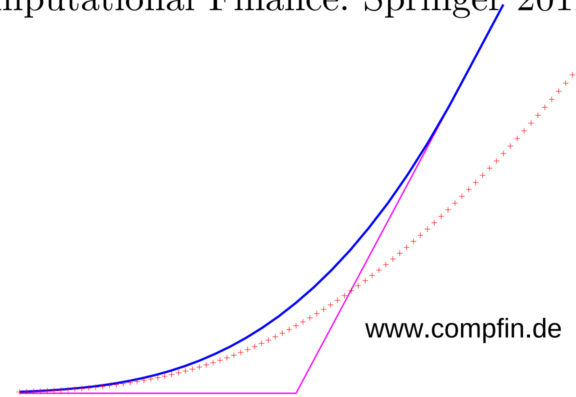
Modeling Tools

Rüdiger Seydel

Lectures on Computational Finance

following Chapter 1 of R. Seydel: Tools for Computational Finance. Springer 2012



University of Cologne
www.compfin.de
Version April 2012

1.1 Options

An option is the right (but not the obligation) to buy or sell a risky asset at a prespecified fixed price within a specified period.

underlying: stocks, indices, currencies, commodities

agreement between two parties about trading the asset at a certain future time.

The *writer* fixes the terms of the option contract and sells the option.

The *holder* purchases the option, paying the market price *(premium)*.

Terminology

The **call** option gives the holder the right to *buy* the underlying for an agreed price K by the date T. The **put** option gives the holder the right to *sell* ...

maturity date T : At time T the rights of the holder expire.

S, or S_t or $S(t)$ price per share of the underlying

The price K of the contract is called **strike** or **exercise price**.

For **European options** exercise is only permitted at expiry date T.

American options can be exercised early.

The dependence of V on S and t is written $V(S,t)$.

Payoff Function

At maturity $t = T$, the rational holder of a European **call** will exercise (get the stock for the strike price K), when $S > K$. (He can immediately sell the asset for the spot price S and makes a gain of $S - K$ per share.)

Then the value of the option is $V = S - K$.

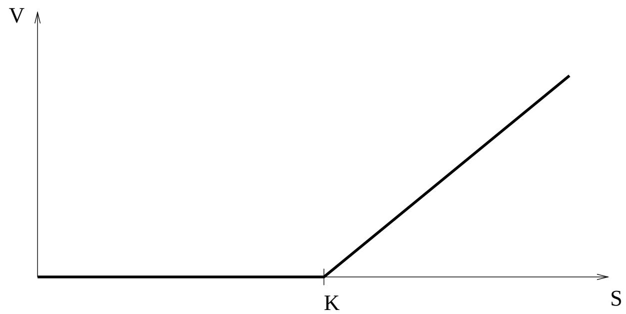
In case $S < K$ the holder will not exercise, (the asset can be purchased on the market for the cheaper price S) hence $V = 0$.

$$V(S_T, T) = \begin{cases} 0 & \text{in case } S_T \leq K \text{ (option expires worthless)} \\ S_T - K & \text{in case } S_T > K \text{ (option is exercised)} \end{cases}$$

or

$$V(S_T, T) = \max\{S_T - K, 0\} = (S_T - K)^+. \tag{1.1C}$$

(**payoff function**, *intrinsic value, cashflow*)



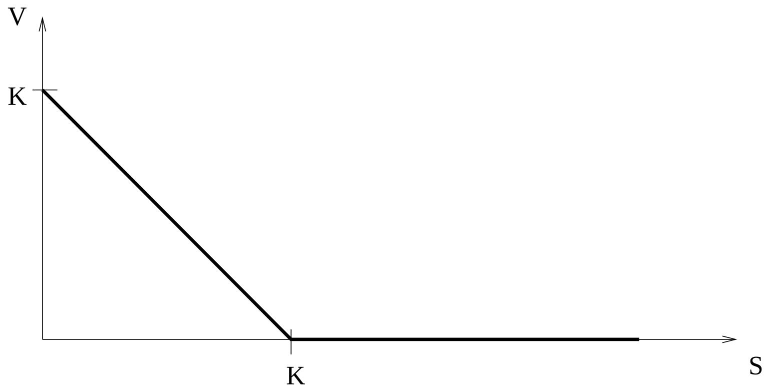
For a European **put** exercising only makes sense in case $S < K$.

The payoff $V(S, T)$ of a put at expiration time T is

$$V(S_T, T) = \begin{cases} K - S_T & \text{in case } S_T < K \text{ (option is exercised)} \\ 0 & \text{in case } S_T \geq K \text{ (option is worthless)} \end{cases}$$

or

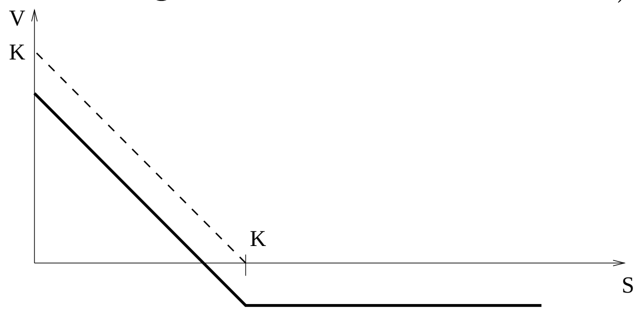
$$V(S_T, T) = \max\{K - S_T, 0\} = (K - S_T)^+ \qquad (1.1\text{P})$$



Profit: The initial costs paid when buying the option at $t = t_0$ must be subtracted.

The initial costs consist of the premium and the transaction costs. Both are multiplied by $e^{r(T-t_0)}$ to take account of the time value; r is the interest rate.

(negative profit for some range of S-values means a loss.)



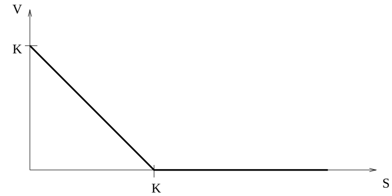
The payoff function for an American call is $(S_t - K)^+$ and for an American put $(K - S_t)^+$ for any $t \leq T$.

The situation for the writer (short position) is reverse. For him the above payoff curves as well as the profit curves are reflected on the S-axis. The writer's profit or loss is the reverse of that of the holder.

A Priori Bounds / Arbitrage

The value $V(S,t)$ of an American option can never fall below the payoff.

This bound follows from the **no-arbitrage principle**.



Assume for an American put that its value is below the payoff. $V < 0$ contradicts the definition of the option. Hence $V \geq 0$, and S and V satisfy $S < K$ and $0 \leq V < K - S$.

This scenario would allow arbitrage as follows: Borrow the cash amount of $S + V$, and buy both the underlying and the put. Then immediately exercise the put, selling the underlying for the strike price K. The profit of this arbitrage strategy is $K - S - V > 0$. This is in conflict with the no-arbitrage principle. We conclude

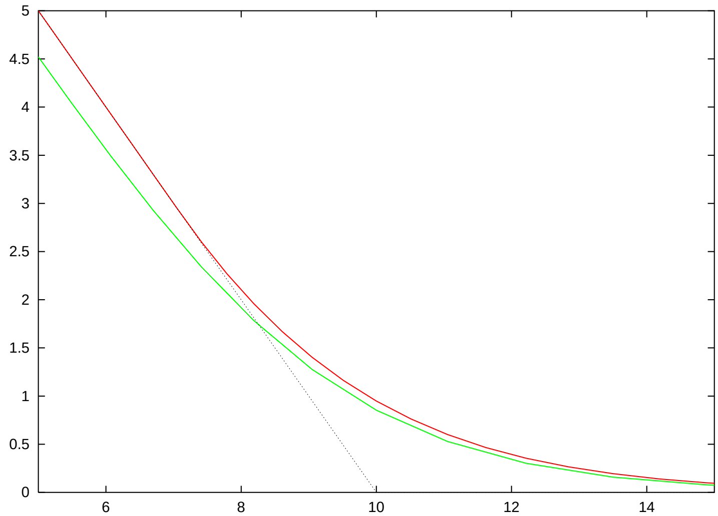
$$V_{\mathrm{P}}^{\mathrm{am}}(S,t) \geq (K-S)^+ \quad \text{for all } S,t \ .$$

Similarly,

$$V_{\mathrm{C}}^{\mathrm{am}}(S,t) \geq (S-K)^+ \quad \text{for all } S,t \ .$$

The value of an American option should never be smaller than that of a European option because the American type includes the European type exercise at $t = T$ and in addition *early exercise* for $t < T$, hence

$$V^{\text{am}} \geq V^{\text{eur}}$$



$V(S, 0)$ of a put American option (red), European option (green), payoff

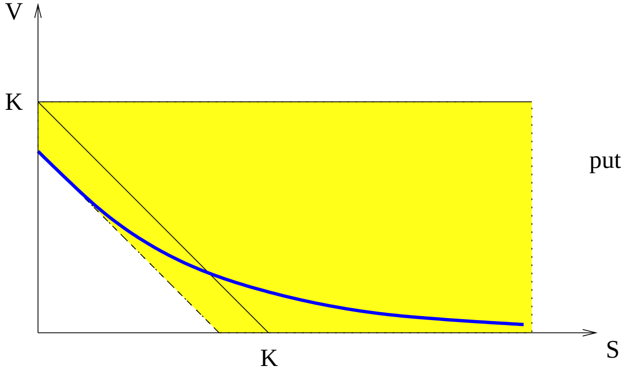
For European options the values of put and call are related by the **put-call parity**

$$S + V_{\text{P}} - V_{\text{C}} = Ke^{-r(T-t)} \ .$$

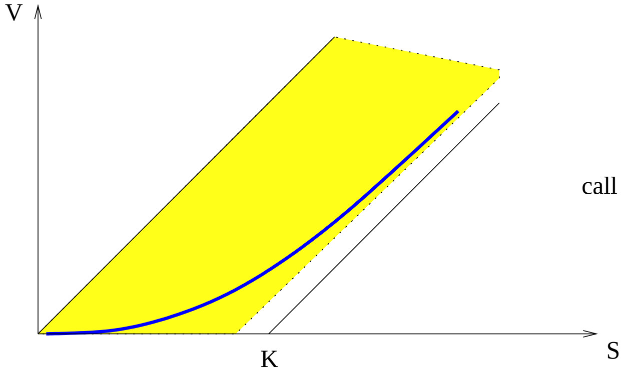
(assumes no dividend payment for $0 \leq t \leq T$, and no transaction costs)

Bounds on European-Style Options

argument: arbitrage



$$V^{\mathrm{am}}(S,t) \geq V^{\mathrm{eur}}(S,t)$$
$$V^{\mathrm{am}}(S,t) \geq \text{payoff}$$



provided that no dividend is paid:

$$V_{\mathrm{C}}^{\mathrm{eur}}(S,t) \geq S - Ke^{-r(T-t)}$$
$$V_{\mathrm{P}}^{\mathrm{eur}}(S,t) \geq Ke^{-r(T-t)} - S$$

Options in the Market

The features of the options imply that an investor purchases puts when he expects the price of the underlying is expected to fall, and buys calls when the prices are about to rise.

$V(S, t)$ also depends on:

the strike price K and the maturity T;

market parameter **interest** rate r, risk-free, continuously compounded, per year;

dividends in case of a dividend-paying asset;

market parameter **volatility** σ of the price S_t

(σ defined as standard deviation of the fluctuations in S_t, for scaling divided by the square root of the observed time period; Writing $\sigma = 0.2$ means a volatility of 20%.)

The time period of interest is $t_0 \leq t \leq T$. We set $t_0 = 0$ in the role of "today." The interval $0 \leq t \leq T$ represents the remaining life time of the option.

In real markets $r(t)$ and $\sigma(t)$. We mostly assume r and σ to be constant on $0 \leq t \leq T$. Further suppose that all variables are arbitrarily divisible and consequently can vary continuously. ($\in \mathbb{R}$)

t	current time, $0 \leq t \leq T$
T	expiration time, maturity
$r > 0$	risk-free interest rate
$S,\ S_t$	spot price, current price per share of stock/asset/underlying
σ	annual volatility
K	strike price, exercise price per share
$V(S,t)$	value of an option at time t and underlying price S

The Geometry of American-Style Standard Options

Standard options are options on one underlying with one of the above two payoffs $\Psi(S) := (K-S)^+$ or $\Psi(S) := (S-K)^+$. All other options are called *exotic*.

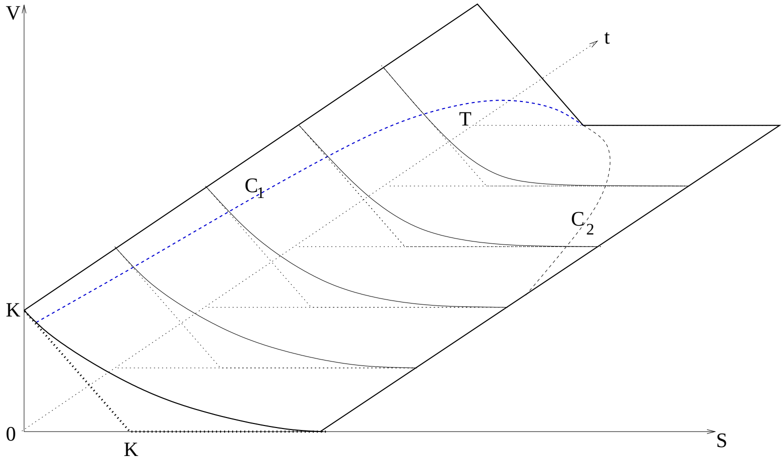
Exotic options include options on a basket of several underlyings, or other payoffs (example: binary option), or *path-dependent* options where the value depends on the entire path S_t for $0 \leq t \leq T$ (example: barrier option).

In what follows, we stick to standard options.

The values $V(S,t)$ for fixed values of K, T, r, σ can be interpreted as a piece of surface over the subset

$$S > 0\ ,\ \ 0 \leq t \leq T\ .$$

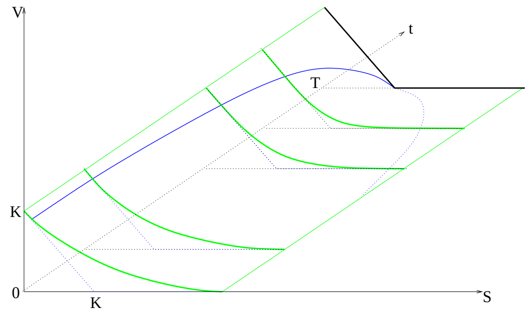
Shifting the payoff parallel for all $0 \leq t < T$ creates another surface, which consists of the two planar pieces $V = 0$ (for $S \geq K$) and $V = K - S$ (for $S < K$). This *payoff surface* created by $(K-S)^+$ is a lower bound to the option surface, $V(S,t) \geq (K-S)^+$.

