pm} + \beta_t^{\pm}.$$

Then, for a constant $\bar{b}$, the process

$$\xi_t = \bar{b}t + \sigma w_t + \xi_t^{+} + \xi_t^{-}$$

is an infinitely divisible process satisfying (1).

Proof. Observe that

$$\int_{\mathbb{R}} f^2(\pm x) I_{|f(\pm x)|<1}\, \Lambda_0(dx) = \int_{(-1,1)} x^2\, \Lambda(dx) < \infty.$$

Therefore, the processes $\alpha_t^{\pm}$ are well defined. In addition,

$$\Lambda_0(\{x > 0 : |f(\pm x)| \geq 1\}) \leq \Lambda_0(\{x : |f(x)| \geq 1\}) = \Lambda(|x| \geq 1) < \infty.$$

Hence, $\beta_t^{\pm}$ is well defined due to Lemma 2.

Next, in order to find the characteristic function of $\xi_t^{\pm}$, notice that $fI_{|f|<a} \to 0$ in $L_2(\mathfrak{B}(\mathbb{R}), \Lambda_0)$ as $a \downarrow 0$, so that upon remembering the properties of stochastic integrals, in particular, Exercise 5.6, we obtain

$$\alpha_t^{\pm} = \underset{a\downarrow 0}{\text{l.i.m.}} \Big(\int_{\mathbb{R}\setminus\{0\}} f(\pm x) I_{a\leq|f(\pm x)|<1}\, p_t^{\pm}(dx)$$
$$- t\int_0^{\infty} f(\pm x) I_{a\leq|f(\pm x)|<1}\, \Lambda_0(dx)\Big).$$

It follows that

$$\xi_t^{\pm} = P\text{-}\lim_{a\downarrow 0} \Big(\int_{\mathbb{R}\setminus\{0\}} f(\pm x) I_{a\leq|f(\pm x)|}\, p_t^{\pm}(dx)$$
$$- t\int_0^{\infty} f(\pm x) I_{a\leq|f(\pm x)|<1}\, \Lambda_0(dx)\Big).$$

Now Lemma 2 implies that $\xi_t^{\pm}$ are infinitely divisible and

$$Ee^{i\lambda\xi_t^{\pm}}$$
$$= \lim_{a\downarrow 0} \exp t\int_0^{\infty} \big\{(e^{i\lambda f(\pm x)} - 1) I_{a\leq|f(\pm x)|} - i\lambda f(\pm x) I_{a\leq|f(\pm x)|<1}\big\}\, \Lambda_0(dx).$$

In the next few lines we use the fact that $|e^{i\lambda x}-1-i\lambda xI_{|x|<1}|$ is less than λ^2x^2 if $|x|<1$ and less than 2 otherwise. Then, owing to Remark 3, we find that

$$g(\lambda,a):=\int_0^\infty\big\{(e^{i\lambda f(x)}-1)I_{a\leq|f(x)|}-i\lambda f(x)I_{a\leq|f(x)|<1}\big\}\,\Lambda_0(dx)$$

$$+\int_0^\infty\big\{(e^{i\lambda f(-x)}-1)I_{a\leq|f(-x)|}-i\lambda f(-x)I_{a\leq|f(-x)|<1}\big\}\,\Lambda_0(dx)$$

$$=\int_{\mathbb{R}}\big\{(e^{i\lambda f(x)}-1)I_{a\leq|f(x)|}-i\lambda f(x)I_{a\leq|f(x)|<1}\big\}\,\Lambda_0(dx)$$

$$=\int_{\mathbb{R}\setminus(-a,a)}\big\{e^{i\lambda x}-1-i\lambda xI_{|x|<1}\big\}\,\Lambda(dx).$$

This along with the dominated convergence theorem implies that

$$g(\lambda,a)\to\int_{\mathbb{R}}(e^{i\lambda x}-1-i\lambda xI_{|x|<1})\,\Lambda(dx)=\int_{\mathbb{R}}(e^{i\lambda x}-1-i\lambda\sin x)\,\Lambda(dx)+i\lambda\tilde{b},$$

where

$$\tilde{b}=\int_{\mathbb{R}}(\sin x-xI_{|x|<1})\,\Lambda(dx)$$

is a well-defined constant because $|\sin x-xI_{|x|<1}|\leq 2\wedge x^2$.

