

# CHAPTER 4

# Asset Pricing

The next chapters cover the development of Monte-Carlo simulations. This chapter introduces the theoretical context. We cover asset pricing in general terms, introduce derivatives markets and the asset pricing theory, and establish the necessary formalism, abstractions, definitions, and notations for the rest of the publication. Derivatives markets and asset pricing are vast subjects, covered in a multitude of publications. Our short introduction is necessarily partial and dense. Derivatives markets are covered in detail in Hull's classic [\[26\]](#), updated in 2017. Asset pricing theory is covered in many textbooks, like [\[27\]](#).

## 4.1 FINANCIAL PRODUCTS

In order to design *generic* pricing libraries, we must abstract the notions of financial products and models and define them in a general manner. Starting with financial products, we need a formalism encompassing underlying assets (shares, currencies, zero-coupon bonds,...), linear transactions (forward contracts, forward rate agreements, interest rate swaps,...), European options (calls and puts, caps and floors, swaptions,...), path-dependent and early exerciseable options (collectively known as *exotics*), multi-underlying options, including derivatives written on assets of different classes (known as *hy-*

*brids*), as well as portfolios of transactions and regulatory amounts like xVA.

## Products and cash-flows

We therefore define a financial product as a collection of cash-flows paid over a given schedule. The payment schedule can be either discrete, in which case the product is a set  $\{CF_p, 1 \leq p \leq P\}$  of cash-flows paid on dates  $\{T_p\}$ , or continuous, in which case the product is a continuous set of cash-flows ( $dCF_u = cf(u)du$ ) paid over the time interval  $[0, T]$ ,  $T$  being the payment date of the last cash-flow, called the product's maturity.

Valuation is linear, so the value  $V_t$  of a product at a time  $t \leq T$  is the sum of the values of its subsequent cash-flows:

$$V_t = \sum_{T_p \geq t} V_t^p \text{ or } \int_t^T v_t^u du .$$

Cash-flows may be deterministic, linear in underlying asset prices, optional, or exotic. In its most general form, a cash-flow  $CF_T$  paid at time  $T$  is a  $\mathcal{F}_T$ -measurable random variable. This means that it may depend on the state of the market on and before its payment date  $T$ , and is fully determined at  $T$ . Formally,  $CF_T = h(S_t, 0 \leq t \leq T)$  is a functional  $h$  of the *path* (which we also call *scenario*) and denote:

$$\Omega = (S_t, 0 \leq t \leq T)$$

where  $S_t$  is the “state of the market” at time  $t$  (which we call *sample*): the values at time  $t$  of the market variables that affect the cash-flow.

A European call of strike  $K$  and maturity  $T$  on some underlying asset  $s$  pays one cash-flow:

$$CF_T = (s_T - K)^+$$

at maturity. A call  $K$  up and out on a barrier  $B$  also pays one cash-flow but it is contingent on not breaching the barrier before maturity:

$$CF_T = 1_{\{\max(s_t, 0 \leq t \leq T) < B\}} (s_T - K)^+$$

The floating leg of an interest rate swap (IRS) pays  $N$  Libor coupons over a discrete schedule; for instance, a  $10y$  quarterly floating leg pays 40 cash-flows:

$$CF_i = L(T_i^{start}, T_i^{end}) (T_i^{end} - T_i^{start})$$

where  $L(T_1, T_2)$  is the Libor rate of maturity  $T_2$  fixed at  $T_1$ .<sup>1</sup>

## Scenarios

A cash-flow  $CF_T$  paid at time  $T$  depends on the scenario, generally defined as the evolution of the market up to the maturity, but its specific meaning depends on the cash-flow: for a European call, the scenario is the underlying asset price at maturity. For a barrier option, it is the continuous path of the underlying price series from now to maturity. For an interest rate swap (IRS), the scenario is the discrete set of Libor fixings on the floating schedule. Cash-flows don't depend on the complete state of the world market at all times in  $[0, T]$ . Depending on the particular cash-flow, the scenario is *specified* by two characteristics:

1. The subset of  $TL \subset [0, T]$  of dates where the state of the market affects the cash-flow, called *timeline*. For a European call, the timeline is the maturity date:  $TL = \{T\}$ . For a barrier, it is the entire interval  $[0, T]$ .

A product's timeline is the union of its cash-flows timelines. The dates  $t \in TL$  in the timeline are called *event dates*; they are the contractual dates where something meaningful happens to determine the cash-flows, in reference to the state of the market on those dates: the exercise of a call, the monitoring of a barrier, the fixing of a Libor, or the payment of a cash-flow. For the floating leg of an IRS, the timeline is the discrete set of fixing and payment dates on the floating schedule.

2. The nature of the sample  $s_t$  for every event date  $t \in TL$  on the timeline. For a European call or a barrier option, this is the under-

lying asset price on time  $t$ . For a floating leg cash-flow, it is the Libor of a given maturity fixed at  $t$ . In general, a sample  $s_t$  for the event date  $t$  is a collection of market variables (asset prices, rates,...) observed at time  $t$ . It is the snapshot of the fraction of the market at time  $t$  that is useful for the determination of the cash-flows. In principle, its dimension could be infinite or even continuous. The dimension and contents of different samples on the timeline may be different from one another.

The time  $t$  sample for a product is the union of the time  $t$  samples of its cash-flows. The collection of all the samples  $\{S_t, t \in TL\}$  across the timeline is called *scenario*, and denoted  $\Omega$ .

The scenario is therefore a multidimensional, and sometimes continuous, collection of samples over a timeline. The timeline, and the nature and the dimension of its samples, depend on the product. The cash-flows are a function(al) of the scenario, more precisely, of the scenario before the payment date:

$$CF_T = b(\{S_t, t \in TL, t \leq T\})$$

## Underlying assets

The notions of products, cash-flows, timelines, samples, and scenarios are the building blocks of the product side of a generic financial library (the other sides being models, discussed next, and algorithms that articulate the two to produce values and risk, like the Monte-Carlo simulations of the next chapter). We illustrate these notions with a few examples, starting with the underlying assets. Underlying assets, like stocks or currencies, are not exactly financial products, but they can be modeled as financial products with the introduction of a holding horizon  $T$ . A stock is not a financial product, but the strategy of buying the stock and holding it to some horizon  $T$  is one.

A stock without dividends, with value  $s_t$  at time  $t$ , can be assimilated, given a holding horizon  $T$ , to a financial product with a single cash-

flow  $CF_T = s_T$  on the horizon date  $T$ . The timeline is therefore the singleton  $\{T\}$  and the scenario is  $\Omega = \{S_T = s_T\}$ .

With a discrete schedule of dividends  $\{d_k\}$  paid on times  $\{T_k\}$ , a stock is modeled as a schedule of cash-flows  $CF_p = \{d_p, T_p \leq T\}$ , plus a final cash-flow  $CF_T = s_T$ . The timeline is the set of ex-dividend dates before horizon  $\{T_k < T\}$ , plus the horizon date  $T$ , and the scenario is  $\Omega = \{S_{T_k} = d_k\} \cup \{S_T = s_T\}$ .

A stock or stock index with a continuous dividend yield  $(y_t)$  is modeled as a continuous schedule of cash-flows ( $cf(t) = s_t y_t, t < T$ ) and a final cash-flow  $CF_T = s_T$ . The timeline is the interval  $[0, T]$ ; the scenario is  $\{(s_t, y_t), t \in [0, T]\}$ .

A foreign currency, which is economically identical to a stock with dividend yield the foreign short rate, is modeled as a financial product in the same way.

In interest rate markets, the notion of “underlying assets” is typically associated with zero-coupon bonds, although they are not directly traded instruments. The zero-coupon bond of maturity  $T$  is a financial product that pays a unique cash-flow of one monetary unit at time  $T$ . Its timeline is  $\{T\}$ . Its cash-flow is deterministic and independent of a scenario.

The value of a zero-coupon bond of maturity  $T$  at a time  $t < T$ , denoted  $DF(t, T)$ , acts as a conversion rate between payments at  $t$  and  $T$ , the value at  $t$  of a cash-flow  $X_t$  known at  $t$  ( $\mathcal{F}_t$ -measurable) and paid at  $T$  being  $DF(t, T)X_t$ . For this reason, zero-coupon bond prices are also called *discount factors*. Evidently,  $DF(T, T) = 1$ .

The collection  $DF(t, T)$  of discount factors of all maturities  $T$  at time  $t$  is called the *discount curve* at  $t$ . The discount curve is part of the market state at  $t$ . Discount factors are market primitives and may be part of samples in a scenario. What discount maturities  $T$  are in-

cluded in what samples depends on the product and its cash-flows. In general, a cash-flow doesn't depend on the whole collection of discount factors  $DF(t, T)$  on an event date  $t$ , but on a smaller number of discounts of a specific set of maturities  $T \in \Theta_t^{DF}$ . The set  $\Theta_t^{DF}$  of discount maturities in the sample  $S_t$  is part of the product-specific definition of the scenario.

Finally, the discount curve can be expressed in price units, discount factors, or in rate units,

$$R(t, T) \equiv -\frac{\log[DF(t, T)]}{T - t} \iff DF(t, T) = \exp[-R(t, T)(T - t)]$$

is called *discount rate* of maturity  $T$ . The quantity:

$$f(t, T) \equiv \frac{\partial}{\partial T}[R(t, T)(T - t)]$$

is called *instantaneous forward rate* (IFR).<sup>2</sup> It follows immediately that:

$$f(t, T) = -\frac{\partial \log[DF(t, T)]}{\partial T} \iff DF(t, T) = \exp \left[ - \int_t^T f(t, u) du \right]$$

and the short rate at time  $t$  is defined as  $r_t \equiv f(t, t)$ .

## Linear transactions

A *forward contract* with strike  $K$  and maturity  $T$  on an underlying asset  $s$  is a financial product paying a unique cash-flow  $s_T - K$  at date  $T$ . Its timeline is the singleton  $\{T\}$  and the time  $T$  sample is the scalar  $\{s_T\}$ .

On a date  $t \leq T$ , the unique strike  $K$  such that the value of the corresponding forward contract is 0 is called the *forward price* of  $s$  and denoted  $F(t, T)$ . Forward prices  $F(t, T)$  are part of the market state at time  $t$ , and it is best, in order to accommodate a wide range of products and models, to consider forwards as market primitives and parts of samples in the scenario. In this case, we don't need spot prices in the samples, since  $s_t \equiv F(t, t)$ .

What forwards of what maturities are part of the samples depends on the exact nature of the cash-flows, same as discount factors. For instance, a forward contract depends on the sample

$$S_T = s_T = F(T, T) \text{ so in this case, } \Theta_T^F = \{T\}.$$

In interest rate markets, a forward rate agreement (FRA) pays  $[L(T_1, T_2) - K](T_2 - T_1)$  on date  $T_2$ ,<sup>3</sup> where  $L(T_1, T_2)$  is the *Libor* rate fixed at  $T_1$  for maturity  $T_2$ . The FRA's timeline is the union of its fixing and payment dates  $\{T_1, T_2\}$  and its dependency is on the sample  $S_{T_1} = L(T_1, T_2)$  on the fixing date  $T_1$ . Why  $T_2$  is on the timeline without a sample will be clarified when we move on to pricing, and realize that we are missing “something” on the payment dates samples.

On a date  $t \leq T$ , the unique strike  $K$  such that the value of the corresponding FRA is 0 is called forward rate, or *forward Libor*, and denoted  $F(t, T_1, T_2)$ . We consider forward Libors as market primitives and part of samples in the scenario. Spot Libors are therefore not necessary, since  $L(T_1, T_2) \equiv F(T_1, T_1, T_2)$ .<sup>4</sup>

What forward Libors of what start and end dates and what index are included in the samples is something specified from the nature of the cash-flows. For instance, a FRA of maturity  $T$  on a  $3m$  Libor (“ $3m$ ” referring both to the Libor’s duration and to its index) pays

$CF_{T+3m} = 0.25[F(T, T, T + 3m, 3m) - K]$  on time  $T + 3m$ . Its timeline is  $\{T, T + 3m\}$ , and it is determined by the time  $T$  sample  $S_T = F(T, T, T + 3m, 3m)$ . It follows that  $\Theta_T^L = \{(T, T + 3m, 3m)\}$ , one Libor being defined, in addition to its fixing date, by two dates and an index.