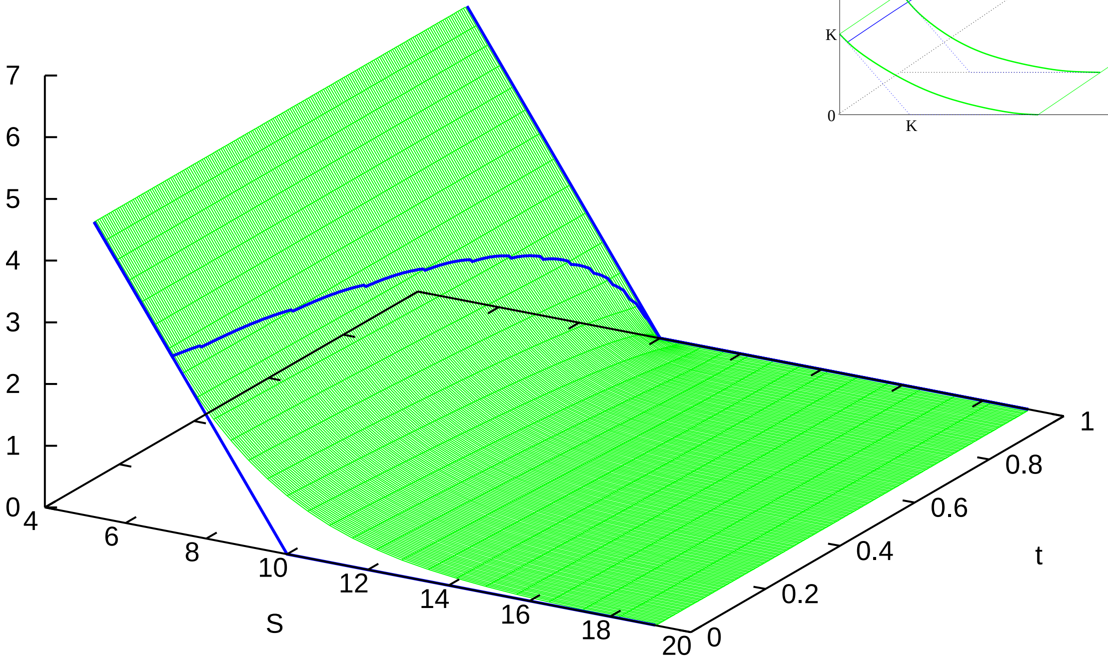


C_1: **early-exercise curve**
When S_t reaches C_1, then immediate exercising is optimal: invest K for the rate r.

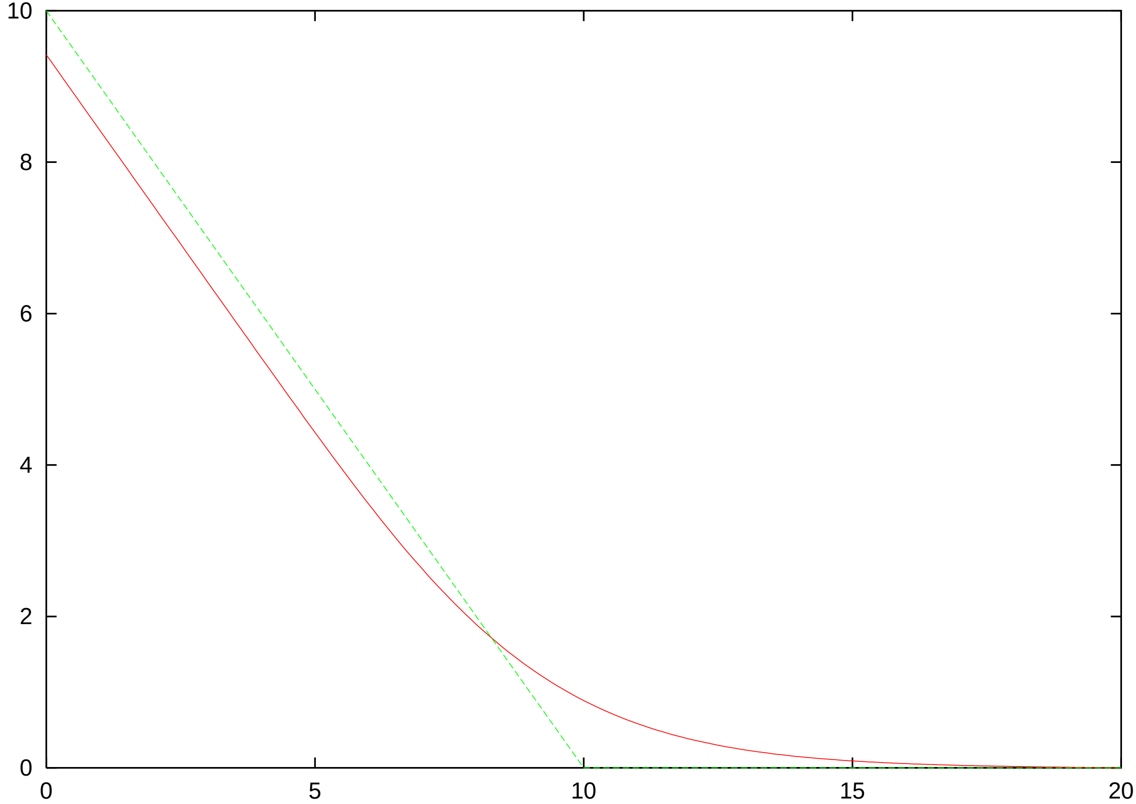
Within the area limited by the curves C_1, C_2, the option surface obeys $V(S,t) > (K-S)^+$. Outside that area, both surfaces coincide. This is strict above C_1, where $V(S,t) = K - S$, and holds approximately for S beyond C_2, where $V(S,t) \approx 0$ or $V(S,t) < \varepsilon$ for a small value of $\varepsilon > 0$.

The locations of C_1 and C_2 are not known, these curves are calculated along with the calculation of $V(S,t)$. Of special interest is $V(S,0)$, the value of the option "today."

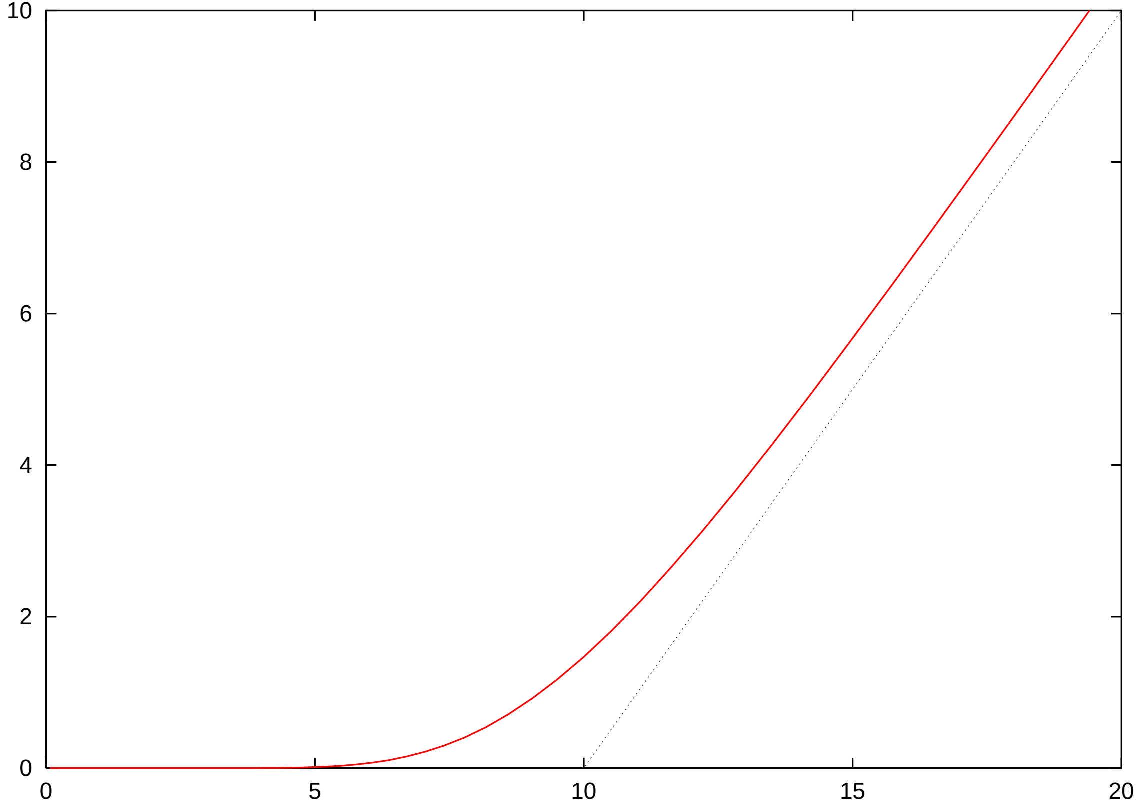




American put, $r = 0.06$, $\sigma = 0.30$



European put $V(S,0)$ for $T=1$, $K=10$, $r=0.06$, $\sigma=0.3$.



European call $V(S, 0)$ for $T = 1$, $K = 10$, $r = 0.06$, $\sigma = 0.3$.

1.2 Model of the Financial Market

classical model after Black, Merton and Scholes (1973)

attractive: option surfaces $V(S,t)$ on the half strip $S > 0$, $0 \leq t \leq T$ as *solutions of suitable equations.*

Then calculating V amounts to solving the equations.

Definition 1.1 (Black-Scholes equation)

$$\boxed{\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0} \tag{1.2}$$

partial differential equation (PDE) for $V(S,t)$, linear

terminal condition for $t = T$

$$V(S,T) = \text{payoff},$$

with payoff function depending on the type of option.

Assumptions 1.2 (B-M-S model of the market)

(a) *The market is frictionless.*
no transaction costs (fees or taxes), interest rates for borrowing and lending money are equal, all parties have immediate access to any information, all securities and credits are available at any time and in any size. (Consequently, all variables are perfectly divisible.) Individual trading will not influence the price.

(b) *There are no arbitrage opportunities.*

(c) *The asset price follows a geometric Brownian motion.*

(d) Technical assumptions (preliminary):
r and σ are constant for $0 \leq t \leq T$. No dividends are paid in that time period. The option is European.

These assumptions lead to the Black-Scholes equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

Solutions $V(S,t)$ of European standard options are functions satisfying this equation with terminal condition for all S and t.

domain: half strip $0 < S$, $0 \leq t \leq T$

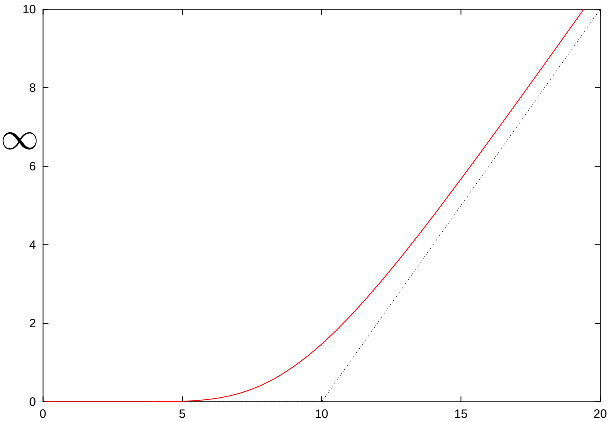
boundary conditions

For numerical purposes, the infinite interval for S must be truncated to $S_{\min} \leq S \leq S_{\max}$, which requires boundary conditions for $S_{\min}$ and $S_{\max}$. Sometimes boundary conditions are not clear and are selected in an artificial way.

example: for a European call the boundary conditions are straightforward, they will be based on

$$V(0,t) = 0\,;$$

$$V(S,t) \to S - Ke^{-r(T-t)} \text{ for } \; S \to \infty$$



transaction costs or **feedback** lead to nonlinear BS-type PDEs.

1.3 Numerical Methods

inevitable in all fields of technology including financial engineering.

Stochastic approaches are natural tools to simulate prices.

stochastic differential equations, Monte Carlo methods, simulate randomness (performed in a deterministic manner)

More efficient methods are preferred *provided their use can be justified* by the validity of the underlying models.

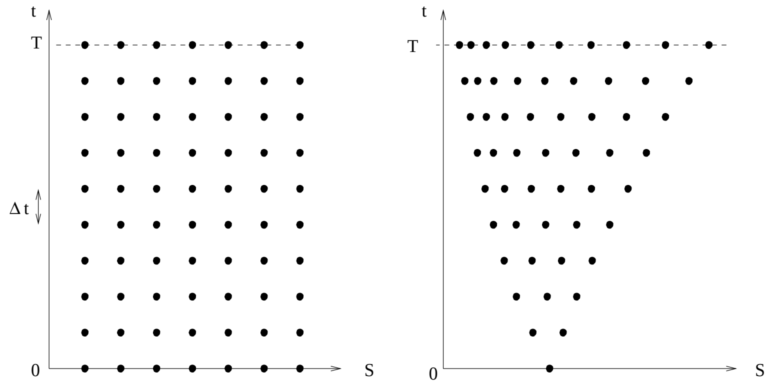
partial differential equations of the Black-Scholes type

choice among finite-difference methods and finite-element methods.

The numerical treatment of exotic options requires a more careful consideration of stability issues.

Efficiency and reliability are key demands.

Discretization



The assumption that all variables $\in \mathbb{R}$ allows to impose *artificial* discretizations convenient for the numerical methods.

The hypothesis of a continuum applies to the (S, t)-domain of the half strip $0 \leq t \leq T$, $S > 0$, and to the differential equations. The artificial discretization introduced by numerical methods is at least twofold:

1.) (S, t)-domain replaced by a **grid** of a finite number of (S, t)-points.

2.) differential equations replaced by a finite number of algebraic equations.

Discretization errors depend on the coarsity of the grid, on Δt and on ΔS. It is one of the aims of numerical algorithms to control the errors.

1.4 The Binomial Method

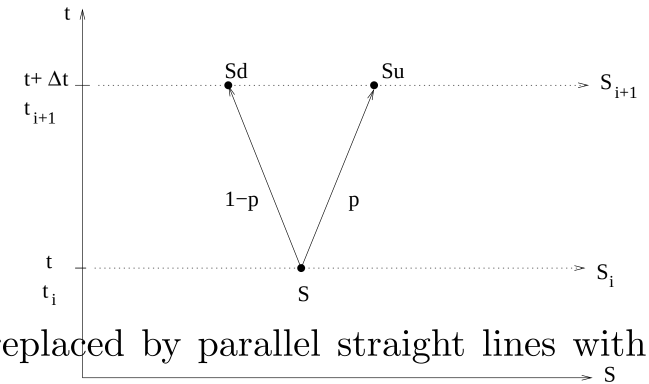
robust and widely applicable.

In practice one is often interested in the one value $V(S_0, 0)$. The binomial method is based on a tree-type grid applying appropriate binary rules at each grid point. The grid is not predefined but is constructed by the method.

