Central European University
Department of Applied Mathematics

THESIS

Pathwise approximation of the Feynman-Kac formula

Author:
Denis Nagovitcyn

Supervisor:
Tamás Szabados

May 17, 2015

Contents

Introduction **3**

1 Discretization of an Itô diffusion **5**
1.1 Twist & Shrink construction 5
1.2 Convergence results 8

2 Discretization of the Feynman-Kac formula **14**
2.1 Solving real-valued Schrödinger equation 14
2.2 Application to the Black-Scholes model 23

Conclusion **27**

Introduction

In the middle of the twentieth century there were essentially two different mathematical formulations of quantum mechanics: the differential equation of Schrödinger, and the matrix algebra of Heisenberg. The two, apparently dissimilar approaches, were proved to be mathematically equivalent and the paradigm of non-relativistic quantum theory remained unchanged before an article written by Richard Feynman [7] was published in 1948. His article was an attempt to present the third approach.

In essence, the new idea was about associating a quantity known as a probability amplitude (which describes a position of an elementary particle) with an entire path of a particle in space-time. Prior to that idea, probability amplitude was considered to be dependent on the position of the particle at a specific points in time. As was noted by the author: *"There is no practical advantage to this, but the formula* (1) *is very suggestive if a generalization to a wider class of action functionals is contemplated."* We will see in an instant that author's remark indeed appeared to be true.

Based on physics intuition Feynman embodied his thoughts in the following formula which sometimes referred as Feynman integral:

$$\psi(x,t) = \frac{1}{A}\int_{\Omega^x} \exp\left\{\frac{i}{\hbar}\int_0^t \left[\frac{1}{2}\left(\frac{\partial w}{\partial s}\right)^2 - V(w_s)\right] \mathrm{d}s\right\} g(w_t)\prod_t \mathrm{d}w_t\,, \tag{1}$$

where $\hbar$ is the Plank constant, $g(w) = \psi(w,0)$ is an initial condition and A is a normalizing constant for the Lebesgue-type infinite product measure $\prod_{0\leq s\leq t} \mathrm{d}w(s)$ over the space of trajectories $\Omega^x[0,t]$. It should be noted that the quantity $L(\dot{w},t) = \frac{1}{2}\left(\frac{\partial w}{\partial s}\right)^2 - V(w_s)$ is known as the Lagrangian and the integral $\int_0^t L(\dot{w},s)\,\mathrm{d}s$ is the classical action integral along the path $W = (w_\tau, 0 < \tau \leq t)$.

Notice that the infinite product measure is not a well-defined mathematical object and Feynman was well aware of that. His hope was that cleverly defined normalizing constant will make it sensible. Unfortunately, he has not presented a mathematically valid form for the constant, instead he proceeded with a famous conjecture that the function (1) solves a complex-valued Schrödinger equation:

$$\frac{1}{i}\frac{\partial\psi}{\partial t} = \frac{\hbar}{2}\frac{\partial^2\psi}{\partial x^2} - V(x)\,\psi\,. \tag{2}$$

As was noted in 1960 by Kiyosi Itô [8]: *"It is easy to see that* (1) *solves* (2) *unless we require mathematical rigor"*. It should be emphasized that despite numerous attempts mathematically valid solution of (2) still hasn't been presented. Nevertheless there is a silver lining, having in mind real-valued Feynman integral:

$$\psi(x,t) = \frac{1}{A}\int_{\Omega^x} \exp\left\{-\int_0^t \left[\frac{1}{2}\left(\frac{\partial w}{\partial s}\right)^2 - V(w_s)\right] \mathrm{d}s\right\} g(w_t)\prod_t \mathrm{d}w_t\,, \tag{3}$$

Mark Kac [11] was able to solve a real-valued Schrödinger equation (also referred as the heat equation):

$$\frac{\partial \psi}{\partial t} = \frac{1}{2}\,\nabla^2 \psi - V(x)\,\psi\,,$$

where ∇^2 stands for the Laplace operator. And the solution is the celebrated Feynman-Kac formula:

$$\psi(x,t) = \int_{C[0,t]} \exp\left\{ \int_0^t V[w(s)]\,\mathrm{d}s \right\} g[w(t)]\,\mathrm{d}\mathbb{P}^x(w)\,, \tag{4}$$

where $C[0,t]$ is the space of continuous trajectories and $\mathbb{P}^x$ is the Wiener measure, which came out from (3) by moving the term $e^{\left(-\frac{1}{2}\left(\frac{\partial w}{\partial t}\right)^2\right)}$ into the integral measure.

Feynman-Kac formula plays very important role in science, it was applied to variety of problems across disciplines. As an illustrative example in the last section we will demonstrate derivation of an option pricing formula.

The aim of the current thesis is to provide a discrete analog of the formula (4) based on the strong approximation of Brownian motion by simple symmetric random walks. An underlying process $w(t)$ will take form of an Itô diffusion with nonconstant coefficients. Weakly convergent (in distribution) approximations were given earlier by M. Kac [10] and E. Csáki [6]. This research may be considered as an extension of a specific case introduced in [21], where underlying stochastic process $w(t)$ was assumed to have constant coefficients.

Results of the current thesis may be useful to give a rigorous prove for the complex-valued case (2). According to the work in progress by Tamás Szabados, the normalizing constant A from the formula (1) might be set in such a way that a discrete analog of the Lebesgue-type product measure $\frac{dw}{A}$ has a binomial distribution, which justifies application of a construction similar to the one presented in the current thesis. There exists vast amount of articles intended to solve the complex-valued case, the most significant ones are [1], [3], [4], [8], [15]. For example in [8] Itô solves equation (2) for the case of function $V(\cdot)$ being a constant, despite substantial degree of mathematical rigor, his solution uses heuristic arguments. The same might be concluded about the majority of existing researches on this topic.

In order to fulfill the stated goal we will follow the following outline: using "twist & shrink" [18] construction of Brownian motion we will establish a discrete analog of a real-valued Schrödinger equation and its solution, which converges to the continuous case.

Chapter 1

Discretization of an Itô diffusion

In current chapter we are going to present a discrete version of time homogeneous Itô diffusion. In the first section discretization of a Wiener process will be given, based on which in the second section several approximation schemes will be discussed and their convergence to the continuous process will be provided.

1.1 Twist & Shrink construction

A basic tool of the present paper is an elementary construction of Brownian motion. This construction, taken from [18], is based on a nested sequence of simple, symmetric random walks that uniformly converges to Brownian motion on bounded time intervals with probability 1. This will be called *"twist & shrink"* construction. This method is a modification of the one given by Frank Knight in 1962 [13] and its simplification by Pal Revesz in 1990 [16].

We summarize the major steps of the "twist & shrink" construction here. We start with a sequence of independent, symmetric random walks (RW):

$$S_m(0) = 0, \quad S_m(n) = \sum_{k=1}^{n} X_m(k) \quad (n \geq 1) \quad ,$$

based on an infinite matrix of independent and identically distributed random variables $X_m(k)$,

$$\mathbb{P}\{X_m(k) = \pm 1\} = \frac{1}{2} \quad (m \geq 0, k \geq 1) \quad ,$$

defined on the same complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

For the first, we would like to decrease the size of a step in time by a factor of two as we move along the sequence, since ultimately we are pursuing Brownian motion which take values for each real time. That is for a m^{th} random walk $S_m(n)$ we would like to have $\Delta n = \frac{1}{2^m}$, but then a natural question arises: how much do we have to decrease correspondingly the size of a step in space to preserve essential properties of a random walk? Surprisingly, the answer to this question is that we have two care about only one property, namely we have to preserve square root of the expected squared distance from the origin, which is a standard deviation of a simple symmetric RW:

$$\sqrt{Var[S_m(n))]} = \sqrt{\mathbb{E}\sum_i (S_{m,i}(n) - \mathbb{E}[S_m(n)])^2} = \sqrt{\mathbb{E}\sum_i (S_{m,i}(n) - 0)^2} = \sqrt{n}\,.$$

Thus we may conclude that after n steps in time on average our random walk will be at $\sqrt{n}$ distance from the origin. So it follows that in order to have n steps in one time unit, the step size in space have to be $1/\sqrt{n}$. And this is called shrinking.

Next, from the independent RW's we want to create dependent ones in such a way that after shrinking each consecutive RW becomes a refinement of the previous one. Since the spatial unit will be halved at each consecutive row, we define stopping times by $T_m(0) = 0$, and for $k \geq 0$,

$$T_m(k+1) = min\{n : n > T_m(k), |S_m(n) - S_m(T_m(k))| = 2\} \quad (m \geq 1)$$

These are random time instants when a RW visits even integers. After shrinking the spatial unit by half, a suitable modification of this RW will visit the same integers in the same order as the previous RW. And this is called twisting.

We operate here on each point $\omega \in \Omega$ of the sample space separately, i.e. we fix a sample path of each RW. We define twisted RW's $\tilde{S}_m$ recursively for k=1,2,... using $\tilde{S}_{m-1}$ starting with $\tilde{S}_0(n) = S_0(n)$ $(n \geq 0)$ and $\tilde{S}_m(0) = 0$ for any $m \geq 0$. With each fixed m we procedd for k=0,1,2,... successively, and for every n in the corresponding bridge, $T_m(k) < n < T_m(k+1)$. Each bridge is flipped if its sign differs from the desired: $\tilde{X}_m(n) = \pm X_m(n)$, depending on whether $S_m(T_m(k+1)) - S_m(T_m(k)) = 2\tilde{X}_{m-1}(k+1)$ or not. So $\tilde{S}_m(n) = \tilde{S}_m(n-1) + \tilde{X}_m(n)$.

Then $(\tilde{S}_m(n))_{n\geq 0}$ is still a simple symmetric RW [18, Lemma 1]. The twisted RW's have the desired refinement property:

$$\tilde{S}_{m+1}(T_{m+1}(k)) = 2\tilde{S}_m(k) \quad (m \geq 0, k \geq 0).$$

The sample paths of $\tilde{S}_m(n)$ $(n \geq 0)$ can be extended to continuous functions by linear interpolation, this way one gets $\tilde{S}_m(t)$ $(t \geq 0)$ for $\forall t \in \mathbb{R}^+$.

Putting all together, the m^{th} "twist & shrink" RW is defined by

$$B_m(t) = 2^{-m}\tilde{S}_m(t2^{2m}).$$

As an illustration of the "twist & shrink" construction we will present step 1 approximation of the initially simulated random walk. For two independent simple symmetric random walks please refer to the Figure 1.1. As reader may anticipate our goal is to refine random walk $S_0(t)$ with another random walk $S_1(t)$ using the described method. Notice that we took $S_0(t)$ up to time 4 and to refine it, in the case of this particular simulation, we needed $S_1(t)$ up to time 10. In the pictures below dots on the graph correspond to the values of the initial random walk.

We start our process with $S_1(0)$ and wait until it hits $2 = 2^m$, for $m = 1$, it happens at time $t = 2$. Then we wait for another move of $S_1(t)$ of magnitude 2 and it happens at $t = 4$, however corresponding move of $S_0(t)$ was down whereas $S_1(t = 4)$ moved up. So we reflect the whole path of $S_1(t)$ starting at $t = 2$ until the end and as it turns out one reflection was enough to mimic direction of moves of $S_0(t)$. Thus we are ready to properly scale twisted process $\tilde{S}_1(t)$, namely we divide its time units by 2^2 and spacial units by 2^1. As a result we have got $B_1(t)$ the first step approximation of a Brownian motion.

Convergence of the "twist & shrink" sequence to the Brownian motion is provided by the next theorem.

