Stochastic Modelling and Forecasting

Xuerong Mao

Department of Statistics and Modelling Science
University of Strathclyde
Glasgow, G1 1XH

RSE/NNSFC Workshop on
Management Science and Engineering and Public Policy
Edinburgh, 17-18 March 2008

Outline

1. Stochastic Modelling
2. Well-known Models
3. Stochastic verse Deterministic
   - Exponential growth model
   - Logistic Model
4. Forecasting and Monte Carlo Simulations
   - EM method
   - EM method for financial quantities

Outline

1. Stochastic Modelling
2. Well-known Models
3. Stochastic verse Deterministic
   - Exponential growth model
   - Logistic Model
4. Forecasting and Monte Carlo Simulations
   - EM method
   - EM method for financial quantities

Outline

1. Stochastic Modelling
2. Well-known Models
3. Stochastic verse Deterministic
   - Exponential growth model
   - Logistic Model
4. Forecasting and Monte Carlo Simulations
   - EM method
   - EM method for financial quantities

Outline

1. Stochastic Modelling
2. Well-known Models
3. Stochastic verse Deterministic
   - Exponential growth model
   - Logistic Model
4. Forecasting and Monte Carlo Simulations
   - EM method
   - EM method for financial quantities

One of the important problems in many branches of science and industry, e.g. engineering, management, finance, social science, is the specification of the stochastic process governing the behaviour of an underlying quantity. We here use the term *underlying quantity* to describe any interested object whose value is known at present but is liable to change in the future. Typical examples are

- shares in a company,
- commodities such as gold, oil or electricity,
- number of working people,
- number of pupils in primary schools.

One of the important problems in many branches of science and industry, e.g. engineering, management, finance, social science, is the specification of the stochastic process governing the behaviour of an underlying quantity. We here use the term *underlying quantity* to describe any interested object whose value is known at present but is liable to change in the future. Typical examples are

- shares in a company,
- commodities such as gold, oil or electricity,
- number of working people,
- number of pupils in primary schools.

One of the important problems in many branches of science and industry, e.g. engineering, management, finance, social science, is the specification of the stochastic process governing the behaviour of an underlying quantity. We here use the term *underlying quantity* to describe any interested object whose value is known at present but is liable to change in the future. Typical examples are

- shares in a company,
- commodities such as gold, oil or electricity,
- number of working people,
- number of pupils in primary schools.

One of the important problems in many branches of science and industry, e.g. engineering, management, finance, social science, is the specification of the stochastic process governing the behaviour of an underlying quantity. We here use the term *underlying quantity* to describe any interested object whose value is known at present but is liable to change in the future. Typical examples are

- shares in a company,
- commodities such as gold, oil or electricity,
- number of working people,
- number of pupils in primary schools.

Now suppose that at time t the underlying quantity is $x(t)$. Let us consider a small subsequent time interval dt, during which $x(t)$ changes to $x(t) + dx(t)$. (We use the notation $d\cdot$ for the small change in any quantity over this time interval when we intend to consider it as an infinitesimal change.) By definition, the intrinsic growth rate at t is $dx(t)/x(t)$. How might we model this rate?

If, given $x(t)$ at time t, the rate of change is *deterministic*, say $R = R(x(t), t)$, then

$$\frac{dx(t)}{x(t)} = R(x(t), t)dt.$$

This gives the ordinary differential equation (ODE)

$$\frac{dx(t)}{dt} = R(x(t), t)x(t).$$

However the rate of change is in general not deterministic as it is often subjective to many factors and uncertainties e.g. system uncertainty, environmental disturbances. To model the uncertainty, we may decompose

$$\frac{dx(t)}{x(t)} = \text{deterministic change} + \text{random change}.$$

The deterministic change may be modeled by

$$\bar{R}dt = \bar{R}(x(t), t)dt$$

where $\bar{R} = \bar{r}(x(t), t)$ is the average rate of change given $x(t)$ at time t. So

$$\frac{dx(t)}{x(t)} = \bar{R}(x(t), t)dt + \text{random change}.$$

How may we model the random change?