Finally, upon remembering that the processes w_t, p_t^+, p_t^- are independent, we conclude that ξ_t is infinitely divisible and

$$Ee^{i\lambda\xi_t}=\lim_{a\downarrow 0}\exp t\big(i\lambda\bar{b}-\sigma^2\lambda^2/2+g(\lambda,a)\big),$$

which equals the right-hand side of (1) if $\bar{b}+\tilde{b}=b$. The theorem is proved.

The theory in this chapter admits a very natural generalization for vector-valued infinitely divisible processes, which are defined in the same way as in Sec. 2. Also as above, having an infinitely divisible process with jumps of all sizes in all directions allows one to construct all other infinitely divisible processes. In connection with this we set the reader the following.

5. Exercise. Let $w_t,w_t^1,...,w_t^d$ be independent Wiener processes. Define $\tau_t=\inf\{s\geq 0:w_s\geq t\}$ and

$$\eta_t=(w_t^1,...,w_t^d),\quad \xi_t=\eta_{\tau_t}.$$

Prove that:

(i) The process ξ_t is infinitely divisible.

(ii) $E\exp(i\lambda\cdot\xi_t) = \exp(-ct|\lambda|)$ for any $\lambda \in \mathbb{R}^d$, where $c > 0$ is a constant, so that ξ_t has a multidimensional Cauchy distribution.

(iii) It follows from (ii) that the components of ξ_t are not independent. On the other hand, the components of η_t are independent random processes and we do a change of time in η_t, random but yet independent of η. Explain why this makes the components of $\xi_t = \eta_{\tau_t}$ depend on each other. What kind of nontrivial information about the trajectory of ξ_t^2 can one get if one knows the trajectory $\xi_t^1, t > 0$?

7. Hints to exercises

1.8 Assume the contrary.

1.12 For any cadlag modification $\tilde{\xi}_t$ of a process ξ_t we have $\xi_t \xrightarrow{P} \tilde{\xi}_s$ as $t \downarrow s$.

2.10 Use $\int_{\mathbb{R}}(\lambda\sin(x/\lambda) - \sin x)x^{-2}\,dx = 0$, which is true since $\sin x$ is an odd function.

2.11 To show that $a = b = 1$, observe that

$$\Psi(z) := \int_0^\infty x^{-3/2}(e^{-zx} - 1)\,dx$$

is an analytic function for $\operatorname{Re} z > 0$ which is continuous for $\operatorname{Re} z \geq 0$. Furthermore, for real z, changing variables, prove that $\Psi(z) = \sqrt{z}\Psi(1)$ and express $\Psi(1)$ through the gamma function by integrating by parts. Then notice that $\sqrt{2\pi}\Psi(i) = -a - ib$.

3.15 (ii) $P(\eta_t(a) = 0) = P(p[a,\infty) = 0)$. (iii) Use that $\xi_t - \eta_t(a)$ and $\eta_t(a)$ are independent and their sum is positive. (iv)&(v) Put $\alpha = 0$ in (3.15) to get the characteristic function of $\eta_t(0+)$ and also the fact that

$$\lim_{a\downarrow 0}\int_{[a,\infty)}(e^{i\lambda x} - 1)\,\Lambda(dx)$$

exists.

4.1 Corollary 3.10 says that the finite measures

$$\nu_{\varepsilon,T}(\Gamma) := Ep\{((0,T] \times (\mathbb{R} \setminus (-\varepsilon,\varepsilon))) \cap \Gamma\}$$

and

$$(\ell \times \Lambda)\{((0,T] \times (\mathbb{R} \setminus (-\varepsilon,\varepsilon))) \cap \Gamma\}$$

coincide on sets Γ of the form $(0,t] \times (a,b]$.

