Chapter 1 Modeling Tools for Financial Options

1.1 Options

What do we mean by option? 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. An option is a financial instrument that allows —amongst other things— to make a bet on rising or falling values of an underlying asset. The **underlying** asset typically is a stock, or a parcel of shares of a company. Other examples of underlyings include stock indices (as the Dow Jones Industrial Average), currencies, or commodities. Since the value of an option depends on the value of the underlying asset, options and other related financial instruments are called *derivatives* ($\longrightarrow$ Appendix A1). An option is an agreement between two parties about trading the asset at a certain future time. One party is the *writer*, often a bank, who fixes the terms of the option contract and sells the option. The other party ist the *holder*, who purchases the option, paying the market price, which is called *premium*. How to calculate a fair value of the premium is a central theme of this book. The holder of the option must decide what to do with the rights the option contract grants. The decision will depend on the market situation, and on the type of option. There are numerous different types of options, which are not all of interest to this book. In Chapter 1 we concentrate on standard options, also known as *plain-vanilla options*. This Section 1.1 introduces important terms.

Options have a limited life time. The *maturity date* T fixes the time horizon. At this date the rights of the holder expire, and for later times ($t > T$) the option is worthless. There are two basic types of option: 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* the underlying for the price K by the date T. The previously agreed price K of the contract is called **strike** or **exercise price**[1]. It is important to note that the holder is not obligated to *exercise* —that is, to buy or sell the underlying according to the terms of the contract. The holder may wish to close his position by selling the option. In summary, at time t the holder of the option can choose to

[1] The price K as well as other prices are meant as the price of one unit of an asset, say, in \$.

- sell the option at its current market price on some options exchange (at $t < T$),
- retain the option and do nothing,
- exercise the option ($t \leq T$), or
- let the option expire worthless ($t \geq T$).

In contrast, the writer of the option has the obligation to deliver or buy the underlying for the price K, in case the holder chooses to exercise. The risk situation of the writer differs strongly from that of the holder. The writer receives the premium when he issues the option and somebody buys it. This up-front premium payment compensates for the writer's potential liabilities in the future. The asymmetry between writing and owning options is evident. This book mostly takes the standpoint of the holder.

Not every option can be exercised at any time $t \leq T$. For **European options** exercise is only permitted at expiry date T. **American options** can be exercised at any time until the expiration date. For options the labels American or European have no geographical meaning. Both types are traded in every continent. Options on stocks are mostly American style.

The value of the option will be denoted by V. The value V depends on the price per share of the underlying, which is denoted S. This letter S symbolizes stocks, which are the most prominent examples of underlying assets. The variation of the asset price S with time t is expressed by writing S_t or $S(t)$. The value of the option also depends on the remaining time to expiry $T-t$. That is, V depends on time t. The dependence of V on S and t is written $V(S,t)$. As we shall see later, it is not easy to calculate the fair value V of an option for $t < T$. But it is an easy task to determine the terminal value of V at expiration time $t = T$. In what follows, we shall discuss this topic, and start with European options as seen with the eyes of the holder.

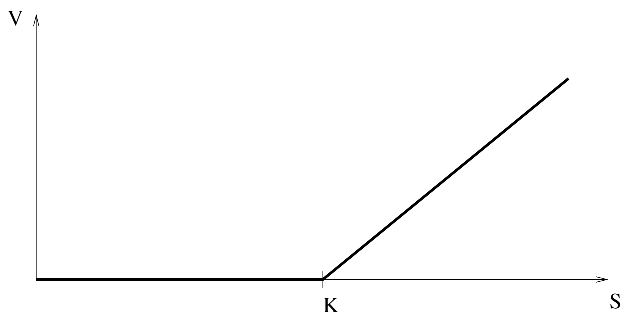

Fig. 1.1. Intrinsic value of a call with exercise price K (payoff function)

The Payoff Function

At time $t = T$, the holder of a European call option will check the current price $S = S_T$ of the underlying asset. The holder will exercise the call (buy the stock for the strike price K), when $S > K$. For then the holder can immediately sell the asset for the spot price S and makes a gain of $S - K$ per share. In this situation the value of the option is $V = S - K$. (This reasoning ignores transaction costs.) In case $S < K$ the holder will not exercise, since then the asset can be purchased on the market for the cheaper price S. In this case the option is worthless, $V = 0$. In summary, the value $V(S, T)$ of a call option at expiration date T is given by

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

Hence

$$V(S_T, T) = \max\{S_T - K, 0\}.$$

Considered for all possible prices $S_t > 0$, $\max\{S_t - K, 0\}$ is a function of S_t. This **payoff function** *(intrinsic value, cashflow)* is shown in Figure 1.1. Using the notation $f^+ := \max\{f, 0\}$, this payoff can be written in the compact form $(S_t - K)^+$. Accordingly, the value $V(S_T, T)$ of a call at maturity date T is

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

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

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

Hence

$$V(S_T, T) = \max\{K - S_T, 0\},$$

or

$$V(S_T, T) = (K - S_T)^+, \tag{1.1P}$$

compare Figure 1.2.

The curves in the payoff diagrams of Figures 1.1, 1.2 show the option values from the perspective of the holder. The profit is not shown. For an illustration of the profit, the initial costs paid when buying the option at $t = t_0$ must be subtracted. The initial costs basically 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. Substracting this amount leads to shifting the curves in Figures 1.1, 1.2 down. The resulting *profit diagram* shows a negative profit for some range of S-values, which of course means a loss.

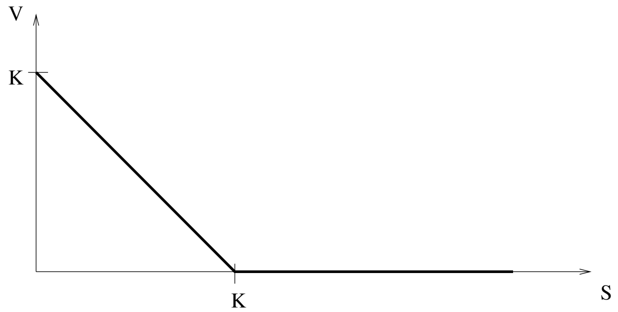

Fig. 1.2. Intrinsic value of a put with exercise price K (payoff function)

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 Figures 1.1, 1.2 as well as the equations (1.1C), (1.1P) remain valid for American type options.

The payoff diagrams of Figures 1.1, 1.2 and the corresponding profit diagrams show that a potential loss for the purchaser of an option (long position) is limited by the initial costs, no matter how bad things get. The situation for the writer (short position) is reverse. For him the payoff curves of Figures 1.1, 1.2 as well as the profit curves must be reflected on the S-axis. The writer's profit or loss is the reverse of that of the holder. Multiplying the payoff of a call in Figure 1.1 by (-1) illustrates the potentially unlimited risk of a short call. Hence the writer of a call must carefully design a strategy to compensate for his risks. We will came back to this issue in Section 1.5.

A Priori Bounds

No matter what the terms of a specific option are and no matter how the market behaves, the values V of the options satisfy certain bounds. These bounds are known a priori. For example, the value $V(S,t)$ of an American option can never fall below the payoff, for all S and all t. These bounds follow from the *no-arbitrage principle* ($\longrightarrow$ Appendix A1). To illustrate the strength of these arguments, we 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 would be in the triangle seen in Figure 1.2. That is, $S < K$ and $0 \leq V < K - S$. This scenario would allow arbitrage. The strategy would be 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. Hence the assumption that the value of an American put is below the payoff must be wrong. We conclude

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

Similarly,

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

Other bounds are listed in Appendix 7. For example, a European put on an asset that pays no dividends until T may also take values below the payoff, but is always above the lower bound $Ke^{-r(T-t)} - S$. 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 $t < T$. That is

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

as long as all other terms of the contract are identical. For European options the values of put and call are related by the *put-call parity*

$$S + V_P - V_C = Ke^{-r(T-t)} \; ,$$

which can be shown by applying arguments of arbitrage ($\longrightarrow$ Exercise 1.1).

Options in the Market

The features of the options imply that an investor purchases puts when the price of the underlying is expected to fall, and buys calls when the prices are about to rise. This mechanism inspires speculators. An important application of options is hedging ($\longrightarrow$ Appendix A1).

The value of $V(S,t)$ also depends on other factors. Dependence on the strike K and the maturity T is evident. Market parameters affecting the price are the interest rate r, the **volatility** σ of the price S_t, and dividends in case of a dividend-paying asset. The interest rate r is the risk-free rate, which applies to zero bonds or to other investments that are considered free of risks ($\longrightarrow$ Appendix A1). The dependence of V on the volatility σ is very sensitve. This critically important parameter σ can be defined as standard deviation of the fluctuations in S_t, for scaling divided by the square root of the observed time period. The volatility σ measures the uncertainty in the asset.

The units of r and σ^2 are per year. Time is measured in years. Writing $\sigma = 0.2$ means a volatility of 20%, and $r = 0.05$ represents an interest rate of 5%. The Table 1.1 summarizes the key notations of option pricing. The notation is standard except for the strike price K, which is sometimes denoted X, or E.

The time period of interest is $t_0 \leq t \leq T$. One might think of t_0 denoting the date when the option is issued and t as a symbol for "today." But this book mostly sets $t_0 = 0$ in the role of "today," without loss of generality. Then the interval $0 \leq t \leq T$ represents the remaining life time of the option. The price S_t is a stochastic process, compare Section 1.6. In real markets, the interest rate r and the volatility σ vary with time. To keep the models and the analysis simple, we assume r and σ to be constant on $0 \leq t \leq T$. Further we suppose that all variables are arbitrarily divisible and consequently can vary continuously —that is, all variables vary in the set $\mathbb{R}$ of real numbers.

Table 1.1. List of important variables

t	current time, $0 \le t \le T$
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, exercise price per share
$V(S,t)$	value of an option at time t and underlying price S

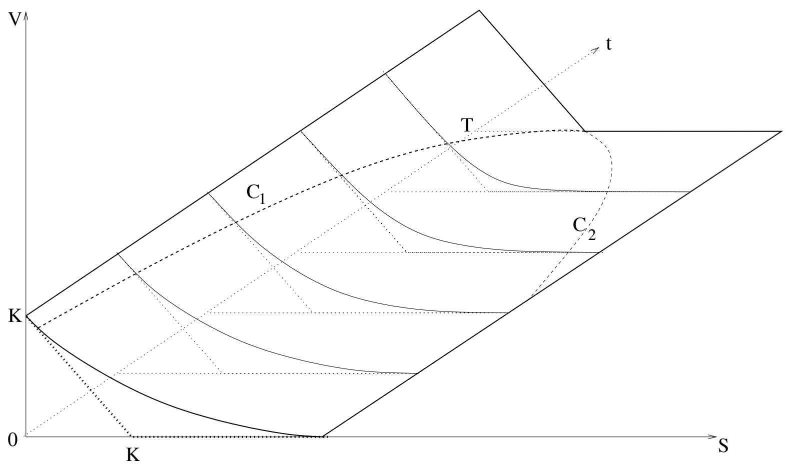

Fig. 1.3. Value $V(S,t)$ of an American put, schematically

The Geometry of Options

As mentioned, our aim is to calculate $V(S,t)$ for fixed values of K, T, r, σ. The values $V(S,t)$ can be interpreted as a piece of surface over the subset

$$S > 0\ ,\ \ 0 \le t \le T$$

of the (S,t)-plane. The Figure 1.3 illustrates the character of such a surface for the case of an American put. For the illustration assume $T = 1$. The figure depicts six curves obtained by cutting the *option surface* with the planes $t = 0, 0.2, \ldots, 1.0$. For $t = T$ the payoff function $(K - S)^+$ of Figure 1.2 is clearly visible.

