2

Real Options in Theory and Practice

2.1 Introduction

In this chapter, the theory and practice of real options will be introduced. Although their application belongs to the field of Corporate Finance, the valuation methods origined in the option pricing theory for financial securities. In 1977, Stewart C. Myers introduced real options as a new area of financial research with the publication of his famous article[1] on this subject. He also coined the term *real option* and, as a result, is often referred to as the *father of real options theory*.

The introduction to the concept of real options is the topic of Section 2.2. Sections 2.3, 2.4, and 2.5 are devoted to Financial Engineering. The language of Financial Engineering and hence that of real options is stochastic calculus, which will be used throughout this book when quantifying real options. Section 2.4 introduces the basics of this theory in order to provide an understanding of the mathematical language used to describe real options valuation and stochastic term structure models. The term *diffusion process* will take center stage in the following chapters. However, since the meaning of diffusion processes can be best explained by applying methods of classical probability theory, a mathematical detour will be taken in Section 2.3 about classical probability theory for diffusion processes. Section 2.5 is devoted to numerical simulation methods of continuous-time stochastic processes as needed for the computer simulation program. The available literature on the quantitative description of real options models is the topic of Section 2.6. This section comprises four subsections: The first section (2.6.1) deals with decision-tree analysis. Subsequently, Section 2.6.2 introduces the idea of contingent-claims analysis and Section 2.6.3 categorizes the various real options valuation methods, distinguishing between analytical and numerical methods. The flexibility of an investment project due to interest rate uncertainty is the topic of Section 2.6.4. The final section (2.7) gives a summary of Chapter 2.

[1] See Myers [99].

M. Schulmerich, *Real Options Valuation: The Importance of Interest Rate Modelling in Theory and Practice*, DOI 10.1007/978-3-642-12662-8_2, © Springer-Verlag Berlin Heidelberg 2010

2.2 Basics of Real Options

This section provides an introduction to the world of real options. It outlines what they are and how they have developed in the past 27 years. The origin of real options theory is the theory of financial options, options that are written on an exchange-traded underlying. The valuation of financial options has been an area of research for more than three decades. The important breakthrough was achieved by Myron C. Scholes[2] and Fisher Black[3] in the early 70s at the MIT Sloan School of Management with their famous Black-Scholes formula[4] for European options. An important contribution to this formula was made by Robert C. Merton[5] with his no-arbitrage argument. A summary of the historical development of the Black-Scholes formula can be found in Black [7].

Although the Black-Scholes pricing formula was applied immediately at Wall Street, it was not obvious at first how to apply this theory to Corporate Finance. The transition from Capital Markets theory to Corporate Finance theory took place in 1977 when Stewart C. Myers published his idea of perceiving discretionary investment opportunities as *growth options*[6]. The term *real option* was coined by Stewart C. Myers as well: *Strategic planning needs finance. Present value calculations are needed as a check on strategic analysis and vice versa. However, standard discounted cash flow techniques will tend to understate the option value attached to growing profitable lines of business. Corporate Finance theory requires extension to deal with "real options".*[7]

Since 1977 academic researchers started to publish a number of articles[8] on this subject. However, it was not before 1996 that the first successful attempt to make the real options theory accessible to financial practitioners was made when Lenos Trigeorgis published his famous book *Real Options*[9]. Trigeorgis presented the theory and the practice of real options, explained various valuation methods and presented several case studies where the real options approach creates value. This book is well suited for both practitioners and Corporate Financial Engineers. From a more applied point of view, Amram

[2] Myron C. Scholes, Professor of Finance, Emeritus, at the Graduate School of Business at Stanford University. Professor at the MIT Sloan School of Management 1968-1973.

[3] Fisher Black (1938-1995), Professor at the MIT Sloan School of Management 1971 - 1984.

[4] See Black & Scholes [9].

[5] Robert C. Merton, Professor of Finance at the Harvard Business School. Professor at the MIT Sloan School of Management 1970-1988.

[6] See Myers [99] and Trigeorgis [133], page 15.

[7] See Myers [99], page 147, or Myers [100], page 137.

[8] See Section 2.6 for an overview.

[9] See Trigeorgis [133].

and Kulatilaka approached this area in their book *Real Options*[10], published in 1999. Copeland and Antikarov offered an applied, less quantitative, approach to real options as well[11].

Real options practitioners and academics have also contributed to the diffusion of the real options concept. A book published by Hommel, Scholich, and Vollrath[12] in 2001 on the real options approach in Corporate Finance practice and theory is one of the most successful books in Germany on this subject. Another book on real options was published in 2003 by Hommel, Scholich, and Baecker[13]. Both books contain articles from various authors, from the corporate sector as well as from the academic world.

Although all these books are a valid introduction to real options for a broad audience (and this list is by no means complete), the main valuation tool in capital budgeting is still the DCF method. Until recently, capital budgeting problems, i.e. the valuation of investment projects, were almost exclusively done via the DCF method. This method calculates the project's net present value NPV given a deterministic cash flow structure of the project and a known discount factor k as

$$NPV \quad = \quad \sum_{i=0}^{N} \frac{C_{t_i}}{(1+k)^{t_i}} \, .$$

A positive cash flow $C_{t_i} > 0$ indicates a cash inflow, a negative cash flow $C_{t_i} < 0$ indicates a cash outflow. In total the project is assumed to have $N+1$ discrete cash in- and outflows. In order to apply the formula above very strong restrictions have to be fulfilled. Especially, the cash flow structure of the whole project needs to be known at the beginning of the project and the discount factor k needs to be constant over the lifetime of the project. However, many projects do not fulfill these criteria.

The theory of the DCF method excludes management from making decisions and capitalizing on emerging opportunities during the lifetime of the project. However, in practice those decisions are made and change the project's cash flow structure and the discount factor that should be applied for the project valuation. Unfortunately, neither the DCF approach nor any other traditional approach of capital budgeting is apt to integrate these changes and capture the asymmetric information embedded in these investment opportunities. To quote Dixit and Pindyck: *The simple NPV rule is not just wrong; it is often very wrong*[14].

[10] See Amram & Kulatilaka [2].
[11] See Copeland & Antikarov [33].
[12] See Hommel, Scholich & Vollrath [61].
[13] See Hommel, Scholich & Baecker [60].
[14] See Dixit & Pindyck [40], page 136.

However, it has to be stressed that a real options approach is not necessary in each and every investment situation. Amram and Kulatilaka developed a list of criteria that show under which circumstances the real options approach is fruitful[15]: *A real options analysis is needed in the following situations:*

1. *When there is a contingent investment decision. No other approach can correctly value this type of opportunity.*

2. *When uncertainty is large enough that it is sensible to wait for more information, avoiding regret for irreversible investment.*

3. *When the value seems to be captured in possibilities for future growth options rather than current cash flow.*

4. *When uncertainty is large enough to make flexibility a consideration. Only the real options approach can correctly value investments in flexibility.*

5. *When there will be project updates and mid-course strategy corrections.*

Both authors also categorize strategic investments and put them into the perspective of real options[16]. The various types of strategic investments are as follows:

1. **Irreversible investments:**
   Once these investments are in place, they cannot be reversed without losing much of their value. The value of an irreversible investment with its associated options is greater than recognized by traditional tools because the options truncate the loss.
2. **Flexibility investments:**
   These are investments that incorporate flexibility in the form of options into the initial design. This is hardly achievable with traditional valuation tools.
3. **Insurance investments:**
   These investments are investments that reduce the exposure to uncertainty, i.e., a put option.
4. **Modular investments:**
   Modular investments create options through product design. Each module has a tightly specified interface to the other, allowing modules to be independently changed and upgraded. A modular product can be viewed as a portfolio of options to upgrade.

[15] See Amram & Kulatilaka [2], page 24.
[16] See Amram & Kulatilaka [2], pape 25-27.

5. **Platform investments:**
   These investments create valuable follow-on contingent investment opportunities. This is particularly important in R&D, *the* classic platform investment. The value of such an investment results from products it releases for further development that may in turn lead to marketable products. Traditional tools greatly undervalue these investments, whereas the real options approach is ideally suited to value such an investment.
6. **Learning investments:**
   Learning investments are investments that are made to obtain information that is otherwise unavailable. A typical learning investment is the exploration work undertaken on a site prior to exploiting it.

Based on this list of different investment types, a categorization was developed and is currently widely used in Finance. This list is based on Trigeorgis [133], pages 1-14. Trigeorgis groups real options into the following classes:

- **Option to defer (learning option):**
  This includes options where the time point of an investment is not determined but flexible allowing this time point to be optimized. Those options can also arise from changes in the term structure of interest rates over time even if the future cash flow is deterministic. Given the sharp interest rates decline in the U.S. in 2001 as seen in Section 1.1 (Figure 1.3), this option is of special use for firms in volatile capital markets: The NPV increases as interest rates decrease and it may be optimal to invest at a later point of time.
- **Time-to-build option:**
  This is an option that allows to stop a step-by-step investment within a project if market conditions turn unfavorable. Such an option is particularly valuable in R&D.
- **Option to alter the operating scale:**
  This is the option to react upon a changing market and to expand operations (favorable market conditions) or to scale down operations (unfavorable market conditions). Such an option can be implemented when a firm wants to introduce a new product or would like to enter a new market, for example in the consumer goods industry.
- **Option to abandon (put option, insurance):**
  The option to sell a project is called *option to abandon*. The value that can be regained by selling the project (*salvage value*) is included in the pricing of the project and can substantially alter the project's NPV calculation.
- **Option to switch:**
  This option comprises the possibility to react upon changed market conditions by changing the input and output factors via input shifts and/or output shifts. This is *the* classical real option.

- **Growth option:**
  Growth options are strategic options. They are particularly revelant for projects which are not advantageous in themselves but may generate lucrative opportunities in the future. This type of option can especially be found in R&D[17]. In the pharmaceutical industry, e.g., it takes more than ten years for a product to develop from the original idea to the final product with a success probability of only a few percent[18]. However, during the course of a project the original investment may generate various other applications which are profitable and generate a positive NPV for the whole investment.
- **Multiple interacting options:**
  Multiple interacting options are combinations of real options of the types described above. In practice, of course, they are the most frequent ones.

To price (real) options different methods can be applied. An overview is given in Figure 2.1. A broad classification and qualitative discussion of these can be found in Hommel & Lehmann [59], Chapter 3. A similar overview with an in-depth discussion and mathematical descriptions of some specific methods can be found in Schulmerich [118], pages 64-67.

Although the real options theory has been a topic of research for over 25 years it has been made broadly accessible to practitioners in Corporate Finance only since the mid 1990s. Therefore, the diffusion of this theory in capital budgeting has still a long way to go. These are the main results of an article by Pritsch and Weber[19] who found that real options pricing models are not well known among senior managers accustomed to the NPV method. The importance of the real options approach in the academic field has yet to be gained in corporate financial practice[20]. As Pritsch and Weber point out, this will not be easily achieved. However, they also emphasize that the same held true for the NPV method in the last decades[21]. In Figure 2.2 they developed an overview of the diffusion of the NPV method in U.S. companies in the last decades[22].

[17] For a detailed analysis of R&D and the valuation approach with multi-dimensional models on American compound options in that area see Pojezny [107].
[18] See Schwartz & Moon [122], pages 87-89.
[19] See Pritsch & Weber [108].
[20] See Pritsch & Weber [108], page 38.
[21] See Pritsch & Weber [108], page 31.
[22] Figure 2.2 is based on the articles Klammer & Walker [76] (for 1955 and 1984), Klammer [75] (for 1959, 1964, and 1970), Gitman & Forrester [47] (for 1970 and 1977) as well as Reichert, Moore & Byler [110] (for 1992).

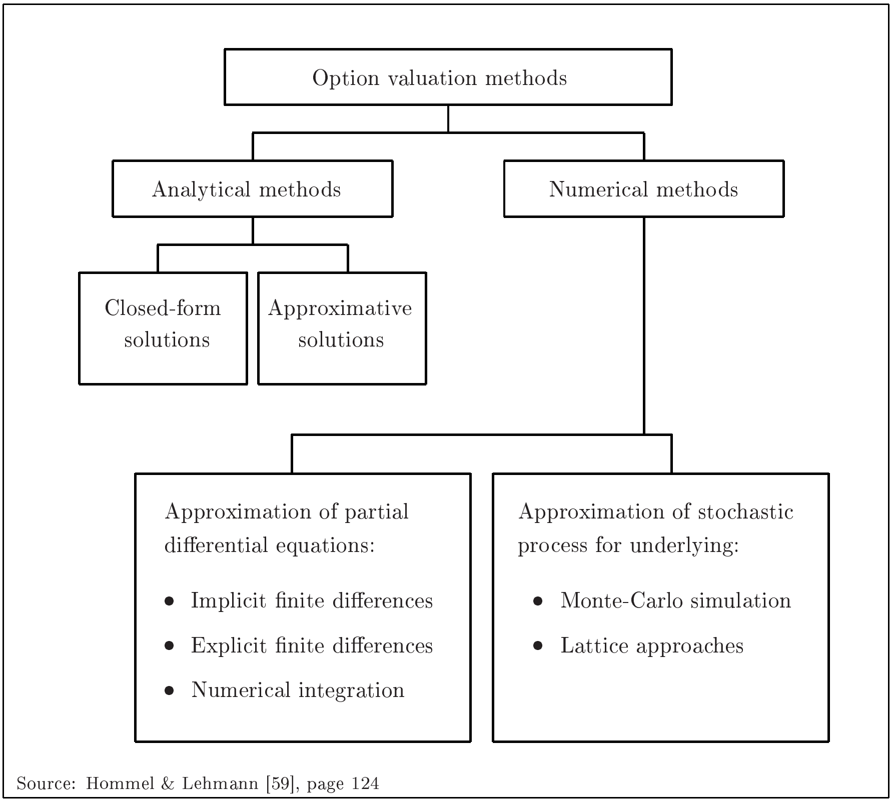

Fig. 2.1. Classification of valuation methods for real options.

A thorough analysis of the importance of the real options approach in capital budgeting practice was undertaken only by a few others[23]. One of the latest and best empirical studies was done by Vollrath[24] who in 2000 investigated a sample of companies with headquarters in Germany. He found that although the real options approach is theoretically superior to traditional capital budgeting tools, it is not widespread in companies.