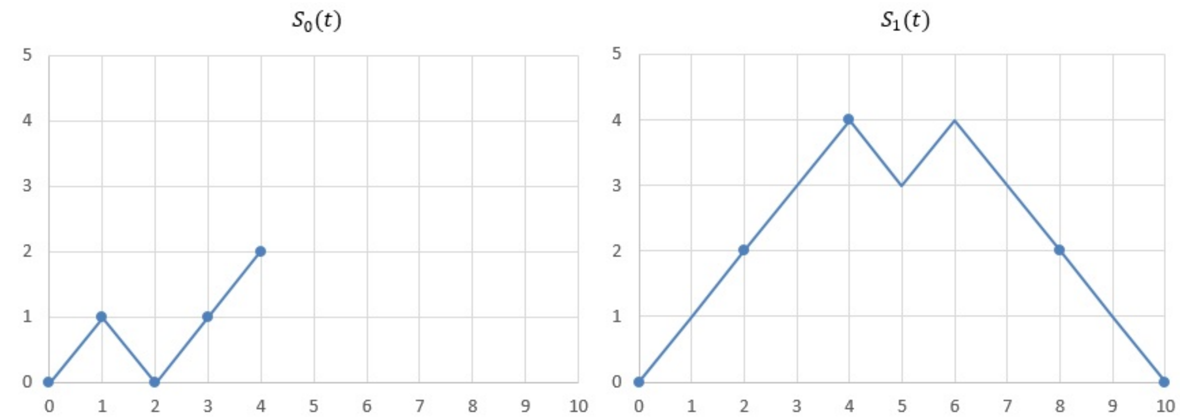

Figure 1.1: Two independent random walks

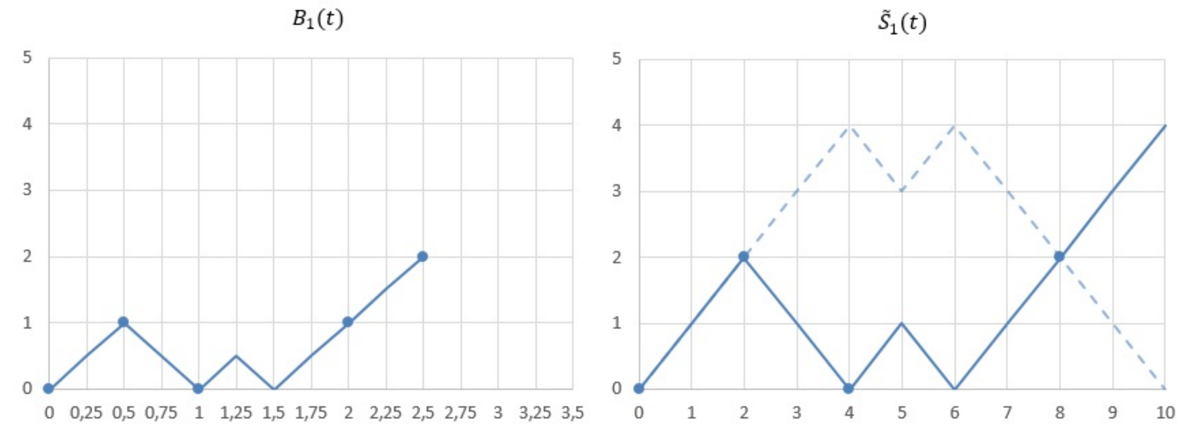

Figure 1.2: Twisted and shrunk random walk $S_1(t)$

Theorem A. *The sequence of random walks $B_m(t)$ uniformly converges to the Brownian motion $W(t)$ on bounded intervals of time with probability 1 as $m \to \infty$:*

$$\sup_{0 \le t \le T} |W(t) - B_m(t)| = O(m^{3/4} 2^{-m/2}), \quad \textit{for } \forall T \ge 0.$$

The proof of which may be found in [19, p. 84].

Conversely, with a given Wiener process $W(t)$, one can define the stopping times which yield to the Skorohod embedded random walks $\tilde{B}_m(k2^{-2m})$ into $W(t)$. For every $m \ge 0$ let $\tau_m(0) = 0$ and

$$s_m(k+1) = \inf\{\tau \,:\, \tau > \tau_m(k), |W(s) - W(s_m(k))| = 2^{-m}\} \quad (k \ge 0).$$

With these stopping times the embedded dyadic walks by definition are

$$\tilde{B}_m(k2^{-2m}) = W(\tau_m(k)) \quad (m \ge 0, k \ge 0).$$

This definition of $\tilde{B}_m$ can be extended to any real $t \ge 0$ by pathwise linear interpolation.

If a Wiener process is built by the "twist & shrink" construction described above using a sequence B_m of nested random walks and then one constructs the Skorohod embedded random walks $\tilde{B}_m$, it is natural to ask about their relationship. The next theorem demonstrates that they are asymptotically equivalent, the proof may be found in [18, p. 24-31].

Theorem B. *For any $E > 1$, and for any $F > 0$ and $m \geq 1$ such that $F\,2^{2m} \geq N(E)$ take the following subset of Ω :*

$$A^*_{F,m} = \left\{ \sup_{n>m} \sup_{k} |2^{-2n} T_{m,n}(k) - k\,2^{-2m}| < C_{E,F} m^{1/2} 2^{-m} \right\},$$

where $C_{E,F}$ is just a constant factor dependent on E and F, $T_{m,n}(k) = T_n \circ T_{n-1} \circ \cdots \circ T_m(k)$ for $n > m \geq 0$ and $k \in [0, F\,2^{2m}]$. Then

$$\mathbb{P}\{(A^*_{F,m})^c\} \leq \frac{2}{1 - 4^{1-E}} (F 2^{2m})^{1-E}.$$

Moreover, $\lim_{n\to\infty} 2^{-2n} T_{m,n}(k) = t_m(k)$ exists almost surely and on $A^*_{F,m}$ we have

$$\tilde{B}_m(k 2^{-2m}) = W(t_m(k)) \quad (0 \leq k 2^{-2m} \leq F).$$

Further, almost everywhere on $A^*_{F,m}$ and any $0 < \delta < 1$, we have $\tau_m(k) = t_m(k)$,

$$\sup_{0 \leq k 2^{-2m} \leq K} |\tau_m(k) - k 2^{-2m}| \leq C_{E,F} m^{1/2} 2^{-m},$$

and

$$\max_{1 \leq k 2^{-2m} \leq K} |\tau_m(k) - \tau_m(k-1) - 2^{-2m}| \leq (7/\delta) 2^{-2m(1-\delta)}.$$

Essentially the second theorem tells us two important things. First, distance between time instances of "twist & shrink" process $B_m(k)$ and Skorokhod embedded process $\tilde{B}_m(k)$ approaches 0. In other words for m big enough two partitions of the time axis have the same mesh. Second, whenever we are in the subset $A^*_{F,m}$ we may use stopped Brownian motion instead of "twist & shrink" process.

1.2 Convergence results

Consider the following Itô diffusion:

$$\mathrm{d}X_t = a(X_t)\,\mathrm{d}t + b(X_t)\,\mathrm{d}W_t,$$

or in the integral form:

$$X_t = x_0 + \int_0^t a(X_s)\,\mathrm{d}s + \int_0^t b(X_s)\,\mathrm{d}W_s, \tag{1.1}$$

where x_0 is an initial state which we assume to be deterministic for the sake of simplicity, W_t stands for the Wiener process and $a(X_t)$ and $b(X_t)$ are drift and diffusion coefficients respectively satisfying certain conditions which we will specify below.

All the requirements that we are going to list are very natural since we would like every term (1.1) to make sense.

First, we will describe a class of functions for which the Itô integral is defined. Using the notation stated above, function $b(\cdot)$ has to satisfy the following properties:

(i) $(t, w) \to b(t, w)$ is $\mathcal{B} \times \mathcal{F}$-measurable, where $\mathcal{B}$ denotes the Borel σ-algebra on $[0, \infty)$

(ii) $b(t, w)$ is $\mathcal{F}_t$-adopted, where $\mathcal{F}_t$ stands for the natural filtration generated be the Brownian motion

(iii) $\mathbb{E}\left[\int_0^t b^2(s,w)\,\mathrm{d}s\right] < \infty$.

Reader should note that class of functions satisfying conditions *(i)-(iii)* is not the largest one for which Itô integral is defined, for the detailed discussion refer to [9, p. 25].

Second, we would like function $a(\cdot)$ to be such that $\mathbb{E}[\int_0^t |a(s,w)|\,\mathrm{d}s] < \infty$.

As known from the theory of stochastic differential equations global Lipschitz continuity is sufficient to guarantee existence and uniqueness of the solution of an SDE: *Let $T \geq 0$ and $a(\cdot)\colon \mathbb{R} \to \mathbb{R},\ b(\cdot) : \mathbb{R} \to \mathbb{R}$ be measurable functions satisfying*

$$|a(x) - a(y)| \leq K|x-y| \tag{1.2}$$

$$|b(x) - b(y)| \leq K|x-y|, \tag{1.3}$$

for some positive constant K and any $x, y \in \mathbb{R}$.

It is worth mentioning that linear growth condition for functions $a(\cdot), b(\cdot)$ is a consequence of global Lipschitz continuity taking $y = 0$, more precisely:

$$|a(x)| \leq C(1+|x|) \tag{1.4}$$

$$|b(x)| \leq C(1+|x|), \tag{1.5}$$

for some positive constant C and any $x \in \mathbb{R}$.

To proceed we need a useful upper bound for the second moment of Itô diffusion X_t. We are going to give a proof for the case of one dimensional Itô diffusion. Reader may refer to the Theorem 4.5.4 of [12, p. 136] for the case of general Itô process. The proof that we are about to give uses very straightforward approach and as a result gives an upperbound which is a little different from the conventional one, however it is still usable.

Lemma 1. *Assume that we have an Itô process X_t satisfying conditions (1.2)-(1.3) then*

$$\mathbb{E}|X_t|^2 \leq e^{t\,C_1}(3x_0^2+1), \qquad \forall x_0 \in \mathbb{R},\ \forall t \in [0,T],$$

where $C_1 = 6\,C^2(T+1)$

Proof. Take a process X_t in the form of equation (1.1) and square it to get:

$$X_t^2 \leq 3\left(x_0^2 + \left(\int_0^t a(X_s)\,\mathrm{d}s\right)^2 + \left(\int_0^t b(X_s)\,\mathrm{d}W_s\right)^2\right).$$

Now taking expectation of both sides of the above inequality and using some basic results will lead us to the desired upperbound:

$$\begin{aligned}
\mathbb{E}|X_t|^2 &\leq 3\,x_0^2 + 3\,t\,\mathbb{E}\int_0^t a(X_s)^2\,\mathrm{d}s + 3\,\mathbb{E}\int_0^t b(X_s)^2\,\mathrm{d}s \\
&\leq 3\,x_0^2 + 6\,T\,C^2\mathbb{E}\int_0^t (1+|X_s|^2)\,\mathrm{d}s + 6\,C^2\,\mathbb{E}\int_0^t (1+|X_s|^2)\,\mathrm{d}s \\
&\leq 3\,x_0^2 + 6\,t\,C^2(T+1) + 6\,C^2(T+1)\int_0^t \mathbb{E}|X_s|^2\,\mathrm{d}s,
\end{aligned}$$

notice that to get the first inequality we have used Cauchy-Schwartz inequality and Itô isometry, while to obtain the second we used linear growth condition for functions $a(\cdot)$, $b(\cdot)$. Now to complete the proof use Gronwall's lemma to get:

$$\mathbb{E}|X_t|^2 \leq 3\,x_0^2 + 6\,t\,C^2(T+1) + 6\,C^2(T+1)\int_0^t e^{6\,C^2(T+1)(t-s)}(3\,x_0^2 + 6\,s\,C^2(T+1))\,\mathrm{d}s.$$

The result follows after easy integral calculus. □

Now let's see how we can construct discrete approximation that converges strongly to the Itô diffusion X_t. Firstly we are going to use Skorokhod embedded random walks. Let $\tilde{B}^m_\tau$ be m^{th} step simple symmetric random walk and $\tilde{X}^m$ be m^{th} approximation of X given by

$$\tilde{X}^m_{\tau_{n+1}} = \tilde{X}^m_{\tau_n} + a(\tilde{X}^m_{\tau_n})\Delta\tau_{n+1} + b(\tilde{X}^m_{\tau_n})\Delta\tilde{B}^m_{\tau_{n+1}},$$

where $a(\cdot)$ and $b(\cdot)$ are functions specified above, $\Delta\tau_{n+1} = \tau_{n+1} - \tau_n$, with τ_i being Skorokhod embedded times, $\Delta\tilde{B}^m_{t_{n+1}} = \tilde{B}^m_{t_{n+1}} - \tilde{B}^m_{t_n}$ and initial condition $\tilde{X}^m_0 = x$ for $\forall m$. Alternatively we may rewrite our approximation in integral form:

$$\tilde{X}^m_{\tau_n} = x + \sum_{i=1}^{n} a(\tilde{X}^m_{\tau_i})\Delta\tau_{i+1} + \sum_{i=1}^{n} b(\tilde{X}^m_{\tau_i})\Delta\tilde{B}^m_{\tau_{i+1}}.$$

Furthermore instead of working on the whole probability space let's restrict ourselves to the subspace $A^*_{F,m}$ which was defined before. Then $\tilde{B}^m_{\tau_i} = W_{\tau_i}$ and we may rewrite our discrete approximation as:

$$\tilde{X}^m_{\tau_n} = x + \sum_{i=1}^{n} a(\tilde{X}^m_{\tau_i})\Delta\tau_{i+1} + \sum_{i=1}^{n} b(\tilde{X}^m_{\tau_i})\Delta W_{\tau_{i+1}}. \tag{1.6}$$

Observe that by by linear interpolation we can extend our approximation to an arbitrary time $t \in [0, T]$, then

$$\tilde{X}^m_t = x + \int_0^{\tau_{N_t}} a(\tilde{X}^m_s)\,\mathrm{d}s + \int_0^{\tau_{N_t}} b(\tilde{X}^m_s)\,\mathrm{d}W_s, \tag{1.7}$$

where $\tau_{N_t} := \inf\{\tau_i : \tau_i > t\}$.

