

A path integral approach to Bayesian inference in Markov processes

Toshiyuki Fujii

*Graduate School of Integrated Arts and Sciences,
Hiroshima University.
Department of Physics, Asahikawa Medical University,
Midorigaoka-higashi, Asahikawa, 078-8510, Japan.*

Noriyuki Hatakenaka

*Graduate School of Integrated Arts and Sciences,
Hiroshima University, Higashi-Hiroshima 739-8521, Japan.*

Abstract

We formulate Bayesian updates in Markov processes by means of path integral techniques and derive the imaginary-time Schrödinger equation with likelihood to direct the inference incorporated as a potential for the posterior probability distribution.

AMS subject classification:

Keywords:

1. Introduction

Statistical inference is a fundamental technology indispensable in modern society. In particular, Bayesian inference is explosive growth in a wide range of all kind of human activities, including philosophy, medicine, sport, law as well as science and technology [9, 7]. This uses a procedure to update the probability sequentially based on the Bayes' theorem that incorporates new evidence into prior probabilities to form updated probability estimates. The form of this theorem is

$$P(H|E) = \frac{P(E|H)}{P(E)} P(H), \quad (1.1)$$

where $P(E|H)$ is the probability for the evidence E given a hypothesis H , $P(H)$ is the prior probability that the hypothesis is true, $P(H|E)$ is the posterior probability that the hypothesis is true given the evidence, and $P(E) = \sum_H P(E|H)P(H)$ is the probability for the evidence averaged over all hypotheses. In this way, the Bayes' theorem relates the prior probability of the hypothesis before getting the evidence $P(H)$ to the posterior probability of the hypothesis after getting the evidence $P(H|E)$ through the likelihood ratio $P(E|H)/P(E)$.

As is well known, the time evolution of the conditional probability (the transition probability) of the Markov process follows the Chapman-Kolmogorov equation [3, 6];

$$p(x_{i+1}, t_{i+1} | x_{i-1}, t_{i-1}) = \int p(x_{i+1}, t_{i+1} | x_i, t_i) p(x_i, t_i | x_{i-1}, t_{i-1}) dx_i, \quad (1.2)$$

where $p(x_i, t_i | x_j, t_j)$ is the conditional probability density function defined as

$$p(x_i, t_i | x_j, t_j) dx_i dx_j \equiv P(x_i < X_i \leq x_i + dx_i | x_j < X_j \leq x_j + dx_j)$$

with x_j and X_j being the realization and the random variable at a certain time t_j . This expresses the chain condition of the Markov process that the state in the near future depends only on the state at the present time, not on past memories. This is consistent with the idea of infinitesimal time evolution in path integral techniques of quantum mechanics. In fact, Onsager and Machlup later incorporated Markov chain conditions into path integral techniques [10]. Along this line, this paper extends the Onsager-Machlup theory by incorporating the likelihood required by Bayesian inference into the Markov chain conditions.

2. A path integral representation of Bayesian updates

Suppose the stochastic processes under the probability distribution density functions $p_i(q_i)$ and $p_i(x_i)$ that are defined as $p_i(q_i) dq_i \equiv P(q_i < Q_i \leq q_i + dq_i)$ and $p_i(x_i) dx_i \equiv P(x_i < X_i \leq x_i + dx_i)$, respectively. Here the continuous random variables Q_i and X_i act as the roles of H and E in the Bayes' theorem in the i -th update. Starting with the law of total probability

$$p_{i+1}(q_{i+1}) = \int dq_i p(q_{i+1} | q_i) p_i(q_i), \quad (2.1)$$

the first Bayesian update is expressed as

$$p_1(q_1 | x_1) = \int dq_0 \frac{p(x_1 | q_1)}{p(x_1)} p(q_1 | q_0) p_0(q_0). \quad (2.2)$$

After repeating Bayesian updates of N times, the posterior probability distribution density function is given as

$$p_N(q_N | x_1, x_2, \dots, x_N) = \int dq_0 K(q_N, q_0; x_1, x_2, \dots, x_N) p_0(q_0), \quad (2.3)$$

where the kernel $K(q_N, q_0; x_1, x_2, \dots, x_N)$ is given as

$$\begin{aligned}
 & K(q_N, q_0; x_1, x_2, \dots, x_N) \\
 &= \int dq_1 \int dq_2 \cdots \int dq_{N-1} \prod_{i=1}^N \frac{p(x_i | q_i)}{p(x_i)} p(q_i | q_{i-1}). \tag{2.4}
 \end{aligned}$$