A Discrete Model

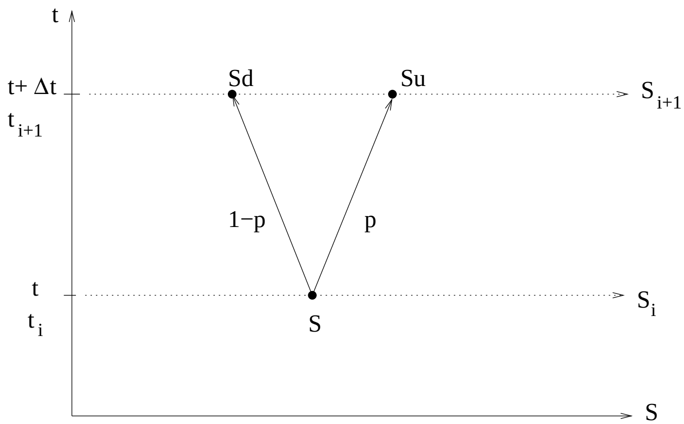
First discretize the continuous time t.

M: number of time steps
$\Delta t := \frac{T}{M}$
$t_i := i \cdot \Delta t, \quad i = 0, ..., M$
$S_i := S(t_i)$



So far the domain of the (S, t) half strip is replaced by parallel straight lines with distance Δt apart.

Next replace the continuous values S_i along the parallel $t = t_i$ by discrete values S_{ji}, for all i and appropriate j. The figure shows a mesh of the grid, namely the transition from t to $t + \Delta t$, or from t_i to t_{i+1}.



Assumptions 1.3 (binomial method)

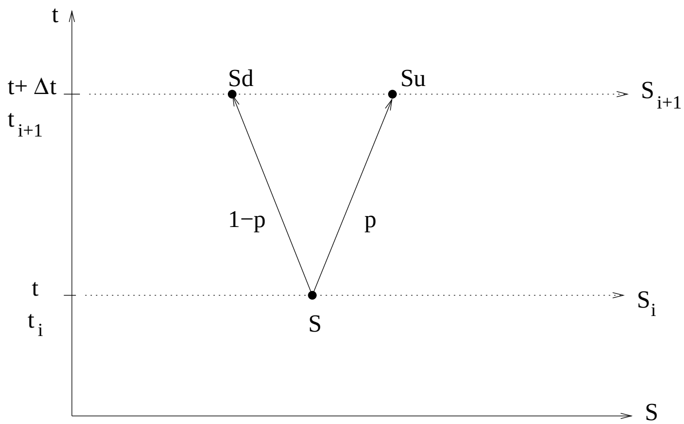
(Bi1) The price S over each period of time Δt can only have two possible outcomes: An initial value S either evolves up to Su or down to Sd with $0 < d < u$.

(Bi2) The probability of an up movement is p, $\mathsf{P}(\text{up}) = p$, with $0 < p < 1$.

(Bi3) Expectation and variance match their continuous-time counterparts.

(Temporarily assume that no dividend is paid within the time period of interest.)

The rules (Bi1), (Bi2) represent a binomial process with probability.



For (Bi3), we compare to an asset price S_t that develops randomly from a value S_i at $t = t_i$ to S_{i+1} at $t = t_{i+1}$, following a continuous-time geometric Brownian motion S_t (see below), with growth rate being the risk-free interest rate r. The expectation is

$$\mathsf{E}(S_{i+1}) = S_i \cdot e^{r\Delta t}\,, \tag{1.4}$$

analogously for the variances.

The probability P of (Bi2) is an artificial risk-neutral probability matching (Bi3). The expectation in $\mathsf{E}(S_{i+1}) = S_i \cdot e^{r\Delta t}$ refers to this probability; this is sometimes written E_{P}.

The parameters u, d and p are unknown.

A consequence of (Bi1) and (Bi2) for the discrete model is

$$\mathsf{E}(S_{i+1}) = pS_i u + (1-p)S_i d\,.$$

Here S_i is an arbitrary value for t_i, which develops randomly to S_{i+1}, following (Bi1), (Bi2). Equating with $\mathsf{E}(S_{i+1}) = S_i \cdot e^{r\Delta t}$ gives

$$e^{r\Delta t} = pu + (1-p)d\,. \tag{1.5}$$

(first equation to fix u, d, p)

Solved for the risk-neutral probability p leads to

$$p = \frac{e^{r\Delta t} - d}{u - d}\,. \tag{1.6}$$

To be a valid model of probability, $0 \le p \le 1$ must hold, or

$$d \le e^{r\Delta t} \le u\,. \tag{1.7}$$

To prevent arbitrage, $d < e^{r\Delta t} < u$ must hold, which is assumed in (Bi2).

equate variances: Via the variance the volatility σ enters the model. From the *continuous model* we apply

$$\mathsf{E}(S_{i+1}^2) = S_i^2 e^{(2r+\sigma^2)\Delta t}\,. \tag{1.8}$$

Recall $\mathsf{Var}(S) = \mathsf{E}(S^2) - (\mathsf{E}(S))^2$. Equations $\mathsf{E}(S_{i+1}) = S_i \cdot e^{r\Delta t}$ and (1.8) combine to

$$\mathsf{Var}(S_{i+1}) = S_i^2 e^{2r\Delta t}(e^{\sigma^2\Delta t} - 1)\,.$$

The *discrete model* satisfies

$$\begin{aligned} \mathsf{Var}(S_{i+1}) &= \mathsf{E}(S_{i+1}^2) - (\mathsf{E}(S_{i+1}))^2 \\ &= p(S_i u)^2 + (1-p)(S_i d)^2 - S_i^2(pu + (1-p)d)^2\,. \end{aligned}$$

Equating variances of the continuous and the discrete model, and applying $e^{r\Delta t} = pu + (1-p)d$ leads to

$$\begin{aligned} e^{2r\Delta t}(e^{\sigma^2\Delta t} - 1) &= pu^2 + (1-p)d^2 - (e^{r\Delta t})^2 \\ e^{2r\Delta t + \sigma^2\Delta t} &= pu^2 + (1-p)d^2 \end{aligned} \tag{1.9}$$

This equation and the above $e^{r\Delta t} = pu + (1-p)d$ constitute two relations for the three unknowns u, d, p.

We are free to impose an arbitrary third equation.

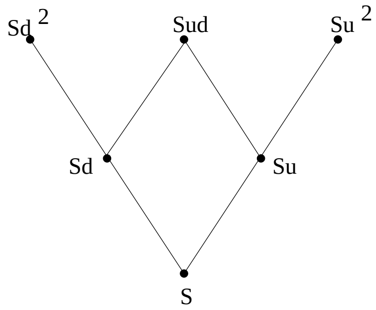
One example is the plausible assumption

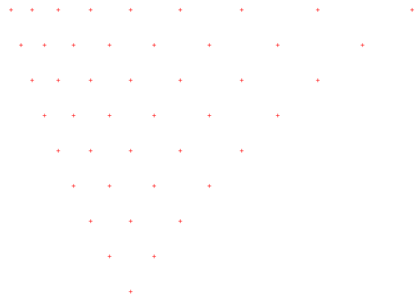
$$u \cdot d = 1\,, \tag{1.10}$$

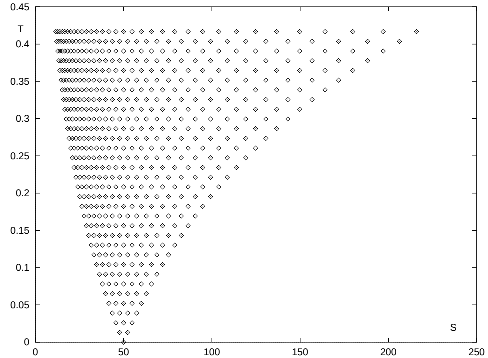
which reflects a symmetry between upward and downward movement.

Now the parameters u, d and p are fixed.
They depend on r, σ and Δt.
So does the grid.

Other choices:
$ud = \gamma$ with suitably chosen γ, or
$p = 0.5$

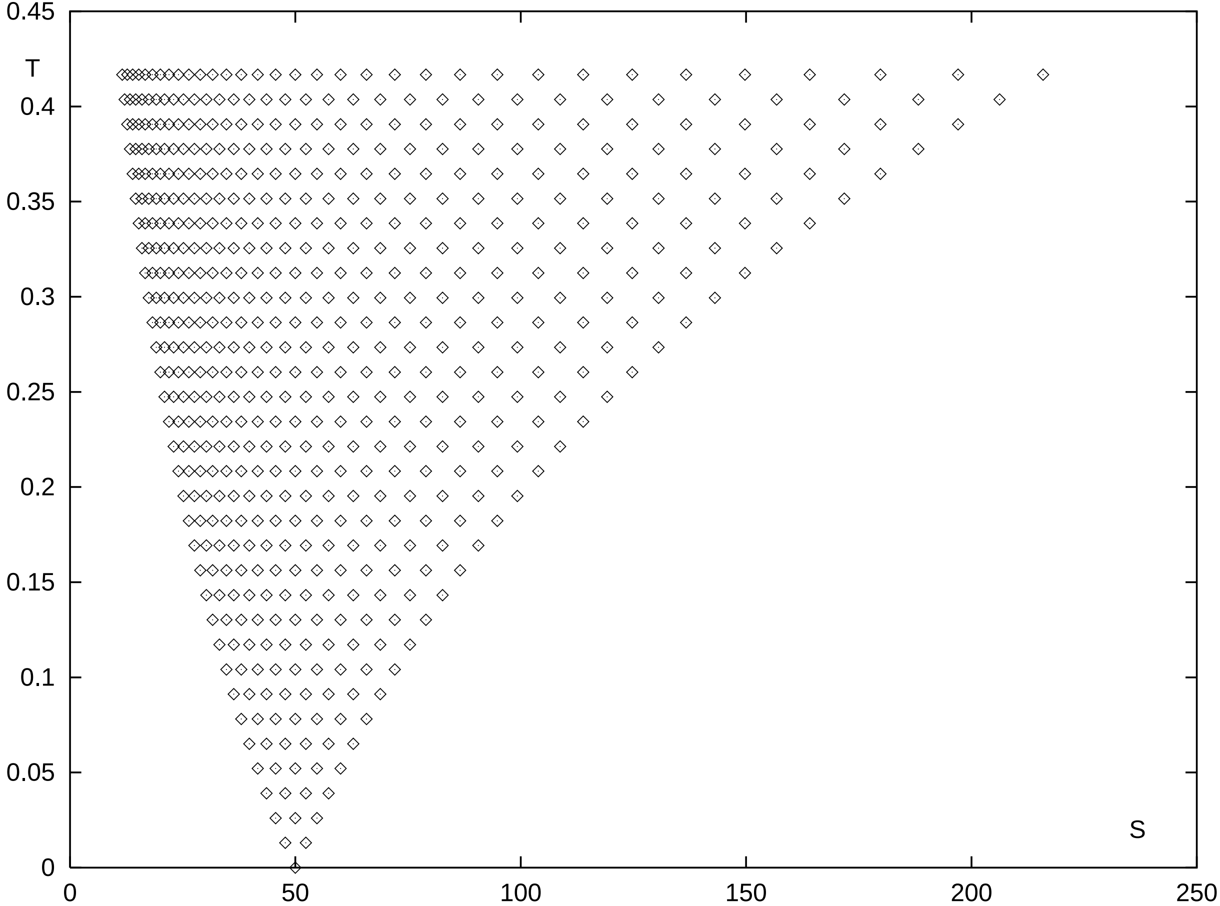






The above rules are applied to each grid line $i = 0, \ldots, M$, starting at $t_0 = 0$ with the specific value $S = S_0$. Attaching meshes for subsequent values of t_i builds a tree with values $Su^j d^k$ and $j + k = i$. In this way, specific discrete values S_{ji} of S_i are defined.

Since the same constant factors u and d underlie all meshes and since $Sud = Sdu$ holds, the tree is *recombining*. It does not matter which of the two paths we take to reach Sud. Consequently the binomial process defined by Assumption 1.3 is *path independent*. Accordingly at expiration time $T = M\Delta t$ the price S can take only the $(M + 1)$ discrete values $Su^j d^{M-j}$, $j = 0, 1, ..., M$. By $ud = 1$ these are the values $Su^j u^{j-M} = Su^{-M} u^{2j} =: S_{jM}$.



T
S
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
50
100
150
200
250

Solution of (1.5), (1.9), (1.10)

Use $\alpha := e^{r\Delta t}$ to obtain the quadratic

$$0 = u^2 - u(\underbrace{\alpha^{-1} + \alpha e^{\sigma^2 \Delta t}}_{=:2\beta}) + 1,$$

with solutions $u = \beta \pm \sqrt{\beta^2 - 1}$. By virtue of $ud = 1$ and Vieta's Theorem, d is the solution with the minus sign.

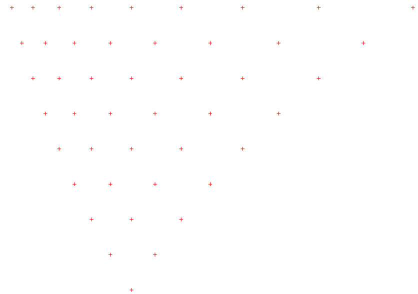
$$\boxed{\begin{aligned} \beta &:= \frac{1}{2}(e^{-r\Delta t} + e^{(r+\sigma^2)\Delta t}) \\ u &= \beta + \sqrt{\beta^2 - 1} \\ d &= 1/u = \beta - \sqrt{\beta^2 - 1} \\ p &= \frac{e^{r\Delta t} - d}{u - d} \end{aligned}} \qquad (1.11)$$

Forward Phase: Initializing the Tree

The current spot price $S = S_0$ for $t_0 = 0$ is the root of the tree. (also denoted S_{00})

$$\begin{array}{l} \textit{For } i = 1, 2, ..., M \ \textit{ calculate}: \\ S_{ji} := S_0 u^j d^{i-j}, \quad j = 0, 1, ..., i \end{array}$$

Now the grid points (t_i, S_{ji}) are fixed,
on which the option values $V_{ji} := V(t_i, S_{ji})$ are calculated.



Calculating the Option Values V, Valuation of the Tree

For t_M the payoff $V(S, t_M) = \Psi(S)$ is known. This defines the values V_{jM}:

$$V_{jM} := \Psi(S_{jM}) \tag{1.12}$$

with $\Psi(S) = (S - K)^+$ for a call, and $\Psi(S) = (K - S)^+$ for a put.

Backward Phase

Calculate recursively for $t_{M-1},\ t_{M-2}, ...$ the option values V for all t_i, starting from V_{jM}. The recursion is based on Assumption 1.3, (Bi3). Repeating the equation that corresponds to (1.5) with double index leads to

$$S_{ji}e^{r\Delta t} = pS_{ji}u + (1-p)S_{ji}d,$$

and

$$S_{ji}e^{r\Delta t} = pS_{j+1,i+1} + (1-p)S_{j,i+1}.$$

Relating the Assumption 1.3, (Bi3) of risk neutrality to V, $\quad V_i = e^{-r\Delta t}\mathsf{E}(V_{i+1})$, we obtain

$$V_{ji} = e^{-r\Delta t} \cdot \left(pV_{j+1,i+1} + (1-p)V_{j,i+1}\right). \tag{1.13}$$

For **European options** this is a recursion for $i = M-1, \ldots, 0$, starting from the payoff, and terminating with V_{00}. The obtained value V_{00} is an approximation to the value $V(S_0, 0)$ of the continuous model, which results in the limit $M \to \infty$ ($\Delta t \to 0$).