Shifting this payoff parallel for all $0 \le t < T$ creates another surface, which consists of the two planar pieces $V = 0$ (for $S \ge K$) and $V = K - S$ (for $S < K$). This *payoff surface* created by $(K - S)^+$ is a lower bound to the option surface, $V(S,t) \ge (K - S)^+$. The Figure 1.3 shows two curves C_1 and C_2 on the option surface. Within the area limited by these two curves the option surface is clearly above the payoff surface, $V(S,t) > (K - S)^+$.

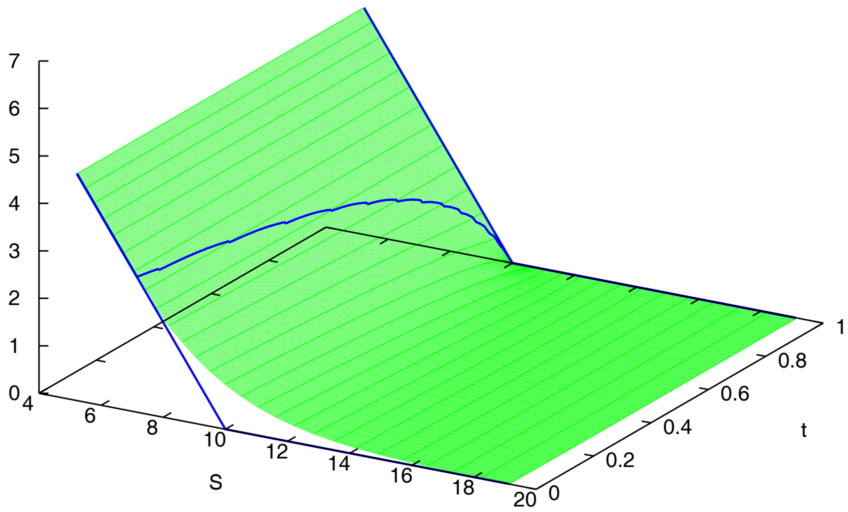

Fig. 1.4. Value $V(S,t)$ of an American put with $r = 0.06$, $\sigma = 0.30$, $K = 10$, $T = 1$

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 $\varepsilon > 0$. These topics will be analyzed in Chapter 4. The location of C_1 and C_2 is not known, these curves must be calculated along with the calculation of $V(S,t)$. Of special interest is $V(S,0)$, the value of the option "today." This curve is seen in Figure 1.3 for $t = 0$ as the front edge of the option surface. This front curve may be seen as smoothing the corner in the payoff function. The schematic illustration of Figure 1.3 is completed by a concrete example of a calculated put surface in Figure 1.4. An approximation of the curve C_1 is shown.

The above was explained for an American put. For other options the bounds are different ($\longrightarrow$ Appendix A7). As mentioned before, a European put takes values above the lower bound $Ke^{-r(T-t)} - S$, compare Figure 1.5.

1.2 Model of the Financial Market

Mathematical models can serve as approximations and idealizations of the complex reality of the financial world. For modeling financial options the models named after the pioneers Black, Merton and Scholes are both successful and widely accepted. This Section 1.2 introduces some key elements of the models.

The ultimate aim is to be able to calculate $V(S,t)$. It is attractive to define the option surfaces $V(S,t)$ on the half strip $S > 0$, $0 \leq t \leq T$ as solutions

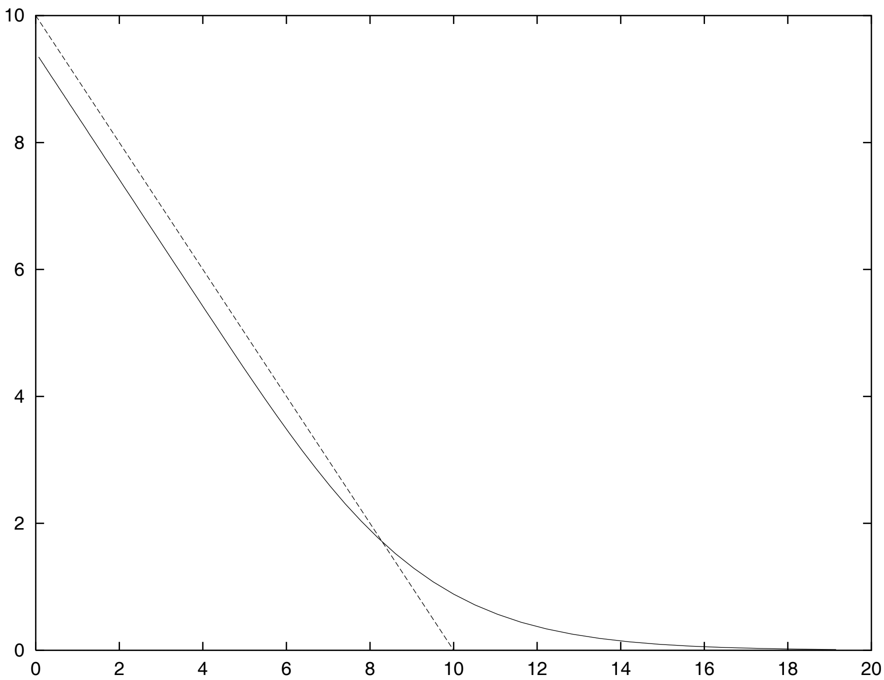

Fig. 1.5. Value of a European put $V(S, 0)$ for $T = 1$, $K = 10$, $r = 0.06$, $\sigma = 0.3$. The payoff $V(S, T)$ is drawn with a dashed line. For small values of S the value V approaches its lower bound, here $9.4 - S$.

of suitable equations. Then calculating V amounts to solving the equations. In fact, a series of assumptions allows to characterize the functions $V(S, t)$ as solutions of certain partial differential equations or partial differential inequalities. The model is represented by the famous Black-Scholes equation, which was suggested 1973.

Definition 1.1 (Black-Scholes equation)

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

The equation (1.2) is partial differential equation for the value $V(S, t)$ of options. This equation is a symbol of the market model. But what are the assumptions leading to the Black-Scholes equation?

Assumptions 1.2 (model of the market)

(a) *The market is frictionless.*

This means that there are no transaction costs (fees or taxes), the interest rates for borrowing and lending money are equal, all parties have immediate access to any information, and all securities and credits are available at any time and in any size. Consequently, all variables are perfectly divisible —that is, may take any real number. Further, individual trading will not influence the price.

(b) *There are no arbitrage opportunities.*

(c) *The asset price follows a geometric Brownian motion.*
(This stochastic motion will be discussed in Sections 1.6–1.8.)

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

These are the assumptions that lead to the Black-Scholes equation (1.2). A derivation of this partial differential equation is given in Appendix A3.

Solutions $V(S,t)$ are functions which satisfy this equation for all S and t. In addition to solving the partial differential equation, the function $V(S,t)$ must satisfiy a terminal condition and boundary conditions. The **terminal condition** for $t = T$ is

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

with payoff function (1.1C) or (1.1P), depending on the type of option. The boundaries of the half strip $0 < S$, $0 \le t \le T$ are defined by $S = 0$ and $S \to \infty$. At these boundaries the function $V(S,t)$ must satisfy **boundary conditions**. For example, a European call must obey

$$V(0,t) = 0; \quad V(S,t) \to S - Ke^{-r(T-t)} \text{ for } S \to \infty. \tag{1.3C}$$

In Chapter 4 we will come back to the Black-Scholes equation and to boundary conditions. For (1.2) an analytic solution is known (equation (A3.5) in Appendix A3). This does not hold for more general models. For example, considering transaction costs as k per unit would add the term

$$-\sqrt{\frac{2}{\pi}}\ \frac{k\sigma S^2}{\sqrt{\sigma t}}\ \left|\frac{\partial^2 V}{\partial S^2}\right|$$

to (1.2), see [WDH96], [Kw98]. In the general case, closed-form solutions do not exist, and a solution is calculated numerically, especially for American options. For numerically solving (1.2) a variant with dimensionless variables is used ($\longrightarrow$ Exercise 1.2).

At this point, a word on the notation is appropriate. The symbol S for the asset price is used in different roles: First it comes without subscript in the role of an independent real variable $S > 0$ on which $V(S,t)$ depends, say as solution of the partial differential equation (1.2). Second it is used as S_t with subscript t to emphasize its random character as stochastic process.

1.3 Numerical Methods

Applying numerical methods is inevitable in all fields of technology including financial engineering. Often the important role of numerical algorithms is not noticed. For example, an analytical formula at hand (such as the Black-Scholes formula (A3.5) in Appendix A3) might suggest that no numerical procedure is needed. But closed-form solutions may include evaluating the logarithm or the computation of the distribution function of the normal distribution. Such elementary tasks are performed using sophisticated numerical algorithms. In pocket calculators one merely presses a button without being aware of the numerics. The robustness of those elementary numerical methods is so dependable and the efficiency so large that they almost appear not to exist. Even for apparently simple tasks the methods are quite demanding ($\longrightarrow$ Exercise 1.3). The methods must be carefully designed because inadequate strategies can easily produce inaccurate results ($\longrightarrow$ Exercise 1.4).

Spoilt by generally available black-box software and graphics packages we take the support and the success of numerical workhorses for granted. We make use of the numerical tools with great respect but without further comments. We just assume an elementary education in numerical methods. An introduction into important methods and hints on the literature are given in Appendix A4.

Since financial markets undergo apparently stochastic fluctuations, stochastic approaches will be natural tools to simulate prices. These methods are based on formulating and simulating stochastic differential equations. This leads to Monte Carlo methods ($\longrightarrow$ Chapter 3). In computers, related simulations of options are performed in a deterministic manner. It will be decisive how to simulate randomness ($\longrightarrow$ Chapter 2). Chapters 2 and 3 are devoted to tools for simulation. These methods can be applied even in case the Assumptions 1.2 are not satisfied.

More efficient methods will be preferred provided their use can be justified by the validity of the underlying models. For example it may be advisable to solve the partial differential equations of the Black-Scholes type. Then one has to choose among several methods. The most elementary ones are finite-difference methods ($\longrightarrow$ Chapter 4). A somewhat higher flexibility concerning error control is possible with finite-element methods ($\longrightarrow$ Chapter 5). The numerical treatment of exotic options requires a more careful consideration of stability issues ($\longrightarrow$ Chapter 6). The methods based on differential equations will be described in the larger part of this book.

The various methods are discussed in terms of accuracy and speed. Ultimately the methods must give quick and accurate answers to real-time problems posed in financial markets. Efficiency and reliability are key demands. Internally the numerical methods must deal with diverse problems such as convergence order or stability.

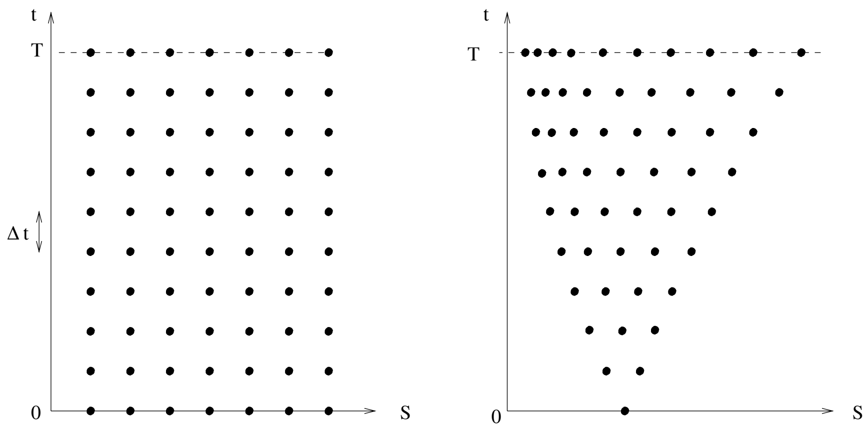

Fig. 1.6. Grid points in the (S, t)-domain

The mathematical formulation benefits from the assumption that all variables take values in the continuum $\mathbb{R}$. This idealization is practical since it avoids initial restrictions of technical nature. This gives us freedom 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. In contrast to the hypothesis of a continuum, the financial reality is rather discrete: Neither the price S nor the trading times t can take any real value. The artificial discretization introduced by numerical methods is at least twofold:

1.) The (S, t)-domain is replaced by a **grid** of a finite number of (S, t)-points, compare Figure 1.6.
2.) The differential equations are adapted to the grid and replaced by a finite number of algebraic equations.