In general, the random change is affected by *many factors independently*. By the well-known central limit theorem this change can be represented by a normal distribution with mean zero and and variance $V^2 dt$, namely

$$\text{random change} = N(0, V^2 dt) = V\ N(0, dt),$$

where $V = V(x(t), t)$ is the standard deviation of the rate of change given $x(t)$ at time t, and $N(0, dt)$ is a normal distribution with mean zero and and variance dt. Hence

$$\frac{dx(t)}{x(t)} = \bar{R}(x(t), t)dt + V(x(t), t)N(0, dt).$$

A convenient way to model $N(0, dt)$ as a process is to use the *Brownian motion* $B(t)$ ($t \geq 0$) which has the following properties:

- $B(0) = 0$,
- $dB(t) = B(t + dt) - B(t)$ is independent of $B(t)$,
- $dB(t)$ follows $N(0, dt)$.

The stochastic model can therefore be written as

$$\frac{dx(t)}{x(t)} = \bar{R}(x(t), t)dt + V(x(t), t)dB(t),$$

or

$$dx(t) = \bar{R}(x(t), t)x(t)dt + V(x(t), t)x(t)dB(t)$$

which is a stochastic differential equation (SDE).

Linear model

If both $\bar{R}$ and V are constants, say

$$\bar{R}(x(t),t) = \mu, \quad V(x(t),t) = \sigma,$$

then the SDE becomes

$$dx(t) = \mu x(t)dt + \sigma x(t)dB(t).$$

This is

- the Black–Scholes geometric Brownian motion in finance,
- the exponential growth model in engineering and population.

Linear model

If both $\bar{R}$ and V are constants, say

$$\bar{R}(x(t), t) = \mu, \quad V(x(t), t) = \sigma,$$

then the SDE becomes

$$dx(t) = \mu x(t)dt + \sigma x(t)dB(t).$$

This is

- the Black–Scholes geometric Brownian motion in finance,
- the exponential growth model in engineering and population.

Logistic model

If

$$\bar{R}(x(t),t) = b + ax(t), \quad V(x(t),t) = \sigma x(t),$$

then the SDE becomes

$$dx(t) = x(t)\Big([b + ax(t)]dt + \sigma x(t)dB(t)\Big).$$

This is the well-known Logistic model in population.

Square root process

If

$$\bar{R}(x(t), t) = \mu, \quad V(x(t), t) = \frac{\sigma}{\sqrt{x(t)}},$$

then the SDE becomes the well-known square root process

$$dx(t) = \mu x(t)dt + \sigma\sqrt{x(t)}dB(t).$$

This is used widely in engineering and finance.

Mean-reverting model

If

$$\bar{R}(x(t), t) = \frac{\alpha(\mu - x(t))}{x(t)}, \quad V(x(t), t) = \frac{\sigma}{\sqrt{x(t)}},$$

then the SDE becomes

$$dx(t) = \alpha(\mu - x(t))dt + \sigma\sqrt{x(t)}dB(t).$$

This is

- the mean-reverting square root process in population,
- the CIR model for interest rate in finance.

Mean-reverting model

If

$$\bar{R}(x(t),t) = \frac{\alpha(\mu - x(t))}{x(t)}, \quad V(x(t),t) = \frac{\sigma}{\sqrt{x(t)}},$$

then the SDE becomes

$$dx(t) = \alpha(\mu - x(t))dt + \sigma\sqrt{x(t)}dB(t).$$

This is

- the mean-reverting square root process in population,
- the CIR model for interest rate in finance.

Theta process

If

$$\bar{R}(x(t), t) = \mu, \quad V(x(t), t) = \sigma(x(t))^{\theta-1},$$

then the SDE becomes

$$dx(t) = \mu x(t)dt + \sigma(x(t))^{\theta} dB(t),$$

which is known as the theta process.

Outline

1 Stochastic Modelling

2 Well-known Models

3 Stochastic verse Deterministic
- Exponential growth model
- Logistic Model

4 Forecasting and Monte Carlo Simulations
- EM method
- EM method for financial quantities

In the classical theory of population dynamics, it is assumed that the grow rate is constant μ. Thus

$$\frac{dx(t)}{x(t)} = \mu dt,$$

which is often written as the familiar ordinary differential equation (ODE)

$$\frac{dx(t)}{dt} = \mu x(t).$$

This linear ODE can be solved exactly to give exponential growth (or decay) in the population, i.e.

$$x(t) = x(0)e^{\mu t},$$

where $x(0)$ is the initial population at time $t = 0$.