4.2 Assume $f \geq 0$, approximate f by the functions $f([tn]/n, x)$, and prove that

$$E\int_{(k/n,(k+1)/n]\times\mathbb{R}} f(k/n,x)\,p(dsdx)$$
$$= E\int_{(k/n,(k+1)/n]\times\mathbb{R}} (Ef(k/n,x))\,p(dsdx).$$

To do this step, use π- and λ-systems in order to show that it suffices to take $f(k/n,x)$ equal to $I_{A\times\Gamma}(\omega,x)$, where A and $p_{(k+1)/n}(\Gamma)-p_{k/n}(\Gamma)$ are independent.

4.3 First let there be an integer n such that $f(t,x)=f((k+1)2^{-n},x)$ whenever k is an integer and $t\in(k2^{-n},(k+1)2^{-n}]$, and let $f((k+1)2^{-n},x)$ be continuous in x. In that case use Lemma 3.8. Then derive the result for any continuous function $f(t,x)$ vanishing for $|x|<\varepsilon$. Finally, pass to the limit from continuous functions to arbitrary ones by using (4.1).

4.5 Take some constants α and β and define $\zeta_t=\alpha\xi_t+\beta\eta_t$, $\varphi(t)=Ee^{i\zeta_t}$. Notice that

$$e^{i\zeta_t}=1+\int_{(0,t]} e^{i\zeta_{t-}}\{[e^{i\alpha}-1]\,d\xi_t+[e^{i\beta}-1]\,d\eta_t\},$$

where on the right we just have a telescoping sum. By taking expectations derive that

$$\varphi(t)=1+\int_0^t \varphi(s)\{[e^{i\alpha}-1]\mu+[e^{i\beta}-1]\nu\}\,ds.$$

This will prove the independence of ξ_t and η_t for any t. To prove the independence of the processes, repeat part of the proof of Lemma 3.7.

5.5 First check (5.5) for $f=I_{(a,b]}$ with $ab>0$, and then use Corollary 3.10, the equality $L_2(\Pi,\Lambda)=L_2(\sigma(\Pi),\Lambda)$, and (2.7), which shows that $\Lambda(\mathbb{R}\setminus(-a,a))<\infty$ for any $a>0$.

5.6 The functions $(n\wedge f)I_{|x|>1/n}$ converge to f in $L_1(\mathfrak{B}(\mathbb{R}),\Lambda)$ and in $L_2(\mathfrak{B}(\mathbb{R}),\Lambda)$.

6.5 (i) Use Theorem 2.6.1. (ii) Add that

$$E\exp(i\lambda\cdot\xi_t)=\int_0^\infty E\exp(i\lambda\cdot\eta_s)\,P(\tau_t\in ds).$$

(iii) Think of jumps.

Chapter 6

Itô Stochastic Integral

The reader may have noticed that stochastic integrals or stochastic integral equations appear in every chapter in this book. Here we present a systematic study of the Itô stochastic integral against the Wiener process. This integral has already been introduced in Sec. 2.7 by using an approach which is equally good for defining stochastic integrals against martingales. This approach also exhibits the importance of the σ-field of predictable sets. Traditionally the Itô stochastic integral against dw_t is introduced in a different way, with discussion of which we start the chapter.

1. The classical definition

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, $\mathcal{F}_t$, $t \geq 0$, an increasing filtration of σ-fields $\mathcal{F}_t \subset \mathcal{F}$, and $w_t, t \geq 0$, a Wiener process relative to $\mathcal{F}_t$.

