(Lesson 4 of 21/10/19) Compiled: December 4, 2019

0.1 Wiener's integral

Consider an *unconstrained* Brownian particle, moving on the real line, starting in x_0 at t_0. By solving the diffusion equation we found that the probability of finding the particle in $[x, x + \mathrm{d}x]$ at time $t > t_0$ is given by the **propagator**:

$$\mathbb{P}\{x(t) \in [x, x + \mathrm{d}x] | x(t_0) = x_0\} = W(x, t|x_0, t_0)\, \mathrm{d}x = \\ = \frac{1}{\sqrt{4\pi D(t - t_0)}} \exp\left(-\frac{(x - x_0)^2}{4D(t - t_0)}\right) \mathrm{d}x \tag{1}$$

By integrating (1) we can then find the probability of finding the particle inside an interval $[A, B]$ at time t:

$$\mathbb{P}\{x(t) \in [A, B] | x(t_0) = x_0\} = \int_A^B \mathrm{d}x\, W(x, t|x_0, t_0) \qquad t > t_0$$

We are now interested in computing the expected value $\langle f \rangle$ of **functionals** f of the trajectory, i.e. of quantities depending on several (or all) points of the trajectory $x(\tau)$ of a Brownian particle.

- The simplest example is the **correlation function**, which is defined as the product of the particle's position at two different times $t_1 < t_2$:

$$f(\{x(t_1), x(t_2)\}) = x(t_1)x(t_2) \quad t_1 < t_2$$

- A more general (and difficult) case is given by a function of the *entire* trajectory, such as:

$$f(\{x(\tau) \colon 0 < \tau \leq \tau\}) = g\left(\int_0^t x(\tau)a(\tau)\, \mathrm{d}\tau\right) \qquad a,\ g \colon \mathbb{R} \to \mathbb{R}$$

0.1.1 Functions of a discrete number of points

Let's start from the simplest case, and consider the **correlation function**:

$$f(\{x(t_1), x(t_2)\}) = x(t_1)x(t_2) \qquad t_1 < t_2$$

To compute $\langle f \rangle$ we will need the *joint probability distribution* $g(x_1, x_2)$ that gives the probability of $x(t_1)$ being "close to" x_1 **and** $x(t_2)$ "close to" x_2 *for the same trajectory.* Let us denote the three events of interest:

$$\begin{aligned} &A \colon \text{Particle starts in } x_0 \text{ at } t_0 \\ &B \colon \text{Particle is close to } x_1 \text{ at } t_1 \ (x(t_1) \in [x_1, x_1 + \mathrm{d}x_1]) \\ &C \colon \text{Particle is close to } x_2 \text{ at } t_2 \ (x(t_2) \in [x_2, x_2 + \mathrm{d}x_2]) \end{aligned}$$

We are interested in the joint probability $\mathbb{P}(C, B|A)$ (the order is defined by $t_2 > t_1 > t_0$). From probability theory:

$$\mathbb{P}(C, B|A) = \mathbb{P}(C|B, A)\mathbb{P}(B|A)$$

We already now how to compute probabilities like $\mathbb{P}(B|A)$, but not like $\mathbb{P}(C|B, A)$. Fortunately, that is not needed.
Recall, in fact, that Brownian motion is a *Markovian process*, meaning that the *future* depends only on the *present state*, i.e. the particle *has no memory*. So, **subsequent displacements are independent**: the probability of the particle going from x_1 to x_2 is the same whether it has started at x_0 or at any other point $\tilde{x}_0$. In other words, if we take the *present state* as the particle being in x_1 at t_1, the future (position at $t_2 > t_1$) depends only on that, and not on the past (position at t_0). So:

$$\mathbb{P}(C|B, A) = \mathbb{P}(C|B)$$

leading to:

$$\mathbb{P}(C, B|A) = \mathbb{P}(C|B)\mathbb{P}(B|A)$$

Inserting the *propagators* (1):

$$\mathrm{d}\mathbb{P}_{t_1,t_2}(x_1, x_2) \equiv W(x_2, t_2|x_1, t_1)W(x_1, t_1|x_0, t_0)\,\mathrm{d}x_1\,\mathrm{d}x_2$$