We observe:

- If $\mu > 0$, $x(t) \to \infty$ exponentially, i.e. the population will grow exponentially fast.
- If $\mu < 0$, $x(t) \to 0$ exponentially, that is the population will become extinct.
- If $\mu = 0$, $x(t) = x(0)$ for all t, namely the population is stationary.

We observe:

- If $\mu > 0$, $x(t) \to \infty$ exponentially, i.e. the population will grow exponentially fast.
- If $\mu < 0$, $x(t) \to 0$ exponentially, that is the population will become extinct.
- If $\mu = 0$, $x(t) = x(0)$ for all t, namely the population is stationary.

We observe:

- If $\mu > 0$, $x(t) \to \infty$ exponentially, i.e. the population will grow exponentially fast.
- If $\mu < 0$, $x(t) \to 0$ exponentially, that is the population will become extinct.
- If $\mu = 0$, $x(t) = x(0)$ for all t, namely the population is stationary.

However, if we take the uncertainty into account as explained before, we may have

$$\frac{dx(t)}{x(t)} = rdt + \sigma dB(t).$$

This is often written as the linear SDE

$$dx(t) = \mu x(t)dt + \sigma x(t)dB(t).$$

It has the explicit solution

$$x(t) = x(0)\exp\left[(\mu - 0.5\sigma^2)t + \sigma B(t)\right].$$

Recall the properties of the Brownian motion

$$\limsup_{t\to\infty} \frac{B(t)}{\sqrt{2t\log\log t}} = 1 \quad a.s.$$

and

$$\liminf_{t\to\infty} \frac{B(t)}{\sqrt{2t\log\log t}} = -1 \quad a.s.$$

- If $\mu > 0.5\sigma^2$, $x(t) \to \infty$ exponentially with probability 1, i.e. the population will grow exponentially fast.
- If $\mu < 0.5\sigma^2$, $x(t) \to 0$ exponentially with probability 1, that is the population will become extinct.
    - In particular, this includes the case of $0 < \mu < 0.5\sigma^2$ where the population will grow exponentially in the corresponding ODE model but it will become extinct in the SDE model.
    - This reveals the important feature - *noise may make a population extinct.*
- If $\mu = 0.5\sigma^2$, $\limsup_{t\to\infty} x(t) = \infty$ while $\liminf_{t\to\infty} x(t) = 0$ with probability 1.

- If $\mu > 0.5\sigma^2$, $x(t) \to \infty$ exponentially with probability 1, i.e. the population will grow exponentially fast.
- If $\mu < 0.5\sigma^2$, $x(t) \to 0$ exponentially with probability 1, that is the population will become extinct.
  - In particular, this includes the case of $0 < \mu < 0.5\sigma^2$ where the population will grow exponentially in the corresponding ODE model but it will become extinct in the SDE model.
  - This reveals the important feature - *noise may make a population extinct*.
- If $\mu = 0.5\sigma^2$, $\limsup_{t\to\infty} x(t) = \infty$ while $\liminf_{t\to\infty} x(t) = 0$ with probability 1.

- If $\mu > 0.5\sigma^2$, $x(t) \to \infty$ exponentially with probability 1, i.e. the population will grow exponentially fast.
- If $\mu < 0.5\sigma^2$, $x(t) \to 0$ exponentially with probability 1, that is the population will become extinct.
  - In particular, this includes the case of $0 < \mu < 0.5\sigma^2$ where the population will grow exponentially in the corresponding ODE model but it will become extinct in the SDE model.
  - This reveals the important feature - *noise may make a population extinct*.
- If $\mu = 0.5\sigma^2$, $\limsup_{t\to\infty} x(t) = \infty$ while $\liminf_{t\to\infty} x(t) = 0$ with probability 1.

- If $\mu > 0.5\sigma^2$, $x(t) \to \infty$ exponentially with probability 1, i.e. the population will grow exponentially fast.
- If $\mu < 0.5\sigma^2$, $x(t) \to 0$ exponentially with probability 1, that is the population will become extinct.
    - In particular, this includes the case of $0 < \mu < 0.5\sigma^2$ where the population will grow exponentially in the corresponding ODE model but it will become extinct in the SDE model.
    - This reveals the important feature - *noise may make a population extinct*.
- If $\mu = 0.5\sigma^2$, $\limsup_{t\to\infty} x(t) = \infty$ while $\liminf_{t\to\infty} x(t) = 0$ with probability 1.