Since the cash-flow is known at time  $T$  and paid  $3m$  later, we can equivalently model this FRA as the cash-flow:

$$CF_T = 0.25[F(T, T, T + 3m, 3m) - K]DF(T, T + 3m)$$

paid at time  $T$ . We used the discount factor to *convert* an amount paid in  $3m$  into an amount paid immediately. This allows to reduce the timeline to the singleton  $\{T\}$  but increases the sample to include the discount factor:  $S_T = \{F(T, T, T + 3m, 3m), DF(T, T + 3m)\}$ . In this case,  $\Theta_T^{DF} = \{T + 3m\}$  and  $\Theta_T^L = \{(T, T + 3m, 3m)\}$ .

An interest rate swap (IRS) is a financial product that exchanges a fixed leg for a floating leg. The payment schedules of both legs share a common start date  $T_s$  and end date  $T_e$  but they typically have different periods  $\delta_{fix}$  and  $\delta_{float}$ , hence a different number of *coupons*  $N_{fix}$  and  $N_{float}$ . In a receiver swap, the fixed leg pays the coupons  $K\delta_{fix}$  on dates  $T_0 + i\delta_{fix}$  and the floating leg pays the coupons  $-L(T_0 + (j - 1)\delta_{float}, T_0 + j\delta_{float})\delta_{float}$  on dates  $T_0 + j\delta_{float}$ . A payer swap pays the opposite coupons. The timeline of an IRS is the union of the fixed and floating schedules. The scenario that determines an IRS is the sequence of Libor fixings over the floating schedule.

The fixed coupon  $K$  such that the value of the swap is 0 on a date  $t \leq T_s$  is called “par swap rate,” or sometimes “forward swap rate.” Par swap rates are not market primitives; we will see shortly that they are expressed in a simple, universal, and model-independent manner from forward Libors and discount factors.

Forward contracts, FRAs, and IRS are all examples of a particular type of financial products called *linear* products. Linear products are defined as schedules of linear cash-flows. Linear cash-flows are often defined as linear functions of the scenario, although this definition is approximate and in some cases incorrect. The definition of linear cash-flows is related to valuation, as we will see shortly.

## European options

Traditionally, European options are those that depend on a single observation of an underlying market variable on a single event date. We go with this definition for now, although it is somewhat blurry in our context. We will correctly define European products later, when we move on to valuation.

For now, a cash-flow  $CF$  is called European if it depends on a single observation, on a single event date, of a single market variable:

$CF = h(s_T)$  where  $s_T$  is a scalar market variable. The timeline of a European cash-flow is a singleton  $\{T\}$ , called “maturity,” “expiry,” or “exercise date” depending on the product. The cash-flow is a function (not a functional)  $h$  of the scalar  $s_T$ .

When  $h(x) = \max(0, x - K)$ , the cash-flow is a *European call*. When  $h(x) = (K - x)^+$ , it is a European put. More generally, any smooth function  $h$  can be written as:

This is known as Carr-Madan's formula [28] and shows that *all* European cash-flows are combinations of calls and puts. Carr-Madan's formula is demonstrated with simple calculus and remains valid when  $h$  is discontinuous, as long as derivatives are read in the distributional sense of Laurent Schwartz.

A European call (respectively put) on a Libor is called a caplet (floorlet). A succession of caplets (floorlets) over a floating schedule is called a cap (floor).

A cash-settled swaption pays a single cash-flow on its exercise date  $T$ :

for a payer swaption, or replace the right-hand side with  $\max(0, \dots)$  for a receiver swaption, where  $\dots$  is the par swap rate on the exercise date. A swaption's cash-flow is a function of the par swap rate, which is not a market primitive, but an aggregate of a number of discount factors and forward Libors as we will see shortly. A swaption may therefore appear not *stricto sensu* European.<sup>5</sup> We will revisit swaptions shortly.

A financial product is called European if all its cash-flows are European (European products may also include multiple European cash-flows on the same date, on different underlying variables). All the linear products we introduced previously are European products. Nonlinear European products are called European options.

## Path-dependent exotics

Cash-flows that are not linear or European depend on multidimensional samples on multiple event dates, in a nonlinear manner.

Products that involve such cash-flows are collectively known as *exotics*. Cash-flows that depend nonlinearly on multidimensional samples are called “basket options.” Those depending on multiple samples prior to their payment dates are called “path-dependent.” When a party holds the contractual right to cancel, redefine, or alter cash-flows in any way, the product is said to be “callable” or “early exercisable.” Those categories are not mutually exclusive.

The timeline of path-dependent cash-flows includes, by definition, multiple event dates. A textbook example is a barrier option<sup>6</sup> on some underlying asset  $\dots$ . A barrier delivers the cash-flow of a European call or put at maturity, but only when the underlying asset price remained below (up and out) or above (down and out) a predefined level called *barrier* and denoted  $\dots$ , at all times before maturity. The cash-flow of an up-and-out call, for example, is:

where . Its timeline is therefore the continuous interval and it is a functional of the continuous scenario . Many other variations exist, like knock-in barriers, which pay only if the barrier was breached, *American digitals*, which pay a different amount depending on whether the barrier was breached or not, double barriers (those with a down barrier and an up barrier, sometimes called *corridors*), and all the imaginable combinations of these.

One important variation for Monte-Carlo simulations is the *discretely monitored* barrier. In this case, the payment is subject to not breaching (knock-out) or breaching (knock-in) the barrier over a *discrete* timeline, irrespective of what happens in between the observation dates. The timeline is the discrete union of the monitoring schedule with the fixing and payment dates of the cash-flow. The scenario is the spot price on the event dates in the timeline.

## American options and cancellable products

Moving on to callable products, we introduce the two most common types of early exerciseable transactions, American options and cancellable products.

American options are those exerciseable at any time *before maturity* into a cash-flow for an American call or for an American put. The party who holds the right to exercise does so on the first occasion when the exercise cash-flow exceeds the *continuation value* of the option (the value immediately after if not exercised). An American option's payment schedule and its timeline are therefore both the continuous interval and its cash-flows are defined by:

where  $\delta$  tracks whether the option is still alive and  $\alpha$  defines the exercise cash-flow. We have a product which cash-flow at  $t$  depends on the product's value on and before  $t$ . We have seen that the product's value at time  $t$  is the sum of the values at  $t'$  of all the cash-flows paid after  $t$ . Hence, values depend on cash-flows and cash-flows depend on values. In addition, the continuation value  $\beta$  depends not only on the state of the market at  $t$ , but also on the option pricing model.

American options therefore involve an undesirable coupling between the product and the model and a recursive relationship between values and cash-flows.

Cancellable products are those where a party has the right to cancel all the subsequent cash-flows on a set of exercise dates. A cancellable product's timeline is the union of the timeline of its non-callable counterpart with the exercise dates. When the non-callable product pays the cash-flows  $\{C_t\}$  on times  $\{t\}$ , the callable product's cash-flows are defined with:

where  $\delta$  is 1 before cancellation and 0 thereafter, hence  $\delta_t = 1$  and  $\delta_t = 0$  and is updated on exercise dates  $t$  with:

where  $\beta_t$  is the time  $t$  value of all subsequent *cancellable* cash-flows. Again, cash-flows depend on future values in a recursive manner, and the definition of the cash-flows depends on the model. This is a defining characteristic of callable products; it is what makes them special and the reason why they require special modeling.

In the somewhat simplified case of a cancelable product with *one* exercise date, sometimes called “one-time callable,” the cash-flows are the cash-flows of the non-callable counterpart before the exercise date , and:

after the exercise date, so the cash-flows are still dependent on a future value, but this is the future value of the non-callable counterpart, not the value of the callable product itself. The recursive relationship is gone, but the dependency on a future value and a model remains.

For instance, a physically settled swaption is the right to enter the underlying swap, or, equivalently, the right to cancel its cash-flows. It is therefore a one-time callable swap, which pays the cash-flows of the underlying swap, provided its present value on the exercise date is positive:

We will see shortly that  $\frac{C}{P}$  is an explicit function of the forward Libors  $L_1, L_2, \dots, L_n$  on the exercise date with maturities the start and end dates on the floating schedule, and the discount factors  $d_1, d_2, \dots, d_n$  to the payment dates of both legs. It follows that:

All these discounts and forward Libors are therefore included in the swaption's sample on the exercise date:  $(d_1, d_2, \dots, d_n, L_1, L_2, \dots, L_n)$ ,

, and:

so:

Seen from this angle, the swaption is no longer a callable product, but fits the definition of path-dependent products. We will see shortly that

we can also consider it a European option that delivers the present value of the underlying swap on the exercise date.

## Turning callables into path-dependents

Callable products are therefore those whose cash-flows depend on the future values of themselves, and/or the future values of financial products other than primary market variables. How algorithms deal with them depends on the algorithm. We will see shortly that finite difference methods (FDM) cope with callable features in a native and natural manner. On the contrary, Monte-Carlo naturally handles path dependency, but doesn't know future transaction values and cannot handle callable features. In order to price a callable transaction with Monte-Carlo, we must first transform the transaction into an equivalent path-dependent, like we just did for a swaption.

In general, unlike for the swaption, we don't have an analytic expression of the future values, so we must somehow estimate them as a function of the market:

where  $\pi$  is called a *proxy* function and the sample  $\Omega$  is the set of market variables on which it accepts as arguments, which must be included in the scenario in order to use the proxy.<sup>7</sup> Using proxies, the cash-flows of the American option simplify into:

and the cash-flows of an early cancellable transaction simplify into:

In both cases, the explicit dependency on future values vanished and the modified cash-flows fit the definition of classic path-dependents. The injection of proxies effectively turns callables into path-dependents.

It is also noticeable that proxies only appear in exercise indicators. On an exercise date, the set of sample values that cause early exercise is called the exercise region. The border of the exercise region is called the exercise boundary. The problem of finding effective proxies boils down to the estimation of exercise boundaries.

In the context of FDM, exercise boundaries are directly estimated on the FDM grid, as we will see shortly. In the context of Monte-Carlo, proxies are estimated by regression over pre-simulations, with an algorithm known as the *least squares method* or LSM, pioneered by Carriere in [13] and Longstaff and Schwartz in [14]. Proxies obtained with LSM are called *regression proxies*. The LSM algorithm is the established best practice for pricing callables with Monte-Carlo since the early 2000s. We briefly introduce it in the next chapter.

## CVA

Finally, it is important to understand that counterparty value adjustment (CVA), along with most other value adjustments (xVA) and other bank-wise regulatory calculations, can be modeled as a financial product. This is crucial, because it follows that the models we develop for financial products are also suitable for these calculations. CVA is a real option a bank gives away whenever it trades with a defaultable counterparty. If the counterparty defaults on a future date when the net sum of all ongoing transactions with that counterparty (called *netting set*) is positive, the bank loses the positive value of the netting set.

This is a one-way street, hence, really an option: if the liquidator will insist that the bank refunds its value. CVA is therefore a zero-strike put, contingent to default, on the entire netting set, consisting itself in thousands of transactions, each one of which is a financial product in its own right, perhaps exotic or hybrid. The difficulty of computing CVA in a practical manner, and in reasonable time, is directly related to the high dimension of the “underlying” netting set.

In its simplest form, an uncollateralized CVA, defined in a discrete manner and without recursion, is modeled as a financial product that pays, over an *exposure schedule*, where and , the payment date of the last cash-flow in the netting set, the cash-flows:

where is the proportion defaulted between and  
<sup>8</sup>  
-

The cash-flows of a CVA therefore depend on the future values of the netting set, so everything we said about callable products applies to CVA, too.<sup>9</sup> In addition, CVA cannot be valued with FDM due to its high dimension, so proxies are necessary there, too.<sup>10</sup>

In addition, CVA depends on credit, so the scenarios must be enlarged to incorporate credit variables and credit events, something we are not doing in this publication. The future values of the netting set typically depend on a vast number of market variables from different asset classes: stocks, rates, currencies, etc. CVA is therefore a *hybrid* option, which valuation necessitates a hybrid model (essentially, an assembly of multiple models driving the different underlying markets, brought together and made consistent and arbitrage-free; see [7]).