This is the joint probability we need to compute $\langle f \rangle$. Of course, nothing stops us from considering N "jumps" instead of only 2:

$$\begin{aligned}\mathrm{d}\mathbb{P}_{t_1,\dots,t_n}(x_1, \dots, x_n|x_0, t_0) &\equiv W(x_n, t_n|x_{n-1}, t_{n-1}) \dots W(x_1, t_1|x_0, t_0)\,\mathrm{d}x_1\,\mathrm{d}x_2 \dots \mathrm{d}x_n = \\ &= \exp\left(-\sum_{i=1}^n \frac{(x_i - x_{i-1})^2}{4\pi D \Delta t_i}\right) \prod_{i=1}^n \frac{\mathrm{d}x_i}{\sqrt{4\pi D\Delta t_i}}\end{aligned} \tag{2}$$

Then, the average of a generic function $f(x(t_1), \dots, x(t_n))$ of the positions of the particle at times $t_1 < t_2 < \dots < t_n$ is defined as:

$$\langle f(x(t_1), \dots, x(t_n))\rangle_w = \int_{\mathbb{R}^n} f(x_1, \dots, x_n)\,\mathrm{d}\mathbb{P}_{t_1,\dots,t_n}(x_1, \dots, x_n|x_0, t_0)$$

0.1.2 Functionals of the whole trajectory

The quantity in (2) can be interpreted as the *infinitesimal volume element* spanned by all the *trajectories* passing through a set of *tiny gates*, as represented in figure 1.
The underlying idea is that *probabilities* satisfy the axioms of *measures*, that is *functions that assign a measure* , i.e. a generalization of "size", *to all sets included in a specific collection.*

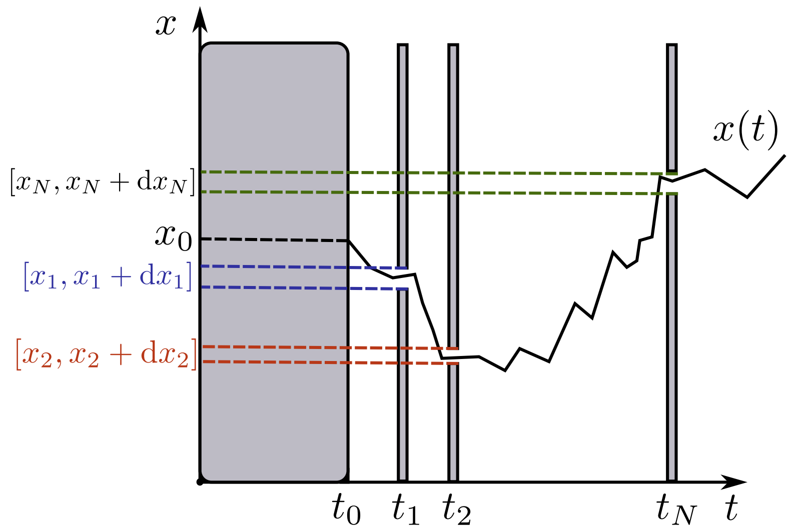

Figure (1) – All trajectories that pass through the set of *gates* $[x_i, x_i + \mathrm{d}x_i]$ at times t_i (such as the $x(t)$ here represented) contribute to the *volume* $\mathrm{d}\mathbb{P}_{t_1,\dots,t_n}(x_1, \dots, x_N)$

.

We now try to formalize this idea in order to extend the results of the previous section to the case of functions depending on a *infinite* number of trajectory points.

1. **Space definition**. Let $T \subset \mathbb{R}$ (**index set**), denote with $\mathbb{R}^T$ the set of all functions (**stochastic processes**) $k\colon T \to \mathbb{R}$. The idea is that an element of $\mathbb{R}^T$ is a collection of *random variables* indexed by T.