A detailed study that reveals similar findings was undertaken by Bubsy and Pitts[25] in 1997 for all firms in the FTSE-100 index in the U.K.[26]. This study shows that decision-makers in those firms intuitively realize the importance

[23] See Vollrath [135], pages 46-52.
[24] See Vollrath [135].
[25] See Bubsy & Pitts [24].
[26] Although all FTSE-100 firms were asked to participate, only 44% finally took part and answered the questionnaire.

of flexibility but are not fully aware of the academic theory of real options[27]. Moreover, only about 50% of the managers interviewed in this study saw the necessity to quantitatively implement a real options approach for their capital budgeting. Vollrath's main results are summarized[28] in the following:

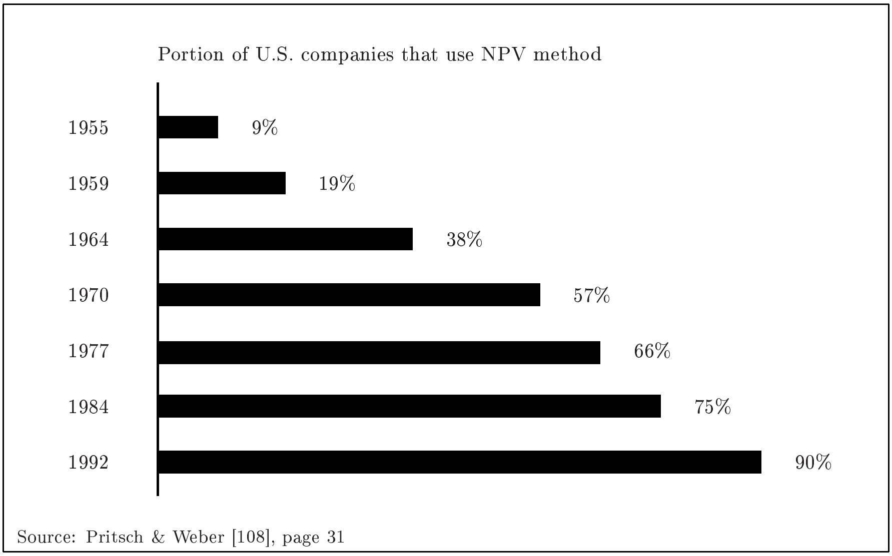

Fig. 2.2. Diffusion of the NPV method in U.S. companies.

1. **Implemented methods for capital budgeting:**
   The major investment decisions are made at the corporate headquarters with several valuation tools used in parallel. However, the real options approach is not applied.

2. **Implementation of methods:**
   Investment projects will be undertaken due to a strategic reason even if the NPV is negative.

3. **Intuitive realization of real options values:**
   Flexibility is intuitively incorporated and assigned a positive value. However, there is no formal option valuation due to the complexity of options pricing. Furthermore, the influence of various parameters on the real option cannot be evaluated sufficiently.

[27] See Vollrath [135], page 50.
[28] See Vollrath [135], pages 72-73, for the complete analysis.

This book intends to bridge the gap between the theoretical and the practical importance of real options by providing various real options valuation tools that can readily be used in practice[29]. Moreover, by applying a stochastic term structure instead of a constant risk-free interest rate, the book intends to improve these numerical methods in order to model reality more accurately.

2.3 Diffusion Processes in Classical Probability Theory

When dealing with real options pricing and term structure models the mathematical tool called Ito calculus or stochastic calculus has to be used. Term structure models are just one example that uses diffusion processes to describe the movement of the yield curve over time. Diffusion processes are also used to model the underlying of options or to model cost processes as done in the Schwartz-Moon model that will be described in Section 4.3. Diffusion processes are best described with stochastic calculus since classical probability theory can only give limited insight into this type of stochastic process. Nevertheless, it is advisable to introduce diffusion processes using classical probability theory since the parameters of a diffusion process can be best explained with classical probability theory. Therefore, this section is devoted to the introduction of a diffusion process using classical probability theory.

This section closely follows Schulmerich [117], pages 3-14, and assumes that the theory of continuous-time Markov processes is known. The state space of a process is always an interval I with left interval point l and right interval point r. The interval I has one of the four forms $(l,r), (l,r], [l,r)$ or $[l,r]$ with $l,r \in \mathbb{R} \cup \{\pm\infty\}$.

2.3.1 Definition: Diffusion Process - 1st Version

A continuous-time stochastic process $(X_t)_{t\geq 0}$, which has the strong Markov property[30], is called (one-dimensional) *diffusion process*, if all paths of the process are smooth functions in t in the almost sure sense.

This definition is called *1st version* since another definition of a diffusion process (called *2nd version*) will be given in connection with *Ito calculus*. The relationship between these two definitions will then be explained in detail.

[29] This has been pointed out as an important point of critique and as an area of future research by Trigeorgis, see Trigeorgis [133], page 375, Section 12.3 - Future Research Directions: *Developing generic options-based user-friendly software packages with simulation capabilities that can handle multiple real options as a practical aid to corporate planners.* This is a major purpose of this book which is realized with the computer simulation program that was developed for all analyses in Chapter 5 based on the models presented in Chapters 3 and 4.

[30] See Karlin & Taylor [73], page 149.

2.3.2 Definitions: Infinitesimal Expectation and Variance

Let $(X_t)_{t\geq 0}$ be a diffusion process with real limits

$$\lim_{h\downarrow 0} \frac{1}{h} E(\Delta_h X_t \,|\, X_t = x) \quad =: \quad \mu(x,t), \tag{2.1}$$

$$\lim_{h\downarrow 0} \frac{1}{h} E([\Delta_h X_t]^2 \,|\, X_t = x) \quad =: \quad \sigma^2(x,t) \tag{2.2}$$

for each $l < x < r$ and $t \geq 0$, where l and r are the left and right interval points of the state space I and $\Delta_h X_t := X_{t+h} - X_t$. The functions $\mu(x,t)$ and $\sigma^2(x,t)$ are the infinitesimal parameters of the diffusion process; $\mu(x,t)$ is called *infinitesimal expectation*, *expected infinitesimal change* or *(infinitesimal) drift*; $\sigma^2(x,t)$ is called *infinitesimal variance* or *diffusion parameter*. Functions μ and σ^2 are called *infinitesimal expectation* and *infinitesimal variance*, respectively, since from (2.1) and (2.2) it is known[31]:

$$E(\Delta_h X_t \,|\, X_t = x) = \mu(x,t)\, h \;\; + \;\; o_1(h), \qquad \lim_{h\downarrow 0} \frac{o_1(h)}{h} = 0,$$

$$Var(\Delta_h X_t \,|\, X_t = x) = \sigma^2(x,t)\, h \;\; + \;\; o_2(h), \qquad \lim_{h\downarrow 0} \frac{o_2(h)}{h} = 0.$$

Here, $o_i, i = 1, 2$, is the Landau symbol. Generally, $\mu(x,t)$ and $\sigma^2(x,t)$ are smooth functions of x and t.

In this book various types of diffusion processes will be used. Some of them are *time-homogeneous* diffusion processes[32], i.e., $\mu(x,t) = \mu(x)$ and $\sigma^2(x,t) = \sigma^2(x)$ are independent of t (see Karlin & Taylor [73], page 160). Other diffusion processes have parameters that depend only on time t but not on the state of the process, such as the short-rate process in the Ho-Lee model. Other processes depend both on time t and state x like in the Hull-White one-factor model. In some cases (see the Hull-White two-factor model) the infinitesimal drift even contains a random component that is modelled via a second diffusion process, an Ornstein-Uhlenbeck process.

In the remainder of Section 2.3 the focus is on time-homogeneous processes, i.e., diffusion processes with $\mu(x,t) = \mu(x)$ and $\sigma^2(x,t) = \sigma^2(x)$ independent of t. It is assumed that μ and σ are smooth in x, and $I = (l, r)$ is assumed to be the state space of the diffusion process. This provides the mathematical tools to handle the Ornstein-Uhlenbeck process, the Vasicek model, and the Cox-Ingersoll-Ross model in Chapter 3. The goal of this section is to derive the stationary distribution of those diffusion processes (see 2.3.4).

[31] See Karlin & Taylor [73], page 160.

[32] The short-rate processes in the Vasicek model and in the Cox-Ingersoll-Ross model are examples of time-homogeneous diffusion processes as will be seen later.

2.3.3 Denominations for Diffusion Processes

Let $(X_t)_{t\geq 0}$ be a diffusion process with the infinitesimal parameters $\mu(x)$ and $\sigma^2(x)$, $l < x < r$, where $I = (l, r)$ is the state space of the process. Via

$$S(x) \;:=\; \int_c^x \exp\left\{ -\int_{\eta_0}^{\eta} \frac{2\mu(\xi)}{\sigma^2(\xi)} d\xi \right\} d\eta, \quad l < x < r,$$

the function $S : I \longrightarrow \mathbb{R}$ is defined and referred to as the *scale function* of the diffusion process. Parameters c and η_0 are flexible constants[33] from the state space I. The first derivative of S,

$$S'(\eta) \;=\; \exp\left\{ -\int_{\eta_0}^{\eta} \frac{2\mu(\xi)}{\sigma^2(\xi)} d\xi \right\}, \quad l < \eta < r,$$

will be denoted with $s(\eta)$. Via

$$M(x) \;:=\; \int_{\tilde{c}}^{x} \frac{2}{\sigma^2(\eta)\, s(\eta)}\, d\eta, \quad l < x < r,$$

the function $M : I \longrightarrow \mathbb{R}$ is defined and referred to as the *speed function* of the diffusion process. Here, $\tilde{c}$ is a flexible constant from the state space I. In a natural manner a measure on I can be defined, which will be (as common in literature) denoted with M:

$$M((a, b]) \;:=\; M(b) - M(a), \quad l < a < b < r.$$

The measure M on I is called *speed measure* of the diffusion process[34]. Via

$$m(x) \;:=\; M'(x) \;=\; \frac{2}{\sigma^2(x)s(x)}, \quad l < x < r,$$

the function $m : I \longrightarrow \mathbb{R}$ is defined, which is called *speed density* of the diffusion process. This density is a *Radon-Nikodym density.*

The functions S and s are unique only up to an affine transformation with a monotonically increasing transformation function in the case of S and up to a positive multiplicative factor in the case of s. For a given S, the

[33] The term *flexible constant* refers to the mathematical idea that the choice of c and η does not matter. However, if certain values for c and η_0 are chosen, these values cannot be changed any more.

[34] It is common in literature to use the M for the speed measure and for the speed function. This does not create confusion since both functions operate on different sets.

speed density m is unique. A process with scale function S in the form of $S(x) = C_1 + C_2 x$, $x \in I$ with $C_1, C_2 \in \mathbb{R}$ constant, is called *in natural scale/canonical scale.*

According to Karlin & Taylor [73], page 194, the speed measure M is a measure $M : I \longrightarrow \mathbb{R}$, such that for each $a < x < b$ with $a, x, b \in I$ it holds:

$$E(T_{a \wedge b} | X_0 = x) = \int\limits_{(a,b)} G_{S(a),S(b)}\left(S(x), S(y)\right) dM(y) \quad \text{with} \tag{2.3}$$

$$G_{q_1,q_2}(u,v) := \left\{ \begin{array}{ll} \dfrac{[u-q_1]\,[q_2-v]}{q_2-q_1}, & q_1 < u \le v < q_2 \\ \dfrac{[q_2-u]\,[v-q_1]}{q_2-q_1}, & q_1 < v \le u < q_2 \end{array} \right\} \tag{2.4}$$

as the *Green function* on $(q_1, q_2) \subseteq \mathbb{R}, q_1 < q_2$.

Knowing this relationship helps to understand the meaning of the speed density of a diffusion process $(X_t)_{t \ge 0}$, which is in natural scale, with state space I. For this let $x \in I$ be flexible but constant and let $\varepsilon > 0$ be so small that $[x-\varepsilon, x+\varepsilon]$ lies in interval I. Then:

$$\begin{aligned}
& E(T_{x-\varepsilon \wedge x+\varepsilon} | X_0 = x) \\
\overset{(2.3)}{=} & \int\limits_{(x-\varepsilon,x+\varepsilon)} G_{S(x-\varepsilon),S(x+\varepsilon)}\left(S(x), S(y)\right) dM(y) \\
= & \int\limits_{(x-\varepsilon,x+\varepsilon)} G_{x-\varepsilon,x+\varepsilon}(x,y)\, dM(y) \\
= & \int\limits_{(x-\varepsilon,x)} G_{x-\varepsilon,x+\varepsilon}(x,y)\, dM(y) \;+\; \int\limits_{[x,x+\varepsilon)} G_{x-\varepsilon,x+\varepsilon}(x,y)\, dM(y) \\
\overset{(2.4)}{=} & \int\limits_{x-\varepsilon}^{x} \frac{([x+\varepsilon]-x)(y-[x-\varepsilon])}{[x+\varepsilon]-[x-\varepsilon]}\, m(y)\, dy \;+ \\
& \qquad \int\limits_{x}^{x+\varepsilon} \frac{(x-[x-\varepsilon])([x+\varepsilon]-y)}{[x+\varepsilon]-[x-\varepsilon]}\, m(y)\, dy \\
= & \frac{1}{2} \left\{ \int\limits_{x-\varepsilon}^{x} (y-x+\varepsilon)\, m(y)\, dy \;+\; \int\limits_{x}^{x+\varepsilon} (x+\varepsilon-y)\, m(y)\, dy \right\}.
\end{aligned}$$

Therefore:

$$
\begin{aligned}
& \frac{1}{\varepsilon^2}\, E(T_{x-\varepsilon \wedge x+\varepsilon} | X_0 = x) \\
= \quad & \frac{1}{2} \left\{ \frac{1}{\varepsilon^2} \int\limits_{x-\varepsilon}^{x} (y - x + \varepsilon)\, m(y)\, dy \;\; + \;\; \frac{1}{\varepsilon^2} \int\limits_{x}^{x+\varepsilon} (x + \varepsilon - y)\, m(y)\, dy \right\} \\
= \quad & \frac{1}{2} \left\{ \frac{1}{\varepsilon^2} (x - \delta_1(\varepsilon)\varepsilon - x + \varepsilon)\, m(x - \delta_1(\varepsilon)\varepsilon)\, (x - [x - \varepsilon]) \right. \\
& \quad + \quad \left. \frac{1}{\varepsilon^2} (x + \varepsilon - (x + \delta_2(\varepsilon)\varepsilon))\, m(x + \delta_2(\varepsilon)\varepsilon)\, ([x + \varepsilon] - x) \right\}, \\
& \quad \delta_1(\varepsilon) \text{ and } \delta_2(\varepsilon) \in [0,1] \text{ appropriate with} \\
& \quad \delta_i \xrightarrow[\varepsilon \to 0]{} 0 \;\; (i = 1, 2) \\
= \quad & \frac{1}{2} \left\{ (1 - \delta_1(\varepsilon))\, m(x - \delta_1(\varepsilon)\varepsilon) \;\; + \;\; (1 - \delta_2(\varepsilon))\, m(x + \delta_2(\varepsilon)\varepsilon) \right\} \\
\xrightarrow[\varepsilon \to 0]{} \; & m(x), \quad (m \text{ is smooth, see 2.3.2 and definition of } m).
\end{aligned}
$$

This means that the speed density $m(x)$ is the speed with which the process $(X_t)_{t \geq 0}$ moves forward if it is in state $x \in I$.