Another kind of discretization is that computers replace the real numbers by a finite number of of rational numbers, namely the floating-point numbers. The resulting rounding error will not be relevant for much of our analysis, except for investigations of stability.

The restriction of the differential equations to the grid causes **discretization errors**. The errors depend on the coarsity of the grid. In Figure 1.6, the distance between two consecutive t-values of the grid is denoted Δt.[2] So the errors will depend on Δt and on ΔS. It is one of the aims of numerical algorithms to control the errors. The left-hand figure in Figure 1.6 shows a

[2] The symbol Δt denotes a small increment in t (analogously $\Delta S, \Delta W$). In case Δ would be a number, the product with u would be denoted $\Delta \cdot u$ or $u\Delta$.

simple rectangle grid, whereas the right-hand figure shows a tree-type grid as used in Section 1.4. The type of the grid matches the kind of underlying equations. Primarily the values of $V(S,t)$ are approximated at the grid points. Intermediate values can be obtained by interpolation.

The continuous model is an idealization of the discrete reality. But the numerical discretization does not reproduce the original discretization. For example, it would be a rare coincidence when Δt represents a day. The derivations that go along with the twofold transition

$$\text{discrete} \longrightarrow \text{continuous} \longrightarrow \text{discrete}$$

do not compensate.

1.4 The Binomial Method

The major part of the book is devoted to continuous models and their discretizations. With much less effort a discrete approach provides us with a short way to establish a first algorithm for calculating options. The resulting *binomial method* due to Cox, Ross and Rubinstein is robust and widely applicable.

In practice one is often interested in the one value $V(S_0,0)$ of an option at the current spot price S_0. Then it is unnecessarily costly to calculate the surface $V(S,t)$ for the entire domain to extract the required information $V(S_0,0)$. The relatively small task of calculating $V(S_0,0)$ can be comfortably solved using the binomial method. This 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. For illustration see the right grid in Figure 1.6, and Figure 1.9.

A Discrete Model

We begin with discretizing the continuous time t, replacing t by equidistant time instances t_i. Let us use the notations

$$\begin{aligned}
&M\text{: number of time steps}\\
&\Delta t := \tfrac{T}{M}\\
&t_i := i\cdot\Delta t,\quad i=0,...,M\\
&S_i := S(t_i)
\end{aligned}$$

So far the domain of the (S,t) half strip is replaced by parallel straight lines with distance Δt apart. In the next step we replace the continuous values S_i along the parallel $t=t_i$ by discrete values S_{ji}, for all i and appropriate j. For a better understanding of the S-discretisation compare Figure 1.7. This

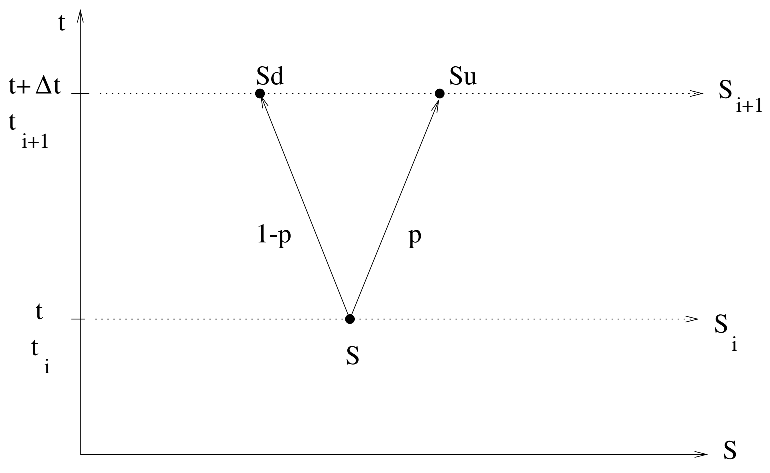

Fig. 1.7. The principle of the binomial method

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)

(A1) 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 $0 < d < u$. Here u is the factor of an upward movement and d is the factor of a downward movement.

(A2) The probability of an up movement is p, $\mathsf{P}(\text{up}) = p$.

(A3) The expected return is that of the risk-free interest rate r. For the asset price S that develops randomly from a value S_i at t_i to S_{i+1} at t_{i+1} this means

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

Let us further assume that no dividend is paid within the time period of interest. This assumption simplifies the derivation of the method and can be removed later.

An asset price following the above rules (A1), (A2) is an example of a binomial process. Such a process behaves like tossing a biased coin where the outcome "head" (up) occurs with probability p. We shall return to the assumptions (A1)-(A3) in the subsequent Section 1.5. The probability P of (A2) does not reflect the expectations of an individual in the market. Rather P is an artificial risk-neutral probability that matches (A3). The expectation in (1.4) refers to this probability; this is sometimes written E_{P}.

At this stage of the modeling the values of the parameters u, d and p are unknown. They will be fixed by suitable equations or further assumptions.

A first equation follows from Assumptions 1.3. A basic idea of the approach will be to equate the variances of the discrete and the continuous model. This will lead to a second equation. Proceeding in this manner will introduce properties of the continuous model. (The continuous model will be described in Section 1.7.) Let us start the derivation.

A consequence of (A1) and (A2) 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 the assumptions (A1), (A2). Equating with (1.4) gives

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

This is the first of three required equations to fix u, d, p. Solved for the risk-neutral probability p we obtain

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

To be a valid model of probability, $0 \le p \le 1$ must hold. This is equivalent to

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

These inequalities relate the upward and downward movement of the asset price to the riskless interest rate r. The inequalities (1.7) are no new assumption but follow from the no-arbitrage principle. The assumption $0 < d < u$ remains valid.

Next we equate variances. Via the variance the volatility σ enters the model. From the continuous model we apply the relation

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

For the relations (1.4) and (1.8) we refer to Section 1.8 ($\longrightarrow$ Exercise 1.12). Recall that the variance satisfies $\mathsf{Var}(S) = \mathsf{E}(S^2) - (\mathsf{E}(S))^2$ ($\longrightarrow$ Appendix A2). Equations (1.4) and (1.8) combine to

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

On the other hand 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 (1.5) leads to

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

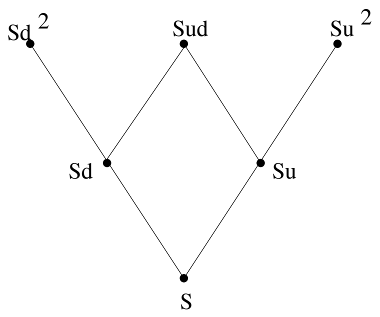

Fig. 1.8 Sequence of several meshes (schematically)

The equations (1.5), (1.9) constitute two relations for the three unknowns u, d, p. We are free to impose an arbitrary third equation. The plausible assumption

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

reflects a symmetry between upward and downward movement of the asset price. Now the parameters u, d and p are fixed. They depend on r, σ and Δt. So does the grid, which is analyzed next (Figure 1.8).

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 of the kind depicted in Figure 1.7 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, after the time period $2\Delta t$ the asset price can only take three values rather than four: The tree is recombining. It does not matter which of the two paths we take to reach Sud. This property extends to more than two time periods. 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 (1.10) these are the values $Su^j u^{j-M} = Su^{-M}u^{2j} =: S_{jM}$. The number of nodes in the tree grows quadratically in M. (Why?)

The symmetry of the choice (1.10) becomes apparent in that after two time steps the asset value S repeats. (Compare also Figure 1.9.) In the (t, S)-plane the tree can be interpreted as a grid of exponential-like curves. The binomial approach defined by (A1) with the proportionality between S_i and S_{i+1} reflects exponential growth or decay of S. So all grid points have the desirable property $S > 0$.

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

Using the abbreviation $\alpha := e^{r\Delta t}$ we obtain by elimination (which the reader may check) the quadratic equation

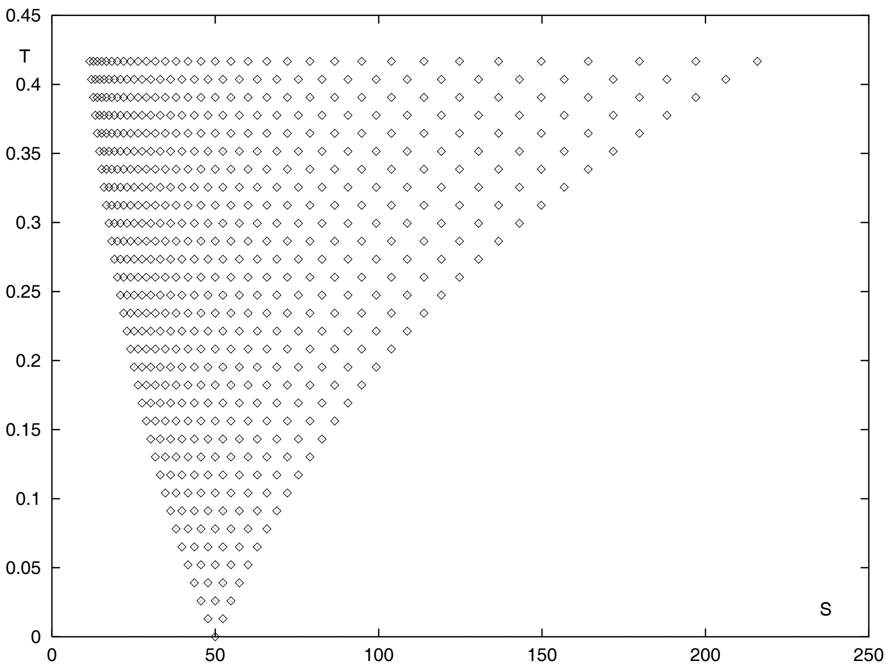

Fig. 1.9. Tree in the (S,t)-plane for $M = 32$ (data of Example 1.6)

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

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. In summary the three parameters u, d, p are given by

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

A consequence of the approach is that up to terms of higher order the relation $u = e^{\sigma\sqrt{\Delta t}}$ holds ($\longrightarrow$ Exercise 1.6). Therefore the extension of the tree in S-direction matches the volatility of the asset. So the tree will cover the relevant range of S-values.

Forward Phase: Initializing the Tree

Now the factors u and d can be considered as known and the discrete values of S for each t_i until $t_M = T$ can be calculated. The current spot price $S = S_0$ for $t_0 = 0$ is the root of the tree. To adapt the notation to the two-dimensional grid of the tree, this initial price is also denoted S_{00}. Each initial price S_0 leads to another tree of values S_{ji}.

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

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

Calculating the Option Values V, Valuation of the Tree

For t_M the payoff $V(S, t_M)$ is known from (1.1C), (1.1P). This payoff is valid for each S, including $S_{jM} = Su^j d^{M-j},\ j = 0, ..., M$. This defines the values V_{jM}:

Call: $V(S(t_M), t_M) = \max\{S(t_M) - K, 0\}$, hence:

$$V_{jM} := (S_{jM} - K)^+ \tag{1.12C}$$

Put: $V(S(t_M), t_M) = \max\{K - S(t_M), 0\}$, hence:

$$V_{jM} := (K - S_{jM})^+ \tag{1.12P}$$

The **backward phase** calculates 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, (A3). 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, (A3) of risk neutrality to V, $V_i = e^{-r\Delta t}\mathsf{E}(V_{i+1})$, we obtain using the double-index notation the recursion

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

For **European options** this is a recursion for $i = M-1, \ldots, 0$, starting from (1.12), 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 $M \to \infty$ ($\Delta t \to 0$). The accuracy of the approximation V_{00} depends on M. This is reflected by writing $V_0^{(M)}$ ($\longrightarrow$ Exercise 1.7). The basic idea of the approach implies that the limit of $V_0^{(M)}$ for $M \to \infty$ is the Black-Scholes value $V(S_0, 0)$ ($\longrightarrow$ Exercise 1.8).