   In our case T is a collection of time instants (e.g. $T = [0, +\infty)$) and a generic element of $\mathbb{R}^T$ is made of *all the traversed points of a trajectory at times* T:

   $$\{x(t)\colon t \in T\} \in \mathbb{R}^T$$

2. **Probability measure on finite points**. The expression in (2), as observed, allows us to *measure* the volumes spanned by trajectories traversing a set of gates. Let's formalize this idea. Consider a *finite* set of times $T = \{t_i\}_{i=1,\dots,n}$ with $n \in \mathbb{N}, t_i \in \mathbb{R}$ and $t_1 < t_2 < \dots t_n$, each associated to a *gate* $H_i = [a_i, b_i]$, with $a_i, b_i \in \mathbb{R}$ and $a_i < b_i$. All the trajectories $\mathbb{R}^T$ traversing each H_i at a time $t_i \in T$ span a **cylindrical set** A of the form:

   $$A = \{x(t)\colon x(t_1) \in H_1, \dots, x(t_n) \in H_n\} \subset \mathbb{R}^T$$

   Using (2) and integrating over the *gates* we can define the measure of A-like sets as:

   $$P_w(A) \equiv \int_{\mathbb{R}^n} \mathrm{d}\mathbb{P}_{t_1,\dots,t_n}(x_1, \dots, x_n | x_0, t_0) \mathbb{I}_{H_1}(x_1) \dots \mathbb{I}_{H_n}(x_n)$$

   where $\mathbb{I}_{H_i}(x_i)$ are **characteristic functions** of the gates:

   $$\mathbb{I}_{H_i}(x) = \begin{cases} 1 & x \in H_i \\ 0 & \text{otherwise} \end{cases}$$

In our case H_i are just intervals, and so:

$$P_w(A) \equiv \mathbb{P}_{t_1,\dots,t_n}(A) \underset{(2)}{=} \int_{H_1} \mathrm{d}x_1 \int_{H_2} \mathrm{d}x_2 \cdots \int_{H_n} \mathrm{d}x_n \prod_{i=1}^n \frac{1}{\sqrt{4\pi D \Delta t_i}} \exp\left(-\frac{(x_i - x_{i-1})^2}{4D\Delta t_i}\right) \tag{3}$$

with $\Delta t_i = t_i - t_{i-1}$.

3. **Generalization on infinite points**. Note that (3) holds for any n. So, using A-like sets, we can construct a σ-algebra[1] $\mathcal{F}$ of $\mathbb{R}^T$. Then, by applying **Kolmogorov extension theorem** we can extend the measure P_w we just found to the entire $\mathcal{F}$.

4. **Probability space**. We now have a set of *all possible outcomes* $\mathbb{R}^T$ (in our case, all the possible trajectories that can be produced by a Brownian motion). We also have the collection of all *events* $\mathcal{F}$, that is subsets of $\mathbb{R}^T$ for which is meaningful to assign a *probability measure* $P_w\colon \mathcal{F} \to [0,1]$. The triad $(\mathbb{R}^T, \mathcal{F}, P_w)$ forms a **probability space**, that gives a rigorous meaning to the concept of "computing the probability of a trajectory".

The **measure** so obtained is called **Wiener measure**, and denoted as the following:

$$P_w(A) \equiv \int_A \mathrm{d}_w\, x(\tau)$$

Then we can compute expected values. For example, if $f(\{x(\tau)\colon \tau \in T\})$ is a function depending on the points traversed at times in a set T, then:

$$\langle f \rangle_w \equiv \int_{\mathbb{R}^T} f(x(\tau))\, \mathrm{d}_w\, x(\tau) \qquad T = [0, \infty)$$

Note that the Weiner measure *exists* and it's well defined (Kolmogorov's theorem), but we know it explicitly only in specific *finite cases*. So, to compute the expected value of functionals $F(\{x(t)\})$ over *continuous trajectories* we first *discretize* the trajectory, and then take a *continuum limit*.

1. Suppose we have a functional $F(\{x(\tau)\colon 0 < \tau < t\})$, and we want to compute $\langle F \rangle$.

2. We *discretize* the problem by *arbitrarily* subdividing the time interval $[0,t]$ in n parts $0 = t_0 < t_1 < t_2 < \cdots < t_n = t$. Then we consider an *approximated* functional $F_N(\{x(t_0), \dots, x(t_n)\})$ (for example approximating the path $x(\tau)$ with a piecewise linear function, depending only on $x(t_0), \dots, x(t_n)$), so that:

$$F = \text{``} \lim_{N\to\infty} \text{''} F_N$$

where $N \to \infty$ means that $\max \Delta t_i \to 0$, with $\Delta t_i = t_i - t_{i-1}$. This limit needs to be properly defined (by using the Weiner measure to define a norm in a space of *integrable* functionals, etc.), but we will not do that here.