Notice that approximations (1.6) and (1.7) are very close to each other, namely their $\mathcal{L}^2$-convergence might be rigorously checked by techniques similar to the one in **Theorem 2** below. Intuitively, since by **Theorem B** $\Delta\tau_i$ is arbitrarily close to Δt_i for m large enough, which means that the first sum in (1.6) is a Riemann sum and the second sum approaches Itô integral as $m \to \infty$ by construction. We are going to check convergence only in case of the scheme (1.7).

Let us also introduce notation $\delta := \max_i \{\tau_{i+1} - \tau_i\}$ which stands for the maximum mesh size for a given approximation at level m.

Theorem 1. *Given an Itô diffusion X_t satisfying conditions (1.2), (1.3) and discrete approximation $\tilde{X}^m_t$ described above as $m \to \infty$*

$$\sup_{t\le T} \mathbb{E}\,[X_t - \tilde{X}^m_t]^2 \to 0\,.$$

Proof. Let's denote

$$Z(T) := \sup_{t\le T} \mathbb{E}\,|X_t - \tilde{X}^m_t|^2 \le \mathbb{E}\sup_{t\le T}|X_t - \tilde{X}^m_t|^2$$

then

$$\begin{aligned} Z(T) \le \mathbb{E}\sup_{t\le T}\bigg| &\int_0^{\tau_{N_t}} [a(X_s) - a(\tilde{X}^m_s)]\,\mathrm{d}s + \int_0^{\tau_{N_t}} [b(X_s) - b(\tilde{X}^m_s)]\,\mathrm{d}W_s \\ &+ \int_{\tau_{N_t}}^{t} a(X_s)\,\mathrm{d}s + \int_{\tau_{N_t}}^{t} b(X_s)\,\mathrm{d}W_s\bigg|^2 \\ \le{}& 3\,(R_1 + R_2 + R_3), \end{aligned}$$

where R_1, R_2 and R_3 will be given explicitly below.

$$\begin{aligned}
R_1 &:= \mathbb{E} \sup_{t\le T} \left| \int_0^{\tau_{N_t}} [a(X_s) - a(\tilde{X}_s^m)] \,\mathrm{d}s \right|^2 \\
&\le \mathbb{E} \sup_{t\le T} \tau_{N_t} \int_0^{\tau_{N_t}} |a(X_s) - a(\tilde{X}_s^m)|^2 \,\mathrm{d}s \le K^2\, \tau_{N_T}\, \mathbb{E} \int_0^{\tau_{N_T}} |X_s - \tilde{X}_s^m|^2 \,\mathrm{d}s \\
&\le K^2\, (T+\epsilon) \int_0^{T+\epsilon} \sup_{t\le s} \mathbb{E}\, |X_t - \tilde{X}_t^m|^2 \,\mathrm{d}s = K^2\, (T+\epsilon) \int_0^{T+\epsilon} Z(s)\,\mathrm{d}s\,,
\end{aligned}$$

since for m large enough $\tau_{N_T} = T + \epsilon$ by **Theorem B** and during the derivation we have used Cauchy-Schwartz inequality, Lipschitz continuity and Fubini theorem to get the result. In a similar manner we are going to deal with the second term.

$$\begin{aligned}
R_2 &:= \mathbb{E} \sup_{t\le T} \left| \int_0^{\tau_{N_t}} [b(X_s) - b(\tilde{X}_s^m)] \,\mathrm{d}W_s \right|^2 \\
&\le 4\,\mathbb{E} \left| \int_0^{\tau_{N_T}} [b(X_s) - b(\tilde{X}_s^m)] \,\mathrm{d}W_s \right|^2 \le 4\,K^2 \int_0^{\tau_{N_T}} \mathbb{E}|X_s - \tilde{X}_s^m|^2 \,\mathrm{d}s \\
&\le 4\,K^2 \int_0^{T+\epsilon} \sup_{t\le s} \mathbb{E}\, |X_t - \tilde{X}_t^m|^2 \,\mathrm{d}s = 4\,K^2 \int_0^{T+\epsilon} Z(s)\,\mathrm{d}s\,,
\end{aligned}$$

where we have used Doob's inequality, Itô isometry and Lipschitz continuity.

As reader may have noticed the first two terms involve integrating up to time τ_{N_t} meanwhile the third term takes care of the remainder:

$$\begin{aligned}
R_3 &:= \mathbb{E} \sup_{t\le T} \left| \int_{\tau_{N_t}}^t a(X_s)\,\mathrm{d}s + \int_{\tau_{N_t}}^t b(X_s)\,\mathrm{d}W_s \right|^2 \\
&\le 2\,\mathbb{E} \sup_{t\le T} \left\{ \left(\int_{\tau_{N_t}}^t a(X_s)\,\mathrm{d}s \right)^2 + \left(\int_{\tau_{N_t}}^t b(X_s)\,\mathrm{d}W_s \right)^2 \right\} \\
&\le 2\,\mathbb{E} \sup_{t\le T} \left[(t - \tau_{N_t}) \int_{\tau_{N_t}}^t a^2(X_s)\,\mathrm{d}s \right] + 8\,\mathbb{E} \int_{\tau_{N_T}}^T b^2(X_s)\,\mathrm{d}s \\
&\le (4\,\delta\,C^2 + 16\,C^2) \int_{\tau_{N_T}}^T (1 + \mathbb{E}X_s^2)\,\mathrm{d}s \le (4\,T\,C^2 + 16\,C^2) e^{T\,C_1} (3\,x_0^2 + 1)\,\delta \\
&\le C_{1,T}\,\delta\,.
\end{aligned}$$

Denoting $C_{1,T} := (4\,T\,C^2 + 16\,C^2) e^{T\,C_1} (3\,x_0^2 + 1)$ and $C_{2,T} := K^2\,(T + \epsilon + 4)$ and collecting all the estimates, for constants $C_{1,T}$ and $C_{2,T}$ being solely dependent on T and not δ, we have

$$Z(T) \le C_{1,T}\,\delta + C_{2,T} \int_0^{T+\epsilon} Z(s)\,\mathrm{d}s\,.$$

Now using Gronwall's lemma one can get:

$$Z(T) \le C_T\,\delta\,.$$

Next, by the consequence of Jensen's inequality known as the Lyapunov's inequality, for $0 < s < t$

$$\left(\mathbb{E}|f|^s\right)^{1/s} \le \left(\mathbb{E}|f|^t\right)^{1/t}\,,$$

setting $s = 1$, $t = 2$ and $f = |X_t - X_t^m|$we have:

$$\sup_{t\le T} \mathbb{E}\,|X_t - X_t^m| \le \sqrt{Z(T)} \le C_T\,\sqrt{\delta}\,.$$

Thus from the **Theorem B** we may infer that as $m \to \infty$, $\delta \to 0$ which gives convergence on the subspace $A^*_{F,m}$ of Ω. Moreover since $\mathbb{P}\{(A^*_{F,m})^c\}$ goes to zero, claim of the theorem holds for the whole space. $\square$

For discussions in the sequel of the paper it is unfortunate to have random time instances. In fact, it is more convenient to have a dyadic rational points instead of a sequence of stopping times. That is why we are motivated to upgrade the discrete approximation in (1.6).

Let us use "twist & shrink" construction of a Brownian motion and denote our new approximation as X^m then:

$$\Delta X^m_{t_i} = a(X^m_{t_i})\Delta t_{i+1} + b(X^m_{t_i})\Delta B^m_{t_{i+1}} \ ,$$

with $X^m_0 = x$ for $\forall m$. Once again we may rewrite it as

$$X^m_{t_n} = x + \sum_{i=0}^{n} a(X^m_{t_i})\Delta t_{i+1} + \sum_{i=0}^{n} b(X^m_{t_i})\Delta B^m_{t_{i+1}}. \tag{1.8}$$

Using linear interpolation we may have a discrete scheme given in the following form

$$X^m_t = x + \int_0^{t_{N_t}} a(X^m_s)\,\mathrm{d}s + \sum_{i=0}^{N_t} b(X^m_{t_i})\Delta B^m_{t_{i+1}} \ , \tag{1.9}$$

where N_t is such that $t_{N_t} = \inf\{t_i : t_i > t\}$. Note that discrete schemes (1.8) and (1.9) are also very close because now we have Riemann sum in (1.8) from the start.

Now it is natural to ask about closeness of two discrete approximations for $\tilde{X}^m_t$ and X^m_t given by (1.7) and (1.9). To demonstrate that they are asymptotically (in $\mathcal{L}^1$ sense) close, we will use technique similar to the one in **Theorem 1**.

Theorem 2. *Given two discrete approximations X^m_t and $\tilde{X}^m_t$ described above, as $m \to \infty$*

$$\sup_{t \le T} \mathbb{E}\,[X^m_t - \tilde{X}^m_t]^2 \to 0.$$

Proof. Let's restrict our working space to the subspace $A^*_{F,m}$ of Ω, so $W_\tau = \tilde{B}^m_\tau$. Without loss of generality we may assume that $\tau_{N_t} \ge t_{N_t}$.