- If $\mu > 0.5\sigma^2$, $x(t) \to \infty$ exponentially with probability 1, i.e. the population will grow exponentially fast.
- If $\mu < 0.5\sigma^2$, $x(t) \to 0$ exponentially with probability 1, that is the population will become extinct.
  - In particular, this includes the case of $0 < \mu < 0.5\sigma^2$ where the population will grow exponentially in the corresponding ODE model but it will become extinct in the SDE model.
  - This reveals the important feature - *noise may make a population extinct*.
- If $\mu = 0.5\sigma^2$, $\limsup_{t\to\infty} x(t) = \infty$ while $\liminf_{t\to\infty} x(t) = 0$ with probability 1.

Outline

1. Stochastic Modelling
2. Well-known Models
3. Stochastic verse Deterministic
   - Exponential growth model
   - Logistic Model
4. Forecasting and Monte Carlo Simulations
   - EM method
   - EM method for financial quantities

The logistic model for single-species population dynamics is given by the ODE

$$\frac{dx(t)}{dt} = x(t)[b + ax(t)]. \tag{3.1}$$

- If $a < 0$ and $b > 0$, the equation has the global solution

$$x(t) = \frac{b}{-a + e^{-bt}(b + ax_0)/x_0} \quad (t \geq 0),$$

which is not only positive and bounded but also

$$\lim_{t \to \infty} x(t) = \frac{b}{|a|}.$$

- If $a > 0$, whilst retaining $b > 0$, then the equation has only the local solution

$$x(t) = \frac{b}{-a + e^{-bt}(b + ax_0)/x_0} \quad (0 \leq t < T),$$

which explodes to infinity at the finite time

$$T = -\frac{1}{b} \log\left(\frac{ax_0}{b + ax_0}\right).$$

- If $a < 0$ and $b > 0$, the equation has the global solution

$$x(t) = \frac{b}{-a + e^{-bt}(b + ax_0)/x_0} \quad (t \geq 0),$$

which is not only positive and bounded but also

$$\lim_{t \to \infty} x(t) = \frac{b}{|a|}.$$

- If $a > 0$, whilst retaining $b > 0$, then the equation has only the local solution

$$x(t) = \frac{b}{-a + e^{-bt}(b + ax_0)/x_0} \quad (0 \leq t < T),$$

which explodes to infinity at the finite time

$$T = -\frac{1}{b} \log\left(\frac{ax_0}{b + ax_0}\right).$$

Once again, the growth rate b here is not a constant but a stochastic process. Therefore, bdt should be replaced by

$$bdt + N(0, v^2 dt) = bdt + vN(0, dt) = bdt + vdB(t),$$

where v^2 is the variance of the noise intensity. Hence the ODE evolves to an SDE

$$dx(t) = x(t)\Big([b + ax(t)]dt + vdB(t)\Big). \tag{3.2}$$

The variance may or may not depend on the state $x(t)$. We consider the latter, say

$$v = \sigma x(t).$$

Then the SDE (3.2) becomes

$$dx(t) = x(t)\Big([b + ax(t)]dt + \sigma x(t)dB(t)\Big). \tag{3.3}$$

How is this SDE different from its corresponding ODE?

Significant Difference between ODE (3.1) and SDE (3.3)

- ODE (3.1): The solution explodes to infinity at a finite time if $a > 0$ and $b > 0$.
- SDE (3.3): With probability one, the solution will no longer explode in a finite time, even in the case when $a > 0$ and $b > 0$, as long as $\sigma \neq 0$.

Significant Difference between ODE (3.1) and SDE (3.3)

- ODE (3.1): The solution explodes to infinity at a finite time if $a > 0$ and $b > 0$.
- SDE (3.3): With probability one, the solution will no longer explode in a finite time, even in the case when $a > 0$ and $b > 0$, as long as $\sigma \neq 0$.