This describes repeated Bayesian updates representing the time evolution of stochastic processes starting from the prior probability distribution density function $p_0(q_0)$ inferred from the repeated conditional probabilities. Hereafter, we simply represent $p_i(q_i | x_1, x_2, \dots, x_{i-1}) = p_i(q_i | \mathbf{x})$ where \mathbf{x} denotes the set of the realizations.

Here, we introduce the parameter $t_i = i\epsilon$ with a small quantity ϵ and an integer i , in order to prepare for converting discrete time to continuous time such as $q_i = q(t_i)$. In this notation, for example, the probability distribution density function and the transition probability distribution density function are expressed as $p_i(q_i = q | \mathbf{x}) = p(q, t_i | \mathbf{x})$ and $p(q_j = q' | q_i = q) = p(q', t_j | q, t_i)$, respectively. Let us consider the infinitesimal time evolution of $p(q', t_j | q, t_i)$ from time t_i to $t_j = t_{i+1}$. Since $p(q', t_{i+1} | q, t_i) = p(q', t_i + \epsilon | q, t_i)$, then

$$p(q', t_i + \epsilon | q, t_i) = p(q', t_i | q, t_i) + \epsilon \left. \frac{\partial p(q', t'_i | q, t_i)}{\partial t'_i} \right|_{t'_i=t_i} + O(\epsilon^2). \tag{2.5}$$

Suppose that the stochastic process that we are assuming is a Wiener process. If the stochastic process is a Markov process, the random variable q obeys the Ito-type stochastic partial differential equation and the probability distribution density function based on this random variable follows the Fokker-Planck (FP) equation. Furthermore, the FP equation results in a diffusion equation in the case of a Wiener process leading to Brownian motion which has no driving force. Therefore, the second term on the right hand side of Eq. (2.5) is replaced by the diffusion equation;

$$\left. \frac{\partial p(q', t' | q, t)}{\partial t'} \right|_{t'=t_i} = \frac{1}{2} D \frac{\partial^2 p(q', t' | q, t)}{\partial q'^2}, \tag{2.6}$$

where D is the diffusion constant. The resulting infinitesimal time-evolution equation is

$$p(q', t_i + \epsilon | q, t_i) = \left(1 + \frac{\epsilon}{2} D \frac{\partial^2}{\partial q'^2} \right) p(q', t'_i | q, t_i) \Big|_{t'_i=t_i} + O(\epsilon^2). \tag{2.7}$$

By using the relation $p(q', t'_i | q, t_i) |_{t'_i=t_i} = \delta(q' - q)$, this is reduced to

$$\begin{aligned}
 p(q', t_i + \epsilon | q, t_i) &= \left(1 + \frac{\epsilon}{2} D \frac{\partial^2}{\partial q'^2}\right) \delta(q' - q) + O(\epsilon^2) \\
 &= \left(1 + \frac{\epsilon}{2} D \frac{\partial^2}{\partial q'^2}\right) \left(\int \frac{ds}{2\pi} e^{i(q'-q)s}\right) + O(\epsilon^2) \\
 &= \int \frac{ds}{2\pi} \left(1 + \frac{\epsilon}{2} D(-s^2)\right) e^{i(q'-q)s} + O(\epsilon^2) \\
 &= \int \frac{ds}{2\pi} e^{-\frac{\epsilon}{2} D s^2} e^{i(q'-q)s} + O(\epsilon^2) \\
 &= \frac{1}{\sqrt{2\pi D \epsilon}} e^{-\frac{(q'-q)^2}{2D\epsilon}} + O(\epsilon^2). \tag{2.8}
 \end{aligned}$$

Here we define the velocity of the random variable by $\dot{q} = (q' - q)/\epsilon$. The equation is then given as

$$p(q', t_i + \epsilon | q, t_i) = \frac{1}{\sqrt{2\pi D \epsilon}} e^{-\frac{\dot{q}^2}{2D}\epsilon} + O(\epsilon^2). \tag{2.9}$$

This is consistent with the Onsager-Machlup formula [10] in the path integral representation of stochastic processes.