Due to the dimension of the model and the size of the market samples, Monte-Carlo is the only practical means of pricing CVA. To generate high-dimensional market samples in reasonable time may be challeng-

ing; see [7]. The netting set's timeline, which is the union of the timelines of all the transactions, is typically dense, further slowing down evaluation, unless compression techniques, introduced in [30] and [11], are implemented. A reasonably fast generation of scenarios is not enough. To evaluate the large number of cash-flows in the netting set in every scenario is another bottleneck. Our publication dedicated to scripting [11] covers general, efficient means of representing and manipulating cash-flows, which helps with the aggregation and compression of the cash-flows, as well as a fast evaluation over the generated scenarios.

It follows that even though CVA may be seen as another financial product and evaluated in a hybrid platform, it remains that its size raises specific problems when we wish to price it in a practical manner and in reasonable time. Although this publication covers major ground for the efficient management of CVA, including generic simulation, parallelism, and AAD, it does not cover *all* the pieces, in particular, the representation and manipulation of cash-flows, the application of regression proxies, and some important algorithms specifically designed for the acceleration of xVA and explained in [31].

## 4.2 THE ARBITRAGE PRICING THEORY

Having defined financial products in an abstract manner suitable for the development of generic libraries, and having explored in detail the key notions of cash-flow, timeline, scenario, and sample,<sup>11</sup> we are ready to discuss their valuation. Valuation is permitted by the keystone theorem of modern finance, underlying financial theory and practice for almost 50 years, established by Harrisson, Kreps, and Pliska in [32] and [33], following the ground-breaking work of Black and Scholes in [22].

## The Asset Pricing Theorem

We state the fundamental asset pricing theorem here, referring to the original articles or textbooks for demonstrations. In the absence of arbitrage, and in complete markets (both of which we assume in all that follows):

1. Financial cash-flows can be replicated by *self-financing strategies*.

A self-financing strategy is a dynamic trading strategy over a time interval, without injection or withdrawal of cash or assets (hence, self financing), that consists in holding, at time  $t$ , the underlying assets  $\{S_i\}$  that constitute the market, of prices  $\{p_i\}$ , in the amounts  $\{x_i(t)\}$ .

A self-financing strategy is specified by the vector process  $\{x_i(t)\}$ . The vector  $\{x_i(t)\}$ , the amounts of assets held at time  $t$ , must be  $\mathcal{F}_t$ -measurable, which means that they are determined at time  $t$ , from the state of the market on a before  $t$ , but not after.<sup>12</sup>

It follows from this definition that the value series of a self-financing strategy is an adapted process satisfying:

with  $x_i(0) = 0$ . One trivial, yet important, self-financing strategy is to leave  $S_i$  compounding at the short rate  $r$  on a bank account. This strategy is traditionally called “bank account,” and its value, denoted  $B_i(t)$ , is of course:

What the theorem says is that for any financial cash-flow  $\{C_i(t)\}$  (defined as a  $\mathcal{F}_t$ -measurable variable) paid at time  $t$ , there exists a self-financing strategy such that the value of the strategy replicates the cash-flow at time  $t$  with probability 1:<sup>13</sup>

The amounts of assets  $\{x_i(t)\}$  held in the replication strategy are called replication coefficients, *hedge coefficients*, or just “deltas.”

The value of the cash flow at a time  $t$  is the value of the replicating strategy.

2. There exists a unique probability measure, equivalent to the “real” probability measure,<sup>14</sup> denoted  $\tilde{\mathbb{P}}$  and called *risk-neutral measure*, such that the discounted values  $\tilde{V}_t$  of all financial cash-flows are martingales under  $\tilde{\mathbb{P}}$ :

for any  $t < T$ .

One consequence is that the value processes  $\tilde{V}_t$  of all cash-flows have a drift equal to  $\mu$  under  $\tilde{\mathbb{P}}$ .<sup>15</sup> It is also the case for financial products, in between the payment of cash-flows (the value of a product dropping by the cash-flow on a payment date).

Another consequence is the fundamental pricing formula:

3. The value  $V_t$  is a function of the market state on and before  $t$ , including the current values  $S_t$  of the underlying assets in the current market. The replication strategy, the amounts of underlying assets to trade at time  $t$  to replicate a cash-flow  $C_t$ , are the derivative sensitivities of  $V_t$  to these underlying prices:

4. More generally, the drift of the value process of any self-financing strategy under  $\tilde{\mathbb{P}}$  is  $\mu$ , and conversely, any adapted process with drift  $\mu$ , under  $\tilde{\mathbb{P}}$  is the value process of a self-financing strategy, and therefore, the price series of a tradable asset.

This theorem guarantees that the market risk of financial products can be covered with trading strategies, justifies the universal practice of measuring risk by taking derivatives of transaction values to current market prices, and provides a general pricing formula for all financial products.

Later work precised, confirmed, and extended this fundamental theorem. For instance, it was no earlier than 2009 that it was finally demonstrated, indisputably and in the most general case, that derivative sensitivities effectively always correspond to hedge coefficients. The demonstration is due to Dupire, who had to develop a new branch of stochastic mathematics, called “stochastic functional calculus,” to obtain this result [34].

## Pricing with a numeraire

Another extension, extremely useful for analytic calculations and numerical implementations, including Monte-Carlo, was introduced in 1995 by El-Karoui, Geman, and Rochet [35]. The bank account  $\pi$  is not the only choice of *numeraire*: any (strictly positive) self-financing strategy may be used as numeraire, including positive financial cash-flows and products, before payment.

For every such choice of a numeraire  $\pi$ , there exists a unique probability measure  $\pi^*$ , equivalent to  $\pi$ , such that all asset values (in between payments) in terms of  $\pi$  are martingales under  $\pi^*$ :

where  $\pi^*$ . In particular, the value of a cash-flow  $\pi^*$  paid at time  $t$  satisfies:

and with  $\pi^*$  without loss of generality:<sup>16</sup>

Remembering that cash-flows are defined as functionals of scenarios:

and that product values are the sum of the values of their cash-flow, it follows that the value  $\mathbb{V}$  of any financial product with cash-flows paid on dates  $t_1, t_2, \dots$  satisfies:<sup>17</sup>

We augment the definition of scenarios and samples<sup>18</sup> to include the numeraire on payment dates, so that:

The functional  $\mathbb{V}$  of the scenario, including the numeraire on payment dates, is called *payoff*, and we get the classic result that the value of a financial product is its expected payoff:

## Forward measures

Besides the bank account  $\mathbb{B}$ , a number of useful numeraires have been widely used in investment banks and universities to resolve analytic and numerical problems. We will see a few in this chapter, and demonstrate a simple change of numeraire in the code of [Chapter 6](#).

For now, we introduce a particularly useful family of numeraires: the zero-coupon bonds. A zero-coupon bond of maturity  $T$  satisfies the criteria of a numeraire. The associated martingale measure is called *forward measure* for a reason that will be made apparent shortly. The forward measure is often denoted  $\mathbb{F}$  and expectations taken under this measure are denoted  $\mathbb{E}^{\mathbb{F}}$ .

One of the multiple benefits of forward measures is that they provide the means to price linear products directly and independently of a model.

**Forward prices and forward contracts** Let us revisit a forward contract of maturity  $T$  and strike  $K$  on an asset  $S$ , and price it under the measure. From the fundamental pricing formula, we immediately get:

since  $S_T = S_0 e^{(r_f - r_s)T}$ . Solving for  $S_T$  in the strike  $K$  to find the forward price, we get:

Forwards of maturity  $T$  are expectations under the martingale measure associated with the zero-coupon bond of maturity  $T$ , hence its name “forward measure.” It also follows that forwards of maturity are martingales under  $\mathbb{P}^F$ , in particular:

These formulas are not particularly helpful for pricing, since today's forwards trade in the market. We know their prices. Where they are helpful is the other way around: the forwards that trade on the market today *imply the means of forward-neutral measures*.

Another trivial consequence is that the value of a forward contract is  $S_0 e^{(r_f - r_s)T}$ . Setting  $K = S_0 e^{(r_f - r_s)T}$ , we get the time  $T$  value of an asset delivered at time  $T$ :

**Forward rates and FRAs** The same can be said of FRAs in interest rate markets. A FRA over the period  $[T, T+1]$  pays a cash-flow on date  $T+1$ , hence, we value it under  $\mathbb{P}^F$ :

Solving  $S_{T+1} = S_T e^{(r_f - r_s)(T+1)}$  in  $r_f$ , we find that the forward Libor is the forward-neutral expectation of the Libor:

and a martingale under  $\mathbb{P}$ . Today's Libor curve implicitly defines the expected values of future Libors under the forward-neutral measures for their end dates. In addition, the value of a FRA is

Setting  $\tau = T$ , we see that the time  $t$  value of a FRA paid at  $T$  is:

### Forward discount factors and instantaneous forward rates

Consider now a product that delivers a zero-coupon bond of maturity  $T$  at a time  $t$ . Its value is obviously  $e^{-rt}$ . From the valuation formula under  $\mathbb{P}$ , we get:

We know, now, that forward-neutral expectations are forwards. It follows that the forward price  $F_{t,T}$  of maturity  $T$  of the zero-coupon bond of maturity  $T$  satisfies:

In addition, the value series  $(F_{t,T})_{t \in [0, T]}$  of a forward discount factor of maturities  $T$  is a martingale under  $\mathbb{P}$ . It follows that today's discount curve fixes the forward-neutral expectations of future discount factors.

The amount

is called *forward discount rate*, and effectively corresponds to the rate of a loan from  $t$  to  $T$  that may be locked at time  $t$  by selling zero-coupon bonds of maturity  $T$  and buying zero-coupon bonds of maturity  $t$  for the same total price.

Denoting  $\mathbb{P}$  and considering the limit case where  $\delta \rightarrow 0$ , we get the time  $t$  instantaneous forward rate of maturity  $\bar{r}_t$  :

which we had defined before, only now we can see where this definition comes from and why the short (discount) rate is effectively

**Forward bank account** Another example is the bank account  $B_t$ . It is a martingale under  $\mathbb{P}$ , therefore:

Since forward-neutral expectations are forwards, it follows that  $\mathbb{E}^{\mathbb{P}}[B_{t+T}] = B_t$  is the forward value of the bank account  $B_t$  at time  $t$  for the maturity  $T$ . In particular, we have today's value of bank account forwards of all maturities:

**Interest rate swaps** It immediately follows from what precedes that the value on an IRS (a receiver swap in this instance) is:

and, solving for  $\bar{r}_t$  in  $\mathbb{E}^{\mathbb{P}}[B_{t+T}] = B_t$  to get the par swap rate  $\bar{r}_t$ , we get:

This formula is universal and model independent, which implies that the par swap rate is a redundant market primitive, hence, not directly included in scenarios. The denominator is called the *annuity* of the fixed leg. The annuity is a combination of zero-coupon bonds, hence, a tradable asset, with time  $t$  value:

On the par swap rate formula, we recognize that  $\bar{A}_t$  is the value at time  $t$  of the floating leg *in terms of the annuity*:

It follows that  $\bar{A}_t$  is a martingale under the martingale measure associated with the annuity.

Also note that the formula of the value of a (receiver, in this instance) IRS can be rearranged into:

**Swaption pricing** A *physically settled* payer swaption (right to enter the underlying payer swap) of strike  $K$  on the exercise date  $T$  is the value, if positive, of the underlying swap on the exercise date:

We can therefore price it under the annuity-neutral measure:

where the payoff is that of a standard put (for a payer swaption, it is a call) on the swap rate, which itself is a martingale. This is the theoretical basis for pricing swaptions under models like Black-Scholes [22] or SABR [36]. These established market practices therefore implicitly, and correctly, manage swaptions under the annuity-neutral measure.

Assuming a constant volatility  $\sigma$  as in Black and Scholes's specification,  $\mathbb{P}$  is given by Black and Scholes' formula [22].