[1] ∧A σ-algebra on a set X is a collection Σ of subsets of X that includes X itself, is closed under complement, and is closed under countable unions

3. Then the **Weiner path integral** is defined as:

$$\langle F\rangle_w = \int \mathrm{d}_w\, x(\tau) F(\{x(\tau)\colon 0<\tau<t\}) \equiv \text{“}\lim_{N\to\infty}\text{”} \langle F_N\rangle_w =$$
$$= \text{“}\lim_{N\to\infty}\text{”} \int \mathrm{d}\mathbb{P}_{t_1,\dots,t_n}(x_1,\dots,x_n|x_0,t_0) F_N(x(t_0),\dots,x(t_n))$$

Geometrically, we are evaluating F for every possible Brownian path $x(\tau)$, and then averaging all these results, each weighted by the *probability* of the corresponding path.

Example 1 (Correlation function and ESCK property)**:**

As expected, the more general definition of the Weiner measure - involving the continuum limit $N\to\infty$ - reduces to (2) when evaluated for a function depending only on a finite set of particle's positions.
For example, consider the expected value of the *correlation function* (assume the particle starting in 0 at time 0 for simplicity):

$$\langle x(t_1')x(t_2')\rangle = \int \mathrm{d}_w x\, x_1(t_1')x_2(t_2') =$$
$$= \text{“}\lim_{N\to\infty}\text{”} \int \mathrm{d}\mathbb{P}_{t_1,\dots,t_N}(x_1,\dots,x_N|0,0) x(t_k)x(t_n) =$$

where we chose the discretization so that $t_k = t_1'$ and $t_n = t_2'$. Then, by expanding the measure and applying the ESCK property we get (the limit is omitted):

$$= \int \mathrm{d}x_1 \dots \mathrm{d}x_N\, W(x_N,t_N|x_{N-1},t_{N-1})\dots W(x_1,t_1|0,0)x_k x_n =$$
$$\underset{(a)}{=} \int \mathrm{d}x_k\, \mathrm{d}x_n\, W(x_n,t_n|x_k,t_k)W(x_k,t_k|0,0)x_k x_n$$

which is the same result we could have obtained directly from (2):

$$\langle x(t_1')x(t_2')\rangle = \int \mathrm{d}x_1'\, \mathrm{d}x_2'\, W(x_2',t_2'|x_1',t_1')W(x_1',t_1'|0,0)x_1'x_2'$$

0.2 Change of random variables

Consider a random variable $X \sim q(x)$, with $q(x)$ being a generic distribution (e.g. $q(x) = \mu e^{-\mu x}$). Now consider a function $y(x)$, e.g. $y(x) = x^2$. Y is then a new random variable, with a certain distribution $p(y)$. We now want to compute $p(y)$ starting from $q(x)$ and $y(x)$.
Suppose that $y(x)$ is invertible. Then, if we extract a value from X, it will be inside $[x, x+\mathrm{d}x]$ with a probability $q(x)\,\mathrm{d}x$. Knowing X, we can use the relation $y(x)$ to uniquely determine Y, that will be in $[y, y+\mathrm{d}y]$ with the same probability. So, the following holds:

$$q(x)\,\mathrm{d}x = p(y)\,\mathrm{d}y \tag{4}$$

We can compute $\mathrm{d}y$ by *nudging* $y(x)$, and expanding in Taylor series:

$$y(x + \mathrm{d}x) \equiv y + \mathrm{d}y + O(\mathrm{d}y^2) = y(x) + \underbrace{\mathrm{d}x\, y'(x)}_{\mathrm{d}y} + O(\mathrm{d}x^2)$$

and so $\mathrm{d}y = \mathrm{d}x\, y'(x)$. Substituting in (4) we get:

$$q(x)\,\mathrm{d}x = p(y)\,\mathrm{d}y = p(y(x))y'(x)\,\mathrm{d}x \Rightarrow p(y) = q(x(y))\frac{\mathrm{d}x}{\mathrm{d}y}(x(y)) \tag{5}$$