For **American options** the above recursion must be modified by adding a test whether early exercise is preferred.

The values V_{ji} of (1.13) are the "continuation" values V_{ji}^{cont} applicable when no early exercise is due. For each t_i the holder optimizes his position by choosing the best of

$$\{\, exercise,\ continue \,\},$$

or

$$\max\{\Psi(S),\ V^{\text{cont}}\}.$$

Hence, the equations (1.12) for i rather than M, combined with (1.13), become

Call:

$$V_{ji} = \max\left\{(S_{ji} - K)^+,\ e^{-r\Delta t} \cdot (pV_{j+1,i+1} + (1-p)V_{j,i+1})\right\} \tag{1.14C}$$

Put:

$$V_{ji} = \max\left\{(K - S_{ji})^+,\ e^{-r\Delta t} \cdot (pV_{j+1,i+1} + (1-p)V_{j,i+1})\right\} \tag{1.14P}$$

Algorithm 1.4 (binomial method)

Input: $r,\ \sigma,\ S = S_0,\ T,\ K,$ choice of put or call, European or American, M

calculate: $\Delta t := T/M,\ u,\ d,\ p$ from (1.11)

$S_{00} := S_0$

$S_{jM} = S_{00} u^j d^{M-j},\ j = 0, 1, ..., M$

(for American options, also $S_{ji} = S_{00} u^j d^{i-j}$ for $0 < i < M$, $j = 0, 1, ..., i$)

V_{jM} from (1.12)

$$V_{ji} \text{ for } i < M \begin{cases} \text{from (1.13) for European options} \\ \text{from (1.14) for American options} \end{cases}$$

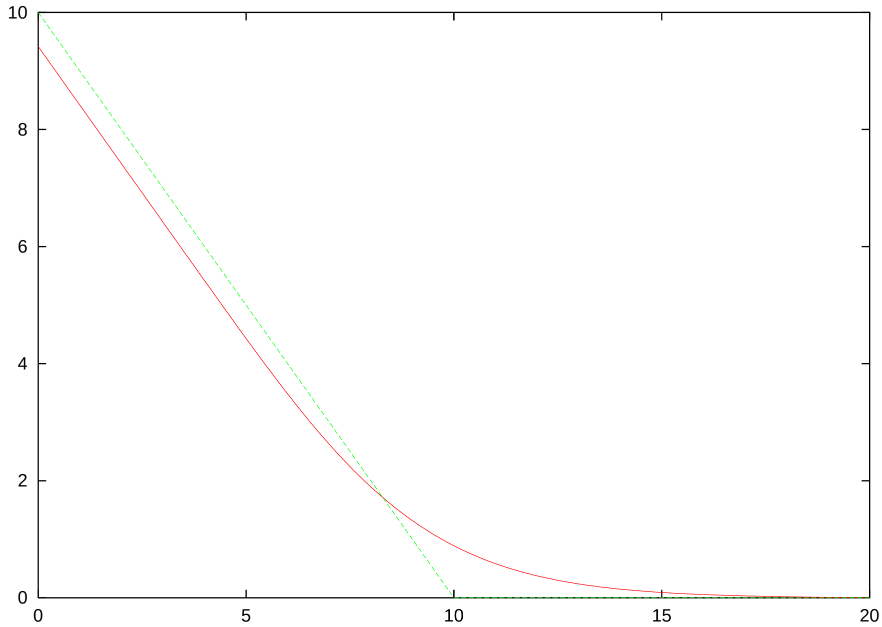
Output: V_{00} is the approximation $V_0^{(M)}$ of $V(S_0, 0)$

Example 1.5 European put

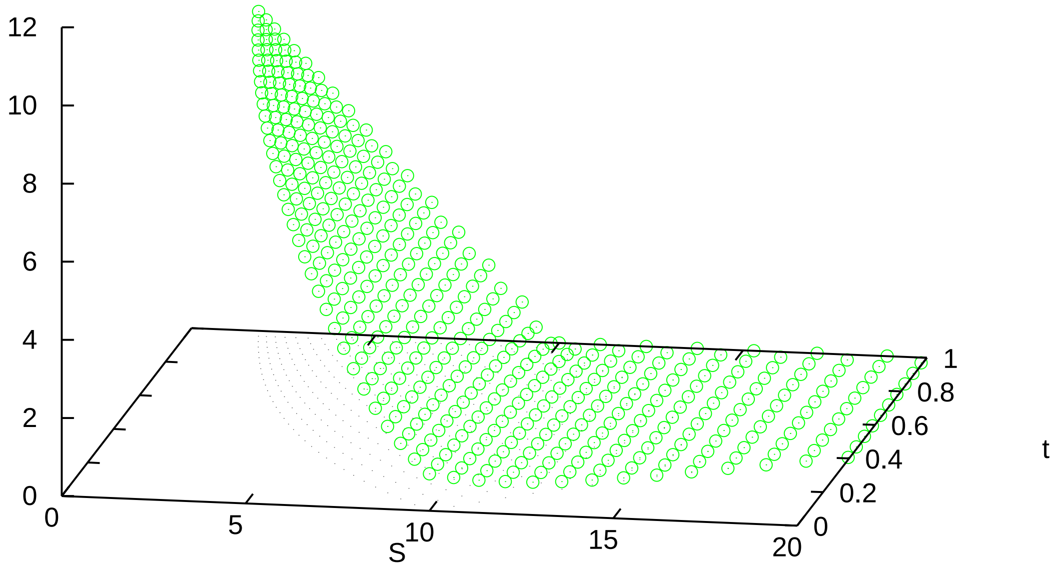
$K = 10,\ S = 5,\ r = 0.06,\ \sigma = 0.3,\ T = 1.$

approximations $V^{(M)}$ to $V(5,0)$

M	$V^{(M)}(5,0)$
8	4.42507
16	4.42925
32	4.429855
64	4.429923
128	4.430047
256	4.430390
2048	4.430451
Black-Scholes	4.43046477621



Example 1.6 American put



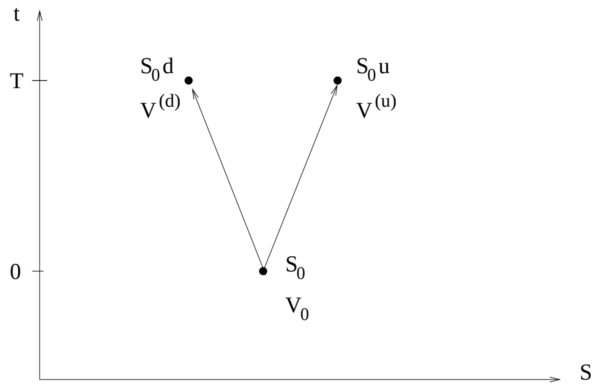
Extensions

dividends: If dividends are paid at t_k the price of the asset drops by the same amount. To take into account this jump, the tree is cut at t_k and the S-values are reduced appropriately.

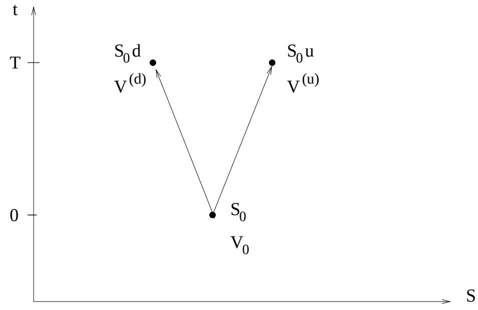
trinomial model: three outcomes, with probabilities $p_1,\ p_2,\ p_3$ and $p_1 + p_2 + p_3 = 1$.

1.5 Risk-Neutral Valuation

The situation of a path-independent binomial process with the two factors u and d $(0 < d < u)$ continues to be the basis of the argumentation.



one-period model: The time period is the time to expiration T. The one-period model has two clearly defined values of the payoff, namely $V^{(d)}$ (corresponds to $S_T = S_0 d$) and $V^{(u)}$ (corresponds to $S_T = S_0 u$). In contrast to the Assumptions 1.3 we neither assume the risk-neutral world (Bi3) nor the corresponding probability $\mathsf{P}(\text{up}) = p$ from (Bi2). Instead we derive the probability using another argument. In this section the factors u and d are assumed to be given.



Construct a portfolio of an investor with a short position in one option and a long position consisting of Δ shares of an asset, where the asset is the underlying of the option. The portfolio manager must **choose the number Δ of shares such that the portfolio is riskless.** (hedging strategy)

Π_t denotes the wealth of this portfolio at time t. Initially,

$$\Pi_0 = S_0 \cdot \Delta - V_0 \ , \tag{1.15}$$

where the value V_0 of the written option is not yet determined. At the end of the period the value V_T either takes the value $V^{(u)}$ or the value $V^{(d)}$. So the value of the portfolio Π_T is either

$$\Pi^{(u)} = S_0 u \cdot \Delta - V^{(u)}$$

or

$$\Pi^{(d)} = S_0 d \cdot \Delta - V^{(d)} \ .$$

$$\Pi^{(u)} = S_0 u \cdot \Delta - V^{(u)}$$

or

$$\Pi^{(d)} = S_0 d \cdot \Delta - V^{(d)} \; .$$

In case Δ is chosen such that the value Π_T is riskless, all uncertainty is removed and $\Pi^{(u)} = \Pi^{(d)}$ must hold. This is equivalent to

$$(S_0 u - S_0 d) \cdot \Delta = V^{(u)} - V^{(d)} \; ,$$

which defines the strategy

$$\Delta = \frac{V^{(u)} - V^{(d)}}{S_0(u-d)} \; . \tag{1.16}$$

With this value of Δ the portfolio with initial value Π_0 evolves to the final value $\Pi_T = \Pi^{(u)} = \Pi^{(d)}$, regardless of whether the stock price moves up or down. Consequently the portfolio is riskless.

Ruling out early exercise, the final value Π_T is reached with certainty. The value Π_T must be compared to the **alternative risk-free investment** of an amount of money that equals the initial wealth Π_0, which after the time period T reaches the value $e^{rT}\Pi_0$.

By arbitrage arguments, both portfolios must be equal: the initial value Π_0 of the portfolio equals the discounted final value Π_T, discounted at the interest rate r,

$$\Pi_0 = e^{-rT}\Pi_T \ .$$

This means

$$S_0 \cdot \Delta - V_0 = e^{-rT}(S_0 u \cdot \Delta - V^{(u)}) \ ,$$

which upon substituting (1.16) leads to the value V_0 of the option:

$$V_0 = e^{-rT}\{V^{(u)}q + V^{(d)} \cdot (1-q)\}$$

with

$$q := \frac{e^{rT} - d}{u - d} \ . \tag{1.17}$$

We have shown

$$q = \frac{e^{rT} - d}{u - d} \quad \Longrightarrow \quad V_0 = e^{-rT}\{V^{(u)}q + V^{(d)} \cdot (1 - q)\} \,. \tag{1.18}$$

The expression for q is identical to the formula for p in (1.6), which was derived in the previous section. Note again

$$0 < q < 1 \quad \Longleftrightarrow \quad d < e^{rT} < u \,.$$

This condition is equivalent to ruling out arbitrage in the model. Presuming these bounds, q can be interpreted as a probability Q. Then $qV^{(u)} + (1-q)V^{(d)}$ is the expected value of the payoff with respect to this probability,

$$\mathsf{E}_{\mathsf{Q}}(V_T) = qV^{(u)} + (1-q)V^{(d)} \,.$$

Now (1.18) can be written

$$V_0 = e^{-rT}\mathsf{E}_{\mathsf{Q}}(V_T) \,. \tag{1.19}$$

That is, the value of the option is obtained by discounting the expected payoff (with respect to q) at the risk-free interest rate r. An analogous calculation shows

$$\mathsf{E}_{\mathsf{Q}}(S_T) = qS_0u + (1-q)S_0d = S_0e^{rT} \,.$$

The probabilities p of Section 1.4 and q from (1.17) are defined by identical formulas (with T corresponding to Δt). Hence $p = q$, and $\mathsf{E}_\mathsf{P} = \mathsf{E}_\mathsf{Q}$. But the underlying arguments are different. Recall that in Section 1.4 we showed the implication

$$\mathsf{E}(S_T) = S_0 e^{rT} \quad \Longrightarrow \quad p = \mathsf{P}(\text{up}) = \frac{e^{rT} - d}{u - d} \ ,$$

whereas in this section we arrive at the implication

$$p = \mathsf{P}(\text{up}) = \frac{e^{rT} - d}{u - d} \quad \Longrightarrow \quad \mathsf{E}(S_T) = S_0 e^{rT} \ .$$

So both statements must be equivalent. Setting the probability of the up movement equal to p is equivalent to assuming that the expected return on the asset equals the risk-free rate. This can be rewritten as

$$e^{-rT} \mathsf{E}_\mathsf{P}(S_T) = S_0 \ . \tag{1.20}$$

This is the important property of a **martingale**: The random variable $e^{-rT} S_T$ of the left-hand side has the tendency to remain at the same level. (“fair game”) A martingale displays no trend, where the trend is measured with respect to E_P. In the martingale property of (1.20) the discounting at the risk-free interest rate r exactly matches the risk-neutral probability $\mathsf{P}(= \mathsf{Q})$ of (1.6)/(1.17). The specific probability for which (1.20) holds is also called *martingale measure*.

Summary of results for the one-period model:

Under the Assumptions 1.2 of the market model, the choice $\Delta = \frac{V^{(u)} - V^{(d)}}{S_0(u-d)}$ eliminates the random-dependence of the payoff and makes the portfolio riskless. There is a specific probability Q $(= \mathsf{P})$ with $\mathsf{Q}(\text{up}) = q$, $q := \frac{e^{rT} - d}{u-d}$, such that the value V_0 satisfies $V_0 = e^{-rT}\mathsf{E}_{\mathsf{Q}}(V_T)$ and S_0 the analogous property $e^{-rT}\mathsf{E}_{\mathsf{P}}(S_T) = S_0$. These properties involve the risk-neutral interest rate r. That is, the option is valued in a risk-neutral world, and the corresponding Assumption 1.3 (Bi3) is meaningful.

The $\Delta = \frac{V^{(u)} - V^{(d)}}{S_0(u-d)}$ is the hedge parameter **delta**, which eliminates the risk exposure of our portfolio caused by the written option. In multi-period models and continuous models Δ must be adapted dynamically. The general definition is

$$\Delta = \Delta(S,t) = \frac{\partial V(S,t)}{\partial S} \ .$$

1.6 Stochastic Processes

Brown (1827): erratic motion of a particle

Bachelier (1900): applied Brownian motion to model the motion of stock prices.

Einstein (1905)

Wiener (1923): mathematical model for this motion

A *stochastic process* is a family of random variables X_t, which are defined for a set of parameters t. For the *time-continuous* situation, $t \in \mathbb{R}$ varies continuously in a time interval I, typically $0 \leq t \leq T$. Let the chance play, then the resulting function X_t is called *realization* or *path* of the stochastic process.