Assuming  $\sigma$  is a stochastic volatility martingale as in SABR's specification,  $\mathbb{P}$  is given by Hagan's formula [36].

**Natural numeraires** We demonstrated that forward prices  $F_{t,T}$ , observed at time  $t$  with maturity  $T$ , are martingales under the forward measure  $\mathbb{F}$  associated with the zero bond of maturity  $T$ , the maturity date of the forward. This zero bond and forward measure are respectively called the *natural numeraire* and the *natural measure* of the forward.

In particular, the spot price  $S_t$  on time  $t$  coincides with its forward  $F_{t,t}$ . Its natural numeraire is therefore the zero bond of maturity  $T$ , and its natural measure is  $\mathbb{F}$ . *Forwards are expectations of asset prices under their natural measure.*

This is true of all asset prices. In particular, the forward price of the bank account  $B_t$  on time  $t$  is its expectation under its natural measure

The forward discount factor:

$d_t$  is a martingale under  $\mathbb{F}$ . Therefore:

The natural measure of the discount factor  $\delta_t$ , observed at time  $t$  with maturity  $T$ , is therefore  $\delta_t(T)$ . The expectation of a discount factor under its natural measure is the corresponding forward discount factor observed today. Its natural numeraire is the zero bond of maturity  $\delta_0$ .

A forward Libor  $L_{t,T}$  observed at time  $t$  with maturities  $T$  is a martingale under its natural measure  $\delta_t$ , associated with its natural numeraire, the zero bond of maturity  $\delta_0$ . In particular, a spot Libor observed at  $t$  with maturity  $T$  coincides with the forward on the fixing date. It follows that the expectation of a Libor under its natural measure is the forward Libor. Its natural numeraire is the zero bond of maturity the Libor's payment date.

Finally, we have seen that the par rate  $r_p$  of a given swap, paid over the annuity  $a_p$  of the swap's fixed leg, is a martingale under  $\delta_0$ .  $\delta_0$  is therefore the swap rate's natural measure, and the fixed leg's annuity  $a_p$  is its natural numeraire.

Every observation of a market variable  $x_t$  fixed at time  $t$  admits a natural measure, such that conditional expectations  $E_t[x_s]$  of  $x_s$  under the natural measure are, by definition, the forward values observed at time  $t$  of the cash-flow  $x_s$  at time  $s$ . It follows that those forward values are martingales under  $\delta_t$ 's natural measure. The numeraire associated to the natural measure of an observation  $x_t$  is called its natural numeraire. The notion of natural numeraire and measure depends on the nature of the market variables, and we have seen that prices, Libors, discounts, bank accounts, and swap rates define their natural measures in a different way. In addition, natural measures relate to observations: two observations on different dates of the same variable have different natural measures and numeraires.

In economic terms, the observation of a market variable always expresses the value of an asset in terms of a unit. This unit is the observation's natural numeraire. The market price of an Apple share on 30 September 2025 expresses the value of a fraction of Apple in terms of 30/09/25 USD. Its natural numeraire is the USD denominated zero bond of maturity 30/09/25.

In formal terms, the natural numeraire of an observation is the numeraire such that under any complete and arbitrage-free probabilistic specification with correct initial values, corresponds to the forward price observed at of the cash-flow . This is rather abstract, but in practical terms, for a given observation of a market variable, it is generally quite clear what its natural numeraire and measure are. In particular, for the market variables that populate market samples:

1. The natural measure for a forward price is , including spot prices with natural measure .
2. It directly follows that the natural measure for a discount factor , or the bank account observed at , is also .
3. The natural measure for a forward Libor is , including spot Libors with natural measure .

**Constant Maturity Swaps** We computed the values of all the linear transactions introduced earlier: forward contracts, FRAs, IRS, without specifying a model, directly, as a function of the forward prices, forward rates, and discount factors trading on the current market. This is a defining characteristic of linear products: linear products are financial products whose value depends on the current market, independently of a model. It follows that today's value of linear products is the same in any arbitrage-free model that respects today's market. We just saw that forwards are expectations under natural measures, so we have a clean definition of linear products:

Linear products are those whose value depends on *expectations*, but not *distributions*. It follows that the value of any linear product is a linear combination of the present values of a number of forward contracts, hence the name “linear” products.

This (correct) definition of linear products conflicts with the (approximate but widely accepted) one of products where cash-flows are linear functions of scenarios. Consider a cash-flow  $\tilde{C}_t$ , linear in the market variable  $S_t$  fixed on  $t$ . We now know that its value satisfies:

When  $\tilde{S}_t$  is  $S_t$ 's natural measure,  $\tilde{C}_t$  is its forward, therefore:

and it follows that the cash-flow is indeed a linear one. When this is not the case, and  $\tilde{S}_t$ 's natural measure  $\tilde{\mathbb{P}}$ :

The difference between the two expectations  $\mathbb{E}[\tilde{C}_t]$  and  $\tilde{\mathbb{E}}[\tilde{C}_t]$  is called *convexity adjustment* and can be calculated with Girsanov's theorem, but only under dynamic assumptions on the future behavior of  $S_t$ , in particular its volatility. It follows that the cash-flow is *not* a linear one, but a European one (because it depends on a single variable on a single date).

To see that clearly on an example, a Libor  $L(t, T)$  is a linear cash-flow when paid at  $T$ . Its value is then:

as seen earlier. When paid at  $T$ , its value is:

and:

Ignoring discount basis and assimilating:

the cash-flow                  paid at      is equivalent to a      payment, under the underlying Libor's natural numeraire, of:

which clearly identifies the cash-flow as a nonlinear quadratic function of the underlying Libor. Its linearity was only apparent. It is actually a European cash-flow. Under a simple model a la Bachelier [37] with Gaussian first increments with annual standard deviation  $\sigma$ , its value is:

(where we skipped trivial calculation steps). In a more general, non-Gaussian context, we can statically replicate this cash-flow, like any European cash-flow written on a Libor, with combinations of caps and floors, by an application of Carr-Madan's formula.

This is the cash-flow of a well-known, actively traded financial product called “Libor in arrears.” More generally, an actively traded family of financial products, called constant maturity swaps or CMS, pay cash-flows linked to libor or swap rates, over schedules different than their natural numeraire, resulting in nonlinear products, which value depends on volatility and dynamic assumptions. Like Libor in arrears, swap-based CMS are statically replicated with combinations of swap-tions, with techniques similar to Carr-Madan's, as explained in [38].

Nonlinear products, like swaptions or CMS, cannot be valued out of the underlying market alone. Their value depends on volatility assumptions, and more generally, dynamic assumptions of how the market is going to evolve. This is the purpose of derivatives *models*.

## 4.3 FINANCIAL MODELS

Having defined financial products in a general manner, we repeat the exercise for financial models. The asset pricing theory gave us the pricing formula for all financial products:

where the payoff function  $\pi$  and the definition of the scenario (what market primitives, including numeraire, on what dates) are fully specified by the product.<sup>19</sup> In order to apply this formula and produce prices, we must specify the distribution of  $\pi$  under  $\mathbb{P}$ . This is the defining purpose of a model.

We call “model” a specification of the probability distribution of the scenario  $\mathbb{P}$ . This specification may either be direct, or implied from dynamic assumptions regarding future market behavior, in which case we say that the model is *dynamic*.

More precisely, the scenario:

is a collection of market samples  $\pi$  on a collection of event dates, and each sample is itself a collection:

of market variables on the event date  $t$ . A sample may be a singleton, or a discrete, infinite, or continuous collection of market variables.

Hence, the scenario  $\mathbb{P}$  is a collection of market *observations*:

where the  $s$  are all the individual market variables on all the event dates. Different market variables on the same date are different  $s$  and the same market variable on different dates are also separate  $s$ . Each  $s$  belongs to one sample  $\mathcal{S}$ , and we denote  $t_s$  its event date, also called observation date or *fixing date*. We now have a more precise definition of models.

The defining purpose of a model is to specify the joint distribution of the observations  $s$  in the scenario  $\mathcal{S}$ .

Note the separation of concerns: the payoff function  $\pi$  and the definition of the scenario  $\mathcal{S}$  are specified by the product outside of any kind of model logic. The model specifies the probability distribution of  $s$  outside of any cash-flow logic.<sup>20</sup>

A model is not free to specify *any* joint distribution for the  $s$ . It must respect its defining property that asset prices in terms of the numeraire are martingales; otherwise the model is arbitrageable, produces inconsistent prices, and is generally not acceptable. All common models introduced shortly are arbitrage-free. In addition, the model's specification of  $\mathcal{S}$  must respect the current market prices. For instance, we have modeled underlying stocks, indices, and currencies as financial products. Their value under  $\mathcal{S}$  as financial products must correspond to their current market price. More generally, we have seen that the forward prices and rates, traded on today's market, fix expectations under forward-neutral measures, imposing a constraint on the model, called *initial value conditions*.

We shall review linear and European models, appropriate for the valuation of respectively linear and European transactions, before we discuss dynamic models, the most general class of models, those appropriate for the pricing of all products.

## Linear models

Linear models refer to what we have been calling so far “today’s market.” They contain all the underlying asset prices, including dividend and repo curves, all the discount and Libor curves, and other variables that define the current state of the market. They “know,” among other things, the forward prices and discount factors of all maturities , as well as forward Libors of all maturities and , for all Libor indices.

Linear markets achieve this capability by interpolation of the available market information: spots, futures, par swap rates, basis and currency swaps, and so forth. This is not as easy as it may sound, especially since the multiplication of basis and discount curves in 2008–2011. To properly construct a linear market is a sophisticated, challenging exercise that occupies some of the greatest minds in the industry. We refer to [39] for a review of linear markets, including a presentation of the difficulties and the sophisticated solutions implemented in modern systems.

Mathematically, linear markets, who know all forwards, define the *expectations* of the s under their respective natural measures, but they know nothing else of their *distribution* , because they have no knowledge of their future dynamics. This transpires in the notations: linear markets know all the for all maturities , but nothing of the when . Linear markets are initial value conditions. They are the modern counterpart of the “initial spot ” found in traditional literature.

Why then do we call them linear *models* and discuss them in the model section here? We have defined a model by its ability to assign a probability distribution to a scenario. Linear models assign probability 1 to the *forward scenario* where every market variable lands, on its event date, on its forward value seen from today, which we have seen is its expectation under its natural measure :

This “probability distribution” of the scenario obviously respects initial conditions, and is trivially arbitrage-free, the discounted price series of all assets being (constant) martingales between payments. This is actually the simplest possible arbitrage-free model consistent with today's market.

It follows that the price of any financial product defined by its payoff and scenario in the linear model is:

which may explain why those models are called linear, although:

so the linearity of linear models is an approximate notion.

We priced a number of linear products in the linear model in [Section 4.2](#), although we did not call it linear model at the time. The value in the linear model is called the *intrinsic value* of the product. For European options, this definition coincides with the traditional definition. For instance, for a European call,

. Our definition extends the notion of intrinsic value to all financial products. The difference between the actual value and the intrinsic value is called *time value*: .

Linear products are those whose time value is zero.<sup>21</sup> The intrinsic value prices the linear part of a product and is independent of volatility or dynamic assumptions. The time value prices the nonlinear part of a product and depends on volatility and other dynamic assumptions.

From a practical point of view, pricing in a linear model is extremely fast, usually a fraction of a millisecond, because it takes a unique scenario. It is also very accurate because no numerical algorithm is involved. Linear models are suitable for Monte-Carlo simulations; in particular they fit the simulation model interface we build in [Chapter 6](#), albeit in a trivial manner, since they generate only one scenario: the

forward scenario . Generic, well-designed valuation systems price linear products (and the intrinsic value of other products, an interesting information in its own right) as a simulation in a single scenario.

More sophisticated models used for the valuation of nonlinear transactions can also price linear products, as a special case of options or exotics. Provided that the models respect today's market, the prices coincide with the linear model, but with lower speed and accuracy.

Finally, note that those more sophisticated models coincide with linear models, not only for linear products, but for all products, when their volatility parameters are set to zero.<sup>22</sup> Therefore linear models are also zero-volatility arbitrage-free dynamic models.