Let's denote

$$U(T) := \sup_{t \le T} \mathbb{E}|\tilde{X}^m_t - X^m_t|^2 \le \mathbb{E}\sup_{t \le T}|\tilde{X}^m_t - X^m_t|^2$$

then

$$\begin{aligned} U(T) \le \mathbb{E}\sup_{t \le T}\Bigg| &\int_0^{t_{N_t}} [a(\tilde{X}^m_s) - a(X^m_s)]\,\mathrm{d}s + \sum_{i=0}^{N_t}\int_{\tau_i}^{\tau_{i+1}} [b(\tilde{X}^m_s) - b(X^m_{t_i})]\,\mathrm{d}W_s \\ &+ \int_{t_{N_t}}^{\tau_{N_t}} a(\tilde{X}^m_s)\,\mathrm{d}s + \sum_{i=0}^{N_t} b(X^m_{t_i})\Delta\tilde{B}^m_{t_{i+1}} - \sum_{i=0}^{N_t} b(X^m_{t_i})\Delta B^m_{t_{i+1}}\Bigg|^2 \\ &\le 4\,(E_1 + E_2 + E_3 + E_4), \end{aligned}$$

where E_1, E_2, E_3 and E_4 will be given explicitly below.

$$\begin{aligned} E_1 &:= \mathbb{E}\,\sup_{t \le T}\left|\int_0^{t_{N_t}} [a(\tilde{X}^m_s) - a(X^m_s)]\,\mathrm{d}s\right|^2 \\ &\le \mathbb{E}\,\sup_{t \le T} t_{N_t}\int_0^{t_{N_t}} |a(\tilde{X}^m_s) - a(X^m_s)|^2\,\mathrm{d}s \le K^2\,t_{N_T}\int_0^{t_{N_T}} \mathbb{E}\,|\tilde{X}^m_s - X^m_s|^2\,\mathrm{d}s \\ &\le K^2\,T\int_0^T \sup_{t \le s} \mathbb{E}\,|\tilde{X}^m_t - X^m_t|^2\,\mathrm{d}s = K^2\,T\int_0^T U(s)\,\mathrm{d}s\,, \end{aligned}$$

where we have used Cauchy-Schwartz inequality, Lipschitz continuity and Fubini theorem.

$$E_2 := \mathbb{E}\sup_{t\le T}\left|\sum_{i=0}^{N_t}\int_{\tau_i}^{\tau_{i+1}}[b(\tilde{X}_s^m)-b(X_{t_i}^m)]\,\mathrm{d}W_s\right|^2 \le 4\,\mathbb{E}\left|\sum_{i=0}^{N_T}\int_{\tau_i}^{\tau_{i+1}}[b(\tilde{X}_s^m)-b(X_{t_i}^m)]\,\mathrm{d}W_s\right|^2$$

$$\le 4\,\mathbb{E}\left\{\sum_{i=0}^{N_T}\left|\int_{\tau_i}^{\tau_{i+1}}[b(\tilde{X}_s^m)-b(X_{t_i}^m)]\,\mathrm{d}W_s\right|^2\right.$$

$$\left.+2\sum_{i<j}\int_{\tau_i}^{\tau_{i+1}}[b(\tilde{X}_s^m)-b(X_{t_i}^m)]\,\mathrm{d}W_s\int_{\tau_j}^{\tau_{j+1}}[b(\tilde{X}_s^m)-b(X_{t_j}^m)]\,\mathrm{d}W_s\right\},$$

where we have used Doob's martingale inequality to get rid of the supremum. It is easy to show that expectation of the second sum is 0 introducing conditional expectation and using property of Itô integral. Thus we have

$$E_2 \le 4\,\mathbb{E}\sum_{i=0}^{N_T}\left|\int_{\tau_i}^{\tau_{i+1}}[b(\tilde{X}_s^m)-b(X_{t_i}^m)]\,\mathrm{d}W_s\right|^2 = 4\sum_{i=0}^{N_T}\int_{\tau_i}^{\tau_{i+1}}\mathbb{E}\,|b(\tilde{X}_s^m)-b(X_{t_i}^m)|^2\,\mathrm{d}s$$

$$\le 4\,K^2\sum_{i=0}^{N_T}\int_{\tau_i}^{\tau_{i+1}}\sup_{t\le s}\mathbb{E}\,|\tilde{X}_t^m-X_t^m|^2\,\mathrm{d}s = 4\,K^2\int_0^{T+\epsilon}U(s)\,\mathrm{d}s\,,$$

where we have used Itô isometry, Fubini theorem and Lipschitz continuity.

Reader should note that in the current case third term takes care of the remainder:

$$E_3 := \mathbb{E}\,\sup_{t\le T}\left|\int_{t_{N_t}}^{\tau_{N_t}}a(\tilde{X}_s^m)\,\mathrm{d}s\right|^2 \le \mathbb{E}\,\sup_{t\le T}(\tau_{N_t}-t_{N_t})\int_{t_{N_t}}^{\tau_{N_t}}|a(\tilde{X}_s^m)|^2\,\mathrm{d}s$$

$$\le 2\,C^2\rho\int_{t^*}^{\tau^*}(1+\mathbb{E}(\tilde{X}_s^m))^2\,\mathrm{d}s \le 2\,C^2\rho^2 e^{T\,C_1}(3\,x^2+1)$$

$$\le C_{3,T}\rho^2,$$

where we have used Cauchy-Schwartz inequality, consequence of Lipschitz continuity, **Lemma 1**, $\rho := \sup_{t\le T}(\tau_{N_t}-t_{N_t})$ and where limits of integration denoted as t^* and τ^* corresponds to ρ. Also notice that by **Theorem B** $\rho \le C_{E,F}\,m^{\frac{1}{2}}2^{-m}$, so $\rho \to 0$ as $m \to \infty$. Fortunately the last term could be handled easily:

$$E_4 := \left|\sum_{i=0}^{N_t}b(X_{t_i}^m)\Delta\tilde{B}_{t_{i+1}}^m - \sum_{i=0}^{N_t}b(X_{t_i}^m)\Delta B_{t_{i+1}}^m\right|^2 = 0\,,$$

since two sums are identical. And using the same logic as in the end of the previous theorem we have the claim. □

Theorem 3. *Given the discrete approximation scheme $X_{t_k}^m$ and Lipschitz continuous functions $a(\cdot)$, $b(\cdot)$ as $m \to \infty$*

$$\sup_{t\in[0,T]}\mathbb{E}[X_t - X_{t_k}^m]^2 \to 0$$

Proof. By applying simple estimate for the square of the sum we have

$$\sup_{t\in[0,T]}\mathbb{E}[X_t - X_{t_k}^m]^2 \le 2\left(\sup_{t\in[0,T]}\mathbb{E}[X_t-\tilde{X}_{t_k}^m]^2 + \sup_{t\in[0,T]}\mathbb{E}[\tilde{X}_{t_k}^m - X_{t_k}^m]^2\right).$$

Take limit of both sides as $m \to \infty$ and use **Theorem 1** and **2**. □

Thus we may choose scheme $X_{t_k}^m$ as a discrete approximation of X_t for further analysis because it is more advantageous to have deterministic partition in time.

Chapter 2

Discretization of the Feynman-Kac formula

In this chapter we are going to prove existence part of the solution for the real-valued Schrödinger equation based on the discrete approximation. In the first section discrete version of the Feynman-Kac formula will be given and its convergence to the continuous case will be shown. In the second section application of the Feynman-Kac formula to the option pricing theory of mathematical finance will be demonstrated.

2.1 Solving real-valued Schrödinger equation

We are going to prove that if functions $g(\cdot), r(\cdot) \in C_0^2(\mathbb{R})$ then the differential equation

$$\begin{aligned} \frac{\partial f}{\partial t} &= A\,f - r\,f; \qquad t > 0,\, x \in \mathbb{R} \\ f(0,x) &= g(x); \qquad x \in \mathbb{R}, \end{aligned} \tag{2.1}$$

has a solution known as Feynman-Kac functional:

$$f(t, x_0) = \mathbb{E}^{x_0}\left[e^{-\int_0^t r(X_s)\,\mathrm{d}s} g(X_t)\right].$$

Where A is an operator which is called infinitesimal generator and it is defined by

$$A\,f(x) = \lim_{t\to 0} \frac{\mathbb{E}^{x_0}[f(X_t)] - f(x_0)}{t},$$

given that $\mathcal{D}^A$ is the domain of A and $f \in \mathcal{D}^A$. For the brief introduction to the concept of infinitesimal generators reader may refer to [9, p. 117] and for more extensive view to [14, p. 216].

For arbitrary function $f \in \mathcal{D}^A$ there is no direct way to express generator A in a usual sense as a sum of derivative operators, it is only possible using results from distribution theory. A brief discussion of this approach may be found in [17, p.108] however handling this case is not an easy task. That is why we will analyze heat equation with dissipation term given in the most general form by (2.1).

For a special case when $f \in C_0^2$ there exist a way to compute A and it turns out that:

$$A\,f(x) = a(x)\,\partial_x f + \frac{1}{2} b(x)^2\,\partial_{xx} f,$$

where ∂_x and ∂_{xx} denotes first and second derivative with respect to x.

If one can show that Feynman-Kac functional $f(t,x_0)$ is in C_0^2 then differential equation (2.1) is equivalent to

$$\frac{\partial f}{\partial t} = a(x)\,\partial_x f + \frac{1}{2}b(x)^2\partial_{xx}f - r(x)f. \tag{2.2}$$

Unfortunately as pointed by [17, p. 119] smoothness of $f(t,x_0)$ is not something easy to check in case of X_t being a general Itô diffusion. For the simple case of an Itô diffusion with constant coefficients a, b reader should refer to [22] to see that in this situation Feynman-Kac functional and its discrete analog are C_0^2 functions and hence heat equation has solution of the form (2.2).

Since we are working with more general case than in [22] we are going to solve the differential equation (2.1), where generator A is not specified.

We would like to continue current section by presenting a discrete version of the heat equation (2.1).

Lemma 2 (Discrete Feynman-Kac formula). *Given time-homogeneous discrete Itô diffusion $X^m_{t_k}$ defined above, with coefficients $a(\cdot)$, $b(\cdot)$ Lipschitz continuous. For functions $r(\cdot)$, $g(\cdot) \in C_0^2$, discrete Feynman-Kac functional*

$$f_m(t_k,x_0) = \mathbb{E}^{x_0}\left[e^{-\sum_{i=0}^{k} r(X^m_{t_i})\Delta t} g(X^m_{t_k})\right], \tag{2.3}$$

is the unique solution of the following difference equation:

$$\begin{aligned}\frac{f_m(t_{k+1},x_0) - f_m(t_k,x_0)}{\Delta t} &= \frac{\mathbb{E}^{x_0}[f_m(t_k,X_{t_1})] - f_m(t_k,x_0)}{\Delta t} - \\ &- \frac{e^{r(x_0)\Delta t}-1}{\Delta t} f_m(t_{k+1},x_0) \\ f_m(0,x_0) &= g(x_0),\end{aligned} \tag{2.4}$$

where $t_{k+1} = t_k + \Delta t$.

Proof. (Existence)

Consider difference quotient from the definition of the generator of a discrete Itô diffusion and use $Z(t_k)$ notation:

$$\begin{aligned}\frac{\mathbb{E}^{x_0}[f_m(t_k,X_{t_1})] - f_m(t_k,x_0)}{\Delta t} &= \frac{1}{\Delta t}\mathbb{E}^{x_0}\{\mathbb{E}^{X_{t_1}}[Z(t_k)g(X^m_{t_k})] - \mathbb{E}^{x_0}[Z(t_k)g(X^m_{t_k})]\} \\ &= \frac{1}{\Delta t}\mathbb{E}^{x_0}\left\{\mathbb{E}^{x_0}\left[g(X^m_{t_{k+1}})e^{-\sum_{i=0}^{k} r(X^m_{t_{i+1}})\Delta t}|\mathcal{F}_{t_1}\right] - \right. \\ &\left. -Z(t_k)g(X^m_{t_k})\right\} \\ &= \frac{1}{\Delta t}\mathbb{E}^{x_0}\left[g(X^m_{t_{k+1}})Z(t_{k+1})e^{r(X^m_{t_0})\Delta t} - Z(t_k)g(X^m_{t_k})\right] \\ &= \frac{1}{\Delta t}\mathbb{E}^{x_0}[g(X^m_{t_{k+1}})Z(t_{k+1}) - g(X^m_{t_k})Z(t_k)] + \\ &+ \mathbb{E}^{x_0}\left[g(X^m_{t_{k+1}})Z(t_{k+1})\frac{e^{r(x_0)\Delta t}-1}{\Delta t}\right].\end{aligned}$$

Rearrange the terms to finish the proof of the existence part.

(Uniqueness)

Consider the following version of the difference equation (2.4):

$$e^{r(x_0)\Delta t} f_m(t_{k+1},x_0) = \mathbb{E}^{x_0}[f_m(t_k,X_{t_1})] \tag{2.5}$$

and suppose that there exists another solution $w(t_k,x_0)$ satisfying equation (2.5) with initial condition $w(0,x_0) = g(x_0)$. We are going to prove the claim by induction on k.