Now, let us incorporate the likelihood which is the most significant factor in Bayesian update into the path integral form. We adopt the exponential function used in the previous studies [2] as likelihood $p(x, t_i | q, t_i)$;

$$p(x, t_i | q, t_i) = M \exp[-V(q(t_i) - x(t_i))\epsilon], \tag{2.10}$$

where M is a normalization constant. This form satisfies the requirement for the probability density function that it has a non-negative value in the range from $-\infty$ to ∞ . Thus, the kernel is represented as

$$\begin{aligned}
 &K(q(t_N), q(t_0), \{x(t_i), i = 1, 2, \dots, N\}) \\
 &= M^N \int dq(t_1) \int dq(t_2) \cdots \int dq(t_{N-1}) \\
 &\quad \times \exp\left[-\epsilon \sum_{i=1}^N \left\{ \frac{\dot{q}(t_i)^2}{2D} + V(q(t_i) - x(t_i)) \right\}\right]. \tag{2.11}
 \end{aligned}$$

Now let us take a limit $\epsilon \rightarrow 0$ under the condition $t_N = N\epsilon \equiv t'$. The discrete variable $t_i = i\epsilon$ becomes dense at time interval $[0, t']$. The set of realizations in stochastic processes \mathbf{q} and \mathbf{x} can be represented by continuous functions $q(t)$ and $x(t)$. Thus, the probability distribution density function $p_i(q_i)$ is also expressed as $p(q, t)$. In this limit, Eq. (2.11) is just a path integral description of Bayesian updates;

$$K(q(t'), q(t), x(t)) = \int \mathcal{D}[q] e^{-S[q(t), \dot{q}(t), x(t)]}, \tag{2.12}$$

with

$$S[q(t), \dot{q}(t), x(t)] = \int_0^{t'} \left\{ \frac{\dot{q}(t)^2}{2D} + V(q(t) - x(t)) \right\} dt, \quad (2.13)$$

where

$$\int \mathcal{D}[q] \equiv M^N \int dq(t_1) \int dq(t_2) \cdots \int dq(t_{N-1}).$$

This action S describes a particle moving in the random potential V . This is a main result of this paper. The likelihood, which is the most important factor in Bayesian update, is incorporated into the path integral form as a potential. This result seems to be quite natural. As is well known, the potential leads the force, which determines the direction of particle movement in mechanics. It is exactly consistent with the role of likelihood in Bayesian update.

3. An imaginary-time Schödinger equation

Let us derive the dynamical equation of the posterior probability distribution density associated with the Bayesian update based on the path integral formula for Bayesian updates derived in the previous section. Under an infinitesimal interval ϵ from t to t' ($t' = t + \epsilon$),

$$p(q', t' | \mathbf{x}) = M \int dq e^{-\epsilon \left[\frac{1}{2D} \left\{ \frac{q' - q}{\epsilon} \right\}^2 + V(q - x(t)) \right]} p(q, t | \mathbf{x}). \quad (3.1)$$

Let us take the Taylor series expansions for probability distribution density functions on both sides of the equation. The probability distribution density function $p(q', t' | \mathbf{x})$ is expanded with respect to ϵ as

$$p(q', t + \epsilon | \mathbf{x}) = p(q', t | \mathbf{x}) + \frac{\partial p(q', t | \mathbf{x})}{\partial t} \epsilon + O(\epsilon^2).$$

On the other hand, the probability distribution density function $p(q, t | \mathbf{x})$ on the right hand side can be expressed as $p(q' - \eta, t | \mathbf{x})$ to meet the variable on the left side by introducing a variable $\eta = q' - q$, and is expanded with respect to η ,

$$\begin{aligned} & p(q' - \eta, t | \mathbf{x}) \\ &= p(q', t | \mathbf{x}) - \frac{\partial p(q', t | \mathbf{x})}{\partial q'} \eta + \frac{1}{2} \frac{\partial^2 p(q', t | \mathbf{x})}{\partial q'^2} \eta^2 + O(\eta^3). \end{aligned}$$