Gaussian process: All finite-dimensional distributions $(X_{t_1}, \ldots, X_{t_k})$ are Gaussian. Hence specifically X_t is distributed normally for all t.

Markov process: Only the present value of X_t is relevant for its future motion. That is, the past history is fully reflected in the present value.

An example of a process that is both Gaussian and Markov, is the Wiener process.

Wiener Process

Definition 1.7 (Wiener process, standard Brownian motion)

A Wiener process (or standard Brownian motion; notation W_t or W) is a time-continuous process with the properties

(a) $W_0 = 0$

(b) $W_t \sim \mathcal{N}(0, t)$ for all $t \geq 0$.

That is, W_t is normally distributed with $\mathsf{E}(W_t) = 0$ and $\mathsf{Var}(W_t) = \mathsf{E}(W_t^2) = t$.

(c) All increments $\Delta W_t := W_{t+\Delta t} - W_t$ on non-overlapping time intervals are independent.

That is, the displacements $W_{t_2} - W_{t_1}$ and $W_{t_4} - W_{t_3}$ are independent for all $0 \leq t_1 < t_2 \leq t_3 < t_4$.

(d) W_t depends continuously on t.

Generally $W_t - W_s \sim \mathcal{N}(0, t - s)$ holds,

$$\mathsf{E}(W_t - W_s) = 0 \ , \quad \mathsf{Var}(W_t - W_s) = \mathsf{E}((W_t - W_s)^2) = t - s. \qquad (1.21\text{a, b})$$

These relations can be derived from Definition 1.7. The second is also known as

$$\mathsf{E}((\Delta W_t)^2) = \Delta t \ . \qquad (1.21\text{c})$$

The independence of the increments according to Definition 1.7(c) implies for $t_{j+1} > t_j$ the independence of W_{t_j} and $(W_{t_{j+1}} - W_{t_j})$, but not of $W_{t_{j+1}}$ and $(W_{t_{j+1}} - W_{t_j})$.

Discrete-Time Model

Let $\Delta t > 0$ be a constant time increment. For the discrete instances $t_j := j\Delta t$ the value W_t can be written as

$$W_{j\Delta t} = \sum_{k=1}^{j} \underbrace{\left(W_{k\Delta t} - W_{(k-1)\Delta t}\right)}_{=:\Delta W_k} .$$

The ΔW_k are independent and because of (1.21) normally distributed with $\mathsf{Var}(\Delta W_k) = \Delta t$. Increments ΔW with such a distribution can be calculated from standard normally distributed random numbers Z. The implication

$$Z \sim \mathcal{N}(0,1) \implies Z \cdot \sqrt{\Delta t} \sim \mathcal{N}(0, \Delta t)$$

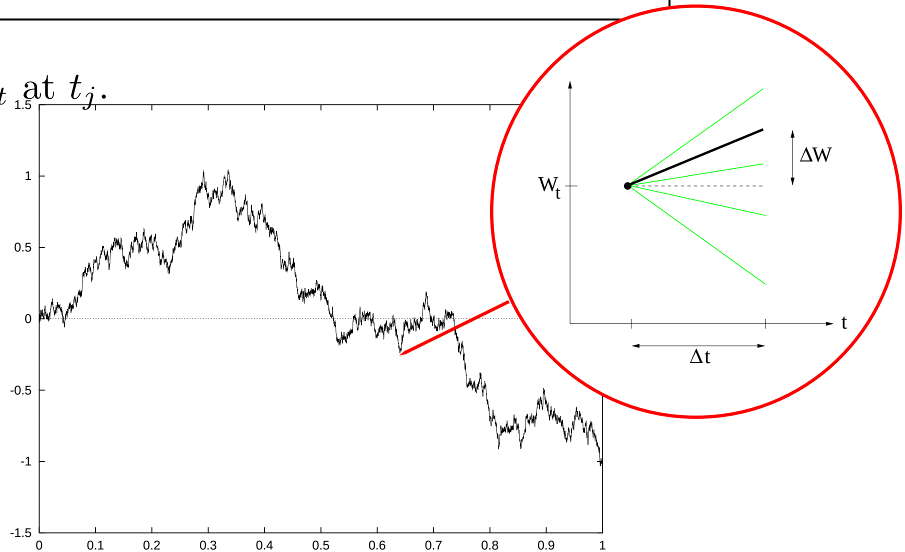
leads to the discrete model of a Wiener process

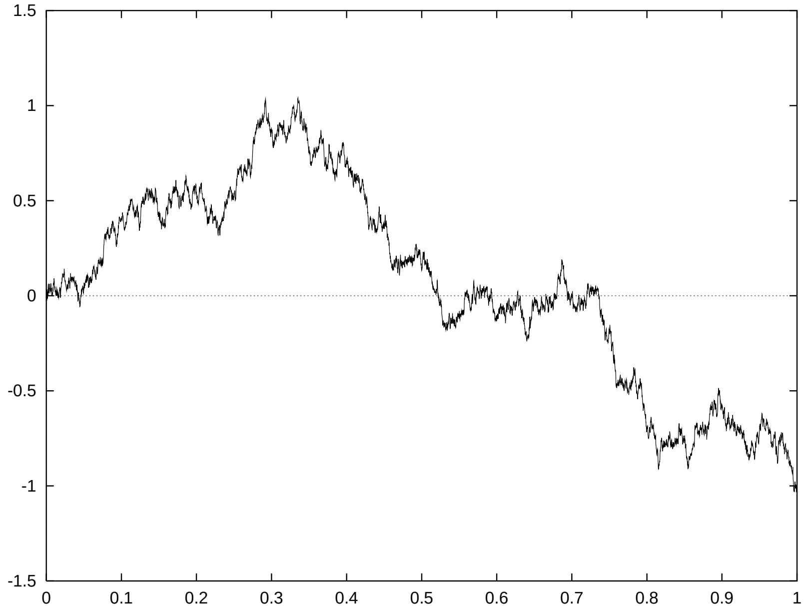
$$\Delta W_k = Z\sqrt{\Delta t} \;\text{ for }\; Z \sim \mathcal{N}(0,1) \text{ for each } k \;. \tag{1.22}$$

Algorithm 1.8 (simulation of a Wiener process)

$$
\begin{aligned}
&\textit{Start:} \quad t_0 = 0,\; W_0 = 0;\; \Delta t \\
&\textit{loop} \;\; j = 1, 2, \ldots : \\
&\qquad\qquad t_j = t_{j-1} + \Delta t \\
&\qquad\qquad \text{draw} \;\; Z \sim \mathcal{N}(0,1) \\
&\qquad\qquad W_j = W_{j-1} + Z\sqrt{\Delta t}
\end{aligned}
$$

W_j is a realization of W_t at t_j.





$\Delta t = 0.0002$

Almost all realizations of Wiener processes are *nowhere differentiable*. This becomes intuitively clear when the difference quotient

$$\frac{\Delta W_t}{\Delta t} = \frac{W_{t+\Delta t} - W_t}{\Delta t}$$

is considered. Because of relation (1.21b) the standard deviation of the numerator is $\sqrt{\Delta t}$. Hence for $\Delta t \to 0$ the normal distribution of the difference quotient disperses and no convergence can be expected.

Stochastic Integral

Suppose that the price development of an asset is described by a Wiener process W_t. Let $b(t)$ be the number of units of the asset held in a portfolio at time t. Start with the simplifying assumption that trading is only possible at discrete time instances t_j, which define a partition of the interval $0 \leq t \leq T$. Then the trading strategy b is piecewise constant,

$$\begin{aligned} & b(t) = b(t_{j-1}) \quad \text{for} \quad t_{j-1} \leq t < t_j \\ & \text{and} \quad 0 = t_0 < t_1 < \ldots < t_N = T \ . \end{aligned} \tag{1.23}$$

(step function)

The trading gain for the subinterval $t_{j-1} \leq t < t_j$ is given by $b(t_{j-1})(W_{t_j} - W_{t_{j-1}})$, and

$$\sum_{j=1}^{N} b(t_{j-1})(W_{t_j} - W_{t_{j-1}}) \tag{1.24}$$

represents the trading gain over the time period $0 \leq t \leq T$. The trading gain (possibly < 0) is determined by the strategy $b(t)$ and the price process W_t.

Now drop the assumption of fixed trading times t_j and allow b to be arbitrary continuous functions. This leads to the question whether (1.24) has a limit when with $N \to \infty$ the size of the subintervals tends to 0. If W_t would be of bounded variation than the limit exists and is called *Riemann-Stieltjes integral*

$$\int_0^T b(t)dW_t \ .$$

But this integral generally does not exist because almost all Wiener processes are not of bounded variation. That is, the *first variation* of W_t, which is the limit of

$$\sum_{j=1}^{N} |W_{t_j} - W_{t_{j-1}}| \ ,$$

is unbounded even in case the lengths of the subintervals vanish for $N \to \infty$.

important assertion $(dW_t)^2 = dt$.

For an arbitrary partition of the interval $[0, T]$ into N subintervals the inequality

$$\sum_{j=1}^{N} |W_{t_j} - W_{t_{j-1}}|^2 \le \max_j(|W_{t_j} - W_{t_{j-1}}|) \sum_{j=1}^{N} |W_{t_j} - W_{t_{j-1}}| \qquad (1.25)$$

holds. The left-hand sum is the *second variation* and the right-hand sum the first variation of W for a given partition into subintervals.

The expectation of the left-hand sum is calculated using properties of Wiener's process,

$$\sum_{j=1}^{N} \mathsf{E}(W_{t_j} - W_{t_{j-1}})^2 = \sum_{j=1}^{N} (t_j - t_{j-1}) = t_N - t_0 = T \ .$$

Even convergence in the mean holds:

Lemma 1.9 (second variation: convergence in the mean)

Let $t_0 = t_0^{(N)} < t_1^{(N)} < \ldots < t_N^{(N)} = T$ be a sequence of partitions of the interval $t_0 \le t \le T$ with $\delta_N := \max_j(t_j^{(N)} - t_{j-1}^{(N)})$. Then (dropping the $^{(N)}$)

$$\underset{\delta_N \to 0}{\text{l.i.m.}} \sum_{j=1}^{N} (W_{t_j} - W_{t_{j-1}})^2 = T - t_0 \qquad (1.27)$$

Proof: The statement (1.27) means convergence in the mean. Because of $\sum \Delta t_j = T - t_0$ we show

$$\mathsf{E}\left(\sum_j ((\Delta W_j)^2 - \Delta t_j)\right)^2 \to 0 \quad \text{for} \quad \delta_N \to 0 \; .$$

Carrying out the multiplications and taking the mean gives $2\sum_j (\Delta t_j)^2$ This can be bounded by $2(T - t_0)\delta_N$, which completes the proof.

Part of the derivation can be summarized to

$$\mathsf{E}((\Delta W_t)^2 - \Delta t) = 0 \quad , \quad \mathsf{Var}((\Delta W_t)^2 - \Delta t) = 2(\Delta t)^2 \; ,$$

hence $(\Delta W_t)^2 \approx \Delta t$. This property of a Wiener process is symbolically written

$$\boxed{(dW_t)^2 = dt} \tag{1.28}$$

It will be needed in subsequent sections.

Now turn to the right-hand side of inequality

$$\sum_{j=1}^{N} |W_{t_j} - W_{t_{j-1}}|^2 \leq \max_j (|W_{t_j} - W_{t_{j-1}}|) \sum_{j=1}^{N} |W_{t_j} - W_{t_{j-1}}|$$

The continuity of W_t implies

$$\max_j |W_{t_j} - W_{t_{j-1}}| \to 0 \qquad \text{for} \quad \delta_N \to 0 \ .$$

Convergence in the mean shows that the vanishing of this factor must be compensated by an unbounded growth of the other factor, so

$$\sum_{j=1}^{N} |W_{t_j} - W_{t_{j-1}}| \to \infty \quad \text{für} \quad \delta_N \to 0 \ .$$

In summary, Wiener processes are not of bounded variation, and the integration with respect to W_t can not be defined as an elementary limit of $\sum_{j=1}^{N} b(t_{j-1})(W_{t_j} - W_{t_{j-1}})$.

Construct a Stochastic Integral

$$\int_{t_0}^{t} f(s)dW_s$$

for general stochastic integrands $f(t)$. (sketch the Itô integral)

For a step function b an integral can be defined as

$$\int_{t_0}^{t} b(s)dW_s := \sum_{j=1}^{N} b(t_{j-1})(W_{t_j} - W_{t_{j-1}}) \ . \tag{1.29}$$

(Itô integral over a step function b) In case the $b(t_{j-1})$ are random variables, b is called a *simple process*. Then the Itô integral is again defined by (1.29). Stochastically integrable functions f can be obtained as limits of simple processes b_n in the sense

$$\mathsf{E}\left[\int_{t_0}^{t} (f(s) - b_n(s))^2 ds\right] \to 0 \quad \text{for} \quad n \to \infty \ . \tag{1.30}$$

Convergence in terms of integrals $\int ds$ carries over to integrals $\int dW_t$. This is achieved by applying Cauchy convergence $\mathsf{E}\int(b_n - b_m)^2 ds \to 0$ and the *isometry*

$$\mathsf{E}\left[\left(\int_{t_0}^{t} b(s)dW_s\right)^2\right] = \mathsf{E}\left[\int_{t_0}^{t} b(s)^2 ds\right] \ .$$

Hence the integrals $\int b_n(s)dW_s$ form a Cauchy sequence with respect to convergence in the mean. Accordingly the Itô integral of f is defined as

$$\int_{t_0}^{t} f(s)dW_s := \text{l.i.m.}_{n\to\infty} \int_{t_0}^{t} b_n(s)dW_s \ ,$$

for simple processes b_n defined by (1.30). The value of the integral is independent of the choice of the b_n in (1.30). The Itô integral as function in t is a stochastic process with the martingale property.

If an integrand $a(x,t)$ depends on a stochastic process X_t, the function f is given by $f(t) = a(X_t, t)$. For the simplest case of a constant integrand $a(X_t, t) = a_0$ the Itô integral can be reduced to a Riemann-Stieltjes integal

$$\int_{t_0}^{t} dW_s = W_t - W_{t_0}.$$

For the "first" nontrivial Itô integral consider $X_t = W_t$ and $a(W_t, t) = W_t$.

1.7 Stochastic Differential Equations

$$x(t) = x_0 + \int_{t_0}^{t} a(x(s), s)ds + \text{ randomness} \,,$$

The integral in this integral equation is an ordinary (Lebesgue- or Riemann-) integral.

The randomness here is modeled by a stochastic integral with respect to a Wiener process. A stochastic process is denoted by X_t or S_t.

$$X_t = X_0 + \int_{t_0}^{t} a(X_s, s)ds + \int_{t_0}^{t} b(X_s, s)dW_s$$

This equation is named after Itô.