For the base case take $k=0$ and use equation (2.5) to deduce that

$$e^{r(x_0)\Delta t}w(t_1,x_0)=\mathbb{E}^{x_0}[w(0,X_{t_1})]=\mathbb{E}^{x_0}[g(X_{t_1})]=\mathbb{E}^{x_0}[f_m(0,X_{t_1})]=e^{r(x_0)\Delta t}f_m(t_1,x_0),$$

hence $w(t_1,x_0)=f_m(t_1,x_0)$. Now assume that $w(t_k,x_0)=f_m(t_k,x_0)$ holds for k let us check whether it is true for $k+1$.

$$e^{r(x_0)\Delta t}w(t_{k+1},x_0)=\mathbb{E}^{x_0}[w(t_k,X_{t_1})]=\mathbb{E}^{x_0}[f_m(t_k,X_{t_1})]=e^{r(x_0)\Delta t}f_m(t_{k+1},x_0)$$

So by induction it follows that $w(t_k,x_0)=f_m(t_k,x_0)$ for all k. $\square$

Recall that our aim is to approximate differential equation (2.1) or written equivalently

$$Af(t,x_0)=\frac{\partial f}{\partial t}(t,x_0)+r(x_0)f(t,x_0). \tag{2.6}$$

For that purpose let us first rewrite difference equation (2.4) in an equivalent form

$$\begin{aligned}\frac{\mathbb{E}^{x_0}[f_m(t_k,X_{t_1})]-f_m(t_k,x_0)}{\Delta t}&=\frac{f_m(t_{k+1},x_0)-f_m(t_k,x_0)}{\Delta t}+\\&+\frac{e^{r(x_0)\Delta t}-1}{\Delta t}f_m(t_{k+1},x_0).\end{aligned} \tag{2.7}$$

So we would like to take limit of the above equation as $m\to\infty$. Our strategy is to establish convergence of the right hand side of the equation (2.7) proceeding term by term.

Let's introduce the following notation:

$$\psi_m(t_k,x):=e^{-\sum_{i=0}^{k}r(X_{t_i}^m)\Delta t}g(X_{t_k}^m)$$
$$\psi(t,x):=e^{-\int_o^t r(X_s)\,\mathrm{d}s}g(X_t).$$

Theorem 4. *Assume that functions $r(\cdot)$, $g(\cdot):\mathbb{R}\to\mathbb{R}$ are from $C_0^2(\mathbb{R})$ also suppose that $a(\cdot)$, $b(\cdot):\mathbb{R}\to\mathbb{R}$ are Lipschitz continuous. Then as $m\to\infty$ we have uniform $\mathcal{L}^2$-convergence on $[0,K]\times\mathbb{R}$:*

$$\sup_{t\in[0,T]}[f_m(t_k,x)-f(t,x)]^2\to 0$$

Proof. By the fact that for any random variable X: $(\mathbb{E}[X])^2\le\mathbb{E}[X^2]$ we have

$$\begin{aligned}\sup_{t\in[0,T]}[f_m(t_k,x_0)-f(t,x_0)]^2&=\sup_{t\in[0,T]}\left[\mathbb{E}^{x_0}\left(\psi_m(t_k,x_0)-\psi(t,x_0)\right)\right]^2\\&\le\sup_{t\in[0,T]}\mathbb{E}^{x_0}\left[\psi_m(t_k,x_0)-\psi(t,x_0)\right]^2.\end{aligned}$$

Hence it is enough to show convergence of the right hand side of the above enequality. For that purpose we are going to use the following simple estimate:

$$\begin{aligned}(e^{-b}d-e^{-c}a)^2&=e^{-2b}(d-e^{-c+b}a)^2=e^{-2b}(d-a+a-e^{b-c}a)^2\\&\le 2\,e^{-2b}((d-a)^2+a^2(1-e^{b-c})^2).\end{aligned} \tag{2.8}$$

Then apply first order Taylor series expansion for function e^{b-c} around 0 to get:

$$e^{b-c}=1+e^s(b-c),$$

where $s\in[0,b-c]$. Substituting instead of function e^{b-c} in (2.8) its expansion one can get:

$$(e^{-b}d-e^{-c}a)^2\le e^{-2b}((d-a)^2+a^2e^{2s}(b-c)^2).$$

So using this estimate and definition of $\psi_m(t_k,x)$ and $\psi(t,x)$ one can get:

$$\sup_{t\in[0,T]} \mathbb{E}\left[\psi_m(t_k,x)-\psi(t,x)\right]^2 = \sup_{t\in[0,T]} \mathbb{E}\left[e^{-\sum_{i=0}^{k} r(X_{t_i}^m)\Delta t} g(X_{t_k}^m) - e^{-\int_0^t r(X_s)\,\mathrm{d}s} g(X_t)\right]^2$$
$$\leq \sup_{t\in[0,T]} \mathbb{E}\left[2e^{-2\sum_{i=0}^{k} r(X_{t_i}^m)\Delta t}\{(g(X_{t_k}^m)-g(X_t))^2+ \right.$$
$$\left. +g^2(X_t)\, e^{2s}\left(\sum_{i=0}^{k} r(X_{t_i}^m)\Delta t - \int_0^t r(X_s)\,\mathrm{d}s\right)^2\right\}\Bigg],$$

where $s\in[0,\sum_{i=0}^{k} r(X_{t_i}^m)\Delta t - \int_0^t r(X_s)\,\mathrm{d}s]$. By assumption $r(\cdot)\in C_0^2$ hence $|r(\cdot)|\leq R_0$, then

$$\sup_{t\in[0,T]} \mathbb{E}\left[\psi_m(t_k,x)-\psi(t,x)\right]^2 \leq 2e^{2\,T\,R_0} \sup_{t\in[0,T]} \mathbb{E}\left[(g(X_{t_k}^m)-g(X_t))^2+ \right.$$
$$\left. +g^2(X_t)\, e^{2s}\left(\sum_{i=0}^{k} r(X_{t_i}^m)\Delta t - \int_0^t r(X_s)\,\mathrm{d}s\right)^2\right]. \quad (2.9)$$

The next step is to take limit of both sides of the above inequality and we are going to do it term by term.

Let us start with the first term and apply first order Taylor series expansion of function $g(X_{t_k}^m)$ around X_t:

$$g(X_{t_k}^m)-g(X_t) = g(z)\,(X_{t_k}^m - X_t),$$

where $z\in[X_t, X_{t_k}^m]$. Then

$$\sup_{t\in[0,T]} \mathbb{E}[g(X_{t_k}^m)-g(X_t)]^2 = \sup_{t\in[0,T]} \mathbb{E}[g^2(z)\,(X_{t_k}^m - X_t)^2] \leq G_0^2 \sup_{t\in[0,T]} \mathbb{E}[X_{t_k}^m - X_t]^2,$$

since by assumption $g(\cdot)\in C_0^2$ hence $|g(\cdot)|\leq G_0$ for some positive constant G_0. Now take limit as $m\to\infty$ and use **Theorem 3** to conclude that:

$$\lim_{m\to\infty}\sup_{t\in[0,T]} \mathbb{E}[g(X_{t_k}^m)-g(X_t)]^2 = 0.$$

For the second term notice that $\sum_{i=0}^{k} r(X_{t_i}^m)\Delta t = \sum_{i=1}^{k}\int_{t_{i-1}}^{t_i} r(X_{t_i}^m)\,\mathrm{d}s$ and $\int_0^t r(X_s)\,\mathrm{d}s = \sum_{i=1}^{k}\int_{t_{i-1}}^{t_i} r(X_s)\,\mathrm{d}s$. Then

$$\sup_{t\in[0,T]} \mathbb{E}\left[\sum_{i=0}^{k} r(X_{t_i}^m)\Delta t - \int_0^t r(X_s)\,\mathrm{d}s\right]^2 = \sup_{t\in[0,T]} \mathbb{E}\left[\sum_{i=1}^{k}\int_{t_{i-1}}^{t_i} [r(X_{t_i}^m)-r(X_s)]\,\mathrm{d}s\right]^2$$
$$\leq T\sum_{i=1}^{k}\int_{t_{i-1}}^{t_i} [r(X_{t_i}^m)-r(X_s)]^2\,\mathrm{d}s,$$

where we have used simple estimate for the square of the sum and Cauchy-Schwartz inequality. As in the case of function $g(\cdot)$, using first order Taylor series expansion of $r(X_{t_k}^m)$ around X_t and taking limit of the above inequality as $m\to\infty$ one can easily derive that

$$\sup_{t\in[0,T]} \mathbb{E}\left[\sum_{i=0}^{k} r(X_{t_i}^m)\Delta t - \int_0^t r(X_s)\,\mathrm{d}s\right]^2 \to 0.$$

To finish the proof notice that in the inequality (2.9) as $m\to\infty$ $e^{2s}=0$ a.s. and in $\mathcal{L}^2$ and since by assumption $g(\cdot)\in C_0^2$ all the terms converge to 0. □

A straightforward corollary of the **Theorem 4** is that $\frac{e^{r(x_0)\Delta t}-1}{\Delta t} f_m(t_{k+1},x_0)$ converges uniformly in $\mathcal{L}^2$ to $r(x_0)f(t,x_0)$, which means convergence of the second term in the equation (2.7).

The next step is to show that the difference quotient $\frac{f_m(t_{k+1},x_0)-f_m(t_k,x_0)}{\Delta t}$ from the equation (2.7) converges to the time derivative of the continuous Feynman-Kac functional as $m \to \infty$. However in order to achieve that we need to develop a theory.

We would like to continue by proving discrete version of Itô formula. Despite the fact that this result will be used later on, it is very important on its own. The proof that we are about to present goes in line with the proof of Theorem 4.1.2 in [9, p. 44], more precisely we are going to consider the case of a discrete process X_{t_k} given by:

$$\Delta X_{t_k} = a(t_{k-1},w)\Delta t + b(t_{k-1},w)\Delta B^m_{t_k},$$

where functions $a(t_{k-1},\cdot)$, $b(t_{k-1},\cdot)$ are simple processes adopted to the filtration $\mathcal{F}_{t_k}$ generated by $B^m_{t_k}$. We are going to define a simple discrete process Y in a standard way used in many textbooks.

Definition. *We say that that Y is a simple process if there exists a sequence of times $0 < t_1 \leq t_2 \leq \dots$ increasing to ∞ a.s. and random variables $\xi_0, \xi_1, \dots$ such that ξ_j is $\mathcal{F}_{t_j}$-measuarble, $\mathbb{E}[\xi_j^2] < \infty$ for all j, and*

$$Y(t) = \xi_0\,\mathbb{I}_{\{0\}}(t) + \sum_{i=1}^{\infty} \xi_{i-1}\,\mathbb{I}_{(t_{i-1},t_i]}(t), \qquad (t \geq 0)$$

Generalization of the next theorem to the case of square-integrable adopted processes $a(\cdot)$, $b(\cdot)$ follows from Lemma 2 of [20, p. 215].

Theorem 5 (Discrete Itô formula). *Let X_{t_k} be a discrete Itô process with coefficients $a(\cdot,\cdot)$, $b(\cdot,\cdot)$ being simple processes. Let function $g \in C_0^2([0,\infty)\times\mathbb{R})$ then $g(t_k, X_{t_k})$ is again a discrete Itô process given by*

$$\Delta g(t_j, X_{t_j}) = \left(\frac{\partial g}{\partial t}(t_j, X_{t_j}) + \frac{\partial g}{\partial x}(t_{j+1}, X^m_{t_j})a(t_j,w) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t_{j+1}, X_{t_j})b^2(t_j,w)\right)\Delta t +$$
$$+\frac{\partial g}{\partial x}(t_{j+1}, X_{t_j})b(t_j,w)\Delta B^m_{t_{j+1}} + o(\Delta t).$$

Proof.

$$\begin{aligned}\Delta g(t_j, X_{t_j}) &= g(t_{j+1}, X_{t_{j+1}}) - g(t_j, X_{t_j})\\ &= g(t_{j+1}, X_{t_{j+1}}) - g(t_{j+1}, X_{t_j}) + g(t_{j+1}, X_{t_j}) - g(t_j, X_{t_j})\end{aligned}$$

Let us start by applying first order Taylor series expansion in time of $g(t_{j+1}, X_{t_j})$ around t_j using $o(\cdot)$ form of the remainder:

$$g(t_{j+1}, X_{t_j}) = g(t_j, X_{t_j}) + \frac{\partial g}{\partial t}(t_j, X_{t_j})\Delta t + o(\Delta t),$$

where $o(\Delta t) := \left[\frac{\partial g}{\partial t}(s_j, X_{t_j}) - \frac{\partial g}{\partial t}(t_j, X_{t_j})\right]\Delta t$ with $s_j \in [t_j, t_{j+1}]$ and since by assumption $g \in C_0^2$ it follows that $\lim_{\Delta t \to 0}\frac{o(\Delta t)}{\Delta t} = 0$.