## European models

Moving one step up to European models, we discuss those models appropriate for the pricing of European products, but not exotics.

We have seen that a European cash-flow of maturity is a combination of calls and puts on the same market variables, same maturity, and same payment date, as demonstrated by Carr-Madan's formula. It follows that a dynamic model is not necessary for the valuation of European cash-flows; all we need are the call prices of all strikes for the maturity .<sup>23</sup> European models, by definition, “know” these prices in the same way linear models “know” forward prices: by interpolation of available market data, in this case, the prices of calls and puts that trade on today's market.<sup>24</sup>

Linear models generally interpolate discount factors not in price, but in rate. Similarly, options are typically interpolated not in their price , but in their Black-Scholes *implied volatility* , defined by:

where is Black and Scholes's formula [22]. We call *implied volatility curve*, or IVC, the continuous collection of Black-Scholes implied

volatilities of a given maturity. An IVC of maturity  $t$ , *combined with a linear model* for the forward and discount in Black and Scholes's formula, can price all European cash-flows of maturity  $t$ . It follows that an interpolated collection of IVCs for all maturities  $t$ , called *implied volatility surface* or IVS, can price all European cash-flows, hence, all European products.<sup>25</sup> In case a European product includes European cash-flows on different underlying assets, we need multiple IVS, one per underlying asset, although, in the interest of clarity, we just call IVS the collection of all call prices on all underlying assets.

European models, or IVS are what we call them now, therefore build on the information contained in linear models, and inject additional information in the form of a continuous, bi-dimensional surface of Black-Scholes implied volatilities. We have seen that linear models define the *expectations* of the observations  $x_t$  in a scenario under their natural measure. European models, or IVS as we call them now, define the *marginal distributions*  $\pi_t(x)$  of these observations under their natural measure. To see this, we apply the valuation formula to a call of strike  $K$ , maturity  $t$ :

Differentiating twice on both sides against the strike  $K$  (in the sense of Laurent Schwartz's distributions,<sup>26</sup> we obtain:

where  $\pi_t(x)$  is the natural numeraire for the observation  $x_t$ , and the normalized distribution  $\pi_t(x)$  is the price of the Dirac mass of  $x_t$  in  $\mathbb{R}$ , called *Arrow-Debreu security*. The call prices of all strikes  $K$  for a maturity  $t$  are equivalent to the natural probability distribution of the underlying observation.<sup>27</sup> Hence, the IVS, which knows the call prices of all strikes and maturities on all relevant assets, also knows the *marginal* probability distributions of all the market variables  $x_t$  in the scenario, called *marginal distributions*.

Because the IVS delegates forwards to the linear model, the means of the natural distributions  $\mu_s$  are the forwards of the observations  $s$ . It follows that the IVS respects today's market and produces arbitrage-free marginal distributions. It does *not* produce arbitrage-free *joint* distributions between underlying asset prices on different dates, but, for European products, all that matters are the marginal distributions of the  $s$ .

The price of a European cash-flow  $C_s$  paid on the natural payment date  $s$  of an underlying observation  $s$  is  $\int_{-\infty}^s C_u dF_u$ , the integral of the cash-flow's payoff against the distribution of its underlying observation specified in the IVS. In practice, this integral may be evaluated by numerical integration. [Chapter 4](#) of *Numerical Recipes* [20] covers some efficient procedures and provides source code. Numerical integration consists in the approximation:

where the weights  $w_k$  and the knots  $k$  are carefully chosen by the numerical integration scheme. To price a European *product*, we may price all its cash-flows and sum them up, or we can price it all together with the help of a change of variable (assuming a discrete schedule of cash-flows and a discrete scenario here to simplify notations, although everything applies in the continuous case too):

where  $F(s)$  is the cumulative probability distribution of  $s$ , known to the IVS, as well as its inverse. The value of the product is the sum of the values of its cash-flows:

The payoff function  $C_s$  of the product is the sum of the payoff functions of its cash-flows, hence:

where, again, the weights  $w_i$  and the knots  $\kappa_i$  are determined by the numerical integration routine, we used the property of European cash-flows where they each depend on exactly one  $\kappa_i$ , and the definition of  $w_i$  should be clear for the context.

A numerical integration is typically extremely fast, of millisecond order, and very accurate, in particular compared to Monte-Carlo. This procedure can also price linear products (although slower and less accurately than linear models). But it cannot correctly price exotics.

The IVS correctly values European products because it knows the (natural) *marginal* distributions of all the observations  $\kappa_i$  in a scenario, but it knows nothing of their *joint* distribution, including the dependence of different underlying assets on the same date, or the dependence of the same underlying asset on different dates. The mathematical mapping of a number of marginal distributions into a single joint distribution is called a *copula* and there exists infinitely many copulas that join the same set of marginals into different joint distributions.

A defining characteristic of European products, and a better definition than the one stated earlier, is that these products depend on marginal distributions, but not on copulas (in the same way that linear products only depend on expectations but not distributions). *European cash-flows are those whose payoff can be written as a function of one observation, under the natural numeraire of the observation.* It follows that a defining characterization of non-European exotics is that they *do* depend on copulas. To see that on a simple example, consider the twice-monitored digital barrier that pays \$1 if two fixings of the underlying asset price on two given dates  $\kappa_1$  and  $\kappa_2$  are both below a barrier  $B$ :

Then:

In this formula, the terms  $\mu_1$  and  $\mu_2$  are European digitals that only depend on marginal distributions, but the term

clearly refers to their dependence.

To price this exotic with the IVS would result in a random and incorrect value. Random because the implementation of the IVS implicitly and arbitrarily chooses a copula: in our numerical integration the same quantiles were sampled simultaneously for all the marginal distributions. And incorrect because a random copula almost certainly results in an arbitrageable joint distributions, even when the marginal distributions are arbitrage-free.

Linear models define the expectations of underlying variables and correctly price linear products. European models inject additional marginal distributions around these expectations, and correctly price the time value of European products, in addition to linear products.

The purpose of dynamic models is to produce an arbitrage-free, credible copula between the marginal distributions of the IVS, so as to correctly price not only linear and European products, but also exotics.

## Dynamic models

To directly specify the copula is a perilous and probably impossible task. The copula must be specified in a way to produce an arbitrage-free, credible joint distribution, and it is unclear how this can be done, working directly on distributions.

To achieve this, dynamic models specify an arbitrage-free stochastic dynamic for the underlying market variables. They fit expectations by respecting the initial values given by the linear model. They fit marginal distributions by *calibration* to the IVS.<sup>28</sup> And they inject copulas, implicitly, with a credible, realistic, and arbitrage-free stochastic dynamics.

All dynamic models calibrated to the IVS produce the same marginals as the IVS, hence the same values for all European transactions. But models also implicitly specify a copula between these marginal distributions. It follows that two different models (say, one with local volatility and one with both stochastic and local volatility) both calibrated perfectly to the same IVS (which in this case is feasible) produce the same marginals but different copulas. It follows that they assign the exact same value for Europeans (modulo calibration and numerical error) but a potentially very different one for exotics. The sensitivity to the copula is a measure of *model risk*. It may be estimated by pricing an exotic or exotic book with different models calibrated to the same IVS. The price dispersion measures “how exotic is this particular exotic.”

For example, it was discovered by Dupire in [45] that a variance swap, which pays the realized statistical variance of an underlying asset price over a period, can be perfectly replicated by a combination of Europeans and a self-financing strategy, irrespective of any dynamic assumption. This theory is the basis of the VIX index published by the CBOE and widely considered as a primary measure of market nervousness. The details are explained, for instance, in the fourth part of

[46]. It follows that the variance swap is completely insensitive to copulas, and priced identically (modulo numerical errors) in different models, as long as they calibrate to the same IVS.

As another example, Peter Carr's well-known static replication of barrier options [47] shows that (under some pretty strong assumptions) simple barrier options are also replicated by combinations of European options and trading strategies. It follows that barriers like down-and-out calls are not that exotic (with the barrier on the strike, the value of a down-and-out call is even linear in the underlying asset price in all continuous models) and their price in different models calibrated to the same IVS is very similar.

On the contrary, tight double no touches (options that pay \$1 if the underlying asset price remains in tight corridor at all times for a period) have high *volga* (second derivative to volatility) and are therefore very sensitive to stochastic volatility. A stochastic volatility model may price such products orders of magnitude higher than models without stochastic volatility, calibrated to the same IVS. It follows that double no touches are very exotic products, with a massive dependency on copulas, and material model risk. Financial institutions should always classify exotics in accordance to model risk, since this is one of the most difficult risks to cover.

The authors of dynamic models generally deliver exact or approximate closed-form solutions for European options, which helps calibration.<sup>29</sup> More sophisticated models like Heston's stochastic volatility model [42] don't offer analytic formulas but can value Europeans very quickly with numerical integrals; see the dedicated chapters of [6]. In the simplest models like Black Scholes and, some exotics like barrier options also admit closed-form solutions. But in a vast majority of contexts, the pricing of an exotic necessitates a numerical algorithm. The ultra-efficient FDM, briefly introduced later in this chapter, can be applied in a limited number of situations, other contexts requiring the slower, but almost universally applicable,

Monte-Carlo algorithm covered in more detail in the rest of this publication.

Note that the different types of models introduced in this chapter form a hierarchy: linear models are self-contained and correctly, accurately, and efficiently price all linear products. European models (IVS) depend on linear models, so they also correctly price linear products. They are also implemented with accurate and efficient numerical integration algorithms, although not as fast or as accurate as linear models. But they correctly price European products in addition to linear products. Dynamic models depend on linear models for their initial conditions, and on IVS for their calibration. Mainly implemented with Monte-Carlo, they are many orders of magnitude slower and less accurate than European models. But they may value exotics in addition to pricing European and linear products, and they do so consistently with European and linear models. Hybrid models are assemblies of single underlying dynamic models, typically joined with a correlation matrix and made consistent and arbitrage-free by drift adjustments in their various components: for example, the risk-neutral drift of underlying assets that don't pay dividends is the short rate. When a stochastic price model is coupled with a stochastic rate model, the drift of the price process is the short rate produced by the rate model. Another example is that in a multi-currency context, the risk-neutral drift of assets denominated in a foreign currency must incorporate a convexity adjustment known as the quanto adjustment. See [27] for details. It follows that in a hybrid model, each (sub-) model is the same exotic model used for single underlying exotics, and that hybrid models value those exotics consistently with exotic models, hence, also, consistently with IVS and linear models. Hybrid models are slower than exotic models, but they can value hybrids and conduct bank-wide regulatory calculations.

This hierarchy is fundamental. The intrinsic value of Europeans, exotics, and hybrids is produced in the exact same model that prices linear products. Calibration guarantees that the European part of exotic

products is valued consistently with the IVS used for the management of Europeans. Hybrids and institution-wide regulatory calculations are conducted in assemblies of the same exotic models that are used to manage exotics. Not only does this provide consistent risk management across an investment bank's trading book, it also satisfies the regulator's justified insistence that regulatory calculations, like credit counterparty risk (CCR), revalue transactions over simulated scenarios *in a manner absolutely consistent with front office pricing.*

We don't implement a model hierarchy in this publication (actually, we don't implement interest rate models at all, not even linear ones, and keep the initial state as parameters in exotic models for simplicity) because that would double the size of the book and because it is not the primary topic here. Model hierarchies, and the consistency they offer, are nonetheless an extremely important subject in modern finance, and something we intend to address in a forthcoming publication.

The rest of this chapter briefly describes some of the most popular dynamic models. Our introduction is brief and without much detail, with the exception of Dupire's model [12], which we use as the main example to demonstrate parallel simulations and AAD, and which is therefore covered in depth in the rest of this publication. We recommend the three volumes of Andersen and Piterbarg's [6] for a deep and broad review of many financial models.

## The Black and Scholes model

The Black and Scholes (BS) model [22] is a single underlying price model on an underlying asset , like a stock, currency, or index, whose value today is . BS copes with a deterministic short rate , which, in a model hierarchy, should be set to the IFR to respect today's market.