For a more general $y(x)$, $x \sim q(x)$ and $y = y(x)$, the expected value of a function f in terms of $q(x)$:

$$\begin{aligned}\langle f(y)\rangle &= \int_{\mathbb{R}} \mathrm{d}x\, f(y(x))q(x) = \\ &= \int_{\mathbb{R}} \mathrm{d}x\, f(y(x))q(x) \underbrace{\int_{\mathbb{R}} \mathrm{d}z\, \delta(z - y(x))}_{=1} = \\ &\underset{(a)}{=} \int_{\mathbb{R}} \mathrm{d}z\, f(z) \underbrace{\int_{\mathbb{R}} \mathrm{d}x\, q(x)\delta(z - y(x))}_{\langle \delta(z-y(x))\rangle_{q(x)}}\end{aligned} \tag{6}$$

where in (a) we used the fact that $\delta(z - y(x)) = 1$ only when $z = y(x)$, and it's 0 otherwise, and so:

$$f(y(x)) = \int_{\mathbb{R}} \mathrm{d}z\, f(z)\delta(z - y(x))$$

Of course we can rewrite $\langle f\rangle$ directly in terms of $p(y)$:

$$\langle f(y)\rangle = \int_{\mathbb{R}} \mathrm{d}y\, f(y)p(y) \tag{7}$$

Comparing (8) with (7) and renaming $y \to z$ leads to:

$$p(z) = \int_{\mathbb{R}} \mathrm{d}x\, q(x)\delta(z - y(x)) = \langle \delta(z - y(x))\rangle_{q(x)} \tag{8}$$

which, in general, is not the same as the previously obtained result:

$$p(z) \neq q(x(z))\frac{\mathrm{d}x(z)}{\mathrm{d}z}$$

To retrieve this special case we must assume $y(x)$ to be invertible, with inverse $x(y)$. This means that $\operatorname{sgn} y'(x) = A$, with $A \in \mathbb{R} \setminus \{0\}$ constant.
We want now to compute $\delta(z - y(x))$ in this case. Recall that $\delta \circ g$, if g is a continuously differentiable function with $g(x_0) = 0$ and $g'(x) \neq 0\, \forall x$ is:

$$\delta(g(x)) = \frac{\delta(x - x_0)}{|g'(x_0)|}$$

So, if we let $g(x) = z - y(x)$, the only zero is at $x = x(z)$, as then $y(x(z)) = z$. So:

$$\delta(z - y(x)) = \frac{\delta(x - x(z))}{|y'(x(z))|}$$

Substituting back in (8):

$$\begin{aligned} p(z) &= \langle \frac{\delta(x - x(z))}{|y'(x(z))|} \rangle_{q(x)} = \int_{\mathbb{R}} \mathrm{d}x\, q(x) \frac{\delta(x - x(z))}{|y'(x(z))|} = q(x(z)) |y'(x(z))|^{-1} = \\ &= q(x(z)) \frac{\mathrm{d}x(y)}{\mathrm{d}y} \Big|_{y=x(z)} \end{aligned} \tag{9}$$

which is the same rule found in (5).

0.3 The 1st integral

Thanks to the Wiener measure we have a way to assign probabilities to *paths* $x(\tau)$. We can recover from this the *transition probabilities* we started from, by considering the functional that *evaluates* a path at an instant t: $x(\tau) \mapsto x(t) \equiv x_t$. Then, by applying (8) we can compute the distribution followed by x_t:

$$p(x_t) = W(x_t, t|0,0) = \langle \delta(x_t - x(\tau)) \rangle_w = \int \mathrm{d}_W x(\tau)\, \delta(x_t - x(\tau)) \tag{10}$$

(The starting condition $x(t_0) = x_0$ is *contained* in the definition of the measure $\mathrm{d}_W\, x(\tau)$).
So we can now write:

$$\begin{aligned} W(x,t|0,0) &= \int \mathrm{d}_W x\, \delta(x(t) - x) = \\ &= \text{“} \lim_{N\to\infty} \text{”} \int \prod_{i=1}^{N+1} \frac{\mathrm{d}x_i}{\sqrt{4\pi D \Delta t_i}} \exp\left(-\sum_{i=1}^{N} \frac{(x_i - x_{i-1})^2}{4D\Delta t_i} \right) \delta(x_{N+1} - x) \end{aligned}$$

where $t_n = t$, $x(t_n) = x_{N+1}$.