Example

$$\frac{dx(t)}{dt} = x(t)[1 + x(t)], \quad t \geq 0, \; x(0) = x_0 > 0$$

has the solution

$$x(t) = \frac{1}{-1 + e^{-t}(1 + x_0)/x_0} \quad (0 \leq t < T),$$

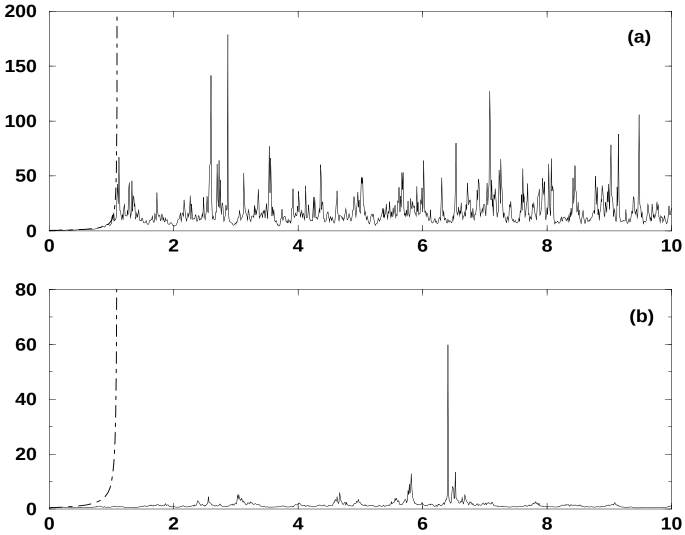
which explodes to infinity at the finite time

$$T = \log\left(\frac{1 + x_0}{x_0}\right).$$

However, the SDE

$$dx(t) = x(t)[(1 + x(t))dt + \sigma x(t)dw(t)]$$

will never explode as long as $\sigma \neq 0$.


200
150
100
50
0
0
2
4
6
8
10
(a)
80
60
40
20
0
0
2
4
6
8
10
(b)

Note on the graphs:

In graph (a) the solid curve shows a stochastic trajectory generated by the Euler scheme for time step $\Delta t = 10^{-7}$ and $\sigma = 0.25$ for a one-dimensional system (3.3) with $a = b = 1$. The corresponding deterministic trajectory is shown by the dot-dashed curve. In Graph (b) $\sigma = 1.0$.

Key Point:
When $a > 0$ and $\varepsilon = 0$ the solution explodes at the finite time $t = T$; whilst conversely, no matter how small $\varepsilon > 0$, the solution will not explode in a finite time. In other words,

stochastic environmental noise suppresses deterministic explosion.

Most of SDEs used in practice do not have explicit solutions. How can we use these SDEs to forecast?

One of the important techniques is the method of Monte Carlo simulations. There are two main motivations for such simulations:

- using a Monte Carlo approach to compute the expected value of a function of the underlying underlying quantity, for example to value a bond or to find the expected payoff of an option;
- generating time series in order to test parameter estimation algorithms.

Most of SDEs used in practice do not have explicit solutions. How can we use these SDEs to forecast?
One of the important techniques is the method of Monte Carlo simulations. There are two main motivations for such simulations:

- using a Monte Carlo approach to compute the expected value of a function of the underlying underlying quantity, for example to value a bond or to find the expected payoff of an option;
- generating time series in order to test parameter estimation algorithms.

Most of SDEs used in practice do not have explicit solutions. How can we use these SDEs to forecast?
One of the important techniques is the method of Monte Carlo simulations. There are two main motivations for such simulations:

- using a Monte Carlo approach to compute the expected value of a function of the underlying underlying quantity, for example to value a bond or to find the expected payoff of an option;
- generating time series in order to test parameter estimation algorithms.

Typically, let us consider the square root process

$$dS(t) = rS(t)dt + \sigma\sqrt{S(t)}dB(t), \quad 0 \le t \le T.$$

A numerical method, e.g. the Euler–Maruyama (EM) method applied to it may break down due to negative values being supplied to the square root function. A natural fix is to replace the SDE by the equivalent, but computationally safer, problem

$$dS(t) = rS(t)dt + \sigma\sqrt{|S(t)|}dB(t), \quad 0 \le t \le T.$$

Outline

1. Stochastic Modelling
2. Well-known Models
3. Stochastic verse Deterministic
   - Exponential growth model
   - Logistic Model
4. Forecasting and Monte Carlo Simulations
   - EM method
   - EM method for financial quantities

Discrete EM approximation

Given a stepsize $\Delta > 0$, the EM method applied to the SDE sets $s_0 = S(0)$ and computes approximations $s_n \approx S(t_n)$, where $t_n = n\Delta$, according to

$$s_{n+1} = s_n(1 + r\Delta) + \sigma\sqrt{|s_n|}\Delta B_n,$$

where $\Delta B_n = B(t_{n+1}) - B(t_n)$.