2.3.4 Stationary Distribution of a Diffusion Process

Let $(X_t)_{t \geq 0}$ be a diffusion process with state space $I = (l, r)$, whereby the process is *in natural scale*. Let

$$
|m| \quad := \quad \int\limits_{l}^{r} m(x)\, dx
$$

be the total mass of the speed density m on I. If $|m| < \infty$ then there exists a stationary distribution of the process $(X_t)_{t \geq 0}$ which has a Lebesgue measure with density

$$
\psi(x) \quad = \quad \frac{m(x)}{|m|}, \quad x \in I.
$$

Since m and $|m|$ are unique up to the same multiplicative constant (see 2.3.3), ψ is uniquely defined. The proof can be found in Rogers & Williams [113], page 303, theorem 54.5. According to Schulmerich [117], Section 1.16, the restriction *in natural scale* on the diffusion process $(X_t)_{t \geq 0}$ above can be omitted.

This is the main result of 2.3. In the following an example is given of a special diffusion process, the one-dimensional Brownian Motion.

2.3.5 Example: One-Dimensional Brownian Motion

A stochastic process $(B_t)_{t\geq 0}$ with $B_0 = 0$ is called *one-dimensional Brownian Motion* (or *Wiener process*) with drift $\mu \in \mathbb{R}$, if for each choice of $0 \leq t_0 < t_1 < \ldots < t_{n-1} < t_n$, $n \in \mathbb{N}$, the joint distribution of $B_{t_0}, B_{t_1}, \ldots, B_{t_n}$ is determined by two conditions:

1. the $(B_{t_k} - B_{t_{k-1}}),\ \ k = 1, \ldots, n$, are independent and
2. $(B_{t_k} - B_{t_{k-1}}) \stackrel{d}{=} N([t_k - t_{k-1}]\mu, [t_k - t_{k-1}]\sigma^2),\ \ k = 1, \ldots, n$, for suitable constants $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$.

If $\mu = 0$ and $\sigma^2 = 1$, this process is called a *Standard Brownian Motion.* A Standard Brownian Motion $(B_t)_{t\geq 0}$ is a time-homogeneous diffusion process with state space $I = (-\infty, \infty)$.

If a Standard Brownian Motion $(B_t)_{t\geq 0}$ is given, a Brownian Motion $Y := (Y_t)_{t\geq 0}$ with drift $\mu \in \mathbb{R}$ and volatility $\sigma \in \mathbb{R}^+$ can be derived via $Y_t := \sigma B_t + \mu t, t \geq 0$. By doing so the state space $I = (-\infty, \infty)$ remains the same. $(Y_t)_{t\geq 0}$ is a time-homogeneous diffusion process as well and has the infinitesimal parameters $\mu_Y(x) = \mu$ and $\sigma_Y^2(x) = \sigma^2 \in \mathbb{R}^+, x \in I$. According to 2.3.3 it holds for each $x \in \mathbb{R}$:

Case I with $\mu \neq 0$:

$$\begin{aligned}
s(x) &= \exp\left\{-\frac{2\mu x}{\sigma^2}\right\}, \quad \text{where } \eta_0 := 0 \in I, \\
S(x) &= \frac{\sigma^2}{2\mu}\left(1 - \exp\left\{-\frac{2\mu x}{\sigma^2}\right\}\right), \quad \text{where } c := 0, \\
M(x) &= \frac{1}{\mu}\left(\exp\left\{\frac{2\mu x}{\sigma^2}\right\} - 1\right), \quad \text{where } \tilde{c} := 0, \\
m(x) &= \frac{2}{\sigma^2}\exp\left\{\frac{2\mu x}{\sigma^2}\right\}.
\end{aligned}$$

Case II with $\mu = 0$:

$$\begin{aligned}
s(x) &= 1, \quad \text{where } \eta_0 := 0 \in I, \\
S(x) &= x, \quad \text{where } c := 0, \\
M(x) &= \frac{2}{\sigma^2}x, \quad \text{where } \tilde{c} := 0, \\
m(x) &= \frac{2}{\sigma^2}.
\end{aligned}$$

The Geometric Brownian Motion $Z = (Z_t)_{t\geq 0}$ can be derived via $Z_t := \exp\{Y_t\},\ \ t \geq 0$. The resulting state space I is then $\mathbb{R}^+$ and, according to the transformation formula in Karlin & Taylor [73], pages 173-175, it holds for the Geometric Brownian Motion:

$$\mu_Z(z) = \left(\mu + \frac{\sigma^2}{2}\right) z \qquad \text{and} \qquad \sigma_Z^2(z) = \sigma^2 z^2, \quad z \in \mathbb{R}^+.$$

2.4 Introduction to the Ito Calculus

This section introduces the language of Financial Engineering, the *stochastic calculus*, referred to as *Ito calculus*. While traditional probability theory is well suited to introduce the notion of a diffusion process as done in the previous section, the main tool for describing and especially for calculating such a process is stochastic calculus. This introduction concentrates on presenting results without giving a detailed proof. However, literature will always be cited where a detailed proof and further information about any of the presented topics can be found. The following introduction closely follows Schulmerich [117], pages 15-28, but focuses on those aspects of the theory that will be needed for this book.

2.4.1 Definition: Ito Integral

Let $(\Omega, \mathcal{F}, P)$ be the probability space and $\mathcal{B}$ the *Borel σ-Field* on $\mathbb{R}_0^+$.

1. Let $\{\mathcal{N}_t | t \geq 0\}$ be a family of increasing σ-fields over Ω. A process $h : \mathbb{R}_0^+ \times \Omega \longrightarrow \mathbb{R}$ is called $(\mathcal{N}_t)_{t \geq 0}$*-adapted* if for each $t \geq 0$ the function $\omega \longmapsto h(t, \omega)$ is $\mathcal{N}_t$-measurable, $\omega \in \Omega$.
2. Let $(B_s)_{s \geq 0}$ be a one-dimensional Standard Brownian Motion. Let $\mathcal{V} := \mathcal{V}(S, T)$, $S, T \in \mathbb{R}_0^+$ with $S < T$ be a class of functions $f : \mathbb{R}_0^+ \times \Omega \longrightarrow \mathbb{R}$ with the following properties:

   (a): $(t, \omega) \longmapsto f(t, \omega)$ is $\mathcal{B} \times \mathcal{F}$-measurable, $t \geq 0$, $\omega \in \Omega$, i.e.: function f is *progressively measurable.*

   (b): f is $\mathcal{F}_t$-adapted, where $\mathcal{F}_t := \sigma(B_s | 0 \leq s \leq t)$, $t \geq 0$.

   (c): $E\left(\int_S^T f(t, \omega)^2 dt\right) < \infty.$
3. For S and T from point 2 above and for fixed $n \in \mathbb{N}$ let $t_j^{(n)}, j \in \mathbb{N}_0$, be defined by

$$t_j^{(n)} \quad := \quad \begin{cases} S, & \text{if } \; j\,2^{-n} < S, \\ j\,2^{-n}, & \text{if } \; S \leq j\,2^{-n} \leq T, \\ T, & \text{if } \; T < j\,2^{-n}. \end{cases}$$

   A function $\lambda_n : \mathbb{R}_0^+ \times \Omega \longrightarrow \mathbb{R}$ from $\mathcal{V}$ is called *elementary function*, if it has the form

$$\lambda_n(t,\omega) \quad = \quad \sum_{j\in\mathbb{N}_0} e_j(\omega)\, 1_{[t_j^{(n)},t_{j+1}^{(n)})}(t), \quad t \geq 0,\ \omega \in \Omega.$$

e_j is a random variable on $(\Omega, \mathcal{F}, P)$ that needs to be $\mathcal{F}_{t_j}$-measurable because of $\lambda_n \in \mathcal{V}$. $1_{[t_j^{(n)},t_{j+1}^{(n)})}$ is the indicator function with respect to the set $[t_j^{(n)}, t_{j+1}^{(n)})$, $j \in \mathbb{N}_0$.

For an elementary function λ_n the *Ito integral* $\int_S^T \lambda_n(t,\omega)\, dB_t(\omega)$ is defined for each $\omega \in \Omega$ as

$$\int_S^T \lambda_n(t,\omega)\, dB_t(\omega) \quad := \quad \sum_{j\in\mathbb{N}_0} e_j(\omega)[B_{t_{j+1}^{(n)}} - B_{t_j^{(n)}}](\omega).$$

According to Øksendal [103], pages 23-25, it holds:
For each $f \in \mathcal{V}$ there exist elementary functions $\lambda_n \in \mathcal{V}, n \in \mathbb{N}$, with

$$E\left(\int_S^T |f - \lambda_n|^2\, dt\right) \quad \xrightarrow[n \to \infty]{} \quad 0.$$

For such a function f the *Ito integral* $\int_S^T f(t,\omega) dB_t(\omega), \omega \in \Omega$, is defined as

$$\int_S^T f(t,\omega)\, dB_t(\omega) \quad := \quad \lim_{n\to\infty} \int_S^T \lambda_n(t,\omega)\, dB_t(\omega).$$

This limit exists as an element of the set $L_2(\Omega, P)$. The Ito integral is well-defined, i.e., independent from the choice of the sequence of elementary functions $(\lambda_n)_{n\in\mathbb{N}}$.

2.4.2 Properties of an Ito Integral

The following rules hold for an Ito integral:

1. $E\left(\left[\int_S^T f\, dB_t\right]^2\right) = E\left(\int_S^T f^2\, dt\right) \quad \forall f \in \mathcal{V}(S,T)$ *(Ito isometry).*

2. For $f, g \in \mathcal{V}(0,T)$ it holds with $0 \leq S < U < T$:

$$(a): \quad \int_S^T f\, dB_t = \int_S^U f\, dB_t + \int_U^T f\, dB_t \quad \text{almost sure.}$$

$$(b): \quad \int_S^T (cf+g)\,dB_t = c\int_S^T f\,dB_t + \int_S^T g\,dB_t \quad \text{almost sure, } c \in \mathbb{R}.$$

$$(c): \quad E\left(\int_S^T f\,dB_t\right) = 0.$$

$$(d): \quad \left(\int_S^T f\,dB_t\right)_{T \geq S} \quad \text{is a martingale.}$$

$$(e): \quad \int_S^T f\,dB_t \quad \text{is } \mathcal{F}_T\text{-measurable.}$$

3. For each $f \in \mathcal{V}(0,T)$ there exists an in t smooth version of the Ito integral

$$\int_0^t f\,dB_s, \quad 0 \leq t \leq T,$$

i.e., there exists an in t smooth process $(X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, P)$ with

$$P\left(X_t = \int_0^t f\,dB_s\right) = 1 \quad \text{for each} \quad 0 \leq t \leq T.$$

In the following only in t smooth Ito integrals will be considered. The proof of these properties can be found in Øksendal [103], pages 26-30.

2.4.3 Definition: A Broader Class of Ito Integrals

1. The Ito integral can also be defined for a broader class of integrands than $\mathcal{V}$: Let $\mathcal{W} \supset \mathcal{V}$ be the class of functions $g : \mathbb{R}_0^+ \times \Omega \longrightarrow \mathbb{R}$ with the following properties:

   a) g is progressively measurable.

   b) There exists a family $(\mathcal{H}_t)_{t\geq 0}$ of increasing σ-fields, such that $(B_t)_{t\geq 0}$ is a martingale with respect to $(\mathcal{H}_t)_{t\geq 0}$ and such that $g(t,\,\cdot\,)$ is adapted on $\mathcal{H}_t$, $t \geq 0$.

   c) $P\left(\left\{\omega \in \Omega \;\middle|\; \int_0^t g(s,\omega)^2\,ds < \infty \quad \forall\, t \geq 0\right\}\right) = 1.$

Analogous to the description in 2.4.1, the Ito integral can also be defined for $g \in \mathcal{W}$. Yet this will not be done here. For details see Øksendal [103], pages 31-32, whose definition will be used in the following. All of the properties in 2.4.2 are fullfilled for each $g \in \mathcal{W}$ as well (see Karatzas & Shreve [72], Chapter 3.2).

2. It is also possible to define the Ito integral from 2.4.1 not only with respect to the Standard Brownian Motion $(B_t)_{t\geq 0}$ as done in the original construction from Ito, but also, under suitable conditions, with respect to each stochastic process $(M_t)_{t\geq 0}$ that is a smooth, quadratic integrable martingale. This broader definition of the stochastic integral can be found in Karatzas & Shreve [72], pages 128-139.

2.4.4 Definition: One-Dimensional Ito Process

Let $(B_t)_{t\geq 0}$ be a one-dimensional Standard Brownian Motion on a probability space $(\Omega, \mathcal{F}, P)$. A *one-dimensional Ito process* or *stochastic integral* is a stochastic process $(X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, P)$ of the form

$$X_t(\omega) = X_0(\omega) + \int_0^t a(s,\omega)ds + \int_0^t b(s,\omega)dB_s(\omega), \;\; \omega \in \Omega \quad \text{and} \quad t \geq 0. \tag{2.5}$$

Here, it is assumed that $b \in \mathcal{W}$ holds and that for a the following equation holds true:

$$P\left(\left\{\omega \in \Omega \;\middle|\; \int_0^t |a(s,\omega)|\, ds \;<\; \infty \quad \forall\, t \geq 0\right\}\right) \quad = \quad 1.$$

Moreover, let a be adapted to $(\mathcal{H}_t)_{t\geq 0}$ where $(\mathcal{H}_t)_{t\geq 0}$ is a family of increasing σ-fields from $\mathcal{W}$. The differential notation for (2.5) is:

$$dX_t(\omega) \;=\; a(t,\omega)\, dt \;+\; b(t,\omega)\, dB_t(\omega), \quad t \geq 0,$$

or in short notation:

$$dX_t \;=\; a\, dt \;+\; b\, dB_t, \quad t \geq 0. \tag{2.6}$$

(2.6) is called *stochastic differential equation* (*SDE*).

The usage of $f \in \mathcal{V}$ in 2.4.1 as integrand in the Ito integral is more specific than the usage of random variables $X_t, t \geq 0$. In particular, each element $f \in \mathcal{V}$ is a stochastic process. The reverse of this statement is not true since not each stochastic process fulfills the conditions (a) and (b) of point 2 in 2.4.1. However, if the notation from definition 2.4.4 is applied, it yields: For each stochastic process $(X_t)_{t\geq 0}$ the parameters $a(t, X_t)$ and $b(t, X_t)$ fulfill the conditions (a) and (b) of point 2 in 2.4.1.