For **American options** the above recursion must be modified by adding a test whether early exercise is to be preferred. To this end the value of (1.13) is compared with the value of the payoff. Then the equations (1.12) for i rather than M, combined with (1.13), read as follows:

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:

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

Let us summarize the algorithm:

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

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

The Table 1.2 lists approximations $V^{(M)}$ to $V(5, 0)$. The convergence towards the Black-Scholes value $V(S, 0)$ is visible. (In this book the number of printed decimals illustrates the attainable accuracy and does not reflect economic practice.) Applying other methods the function $V(S, 0)$ can be approximated for an interval of S-values. The Figure 1.5 shows related results obtained by using the methods of Chapter 4.

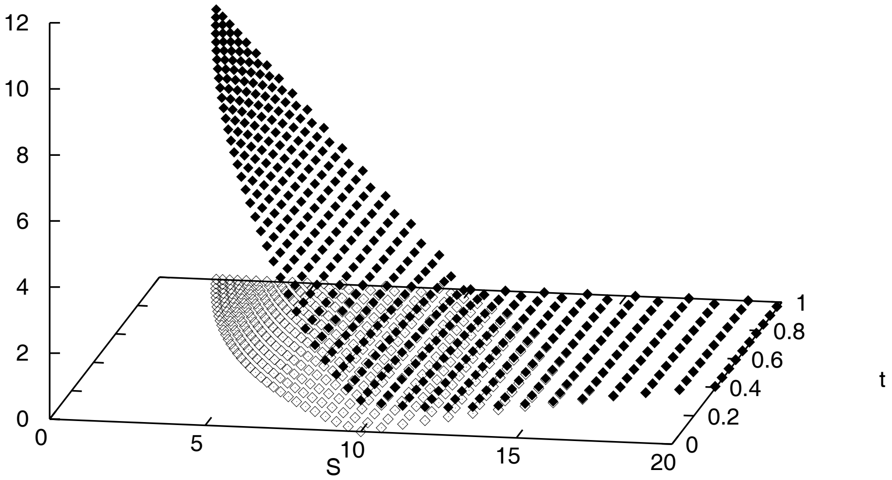

Fig. 1.10. Tree in the (S, t)-plane with (S, t, V)-points for $M = 32$ (data as in Figure 1.4)

Table 1.2. Results of Example 1.5

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

$K = 50,\ S = 50,\ r = 0.1,\ \sigma = 0.4,\ T = 0.41666...$ ($\frac{5}{12}$ for 5 months), $M = 32$.

The Figure 1.9 shows the tree for $M = 32$. The approximation to V_0 is 4.2719. Although the binomial method is not designed to accurately approximate the surface $V(S, t)$, it provides rough information also for $t > 0$. Figure 1.11 depicts for three time instances $t = 0.404$, $t = 0.3$, $t = 0.195$ the obtained approximation of $V(S, t)$; the calculated discrete values are interpolated by straight line segments. The function $V(S, 0)$ can be approximated with the methods of Chapter 4, compare Figure 4.10.

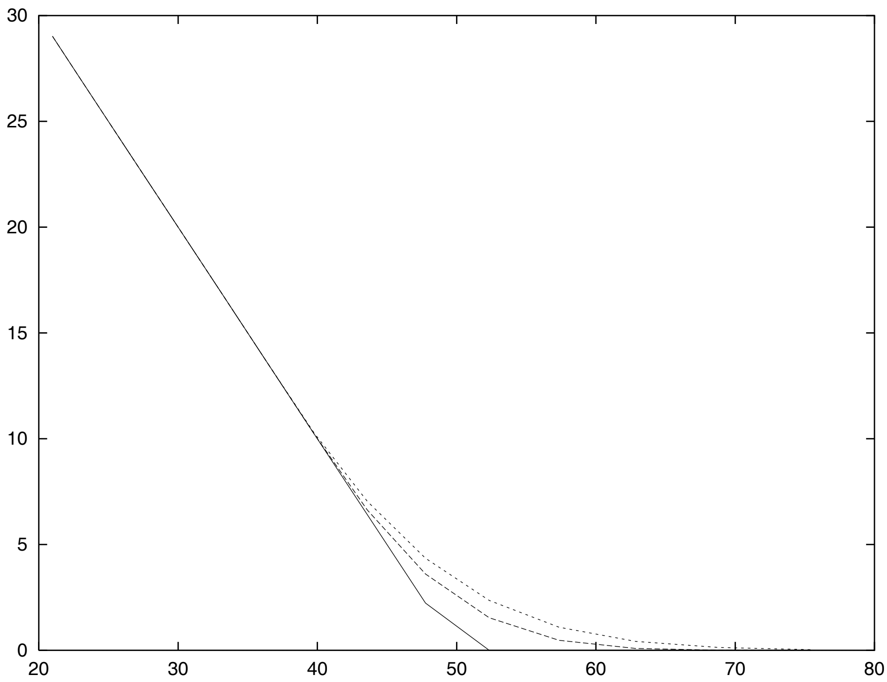

Fig. 1.11. to Example 1.6: Three cuts through the rough approximation of the surface $V(S,t)$ for $t = 0.404$ (solid curve), $t = 0.3$ (dashed), $t = 0.195$ (dotted), approximated with $M = 32$

Extensions

The paying of dividends can be incorporated into the binomial algorithm. 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, see [Hu00, § 16.3], [WDH96].

Correcting the terminal probabilities, which come out of the binomial distribution ($\longrightarrow$ Exercise 1.8), it is possible to adjust the tree to actual market data [Ru94]. Another extension of the binomial model is the *trinomial model*. Here each mesh offers three outcomes, with probabilities $p_1,\ p_2,\ p_3$ and $p_1 + p_2 + p_3 = 1$. The trinomial model allows for higher accuracy. The reader may wish to derive the trinomial method.

1.5 Risk-Neutral Valuation

In the previous section we have used the Assumptions 1.3 to derive an algorithm for valuation of options. This Section 1.5 discusses the assumptions again leading to a different interpretation.

The situation of a path-independent binomial process with the two factors u and d continues to be the basis of the argumentation. The scenario is illustrated in Figure 1.12. Here the time period is the time to expiration T, which replaces Δt in the local mesh of Figure 1.7. Accordingly, this global model is called *one-period model*. The one-period model with only two possible values of S_T 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 (A3) nor the corresponding probability $\mathsf{P}(\text{up}) = p$ from (A2). Instead we derive the probability using another argument. In this section the factors u and d are assumed to be given.

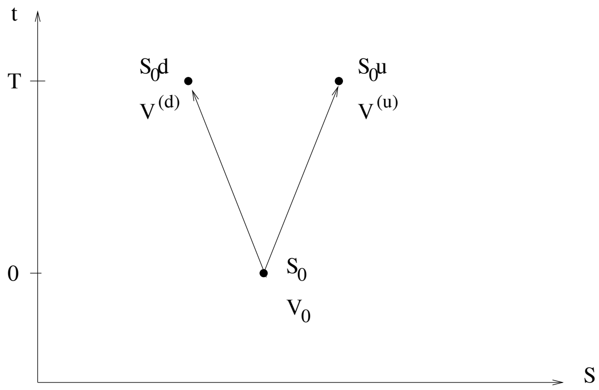

Fig. 1.12 One-period binomial model

Let us 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.** That is, a hedging strategy is needed. To discuss the hedging properly we assume that no funds are added or withdrawn.

By Π_t we denote the wealth of this portfolio at time t. Initially the value is

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

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 at the end of the life of the option is either

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

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.

If we rule 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 $e^{rT}\Pi_0$. Both the assumptions $\Pi_0 e^{rT} < \Pi_T$ and $\Pi_0 e^{rT} > \Pi_T$ would allow a strategy of earning a risk-free profit. This is in contrast to the assumed arbitrage-free world. Hence both $\Pi_0 e^{rT} \geq \Pi_T$ and $\Pi_0 e^{rT} \leq \Pi_T$ and hence equality must hold.[3] Accordingly 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:

$$\begin{aligned}
V_0 &= S_0 \cdot \Delta - e^{-rT}(S_0 u \Delta - V^{(u)}) \\
&= e^{-rT}\{\Delta \cdot [S_0 e^{rT} - S_0 u] + V^{(u)}\} \\
&= \tfrac{e^{-rT}}{u-d}\{(V^{(u)} - V^{(d)})(e^{rT} - u) + V^{(u)}(u-d)\} \\
&= \tfrac{e^{-rT}}{u-d}\{V^{(u)}(e^{rT} - d) + V^{(d)}(u - e^{rT})\} \\
&= e^{-rT}\{V^{(u)} \tfrac{e^{rT}-d}{u-d} + V^{(d)} \tfrac{u-e^{rT}}{u-d}\} \\
&= e^{-rT}\{V^{(u)} q + V^{(d)} \cdot (1-q)\}
\end{aligned}$$

with

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

[3] For an American option it is not certain that Π_T can be reached because the holder may choose early exercise. Hence we have only the inequality $\Pi_0 e^{rT} \leq \Pi_T$.

We have shown that with q from (1.17) the value of the option is given by

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

The expression for q in (1.17) is identical to the formula for p in (1.6), which was derived in the previous section. Again we have

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

Presuming these bounds for u and d, 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 (1.17),

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

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 from (1.17)) 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

$$\mathsf{E}(S_T) = S_0e^{rT} \quad \Longrightarrow \quad p = \mathsf{P}(\mathrm{up}) = \frac{e^{rT}-d}{u-d} \ ,$$

whereas in this section we arrive at the implication

$$p = \mathsf{P}(\mathrm{up}) = \frac{e^{rT}-d}{u-d} \quad \Longrightarrow \quad \mathsf{E}(S_T) = S_0e^{rT} \ .$$

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

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

The important property expressed by equation (1.20) is that of a *martingale*: The random variable $e^{-rT}S_T$ of the left-hand side has the tendency to remain at the same level. That is why a martingale is also called "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 Δ of (1.16) 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 from (1.17), such that the value V_0 satisfies (1.19) and S_0 the analogous property (1.20). 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 (A3) is meaningful.

In the real-world economy, growth rates in general are different from r, and individual subjective probabilities differ from our Q. But the assumption of a risk-neutral world leads to a fair valuation of options. The obtained value V_0 can be seen as a *rational* price. In this sense the resulting value V_0 applies to the real world. The risk-neutral valuation can be seen as a technical tool. The assumption of risk neutrality is just required to define and calculate a rational price or fair value of V_0. For this specific purpose we do not need actual growth rates of prices, and individual probabilities are not relevant. But note that we do not really assume that financial markets are actually free of risk.

The general principle outlined for the one-period model is also valid for the multi-period binomial model and for the continuous model of Black and Scholes ($\longrightarrow$ Exercise 1.8).

The Δ of (1.16) 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. The general definition is

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

the expression (1.16) is a discretized version.

1.6 Stochastic Processes

Brownian motion originally meant the erratic motion of a particle (pollen) on the surface of a fluid, caused by tiny impulses of molecules. Wiener suggested a mathematical model for this motion, the *Wiener process*. But earlier Bachelier had applied Brownian motion to model the motion of stock prices, which instantly respond to the numerous upcoming informations similar as pollen react to the impacts of molecules. The illustration of the *Dow* in Figure 1.13 may serve as motivation.