Definition 1.10 (Itô stochastic differential equation)

An Itô stochastic differential equation is

$$dX_t = a(X_t, t)dt + b(X_t, t)dW_t \,; \tag{1.31a}$$

this together with $X_{t_0} = X_0$ is a symbolic short form of the integral equation

$$X_t = X_0 + \int_{t_0}^{t} a(X_s, s)ds + \int_{t_0}^{t} b(X_s, s)dW_s \,. \tag{1.31b}$$

The terms in (1.31) are named as follows:

$a(X_t, t)$: drift term or drift coefficient

$b(X_t, t)$: diffusion term

solution X_t: Itô process, or stochastic diffusion

A Wiener process is a special case of an Itô process, because from $X_t = W_t$ the trivial SDE $dX_t = dW_t$ follows, hence $a = 0$ and $b = 1$.

The simplest numerical method combines the discretized version of the Itô SDE

$$\Delta X_t = a(X_t, t)\Delta t + b(X_t, t)\Delta W_t \tag{1.32}$$

with the Algorithm 1.8 for approximating a Wiener process, using the same Δt for both discretizations. The result is

Algorithm 1.11 (Euler discretization of an SDE)

Approximations y_j to X_{t_j} are calculated by

Start: $t_0,\ y_0 = X_0,\ \Delta t,\ W_0 = 0.$

loop $j = 0, 1, 2, ...$

$$t_{j+1} = t_j + \Delta t$$

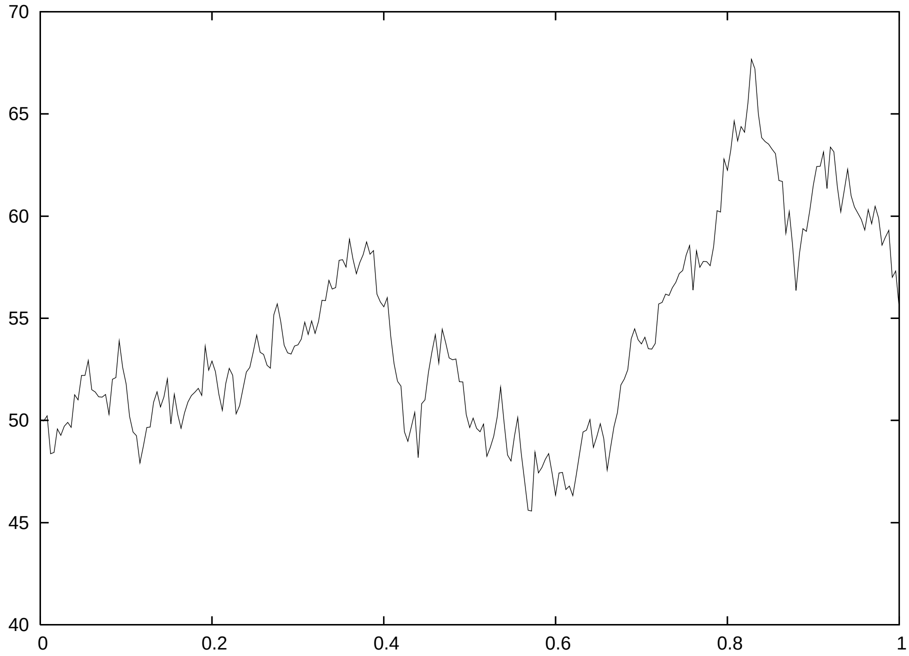
$$\Delta W = Z\sqrt{\Delta t} \quad \text{with} \quad Z \sim \mathcal{N}(0,1)$$

$$y_{j+1} = y_j + a(y_j, t_j)\Delta t + b(y_j, t_j)\Delta W$$

For example, the *step length* Δt is chosen equidistant, $\Delta t = T/m$ for a suitable integer m. Solutions are called *trajectories* or *paths*. By *simulation* of the SDE we understand the calculation of one or more trajectories.

Example 1.12 $dX_t = 0.05X_t\,dt + 0.3X_t\,dW_t$

Without the diffusion term the exact solution would be $X_t = X_0 e^{0.05t}$. For $X_0 = 50$, $t_0 = 0$ and a time increment $\Delta t = 1/250$ the figure depicts a trajectory X_t of the SDE.



Application to the Stock Market

Samuelson (1965/1970): continuous model for motions of the prices S_t of stocks.

This standard model assumes that the relative change (return) dS/S of a security in the time interval dt is composed of a deterministic drift term μ plus stochastic fluctuations in the form σdW_t:

Model 1.13 (geometric Brownian motion)

$$\boxed{dS_t = \mu S_t \, dt + \sigma S_t \, dW_t .} \tag{1.33–GBM}$$

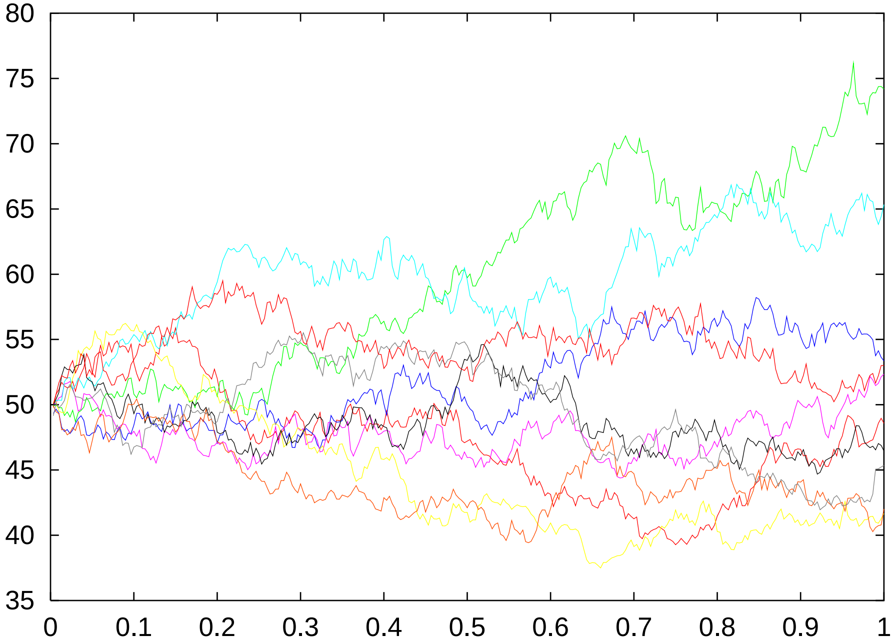
This SDE is linear in $X_t = S_t$. The drift rate is $a(S_t, t) = \mu S_t$, with the expected *rate of return* μ, and $b(S_t, t) = \sigma S_t$, with volatility σ. GMB is the reference model on which the Black–Scholes approach is based. According to Assumption 1.2 μ and σ are assumed constant.

The deterministic part of (GMB) is the ordinary differential equation

$$\dot{S} = \mu S$$

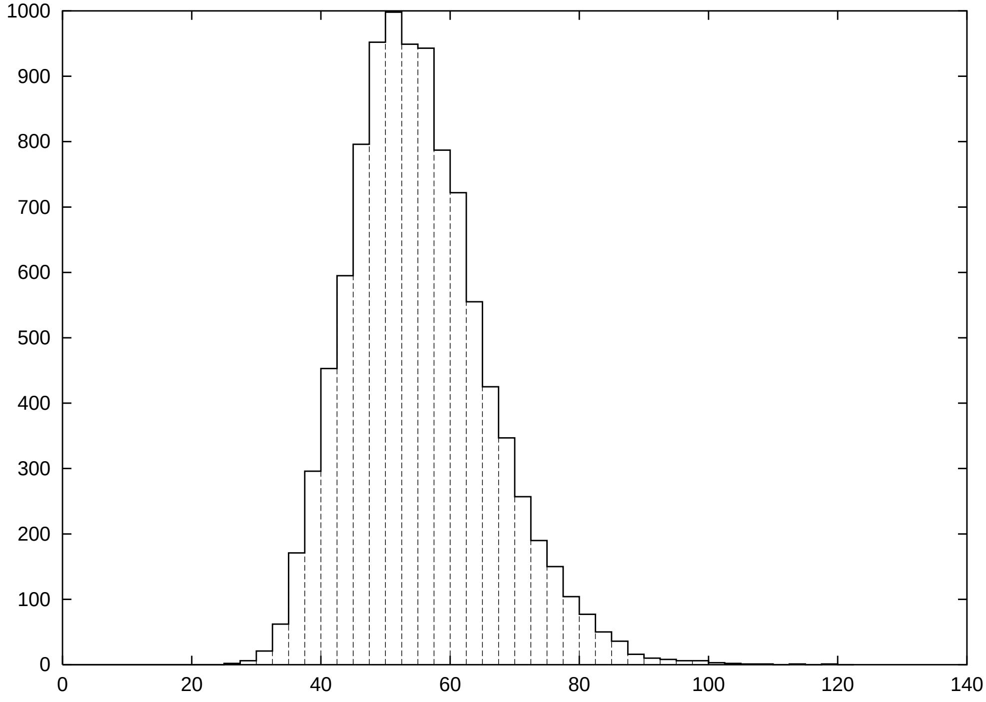
with solution $S_t = S_0 e^{\mu(t-t_0)}$. For the linear SDE of (GMB) the expectation $\mathsf{E}(S_t)$ solves $\dot{S} = \mu S$. Hence $S_0 e^{\mu(t-t_0)}$ is the expectation of the stochastic process and μ is the expected *growth rate*.

The simulated values S_1 of the ten trajectories in the figure group around the value $50 \cdot e^{0.1} \approx 55.26$.



$S_0 = 50$
$\mu = 0.1$
$\sigma = 0.2$

Empirical distribution of the values S_1 about their expected value. For 10000 trajectories count how many of the terminal values S_1 fall into the subintervals $k5 \leq t < (k+1)5$, for $k = 0, 1, 2 \ldots$. The resulting histogram has a skewed distribution.



The discrete version of (GBM) is

$$\frac{\Delta S}{S} = \mu \Delta t + \sigma Z \sqrt{\Delta t}, \qquad (1.34a)$$

compare Algorithm 1.11. The ratio $\frac{\Delta S}{S}$ is called one-period *simple return*, where Δt is interpreted as one period. It satisfies

$$\frac{\Delta S}{S} \sim \mathcal{N}(\mu \Delta t, \sigma^2 \Delta t). \qquad (1.34b)$$

Provided the data match the GBM assumption, this distribution allows to calculate estimates of historical values of the volatility σ. The approximation is valid as long as Δt is small.

For modeling of $V(S_t, t)$, a risk-neutral world is assumed which leads to replace μ by the risk-free rate r.

Risk-Neutral Valuation

$$\begin{aligned} \mathrm{d}S &= \mu S\,\mathrm{d}t + \sigma S\,\mathrm{d}W \\ &= rS\,\mathrm{d}t + (\mu - r)S\,\mathrm{d}t + \sigma S\,\mathrm{d}W \\ &= rS\,\mathrm{d}t + \sigma S\,[\gamma \mathrm{d}t + \mathrm{d}W] \end{aligned}$$

with $\gamma := \frac{\mu - r}{\sigma}$.

Girsanov: For suitable γ (e.g. γ constant) there is a probability Q such that

$$W_t^\gamma := W_t + \int_0^t \gamma \mathrm{d}s$$

is a (standard) Wiener process under Q.

The change of drift $\mu \to r$, whith $W_t \to W_t^\gamma$, adjusts the probability P to Q: With respect to Q, the discounted $e^{-rt}S_t$ is martingale.

Then, by the *fundamental theorem of asset pricing*, the market model is free of arbitrage.

Remark 1.14 (risk-neutral valuation principle)

For modeling options under GBM, the return rate μ is replaced by the risk-free interest rate r.

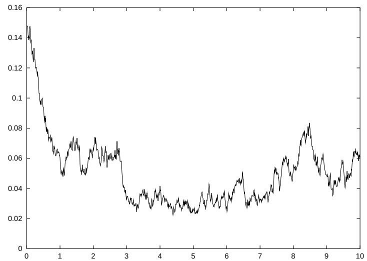
Mean Reversion

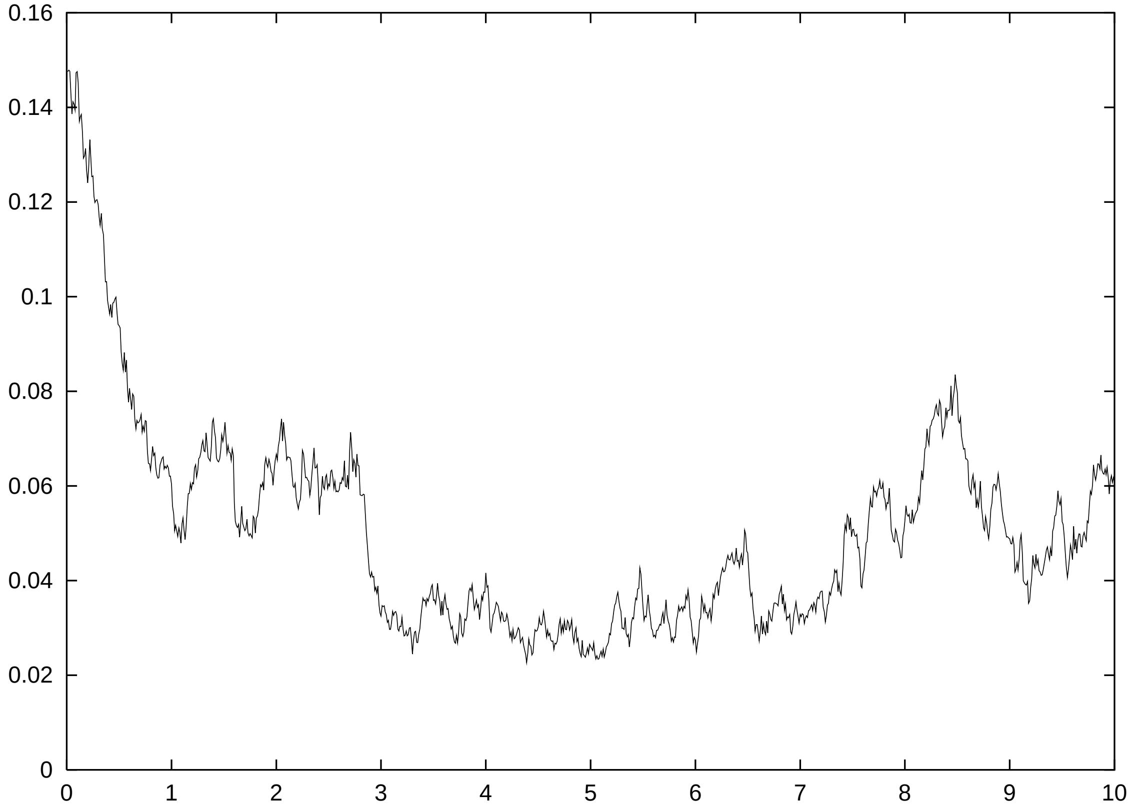
The assumptions of a constant interest rate r and a constant volatility σ are quite restrictive. SDEs for r_t and σ_t have been constructed that control r_t or σ_t stochastically. A class of models is given by the SDE for the process r_t,

$$dr_t = \alpha(R - r_t)dt + \sigma_{\mathrm{r}} r_t^{\beta} dW_t \,, \;\; \alpha > 0. \qquad (1.40)$$

The drift term $\alpha(R - r_t)$ is positive for $r_t < R$ and negative for $r_t > R$, which causes a pull to R. This effect is called *mean reversion*. The parameter R, which may depend on t, corresponds to a long-run mean of the interest rate over time.