2.4.5 Theorem: Existence and Uniqueness of the Solution of a Stochastic Differential Equation

With $T > 0$ let $a : [0,T] \times \mathbb{R} \longrightarrow \mathbb{R}$ and $b : [0,T] \times \mathbb{R} \longrightarrow \mathbb{R}$ be Lebesgue-measurable functions. It is further assumed that the following two conditions hold:

(i) $\quad |a(t,x)| + |b(t,x)| \;\leq\; C\,(1+|x|),\;\; x \in \mathbb{R},\;\; t \in [0,T],$

$\qquad C \in \mathbb{R}^+$ appropriate *(growth condition).*

(ii) $\quad |a(t,x) - a(t,y)| + |b(t,x) - b(t,y)| \;\leq\; D|x-y|,\;\; x,y \in \mathbb{R},$

$\qquad t \in [0,T],\;\; D \in \mathbb{R}^+$ appropriate *(Lipschitz condition).*

Let $(B_t)_{t\geq 0}$ be a one-dimensional Standard Brownian Motion and let Z be a random variable with $E(|Z|^2) < \infty$ that is independent from $\mathcal{F}_\infty := \sigma(B_s|s \geq 0)$. Then the SDE

$$dX_t \;=\; a(t,X_t)\,dt \;+\; b(t,X_t)\,dB_t,\;\; 0 \leq t \leq T,\;\; X_0 = Z, \qquad (2.7)$$

has a unique and in t smooth solution $(X_t)_{t\geq 0}$ that is an element of $\mathcal{V}(0,T)$. More precisely, the process $(X_t)_{t\geq 0}$ is referred to as a *strong solution* since $(B_t)_{t\geq 0}$ is given and the solution process is $\mathcal{F}_t$-adapted, $t \geq 0$. A solution is called a *weak solution* if the Standard Brownian Motion is not given but a pair $((B_t)_{t\geq 0}, (X_t)_{t\geq 0})$ is searched for that fulfills (2.7).

A strong solution $(X_t)_{t\geq 0}$ of (2.7) possesses the strong Markov property and fulfills

$$E\left(|X_t|^2\right) \;<\; \infty \qquad \forall\; t \geq 0.$$

If μ and σ do not depend on t, i.e., if the SDE

$$dX_t \;=\; a(X_t)\,dt \;+\; b(X_t)\,dB_t, \quad t \geq 0,\;\; X_0 = Z,$$

is considered, (i) and (ii) are simplified to

(iii) $\quad |a(x) - a(y)| + |b(x) - b(y)| \;\leq\; D\,|x-y|, \quad x,y \in \mathbb{R}.$

SDE (2.7) is the type of an SDE that describes most of the term structure models in Chapter 3, e.g., the Vasicek model, the Cox-Ingersoll-Ross (CIR) model, the Ho-Lee model, and the Hull-White one-factor model.

The goal for the remainder of this section is to give explicit solutions for some specific SDEs. These formulas will be then further used in Chapter 3 to describe various term structure models.

2.4.6 Coefficients of an Ito Process

According to Karlin & Taylor [73], page 376, it can be shown: If an Ito process from 2.4.4 fulfills the criteria of definition 2.3.1 and also fulfills the conditions 2.4.5 (i), (ii), then $a \equiv \mu$ and $b \equiv \sigma$, using the notation μ and σ from definition 2.3.2. In the following only the notation μ and σ is used. σ is also called *volatility*.

2.4.7 Definition: Diffusion Process - 2nd Version

A one-dimensional diffusion process $(X_t)_{t\geq 0}$ is a stochastic process that fulfills the SDE

$$dX_t \;=\; \mu(t, X_t)\, dt \;+\; \sigma(t, X_t)\, dB_t, \;\; t \geq 0, \;\; X_0 = x_0 \in \mathbb{R} \text{ constant},$$

where $\mu : \mathbb{R}_0^+ \times \mathbb{R} \longrightarrow \mathbb{R}$ and $\sigma : \mathbb{R}_0^+ \times \mathbb{R} \longrightarrow \mathbb{R}$ fulfill the conditions 2.4.5 (i), (ii), and $(B_t)_{t\geq 0}$ is a one-dimensional Standard Brownian Motion. $(X_t)_{t\geq 0}$ is also called *Ito diffusion* or *diffusion*.

According to 2.4.5 a diffusion process is smooth and possesses the strong Markov property. If functions μ and σ are independent from t the diffusion process is a time-homogeneous diffusion process (see Øksendal [103], page 104). Moreover, this definition can be generalized to a multi-dimensional diffusion process, see Schulmerich [117], page 21.

2.4.8 Remark: Relationship of the Two Definitions of a Diffusion Process

The two definitions of a diffusion process given in 2.3.1 (version 1) and 2.4.7 (version 2), respectively, are <u>not</u> equivalent. In particular, a diffusion process described in 2.4.7 is smooth and possesses the strong Markov property. This yields:

$(X_t)_{t\geq 0}$ is a diffusion according to version 2 $\quad\Longrightarrow\quad$ $(X_t)_{t\geq 0}$ is a diffusion according to version 1.

The opposite direction does generally not hold true as the *Feller-McKean process* demonstrates (see Rogers & Williams [113], page 271). This process is a diffusion process according to version 1 but not according to version 2. Therefore, there exist *more general* diffusion processes (version 1) than diffusion processes which are solutions of an SDE (version 2).

In the following, a process is called a diffusion process if it complies with the more general definition 2.3.1 (1st version of a diffusion process). This is common in literature.

2.4.9 Theorem: Solution of a One-Dimensional, Linear Stochastic Differential Equation

A one-dimensional, linear SDE is given via

$$dX_t \;=\; [A(t)X_t + a(t)]\,dt \;+\; [C(t)X_t + c(t)]\,dB_t,\;\; t \geq 0, \qquad (2.8)$$

where X_0 is a one-dimensional random variable with $E(X_0^2) < \infty$ that is independent from $\mathcal{F}_\infty$ (see 2.4.5). The functions $A, a, C,$ and c are assumed to have real values[35], and $(B_t)_{t\geq 0}$ is a one-dimensional Standard Brownian Motion. Functions $A, a, C,$ and c are further assumed to be smooth in t. Then the stochastic differential equation (2.8) has the solution[36]:

$$X_t = \Phi(t)\left[X_0 + \int_0^t \Phi^{-1}(u)\,[a(u) - C(u)c(u)]\,du \;+ \int_0^t \Phi^{-1}(u)\,c(u)\,dB_u\right],$$

where

$$\Phi(t) = \exp\left\{\int_0^t \left[A(u) - \frac{1}{2}C^2(u)\right]\,du \;+ \int_0^t C(u)\,dB_u\right\}, \quad t \geq 0.$$

A specific case is given for $C \equiv 0$ in (2.8). Under the assumption

$$E\left(\int_0^t \left[\Phi^{-1}(u)\sigma(u)\right]^2 du\right) < \infty \quad \forall\, t \geq 0$$

this yields:

$$(i) \quad E(X_t) \;=\; \Phi(t)\left[E(X_0) + \int_0^t \Phi^{-1}(u)a(u)\,du\right], \quad t \geq 0.$$

$$(ii) \;\; Cov(X_s, X_t) \;=\; \Phi(s)\left[Var(X_0) \;+ \int_0^{s\wedge t} \left(\Phi^{-1}(u)\sigma(u)\right)^2 du\right]\Phi(t), \quad s, t \geq 0.$$

For a proof of this theorem see Schulmerich [117], pages 26-28.

[35] The theorem holds also true for a multi-dimensional linear SDE, see Schulmerich [117], Sections 2.16, 2.19, and 2.20. In this case, A and C are real matrices and a and c are real vectors.

[36] Φ is a so-called *fundamental solution* in the one-dimensional case. For a multi-dimensional stochastic process Φ is a *fundamental matrix*. For more details see Schulmerich [117], pages 22-26. The proof of the solution is done with the Ito formula and is explained in detail in Göing [48], pages 60-62.

As mentioned earlier, stochastic calculus is also the language to describe the term structure models that will be explained in Chapter 3. In that context the term *mean reversion* will play a crucial role.

2.4.10 Definition: Mean Reversion Process

Let $(X_t)_{t\geq 0}$ be a diffusion process that is given as the solution of the SDE

$$\left.\begin{aligned} dX_t &= \mu(X_t)dt + \sigma(X_t)dB_t,\ \ t \geq 0,\ \ X_0 = x_0 \in I, \quad \text{where} \\ \mu(x) = \beta - \alpha\, x &= \alpha\left(\frac{\beta}{\alpha} - x\right),\ \ x \in I,\ \ \alpha \in \mathbb{R}^+,\ \ \beta \in \mathbb{R}, \\ \sigma(x) &> 0\ \ \forall\, x \in I. \end{aligned}\right\} \tag{2.9}$$

$I = (l, r)$ denotes the state space of the process (see Section 2.3). Such a process is called a *mean reverting* or *mean reversion process*; parameter α is called *mean reversion force* and $\frac{\beta}{\alpha}$ is called *mean reversion level* . These names are based on the notion that α is the "force" that pulls the process back to its "mean" $\frac{\beta}{\alpha}$. The model is called a *mean reversion model*.

According to this definition only a constant mean reversion level is allowed. However, this can be extended to a time-dependent mean reversion level or even a mean reversion level that contains a stochastic process, e.g., an Ornstein-Uhlenbeck process[37]. This process will be introduced in the following.

2.4.11 Example: Ornstein-Uhlenbeck Process

The Ornstein-Uhlenbeck process is a one-dimensional Ito diffusion that is given as the solution of the following linear SDE in the stronger sense:

$$dX_t = -\alpha X_t\, dt + \sigma\, dB_t,\ \ t \geq 0,\ \ X_0 = x_0 \in \mathbb{R} \text{ constant},\ \ \alpha \in \mathbb{R}^+,\ \ \sigma \in \mathbb{R}^+.$$

This process mean reverts around zero[38]. According to 2.4.9 it holds

$$\Phi(t) = e^{-\alpha t},\ \ t \geq 0,$$

and

$$X_t = x_0\, e^{-\alpha t} + \sigma\, e^{-\alpha t} \int_0^t e^{\alpha s}\, dB_s, \quad t \geq 0.$$

[37] An example is the Hull-White two-factor model which includes a drift term with a stochastic part which is modelled via an Ornstein-Uhlenbeck process.

[38] This property follows directly from 2.4.9 (i) for $t \to \infty$. See also Schulmerich [117], page 30.

2.5 Discretization of Continuous-Time Stochastic Processes

After having described thoroughly the underlying mathematics for real options and term structure modelling in the previous two sections, the next question to be considered is how to implement these continuous-time models in the discrete, numerical world of a computer.

Section 2.5 answers this question by providing three algorithms to simulate a discrete path of a continuous-time stochastic process. These algorithms have different degrees of "accuracy", a term that will be explained as well. Moreover, it will be shown how to generate normally distributed and correlated realizations of random variables as the basis for the three algorithms presented. The principles of this simulation can be found in Kloeden & Platen [77], Chapters 5.5, 9.6, 10.2, 10.3, and 10.4. A similar introduction to this topic can be found in Schulmerich [117], pages 143-146.

For a constant $T > 0$ consider a one-dimensional Ito process $(X_t)_{0 \le t \le T}$ that is given as the solution of the SDE

$$dX_t \;=\; \mu(t, X_t)\, dt \;+\; \sigma(t, X_t)\, dB_t, \;\; 0 < t \le T, \;\; X_0 = x_0 \in \mathbb{R}.$$

$(B_t)_{t \ge 0}$ is a one-dimensional Standard Brownian Motion. To simulate this continuous-time stochastic process $(X_t)_{0 \le t \le T}$ the *simulation time points*

$$0 \;=\; \tau_0 \;<\; \tau_1 \;<\; \ldots \;<\; \tau_N \;=\; T, \;\; N \in \mathbb{N},$$

of the *observation period* $[0, T]$ have to be specified. T is the *time horizon* of the observation period. Moreover, let $\Delta_s^i := \tau_{i+1} - \tau_i$, $0 \le i \le N-1$. Δ_s^i is the i-th *simulation step size*. In the computer simulation program for the book only equidistant simulation time points are used, i.e., $\tau_i = i\Delta_s, i = 0, \ldots, N$, with $\Delta_s := \frac{T}{N}$. Therefore, the following three simulation methods are specified for equidistant simulation time points. The number of simulated points is then $N = \frac{T}{\Delta_s}$, i.e., the total number of data points for one single simulated path is $N+1$ (including the non-random and given starting point x_0 of the process).

The basic tool for the following simulation methods is the Ito-Taylor expansion of a stochastic process $(X_t)_{0 \le t \le T}$ that can be found in Kloeden & Platen [77], Chapter 5.5. This expansion is the stochastic counterpart to the Taylor expansion. Depending on how many terms are used in the Ito-Taylor expansion for the simulation of the process $(X_t)_{0 \le t \le T}$, a different simulation method results. The Ito-Taylor expansion will not be explained here.

2.5.1 Mathematical Methods to Generate Random Variables

Let U_1 and U_2 be two on $(0, 1)$ uniformly distributed and independent random variables. Then

$$\left.\begin{aligned} N_1^{(0,1)} &:= \sqrt{-2\ln(U_1)}\ \cos(2\pi U_2), \\ N_2^{(0,1)} &:= \sqrt{-2\ln(U_1)}\ \sin(2\pi U_2) \end{aligned}\right\} \tag{2.10}$$

are two independent and standard-normally distributed random variables. The superscriptions $(0,1)$ indicate that these random variables are standard-normal (mean zero and variance one). (2.10) is called *Box-Muller-method.* With $\mu_1, \mu_2 \in \mathbb{R}$ and $\sigma_1, \sigma_2 \in \mathbb{R}^+$ the definitions

$$N_1^{(\mu_1,\sigma_1^2)} := \sigma_1 N_1^{(0,1)} + \mu_1 \quad \text{and} \quad N_2^{(\mu_2,\sigma_2^2)} := \sigma_2 N_2^{(0,1)} + \mu_2 \tag{2.11}$$

give two independent, normally distributed random variables with

$$N_1^{(\mu_1,\sigma_1^2)} \stackrel{d}{=} N(\mu_1, \sigma_1^2) \qquad \text{and} \qquad N_2^{(\mu_2,\sigma_2^2)} \stackrel{d}{=} N(\mu_2, \sigma_2^2).$$

To create realizations of these random variables within a computer simulation program, the random number generator of a programming language can be used. Since this random number generator usually produces only realizations with a uniform distribution on interval $(0,1)$ (like in C++), formulas (2.10) and (2.11) can be used to get the appropriate realization of a normally distributed random variable. This is needed for the Euler scheme in 2.5.2 and the Milstein scheme in 2.5.3.

In the Taylor 1.5 scheme that will be presented in 2.5.4 two correlated normally distributed random variables are needed with correlation coefficient $\frac{\sqrt{3}}{2}$. The definitions

$$\Delta B^{(1)} := N_1^{(0,1)}\sigma_1 \quad \text{and} \quad \Delta B^{(2)} := \frac{1}{2}\sigma_1^3\left(N_1^{(0,1)} + \frac{1}{\sqrt{3}}N_2^{(0,1)}\right) \tag{2.12}$$

give two correlated normally distributed random variables with

$$\left.\begin{aligned} &\Delta B^{(1)} \stackrel{d}{=} N(0,\sigma_1^2), \quad \Delta B^{(2)} \stackrel{d}{=} N(0, \tfrac{1}{3}\sigma_1^6), \quad \text{and} \\ &Corr(\Delta B^{(1)}, \Delta B^{(2)}) = \frac{\sqrt{3}}{2}. \end{aligned}\right\} \tag{2.13}$$

2.5.2 Euler Scheme

The sequence $(Y_i)_{i=0}^N \equiv (Y_{\tau_i})_{i=0}^N$ is defined iteratively via the first three terms of the Ito-Taylor expansion of the continuous-time process $(X_t)_{0\le t\le T}$:

$$\left.\begin{aligned} Y_0 &:= x_0, \\ Y_{i+1} &:= Y_i + \mu(\tau_i, Y_i)\,\Delta_s + \sigma(\tau_i, Y_i)\,\Delta B_i^{(1)}, \\ &\quad i = 0, \ldots, N-1, \end{aligned}\right\} \tag{2.14}$$

where $\Delta B_i^{(1)} \overset{d}{=} N(0, \Delta_s)$ for $i = 0, \ldots, N-1$. $(Y_i)_{i=0}^N$ is the *Euler approximation* of the continuous-time process $(X_t)_{0 \le t \le T}$ in the observation period $[0, T]$. (2.14) is called *Euler scheme*. In the computer simulation program a realization of $\Delta B_i^{(1)}$ is calculated via the method presented in (2.10) and (2.11) with $\mu = 0$ and $\sigma_1 = \sqrt{\Delta_s}$.