A *stochastic process* is a family of random variables X_t, which are defined for a set of parameters t ($\longrightarrow$ Appendix 1.2). Here we consider the *time-continuous* situation. That is, $t \in \mathbb{R}$ varies continuously in a time interval I,

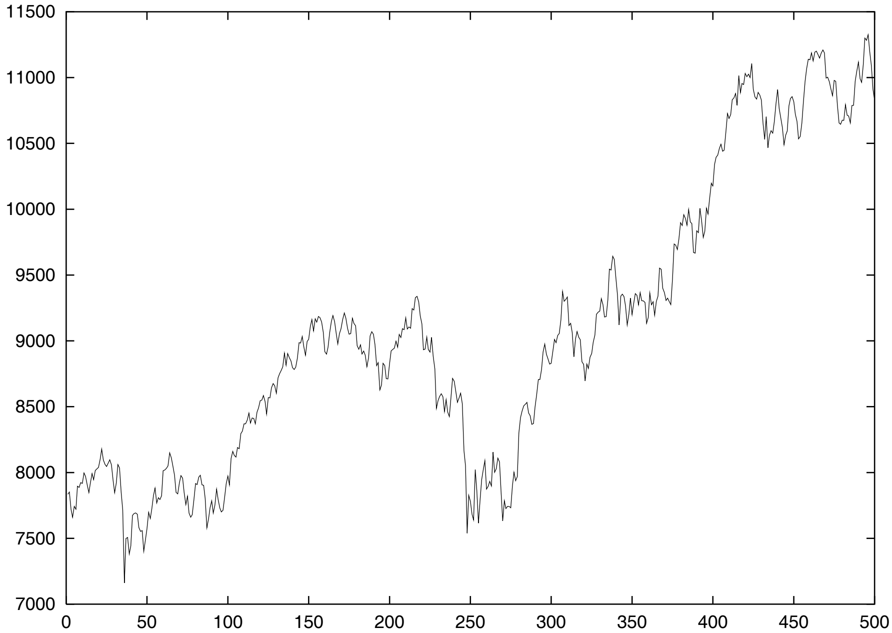

Fig. 1.13 The Dow at 500 trading days from September 8, 1997 through August 31, 1999

which typically represents $0 \leq t \leq T$. A frequent and more complete notation for a stochastic process is $\{X_t, t \in I\}$, or $(X_t)_{0 \leq t \leq T}$. Let the chance play for all t in the interval $0 \leq t \leq T$, then the resulting function X_t is called *realization* or *path* of the stochastic process.

Special properties of stochastic processes have lead to the following names:

Gaussian process: All joint distributions 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.[4]

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

[4] This assumption together with the assumption of an immediate reaction of the market to arriving informations are called *hypothesis of the efficient market* [Bo98].

1.6.1 Wiener Process

Definition 1.7 (Wiener process, Brownian motion)

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

(a) $W_0 = 0$ (with probability one)
(b) $W_t \sim \mathcal{N}(0, t)$ for all $t \geq 0$. That is, for each t the random variable W_t is normally distributed with mean $\mathsf{E}(W_t) = 0$ and variance $\mathsf{Var}(W_t) = \mathsf{E}(W_t^2) = t$.
(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 for $0 \leq s < t$ the property $W_t - W_s \sim \mathcal{N}(0, t - s)$ holds,

$$\mathsf{E}(W_t - W_s) = 0 \, , \tag{1.21a}$$

$$\mathsf{Var}(W_t - W_s) = \mathsf{E}((W_t - W_s)^2) = t - s. \tag{1.21b}$$

The relations (1.21a,b) can be derived from Definition 1.7 ($\longrightarrow$ Exercise 1.9). The relation (1.21b) is also known as

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

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 $(W_{t_{j+1}} - W_{t_j})$. Wiener processes are examples of martingales — there is no drift.

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 a sum of increments ΔW_k,

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

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

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

We summarize the numerical simulation of a Wiener process as follows:

Algorithm 1.8 (simulation of a Wiener process)

Start: $t_0 = 0,\ W_0 = 0;\ \Delta t$

loop $j = 1, 2, \ldots :$

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

draw $Z \sim \mathcal{N}(0, 1)$

$$W_j = W_{j-1} + Z\sqrt{\Delta t}$$

The drawing of Z —that is, the calculation of $Z \sim \mathcal{N}(0,1)$— will be explained in Chapter 2. The values W_j are a realization of W_t at the discrete points t_j. The Figure 1.14 shows a realization of a Wiener process; 5000 calculated points (t_j, W_j) are joined by linear interpolation.

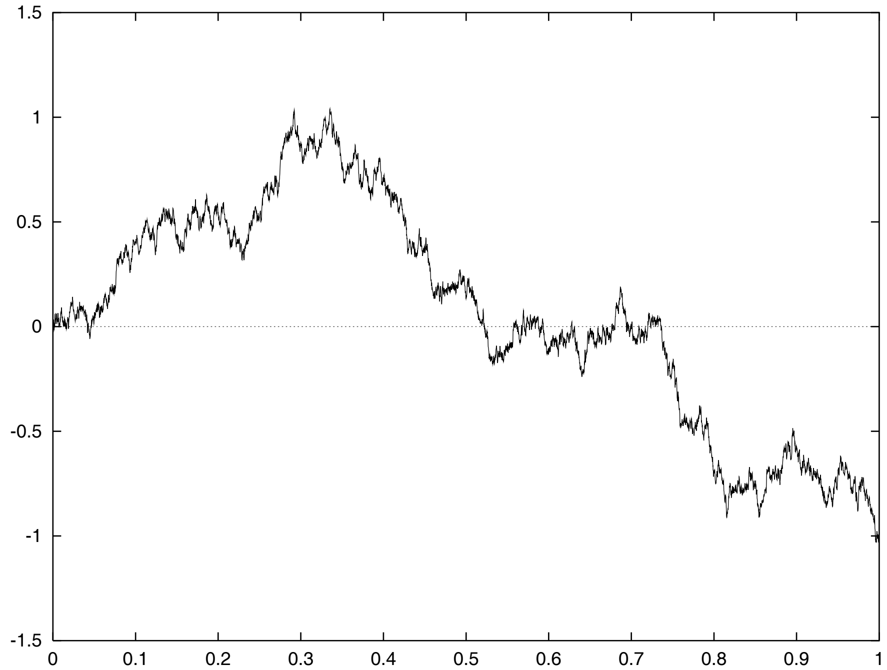

Fig. 1.14. Realization of a Wiener process, with $\Delta t = 0.0002$

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.

1.6.2 Stochastic Integral

Let us 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. We 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,

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

Such a function $b(t)$ is called *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.

We 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*

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

In our situation 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

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

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

Although this statement is not of primary concern for the theme of this book[5], we digress for a discussion because it introduces the important assertion $(dW_t)^2 = dt$. For an arbitrary partition of the interval $[0, T]$ into N subintervals the inequality

[5] The less mathematically oriented reader may like to skip the rest of this subsection.

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

holds. The left-hand sum in (1.25) 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 can be calculated using (1.21),

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

But 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 $t_0 \leq t \leq T$ with

$$\delta_N := \max_j (t_j^{(N)} - t_{j-1}^{(N)}) \ . \tag{1.26}$$

Then (dropping the $^{(N)}$)

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

Proof: The statement (1.27) means convergence in the mean ($\longrightarrow$ Appendix A2). Because of $\sum \Delta t_j = T - t_0$ we must 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

$$2\sum_j (\Delta t_j)^2$$

($\longrightarrow$ Exercise 1.10). 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 we know enough about the convergence of the left-hand sum of (1.25) and turn to the right-hand side of this inequality. 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 applied to (1.25) 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 (1.24).

The aim is to construct a stochastic integral

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

for general stochastic integrands $f(t)$. For our purposes it suffices to briefly sketch the Itô integral, which is the prototype of a stochastic integral.

For a step function b from (1.23) an integral can be defined via the sum (1.24),

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

This is the 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 $a(W_t, t) = W_t$. Its solution will be presented in Section 3.2.

1.7 Stochastic Differential Equations

1.7.1 Itô Process

Many phenomena in nature, technology and economy are modeled by means of deterministic differential equations $\dot{x} = \frac{d}{dt}x = a(x,t)$. This kind of modeling neglects stochastic fluctuations and is not appropriate for stock prices. The easiest way to consider stochastic movements is via an additive term,

$$\frac{dx}{dt} = a(x,t) + b(x,t)\xi_t.$$

Here we use the notations

a: deterministic part,

$b\xi_t$: stochastic part, ξ_t denotes a generalized stochastic process.

An example of a generalized stochastic process is *white noise*. For a brief definition of white noise we note that to each stochastic process a generalized version can be assigned [Ar74]. For generalized stochastic processes derivatives of any order can be defined. Suppose that W_t is the generalized version of a Wiener process, then W_t can be differentiated. White noise ξ_t is then defined as $\xi_t = \dot{W}_t = \frac{d}{dt}W_t$, or vice versa,

$$W_t = \int_0^t \xi_s ds.$$

That is, a Wiener process is obtained by smoothing the white noise. The smoother integral version dispenses with using generalized stochastic processes. Hence the integrated form of $\dot{x} = a(x,t) + b(x,t)\xi_t$ is studied,

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

and we replace $\xi_s ds = dW_s$. The first integral in this integral equation is an ordinary (Lebesgue- or Riemann-) integral. The second integral is an Itô integral to be taken with respect to the Wiener process W_t. The resulting stochastic differential equation (SDE) 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

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 $b = 1$ in (1.31). If $b \equiv 0$ and X_0 is constant, then the SDE becomes deterministic.

An experimental approach may help developing an intuitive understanding of Itô processes. 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 $j = 0, 1, 2, ...$
$t_{j+1} = t_j + \Delta t$
$\Delta W = Z\sqrt{\Delta t}$ with $Z \sim \mathcal{N}(0, 1)$
$y_{j+1} = y_j + a(y_j, t_j)\Delta t + b(y_j, t_j)\Delta W$

In the easiest case the *step length* Δt is chosen equidistant, $\Delta t = T/m$ for a suitable integer m. Of course the accuracy of the approximation depends on the choice of Δt ($\longrightarrow$ Chapter 3). The Algorithm 1.11 is sometimes called after Euler and Maruyama. The evaluation is straightforward. When for some example the functions a and b are easily calculated, the greatest effort may be to calculate random numbers $Z \sim \mathcal{N}(0,1)$ ($\longrightarrow$ Section 2.3). Solutions to the SDE or to its discretized version for a given realization of the Wiener process are called *trajectories* or *paths*. By *simulation* of the SDE we understand the calculation of one or more trajectories. For the purpose of visualization, the discrete data are mostly joined by straight lines.

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

Without the diffusion term the exact solution would be $X_t = X_0e^{0.05t}$. For $X_0 = 50$, $t_0 = 0$ and a time increment $\Delta t = 1/300$ the Figure 1.15 depicts a trajectory X_t of the SDE for $0 \le t \le 1$. For another realization of a Wiener process W_t the solution looks different. This is demonstrated for a similar SDE in Figure 1.16.