The model also allows a deterministic dividend yield  $\mu_d$ . To hold a dividend-paying asset is *not* a self-financing strategy. What is a self-financing strategy is to hold the asset and reinvest dividends into more units of the asset. Denoting  $S_t$  the number of assets held at time  $t$ ,

and it follows from the reinvestment of the dividend paid at  $t$  that:

so

and the value of the strategy is:

It follows that

and that the drift of  $S_t$  is the drift of  $S_0$  plus  $\mu_d$ . Under the risk-neutral measure, associated with the numeraire  $S_t^*$ , the value of the self-financing strategy has a drift  $\mu_{S^*}$ ; hence the risk-neutral drift of the price series  $S_t^*$  is:

The defining assumption of the Black and Scholes model is that asset prices are diffusions with a deterministic volatility  $\sigma$ . It follows that the risk-neutral process of  $S_t^*$  is:

where  $B_t$  is a standard Brownian motion under  $\mathbb{P}^*$ .

## Change of numeraire

Models can be equivalently rewritten under different measures, associated with different numeraires than the classic bank account. A wise choice of numeraire may simplify analytic calculations and improve the efficiency of numerical algorithms. We demonstrate this technique

in the simple case of the Black and Scholes model, although it applies in a vast number of more sophisticated contexts; see [35] for some classic applications.

**Forward measure** First, we note that with deterministic interest rates, all forward-neutral measures are identical and coincide with the risk-neutral measure. To see this, we price a cash-flow paid at date under the risk-neutral measure:

where we moved the deterministic out of the expectation.

Repeating the exercise under the forward measure , we get:

Trivially:

and it follows that , for any -measurable variable . Hence, by definition, the two measures coincide. This is all true for any maturity . Hence, all measures coincide with the risk-neutral measure and are therefore identical.

It follows that forwards are also risk-neutral expectations in this case:

and that all forwards of all maturities are martingales under . In addition, the volatility of a forward is:

so the risk-neutral dynamics of all forwards is simply:

Applying Ito's lemma to the logarithm, we get:

and it follows that for all times  $t$  :

Hence, the *distribution* of  $S_t$  conditional to  $\mathcal{F}_t$  is:

where  $\mu = \ln(S_0) + (\mu - \frac{\sigma^2}{2})t$  and  $\sigma$  is a standard Gaussian variable.

**Spot measure** An algorithm implementing the model under the risk-neutral measure must track the spot price  $S_t$  and the bank account numeraire  $B_t$ .<sup>30</sup> Alternatively, we can equivalently rewrite the model using the underlying asset itself as a numeraire, something known in literature as “spot measure,” so we no longer need to additionally track the numeraire, the underlying asset *being* the numeraire.

With dividends, however, the asset price does not qualify as a numeraire. We fix a horizon  $T$  on or after the last date of interest in our working context, a technique known as “terminal measure.” We consider the present value of a forward of maturity  $T$  :

where we dropped the reference to the fixed horizon to simplify notations, and  $S_T$  trivially qualifies as numeraire. We call spot measure the associated measure  $\mathbb{P}_S$ , even though  $S_T$  is not exactly the spot price of  $S_T$ .

Pricing a  $\mathcal{F}$ -measurable cash-flow under the spot measure:

and comparing with the risk-neutral measure:

(where we moved the deterministic bank account out of the expectation) implies that for any  $\mathcal{F}$ -measurable variable  $X$ ,

It follows that the Radon-Nikodym derivative of the spot-neutral to the risk-neutral measure, conditional to  $\mathcal{F}_t$ , is:

where

is an exponential martingale under  $\mathcal{F}$  with volatility  $\sigma$ . Therefore, by Girsanov's theorem:

is a standard Brownian motion under  $\mathcal{F}$ , and we get the dynamics of the forward (*any* forward) under the spot measure:

Hence, the *distribution* of  $X_t$  conditional to  $\mathcal{F}_s$  under the spot measure is:

where we note that the only difference with the risk-neutral measure is the sign of the left term in the exponential.

## The Black and Scholes formula

Having investigated some fundamental properties of the Black and Scholes model, we can easily derive the analytic formula for call options (the formula for puts following from the call–put parity):

and

where we used the notation  $\Phi$  to refer to a standard Gaussian variable or the cumulative normal distribution, which is clear from context, and the traditional  $\mu$  and  $\sigma$  notations.

One singular beauty in the Black and Scholes formula is that it exposes the hedge coefficients, too. In principle,  $\Delta$  and  $\Gamma$  are functions of  $\mu$  and  $\sigma$  so the differentiation of the formula could be a painful, if trivial, exercise. Noting that an option price is homogeneous in the for-

ward and strike (an option on assets for times the strike is like options on one asset for the strike price), it follows by Euler's theorem that the price is the sum of its derivatives, hence:<sup>31</sup>

It follows that the replication strategy at time  $t$  is hold for-wards and borrow an amount of cash, the value of the replication portfolio being trivially given by Black and Scholes's formula. It follows that the formula does not only express a price: it expresses the *replication strategy*, of which the price is a byproduct.

## Dupire's model

The Black and Scholes model, despite its simplicity, remains a strong reference in the theory and practice of option trading, and not only because of its historical importance or its remarkable analytic tractability. The Black and Scholes model remains a keystone derivatives risk management instrument, useful in many important ways besides converting prices to volatilities and back. To appreciate the robustness of the model and its relevance for risk management, see the recent [48].

But the model suffers from one major flaw: the same assumption of a deterministic volatility that gave it such remarkable analytic tractability prevents it from correctly calibrating to option markets. It should be clear that with a single time-dependent free parameter  $\sigma_t$ , the model may hope to hit, at best, one option price (one strike) per maturity. It follows that the *implied volatility* (obtained by solving in volatility so Black and Scholes's price corresponds to the market price) may be different for different strikes, a phenomenon commonly known as “smile,” the curve of implied volatility as a function of the strike generally rising away from the money on both sides,<sup>32</sup> looking like a smile (with a bit of imagination).

In the early 1990s, Bruno Dupire worked out the simplest possible extension of the Black and Scholes model that can be calibrated to option prices of all strikes and maturities. Like most ground-breaking ideas, Dupire's specification looks trivial after the fact: make volatility a function of the underlying asset price. This is called *local volatility* and presented here under [\[12\]](#), with zero rates and dividends for simplicity:

To calibrate Dupire's model means to solve for all strikes and maturities, in the *local volatility surface*  $\sigma(S, t)$ . We are calibrating a continuous, two-dimensional local volatility surface, to a continuous, two-dimensional IVS. Intuitively, we could find a unique, perfect fit. Dupire's formula [\[12\]](#) confirms the intuition and even provides an explicit formula for  $\sigma(S, t)$  as a function of today's options prices  $C(S, K, T)$ .

We discuss, demonstrate, implement in code, and apply Dupire's formula to risk management in [Chapter 13](#). Dupire's formula is widely considered one of the three most important formulas in finance (the other two being Black and Scholes and Heath-Jarrow-Morton [\[49\]](#), which we introduce shortly). Dupire's model, or its stochastic volatility extensions by Dupire himself [\[50\]](#) and [\[45\]](#) and Bergomi [\[51\]](#), power almost all modern exotic risk management platforms.

## Stochastic volatility models

The milestone stochastic volatility models, where volatility itself is a random process, were developed in 1987 by Hull and White [\[52\]](#), 1993 by Heston [\[42\]](#), and 2002 by Hagan [\[36\]](#). They gained traction in the early 2000s, when the industry realized that “options are hedged with options,” and that stochastic volatility is necessary to measure the consequences on valuation and risk management; see [\[46\]](#).

We briefly introduce Hagan's SABR, which specification is simpler than Heston, although both models produce a similar behavior and mainly differ in their implementation. Heston's model and its implementation are explained in detail in the first volume of [6].

SABR is written directly under a forward or annuity measure, depending on context<sup>33</sup> where the forward is a martingale. With the addition of stochastic volatility, it implements the following dynamics:

The first line is a trivial extension of Black and Scholes with a power coefficient. The rest specifies the dynamics of volatility. The first thing to note is that this model is incomplete, in the sense that two sources of risk, represented by the two Brownian motions, cannot be hedged by trading one underlying forward. In addition, the initial value  $\sigma_0$  is undefined.<sup>34</sup> It follows that we can't apply the asset pricing theorem unless *we complete the model by calibration*. After we calibrate  $\sigma_0$  to hit the market price of some option  $C_0$ ,<sup>35</sup> the model is complete (because we now have *two* hedge instruments: the forward and the option), with well-defined initial conditions and an arbitrage-free dynamics (the forward is a martingale, and so is the numeraire deflated option price, by construction, because  $\mathbb{E}[C_t] = C_0$ ).

A strong dependency on calibration is a defining characteristic of stochastic volatility models. Calibration is always a concern, but stochastic volatility models, unless calibrated, are ill-defined and fundamentally incorrect. The purpose of these models is to measure the impact of hedging with options on values and risk, hence the strong reliance on the initial values of options.

The parameters  $\alpha$  (volatility of volatility),  $\beta$  (an additional local volatility parameter), and  $\rho$  (correlation between the forward and its

volatility) are the free parameters of SABR. Note that the acronym SABR comes from the initials of its parameters: Sigma-Alpha-Beta-Rho.

Hagan produced a closed-form *approximation* for European options under the SABR dynamics using small noise expansion techniques. The result is a very precise approximation. It may be found in the original paper, a number of textbooks, or Wikipedia's SABR article. Hagan's solution involves steps specific to the SABR model and is not easily reusable. For a general presentation of expansion techniques and applications outside SABR, see Andreasen and Huge's [53].

Stochastic volatility is covered in a vast amount of literature, including Lipton's textbook [54], Gatheral's lecture notes [43], and Bergomi's recent [51]. It is also covered in detail in the lecture notes [46].

## Interest rate models

**The Heath-Jarrow-Morton framework** Despite several early attempts to implement models *a la Black and Scholes* in the interest rate world, it was not before 1992 and the ground-breaking work of Heath, Jarrow, and Morton (HJM, [49]) that a consistent, general, and correct theoretical framework was established for modeling stochastic interest rates.

In interest rate markets, the “underlying asset” is the collection of zero-coupon bonds of all maturities, their value at time  $t$  being the *continuous curve*  $\{D(t, \tau)\}_{\tau \geq t}$ . Ignoring stochastic basis for now, the state of the model at time  $t$  is the entire discount curve, or, equivalently, the instantaneous forward rate (IFR) curve, introduced earlier:

Under the risk-neutral measure, the drift of all discount factors must be the short rate  $r_t$  :

where  $\sigma$  is the volatility of discount factors, it is conventionally negative, so that Brownian increments that positively affect rates, negatively affect bonds, and we introduced a one-factor version (one Brownian motion) for simplicity, although the original paper works with the general multifactor case, where equations remain essentially unchanged with matrices in place of scalars.

The celebrated Heath-Jarrow-Morton formula gives the dynamics of the IFR under the risk-neutral measure:

where  $\sigma$  is the volatility of the IFR's increments (not returns). From the definition of the IFR it immediately follows that:

and that

is a martingale under  $\mathbb{P}$ . Further, we prove below that:

is a standard Brownian motion under  $\mathbb{P}$ , and the HJM formula immediately follows.

Pricing a  $\mathcal{F}_t$ -measurable cash-flow  $\mathbf{C}_t$  under  $\mathbb{P}$ :

and comparing with the risk-neutral measure:

implies that for any  $\mathcal{F}_t$ -measurable variable  $x_t$ ,

It follows that the Radon-Nikodym derivative of  $\mathbb{P}$  to  $\mathbb{Q}$ , conditional to  $\mathcal{F}_t$ , is:

where:

is an exponential martingale under  $\mathbb{P}$  with volatility, the volatility of the zero-coupon bond of maturity  $T$ .<sup>36</sup> Therefore, by Girsanov's theorem:

is a standard Brownian motion under  $\mathbb{Q}$ , which completes the demonstration.

When  $\mu$  is deterministic, the distributions of the IFRs are Gaussian and those of the discounts are log-normal. This model, despite being subject to a single Brownian motion, is generally *non-Markov*. This means that there exists no deterministic relationship between

and  $\mathbb{P}$  or between  $\mathbb{Q}$  and  $\mathbb{P}$  when

. The Markov dimension of the model is infinite: its state at time  $t$  is the whole discount or forward curve and it cannot be reduced without loss of information.