2.5.3 Milstein Scheme

As opposed to the Euler scheme, the Milstein scheme uses one more term of the Ito-Taylor formula. In the following, a partial derivative with respect to the local component (=second component) is denoted with a stroke next to the variable. The sequence $(Y_i)_{i=0}^N \equiv (Y_{\tau_i})_{i=0}^N$ can now iteratively be defined via:

$$\left.\begin{aligned} Y_0 &:= x_0, \\ Y_{i+1} &:= Y_i + \mu(\tau_i, Y_i)\,\Delta_s + \sigma(\tau_i, Y_i)\,\Delta B_i^{(1)} \\ &\quad + \frac{1}{2}\sigma(\tau_i, Y_i)\sigma'(\tau_i, Y_i)\left\{(\Delta B_i^{(1)})^2 - \Delta_s\right\}, \\ &\quad i = 0, \ldots, N-1, \end{aligned}\right\} \tag{2.15}$$

where $\Delta B_i^{(1)} \overset{d}{=} N(0, \Delta_s)$ for $i = 0, \ldots, N-1$. $(Y_i)_{i=0}^N$ is the *Milstein approximation* of the continuous-time process $(X_t)_{0 \le t \le T}$ in the observation period $[0, T]$. (2.15) is called *Milstein scheme*. In the computer simulation program a realization of $\Delta B_i^{(1)}$ is calculated via the method presented in (2.10) and (2.11) with $\mu = 0$ and $\sigma_1 = \sqrt{\Delta_s}$.

2.5.4 Taylor 1.5 Scheme

As opposed to the Milstein scheme, the Taylor 1.5 scheme uses several more terms of the Ito-Taylor formula. The sequence $(Y_i)_{i=0}^N \equiv (Y_{\tau_i})_{i=0}^N$ can iteratively be defined, see (2.16). Here, $\Delta B_i^{(1)} \overset{d}{=} N(0, \Delta_s)$ and $\Delta B_i^{(2)} \overset{d}{=} N(0, \frac{1}{3}\Delta^{1.5})$ for $i = 0, \ldots, N-1$. In the computer simulation program realizations of $\Delta B_i^{(1)}$ and $\Delta B_i^{(2)}$ are calculated via the method presented in (2.12), which according to (2.13) yields the appropriate distributions with $\mu = 0$ and $\sigma_1 = \sqrt{\Delta_s}$. $(Y_i)_{i=0}^N$ is the *Taylor 1.5 approximation* of the continuous-time process $(X_t)_{0 \le t \le T}$ in the observation period $[0, T]$. (2.16) is the *Taylor 1.5 scheme*. The denomination of the simulation method (2.16) in the book as *Taylor 1.5 scheme* stems from the following introduction of the term *strong convergence of order* κ with $\kappa \in \mathbb{R}^+$ which will be used for comparing the three simulation methods introduced above with each other.

$$\left.\begin{aligned}
Y_0 \quad &:= \; x_0, \\
Y_{i+1} \quad &:= \; Y_i \;+\; \mu(\tau_i, Y_i)\,\Delta_s \;+\; \sigma(\tau_i, Y_i)\,\Delta B_i^{(1)} \\
&\quad + \tfrac{1}{2}\sigma(\tau_i, Y_i)\sigma'(\tau_i, Y_i)\;\left\{(\Delta B_i^{(1)})^2 - \Delta_s\right\} \\
&\quad + \mu'(\tau_i, Y_i)\sigma(\tau_i, Y_i)\,\Delta B_i^{(2)} \\
&\quad + \tfrac{1}{2}\left(\mu(\tau_i, Y_i)\mu'(\tau_i, Y_i) \;+\; \tfrac{1}{2}\sigma^2(\tau_i, Y_i)\mu''(\tau_i, Y_i)\right)\,\Delta_s^2 \\
&\quad + \;\left(\mu(\tau_i, Y_i)\sigma'(\tau_i, Y_i) \;+\; \tfrac{1}{2}\sigma^2(\tau_i, Y_i)\sigma''(\tau_i, Y_i)\right) \\
&\qquad \left\{\Delta B_i^{(1)}\,\Delta_s \;-\; \Delta B_i^{(2)}\right\} \\
&\quad + \tfrac{1}{2}\sigma(\tau_i, Y_i)\;(\sigma(\tau_i, Y_i)\sigma''(\tau_i, Y_i) \\
&\quad + [\sigma'(\tau_i, Y_i)]^2)\;\left\{\tfrac{1}{3}[\Delta B_i^{(1)}]^2 - \Delta_s\right\}\;\Delta B_i^{(1)}, \\
&\quad i = 0, \ldots, N-1.
\end{aligned}\right\} \tag{2.16}$$

2.5.5 Strong Convergence of Order κ

Let $(X_t)_{t \geq 0}$ be an Ito process that is given as the solution of the SDE

$$dX_t \;=\; \mu(t, X_t)\,dt \;+\; \sigma(t, X_t)\,dB_t, \;\; t \geq 0, \;\; X_0 = x_0.$$

Let $0 = \tau_0 < \tau_1 < \ldots < \tau_N = T$ be the simulation time points (not necessarily equidistant any more) in the time interval $[0, T]$ with the maximal step size

$$\delta \quad := \quad \max\{\tau_{i+1} - \tau_i \,|\, i = 0, \ldots, N-1\}$$

and $(Y_i^\delta)_{i=0}^N$ be a discrete-time approximation of $(X_t)_{t \geq 0}$ on $[0, T]$ with respect to the simulation time points $\{\tau_0, \ldots, \tau_N\}$. The approximation $(Y_i^\delta)_{i=0}^N$ converges strongly with order $\kappa > 0$ towards $(X_t)_{t \geq 0}$ at time T, if there exists a constant $C > 0$ that is independent from δ, and if there exists a constant $\delta_0 > 0$ with

$$E\left(\left|X_T - Y_T^\delta\right|\right) \quad \leq \quad C\,\delta^\kappa \quad \forall\,\delta \in (0, \delta_0).$$

Under appropriate assumptions[39] it can be shown:

[39] See Kloeden & Platen [77], page 323, page 345, and page 351.

1. The Euler scheme is strongly convergent of order $\kappa = 0.5$.
2. The Milstein scheme is strongly convergent of order $\kappa = 1$.
3. The Taylor 1.5 scheme is strongly convergent of order $\kappa = 1.5$.

In this sense the Taylor 1.5 scheme is better suited to simulate an Ito process than the other two schemes. However, all three simulation methods are implemented for the simulation of the short-rate process in the computer simulation program whenever possible. In this book, the approximation scheme that is best suited to simulate the short-rate process is always applied in the test situations 2-5 of Chapter 5.

2.6 Evolution of the Real Options Theory and Models in the Literature

Real options have already been introduced in Section 2.2. An introduction to the mathematics of Financial Engineering was also provided in Sections 2.3, 2.4, and 2.5. The developed theory will now be used to give an overview of the evolution of the real options field in the literature but without focusing on the mathematical derivation of the models in this section. However, it is not the goal to give a plain literature review since this is already done is great detail in, e.g., Trigeorgis [133], pages 14-21. Rather, the goal is to provide a broad overview with classifications of the available articles, theories, and methods and, thereby, set the context for what is done analytically and numerically in Chapters 4 and 5 of this book.

The study of real options *arose in part as a response to the dissatisfaction of corporate practitioners, strategists, and some academics with traditional techniques of capital budgeting*[40]. This refers especially to the DCF method introduced in Section 2.2. Although some of the unsatisfactory valuation results from the DCF method stemmed from its being misapplied in practice, its inappropriateness to the valuation of investments with operating or strategic options was clear[41]. In particular, the DCF method becomes more inappropriate as the number of options in an investment project rises.

Even before the idea of real options was born, economists like Roberts and Weitzman found that in sequential decision-making, even in the case of a negative NPV, it may be worthwhile to undertake an investment[42]. Several authors stressed the limitations of the DCF method for evaluating investment decisions with strategic considerations[43], i.e., undervaluing those investment

[40] See Trigeorgis [133], pages 14-15.
[41] See Myers [100].
[42] See Roberts & Weitzman [112].
[43] See Dean [38], Hayes & Abernathy [49], and Hayes & Garvin [50].

projects. In the words of Trigeorgis[44]: *The basic inadequacy of the NPV approach and other DCF approaches to capital budgeting is that they ignore, or cannot properly capture, management's flexibility to adapt and revise later decisions (i.e., review its implicit operating strategy). The traditional NPV approach, in particular, makes implicit assumptions concerning an "expected scenario" of cash flows and presumes management's commitment to a certain "operating strategy".*

One of the results of this criticism was the development of the decision-tree analysis by Hertz and Magee[45] in 1964. Interestingly enough, the idea of the decision-tree analysis was published by these authors in the Harvard Business Review, which is more focused on application and is not a classical academic journal. In the following the basic idea of decision-tree analysis will be explained and elaborated to show its advantanges and disadvantages.

2.6.1 Decision-Tree Analysis

Decision-tree analysis (DTA) picks up at the point of uncertain future cash flows in describing the project within a tree structure that allows different paths during the life of the project. The decisions that need to be made within the tree are marked by quadratic nodes; those decisions do not need to be made at the time of the valuation but later when more information has arrived. If the project contains several realizations at one point during the project's life that cannot be influenced by management, this is represented by a circular node in the tree. Real probabilities q have to be assigned to all possible realizations (i.e., branches). The superscripts ^{H}, ^{M}, and ^{L} describe the economic condition (high, medium, and low market, respectively). By this method management can visualize the project's inherent options and price them into the project's NPV. Within the tree each alternative must be represented as a branch. Moreover, each final branch in the tree must be assigned a numerical value in case of its realization.

Example 1. A platform investment in R&D is a typical example where DTA can be used. It is graphically explained in Figure 2.3. The numbers on the right next to the final point (denoted with c_t) give the NPVs at the end of year 3 in the respective market stage of the expected discounted cash flows of the years 3 and 4 - 10. It is assumed that at the end of year 10, the last year of production within the project, the plant is closed down at no cost. The cash flows are assumed to occur at the end of the respective year, although the investments have to be made at the beginning of that year.

44 See Trigeorgis [133], page 121.
45 See Hertz [55] and Magee [85].

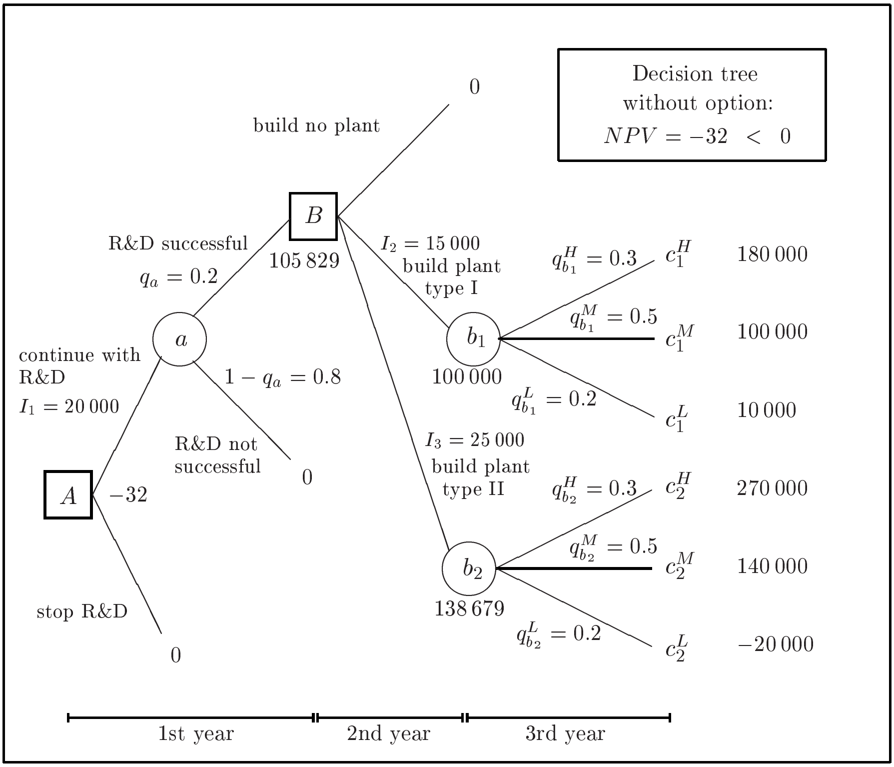

Fig. 2.3. Example of a platform investment in R&D within the framework of decision-tree analysis.

The calculations of the values in the tree are done by discounting with the cost of capital k, which is assumed to be 6%, and taking into consideration the investment cost and the management decision at point B:

$$\text{Point } b_1: \quad 100\,000 = \frac{0.3 \cdot 180\,000 + 0.5 \cdot 100\,000 + 0.2 \cdot 10\,000}{1 + 0.06}$$

$$\text{Point } b_2: \quad 138\,679 = \frac{0.3 \cdot 270\,000 + 0.5 \cdot 140\,000 + 0.2 \cdot (-20\,000)}{1 + 0.06}$$

$$\text{Point } B: \quad 105\,829 = max\left(0, \frac{100\,000}{1 + 0.06} - 15\,000, \frac{138\,679}{1 + 0.06} - 25\,000\right)$$

$$\text{Point } A: \quad -32 = \frac{0.2 \cdot 105\,829 + 0.8 \cdot 0}{1 + 0.06} - 20\,000$$

Applying DTA yields a negative NPV of −32, which means that the platform investment would not be undertaken: stop R&D.