Recall that:

$$W(x_t, t|0,0) = \frac{1}{\sqrt{4\pi D t}} \exp\left(-\frac{x_t^2}{4Dt} \right)$$

If we set $x_t = 0$ (for simplicity), we get:

$$W(0,t|0,0) = \frac{1}{\sqrt{4\pi D t}} \tag{11}$$

As an exercise to get some familiarity with Wiener integrals, we want now to *re-derive* this result, by evaluating the Weiner path integral in (10), with $x_t = 0$:

$$W(0,t|0,0) = \langle \delta(0 - x(\tau)) \rangle_w = \int \mathrm{d}_W x(\tau)\, \delta(x(\tau)) \equiv I_1 \tag{12}$$

First, it is convenient to establish some additional notation. Let $T = [0, \infty)$. We denote with $\mathcal{C}\{0,0;t'\}$ the subset of trajectories in $\mathbb{R}^T$ starting from $x = 0$ at $t = 0$, and lasting a time span t'. Then, $\mathcal{C}\{0,0;x',t'\}$ is the subset of $\mathcal{C}\{0,0;t'\}$ when even the end-point is fixed to be $x(t') = x'$. The following *normalization* property holds:

$$\langle 1 \rangle_w = \int_{\mathcal{C}\{0,0;t\}} 1 \cdot \mathrm{d}_w x(\tau) = 1$$

We can then rewrite (12) as:

$$I_1 = \int_{\mathcal{C}\{0,0;0,t\}} \mathrm{d}_W x(\tau)$$

Geometrically, $W(0,t|0,0) = I_1$ is the probability that a Brownian particle starting at the origin *returns* in it after a finite amount of time t.

The standard way to compute a Wiener integral is to *discretize* it, and then take a *continuum* limit. So, consider for simplicity a **uniform** time discretization $\{t_i\}_{i=1,\dots,N+1}$, with instants ϵ-apart from each other, so that:

$$t_i - t_{i-1} \equiv \epsilon = \frac{t}{N+1} \quad \forall i = 1, \dots, N+1$$

Note that the end-points are $x_0 = x_{N+1} = 0$.
Then:

$$I_1 \equiv \lim_{\substack{\epsilon \to 0 \\ N \to \infty}} I_1^{(N)} \tag{13}$$

$$I_1^{(N)} \equiv \frac{1}{(\sqrt{4\pi D\epsilon})^{N+1}} \int_{-\infty}^{+\infty} \mathrm{d}x_1 \int_{-\infty}^{+\infty} \mathrm{d}x_2 \cdots \int_{-\infty}^{+\infty} \mathrm{d}x_N \exp\left(-\frac{1}{4D\epsilon} \sum_{i=0}^{N} (x_{i+1} - x_i)^2\right) \tag{14}$$

Let's focus on the summation in the exponential:

$$\begin{aligned}
\sum_{i=0}^{N} (x_{i+1} - x_i)^2 &= x_1^2 + \cancel{x_0}^2 - 2x_0x_1 + x_2^2 + x_1^2 - 2x_1x_2 + \cdots + \cancel{x_{N+1}}^2 + x_N^2 - 2x_Nx_{N+1} = \\
&= 2(x_1^2 + \cdots + x_N^2) - 2(x_1x_2 + x_2x_3 + \cdots + x_{N-1}x_N) = \\
&= 2\left(\sum_{i=1}^{N} x_i^2\right) - 2\left(\sum_{i=1}^{N-1} x_i x_{i+1}\right)
\end{aligned}$$

This is a *quadratic form*, i.e. a polynomial with all terms of order 2. So, it can be written in *matrix form*:

$$= \sum_{k,l=1}^{N} x_k A_{kl} x_l = \boldsymbol{x}^T A_N \boldsymbol{x}$$

for an appropriate choice of entries A_{kl} of the $N \times N$ matrix A_N:

$$A_{kk} = 2;\ A_{kl} = -(\delta k, l + 1 + \delta_{k+1,l}) \Rightarrow A_N = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix}$$