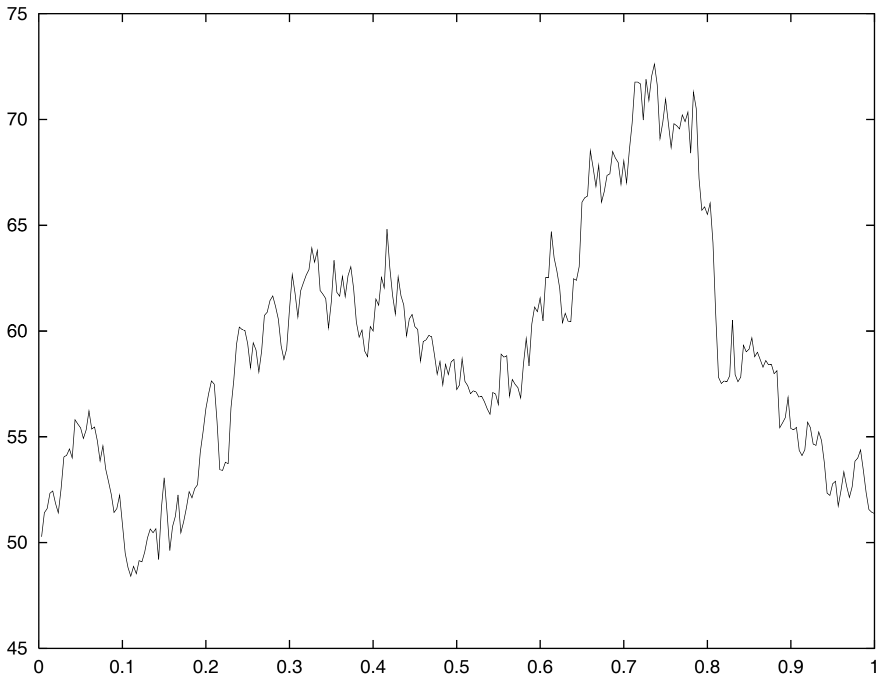

Fig. 1.15. Numerically approximated trajectory of Example 1.12 with $a = 0.05X_t$, $b = 0.3X_t$, $\Delta t = 1/300$, $X_0 = 50$

1.7.2 Application to the Stock Market

Now we discuss one of the most important continuous models 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)

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

This SDE is linear in $X_t = S_t$, $a(S_t, t) = \mu S_t$ is the drift rate with the expected *rate of return* μ, $b(S_t, t) = \sigma S_t$, σ is the volatility. (Compare Example 1.12 and Figure 1.15.) The geometric Brownian motion of (1.33) is the reference model on which the Black-Scholes-Merton approach is based. According to Assumption 1.2 we assume that μ and σ are constant.

A theoretical solution of (1.33) will be given in (1.39). The deterministic part of (1.33) 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 (1.33) 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 continuously compounded return earned by an investor per year. The rate of return μ is also called *growth rate*. The function $S_0 e^{\mu(t-t_0)}$ may be seen as a core about which the process fluctuates. Accordingly the simulated values S_1 of the ten trajectories in Figure 1.16 group around the value $50 \cdot e^{0.1} \approx 55.26$.

Let us test empirically how the values S_1 distribute about their expected value. To this end we calculate, for example, 10000 trajectories and count how many of the terminal values S_1 fall into the subintervals $k5 \le t < (k+1)5$, for $k = 0, 1, 2 \ldots$. The Figure 1.17 shows the resulting histogram. Apparently the distribution is skewed. We revisit this distribution in the next section.

The discrete version of (1.33) is

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

which we know from Algorithm 1.11. Consequently the return satisfies

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

This distribution matches actual market data in a rough approximation, see for instance Figure 1.19. This allows to calculate estimates of historical values

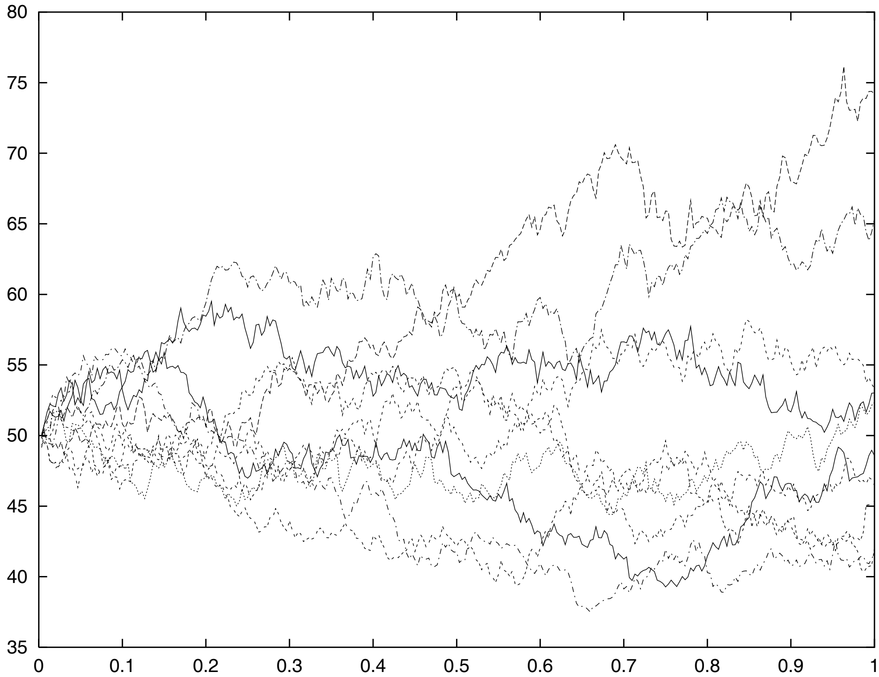

Fig. 1.16. 10 paths of SDE (1.33) with $S_0 = 50$, $\mu = 0.1$ and $\sigma = 0.2$

of the volatility σ.[6] The approximation is valid as long as Δt is small. We will return to this at the end of this chapter.

As Appendix A3 shows for the continuous case, an option can be modeled independent of individual subjective expectations on the growth rate μ. For modeling of $V(S_t, t)$, a risk-neutral world is assumed which allows to replace μ by the risk-free rate r. This was discussed for the one-period model in Section 1.5. For a thorough discussion of the continuous model, martingale theory is used. For this discussion see, for example, [Do53], [HP81], [RY91], [Du96], [Hu00], [Ne96], [MR97]. Let us summarize the situation in a remark:

Remark 1.14 (risk-neutral valuation principle)

For modeling options the return rate μ is replaced by the risk-free interest rate r, $\mu = r$.

In the reality of the market $\mu \neq r$; otherwise nobody would invest in the stock market. The investor expects $\mu > r$ as compensation for the risk that is higher for stocks than for bonds.

---

[6] For the *implied volatility* see Exercise 1.5.

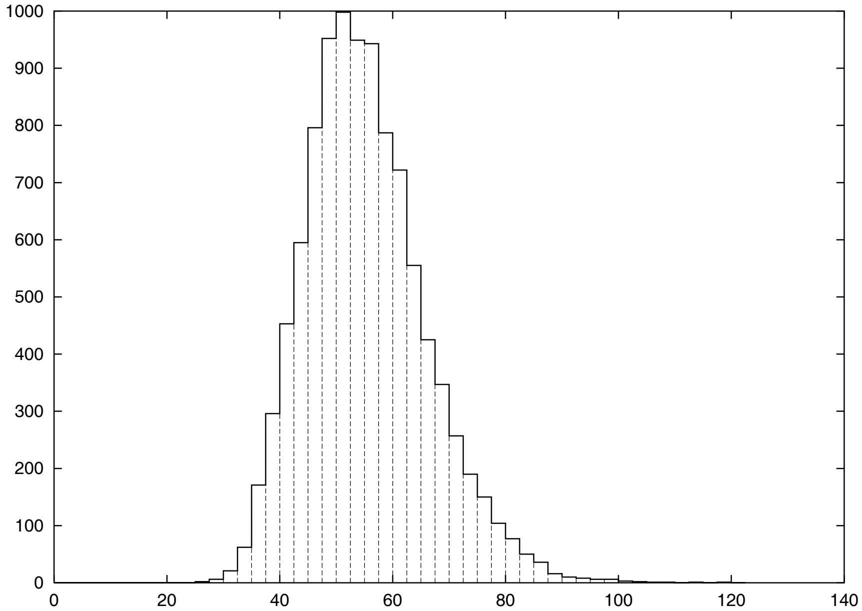

Fig. 1.17. Histogram of 10000 calculated values S_1 corresponding to (1.33), with $S_0 = 50$, $\mu = 0.1$, $\sigma = 0.2$

Mean Reversion

The assumptions of a constant interest rate r and a constant volatility σ are quite restrictive. To overcome this simplification, 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

$$dr_t = \alpha(R - r_t)dt + \sigma_{\mathrm{r}} r_t^\beta dW_t \, , \;\; \alpha > 0. \tag{1.35}$$

W_t is again a Brownian motion. The drift term in (1.35) is positive for $r_t < R$ and negative for $r_t > R$. This 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) equation (1.35) 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 $r_t \geq 0$ for all t. For a discussion of related models we refer to [LL96], [Hu00], [Kw98].

The SDE (1.35) is of a different kind as (1.33). 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. A related example is given by Example 1.15 below.

Vector-Valued SDEs

The Itô equation (1.31) is formulated as scalar equation; our SDE (1.33) 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)

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

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

Here and sometimes later on, we suppress the subscript t, which may be done when the role of the variables as stochastic processes is clear from the context. The rate of return μ is supposed to be zero; $dW^{(1)}$ and $dW^{(2)}$ may be correlated. The stochastic volatility σ follows the mean volatility ζ and is simultaneously perturbed by a Wiener process. The two SDEs for σ and ζ may be seen as a tandem controlling the dynamics of the volatility. We recommend numerical tests. As motivation see Figure 3.1.

Computational Matters

Stochastic differential equations are simulated in the context of Monte Carlo methods. Here the SDE is integrated N times, with N very large, for example, $N = 10000$. Then the weight of any single trajectory is almost neglectable. Expectation and variance are calculated over the N trajectories. Generally this costs an enormous amount of computing time. The required instruments are:

1.) Generating $\mathcal{N}(0, 1)$-distributed random numbers (Chapter 2)
2.) Integration methods for SDEs (Chapter 3)

1.8 Itô Lemma and Implications

Itô's lemma is most fundamental for stochastic processes. It may help, for example, to derive solutions of SDEs ($\longrightarrow$ Exercise 1.11).

Lemma 1.16 (Itô)

Suppose X_t follows an Itô process (1.31), $dX_t = a(X_t,t)dt + b(X_t,t)dW_t$, 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:

$$dY_t = \left(\frac{\partial g}{\partial x} a + \frac{\partial g}{\partial t} + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} b^2 \right) dt + \frac{\partial g}{\partial x} b \, dW_t \tag{1.36}$$

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

For a proof we refer to [Ar74], [Øk98], [Ste01]. Here we confine ourselves to the basic idea. 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 (1.28), $(\Delta W)^2 \approx \Delta t$, this term is also of the order $O(\Delta t)$ and belongs to the linear terms. Taking correct limits (similar as in Lemma 1.9) one obtains (1.36).

Consequences for Stocks and Options

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

$$dV_t = \left(\frac{\partial V}{\partial S} \mu S_t + \frac{\partial V}{\partial t} + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} \sigma^2 S_t^2 \right) dt + \frac{\partial V}{\partial S} \sigma S_t dW_t. \tag{1.37}$$

This SDE is used to derive the Black-Scholes equation, see Appendix A3.

As second application of Itô's lemma we 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 $\sigma^2(t - t_0)$. In view of (1.31b) 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:

$$f(S; t-t_0, S_0) := \frac{1}{S\sigma\sqrt{2\pi(t-t_0)}} \exp\left\{ -\frac{\left(\log(S/S_0) - \left(\mu - \frac{\sigma^2}{2}\right)(t-t_0)\right)^2}{2\sigma^2(t-t_0)} \right\} \tag{1.38}$$

This is the density of the *lognormal* distribution. The stock price S_t is lognormally distributed under the basic assumption of a geometric Brownian motion (1.33). The distribution is skewed, see Figure 1.18. Now the skewed behavior coming out of the experiment reported in Figure 1.17 is clear. Note that the parameters of Figures 1.17 and 1.18 match. Figure 1.17 is an approximation of the solid curve in Figure 1.18. — Having derived the density (1.38), we now can prove equation (1.8), with $\mu = r$ according to Remark 1.14 ($\longrightarrow$ Exercise 1.12).

It is inspiring to test the idealized Model 1.13 of a geometric Brownian motion 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 ($\longrightarrow$ Figure 1.19). The transformation $y = \log(S)$ is most practical ($\longrightarrow$ Exercise 1.13). In view of (1.34b), the data allow 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 (1.38) amounts to *thin tails*. This underestimates extreme events and hence does not match reality. It turns out that the geometric Brownian motion is not suitable to model risks.