1. Definition. Let $f_t = f_t(\omega)$ be a function defined on $\Omega \times [0, \infty)$. We write $f \in H_0$ if there exist nonrandom points $0 = t_0 \leq t_1 \leq ... \leq t_n < \infty$ such that the f_{t_i} are $\mathcal{F}_{t_i}$-measurable, $Ef_{t_i}^2 < \infty$, and $f_t = f_{t_i}$ for $t \in [t_i, t_{i+1})$ if $i \leq n$, whereas $f_t = 0$ for $t \geq t_n$.

2. Exercise. Why does it not make much sense to consider functions satisfying $f_t = f_{t_i}$ for $t \in (t_i, t_{i+1}]$?

For $f \in H_0$ we set

$$If = \sum_{i=0}^{n-1} (w_{t_{i+1}} - w_{t_i}) f_{t_i}.$$

Obviously this definition is independent of the partition $\{t_i\}$ of $[0,\infty)$ provided that $f \in H_0$. In particular, the notation If makes sense, and I is a linear operator on H_0.

3. Lemma. *If $f \in H_0$, then*

$$E(If)^2 = E\int_0^\infty f_t^2\,dt, \quad EIf = 0.$$

Proof. We have (see Theorem 3.1.12)

$$Ef_{t_j}^2(w_{t_{j+1}} - w_{t_j})^2 = Ef_{t_j}^2 E\{(w_{t_{j+1}} - w_{t_j})^2|\mathcal{F}_{t_j}\} = Ef_{t_j}^2(t_{j+1} - t_j),$$

since $w_{t_{j+1}} - w_{t_j}$ is independent of $\mathcal{F}_{t_j}$ and f_{t_j} is $\mathcal{F}_{t_j}$-measurable. This and Cauchy's inequality imply that the first expression in the following relations makes sense:

$$Ef_{t_i}(w_{t_{i+1}} - w_{t_i})f_{t_j}(w_{t_{j+1}} - w_{t_j})$$
$$= Ef_{t_i}(w_{t_{i+1}} - w_{t_i})f_{t_j}E\{(w_{t_{j+1}} - w_{t_j})|\mathcal{F}_{t_j}\} = 0$$

if $i < j$, since $t_{i+1} \le t_j$ and $f_{t_j}, w_{t_{i+1}} - w_{t_i}, f_{t_i}$ are $\mathcal{F}_{t_j}$-measurable, whereas $w_{t_{j+1}} - w_{t_j}$ is independent of $\mathcal{F}_{t_j}$. Hence

$$E(If)^2 = \sum_{j=0}^{n-1} Ef_{t_j}^2(w_{t_{j+1}} - w_{t_j})^2 + 2\sum_{i<j\le n-1} Ef_{t_i}(w_{t_{i+1}} - w_{t_i})f_{t_j}(w_{t_{j+1}} - w_{t_j})$$

$$= \sum_{j=0}^{n-1} Ef_{t_j}^2(t_{j+1} - t_j) = E\int_0^\infty f_t^2\,dt.$$

Similarly, $Ef_{t_j}(w_{t_{j+1}} - w_{t_j}) = 0$ and $EIf = 0$. The lemma is proved.

The next step was not done in Secs. 2.7 and 2.8 because we did not have the necessary tools at that time. In the following lemma we use the notion of continuous time martingale, which is introduced in the same way as in Definition 3.2.1, just allowing m and n to be arbitrary numbers satisfying $0 \le n \le m$.

4. Lemma. *For $f \in H_0$, define $I_s f = I(I_{[0,s)}f)$. Then $(I_s f, \mathcal{F}_s)$ is a martingale for $s \ge 0$.*

Proof. Fix s and without loss of generality assume that $s \in \{t_0, ..., t_n\}$. If $s = t_k$, then

$$I_{[0,s)}f_t = \sum_{i=0}^{k-1} f_{t_i} I_{[t_i,t_{i+1})}(t), \quad I_s f = \sum_{i=0}^{k-1} f_{t_i}(w_{t_{i+1}} - w_{t_i}).$$