Substituting back in (14):

$$I_1^{(N)} = \frac{1}{(\sqrt{4\pi D\epsilon})^{N+1}} \int_{-\infty}^{+\infty} \mathrm{d}x_1 \cdots \int_{-\infty}^{+\infty} \mathrm{d}x_N \exp\left(-\frac{\boldsymbol{x}^T A_N \boldsymbol{x}}{4D\epsilon}\right)$$

Recall the multivariate Gaussian integral:

$$\int_{-\infty}^{+\infty} \mathrm{d}x_1 \cdots \mathrm{d}x_N \exp\left(-\sum_{ij}^{N} B_{ij} x_i x_j\right) = \frac{(\sqrt{\pi})^N}{\sqrt{\det B}}$$

leading to:

$$I_1^{(N)} = \frac{1}{(\sqrt{4\pi D\epsilon})^{N+1}} \sqrt{\frac{\pi^N}{\det\left(A_N \left[\frac{1}{4D\epsilon}\right]^N\right)}} \underset{(a)}{=} \frac{1}{(\sqrt{4\pi D\epsilon})^{N+1}} \frac{\sqrt{4\pi D\epsilon}^N}{\sqrt{\det A_N}} = \frac{1}{\sqrt{4\pi D\epsilon}} \frac{1}{\sqrt{\det A_N}} \tag{15}$$

where in (a) we used the property of the determinant $\det(cA) = c^n \det(A)\, \forall c \in \mathbb{R}$. Now, all that's left is to compute the determinant of A_N. Fortunately, as A_N is a *tri-diagonal* matrix, there is a recurrence relation in terms of the leading principal minors of A_N, which turns out to be multiples of the determinants of A_{N-1} and A_{N-2}.
Explicitly, consider A_N:

$$\det A_N \equiv \begin{vmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 & {\color{purple}-1} \\ 0 & \cdots & 0 & -1 & {\color{red}2} \end{vmatrix}_{N\times N}$$

and start computing the determinant following the last column. The only non-zero contributions are:

$$\det A_N = \underbrace{(-1)^{(N-1)+N}(-1)}_{+1} \begin{vmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 & -1 \\ 0 & \cdots & 0 & 0 & -1 \end{vmatrix}_{(N-1)\times(N-1)} + (-1)^{2N}(2)\det A_{N-1} =$$

$$= (-1)^{2(N-1)}(-1)\det A_{N-2} + 2\det A_{N-1} = 2\det A_{N-1} - \det A_{N-2} \qquad (16)$$

where the terms in blue are just the alternating signs from the determinant expansion, and the other colours identify the matrix entries that are being used. Then, it is just a matter of computing the first two terms of the succession ($|A_N| = \det A_N$ for brevity):

$$|A_1| = 2 \qquad |A_2| = \begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 4 - 1 = 3$$

And now we can use (16) to iteratively compute all $|A_N|$, e.g. $|A_3| = 2 \cdot 3 - 2 = 4$. To find $|A_N|$ for a *generic* N, we need to make an hypothesis, and then verify that it is compatible with (16). In this case, note that $|A_N| = N + 1$ $(*)$ for all the examples we explicitly computed. Then, by induction:

$$|A_{N+1}| \underset{(16)}{=} 2 \cdot |A_N| - |A_{N-1}| \underset{(*)}{=} 2 \cdot (N+1) - (N - 1 + 1) = 2N + 2 - N = (N+1) + 1$$

which is indeed compatible with $(*)$. So, substituting back in (15) we get:

$$I_1^{(N)} = \frac{1}{\sqrt{4\pi D\epsilon}} \frac{1}{\sqrt{N+1}} \underset{(a)}{=} \frac{1}{\sqrt{4\pi Dt}}$$

where in (a) we used $\epsilon = t/(N+1) \Rightarrow N + 1 = t/\epsilon$ from the discretization. Note that this result is *constant* with respect to ϵ or N (recall that t is fixed beforehand) and so taking the *continuum* limit leads immediately to I_1 (13):

$$I_1 \equiv \lim_{\substack{\epsilon \to 0 \\ N \to \infty}} \frac{1}{\sqrt{4\pi Dt}} = \frac{1}{\sqrt{4\pi Dt}}$$

which is coherent with the result we previously computed (11).