Similarly, consider second order Taylor series expansion of $g(t_{j+1}, X_{t_{j+1}})$ in space around X_{t_j}:

$$g(t_{j+1}, X_{t_{j+1}}) = g(t_{j+1}, X_{t_j}) + \frac{\partial g}{\partial x}(t_{j+1}, X_{t_j})\Delta X_{t_j} + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t_{j+1}, X_{t_j})(\Delta X_{t_j})^2$$
$$+o(\Delta X_{t_j}^2),$$

where $o(\Delta X_{t_j}^2) := \frac{1}{2}\left[\frac{\partial^2 g}{\partial x^2}(t_{j+1}, S_{t_j}) - \frac{\partial^2 g}{\partial x^2}(t_{j+1}, X_{t_j})\right](\Delta X_{t_j})^2$ with $S_{t_j} \in [X_{t_j}, X_{t_{j+1}}]$. Using definition of the process ΔX_{t_j}:

$$\begin{aligned}
(\Delta X_{t_j})^2 &= (a(t_j, w)\Delta t + b(t_j, w)\Delta B_{t_{j+1}}^m)^2 \\
&= a^2(t_j, w)(\Delta t)^2 + b^2(t_j, w)(\Delta B_{t_{j+1}}^m)^2 + 2\, a(t_j, w)\, v(t_j, w)\Delta t\, \Delta B_{t_{j+1}}^m \\
&= a^2(t_j, w)(\Delta t)^2 + b^2(t_j, w)\Delta t + 2\, a(t_j, w)\, b(t_j, w)\Delta t\, \Delta B_{t_{j+1}}^m.
\end{aligned}$$

It is clear that $o(\Delta X_{t_j}^2) = o(\Delta t)$ since $|\Delta B_{t_{j+1}}^m| = \sqrt{\Delta t}$. Then

$$(\Delta X_{t_j})^2 = b^2(t_j, w)\Delta t + o(\Delta t).$$

Putting all together we have the claim:

$$\begin{aligned}
\Delta g(t_j, X_{t_j}) = \Bigg(&\frac{\partial g}{\partial t}(t_j, X_{t_j}) + \frac{\partial g}{\partial x}(t_{j+1}, X_{t_j}^m)a(t_j, w) + \\
&+ \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t_{j+1}, X_{t_j})b^2(t_j, w)\Bigg)\Delta t + \\
&+ \frac{\partial g}{\partial x}(t_{j+1}, X_{t_j})b(t_j, w)\Delta B_{t_{j+1}}^m + o(\Delta t).
\end{aligned}$$

□

As pointed in Oksendal [9, p. 46] we may extend **Theorem 5** to the case of $g(\cdot) \in C^2$ by approximating it with sequence of functions $g_n(\cdot) \in C_0^2$.

Use telescopic sum to see connection to the continuous Itô formula:

$$\begin{aligned}
g(t_k, X_{t_k}) &= g(0, x_0) + \sum_{i=0}^{k-1} \Delta g(t_j, X_{t_j}) \\
&= g(0, x_0) + \sum_{i=0}^{k-1}\Bigg(\frac{\partial g}{\partial t}(t_i, X_{t_i}) + \frac{\partial g}{\partial x}(t_{i+1}, X_{t_i}^m)a(t_i, w) + \\
&\quad + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t_{i+1}, X_{t_i})b^2(t_i, w)\Bigg)\Delta t + \sum_{i=0}^{k-1}\frac{\partial g}{\partial x}(t_{i+1}, X_{t_i})b(t_i, w)\Delta B_{t_{i+1}}^m + \\
&\quad + \sum_{i=0}^{k-1} o(\Delta t). \qquad (2.10)
\end{aligned}$$

Using techniques similar to **Theorem 4** upon taking limit of (2.10) as $\Delta t \to 0$ one has convergence to the continuous Itô lemma for simple processes $a(\cdot,\cdot)$, $b(\cdot,\cdot)$. In the sequel of the section we will return to the time homogeneous case of a discrete Itô diffusion with coefficients $a(X_{t_k})$ and $b(X_{t_k})$.

We would like to apply discrete Itô lemma to the discrete Feynman-Kac functional to have its representation in form of expectation of an Itô process. For that purpose recall definition of the discrete Feynman-Kac functional:

$$f_m(t_k, x) = \mathbb{E}^x\left[e^{-\sum_{i=0}^k r(X_{t_i}^m)\Delta t} g(X_{t_k}^m)\right].$$

For further analysis assume that $r(\cdot)$, $g(\cdot) \in C_0^2(\mathbb{R})$ and functions $a(\cdot)$, $b(\cdot)$ are bounded and Lipschitz continuous.

It is clear that $g(X_{t_k}^m)$ is a discrete Itô process since function $g(\cdot)$ satisfies requirements of **Theorem 5**, then it follows that

$$\Delta g(X_{t_j}^m) = \left(\frac{\partial g}{\partial x}(X_{t_j}^m)a(X_{t_j}^m) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(X_{t_j}^m)b^2(X_{t_j}^m)\right)\Delta t + \\ + \frac{\partial g}{\partial x}(X_{t_j}^m)b(X_{t_j}^m)\Delta B_{t_{j+1}}^m + o(\Delta t). \qquad (2.11)$$

Let us introduce a new notation $Z(t_j) := e^{-\sum_{i=0}^{j} r(X_{t_i}^m)(t_{i+1}-t_i)}$ and let's compute $\Delta Z(t_j)$:

$$\begin{aligned}\Delta Z(t_j) &= Z(t_{j+1}) - Z(t_j) = e^{-\sum_{i=0}^{j+1} r(X_{t_i}^m)\Delta t} - e^{-\sum_{i=0}^{j} r(X_{t_i}^m)\Delta t} \\ &= e^{-\sum_{i=0}^{j} r(X_{t_i}^m)\Delta t}\left(e^{-r(X_{t_{j+1}}^m)\Delta t} - 1\right) = e^{-\sum_{i=0}^{j} r(X_{t_i}^m)\Delta t}\left(-r(X_{t_{j+1}}^m)\Delta t + o(\Delta t)\right) \\ &= -Z(t_j)r(X_{t_{j+1}}^m)\Delta t + o(\Delta t), \qquad (2.12)\end{aligned}$$

where $o(\Delta t) := [-e^{-s}+1]\, r(X_{t_{j+1}}^m)\Delta t$ with $s \in [0, r(X_{t_{j+1}}^m)\Delta t]$.

Notice that inside of the Feynman-Kac functional we have the product $Z(t_k)g(X_{t_k}^m)$, so it would be advantageous to express it as an Itô diffusion as well:

$$\begin{aligned}\Delta[Z(t_j)g(X_{t_j}^m)] &= Z(t_{j+1})g(X_{t_{j+1}}^m) - Z(t_j)g(X_{t_j}^m) \\ &= Z(t_{j+1})\left(g(X_{t_{j+1}}^m) - g(X_{t_j}^m)\right) + g(X_{t_j}^m)\left(Z(t_{j+1}) - Z(t_j)\right) \\ &= Z(t_{j+1})\Delta g(X_{t_j}^m) + g(X_{t_j}^m)\Delta Z(t_j). \qquad (2.13)\end{aligned}$$

Now we are ready to present difference quotient $\frac{f_m(t_{k+1},x_0)-f_m(t_k,x_0)}{\Delta t}$ in a useful form.

Lemma 3. *Given that functions $r(\cdot), g(\cdot) \in C_0^2$ and $a(\cdot)$, $b(\cdot)$ are bounded and Lipschitz continuous*

$$\frac{f_m(t_{k+1},x_0) - f_m(t_k,x_0)}{\Delta t} = \mathbb{E}^{x_0}\left[-r(X_{t_{k+1}}^m)Z(t_k)g(X_{t_k}^m) + \frac{o(\Delta t)}{\Delta t} + \right. \\ \left. + Z(t_{k+1})\left\{\partial_x g(X_{t_k}^m)\, a(X_{t_k}^m) + \frac{1}{2}\partial_{xx} g(X_{t_k}^m)\, b^2(X_{t_k}^m)\right\}\right]. \qquad (2.14)$$

Proof. Using expressions for $\Delta Z(t_j)$, $\Delta g(X_{t_j}^m)$ given by (2.11), (2.12) and by taking sum in the equation (2.13) we could write the product $Z(t_k)g(X_{t_k}^m)$ as

$$\begin{aligned}Z(t_k)g(X_{t_k}^m) &= g(x_0) + \sum_{i=0}^{k-1} g(X_{t_i}^m)\Delta Z(t_i) + \sum_{i=0}^{k-1} Z(t_{i+1})\Delta g(X_{t_i}^m) \\ &= g(x_0) - \sum_{i=0}^{k-1} r(X_{t_{i+1}}^m)Z(t_i)g(X_{t_i}^m)\Delta t + \\ &\quad + \sum_{i=0}^{k-1} Z(t_{i+1})\left\{\partial_x g(X_{t_i}^m)\, a(X_{t_i}^m) + \frac{1}{2}\partial_{xx} g(X_{t_i}^m)\, b^2(X_{t_i}^m)\right\}\Delta t + \\ &\quad + \sum_{i=0}^{k-1} Z(t_{i+1})\, \partial_x g(X_{t_i}^m)\, b(X_{t_i}^m)\Delta B_{t_{i+1}}^m + \sum_{i=0}^{k-1} o(\Delta t).\end{aligned}$$

Next take expectation of both parts to get

$$\mathbb{E}^{x_0}\left[Z(t_k)g(X_{t_k}^m)\right] = g(x_0) + \mathbb{E}^{x_0}\left[-\sum_{i=0}^{k-1} r(X_{t_{i+1}}^m)Z(t_i)g(X_{t_{i+1}}^m)\Delta t + \sum_{i=0}^{k-1} o(\Delta t) \right. \\ \left. + \sum_{i=0}^{k-1} Z(t_{i+1})\left\{\partial_x g(X_{t_i}^m)\, a(X_{t_i}^m) + \frac{1}{2}\partial_{xx} g(X_{t_i}^m)\, b^2(X_{t_i}^m)\right\}\Delta t\right].$$

Similarly one can compute expression for $\mathbb{E}^{x_0}[Z(t_{k+1})g(X_{t_{k+1}}^m)]$ and claim follows easily after subtraction. □

For the continuous case existence of the time derivative of the Feynman-Kac functional could be shown by applying logic similar to the one in the proof of Lemma 7.3.2 [9, p. 118]. Moreover, an exact form of the expression could be also inferred (using notations from the Oksendal's book):

$$\frac{\partial}{\partial t} f(t, x_0) = \mathbb{E}^{x_0}\left[-r(X_t)g(X_t)Z_t + \right.$$
$$\left. + Z_t\left(\partial_x g(X_t)\, a(X_t) + \frac{1}{2}\partial_{xx} g(X_t)\, b^2(X_t)\right)\right]. \tag{2.15}$$

In the next theorem we will prove convergence of (2.14) to (2.15) as $m \to \infty$ which is the last step on the way of establishing convergence of a difference heat equation to the continuous case.