Example 2. The picture will be different when a real option is included in the platform investment of example 1. In this case the management estimates the future cash flow in the years 4 - 10 at the end of year 3 and has the option to sell the plant at the end of year 3 for its salvage value X. This salvage value is assumed to be dependent on the state of the economy at the end of year 3 and on the size of the plant[46]:

$$\begin{array}{rclcrcl} X(c_1^H) & = & 12\,000 & \qquad & X(c_2^H) & = & 24\,000 \\ X(c_1^M) & = & 10\,000 & & X(c_2^M) & = & 20\,000 \\ X(c_1^L) & = & 8\,000 & & X(c_2^L) & = & 18\,000 \end{array}$$

Management expects the following NPVs (with respect to the end of year 3, thus index 3) for the cash flows of the years 4 - 10:

$$\begin{array}{rcr} NPV_3(\text{Cash flow of years 4 - 10}, c_1^H) & = & 145\,000 \\ NPV_3(\text{Cash flow of years 4 - 10}, c_1^M) & = & 70\,000 \\ NPV_3(\text{Cash flow of years 4 - 10}, c_1^L) & = & 7\,000 \\ NPV_3(\text{Cash flow of years 4 - 10}, c_2^H) & = & 200\,000 \\ NPV_3(\text{Cash flow of years 4 - 10}, c_2^M) & = & 90\,000 \\ NPV_3(\text{Cash flow of years 4 - 10}, c_2^L) & = & -14\,000 \end{array}$$

This also gives the following cash flows at the end of year 3 for year 3 as $E(c_i^j) = c_i^j - NPV_3(c_i^j)$:

$$\begin{array}{rclcr} E(\text{Cash flow in year } 3, c_1^H) & = & 180\,000 - 145\,000 & = & 35\,000 \\ E(\text{Cash flow in year } 3, c_1^M) & = & 100\,000 - 70\,000 & = & 30\,000 \\ E(\text{Cash flow in year } 3, c_1^L) & = & 10\,000 - 7\,000 & = & 3\,000 \\ E(\text{Cash flow in year } 3, c_2^H) & = & 270\,000 - 200\,000 & = & 70\,000 \\ E(\text{Cash flow in year } 3, c_2^M) & = & 140\,000 - 90\,000 & = & 50\,000 \\ E(\text{Cash flow in year } 3, c_2^L) & = & (-20\,000) - (-14\,000) & = & -6\,000 \end{array}$$

[46] Alternative methods are presented in Trigeorgis [133], pages 61-65.

At the end of year 3 management will only sell the plant if NPV_3 is smaller than the salvage value. For both plant types this is only the case for a low market when plant type I sells for 8 000 and plant type II sells for 18 000. Taking this salvage value into consideration at the end of year 3, gives the following situation (compared with the right side of Figure 2.3):

- Don't change 180 000 at c_1^H since $35\,000 + \max(145\,000,\ 12\,000) = 180\,000$
- Don't change 100 000 at c_1^M since $30\,000 + \max(70\,000,\ 10\,000) = 100\,000$
- Replace 10 000 at c_1^L with $3\,000 + \max(7\,000,\ 8\,000) = \underline{\mathbf{11\,000}}$
- Don't change 270 000 at c_2^H since $70\,000 + \max(200\,000,\ 24\,000) = 270\,000$
- Don't change 140 000 at c_2^M since $50\,000 + \max(90\,000,\ 20\,000) = 140\,000$
- Replace $-20\,000$ at c_2^L with $-6\,000 + \max(-14\,000,\ 18\,000) = \underline{\mathbf{12\,000}}$

The remaining calculations are the same as in example 1. They produce the decision tree of Figure 2.4 in the case of an option to abandon for a salvage value in this example of a platform investment in R&D. It finally yields an NPV of 1 043, so that the investment would be undertaken. This NPV is called a *strategic NPV* since it includes a real option. The NPV of −32 previously calculated in the absence of a real option is called *static NPV*. The value of the real option can therefore be calculated as:

$$\begin{aligned}
\text{Value of the option to abandon} &= \text{option premium}\\
&= \text{strategic NPV} - \text{static NPV}\\
&= 1\,043 - (-32)\\
&= 1\,075.
\end{aligned}$$

In summary, DTA is well suited to price some types of real options. It can price sequential investment decisions in which management decisions are made at discrete points in the future and uncertainty is resolved at discrete points in time as well. DTA is able to handle this embedded flexibility but the practical application has serious limitations: the number of discrete points in time can get large, thus, creating an extremely complex tree. Trigeorgis refers to this as *decision-bush analysis*[47] instead of decision-tree analysis. Realistic corporate budgeting situations cannot be properly handled that way. Second, the applied discount rate poses a problem since it is usually assumed to be constant. So, when uncertainty gets resolved at decision nodes, the DTA method does not use a changed discount factor. DTA cannot therefore reflect this change in the riskiness of the project expressed in the discount rate.

[47] See Trigeorgis [133], page 66.

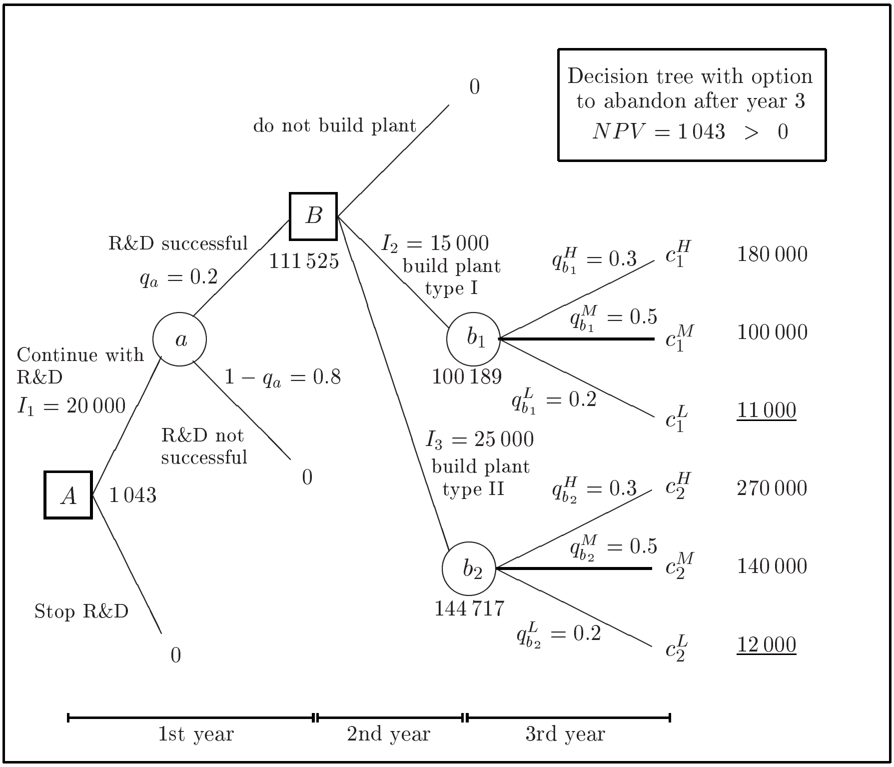

Fig. 2.4. Example of a platform investment in R&D with option to abandon within the framework of decision-tree analysis.

2.6.2 Contingent-Claims Analysis

A comparison of the DTA method and the DCF method shows that DTA improves on the DCF method by including real options. However, the DTA has two major drawbacks. On the one hand it applies a constant discount rate. On the other hand the decision tree gets very complicated even in simple cases. Both reasons limit its application in practice. Contingent-claims analysis (CCA) is a method that gets rid of the discount rate problem which the DTA method exhibits. CCA seeks to replicate the pay-off structure of the project and its real options via financial transactions in order to determine the NPV of the project. This section builds on the ideas of Trigeorgis[48] in illustrating the CCA method with an option to defer.

[48] See Trigeorgis [133], pages 153-161.

The characteristic of a real option is that it alters the risk structure of the project. In example 2 of the previous Section 2.6.1, the option to abandon after year 3 gets rid of negative cash flows in a low-case market. This should result in a new risk-adjusted interest rate as discount factor below 6%. Since, however, the discount factor stays the same, the true NPV is understated. Therefore, the true NPV is above the calculated 1 043.

Trigeorgis characterizes DTA and CCA in a very precise way[49]: *DTA can actually be seen as an advanced version of DCF or NPV - one that correctly computes unconditional expected cash flows by properly taking account of their conditional probabilities given each state of nature. As such, DTA is correct in principle and is particularly useful for analyzing complex sequential investment decisions. Its main shortcoming is the problem of determining the appropriate discount rate to be used in working back through the decision tree.* [...] *The fundamental problem with traditional approaches to capital budgeting lies in the valuation of investment opportunities whose claims are not symmetrical or proportional. The asymmetry resulting from operating flexibility options and other strategic aspects of various projects can nevertheless be properly analyzed by thinking of discretionary investment opportunities as options on real assets (or as real options) through the technique of contingent-claims analysis.*

The CCA method builds upon the DTA method and eliminates the weak point of "constant interest rate". With DTA the basic idea is to calculate with real probabilities and a constant, risk-adjusted interest rate as the discount rate. With CCA the basic idea is to transform the real probabilities into risk-adjusted probabilities such that the algorithm can use a constant, risk-free interest rate that is independent of the project's risk structure. To explain the idea of CCA some definitions are necessary:

V := total value of the project,

S := price of the twin security that is almost perfectly correlated with V,

E := equity value of the project for the shareholder,

k := return of the twin security,

r_f := risk-free interest rate,

p := risk-neutral probability for up-movements of V and S per period,

q := real probability for up-movements of V and S per period,

u := multiplicative factor for up-movements of V and S per period,

d := multiplicative factor for down-movements of V and S per period.

[49] See Trigeorgis [133], page 155.

As the starting point the current values for V and S are assumed to be $V = 100$ and $S = 20$. S is the price of the so-called *twin security* that exhibits the same risk profile as the project with value V and can be traded on the capital markets. Furthermore, the real probability q is assumed to be 0.5. This gives a development of S and V with $u = 1.8$ and $d = 0.6$ as shown in Figure 2.5.

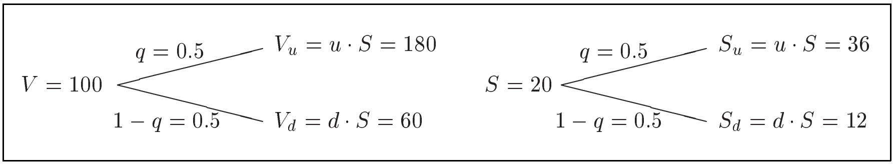

Fig. 2.5. Development of the total value V of a project and of the price S of a twin security.

The goal is to replicate the pay-off structure of the project after one year through the purchase of n shares of the twin security and through issuing of a 1-year Zero bond with nominal value B and risk-free return r_f. This is illustrated in Figure 2.6. The issue of the Zero is equivalent to getting a loan of value B for the interest rate r_f. Therefore, a positive B indicates a bond issue.

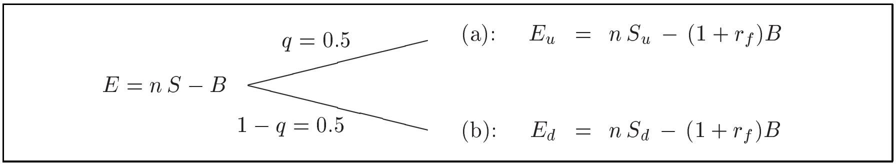

Fig. 2.6. Idea of the replicating portfolio.

$$\mathrm{n} = \frac{E_u - E_d}{S_u - S_d} \qquad \mathrm{E} = \frac{pE_u + (1-p)E_d}{1 + r_f}$$

$$\mathrm{B} = \frac{E_u S_d - E_d S_u}{(S_u - S_d)(1 + r_f)} \qquad \mathrm{p} = \frac{(1 + r_f) - d}{u - d}$$

Fig. 2.7. Solutions of the replicating portfolio.

By choosing the replicating portfolio approach $E = nS - B$ (i.e., E is replicated via n shares of stocks and the issue of a Zero worth nominal B) equations

(a) and (b) in Figure 2.6 have to be solved simultaneously to arrive at n and B. The solutions are summarized in Figure 2.7.

When dealing with a project that starts immediately, it is obvious that $V = E, V_u = E_u$, and $V_d = E_d$ have to hold since the value of the project has to be exactly the same as the value the shareholder gets out of the project. Putting the given data into the formulas, with a risk-free rate of $r_f = 8\%$, yields the following result:

$$\begin{aligned}
p &= \frac{(1+r_f)-d}{u-d} &&= \frac{(1+0.08)-0.6}{1.8-0.6} &&= 0.4\\
V &= \frac{pV_u+(1-p)V_d}{1+r_f} &&= \frac{0.4\cdot 180+(1-0.4)\cdot 60}{1+0.08} &&= 100\\
B &= \frac{V_uS_d-V_dS_u}{(S_u-S_d)(1+r_f)} &&= \frac{180\cdot 12-60\cdot 36}{(36-12)(1+0.08)} &&= 0\\
n &= \frac{V_u-V_d}{S_u-S_d} &&= \frac{180-60}{36-12} &&= 5
\end{aligned}$$

With an assumed investment cost of $I_0 = 104$, the NPV is $V - I_0 = 100 - 104 = -4 < 0$. Thus, in case of an immediate start, this project would not be undertaken.

In the above formulas, the risk-adjusted probability $p = 0.4$ is applied (and not the real probability $q = 0.5$) as well as a risk-free interest rate r_f. This is the crucial difference between CCA and DTA since the latter uses real probabilities and a risk-adjusted interest rate that cannot be easily determined (if at all). If the DTA is used to value the project the return k of the security is needed:

$$k = \frac{qS_u+(1-q)S_d}{S} - 1 = \frac{0.5\cdot 36+(1-0.5)\cdot 12}{20} - 1 = 20\%.$$

This gives:

$$V^{DTA} = \frac{qV_u+(1-q)V_d}{1+k} = \frac{0.5\cdot 180+(1-0.5)\cdot 60}{1+0.2} = 100.$$

In this case, therefore, the same value of V is yielded by DTA and CCA. This will change when the project contains a real option. Only the CCA with its risk-free interest rate gives the correct value for projects with real options, unlike the DTA with its risk-adjusted interest rate. This will be illustrated in the example of an option to defer by calculating the values for V and E if the project starts immediately or one year from today. For the NPV calculation the development of V, S, and E over two years with $S_{uu} = u^2S$, $S_{ud} = udS = S_{du}$, and $S_{dd} = d^2S$ needs to be determined. $S_{ud} = S_{du}$ leads to a recombining tree, i.e., the branches of the tree merge in the next stage and each additional stage creates one new node.

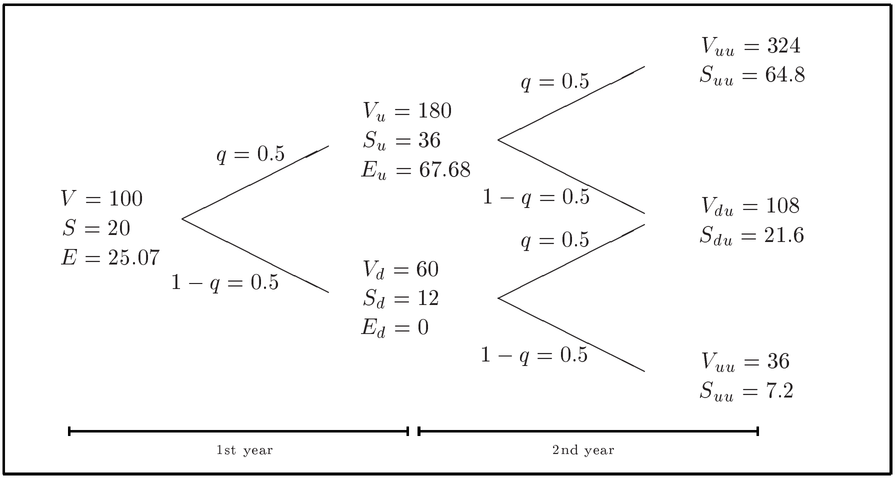

Fig. 2.8. Example of an option to defer the project start by one year within the framework of contingent-claims analysis.