Then, we obtain

$$\begin{aligned} & p(q', t | \mathbf{x}) + \frac{\partial p(q', t | \mathbf{x})}{\partial t} \epsilon \\ &= M \int d\eta e^{-\frac{\eta^2}{2\epsilon}} [1 - \epsilon V(q - x(t))] \\ & \quad \times \left[p(q', t | \mathbf{x}) + \frac{\partial p(q', t | \mathbf{x})}{\partial q'} \eta + \frac{1}{2} \frac{\partial^2 p(q', t | \mathbf{x})}{\partial q'^2} \eta^2 \right]. \quad (3.2) \end{aligned}$$

After Gaussian integration with

$$\int_{-\infty}^{\infty} \eta^{2n} \exp(-\eta^2/2D\epsilon) d\eta = (2n-1)!!(D\epsilon)^n \sqrt{2\pi D\epsilon},$$

we get

$$\begin{aligned} & p(q', t|\mathbf{x}) + \frac{\partial p(q', t|\mathbf{x})}{\partial t} \epsilon \\ &= M\sqrt{2\pi D\epsilon} [1 - \epsilon V(q - x(t))] \left[p(q', t|\mathbf{x}) + \frac{D}{2} \frac{\partial^2 p(q', t|\mathbf{x})}{\partial q'^2} \epsilon \right] \\ &= M\sqrt{2\pi D\epsilon} p(q', t|\mathbf{x}) \\ & \quad + M\sqrt{2\pi D\epsilon^3} \left\{ -V(q - x(t)) + \frac{D}{2} \frac{\partial^2 p(q', t|\mathbf{x})}{\partial q'^2} \right\} + O(\epsilon^2). \end{aligned} \quad (3.3)$$

Comparing the first term on both sides of equation yields $M = (2\pi D\epsilon)^{-1/2}$. The remaining terms eventually form

$$-\frac{\partial p(q, t|\mathbf{x})}{\partial t} = \left\{ -\frac{D}{2} \frac{\partial^2}{\partial q^2} + V(q - x(t)) \right\} p(q, t|\mathbf{x}). \quad (3.4)$$

This describes the time evolution of the posterior probability density function in Bayesian inference in Markov processes. This is nothing but the Schrödinger equation in quantum mechanics except that it is imaginary time. This equation is very advantageous for Bayesian inference in a way different from the path integral form because there are many technological accumulations in the Schrödinger equation along with the development of quantum mechanics. In the next section, we will look at some of these advantages.

3.1. Example: Gaussian likelihood

Let us consider the time evolution of the probability distribution density function using the imaginary-time Schrödinger equation for Bayesian update obtained in the previous section by adopting the Gaussian likelihood as a concrete example given by

$$V = \frac{[q - x(t)]^2}{2\Delta^2}. \quad (3.5)$$

The imaginary-time Schrödinger equation for Bayesian update to be solved is then

$$-\frac{\partial p(q, t|\mathbf{x})}{\partial t} = \left\{ -\frac{D}{2} \frac{\partial^2}{\partial q^2} + \frac{[q - x(t)]^2}{2\Delta^2} \right\} p(q, t|\mathbf{x}). \quad (3.6)$$

This equation is the imaginary time version of the Schrödinger equations of forced harmonic oscillators with a random force $x(t)$

$$F = -\frac{dV}{dq} = -\frac{q}{\Delta^2} q + \frac{q}{\Delta^2} x(t). \quad (3.7)$$

This partial differential equation can simply be solved using the Feynman-Kac formula [5, 4]. However, we will seek the solution using the operator method developed in quantum mechanics to show the usefulness of the imaginary-time Schrödinger equation in Bayesian inference.