Theorem 6. *Suppose that functions $r(\cdot), g(\cdot) \in C_0^2$ and $a(\cdot)$, $b(\cdot)$ are Lipschitz continuous then as $m \to \infty$*

$$\sup_{t\in[0,T]} \left[\frac{f_m(t_{k+1}, x_0) - f_m(t_k, x_0)}{\Delta t} - \frac{\partial}{\partial t} f(t, x_0)\right]^2 \to 0.$$

Proof. Use the fact that $(\mathbb{E}[X])^2 \le \mathbb{E}[X]^2$ for any random variable X. Then

$$\sup_{t\in[0,T]} \left[\frac{f_m(t_{k+1}, x_0) - f_m(t_k, x_0)}{\Delta t} - \frac{\partial}{\partial t} f(t, x_0)\right]^2 \le$$
$$\le \sup_{t\in[0,T]} \mathbb{E}\left[-r(X^m_{t_{k+1}})Z(t_k)g(X^m_{t_k}) + r(X_t)g(X_t)Z_t + \right.$$
$$+ Z(t_{k+1})\,\partial_x g(X^m_{t_k})\, a(X^m_{t_k}) - Z_t\,\partial_x g(X_t)\, a(X_t) +$$
$$\left. + \frac{1}{2} Z(t_{k+1})\,\partial_{xx} g(X^m_{t_k})\, b^2(X^m_{t_k}) - \frac{1}{2} Z_t\,\partial_{xx} g(X_t)\, b^2(X_t) + \frac{o(\Delta t)}{\Delta t}\right]^2.$$

Using simple estimate $(a+b+c+d)^2 \le 4(a^2+b^2+c^2+d^2)$ we have

$$\sup_{t\in[0,T]} \mathbb{E}\left[\frac{f_m(t_{k+1}, x_0) - f_m(t_k, x_0)}{\Delta t} - \frac{\partial}{\partial t} f(t, x_0)\right]^2 \le$$
$$\le 4 \sup_{t\in[0,T]} \mathbb{E}\left[r(X_t)g(X_t)Z_t - r(X^m_{t_{k+1}})Z(t_k)g(X^m_{t_k})\right]^2 +$$
$$+ 4 \sup_{t\in[0,T]} \mathbb{E}\left[Z(t_{k+1})\,\partial_x g(X^m_{t_k})\, a(X^m_{t_k}) - Z_t\,\partial_x g(X_t)\, a(X_t)\right]^2 +$$
$$+ \frac{4}{2} \sup_{t\in[0,T]} \mathbb{E}\left[Z(t_{k+1})\,\partial_{xx} g(X^m_{t_k})\, b^2(X^m_{t_k}) - Z_t\,\partial_{xx} g(X_t)\, b^2(X_t)\right]^2 +$$
$$+ 4 \sup_{t\in[0,T]} \mathbb{E}\left[\frac{o(\Delta t)}{\Delta t}\right]^2$$
$$\le 4\left(I_1 + I_2 + \frac{1}{2} I_3 + \sup_{t\in[0,T]} \mathbb{E}\left[\frac{o(\Delta t)}{\Delta t}\right]^2\right), \tag{2.16}$$

where I_1, I_2, I_3 will be defined below. Let us start with I_1.

$$\begin{aligned} I_1 &:= \sup_{t\in[0,T]} \mathbb{E}\left[r(X_t)g(X_t)Z_t - r(X^m_{t_{k+1}})Z(t_k)g(X^m_{t_k})\right]^2 \\ &= \sup_{t\in[0,T]} \mathbb{E}\left[r(X_t)g(X_t)Z_t - r(X_t)Z(t_k)g(X^m_{t_k}) + \right. \\ &\qquad \left. + r(X_t)Z(t_k)g(X^m_{t_k}) - r(X^m_{t_{k+1}})Z(t_k)g(X^m_{t_k})\right]^2 \\ &\le 2 \sup_{t\in[0,T]} \mathbb{E}\left[r^2(X_t)(g(X_t)Z_t - g(X^m_{t_k})Z(t_k))^2\right] + \\ &\qquad + 2 \sup_{t\in[0,T]} \mathbb{E}\left[Z^2(t_k)g^2(X^m_{t_k})(r(X_t) - r(X^m_{t_{k+1}}))^2\right]. \end{aligned}$$

Since by assumption $|r(\cdot)| \le R_0$ and $|g(\cdot)| \le G_0$ for some positive constants $R_0, G_0 > 0$:

$$\begin{aligned} I_1 &\le 2\, R_0^2 \sup_{t\in[0,T]} \mathbb{E}\left[g(X_t)Z_t - g(X^m_{t_k})Z(t_k)\right]^2 + \\ &\qquad + 2\, G_0^2\, e^{2\,TR_0} \sup_{t\in[0,T]} \mathbb{E}\left[r(X_t) - r(X^m_{t_{k+1}})\right]^2. \end{aligned}$$

So taking limit as $m \to \infty$ by **Theorem 4** we have

$$\lim_{m\to\infty} I_1 = 0.$$

Use similar technique for the second term:

$$\begin{aligned} I_2 &:= \sup_{t\in[0,T]} \mathbb{E}\left[Z(t_{k+1})\,\partial_x g(X^m_{t_k})\,a(X^m_{t_k}) - Z_t\,\partial_x g(X_t)\,a(X_t)\right]^2 \\ &= \sup_{t\in[0,T]} \mathbb{E}\left[Z(t_{k+1})\,\partial_x g(X^m_{t_k})\,a(X^m_{t_k}) - Z(t_{k+1})\,\partial_x g(X^m_{t_k})\,a(X_t) + \right. \\ &\qquad \left. + Z(t_{k+1})\,\partial_x g(X^m_{t_k})\,a(X_t) - Z_t\,\partial_x g(X_t)\,a(X_t)\right]^2 \\ &\le 2 \sup_{t\in[0,T]} \mathbb{E}\left[Z^2(t_{k+1})\,(\partial_x g(X^m_{t_k}))^2(a(X^m_{t_k}) - a(X_t))^2\right] + \\ &\qquad + 2 \sup_{t\in[0,T]} \mathbb{E}\left[a^2(X_t)(Z(t_{k+1})\,\partial_x g(X^m_{t_k}) - Z_t\,\partial_x g(X_t))^2\right]. \end{aligned}$$

Using linear growth of $a(\cdot)$ and assumption $|\partial_x g(\cdot)| \le G_1$ for some positive constant G_1 and for m large enough, we get

$$\begin{aligned} I_2 &\le 2\,K^2\,G_1^2\,e^{2\,TR_0} \sup_{t\in[0,T]} \mathbb{E}\left[X^m_{t_k} - X_t\right]^2 + \\ &\qquad + 4C^2 \sup_{t\in[0,T]} \mathbb{E}\left[\left(1 + X_t^2\right)\left(Z(t_{k+1})\,\partial_x g(X^m_{t_k}) - Z_t\,\partial_x g(X_t)\right)^2\right] \\ &\le 2\,K^2\,G_1^2\,e^{2\,TR_0} \sup_{t\in[0,T]} \mathbb{E}\left[X^m_{t_k} - X_t\right]^2 + \\ &\qquad + 4(1 + K_1^2)\,C^2 \sup_{t\in[0,T]} \mathbb{E}\left[Z(t_{k+1})\,\partial_x g(X^m_{t_k}) - Z_t\,\partial_x g(X_t)\right]^2, \end{aligned}$$

where we have used assumption that function $\partial_x g(X_t)$ has compact support, hence we may conclude that when X_t is bounded by K_1 when it belongs to the support of $\partial_x g(X_t)$ and for otherwise $\partial_x g(X_t)$ and $\partial_x g(X^m_t)$ will be equal to 0 upon taking the limit as $m \to \infty$. Notice that by **Theorem 3** the first term on the right hand side of the inequality goes to 0 as $m \to \infty$ and for the second apply technique similar to the one in **Theorem 4** to see $\mathcal{L}^2$convergence of $Z_{t_k}\partial_x g(X^m_{t_k})$ to $Z_t\,\partial_x g(X_t)$ and so

$$\lim_{m\to\infty} I_2 = 0.$$

For the third term

$$I_3 := \frac{1}{2} \sup_{t\in[0,T]} \mathbb{E}\left[Z(t_{k+1})\,\partial_{xx}g(X^m_{t_k})\,b^2(X^m_{t_k}) - Z_t\,\partial_{xx}g(X_t)\,b^2(X_t)\right]^2$$

apply similar logic and conclude that

$$\lim_{m\to\infty} I_3 = 0.$$

Thus upon taking limit of the inequality (2.16) as $m \to \infty$ the proof is completed. $\square$

We finish current section by presenting convergence result which proves that continuous Feynman-Kac functional solves continuous-time heat equation and might be approximated in $\mathcal{L}^2$ by (2.3).

Theorem 7 (Convergence of a real-valued Schrödinger equation). *Assume that functions $r(\cdot), g(\cdot) \in C_0^2$ and $a(\cdot)$, $b(\cdot)$ are bounded and Lipschitz continuous then as $m \to \infty$*

$$\sup_{t\in[0,T]} \left[\frac{\mathbb{E}^{x_0}[f_m(t_k, X_{t_1}) - f_m(t_k, x_0)]}{\Delta t} - A\,f(t, x_0)\right]^2 \to 0$$

Proof. Use equations (2.13), (2.7) and apply **Theorems 4** and **6**. $\square$

2.2 Application to the Black-Scholes model

In the previous sections we were dealing with so called forward case, its name stems from the fact that the initial condition of the system was given at time $t = 0$. However in some applications of the Feynman-Kac formula (particularly in finance) we have terminal condition instead, which gives value of the functional at time $t = T$. That is why the former is referred in the literature as a backward case. To put it simply, in the first case we know the starting point of the system, whereas in the second case we know the end state only.

Since we will be working with an option pricing model from financial mathematics, we will face the backward case. So it is necessary to define backward version of the discrete Feynman-Kac functional:

$$f^b_m(t_k, x) = \mathbb{E}\left[e^{-\sum_{i=k}^N r(X^m_{t_i})\Delta t} g(X^m_{t_N})\,\middle|X^m_{t_k} = x\right],$$

where $N := \lfloor T\,2^{2m} \rfloor$ and let $Z^b(t_k) := e^{-\sum_{i=k}^N r(X^m_{t_i})\Delta t}$. To proceed further we need version of **Lemma 2** with boundary condition.

Lemma 4. *Given time-homogeneous discrete Itô diffusion $X^m_{t_k}$ defined above, backward discrete Feynman-Kac functional $f^b_m(t_k, x)$ is the unique solution of the following difference equation:*

$$\frac{f^b_m(t_{k+1}, x) - f^b_m(t_k, x)}{\Delta t} = \frac{\mathbb{E}[f^b_m(t_k, X_{t_{k+1}})|X^m_{t_k} = x] - f^b_m(t_k, x)}{\Delta t} -$$
$$-\frac{e^{-r(x)\Delta t} - 1}{\Delta t} f^b_m(t_{k+1}, x) \tag{2.17}$$
$$f^b_m(T, x) = g(x),$$

where $t_{k+1} = t_k + \Delta t$.

Proof. (Existence)

Consider difference quotient from the definition of the generator of a discrete Itô diffusion and use $Z^b(t_k)$ notation:

$$
\begin{aligned}
&\frac{\mathbb{E}[f_m^b(t_k, X_{t_{k+1}})|X_{t_k}^m = x] - f_m^b(t_k, x)}{\Delta t} \\
&= \frac{1}{\Delta t}\mathbb{E}\left[\mathbb{E}[Z^b(t_k)g(X_{t_N}^m)|X_{t_k}^m = X_{t_{k+1}}] - Z^b(t_k)g(X_{t_N}^m)\,\middle|X_{t_k}^m = x\right] \\
&= \frac{1}{\Delta t}\mathbb{E}\left[\mathbb{E}[e^{-r(X_{t_k}^m)}Z^b(t_{k+1})g(X_{t_N}^m)|\mathcal{F}_{t_{k+1}}] - Z^b(t_k)g(X_{t_N}^m)\,\middle|X_{t_k}^m = x\right] \\
&= \frac{1}{\Delta t}\mathbb{E}\left[e^{-r(X_{t_k}^m)}Z^b(t_{k+1})g(X_{t_N}^m) - Z^b(t_k)g(X_{t_N}^m)\,\middle|X_{t_k}^m = x\right] \\
&= \frac{1}{\Delta t}\mathbb{E}\left[Z^b(t_{k+1})g(X_{t_N}^m) - Z^b(t_k)g(X_{t_N}^m)\,\middle|X_{t_k}^m = x\right] + \\
&\quad + \mathbb{E}\left[\frac{e^{-r(X_{t_k}^m)\Delta t} - 1}{\Delta t}Z^b(t_{k+1})g(X_{t_N}^m)\,\middle|X_{t_k}^m = x\right].
\end{aligned}
$$

Rearrange the terms to finish the proof of the existence part.