We conclude this section by deriving the analytical solution of the basic linear SDE (1.33)

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

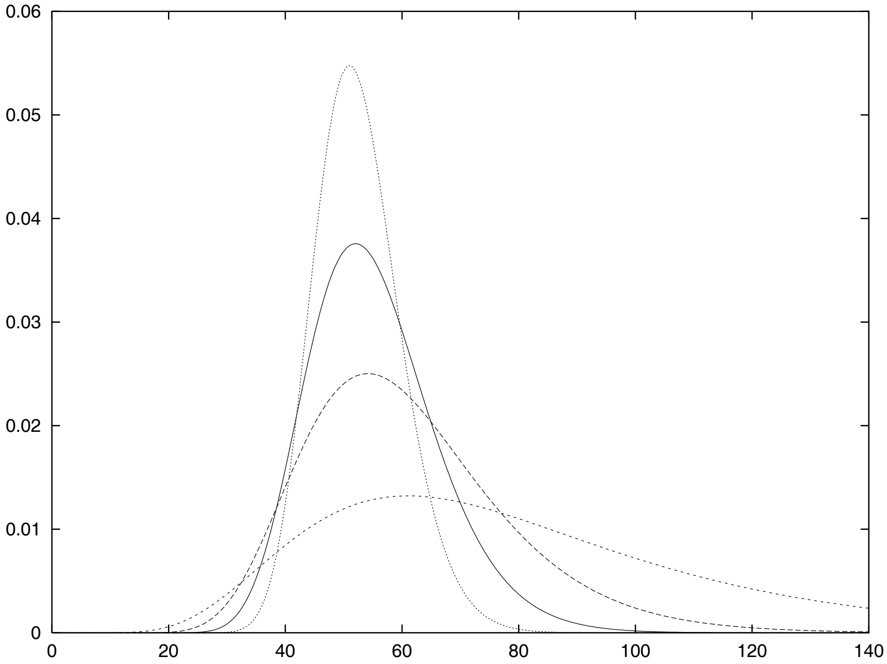

Fig. 1.18. Density (1.38) over S for $\mu = 0.1,\ \sigma = 0.2,\ S_0 = 50,\ t_0 = 0$ and $t = 0.5$ (dotted curve with steep gradient), $t = 1$ (solid curve), $t = 2$ (dashed) and $t = 5$ (dotted with flat gradient)

of Section 1.7.2. Here we again apply Itô's lemma. For an arbitrary Wiener process W_t set $X_t := W_t$ and

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

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

$$S_t := S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W_t\right) \tag{1.39}$$

solves the linear SDE (1.33). We will return to this in Chapter 3.

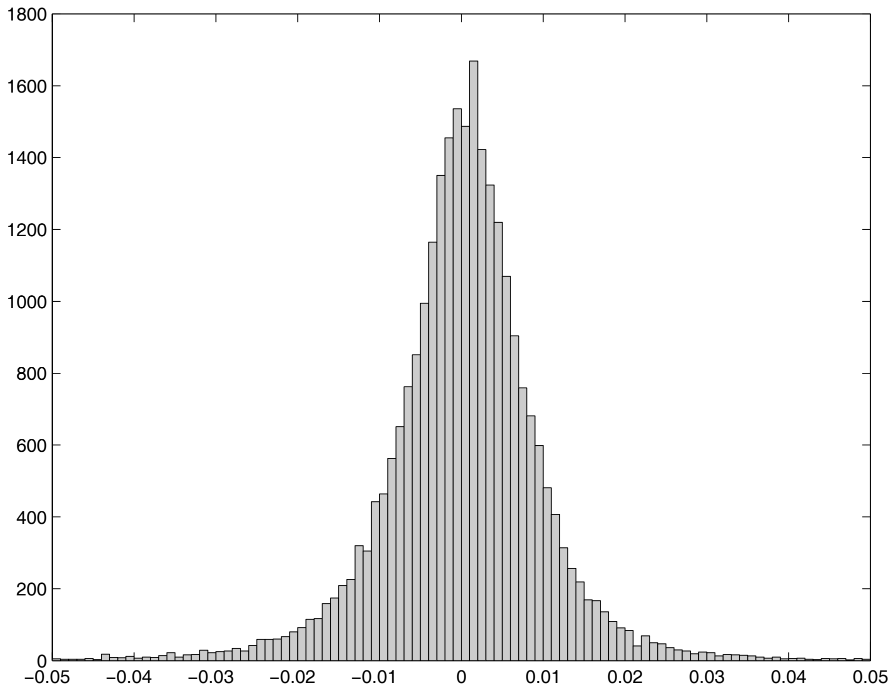

Fig. 1.19. Histogram (compare Exercise 1.13): frequency of daily log-returns $R_{i,i-1}$ of the Dow in the time period 1901-1999.

Notes and Comments

on Section 1.1:

This section presents a brief introduction into standard options. For more comprehensive studies of financial derivatives we refer, for example, to [CR85], [WDH96], [Hu00]. Mathematical detail can be found in [MR97], [KS98], [LL96], [Shi99], [Ep00], [Ste01]. (All hints on the literature are examples; an extensive overview on the many good books in this rapidly developing field is hardly possible.)

on Section 1.2:

Black, Merton and Scholes developed their approaches concurrently, with basic papers in 1973 ([BS73], and [Me90], Chapter 8). Merton and Scholes were awarded the Nobel price for economics in 1997. (Black had died in 1995.) One of the results of these authors is the so-called Black-Scholes equation (1.2) with its analytic solution formula (A3.5). For the Assumption 1.2(c) of a geometric Brownian motion see also the notes and comments on Sections 1.7/1.8.

on Section 1.3:

References on specific numerical methods are given where appropriate. As computational finance is concerned, most quotations refer to research papers. A general text book discussing computational issues is [WDH96]; further hints can be found in [RT97].

on Section 1.4:

The binomial method can sometimes be found under the heading *tree method* or *lattice method*. The binomial method was introduced by Cox, Ross and Rubinstein in 1979 [CRR79], later than the approach of Black, Merton and Scholes. The costs of the binomial method grows quadratically with the number of nodes M. The convergence rate is $O(\Delta t) = O(M^{-1})$, which is seen by plotting $V^{(M)}$ over M^{-1}. As illustrated by Figure 1.9, the described standard version wastes many nodes S_{ji} close to zero and far away from the strike region. Alternatively to the choice $ud = 1$ in equation (1.10) the choice $p = \frac{1}{2}$ is possible, see [Hu00], §16.5. For a detailed account of the binomial method see also [CR85]. When the strike K is not well grasped by the tree and its grid points, the error depending on M may oscillate. The error can be smoothed by special choices of u and d. For advanced binomial methods and speeding up convergence see [Br91], [Kl01].

We have introduced the binomial model for a one-factor model, with S representing one scalar variable. The method can be extended to multi-factor models, where S represents a vector of, say, n factors. Already for $n = 2$ there are more possibilities to have a lattice grow than the natural extensions of binomial or trinomial methods. In [MW99] hexagonal lattices are discussed for $n = 2$, and icosahedral lattices for $n = 3$. These approaches balance the number of nodes and the qualitity of the approximated distribution in an efficient way.

on Section 1.5:

As shown in Section 1.5, a valuation of options based on a hedging strategy is equivalent to the risk-neutral valuation described in Section 1.4. Another equivalent valuation is obtained by a *replication* portfolio. This basically amounts to including the risk-free investment, to which the hedged portfolio of Section 1.5 was compared, into the portfolio. To this end, the replication portfolio includes a bond with the initial value $B_0 := -(\Delta \cdot S_0 - V_0) = -\Pi_0$ and interest rate r. The portfolio consists of the bond and Δ shares of the asset. At the end of the period T the final value of the portfolio is $\Delta \cdot S_T + e^{rT}(V_0 - \Delta \cdot S_0)$. The hedge parameter Δ and V_0 are determined such that the value of the portfolio is V_T, independent of the price evolution. By adjusting B_0 and Δ in the right proportion we are able to replicate the option position. This strategy is *self-financing*: No initial net investment is required. The result of the self-financing strategy with the replicating portfolio is the same as what was derived in Section 1.5. The reader may like to check this.

Frequently discounting is done with the factor $(1 + r \cdot \Delta t)^{-1}$. Our $e^{-r\Delta t}$ or e^{-rT} is consistent with the approach of Black, Merton and Scholes. For references on risk-neutral valuation we mention [Hu00], [MR97], [Kw98] and [Shr00].

The *martingale* property is defined via conditional expected values as

$$\mathsf{E}(X_t|\mathcal{F}_s) = X_s \quad \text{for all} \quad s < t \ .$$

Here $\mathcal{F}_s$ is a filtration and the stochastic process X_t is *adapted*, which means that X_t is $\mathcal{F}_t$ measurable for all t [Do53], [Ne96], [Øk98], [Shi99], [Shr00]. The martingale property means that at time instant s with given information set $\mathcal{F}_s$, all variations of X_t for $t > s$ are unpredictable; X_s is the best forecast.

on Section 1.6:

Introductions into stochastic processes and further hints on advanced literature may be found in [Do53], [Fr71], [Ar74], [Bi79], [RY91], [KP92], [Shi99]. The requirement (a) of Definition 1.7 ($W_0 = 0$) is merely a convention of technical relevance; it serves as normalization. This Brownian motion ist called *standard* Brownian motion.

In contrast to the results for Wiener processes, differentiable functions W_t satisfy for $\delta_N \to 0$

$$\sum |W_{t_j} - W_{t_{j-1}}| \longrightarrow \int |W'_s| ds \quad , \quad \sum (W_{t_j} - W_{t_{j-1}})^2 \longrightarrow 0 \ .$$

The Itô integral and the alternative Stratonovich integral are explained in [Do53], [Ar74], [CW83], [RY91], [KS91], [KP92], [Øk98], [Sc80], [Shr00]. The difference between the two integrals is easy to illustrate for diffusion terms $b(t)dW_t$, for which $b(t)$ is a step function, see (1.23). Then a definition of the stochastic integral $\int_{t_0}^{t} b(s)dW_s$ can be constructed by means of Riemann-Stieltjes sums

$$\sum_{i=1}^{n} b(\tau_i) \left(W_{t_i} - W_{t_{i-1}} \right)$$

for intermediate values $\tau_i \in [t_{i-1}, t_i]$. Itô chooses *non-anticipating* $\tau_i = t_{i-1}$, whereas the Stratonovich integral is defined by choosing $\tau_i = \frac{1}{2}(t_{i-1} + t_i)$. Both integrals have different properties. The Stratonovich integral does not satisfy the martingale property and is not suitable for applications in finance. The class of (Itô-)stochastically integrable functions is characterized by the properties $f(t)$ is $\mathcal{F}_t$ adapted and $\mathsf{E} \int f(s)^2 ds < \infty$.

on Sections 1.7, 1.8:

The connection between white noise and Wiener processes is discussed in [Ar74]. White noise is a Gaussian process ξ_t with $\mathsf{E}(\xi_t) = 0$ and a spectral density that is constant on the entire real axis. White noise is an analogy to white light where all frequencies have the same energy.

The general linear SDE is of the form

$$dX_t = (a_1(t)X_t + a_2(t))dt + (b_1(t)X_t + b_2(t))dW_t.$$

The expectation $\mathsf{E}(X_t)$ of a solution process X_t of a linear SDE satisfies the differential equation

$$\frac{d}{dt}\mathsf{E}(X_t) = a_1\mathsf{E}(X_t) + a_2,$$

see [KP92], p. 113. A similar differential equation holds for $\mathsf{E}(X_t^2)$. This allows to calculate the variance. — The Example 1.15 with a system of three SDEs is taken from [HPS92]. [KP92] gives in Section 4.4 a list of SDEs that are analytically solvable or reducible.