Infinite dimension does not lend itself to an effective practical implementation. Therefore, HJM remains to date the theoretical reference for stochastic interest rate modeling, but what is effectively implemented in modern systems, for the risk management of interest rate exotics, are variations and extensions purposely developed for algorithmic tractability. The most successful variations, developed shortly after HJM and still in activity today with substantial upgrades, are the Libor Market Model (LMM) of Brace, Gatarek, and Musiela ([\[29\]](#))

and Markov versions of HJM, independently established by many researchers in the mid-1990s, most prominently Cheyette [55].

**Libor Market Models** Libor Market Models partition the rate curve up to a final horizon into a finite number of forward Libors of duration : . The original paper postulates a log-normal dynamic a la Black-Scholes with volatility and deduces an HJM-like formula for the risk-neutral dynamics of the forward Libors (which we don't reproduce here, referring readers to the original paper or the dedicated chapter of [6]).

LMM are also generally non-Markov, although their dimension is finite. With horizon 50 years and 3m Libors, its dimension is 200, which is certainly better than infinite, but too elevated for performance.

LMM remain popular nonetheless, partly because they directly model market traded Libors, as opposed to theoretical, shadow variables like discount factors or IFR, and partly because the large number of free parameters, the collection of curves , allows them to correctly calibrate to interest rate option markets, even in their one-factor form.

Importantly, the original paper also delivered a closed-form formula for caps and swaptions in the LMM framework.<sup>38</sup> At the heart of this exercise was the determination of the volatility of swap rates as a function of the volatility and correlation of forward Libors. It later turned out that this new way of reasoning about the dynamics of interest rates was very useful in itself, with applications beyond the implementation of LMM. For instance, closed-form approximations for European options could be found for the general HJM model, something that had not been done before. The mapping of Libor to swap rate dynamics also helped with the structuring and risk management of interest rate exotics during the 1998–2008 decade when rate exotic businesses were growing exponentially in size and complexity.

**Markov specification of HJM** LMM resolved the infinite dimension problem by implementing what is essentially a discrete version of HJM. Another, diametrically opposite approach is to specify HJM in a way to lower its Markov dimension. In its one-factor Gaussian form presented here, HJM is Markov in dimension one if (and only if) for any couple of maturities  $t_1, t_2$ , there exists a deterministic function  $\phi$  such that:

When this is the case, all the forward rates at time  $t_1$  are deduced from one another, so when we know one, for example, the short rate

$\phi(t_1)$ , we know them all, and it follows that the state of the model at time  $t_1$  is contained in one scalar variable, for example  $\phi(t_1)$ , so the Markov dimension of the model is, by definition, one. This would allow to model interest rate curves in the same way we model asset prices, with one-dimensional dynamics, and apply the efficient algorithms designed for these, FDM whenever possible, low-dimension Monte-Carlo simulations otherwise.

Although this is somewhat outside of our scope, we demonstrate here how this is achieved, because this is an interesting exercise, and because it sheds light on *why* HJM was infinite dimensional to start with, even in its one-factor form.

Integrating the HJM formula (and dropping the  $t_1$  subscript on the IFR volatility) we find:

The left-hand side is a deterministic, known risk-neutral expectation and does not matter for our purpose. The right-hand side is the unknown, stochastic part of the IFR, a centred Gaussian variable. The stochastic parts of two IFRs of maturities  $t_1$  and  $t_2$  are:

and we see why the model is non-Markov: the random increments affect the IFRs with different weights and , and it is, in general, impossible to retro-engineer their aggregated impact on one out of their aggregated impact on the other . The model is Markov if and only if one is somehow a deterministic function of the other. Since both are centered Gaussian variables, if such function exists, it has to be linear:

This must be true for all times , so we can differentiate the equation in :

For this equation to be satisfied irrespective of the Brownian path, the derivative in of the right-hand side's left term must be zero, so is not a function of after all, and the equation simplifies into:

which implies that for any , and for any triplet :

It follows that there must exist a function such that

in particular:

and denoting , we find that the model is Markov if and only if:

This requirement restricts the specification of HJM to a so-called *separable* volatility specification, because  $\sigma_{t,x}$  is the product of a function of only  $t$  and a function of only  $x$ :

. It is also clear that the dimension of the free parameters in HJM is reduced by an order of magnitude from a two-dimensional surface  $\sigma_{t,x}$  to two curves  $\sigma_t$  and  $\sigma_x$ . The resulting model is orders of magnitude more tractable in a practical implementation, but its ability to calibrate to the market prices of interest rate options is limited due to the small number of free parameters.<sup>39</sup>

In addition, going back to the integral form of the IFR and setting  $\sigma_x = 0$ , we find the integral form of the short rate:

Injecting the Markov volatility form, differentiating, and rearranging, we quickly find the dynamics of the short rate (we make  $\sigma_t$  a constant here to simplify the equations):

where  $r_t$  is deterministic. The short rate follows a *mean-reverting* process a la Vasicek (1977) [56], so one of this model's many names is "extended Vasicek." It is also widely known as Hull and White's interest rate model [57], although, like Vasicek whose work precedes HJM by 15 years, Hull and White postulated a Markov dynamic for the short rate and worked their way up to the dynamics of the curve, not the other way around. We shall call this model by another of its popular names, Linear Gauss Markov (LGM). Its parameters are  $\sigma_r$ , the volatility of the first differences (not returns) of the short rate, and its mean-reversion  $\theta$  (even though we kept  $\sigma_t$  constant to simplify the presentation, it doesn't have to be).

It is more convenient to work with the variable  $\tilde{S}_t$ , which dynamics is the same as  $S_t$  but drops the reference to today's forwards in the drift:

where  $\tilde{S}_t = S_t e^{-\int_0^t \theta_s ds}$ . The following results may be found with simple (if somewhat painful) calculus.

1. All forward rates at time  $t$  are a deterministic function of  $\tilde{S}_t$ . This is the whole point of the model, of course, but it so happens that this function is explicit:

2. It follows that all discount factors at time  $t$  are also a deterministic function of  $\tilde{S}_t$ :

3. Finally, forward Libors at time  $t$  are a deterministic function of  $\tilde{S}_t$ :

This formula correctly incorporates Libor basis, but keeps it constant.

These formulas, called *reconstruction formulas*, are a fundamental part of the model, because they map a path of its state variable  $\tilde{S}_t$  to a scenario of market samples  $(\tilde{S}_t)$ . Contrary to the other models we introduced so far, the LGM dynamics is not directly written on market variables, so the scenario is not modeled directly, but implicitly, as its samples are deduced from  $\tilde{S}_t$  with the reconstruction formulas.

Remember that the purpose of this model, like any model, is to specify the joint distribution of the market observations  $(\tilde{S}_t)$  in the scenario.

In an interest rate context, these observations are the values on the event dates of forward discounts and Libors, as well as the numeraire on payment dates. What LGM models is its internal state variable  $\theta$ , which joint distribution on event dates is a multidimensional Gaussian distribution, with known expectations and a known covariance matrix (although we are not calculating them here). The discounts and Libors on the event date  $t$  are modeled implicitly as known functions of the state variable  $\theta_t$  on the same event date. But this is not the case for the risk-neutral numeraire  $N_t$ . From its definition:

it appears clearly that  $N_t$  is a functional of the continuous path of over  $\theta_s$  but not of  $\theta_t$ . Of more concern is that  $N_t$  is not a function of a *discrete path of*  $\theta_s$ . It would follow that, in order to fulfill its purpose and specify the distribution of the scenario, the model would need to track the numeraire, in addition to its state variable, on event dates, somewhat defeating the benefit of the one-dimensional Markov specification, and damaging the speed and accuracy of algorithms. This comment is not limited to LGM and concerns all interest rate models. The risk-neutral measure is generally *not* the wisest choice for their implementation.

The solution is to change the numeraire, of course, and use one numeraire  $N_t^*$  where  $N_t^*$  is a known function of  $\theta_t$ . We have seen that this is the case for zero-coupon bonds. In order to use the same numeraire for all cash-flows, we must use a bond of a longer maturity than the financial product we are working with. The preferred maturity is generally the payment date of the last coupon, and the related probability is known as the *terminal measure*.

When we demonstrated the Heath-Jarrow-Morton formula page 169, we proved that:

is a standard Brownian motion under  $\mathbb{P}$ . In LGM:

hence, the dynamics of the state variable  $S_t$  under the terminal measure is:

the rest of the model, in particular, the reconstruction formulas, being unmodified by the change of measure.

Finally, an exact closed-form solution for coupon bond options was derived in this model by El Karoui and Rochet in 1989 [58]. The formula can be applied to price a swaption by turning the underlying swap into a coupon bond, although the exercise is less trivial with Libor basis. It also applies to caplets and floorlets, considered as single-period swaptions, hence to caps and floors.

LGM is a fundamental interest rate model, and, to a large extent, the Black (and Scholes) of the interest rate world. Like Black, LGM is simple, Gaussian, and analytic. Like Black, LGM can be implemented in very fast algorithms, thanks to its one-factor Markov dimension and Gaussian transition probabilities. And like Black, LGM was one of the first operational models in its field, and provided a solid, if somewhat limited, basis for trading the first interest rate exotics.

The fundamental result, that HJM is Markov if and only if its volatility structure is separable, was independently found by multiple researchers in the mid-1990s, including Cheyette in 1992 [55] and Ritchken and Sankarasubramanian in 1995 [59]. Cheyette, in particular, derived the result in a very general context, not only in a multifac-

tor framework, but also with a non-deterministic volatility. Cheyette demonstrated that reconstruction formulas hold with local and even stochastic volatility, as long as it remains separable, even though any other specification than Gaussian increases the Markov dimension of the model.

**Modern implementation of interest rate models** Research in interest rate models peaked in the decade 1998–2008 to cope with the rapidly growing size and complexity of rate exotic markets, the main focus being stochastic volatility and multi-factor extensions. LMM and Cheyette<sup>40</sup> roughly shared the scene and were both upgraded with multiple factors and stochastic volatility.

Cheyette implementations are more efficient due to a low Markov dimension, LMM being more flexible and better calibrated, although the calibration gap significantly narrows when the number of factors in Cheyette is increased.

Research on stochastic rate models continued after 2008–2011, although somewhat less actively, mainly with the incorporation of stochastic Libor basis and collateral discounting, and a focus on a fast implementation for the needs of xVA and regulations, therefore favoring the Cheyette family. We refer to [7] for an application of MFC to xVA.

## A general formulation of dynamic models

The purpose of a model, dynamic or not, is to specify the distribution of the scenario  $\Omega$ , a collection of market observations on a collection of event dates, which definition is independent of the model.

How models fulfill this purpose is down to their internal mechanics and depends on the specific model. Dynamic models are those that determine the distribution of the scenario, implicitly, with the specification of a stochastic dynamics. Some models, like Black and Scholes, directly specify the dynamics of market variables. But this is not always the case. Other models, like LGM, specify the dynamics of some inter-

nal state variables, together with a mapping from state to market, which we called reconstruction formulas.

More generally, dynamic models define the stochastic dynamics of an internal state and specify a mapping from state to market variables. A model is the sum of the dynamics and the mapping. Both may be subject to parameters, like Black and Scholes's volatility, Dupire's bi-dimensional local volatility surface, or LGM's time-dependent volatility and mean-reversion.

More formally, a dynamic model defines by specifying the dynamics of an internal state vector and a mapping:

from a path of to the market samples over a scenario's timeline, including the numeraire on payment dates. Note that we have a mapping from to , not one from to for every event date . This would be generally desirable, but not always possible. For example, we have seen that the risk-neutral numeraire in LGM is a functional of but not a function of . Particular models may have time-wise mappings, but in general, it is a path-wise mapping. This is a detail, but one that affects the architecture of simulation libraries.