Continuous-time EM approximation

$$s(t) := s_0 + r \int_0^t \bar{s}(u))du + \sigma \int_0^t \sqrt{|\bar{s}(u)|}dB(u),$$

where the “step function” $\bar{s}(t)$ is defined by

$$\bar{s}(t) := s_n, \quad \text{for } t \in [t_n, t_{n+1}).$$

Note that $s(t)$ and $\bar{s}(t)$ coincide with the discrete solution at the gridpoints; $\bar{s}(t_n) = s(t_n) = s_n$.

The ability of the EM method to approximate the true solution is guaranteed by the ability of either $s(t)$ or $\bar{s}(t)$ to approximate $S(t)$ which is described by:

Theorem

$$\lim_{\Delta\to 0} E\Big(\sup_{0\le t\le T} |s(t)-S(t)|^2\Big) = \lim_{\Delta\to 0} E\Big(\sup_{0\le t\le T} |\bar{s}(t)-S(t)|^2\Big) = 0.$$

The ability of the EM method to approximate the true solution is guaranteed by the ability of either $s(t)$ or $\bar{s}(t)$ to approximate $S(t)$ which is described by:

Theorem

$$\lim_{\Delta\to 0} E\Big(\sup_{0\le t\le T} |s(t) - S(t)|^2\Big) = \lim_{\Delta\to 0} E\Big(\sup_{0\le t\le T} |\bar{s}(t) - S(t)|^2\Big) = 0.$$

Outline

1. Stochastic Modelling
2. Well-known Models
3. Stochastic verse Deterministic
   - Exponential growth model
   - Logistic Model
4. Forecasting and Monte Carlo Simulations
   - EM method
   - EM method for financial quantities

Bond

If $S(t)$ models short-term interest rate dynamics, it is pertinent to consider the expected payoff

$$\beta := \mathbb{E} \exp \left(-\int_0^T S(t)dt \right)$$

from a bond. A natural approximation based on the EM method is

$$\beta_\Delta := \mathbb{E} \exp \left(-\Delta \sum_{n=0}^{N-1} |s_n| \right), \text{ where } N = T/\Delta.$$

Theorem

$$\lim_{\Delta \to 0} |\beta - \beta_\Delta| = 0.$$

Bond

If $S(t)$ models short-term interest rate dynamics, it is pertinent to consider the expected payoff

$$\beta := \mathbb{E} \exp\left(- \int_0^T S(t)dt \right)$$

from a bond. A natural approximation based on the EM method is

$$\beta_\Delta := \mathbb{E} \exp\left(-\Delta \sum_{n=0}^{N-1} |s_n| \right), \text{ where } N = T/\Delta.$$

Theorem

$$\lim_{\Delta \to 0} |\beta - \beta_\Delta| = 0.$$

Up-and-out call option

An up-and-out call option at expiry time T pays the European value with the exercise price K if $S(t)$ never exceeded the fixed barrier, c, and pays zero otherwise.

Theorem

Define

$$\begin{aligned} V &= \mathbb{E}\left[(S(T)-K)^+ I_{\{0\le S(t)\le c,\ 0\le t\le T\}}\right], \\ V_\Delta &= \mathbb{E}\left[(\bar{s}(T)-K)^+ I_{\{0\le \bar{s}(t)\le B,\ 0\le t\le T\}}\right]. \end{aligned}$$

Then

$$\lim_{\Delta\to 0} |V - V_\Delta| = 0.$$

Up-and-out call option

An up-and-out call option at expiry time T pays the European value with the exercise price K if $S(t)$ never exceeded the fixed barrier, c, and pays zero otherwise.

Theorem

Define

$$\begin{aligned} V &= \mathbb{E}\left[(S(T)-K)^{+} I_{\{0\le S(t)\le c,\ 0\le t\le T\}}\right], \\ V_\Delta &= \mathbb{E}\left[(\bar{s}(T)-K)^{+} I_{\{0\le \bar{s}(t)\le B,\ 0\le t\le T\}}\right]. \end{aligned}$$

Then

$$\lim_{\Delta\to 0} |V - V_\Delta| = 0.$$