By introducing new operators called creation (a) and annihilation (a^\dagger) operators in quantum mechanics;

$$q = \sqrt{\frac{\Delta\sqrt{D}}{2}}(a + a^\dagger), \tag{3.8}$$

$$\frac{\partial}{\partial q} = \sqrt{\frac{1}{2\Delta\sqrt{D}}}(a - a^\dagger), \tag{3.9}$$

with $[a, a^\dagger] \equiv aa^\dagger - a^\dagger a = 1$, the equation is expressed by

$$-\frac{\partial p(q, t|\mathbf{x})}{\partial t} = \left\{ \frac{\sqrt{D}}{\Delta} (a^\dagger a + 1/2) + \frac{x(t)^2}{2\Delta^2} + x(t)\sqrt{\frac{\sqrt{D}}{2\Delta^3}}(a + a^\dagger) \right\} p(q, t|\mathbf{x}). \tag{3.10}$$

The first two terms on the right hand side represent operators corresponding to Hamiltonians for a simple harmonic oscillator and a random force in quantum mechanics. On the other hand, the third term expresses their interaction. We focus on the time evolution of the probability distribution density function $p'(q, t|\mathbf{x})$ due to the interaction using the operator

$$U_0(t) = \exp \left\{ \frac{-\sqrt{D}}{\Delta} \left(a^\dagger a + \frac{1}{2} \right) t - \int_0^t \frac{x(t')^2}{2\Delta^2} dt' \right\}, \tag{3.11}$$

associated with $p(q, t|\mathbf{x}) = U_0 p'(q, t|\mathbf{x})$, this reads

$$-\frac{\partial p'(q, t|\mathbf{x})}{\partial t} = x(t)\sqrt{\frac{\sqrt{D}}{2\Delta^3}}U_0(-t)(a + a^\dagger)U_0(t)p'(q, t|\mathbf{x}). \tag{3.12}$$

According to the Baker-Campbell-Hausdorff formula [1] for any two operators, say A and B ,

$$e^A B e^{-A} = B + [A, B] + \frac{1}{2}[A, [A, B]] + \dots + \frac{1}{n!}[A, [A, \dots [A, B]]] + \dots, \tag{3.13}$$

the terms $U_0(-t)(\hat{a} + \hat{a}^\dagger)U_0(t)$ are expressed as

$$U_0(-t)(a + a^\dagger)U_0(t) = e^{-\frac{\sqrt{Dt}}{\Delta}} a + e^{\frac{\sqrt{Dt}}{\Delta}} a^\dagger. \tag{3.14}$$

Thus, Eq. (3.12) is now represented as

$$\begin{aligned} -\frac{\partial p'(q, t|\mathbf{x})}{\partial t} &= x(t)\sqrt{\frac{\sqrt{D}}{2\Delta^3}} \left(e^{-\frac{\sqrt{D}t}{\Delta}} a + e^{\frac{\sqrt{D}t}{\Delta}} a^\dagger \right) p'(q, t|\mathbf{x}) \\ &\equiv H(t)p'(q, t|\mathbf{x}). \end{aligned} \quad (3.15)$$

The formal solution of this equation is well known and given as

$$p'(q, t|\mathbf{x}) = T \left[\exp \left\{ - \int_0^t H(t') dt' \right\} \right] p'(q, 0), \quad (3.16)$$

where T is the time ordering operator

$$T[A(t), B(t')] = \begin{cases} A(t)B(t') & \text{for } t > t' \\ B(t')A(t) & \text{for } t < t'. \end{cases} \quad (3.17)$$

This is also expressed by [8]

$$p'(q, t|\mathbf{x}) = \exp \left[- \int_0^t H(t') dt' - \frac{1}{2} \int_0^t dt' \int_0^{t'} dt'' [H(t'), H(t'')] \right] p(q, 0), \quad (3.18)$$

since $[H(t), [H(t'), H(t'')]] = 0$ from

$$\begin{aligned} [H(t'), H(t'')] &= x(t')x(t'')\frac{\sqrt{D}}{2\Delta^3} \left[(e^{-\frac{\sqrt{D}t'}{\Delta}} a + e^{\frac{\sqrt{D}t'}{\Delta}} a^\dagger), (e^{-\frac{\sqrt{D}t''}{\Delta}} a + e^{\frac{\sqrt{D}t''}{\Delta}} a^\dagger) \right] \\ &= x(t')x(t'')\frac{\sqrt{D}}{2\Delta^3} \left[(e^{-\frac{\sqrt{D}(t'-t'')}{\Delta}} [a, a^\dagger] + e^{\frac{\sqrt{D}(t'-t'')}{\Delta}} [a^\dagger, a]) \right] \\ &= -x(t')x(t'')\frac{\sqrt{D}}{\Delta^3} \sinh \left(\frac{\sqrt{D}(t' - t'')}{\Delta} \right). \end{aligned} \quad (3.19)$$