An immediate start of the project has so far been assumed with an investment cost of $I_0 = 104$. If the start of the project is in one year from today, an investment cost of $I_1 = (1 + r_f)I_0 = 112.32$ can be assumed[50]. The valuation of the real option is explained in Figure 2.8. The shareholder values in Figure 2.8 can be calculated as follows:

$$\begin{aligned}
E_u &= max(V_u - I_1, 0) &&= max(180 - 112.32, 0) &&= 67.68 \\
E_d &= max(V_d - I_1, 0) &&= max(60 - 112.32, 0) &&= 0 \\
E &= \frac{pE_u + (1-p)E_d}{1 + r_f} &&= \frac{0.4 \cdot 67.68 + (1 - 0.4) \cdot 0}{1 + 0.08} &&= 25.07
\end{aligned}$$

Thus, in the case of a 1-year deferred project start, different values for V and E at $t = 0$ are calculated: $V = 100$ but $E = 25.07$. In this situation the NPV as the equity value to the shareholder is 25.07. The question that arises here is: What is an option worth that allows the start of the project to be deferred by one year? It has to be the difference of both NPVs, i.e., $25.07 - (-4) = 29.07$. This means that management would be willing to pay up to 29.07 to have the option to defer the project start by exactly one year.

Applying the DTA method yields a different value for E and, therefore, for the NPV:

$$NPV^{DTA} = E^{DTA} = \frac{qE_u + (1-q)E_d}{1+k} = \frac{0.5 \cdot 67.68 + 0.5 \cdot 0}{1 + 0.2} = 28.20\,.$$

50 See Trigeorgis [133], page 158.

So the question is which NPV is the correct one? This question can be easily answered by applying the theory of the replicating portfolio, see Figure 2.7.

$$n = \frac{E_u - E_d}{S_u - S_d} = \frac{67.68 - 0}{36 - 12} = 2.82$$

shares of stocks and the issue of a 1-year Zero with nominal value

$$B = \frac{E_u\, S_d \; - \; E_d\, S_u}{(S_u - S_d)\,(1 + r_f)} = \frac{67.68 \cdot 12 \; - \; 0 \cdot 36}{(36 - 12)\,(1 + 0.08)} = 31.33$$

are needed to replicate the equity value in the up- and down-cases for a risk-free interest rate of $r_f = 8\%$. The cost of a replicating portfolio is $2.82 \cdot 20 - 31.33 = 25.07$, so that $E = 25.07$ is the correct answer. DTA, therefore, overestimates the equity value. The reason for this is that the DTA approach uses a constant interest rate, which is not appropriate if real options are present.

In summary, asymmetrical cash flows change the risk structure of the project and, in so doing, the discount rate - something the DTA does not consider. The flaw of the DCF is nicely described in Trigeorgis [133], page 152: *The presence of flexibility embedded in future decision nodes, however, changes the payoff structure and the risk characteristics of an actively managed asset in a way that invalidates the use of a constant discount rate. Unfortunatelly, classic DTA is in no better position than DCF techniques to provide any recommendation concerning the appropriate discount rate.*

This error is corrected in the CCA method since CCA transforms the real probabilities into risk-adjusted probabilities and the risk-adjusted interest rates into risk-free interest rates to correctly mirror the asymmetry within a real option. In the words of Trigeorgis[51]: *Traditional DTA is on the right track, but although mathematically elegant it is economically flawed because of the discount rate problem. An options approach can remedy these problems.*

Therefore, only a CCA approach can capture the flexibility that is inherent in investment projects. *Flexibility*, according to Trigeorgis [133], page 151, *is nothing more than the collection of options associated with an investment opportunity, financial or real.* However, the general question of the justification of an option based approach with replicating portfolios and its limits has to be answered. A good justification for this approach is given in Trigeorgis [133], pages 124-129. While his full justification will not be restated here, a summary of Trigeorgis' most important points is needed[52]: *Can the standard technique of valuing options on the basis of a no-arbitrage equilibrium, using portfolios*

[51] See Trigeorgis [133], page 68.
[52] See Trigeorgis [133], page 127.

of traded securities to replicate the payoff to options, be justifiably applied to capital budgeting where projects may not be traded? As Mason and Merton[53] *in 1985 pointed out, the answer is affirmative if the same assumptions are adopted that are used by standard DCF approaches - including NPV - which attempt to determine what an asset or project would be worth if it were to be traded.* [...] *Given the prices of the project's twin security, management can, in principle, replicate the returns to a real option by purchasing a certain number of shares of its twin security while financing the purchase partly by borrowing at the riskless rate.* [...] *Risk-neutral valuation can be applied, whether the asset is traded or not, by replacing the actual growth rate, α, with a "certainty-equivalent" or risk-neutral growth rate, $\hat{\alpha}$, after substracting an appropriate risk premium ($\hat{\alpha} = \alpha -$ risk premium).*

To summarize the analogies of an investment opportunity (= real option on a project) and a call option on a stock, the comparison of financial and real options given by Trigeorgis[54] will be provided here in Table 2.1.

Table 2.1. Comparison between a real option on a project and a call option.

Call option on stock	Real option on project
current value on stock	(gross) present value of expected future cash flows
exercise price	investment cost
time to expiration	time until opportunity disappears
stock value uncertainty	project value uncertainty
risk-free interest rate	risk-free interest rate

Source: Trigeorgis [133], page 127

In the words of Trigeorgis[55]: *Option valuation can be seen operationally as a special, economically corrected version of decision-tree analysis that is better suited in valuing a variety of corporate operating and strategic options.*

However, the limitation of the analogy of a real option to a call option needs to be considered as well. In 1993, Kester analyzed the main differences in this analogy[56]. Besides the fact that real assets are non-traded and financial securities are traded (such that dividend-like adjustment are necessary[57]) the main differences are:

[53] See Mason & Merton [88].
[54] See Trigeorgis [133], page 125.
[55] See Trigeorgis [133], page 15. Compare also with Mason & Trigeorgis [89].
[56] See Kester [74].
[57] See Trigeorgis [133], page 127.

1. (Non) Exclusiveness of ownership and competitive interaction
2. Non-tradability and preemption
3. Across-time (strategic) interdependence and option compoundness

For a more detailed analysis on this subject see Trigeorgis [133], page 127.

2.6.3 Categorization of Real Options Valuation Methods

The valuation of investment projects with strategic and operating options started in 1977 by Stewart C. Myers. His pioneering idea was to perceive discretionary investment opportunities as *growth options*, giving corporate valuation a totally new direction, the *real options direction*. Since 1977 many articles have been published that deal with the valuation of real options. Two main avenues of real options valuation can be distinguished.

1. **Analytical methods:** approximative analytical solutions and closed-form solutions, including the Schwartz-Moon model[58], which will be introduced in Section 4.3.
2. **Numerical methods:** approximation of the partial differential equation that describes the option and approximation of the underlying stochastic process of the real option.

The following provides an overview of these two methods.

1. **Analytical methods.** Analytical methods can be divided into closed-form solutions and approximative analytical solutions. For the former, according to Trigeorgis [133], page 17, *a series of papers gave a boost to the real options literature by focusing on valuing quantitatively - in many cases deriving analytic, closed-form solutions.* Much work in that area was characterized by these analytical solutions that offer a nice (since closed-form) solution to simplified problems that seldom reflect reality. Trigeorgis offers a good overview of the academic articles taking this approach for various types of real options (see Trigeorgis [133], pages 2-3). In Chapter 6 of Trigeorgis [133] he also discusses in more detail some of the continuous-time models (including the martingale approach in Section 6.6), which are briefly summarized in the following:
   a) Option to defer (McDonald & Siegel [92], 1986; Paddock, Siegel & Smith [104], 1988; Majd & Pindyck [86], 1987): For the gross project value $(V_t)_{t\geq 0}$ McDonald and Siegel use a diffusion process given via the SDE

$$dV_t \;=\; \alpha\, V_t\, dt \;+\; \sigma\, V_t\, dB_t,\;\; t \geq 0, \alpha \in \mathbb{R}^+, \sigma \in \mathbb{R}^+.$$

[58] See Schwartz & Moon [122].

α is the instantaneous expected return on the project and σ its instantaneous standard deviation. Paddock, Siegel, and Smith use a similar process given via the SDE

$$dV_t \;=\; (\alpha - D)\, V_t\, dt \;+\; \sigma\, V_t\, dB_t,\;\; t \geq 0, \alpha \in \mathbb{R}^+, \sigma \in \mathbb{R}^+,$$

with D as the payout rate for the valuation of off-shore petroleum leases. Majd and Pindyck use the same process for the market value of the completed project in valuing the option to defer irreversible construction in a project where a series of outlays must be made sequentially but construction has an upper speed boundary.

b) Option to shut down or abandon (McDonald & Siegel [91], 1985; Myers & Majd [101], 1990): McDonald and Siegel assume that the diffusion process $(P_t)_{t \geq 0}$ for the output price is given by the SDE

$$dP_t \;=\; \alpha\, P_t\, dt \;+\; \sigma\, P_t\, dB_t,\;\; t \geq 0, \alpha \in \mathbb{R}^+, \sigma \in \mathbb{R}^+, \qquad (2.17)$$

i.e., the same process as in McDonald & Siegel [92]. Myers and Majd use a similar process, given by

$$dP_t \;=\; (\alpha - D)\, P_t\, dt \;+\; \sigma\, P_t\, dB_t,\;\; t \geq 0, \alpha \in \mathbb{R}^+, \sigma \in \mathbb{R}^+, \quad (2.18)$$

where D represents the instantaneous cash payout (calculated as cash flow divided by the project value).

c) Option to switch (Margrabe [87], 1978; Stulz [127], 1982): Margrabe analyzes the value of an option to exchange one non-dividend paying risky asset for another where the prices V and S of the risky assets are modelled via the same processes as in (2.17), however with different coefficients for the V process and the S process. Stulz analyzes European options on the minimum or maximum of two risky assets, assuming the same two processes as in (2.17) and (2.18).

d) (Simple) Compound options (Geske [44], 1979): Geske calculates the value of an option on a stock where the stock can be seen as a European call option on the value of the firm's assets. Trigeorgis stresses this application for real options as to value a compound growth option where earlier investments have to be undertaken as prerequisites[59]. The value of the underlying is assumed to follow the same process as in (2.17).

e) Compound options to switch (Carr [27], 1988; Carr [28], 1995): Carr uses the same processes to describe the value of two risky assets to analyze European compound (or sequential) exchange options as in (2.17), however with different coefficients for the two processes.

[59] See Trigeorgis [133], page 213.

In addition to these articles several books were published in the past presenting specific real options models. 17 articles on real options by various authors were compiled in Brennan & Trigeorgis [21], published in 2000. The models presented are mathematically very complex and specific. The analytical analysis is often accompanied by empirical valuation results for the situation analyzed. Important to notice is that all articles assume a constant risk-free rate apart from Miltersen [96]. Another (less quantitative) work on real options was published in 1995 by Trigeorgis[60]. This book also contains 17 separate articles on real options, all assuming a constant risk-free rate.

Two main disadvantages are inherent in all analytical solutions. First, the capital budgeting problem needs to be "nice" in order to be analytically trackable. This means that for a capital budgeting problem one must be in the position to write down the describing partial differential equation with the underlying stochastic process. In practice this is almost never the case. Second, even if a solution to the capital budgeting problem can be found for each single real option, this does not say anything about the value of the total option package. The valuation for complex options, i.e., option packages with many different real options and real option types[61], need to take option interactions into consideration, something almost always impossible to handle in an analytical approach.

2. **Numerical methods.** One criticism of analytical approaches is that they are not suited to valuing complex real options. The ability to value such complex options has been enhanced through various numerical techniques[62]. These can be divided into methods that approximate the partial differential equation and methods that approximate the underlying stochastic process[63]:

   a) **Approximation of the partial differential equation:** This includes numerical integration methods as well as the explicit and implicit finite difference methods.

   b) **Approximation of the underlying stochastic process:** This includes Monte Carlo simulation and lattice approaches like the Cox-Ross-Rubinstein binomial tree method and the Trigeorgis log-transformed binomial tree method.

   In the following the basic characteristics of both numerical approximation methods will be described:

[60] See Trigeorgis [132].
[61] See Section 2.2.
[62] See Trigeorgis [133], pages 20-21.
[63] See Trigeorgis [133], page 21.

a) **Approximation of the partial differential equation.** This type of approximation includes numerical integration methods as well as the explicit and implicit finite difference methods. The methods of finite differences can only be applied if the time development of the option value can be described via a partial differential equation. This means it is sufficient to build a partial differential equation; a closed-form solution does not need to exist. The finite difference approach discretizes this partial differential equation. The name of the specific difference method derives from the way the resulting grid is solved: the implicit finite difference method, the explicit finite difference method or hybrid methods like the Crank-Nicholson finite difference method. Moreover, the underlying can be log-transformed before discretizing the partial differential equation, which yields better mathematical properties for the finite difference method. The log-transformed explicit and implicit methods are both presented in Chapter 4 and numerically analyzed in Chapter 5. Due to the properties *consistency*, *stability*, and *efficiency*[64] this book will primarily focus on the implicit method.

   All finite difference methods can be used to value European options and American options, and they can handle several state variables (multi-dimensional grids). An important point is that the method gives option values for many different start values for the underlying, something lattice approaches do not provide. Moreover, Trigeorgis characterizes the finite difference methods as more mechanical, requiring less intuition than lattice approaches[65].

   Various authors have contributed to this type of approximation. Trigeorgis cites many of the following publications in his book entitled *Real Options*[66]:

   - **Numerical integration:**
     Parkinson [105], 1977.
   - **Finite difference schemes:**
     Brennan & Schwartz [16], 1977; Brennan & Schwartz [17], 1978; Brennan[15], 1979; Majd & Pindyck [86], 1987.

---

[64] *Consistency*, also called *accuracy*, refers to the idea that the discrete-time process used for calculation has the same mean and variance for every time-step size as the underlying continuous process. *Stability*, or *numerical stability*, means that the approximation error in the computations will be dampened rather than amplified. *Efficiency* refers to the number of operations, i.e., the computational time needed for a given approximation accuracy. For more information see Trigeorgis [133], page 308. These terms are addressed in Chapter 4 where the various methods are introduced in detail.

[65] See Trigeorgis [133], page 305.

[66] See Trigeorgis [133], page 306.