For $\beta = 0$ (constant volatility) the SDE specializes to the Vasicek model. The Cox-Ingersoll-Ross model is obtained for $\beta = \frac{1}{2}$. Then the volatility $\sigma_{\mathrm{r}}\sqrt{r_t}$ vanishes when r_t tends to zero. Provided $r_0 > 0$, $R > 0$, this guarantees $r_t \geq 0$.





A simulation r_t of the Cox-Ingersoll-Ross model for $R = 0.05$, $\alpha = 1$, $\beta = 0.5$, $y_0 = 0.15$, $\Delta t = 0.01$

The SDE (1.40) is of a different kind as (GBM). Coupling the SDE for r_t to that for S_t leads to a *system* of two SDEs. Even larger systems are obtained when further SDEs are coupled to define a stochasic process R_t or to calculate stochastic volatilities. Related examples are given by Examples 1.15, 1.16 (Heston's model) below.

Vector-Valued SDEs

The Itô equation (1.31) is formulated as scalar equation; accordingly the SDE (GBM) is a *one-factor model*. The general *multi-factor* version can be written in the same notation. Then $X_t = (X_t^{(1)}, \ldots, X_t^{(n)})$ and $a(X_t, t)$ are n-dimensional vectors. The Wiener process can be m-dimensional, with components $W_t^{(1)}, ..., W_t^{(m)}$. Then $b(X_t, t)$ is an $(n \times m)$-matrix. The interpretation of the SDE systems is componentwise. The scalar stochastic integrals are sums of m stochastic integrals,

$$X_t^{(i)} = X_0^{(i)} + \int_{t_0}^{t} a_i(X_s, s)ds + \sum_{k=1}^{m} \int_{t_0}^{t} b_{ik}(X_s, s)dW_s^{(k)},$$

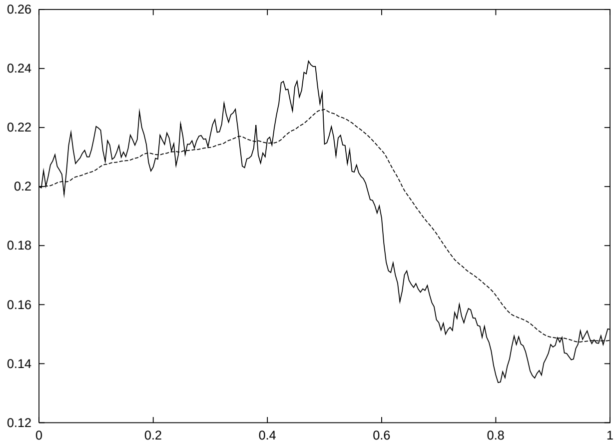
for $i = 1, ..., n$.

Example 1.15 (mean-reverting volatility)

three-factor model with stock price S_t, instantaneous spot volatility σ_t and an averaged volatility ζ_t serving as mean-reverting "parameter":

$$\begin{cases} dS = \sigma S dW^{(1)} \\ d\sigma = -(\sigma - \zeta)dt + \alpha\sigma dW^{(2)} \\ d\zeta = \beta(\sigma - \zeta)dt \end{cases}$$

The stochastic volatility σ follows the mean volatility ζ and is simultaneously perturbed by a Wiener process. $\longrightarrow$ tandem



$\alpha = 0.3$
$\beta = 10$
$\sigma_0 = \zeta_0 = 0.2$
σ_t
ζ_t (dashed)
$\Delta t = 0.004$

Example 1.16 (Heston's model)

Heston [Hes93] uses an Ornstein–Uhlenbeck process to model a stochastic volatility σ_t. Then the variance $v_t := \sigma_t^2$ follows a Cox–Ingersoll–Ross process:

$$\begin{aligned} \mathrm{d}S_t &= \mu S_t \,\mathrm{d}t + \sqrt{v_t}\, S_t \,\mathrm{d}W_t^{(1)} \\ \mathrm{d}v_t &= \kappa(\theta - v_t)\,\mathrm{d}t + \sigma_{\mathrm{v}}\sqrt{v_t}\,\mathrm{d}W_t^{(2)} \end{aligned} \tag{1.43}$$

with two correlated Wiener processes $W_t^{(1)}, W_t^{(2)}$ and suitable parameters μ, κ, θ, σ_{v}, ρ, where ρ is the correlation between $W_t^{(1)}, W_t^{(2)}$. Hidden parameters might be the initial values S_0, v_0, if not available.

This model establishes a correlation between price and volatility.

1.8 Itô Lemma and Implications

Itô's lemma is most fundamental for stochastic processes.

Lemma 1.17 (Itô)

Suppose X_t follows an Itô process, $dX_t = a(X_t, t)dt + b(X_t, t)dW_t$, and let $g(x, t)$ be a function with continuous $\frac{\partial g}{\partial x}$, $\frac{\partial^2 g}{\partial x^2}$, $\frac{\partial g}{\partial t}$. Then $Y_t := g(X_t, t)$ follows an Itô process with the *same* Wiener process W_t:

$$dY_t = \left(\frac{\partial g}{\partial x} a + \frac{\partial g}{\partial t} + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} b^2 \right) dt + \frac{\partial g}{\partial x} b \; dW_t \tag{1.44}$$

where the derivatives of g as well as the coefficient functions a and b in general depend on the arguments (X_t, t).

Sketch of a proof: When t varies by Δt, then X by $\Delta X = a \cdot \Delta t + b \cdot \Delta W$ and Y by $\Delta Y = g(X+\Delta X, t+\Delta t) - g(X, t)$. The Taylor expansion of ΔY begins with the linear part $\frac{\partial g}{\partial x} \Delta X + \frac{\partial g}{\partial t} \Delta t$, in which $\Delta X = a\Delta t + b\Delta W$ is substituted. The additional term with the derivative $\frac{\partial^2 g}{\partial x^2}$ is new and is introduced via the $O(\Delta X^2)$-term of the Taylor expansion. Because of $(\Delta W)^2 \approx \Delta t$, this term is also of the order $O(\Delta t)$ and belongs to the linear terms. Taking correct limits (as in Lemma 1.9) one obtains (1.44).

Consequences for Stocks and Options

Assume the stock price to follow (GBM), hence $X_t = S_t,\ a = \mu S_t,\ b = \sigma S_t,\ \mu, \sigma$ constant. The value V_t of an option depends on S_t. Assuming C^2-smoothness of V_t depending on S and t, apply Itô's lemma. For $V(S,t)$ in the place of $g(x,t)$ the result is

$$dV_t = \left(\frac{\partial V}{\partial S}\mu S_t + \frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\sigma^2 S_t^2 \right) dt + \frac{\partial V}{\partial S}\sigma S_t dW_t. \tag{1.45}$$

This SDE is used to derive the Black-Scholes equation.

As second application of Itô's lemma consider $Y_t = \log(S_t)$, viz $g(x,t) = \log(x)$. This leads to the linear SDE

$$d \log S_t = (\mu - \frac{1}{2}\sigma^2)dt + \sigma dW_t.$$

For this linear SDE the expectation $\mathsf{E}(Y_t)$ satisfies the deterministic part

$$\frac{d}{dt}\mathsf{E}(Y_t) = \mu - \frac{\sigma^2}{2}\ .$$

The solution of $\dot{y} = \mu - \frac{\sigma^2}{2}$ with initial condition $y(t_0) = y_0$ is

$$y(t) = y_0 + (\mu - \frac{\sigma^2}{2})(t - t_0).$$

In other words, the expectation of the Itô process Y_t is

$$\mathsf{E}(\log S_t) = \log S_0 + (\mu - \frac{\sigma^2}{2})(t - t_0) \ .$$

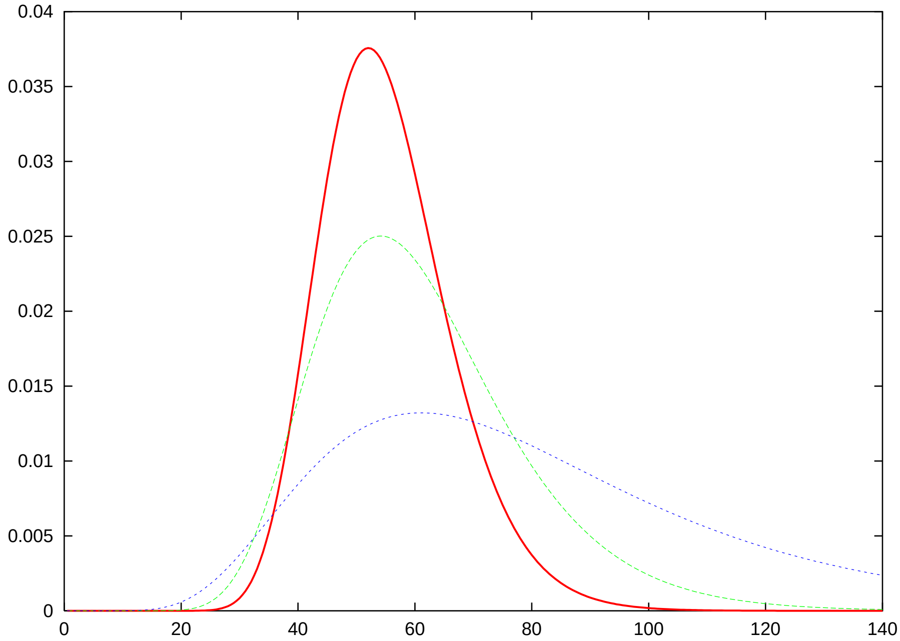
Analogously, we see from the differential equation for $\mathsf{E}(Y_t^2)$ (or from the analytical solution of the SDE for Y_t) that the variance of Y_t is $\sigma^2(t - t_0)$. The simple SDE for Y_t implies that the stochastic fluctuation of Y_t is that of σW_t. So Y_t is normally distributed, with density

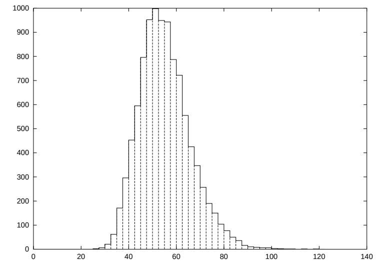
$$\widehat{f}(Y_t) := \frac{1}{\sigma\sqrt{2\pi(t-t_0)}} \exp\left\{ -\frac{\left(Y_t - y_0 - \left(\mu - \frac{\sigma^2}{2}\right)(t-t_0)\right)^2}{2\sigma^2(t-t_0)} \right\} .$$

Back transformation using $Y = \log(S)$ and considering $dY = \frac{1}{S}dS$ and $\widehat{f}(Y)dY = \frac{1}{S}\widehat{f}(\log S)dS = f(S)dS$ yields the density of S_t:

$$f(S; t - t_0, S_0) := \frac{1}{S\sigma\sqrt{2\pi(t-t_0)}} \exp\left\{ -\frac{\left(\log(S/S_0) - \left(\mu - \frac{\sigma^2}{2}\right)(t-t_0)\right)^2}{2\sigma^2(t-t_0)} \right\} \tag{1.48}$$

This is the density of the **lognormal distribution**. The stock price S_t is lognormally distributed under the basic assumption of (GBM). The distribution is skewed. Now the skewed behavior coming out of the experiment reported earlier is clear.





density over S
for $\mu = 0.1$,
$\sigma = 0.2$
$S_0 = 50$
$t_0 = 0$
$t = 1$ (red, solid curve)
$t = 2$ (green, dashed)
$t = 5$ (blue, flat)

Test the idealized Model 1.13 of GBM against actual **empirical data**. Suppose the time series $S_1, ..., S_M$ represents consecutive quotations of a stock price. To test the data, histograms of the returns are helpful. The transformation $y = \log(S)$ is most practical. It leads to the notion of the **log return**, defined by

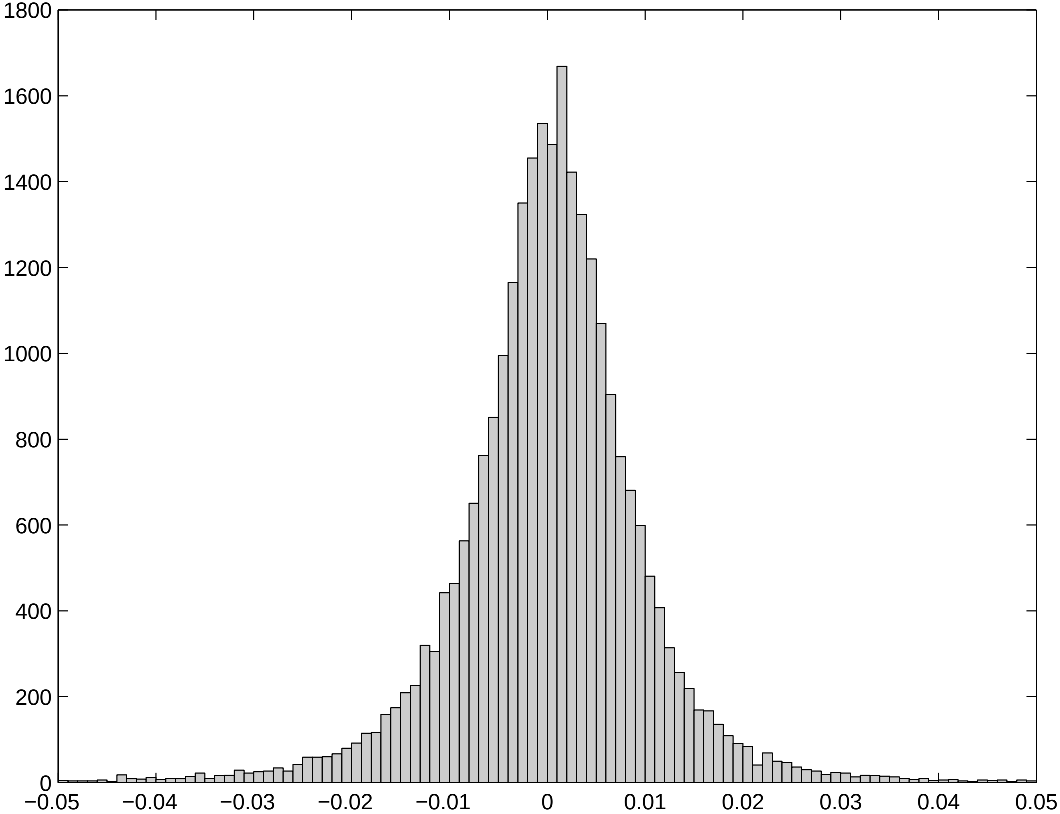
$$R_{i,i-1} := \log \frac{S_i}{S_{i-1}} \; .$$

Since $S_i = S_{i-1} \exp(R_{i,i-1})$, the log return is also called *continuously compounded return* in the ith time interval. Let Δt be the equally spaced sampling time interval between the quotations S_{i-1} and S_i, measured in years. Then (1.48) leads to

$$R_{i,i-1} \sim \mathcal{N}((\mu - \frac{\sigma^2}{2})\Delta t \; , \; \sigma^2 \Delta t) \; .$$

The sample variance $\sigma^2 \Delta t$ of the data allows to calculate estimates of the historical volatility σ.