The time evolution of the probability distribution density function which focused on the interaction is then given as

$$p'(q, t|\mathbf{x}) = e^{f(t)} \exp \left[- \int_0^t dt' x(t') \sqrt{\frac{\sqrt{D}}{2\Delta^3}} \left(a e^{-\frac{\sqrt{D}t'}{\Delta}} + a^\dagger e^{\frac{\sqrt{D}t'}{\Delta}} \right) \right] p'(q, 0), \quad (3.20)$$

where $f(t) \equiv (\sqrt{D}/2\Delta^3) \int_0^t dt' \int_0^{t'} dt'' x(t')x(t'') \sinh(\sqrt{D}(t' - t'')/\Delta)$. This can be reverted to the posterior probability distribution density function using the relation

$p(q, t|\mathbf{x}) = U_0(t)p'(q, t|\mathbf{x})$, resulting in

$$\begin{aligned} p(q, t|\mathbf{x}) &= e^{f(t)} \exp \left[- \int_0^t dt' x(t') \sqrt{\frac{1}{2\Delta^3}} U_0(t) (ae^{-\frac{\sqrt{D}t'}{\Delta}} + a^\dagger e^{\frac{\sqrt{D}t'}{\Delta}}) U_0^{-1} \right] \\ &\quad \times U_0(t) p(q, 0) \\ &= e^{f(t)} \exp \left[- \int_0^t dt' x(t') \sqrt{\frac{\sqrt{D}}{2\Delta^3}} (ae^{-\frac{\sqrt{D}(t'-t)}{\Delta}} + a^\dagger e^{\frac{\sqrt{D}(t'-t)}{\Delta}}) \right] \\ &\quad \times U_0(t) p(q, 0), \end{aligned} \tag{3.21}$$

where the relations $p(q, 0) = p'(q, 0)$ and $U_0(t)F(A)U_0^{-1}(t) = F(U_0(t)AU_0^{-1})$ that holds for any function $F(A)$ of the operator A are used.

The prior probability distribution density function is assumed to be given by the superposition of the harmonic oscillator solutions f_n as follows

$$p(q, 0) = \sum_n c_n f_n(q), \tag{3.22}$$

where

$$f_n(q) = \frac{1}{(\pi \Delta \sqrt{D})^{1/4} \sqrt{2^n n!}} e^{-\frac{q^2}{2\Delta\sqrt{D}}} H_n \left(\frac{q}{\sqrt{\Delta\sqrt{D}}} \right), \tag{3.23}$$

with $H_n(x)$ being the n -th order Hermite polynomial at x . The expansion coefficient c_n is given by $c_n = \int dq f_n(q) P_0(q, 0)$. Starting with this prior probability distribution density function, the posterior probability distribution density function is given as follows;

$$\begin{aligned} p(q, t|\mathbf{x}) &= e^{f(t)} \exp \left[- \int_0^t dt' x(t') \sqrt{\frac{\sqrt{D}}{2\Delta^3}} \left(ae^{-\frac{\sqrt{D}(t'-t)}{\Delta}} + a^\dagger e^{\frac{\sqrt{D}(t'-t)}{\Delta}} \right) \right] \\ &\quad \times U_0(t) \left(\sum_{n=0}^{\infty} c_n f_n(q) \right) \\ &= e^{f(t) - \int_0^t \frac{x(t')^2}{2\Delta^2} dt'} \sum_{n=0}^{\infty} c_n e^{-\frac{\sqrt{D}}{\Delta} (n+\frac{1}{2})t} \\ &\quad \times \exp \left[- \int_0^t dt' x(t') \sqrt{\frac{\sqrt{D}}{2\Delta^3}} \left(ae^{-\frac{\sqrt{D}(t'-t)}{\Delta}} + a^\dagger e^{\frac{\sqrt{D}(t'-t)}{\Delta}} \right) \right] f_n(q). \end{aligned} \tag{3.24}$$