The model of a geometric Brownian motion of equation (1.33) is the classical model describing the dynamics of stock prices. It goes back to Samuelson (1965; Nobel price for economics in 1970). Already in 1900 Bachelier had suggested to model stock prices with Brownian motion. Bachelier used the arithmetic version, which can be characterized by replacing the left-hand side of (1.33) by the absolute change dS. For $\mu = 0$ this amounts to the process $S_t = S_0 + \sigma W_t$. Here the stock price can become negative. Main advantages of the geometric Brownian motion are the success of the approaches of Black, Merton and Scholes, which is based on that motion, and the existence of moments (as the expectation).

It is questionable whether linear models will be lasting in the future. In view of their continuity, Wiener processes are not appropriate to model jumps, which are characteristic for the evolution of stock prices. The jumps lead to relatively *thick tails* in the distribution of empirial returns (see Figure 1.19). As already mentioned, the tails of the lognormal distribution are too thin. Other distributions match empirical data better. One example is the Pareto distribution, which has tails behaving like $x^{-\alpha}$ for large x and a constant $\alpha > 0$. A correct modeling of the tails is an integral basis for *value at risk (VaR)* calculations. For the risk aspect compare [BaN97], [Do98], [EKM97]. For distributions that match empirical data see [EK95], [Shi99], [BP00], [MRGS00], [BTT00]. Estimates of future values of the volatility are obtained by (G)ARCH methods, which work with different weights of the returns [Shi99], [Hu00]. Of great promise are models that consider the market as *dynamical system* [Lu98], [BH98], [CDG00], [BV00], MCFR00], [Sta01], [DBG01]. These systems experience the nonlinear phenomena *bifurcation* and *chaos*, which require again numerical methods. Such methods exist, but are explained elsewhere [Se94].

Exercises

Exercise 1.1 Put-Call Parity

Consider a portfolio consisting of three positions related to the same asset, namely one share (price S), one European put (value V_{P}), plus a short position of one European call (value V_{C}). Put and call have the same expiration date T, and no dividends are paid. Assume a no-arbitrage market without transaction costs. Show

$$S + V_{\mathrm{P}} - V_{\mathrm{C}} = K e^{-r(T-t)}$$

for all t, where K is the strike and r the risk-free interest rate.

Exercise 1.2 Transforming the Black-Scholes Equation

Show that the Black-Scholes equation (1.2)

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

for $V(S,t)$ is equivalent to the equation

$$\frac{\partial y}{\partial \tau} = \frac{\partial^2 y}{\partial x^2}$$

for $y(x,\tau)$. For proving this, you may proceed as follows:

a) Use the transformation $S = Ke^x$ and a suitable transformation $t \leftrightarrow \tau$ to show that (1.2) is equivalent to

$$-\dot{V} + V'' + \alpha V' + \beta V = 0$$

with $\dot{V} = \frac{\partial V}{\partial \tau}$, $V' = \frac{\partial V}{\partial x}$, $\alpha,\ \beta$ depending on r and σ.

b) The next step is to apply a transformation of the type

$$V = K \exp(\gamma x + \delta \tau) y(x,\tau)$$

for suitable $\gamma,\ \delta$.

c) Transform the boundary conditions and the terminal condition of the Black-Scholes equation accordingly.

Exercise 1.3 Standard Normal Distribution Function

Establish an algorithm to calculate

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-\frac{t^2}{2}) dt.$$

Hint: Construct an algorithm to calculate the *error function*

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2)dt$$

and use $\operatorname{erf}(x)$ to calculate $F(x)$. Use quadrature methods ($\longrightarrow$ Appendix A4).

Exercise 1.4 Calculating an Estimate of the Variance

An estimate of the variance of M numbers $x_1, ..., x_M$ is

$$s_M^2 := \frac{1}{M-1} \sum_{i=1}^{M} (x_i - \bar{x})^2, \quad \text{with} \quad \bar{x} := \frac{1}{M} \sum_{i=1}^{M} x_i$$

The alternative formula

$$s_M^2 = \frac{1}{M-1} \left(\sum_{i=1}^{M} x_i^2 - \frac{1}{M} \left(\sum_{i=1}^{M} x_i \right)^2 \right)$$

can be evaluated with only one loop $i = 1, ..., M$, but should be avoided because of the danger of cancellation. The following single-loop algorithm is recommended:

$$\begin{aligned}
&\alpha_1 := x_1, \ \beta_1 := 0 \\
&\mathit{for} \ \ i = 2, ..., M : \\
&\alpha_i := \alpha_{i-1} + \frac{x_i - \alpha_{i-1}}{i} \\
&\beta_i := \beta_{i-1} + \frac{(i-1)(x_i - \alpha_{i-1})^2}{i}
\end{aligned}$$

a) Show $\bar{x} = \alpha_M, \ s_M^2 = \frac{\beta_M}{M-1}$.

b) For the i-th *update* in the algorithm carry out a rounding error analysis. What is your judgement on the algorithm?

Exercise 1.5 Implied Volatility

For European options we take the valuation formula of Black and Scholes of the type $V = v(S, t, T, K, r, \sigma)$. For the definition of the function v see Appendix A3, equation (A3.5). If actual market data of the price V are known, then one of the parameters considered known so far can be viewed as unknown and fixed via the implicit equation

$$V - v(S, t, T, K, r, \sigma) = 0 \, .$$

The unknown parameter can be calculated iteratively as solution of this equation. Consider σ to be in the role of the unknown parameter. The volatility σ determined in this way is called *implied volatility* and is zero of $f(\sigma) := V - v(S, t, T, K, r, \sigma)$.

Assignment: Design, implement and test an algorithm to calculate the implied volatility of a call. Use Newton's method to construct a sequence $x_k \to \sigma$. The derivative $f'(x_k)$ can be approximated by the difference quotient

$$\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} .$$

For the resulting *secant iteration* invent a stopping criterion that requires smallness of both $|f(x_k)|$ and $|x_k - x_{k-1}|$.

Exercise 1.6 Price Evolution for the Binomial Method

For β from (1.11) and $u = \beta + \sqrt{\beta^2 - 1}$ show

$$u = \exp\left(\sigma\sqrt{\Delta t}\right) + O\left(\sqrt{(\Delta t)^3}\right) .$$

Exercise 1.7 Implementing the Binomial Method

Design and implement an algorithm for calculating the value $V^{(M)}$ of a European or American option. Use the binomial method of Algorithm 1.4.

INPUT: r (interest rate), σ (volatility), T (time to expiration in years),
K (strike price), S (price of asset), and the choices
put or *call*, and *European* or *American*.

Control the mesh size $\Delta t = T/M$ adaptively. For example, calculate V for $M = 8$ and $M = 16$ and in case of a significant change in V use $M = 32$ and possibly $M = 64$.

Test examples:
a) put, European, $r = 0.06$, $\sigma = 0.3$, $T = 1$, $K = 10$, $S = 5$
b) put, American, $S = 9$, otherwise as in a)
c) call, otherwise as in a)

Exercise 1.8 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 $V_C(S_0, 0)$ of the continuous Black-Scholes model given by (A3.5) ($\longrightarrow$ Appendix A3). 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. (For an illustration see Figure 1.20.)

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 ($\longrightarrow$ Appendix A2), 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: Use Exercise 1.6: 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.)

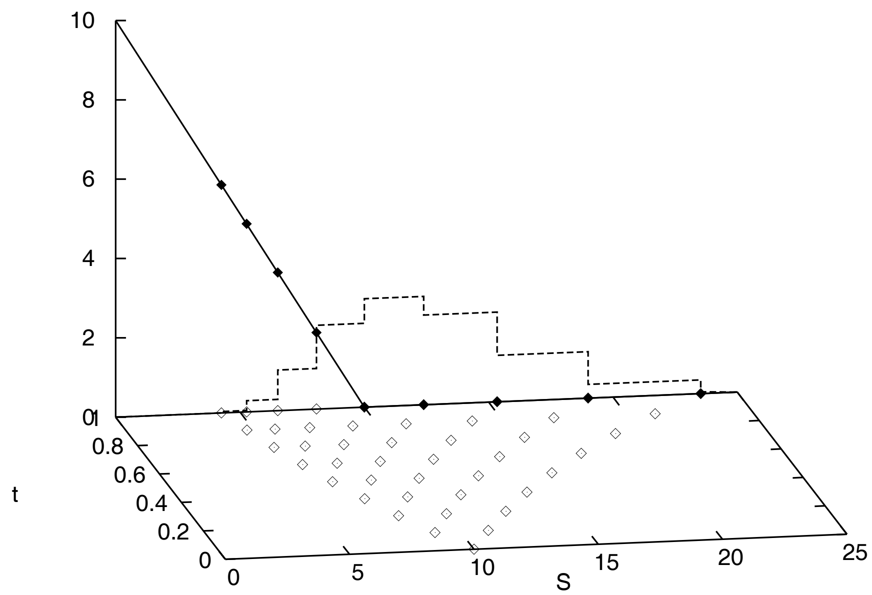

Fig. 1.20. Illustration of a binomial tree and payoff for Exercise 1.8, here for a put, (S, t) points for $M = 8$, $K = S_0 = 10$. The binomial density is shown, scaled with factor 10.

Exercise 1.9

In Definition 1.7 the requirement (a) $W_0 = 0$ is dispensable. Then the requirement (b) reads

$$\mathsf{E}(W_t - W_0) = 0 \quad , \quad \mathsf{E}((W_t - W_0)^2) = t \quad .$$

Use these relations to deduce (1.21).
Hint: $(W_t - W_s)^2 = (W_t - W_0)^2 + (W_s - W_0)^2 - 2(W_t - W_0)(W_s - W_0)$

Exercise 1.10

a) Suppose that a random variable X_t satisfies $X_t \sim \mathcal{N}(0, \sigma^2)$. Use (A2.1) to show

$$\mathsf{E}(X_t^4) = 3\sigma^4 \ .$$

b) Apply a) to show the assertion in Lemma 1.9,

$$\mathsf{E}\left(\sum_j ((\Delta W_j)^2 - \Delta t_j)\right)^2 = 2\sum_j (\Delta t_j)^2$$

Exercise 1.11 Analytical Solution of Special SDEs

Apply Itô's lemma to show

a) $X_t = \exp\left(W_t - \frac{1}{2}t\right)$ solves $dX_t = X_t dW_t$
b) $X_t = \exp\left(2W_t - t\right)$ solves $dX_t = X_t dt + 2X_t dW_t$

Hint: Use suitable functions g with $Y_t = g(X_t, t)$. In (a) start with $X_t = W_t$ and $g(x, t) = \exp(x - \frac{1}{2}t)$.

Exercise 1.12 Moments of the Lognormal Distribution

For the density function $f(S; t - t_0, S_0)$ from (1.38) show

a) $\int_0^\infty S f(S; t - t_0, S_0) dS = S_0 e^{\mu(t-t_0)}$
b) $\int_0^\infty S^2 f(S; t - t_0, S_0) dS = S_0^2 e^{(\sigma^2+2\mu)(t-t_0)}$

Hint: Set $y = \log(S/S_0)$ and transform the argument of the exponential function to a squared term.
In case you still have strength afterwards, calculate the value of S for which f is maximal.

Exercise 1.13 Return of the Underlying

Let a time series $S_1, ..., S_M$ of a stock price be given (for example the data of Figure 1.13 in the internet http://www.mi.uni-koeln.de/numerik/compfin).

The *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

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

has better properties.

a) Show $R_{i,i-1} \approx \hat{R}_{i,i-1}$, and
b) $R_{i,j}$ satisfies $(*)$.
c) For empirical data calculate the $R_{i,i-1}$ and set up histograms.
d) Suppose S is lognormally distributed. How can a value of the volatility be obtained from an estimate of the variance?
e) The mean of the 26866 returns of the time period of 98.66 years of Figure 1.19 is 0.000199 and the standard deviation is 0.01069. Calculate an estimate of the historical volatility σ.