- **Analytic approximation:**
  Johnson [71], 1983; Geske & Johnson [45], 1984; MacMillan [84], 1986; Blomeyer [10], 1986; Barone-Adesi & Whaley [4], 1987; Ho, Stapleton & Subrahmanyam [58], 1997.

b) **Approximation of the underlying stochastic process.** When pricing financial options the Monte Carlo simulation plays an important role for the approximation of the underlying stochastic process. However, this type of simulation is mainly used for the pricing of European options since in the past it could not be used to price American options. Recently, however, several publications have addressed the issue of pricing American type options with Monte Carlo simulation. A good overview is given by Pojezny[67]. He classifies Monte Carlo simulation for American type options into four groups:

- Combination procedures
- Parametrisation of early excercise boundary
- Estimation of bounds
- Approximation of value function

These groups of Monte Carlo simulation are not a focus of this book. For more information on the subject see Pojezny [107], Section 3.4.3. However, the basic idea of Monte Carlo simulation (as used for European option valuation) is described.

The starting point of a Monte Carlo simulation is the stochastic differential equation that describes the underlying. The underlying $(S_t)_{t\geq 0}$ is often described via

$$dS_t \quad = \quad \alpha\, S_t\, dt \;+\; \sigma\, S_t\, dB_t, \quad t \geq 0, \tag{2.19}$$

where $\alpha \in \mathbb{R}^+$ is the infinitesimal or instantaneous return of the underlying, $\sigma \in \mathbb{R}^+$ is the infinitesimal or instantaneous standard deviation of the underlying and dB_t is a normal distributed random variable with variance dt. The parameters in this equation can be estimated from financial data and have already been introduced in the various methods presented in the analytical methods part. One aspect to consider is that α has to be chosen in such a way that the generated path represents the underlying in a risk-neutral world. If the underlying is an exchange traded stock, α has to be replaced by $k - \delta$ with k as the return of the stock and δ as its dividend yield[68].

[67] See Pojezny [107], Section 3.4.3.
[68] See Lieskovsky, Onkey, Schulmerich, Teng & Wee [81], pages 13-14, and Trigeorgis [133], page 310.

The stochastic differential equation (2.19) describes the paths of the underlying for $t \geq 0$. A discretization of the stochastic differential equation allows the simulation of such a path with a computer simulation program. These discretization methods were already described in detail in Section 2.5. The idea is to divide the time interval $[0,T], T > 0$, in $N \in \mathbb{N}$ subintervals with equal length $\Delta_s := \frac{T}{N}$. Δ_s is called *stepsize* of the simulation. The goal is to simulate a path value $S_i := S(\tau_i)$ for each of the time points $\tau_i := i\Delta_s,\ \ i = 0, 1, \dots, N$. To do this a start value S_0 of the process is needed. If this start value is given, the path values can be calculated iteratively, e.g., via

$$S_{i+1} \quad := \quad S_i \ + \alpha\, S_i\, \Delta_s \ + \ \sigma\, S_i\, \Delta B_i, \quad i = 0, 1, \dots, N,$$

where $\Delta B_i \overset{d}{=} N(0, \Delta_s)$. This is the so-called *Euler scheme* as presented in 2.5.2. Having simulated such a path, the further steps depend on the type of option that needs to be priced. In the following a European call option with strike price X will be valued. To do this the final value S_T of the path is needed and has to be compared with the strike price of the option. Let $P_j := max(S_T - X, 0)$ be the price of a European call at maturity date T where index j indicates that this is the price for the j^{th} simulated path. This simulation of the paths has to be done many times with each simulation independent of each other. Let A denote the number of simulations (e.g., $A := 10\,000$). Each j gives a different value P_j. The mean

$$B \quad := \quad \sum_{j=1}^{A} P_j \tag{2.20}$$

then gives the value of the option at time T, and $B\, e^{-r_f T}$ gives the current discounted value of the option with r_f as the current risk-free interest rate. An interesting application of path simulations and option pricing can be found in Lieskovsky, Onkey, Schulmerich, Teng & Wee [81] for pricing and hedging Asian options on copper.

As already mentioned, the latest publications also show how Monte Carlo simulation can be used to price American type options. If the underlying does not pay any dividends, the exercise of an American call option prior to expiration is never optimal[69]. However, if the underlying pays dividends, it can be optimal to exercise the call option prior to maturity. For an American put option it can be optimal to exercise early even in the absence of dividend payments[70].

[69] See Merton [94].

[70] See Schweser [124], pages 609-611.

Another way of approximating the underlying is done in lattices, e.g., by using binomial or trinomial trees that start with the current underlying value as the start value. This approach can also accommodate American options. One of the building blocks of the lattice approach was the work of Cox, Ross, and Rubinstein in 1979 for the pricing of financial options (see Cox, Ross & Rubinstein [34]). Here, the construction of a binomial tree was a breakthrough for option valuation in discrete time, achieved by building on the earlier work (1976) of Cox and Ross[71] who introduced the notion of a *replicating portfolio* to create a *synthetic option* from an equivalent portfolio comprising the underlying of the option and a bond. This risk-neutral valuation allows to value an option using risk-adjusted probabilities and risk-free interest rates. This idea is framed as a contingent-claims analysis, an option-based approach that enables management to quantify the additional value of a project's operating flexibility[72]. This approach was already introduced in Section 2.6.2 and will be elaborated in Chapter 4. The positive characteristics of lattice methods are best summarized by Trigeorgis[73]: *Lattice approaches are generally more intuitive, simpler, and more flexible in handling different stochastic processes, options payoffs, early exercise of other intermediate decisions, several underlying variables, etc.*

The main disadvantage of lattices is that they only give the option value for one single underlying start value. Therefore, the whole procedure has to be run many times with different start values. This is time consuming, especially if the goal is to calculate an option value distribution depending on the start value of the underlying. On the other hand, lattice methods can handle real options and real options packages, compounded options, dividend payments on the underlying and interaction among multiple options within an option package.

Again, various authors contributed to the development of this type of approximation. Trigeorgis cites the following publications[74]:

- **Monte Carlo simulation:**
  Boyle [13], 1977.
- **Lattice methods:**
  Cox, Ross & Rubinstein [36], 1979; Boyle [14], 1988; Hull & White [63], 1988; Trigeorgis [131], 1991.

[71] See Cox & Ross [36].
[72] See Trigeorgis [133], page 155.
[73] See Trigeorgis [133], page 306.
[74] See Trigeorgis [133], page 306.

A final word has to be addressed to the valuation of multiple options. Interactions among multiple options are the reason why option value additivity can usually not be assumed when valuing multiple options. Valuing each option of such an option package separately and then adding up the values does not take into consideration the fact that options usually interact. Consider, for example, an option to abandon and an option to switch. The value of the option to switch is zero if the option to abandon has been exercised. Therefore, the value of the option to switch depends on whether the project is still valid (i.e., option to abandon has not been exercised and is still active) or not. In such a case the values of each single option cannot simply be added together; rather, both options have to be valued together. This gives a multiple option value less than the sum of the single option values, since the sum is less the more interactions there are among the single options. Trigeorgis elaborates on the aspect that the incremental value of an additional option, in the presence of another option, is generally less than its value in isolation (in Chapter 7 of Trigeorgis [133]). An example of value additivity is an option package comprising an option to expand and an option to contract, each one year after the project was started. The underlying value, the present value of the project, can be split into exercise ranges for each of the options. These ranges do not overlap in the case of an option to expand or an option to abandon, which is obvious.

2.6.4 Flexibility Due to Interest Rate Uncertainty

So far the focus has been on real options that were valued using a constant risk-free interest rate. The existence of a non-deterministic future cash flow structure and a deterministic (since constant) future risk-free interest rate was sufficient to create a real option. This situation can also be turned around: Even the case of deterministic future cash flows but uncertain future interest rates creates a real option. As Trigeorgis points out[75], management has the flexibility to wait with the investment (option to defer) or to abandon (option to abandon) the project completely.

An option to defer created by a stochastic interest rate even in the case of a deterministic future cash flow structure was first analyzed by Ingersoll and Ross[76] in their publication *Waiting to Invest: Investment and Uncertainty* in 1992 with the simple case of a 1-year Zero bond that pays 1 \$ and was issued at time $t \geq 0$ with maturity $T = t+1$. This can be seen as a project that starts at time t for an investment cost of P and that guarantees a 1 \$ payment at time $t+1$. The authors show how, with stochastically modelled interest rates

[75] See Trigeorgis [133], page 197.
[76] See Ingersoll & Ross [69].

of a special type[77], the optimal time t to invest can be determined. These pioneering ideas will be introduced in detail in 4.4.1. In 2001, when the U.S. Federal Reserve Bank reduced the risk-free interest rate several times, this idea became especially important since a decreasing interest rate increases a project's present value and might lead to earlier initiation of a project.

Trigeorgis also mentions, if only briefly, the notion of uncertain interest rates creating real options in his book *Real Options*[78]. He addresses this idea in the simple case that Ingersoll & Ross dealt with in their publication mentioned above[79]: *In reality, of course, situations do not present themselves in such simple scenarios. For example, a parallel upward shift in the level of interest rates (favoring early investment) in the presence of rising interest rate uncertainty (favoring project delay) would result in an unclear mixed effect. Moreover, if the cash flows of a project are growing and/or the project's life expires at a specified time (e.g., upon expiration of a patent or upon competitive entry), then delaying the project would involve an opportunity cost analogous to a "dividend" effect, reducing the value of the option to wait and favoring earlier initiation. Furthermore, if the cash flows are uncertain and correlated with the interest rate (and the market return), the cost of capital will differ from the risk-free interest rate. A risk-neutral valuation approach will then be necessary, especially if other real options are also present. The risk-neutral probabilities will again be derived from market price and interest rate information, but now they may vary over time and across states (since they will depend on changing interest rates).*

Other authors who dealt with stochastic interest rate modelling in (real) options valuation are, e.g., Sandmann[80] (1993, European options only, see Section 4.4), Ho, Stapleton, and Subrahmanyam[81] (1997, valuation of American options with a stochastic interest rate by generalizing the Geske-Johnson model), Miltersen[82] (2000) as well as Alvarez and Koskela[83] (2002). Miltersen used a model developed by Miltersen and Schwartz[84] to value natural resource investment projects. He assumed that the multiple product decisions are independent, and hence can be valued as the sum of European options[85]. His work is mathematically very complex, but he also provides various examples.

[77] Ingersoll and Ross used the Cox-Ingersoll-Ross model to describe the term structure of interest rates whereby the parameters α and β are omitted, i.e., the short-rate process has a zero-expected change.
[78] See Trigeorgis [133].
[79] See Trigeorgis [133], page 199.
[80] See Sandmann [115].
[81] See Ho, Stapleton & Subrahmanyam [58].
[82] See Miltersen [96].
[83] See Alvarez & Koskela [1].
[84] See Miltersen & Schwartz [97].
[85] See Miltersen [96], page 196.

The recent article of Alvarez and Koskela builds[86] on Ingersoll & Ross [69] and analyzes irreversible investments under interest rate variability. According to Alvarez and Koskela[87], *the current extensive literature on irreversible investment decisions usually makes the assumption of constant interest rate.*

As Alvarez and Koskela further state[88]: *In these studies dealing with the impact of irreversibility in a variety of problems and different types of frameworks the constancy of the discount rates has usually been one of the most predominant assumptions. The basic motivation of this argument is that interest rates are typically more stable (and consequently, less significant) than revenue dynamics. [...] If the exercise of such investment opportunities takes a long time, the assumed constancy of the interest rate is questionable. This observation raises several questions: Does interest rate variability matter and, if so, in what way and how much?*

Consequently, their studies are motivated by the importance of interest rate variability[89]: *It is known from empirical research that interest rates fluctuate a lot over time and that in the long run these follow a more general mean reverting process (for an up-to-date theoretical and empirical survey in the field see e.g. Björk [6] and Cochrane [32]). Since variability may be deterministic and/or stochastic, we immediately observe that interest rate variability in general can be important from the point of view of exercising real investment opportunities.*

The article of Alvarez and Koskela is very thorough and focuses on the mathematical aspects but does not provide detailed real world examples or historical backtesting. In this book, the focus is on numerical real options pricing with interest rate uncertainty. Although articles about option pricing in general are plentiful in the literature, thorough numerical analyses using simulations and historical backtesting of complex real options situations with stochastic interest rates are rare at best. Therefore, the goal of this book is to fill this gap by providing in-depth insight into how a stochastically modelled interest rate influences the real options value in various real options cases common in practice and to derive rules for application in Corporate Finance.

---

86 See Alvarez & Koskela [1], page 2: *[...] we generalize the important findings by Ingersoll and Ross both by allowing for stochastic interest rate of a mean reverting type and by exploring the impact of combined interest rate and revenue variability on the value and the optimal exercise policy of irreversible real investment opportunities.*

87 See the abstract in Alvarez & Koskela [1].

88 See Alvarez & Koskela [1], page 1.

89 See Alvarez & Koskela [1], page 2.

2.7 Summary

Chapter 2 introduced the idea of real options: qualitatively, quantitatively, and with respect to the historical development of this research field. Various real options approaches have been categorized and explained in terms of their development over time and use in financial practice. While Section 2.2, *Basics of Real Options*, introduced the concept of real options in general, Sections 2.3, 2.4, and 2.5 introduced the necessary financial mathematics needed in the following chapters. Section 2.3, *Diffusion Processes in Classical Probability Theory*, presented the idea of the diffusion process that is necessary to model real options and the term structure of interest rates. Section 2.4, *Introduction to Ito Calculus*, introduced Ito calculus in its basic form, especially with a special focus on diffusion processes. Finally, Section 2.5, *Discretization of Continuous-Time Stochastic Processes*, provided the necessary algorithms to numerically calculate a realized discrete path of a continuous-time stochastic process in a computer simulation program. The *Evolution of the Real Options Theory and Models in the Literature* was the topic of Section 2.6, a section divided into four subsections: *Decision-Tree Analysis* in 2.6.1, *Contingent-Claims Analysis* in 2.6.2, *Categorization of Real Options Valuation Methods* in 2.6.3, and *Flexibility due to Interest Rate Uncertainty* in 2.6.4.

In the remaining chapters of this book the focus is on real options pricing under interest rate uncertainty, which was introduced in 2.6.4. The central thrust will be to thoroughly analyze complex real options situations with non-constant (especially, stochastically modelled) risk-free interest rates using numerical simulations and historical backtesting. The primary valuation tools will be methods used for valuing financial options but modified to value real options problems (i.e. the investment opportunities) since the similarities between these two are manifold. To quote Dixit and Pindyck[90]: *Opportunities are options - rights but no obligations to take some action in the future. Capital investments, then, are essentially about options.*

The stochastic interest rate movement in these models will be described via term structure models. Each term structure model is specified by a diffusion process that will be introduced in the following chapter. The real options valuation methods will be explained in Chapter 4 and numerically analyzed in Chapter 5.

[90] See Dixit & Pindyck [40], page 105.

http://www.springer.com/978-3-642-12661-1

Real Options Valuation
The Importance of Interest Rate Modelling in Theory and Practice
Schulmerich, M.
2010, XVIII, 389 p., Hardcover
ISBN: 978-3-642-12661-1