Note that

$$e^{-\int_0^t dt' x(t') \sqrt{\frac{\sqrt{D}}{2\Delta^3}} \left(a e^{-\frac{\sqrt{D}(t'-t)}{\Delta}} + a^\dagger e^{\frac{\sqrt{D}(t'-t)}{\Delta}} \right)} = e^{-x''(t)q} e^{-x'(t)\frac{\partial}{\partial q}} e^{g(t)}, \quad (3.25)$$

where

$$x'(t) = \frac{\sqrt{D}}{\Delta} \int_0^t dt' x(t') \sinh \left\{ \frac{\sqrt{D}(t-t')}{\Delta} \right\}, \quad (3.26)$$

$$x''(t) = \frac{1}{\Delta^2} \int_0^t dt' x(t') \cosh \left\{ \frac{\sqrt{D}(t-t')}{\Delta} \right\}, \quad (3.27)$$

$$g(t) = \frac{\sqrt{D}}{2\Delta^3} \int_0^t dt' \int_0^{t'} dt'' x(t') x(t'') \times \cosh \left\{ \frac{\sqrt{D}(t-t')}{\Delta} \right\} \sinh \left\{ \frac{\sqrt{D}(t-t'')}{\Delta} \right\}, \quad (3.28)$$

and $e^{c\frac{\partial}{\partial q}} f(q) = f(q+c)$.

As a result, we reach the final expression for the posterior probability distribution density function $p(q, t|\mathbf{x})$ as

$$p(q, t|\mathbf{x}) = e^{f(t)+g(t)-\int_0^t \frac{x(t')^2}{2\Delta^2} dt'} e^{-x''(t)q} \times \sum_{n=0}^{\infty} c_n e^{-\frac{\sqrt{D}}{\Delta} \left(n+\frac{1}{2}\right)t} f_n(q - x'(t)). \quad (3.29)$$

We can find the role of Bayesian update in stochastic process in this solution obtained from stationary solution f_n of the Schrödinger equation. Evidence incorporated in Bayesian update is embedded in the argument of the function f_n . This means that the center of the posterior distribution density function is shifted with every update. In addition, the first n -independent exponent is interpreted as the global phase corresponding to the geometric phase. These features are difficult to extract in the path integral form like Feynman-Kac form. Therefore, this is the greatest advantage of analysis that makes it possible to elucidate elementary processes of stochastic processes using the Schrödinger equation.

4. Summary

We have formulated the Bayesian inference in Markov processes by means of path integral techniques and have succeeded in incorporating the likelihood which is the key in Bayesian update as a potential into the path integral form. Then we have derived the imaginary-time Schrödinger equation from path integral representation of the evolution of the Bayesian inference in Markov processes. Through an example of Gaussian likelihood, we have showed that the operator technique for the Schrödinger equation is effective for elucidating the elementary processes of stochastic processes.

Acknowledgments

We are grateful to T. Honma and K. Inagaki of Asahikawa Medical University for valuable discussions and their continuous encouragement.

References

- [1] HF Baker. Alternants and continuous groups. *Proceedings of the London Mathematical Society*, 2(1):24–47, 1905.
- [2] William Bialek, Curtis G. Callan, and Steven P. Strong. Field theories for learning probability distributions. *Phys.Rev.Lett*, pages 4693–4697, 1996.
- [3] S. Chapman. On the brownian displacements and thermal diffusion of grains suspended in a non-uniform fluid. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 119(781):34–54, 1928.
- [4] Bin Kang Cheng. The exact propagator of time-dependent forced harmonic oscillator. *Revista Brasileira de Física*, 13(1):220–230, 1983.
- [5] Mark Kac. On distributions of certain wiener functionals. *Transactions of the American Mathematical Society*, 65(1):1–13, 1949.
- [6] A. Kolmogoroff. Über die analytischen methoden in der wahrscheinlichkeitsrechnung. *Mathematische Annalen*, 104(1):415–458, Dec 1931.
- [7] John K Kruschke and Torrin M Liddell. The bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a bayesian perspective. *Psychonomic Bulletin & Review*, pages 1–29, 2017.
- [8] Eugen Merzbacher. *Quantum Mechanics*. Wiley, 1997.
- [9] Anthony O’Hagan and Mike West. *The Oxford handbook of applied Bayesian analysis*. Oxford University Press, 2010.
- [10] L. Onsager and S. Machlup. Fluctuations and irreversible processes. *Phys. Rev.*, 91:1505–1512, Sep 1953.