But the tails of the data are not well modeled by the hypothesis of a geometric Brownian motion: The exponential decay expressed by (1.48) amounts to *thin tails*. This underestimates extreme events and hence does not match reality.



Histogram: frequency of daily log returns $R_{i,i-1}$ of the Dow in the time period 1901-1999.

analytical solution of the basic linear constant-coefficient SDE (GBM)

$$dS_t = \mu S_t \, dt + \sigma S_t \, dW_t$$

For an arbitrary Wiener process W_t set $X_t := W_t$ and apply Itô's lemma

$$Y_t = g(X_t, t) := S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma X_t\right)$$

From $X_t = W_t$ follows the trivial SDE with coefficients $a = 0$ and $b = 1$. By Itô's lemma

$$\begin{aligned} dY_t &= \left(\mu - \frac{\sigma^2}{2}\right) Y_t dt + \frac{\sigma^2}{2} Y_t dt + \sigma Y_t dW_t \\ &= \mu Y_t dt + \sigma Y_t dW. \end{aligned}$$

Consequently the process

$$S_t := S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W_t\right) \tag{1.54}$$

solves the linear constant-coefficient SDE (GBM).

1.9 Jump Processes

Rapid asset price movements can be modeled as jumps. Here we discuss Merton's **jump diffusion**, which is based on a Poisson process. One has to pay a price: With a jump process the risk of an option in general can not be hedged away to zero.

Denote the time instances for which a jump occurs τ_j, with

$$\tau_1 < \tau_2 < \tau_3 < \dots$$

Let the number of jumps be counted by the counting variable J_t, where

$$\tau_j = \inf\{t \geq 0 \ , \ J_t = j\}.$$

The probability that a jump occurs is introduced via a Bernoulli experiment. To this end, consider a subinterval of length $\Delta t := \frac{t}{n}$ and allow for two outcomes, jump *yes* or *no*, with the probabilities

$$\begin{aligned} \mathsf{P}(J_t - J_{t-\Delta t} = 1) &= \lambda \Delta t \\ \mathsf{P}(J_t - J_{t-\Delta t} = 0) &= 1 - \lambda \Delta t \end{aligned}$$

for some λ such that $0 < \lambda \Delta t < 1$. The parameter λ is the *intensity* of the jump process.

$$\begin{aligned} \mathsf{P}(J_t - J_{t-\Delta t} = 1) &= \lambda \Delta t \\ \mathsf{P}(J_t - J_{t-\Delta t} = 0) &= 1 - \lambda \Delta t \end{aligned}$$

Consequently k jumps in $0 \le \tau \le t$ have the probability

$$\mathsf{P}(J_t - J_0 = k) = \binom{n}{k} (\lambda \Delta t)^k (1 - \lambda \Delta t)^{n-k} \ ,$$

where the trials in each subinterval are considered independent. A little reasoning reveals that for $n \to \infty$ this probability converges to

$$\frac{(\lambda t)^k}{k!} e^{-\lambda t} \ ,$$

which is known as the Poisson distribution with parameter $\lambda > 0$.

Definition 1.19 (Poisson process)

The process $\{J_t \ , \ t \geq 0\}$ is called Poisson process if the following conditions hold:

(a) $J_0 = 0$

(b) $J_t - J_s$ are integer-valued for $0 \leq s < t < \infty$ and

$$\mathsf{P}(J_t - J_s = k) = \frac{\lambda^k (t-s)^k}{k!} e^{-\lambda(t-s)} \quad \text{for } k = 0, 1, 2 \ldots$$

(c) The increments $J_{t_2} - J_{t_1}$ and $J_{t_4} - J_{t_3}$ are independent for all $0 \leq t_1 < t_2 < t_3 < t_4$.

As consequence of this definition, several properties hold.

Properties 1.20 (Poisson process)

(d) J_t is right-continuous and non-decreasing.

(e) The times between successive jumps are independent and exponentially distributed with parameter λ: $\mathsf{P}(\tau_{j+1} - \tau_j > \Delta\tau) = e^{-\lambda \Delta\tau}$ for each $\Delta\tau$.

(f) J_t is a Markov process.

(g) $\mathsf{E}(J_t) = \lambda t$, $\mathsf{Var}(J_t) = \lambda t$

Simulating jumps: the instant τ_j

two possibilities to calculate jump instances τ_j such that the probabilities

$$\begin{aligned} \mathsf{P}(J_t - J_{t-\Delta t} = 1) &= \lambda \Delta t \\ \mathsf{P}(J_t - J_{t-\Delta t} = 0) &= 1 - \lambda \Delta t \end{aligned}$$

are met.

First, these probablities can be simulated using uniform deviates. In this way a Δt-discretization of a t-grid can be easily exploited to decide whether a jump occurs in a subinterval.

The other alternative is to calculate exponentially distributed random numbers $h_1, h_2, \ldots$ to simulate the intervals $\Delta \tau$ between consecutive jump instances, and set

$$\tau_{j+1} := \tau_j + h_j .$$

(See Chapter 2 for uniform deviates and exponentially distributed random numbers.)

Simulating jumps: the jump magnitude

In addition to the jump instances τ_j another random variable is required to simulate the jump *sizes*. The unit amplitudes of the jumps of the Poisson counting process J_t are not relevant for the purpose of establishing a market model. The jump sizes of the price of a financial asset will be considered random.

Let the random variable S_t jump at τ_j and denote τ^+ the moment after the jump and τ^- the moment before. Then the absolute size of the jump is

$$\Delta S = S_{\tau^+} - S_{\tau^-} \ ,$$

which we model as a **proportional jump**,

$$S_{\tau^+} = qS_{\tau^-} \quad \text{with } q > 0 \ . \tag{1.56}$$

So, $\Delta S = qS_{\tau^-} - S_{\tau^-} = (q-1)S_{\tau^-}$. The jump sizes equal $q-1$ times the current asset price. Accordingly, a jump process depends on a process q_t and is written

$$dS_t = (q_t - 1)S_t dJ_t \ , \text{ where } J_t \text{ is a Poisson process.}$$

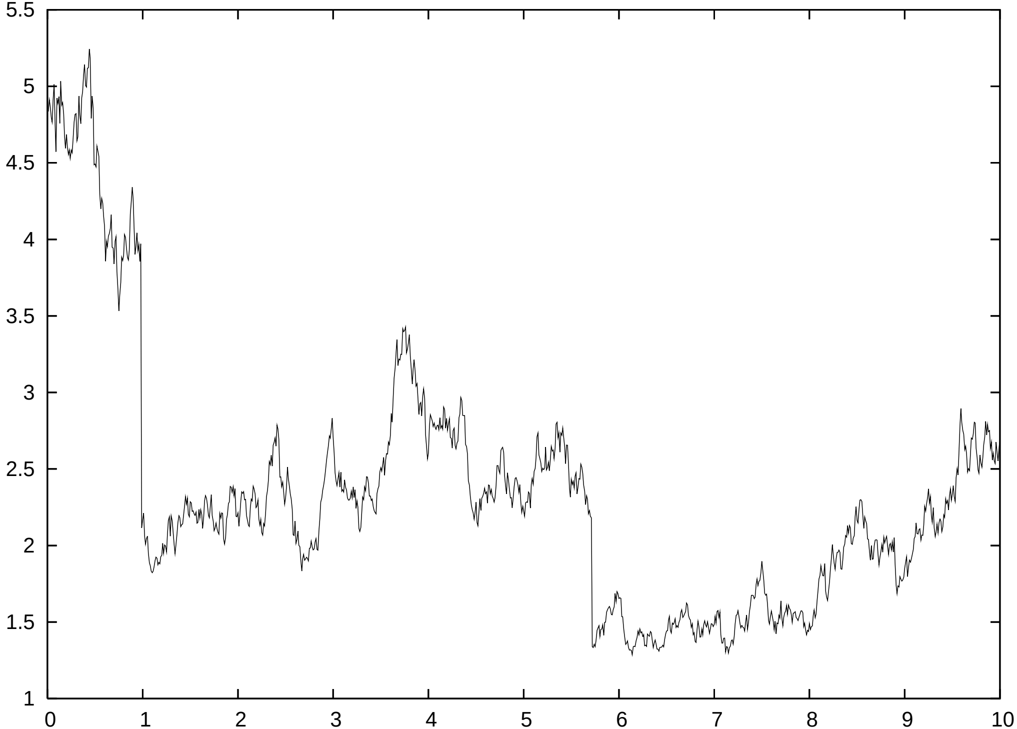
compound Poisson process

Jump Diffusion

Next superimpose the jump process to the continuous Wiener process. The combined geometric Brownian und jump process is given by

$$dS_t = \mu S_t dt + \sigma S_t dW_t + (q_t - 1) S_t dJ_t \, . \tag{1.57}$$

Example 1.21



$r = 0.06$, $\mu = 0.0995$
$\sigma = 0.3$
$\lambda = 0.2$
$\mu_J = -0.3$
$\sigma_J = 0.4$

Assume $\log(q) \sim \mathcal{N}(\mu_J, \sigma_J^2)$. $\exp(\mu_J) = 0.7408$, that is an average drop of 26 % .
A heavy 47% crash occurs for $\tau = 0.99$, with $q = 0.526$.

An analytical solution of (1.57) can be calculated on each of the jump-free subintervals $\tau_j < t < \tau_{j+1}$ where the SDE is just GBM $dS = S(\mu dt + \sigma dW)$.

For example, in the first subinterval until τ_1 the solution is given by (1.54). At τ_1 a jump of size

$$(\Delta S)_1 := (q_{\tau_1} - 1)S_{\tau_1^-}$$

occurs, and thereafter the solution continues with

$$S_t = S_0 \cdot \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t\right) + (q_{\tau_1} - 1)S_{\tau_1^-} \ ,$$

until τ_2. The interchange of continuous parts and jumps proceeds in this way, all jumps are added. So the SDE can be written as

$$S_t = S_0 + \int_0^t S_s(\mu ds + \sigma dW_s) + \sum_{j=1}^{J_t} S_{\tau_j^-}(q_{\tau_j} - 1)\,.$$

The task of minimimizing risks leads to a partial integro-differential equation. This equation reduces to the Black-Scholes equation in the no-jump special case for $\lambda = 0$.

Merton 1976

Notes and Extensions

Black-Scholes Formula as Limiting Case of the Binomial Model

Consider a European Call in the binomial model of Section 1.4.

Suppose the calculated value is $V_0^{(M)}$. In the limit $M \to \infty$ the sequence $V_0^{(M)}$ converges to the value $V_{\mathrm{C}}(S_0, 0)$ of the continuous Black-Scholes model. To prove this, proceed as follows:

a) Let j_K be the smallest index j with $S_{jM} \geq K$. Find an argument why

$$\sum_{j=j_K}^{M} \binom{M}{j} p^j (1-p)^{M-j} (S_0 u^j d^{M-j} - K)$$

is the expectation $\mathsf{E}(V_T)$ of the payoff.

b) The value of the option is obtained by discounting, $V_0^{(M)} = e^{-rT}\mathsf{E}(V_T)$. Show

$$V_0^{(M)} = S_0 B_{M,\tilde{p}}(j_K) - e^{-rT} K B_{M,p}(j_K) \ .$$

Here $B_{M,p}(j)$ is defined by the binomial distribution, and $\tilde{p} := pue^{-r\Delta t}$.

c) For large M the binomial distribution is approximated by the normal distribution with distribution $F(x)$. Show that $V_0^{(M)}$ is approximated by

$$S_0 F\left(\frac{M\tilde{p}-\alpha}{\sqrt{M\tilde{p}(1-\tilde{p})}}\right) - e^{-rT} K F\left(\frac{Mp-\alpha}{\sqrt{Mp(1-p)}}\right) \quad ,$$

where

$$\alpha := -\frac{\log\frac{S_0}{K} + M\log d}{\log u - \log d} \quad .$$

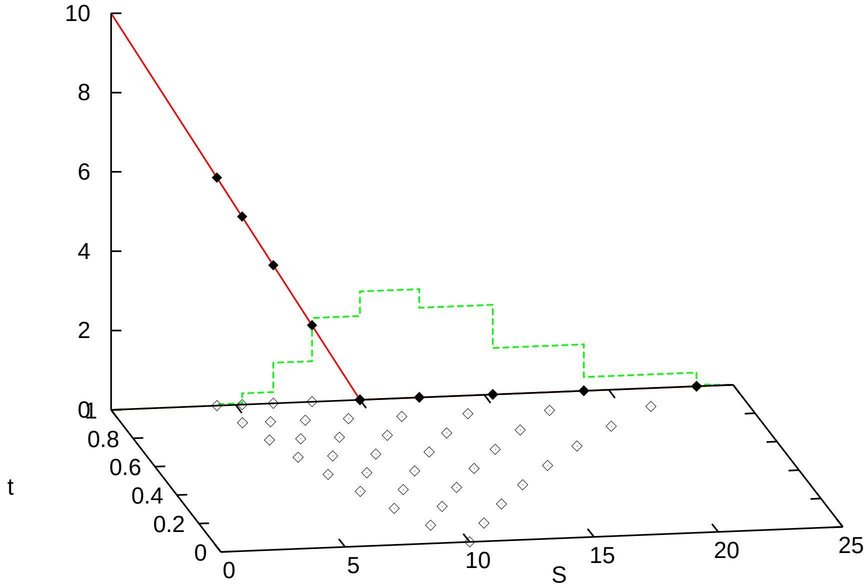
d) Substitute the p, u, d by their expressions from (1.11) to show

$$\frac{Mp-\alpha}{\sqrt{Mp(1-p)}} \longrightarrow \frac{\log\frac{S_0}{K} + (r - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}$$

for $M \to \infty$.

Hint: Up to terms of high order the approximations $u = e^{\sigma\sqrt{\Delta t}}$, $d = e^{-\sigma\sqrt{\Delta t}}$ hold. (In an analogous way the other argument of F can be analyzed.)

Illustration of a Binomial Tree and Payoff



for a put, (S,t)-points for $M = 8$, $K = S_0 = 10$.
The binomial density is shown, scaled with factor 10.

Return of the Underlying

Let a time series $S_1, ..., S_M$ of a stock price be given.

The simple return

$$\hat{R}_{i,j} := \frac{S_i - S_j}{S_j} \ ,$$

an index number of the success of the underlying, lacks the desirable property of additivity

$$R_{M,1} = \sum_{i=2}^{M} R_{i,i-1}. \qquad (*)$$

The log return

$$R_{i,j} := \log S_i - \log S_j$$

satisfies $(*)$ and $R_{i,i-1} \approx \hat{R}_{i,i-1}$