(Uniqueness)

To prove the uniqueness part consider the following version of the difference equation (2.17)

$$
e^{-r(x)\Delta t} f_m^b(t_{k+1}, x) = \mathbb{E}\left[f_m^b(t_k, X_{t_{k+1}})\,\middle|X_{t_k}^m = x\right],
$$

and apply induction starting from the boundary condition at time T and proceeding with step Δt backward in time. □

Next, following similar lines as in the previous chapter, one can show $\mathcal{L}^2$-convergence as $m \to \infty$ of two terms:

$$
f_m^b(t_k, x) \to f^b(t, x), \qquad \textit{where} \qquad f^b(t, x) := \mathbb{E}\left[e^{-\int_t^T r(X_s)\,\mathrm{d}s} g(X_T)\,|X_t = x\right]
$$

$$
\frac{f_m^b(t_{k+1}, x) - f_m^b(t_k, x)}{\Delta t} \to \frac{\partial f^b}{\partial t}(t, x),
$$

hence we conclude that difference equation (2.17) is a discretization of the following differential equation:

$$
\frac{\partial f^b}{\partial t}(t, x) = A\, f^b(t, x) + r(x)\, f^b(t, x),
$$

with boundary condition $f^b(T, x) = g(x)$.

Now we will provide connection of the above equation to the option pricing theory. For simplicity assume that there are two assets trading on the financial market: risky (equity) and risk-less (bond). Let function $r(\cdot)$ be constant, it represents risk-free interest rate in the economy, to be precise let it be return of a zero-coupon default-free bond.

Let S_t be a price process of a risky asset given by

$$
dS_t = a(S_t)\, dt + b(S_t)\, dW_t,
$$

where $a(\cdot)$, $b(\cdot)$ are functions described in the previous sections and W_t is a Brownian motion, we will call $a(\cdot)$ and $b(\cdot)$ drift and diffusion coefficients respectively.

Dynamics of a risk-free bond, let's call it β_t, is entirely deterministic and follows simple differential equation:

$$
d\beta_t = r\, \beta_t dt,
$$

hence $\beta_t = e^{rt}$.

Our task is to find a fair price of the financial instrument known as option issued on an equity, which value at time t depends solely on the price of an underlying asset S_t. In the simplest case the only payment to the holder of an option occurs at a maturity date T and let us denote it as $g(S_T)$. Our task is to determine the price of an option at any time $t \leq T$.

In the spirit of the original paper written by Fischer Black and Myron Scholes [2] we assume that payoff of an option may be replicated by holding necessary amount of stocks $\alpha_1(t)$ and bonds $\alpha_2(t)$ and let's call such a portfolio $V(t)$. Then value of an option at time t denoted by $f(t, S_t)$ should be equal to the value of the replicating portfolio $V(t)$, such condition will prevent an arbitrage opportunity.

We have to assume self-financing condition for a portfolio $V(t)$ which means that until time T no funds were taken out from the portfolio, then value of the portfolio satisfies

$$V(t) = \alpha_1(t) S_t + \alpha_2(t) \beta_t. \tag{2.18}$$

The key step in the process is so called change of measure, reader may find comprehensive theory behind it in [9, p. 153-160], the legitimacy of the procedure essentially guaranteed by the Girsanov second theorem [9, p. 157]. Assume that the Novikov's condition is satisfied

$$\mathbb{E}\left[e^{\frac{1}{2}\int_0^T \left(\frac{a(S_u) - r\, S_u}{b(S_u)}\right)^2 \mathrm{d}u} \right] < \infty,$$

then the new probability measure $\mathbb{Q}$ is defined via Radon-Nykodim derivative as:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = e^{-\int_0^t \frac{a(S_u) - r\, S_u}{b(S_u)} \mathrm{d}W_u - \frac{1}{2}\int_0^t \left(\frac{a(S_u) - r\, S_u}{b(S_u)}\right)^2 \mathrm{d}u}.$$

Under the new probability measure $\tilde{W}_t := \int_0^t \frac{a(S_u) - r\, S_u}{b(S_u)} \mathrm{d}u + W_t$ is a $\mathbb{Q}$-Brownian motion and the price process S_t has the stochastic representation:

$$dS_t = r S_t dt + b(S_t) d\tilde{W}_t. \tag{2.19}$$

One can solve SDE (2.19) by applying general Itô lemma to the function $ln(S_t)$ to get:

$$S_t = S_0\, e^{\int_0^t \left(r - \frac{1}{2}\frac{b^2(S_u)}{S_u^2}\right) \mathrm{d}u + \int_0^t \frac{b(S_u)}{S_u} \mathrm{d}W_u},$$

and hence discounted price process has the following form:

$$\frac{S_t}{\beta_t} = S_0\, e^{\int_0^t \frac{b(S_u)}{S_u} \mathrm{d}W_u - \frac{1}{2}\int_0^t \frac{b^2(S_u)}{S_u^2} \mathrm{d}u}.$$

Notice that if function $b(\cdot)$ satisfies Novikov's condition:

$$\mathbb{E}_{\mathbb{Q}}\left[e^{\frac{1}{2}\int_0^T \frac{b^2(S_u)}{S_u^2} \mathrm{d}u} \right] < \infty,$$

then discounted price process $S(t)/\beta_t$ is a martingale and hence from equation (2.18) it follows that discounted replicating portfolio $V(t)/\beta_t$ is also a martingale. Since there is no arbitrage price of an option is a martingale and we know that at maturity it pays $g(S(T))$ or in discounted form $e^{-r(T-t)} g(S(T))$ then by martingale property it follows that:

$$f(t, S_t) = \mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)} g(S(T))\, | S_t = x \right]. \tag{2.20}$$

Notice that (2.20) defines backward Feynman-Kac functional and hence is a solution of the following differential equation:

$$\frac{\partial f}{\partial t}(t, x) = A\, f(t, x) + r(x)\, f(t, x),$$

or provided that payout function $g(S(T)) \in C_0^2$ we replace generator A with an exact expression:

$$\partial_t f(t,x) = rx\,\partial_x f(t,x) + \frac{1}{2}b^2(x)\partial_{xx} f(t,x) + r\,f(t,x),$$

which is known as Black-Scholes equation.

It is important to know an exact number of equities and bonds in the replicating portfolio at any time $t \leq T$ in order to guarantee that the price of an option given in equation (2.20) will be realized. So we may apply generalized Itô lemma for $f(t, S_t)$ to get:

$$df(t,S_t) = \partial_t f(t,X_t)dt + \partial_s f(t,S_t)dS_t + \frac{1}{2}\partial_{ss} f(t,S_t)(dS_t)^2,$$

substitute (2.19) instead of dS_t to get

$$df(t,S_t) = \left(\partial_t f(t,X_t) + rS_t\,\partial_s f(t,S_t) + \frac{1}{2}b^2(S_t)\partial_{ss} f(t,S_t)\right)dt + b(S_t)\partial_s f(t,S_t)d\tilde{W}_t.$$

On the other hand, using equation (2.18) one can derive

$$\begin{aligned} df(t,S_t) &= \alpha_1(t)dS_t + \alpha_2(t)d\beta_t \\ &= (\alpha_1(t)r\,S_t + \alpha_2(t)r\,\beta_t)\,dt + \alpha_1(t)b(S_t)d\tilde{W}_t. \end{aligned}$$

In order for two expressions for the dynamics of an option price to be equal we must have:

$$\begin{aligned} \alpha_1(t) &= \partial_s f(t,S_t) \\ \alpha_2(t) &= \frac{1}{r\,\beta_t}\left(\partial_t f(t,S_t) + \frac{1}{2}b^2(S_t)\partial_{ss} f(t,S_t)\right). \end{aligned}$$

Conclusion

Using "twist & shrink" construction of a Brownian motion we have provided several discrete approximation schemes for a time-homogeneous Itô diffusion. Assuming that its coefficients satisfy global Lipschitz continuity condition we have established $\mathcal{L}^2$-convergence to the continuous case.

Discrete Itô diffusion with deterministic partition of the time scale then was used to present a discretization of the Feynman-Kac functional. Assuming that functions $r(\cdot)$ and $g(\cdot)$ are twice continuously differentiable with compact support we have proved $\mathcal{L}^2$-convergence of the discrete Feynman-Kac functional and its time derivative to the continuous analogs. Consequently existence of the solution for the original heat equation followed. Special treatment of uniqueness part may be found in [9, p. 137].

In the last section we have given an alternative prove of the Black-Scholes option pricing formula for a case of diffusion price process. A natural continuation of which may be a theory similar to the discrete model of Cox-Ross-Rubenstein [5].

Bibliography

[1] Albeverio S. (1997) *Wiener and Feynman path integral and their applications.* Proc. Symp. Appl. Math. 52, 163-194

[2] Black F., Scholes M. (1973) *The pricing of options and corporate liabilities.* J. Polit. Econ. 81, 637-654

[3] Cameron R. (1960) *A family of integrals serving to connect the Wiener and Feynman integrals.* J. Math. Phys. 39, 126-140.

[4] Cameron R., Storvick D. (1983) *A simple definition of the Feynman integrals with applications.* Mem. A.M.S., Providence, R. I.

[5] Cox J, Ross S., Rubinstein M. (1979) *Option pricing: A simplified approach.* J. Financ. Econ. 7, 229-263

[6] Csáki E. (1993) *A discrete Feynman-Kac formula.* J. Statist. Plann. Inference 34, 63-73

[7] Feynman R. (1948) *Space-Time approach to Non-Relativistic Quantum Mechanics.* Rev. of Mod. Phys. 20, 367-387.

[8] Itô K. (1960) *Wiener integral and Feynman integral.* Proc. Fourth Berkley Symp. on Math. Stat. Probab. II, 228-238

[9] Øksendal B. (2007) *Stochastic Differential Equations.* Sixth Edition, Springer, Berlin.

[10] Kac M. (1949) *On distribution of certain Wiener functionals.* Trans. A.M.S. 65, 1-13.

[11] Kac M. (1951) *On some connections between probability theory and differential and integral equations.* Proc. Second Berkeley Symp. on Math. Stat. Prob., 189-215

[12] Kloeden P., Platen E. (1999) *Numerical Solution of Stochastic Differential Equations.* Springer, Berlin.

[13] Knight F. (1962) *On the random walk and Brownian motion.* Trans. Amer. Math. Soc. 103, 218-228.

[14] Kuo H. (2006) *Introduction to stochastic integration.* Springer.

[15] Nelson E. (1964) *Feynman integrals and the Schrödinger equation.* J. Math. Phys. 5, 332-343

[16] Revesz P. (1990) *Random walk in random and non-random environments.* World Scientific, Singapore.

[17] Schilling R., Partzsch L. (2012) *Brownian Motion* Walter de Gruyter GmbH & Co. KG, Berlin/Boston

[18] Szabados T. (1996) *An Elementary Introduction to the Wiener Process and stochastic integrals.* Studia Sci. Math. Hung. 31, 249-297.

[19] Szabados T., Székely B. (2005) *An elementary approach to brownian local time based on simple, symmetric random walks.* Periodica Mathematica Hungarica Vol. 51 (1), pp. 79-98

[20] Szabados T., Székely B. (2008) *Stochastic integration based on simple, symmetric random walks.* J Theor Probab (2009) 22: 203-219

[21] Szabados T. (2012) *Pathwise approximation of Feynman path integrals using simple random walks.*

[22] Szabados T., Zsolt N. (2014) *Strong approximation of Black-Scholes theory based on simple random walks.* 2010 MSC.