We restrict the presentation to diffusive models, in which the state vector follows, under , a multidimensional diffusion of the type:

The size of the state vector is called the *Markov dimension* of the model. It may be as low as one, but we also introduced models where it may be large (LMM), or infinite and even continuous (HJM). We carefully distinguish the Markov dimension: the size of the state vector ; and the number of factors: the size of the Brownian Motion . It is always the case that , and, in many cases, is vastly superior to .

The drift  $\mu$  is a vector-valued function of the state vector and time, which components are the risk-neutral drifts of the components of (called *state variables*) each as a function of the entire state vector  $x$ . The volatility matrix  $\Sigma$  is an  $n \times n$  matrix that represents the (instantaneous) covariance of the state variables. Since  $\Sigma$  is a standard Brownian Motion of dimension  $n$ , its components are independent; hence, the covariance of  $\Sigma$  is  $\Sigma$ , which also means that  $\Sigma$  is a square root (in matrix terms) of the covariance matrix of  $x$ .

The specification of  $\mu$  and  $\Sigma$ , along with the parameters of the mapping  $\pi$ , constitute the parameter set of the model. The model parameters are either set in reference to initial values (today's spot and forward prices, discounts, or rates), estimated (correlation) or calibrated (typically, volatilities) to match the market prices of a number of active (generally, European) instruments. Derivatives of product values to model parameters are the risk sensitivities, which effective production is a topic of [Part III](#).

Finally, the combination of the dynamics of  $x$  and the mapping  $\pi$  must guarantee that ratios of asset prices to the numeraire are martingales, and that initial values correspond to today's market prices. This guarantees that the model is arbitrage-free and respects today's market. Practically, it also ensures that the model correctly prices linear instruments. When, in addition, the model is correctly calibrated, it also correctly prices European instruments. Dynamic models also implicitly specify the joint distribution of the scenario, allowing them to price exotics, including hybrids and xVA, provided that the model correctly specifies the joint dynamics of all the relevant market variables.

## Finite Difference Methods

We have seen that, with the exception of a small number of particularly simple cases, dynamic models cannot price exotics analytically. Numerical methods must be invoked. There exist essentially two families of generic numerical methods for pricing with dynamic models: fi-

nite difference methods (FDM) and Monte-Carlo simulations. Other methods exist, but they are generally niche and only work in specific contexts like Fourier transforms, or they are inefficient, like binomial or trinomial trees, which can be seen as less stable and less efficient versions of FDM.

We briefly introduce FDM now, and explore Monte-Carlo in detail in the next chapters. We shall not discuss FDM further thereafter, and refer to the dedicated chapters in Andersen and Piterbarg's book [6] or Andreasen's lecture notes [60].

For the purpose of this brief introduction to FDM, we disregard interest rates and work under the risk-neutral measure, where, in this case, all the discount factors, and the numeraire, are constant equal to one, and cash-flows and payoffs are the same thing. Let us start with a European transaction with a single cash-flow of maturity  $T$  with a one-factor, one-state-variable dynamic model, with a time-wise mapping  $\phi$ :

From the Markov property of  $\phi$ , we know that  $\phi(T, \cdot)$ . From Feynman-Kac's theorem, we know that  $\phi(T, \cdot)$  is the solution of the partial differential equation (PDE):

(where the subscripts denote partial derivatives) with the boundary condition:

(the value at maturity is the payoff).

The solution of this PDE may be approached as follows: consider a grid of  $t_i$ , say evenly spaced times  $0 = t_0 < t_1 < \dots < t_N = T$  and  $N+1$  evenly spaced  $s_j$ 's

ranging from, say, the expectation minus five standard deviations at maturity to the expectation plus five standard deviations at maturity.<sup>41</sup>

Now we have a grid where each node represents a scenario where . We are going to populate this grid with the corresponding future values of the transaction. First, we apply the boundary condition, which populates . Then, we approximate the PDE with finite differences over the nodes in the grid in place of the continuous derivatives, which permits to propagate values backwards in time and populate the grid entirely. The node (where is the position of in the grid) is today's value.

Unfortunately, it does not work that easily. We left out the details of the back-propagation process. What do we do exactly once we know the values at to find the values at ? We could compute the derivatives in over the known grid at , and apply time differences to produce the grid at . This is known as an *explicit* scheme and is notoriously unstable. An *implicit* scheme must be applied instead, where finite differences are computed over the *unknown* grid at . This boils down to a tri-diagonal system of equations, something that is solvable efficiently in linear time. *Numerical Recipes* [20] provides the code of an efficient routine for this purpose. Even better, we can populate the grid at with a mix (say, half–half) of the explicit and the implicit solutions, something known as the Crank–Nicolson (CN) scheme. CN not only guarantees the stability of the numerical solution (as long as it is at least 50% implicit), but also its convergence with order .

The complexity of FDM is linear in both the size of the grid and the size of the time grid. Linear complexity combined with quadratic convergence makes FDM *the most efficient valuation method ever invented*.

Of course, FDM is not only adequate for European profiles. The back-propagation equation is independent of the transaction. The cash flows only intervene through boundary conditions. We can incorporate many types of cash-flows by manipulating boundaries and without modification of back-propagation. For instance:

- We can have multiple European cash flows by applying the boundary:

(where we used the C style “+” notation) *after* we solved the grid for .

- We can implement a barrier with the boundary condition that is 0 above or below the barrier. The FDM will even correctly value a continuous barrier with a minor modification of the back-propagation scheme. For this reason, FDM is the method of choice for barrier options.
- Since we are populating the different scenarios on the grid with the future values of the transaction, we can easily incorporate early exercise at time with payoff with the boundary:

again, to be applied after we solved the grid for . We can also determine exercise regions:

and exercise boundaries. FDM can even correctly propagate the exercise between consecutive time steps as a continuous exercise; therefore it is adequate (actually, it is the method of choice) for American options.

What we *cannot* easily incorporate is general path dependence that cannot be interpreted as a boundary. For instance, the payoff of an Asian option (option on average) cannot be interpreted as a boundary. Because of that path-dependence, the future price of an Asian option is not a function of time and spot only, but also of the running average up to that time. Hence, the Markov dimension of the PDE is two in this case, although the Markov dimension of the model is still one. In gen-

eral, FDM *can* handle path-dependency, but at the cost of an increase in dimension, something that FDM does not handle that well.

The major challenge with FDM is that its complexity grows exponentially with dimension for a fixed convergence. As the Markov dimension of the PDE grows to  $\dots$ , the number of nodes to process grows to  $\dots$  to maintain a numerical error in  $\dots$ . If we have 100 steps in each direction, we have 10,000 nodes to process in dimension 1. A modern computer does that in a fraction of a millisecond. In dimension 2, we have 1M nodes, something a modern computer also handles easily. But we are no longer talking milliseconds. In dimension 3, this is 100M nodes. Now we are talking seconds. In dimension 4, it is 10BN nodes we have to process for a single valuation. We may be talking talking minutes. In addition to an exponentially growing complexity, dimension more than one creates difficulties in the treatment of the cross-derivatives, and the solutions (known as Craig-Sneyd's scheme and friends) further increase complexity.

We will see that Monte-Carlo simulations, which have a much lower convergence order of  $\dots$  in the number of simulations and a much higher complexity per scenario, still win because of their linear complexity in the Markov dimension of the model (not of the PDE, which dimension is at least equal to that of the model). FDM is not practical in dimension 4 or higher. That pretty much disqualifies it in the context of general path-dependent transactions and many models of practical relevance. Cheyette's Markov interest rate model with two factors, for example, is of dimension two with a deterministic volatility, but five otherwise. Hence, FDM is out for multi-factor interest rate models, too. More generally, the growing complexity of the exotic transactions in the decade 1998–2008, causing models to grow in sophistication to follow-up, increased the Markov dimension in most pricing contexts, disqualifying FDM and leaving the much slower and less accurate Monte-Carlo as the only solution. This is without even mentioning the hybrid models used for regulatory calculations, the dimension of which routinely stands in the hundreds.

One specific area in modern financial markets where FDM remains relevant is the so-called *first-generation foreign exchange exotics*. These are short-term (typically less than 6 months) barrier options (all sort of barriers, one touch, double no touch, knock-in, reverse knock-out, and so forth) written on currency exchange rates. These exotics trade in high volumes on specialized OTC<sup>42</sup> markets. The short maturity permits to neglect the impact of stochastic rates and keep dimension low, and FDM is ideal for barriers, as mentioned earlier.

Lipton's masterclass [61] provides an overview of forex exotics and analyses models with local volatility, stochastic volatility, and jumps in relation to those markets. One leading model is the so-called SLV (stochastic and local volatility) model defined by:

with Heston's specification for the stochastic volatility, although other specifications are possible; see [51]. This is a two-factor, two-state-variable model suitable for FDM. The model is calibrated with Dupire's second formula (see [50]) implemented in the form of a forward FDM (see [44]), and it values barriers with backward FDM. Pricing takes a few        s of a second, remarkably for a model of such sophistication.

It is, however, Monte-Carlo that accounts for the vast majority of valuation contexts of practical relevance in modern finance. But this is only because Monte-Carlo is the only viable method in such high dimension. Researchers settled for Monte-Carlo despite its slow convergence and CPU-demanding operation. Monte-Carlo is a *fallback* in a context where the superior FDM is no longer relevant.

For this reason, a computationally efficient implementation of Monte-Carlo simulations, including parallelism, and the differentiation of re-

sults, is a critical objective in modern finance. To give an idea of the time scales involved, it takes around one second to compute the full risk against 1,600 local volatilities of a barrier option in Dupire's local volatility model, with FDM, 100 time steps, 150 spot steps, by bumping, on a single modern CPU. To get a similar accuracy with Monte-Carlo takes at least 100,000 simulations over 200 time steps, and the full bump risk on a single core takes about half an hour, around 2,000 times slower than FDM. With parallelism and AAD, FDM risk time is reduced under  $\frac{1}{100}$  of a second (on a quad core laptop). In relative terms, this is a fantastic 100+ speed-up. In absolute terms, 1 second or  $\frac{1}{100}$  of a second does not make much difference. With Monte-Carlo, parallelism and AAD reduce the time spent on the risk report *from half an hour to half a second* (on a quad core laptop or a quarter of a second on an octo-core workstation), without loss of accuracy. The difference this makes for investment banks and their trading desks is massive. This is the difference between a real-time risk report on a trader's workstation or an overnight computation in a data center.

The purpose of the techniques described in the rest of this publication is to bring the performance of FDM to Monte-Carlo simulations.

## NOTES

---

<sup>1</sup> The dates  $\tau_{i-1}$  and  $\tau_i$  are the start and end dates in the  $i$ th period in the floating schedule, and we ignore conventional day counts and fixing lags.

---

<sup>2</sup> We should call it “instantaneous forward *discount rate*” to avoid confusion with Libor or OIS rates, which generally trade with a significant difference called *basis* to discount rates.

---

<sup>3</sup> Ignoring conventional day counts and fixing lags.

---

<sup>4</sup> “Old school” interest rate modeling uses the “identity”:

to “deduce” Libors from discount factors. It would follow that Libors are not additional market primitives, but aggregates of discount factors. However, this formula ignores discount basis and assumes Libor discounting, something completely wrong in modern markets. Since the explosion of basis adjustments, the advent of collateral discounting, and the artificial stabilization of interest rates around zero in 2008–2011, Libor basis (roughly, the difference of interest rates picked on the Libor curve or on the discount curve) largely exceeds interest rates, both in value and daily moves, and accounts for a major fraction of interest rate risk. This is why Libors and discounts are different market primitives. In addition, we have multiple Libor indices per currency, with large, fast-moving bases between indices, so a complete definition of a forward Libor at time  $t$  includes its index:  
, although we often drop the index here to simplify notations.

---

**5** A swaption is an option on a combination of forward Libors, an exotic product known as a “basket option.” The Libor Market Model of [29] builds on this observation to develop an approximate closed-form valuation formula for swaptions as a function of the volatilities and correlations of forward Libors.

---

**6** Options on average, called “Asian options,” also often illustrate the notion of path-dependent products, although they rarely trade in practice, contrary to barriers.

---

**7** In general, for example when the callable product has path-dependent cash-flows, its value depends not only on the current market, but also on the past:

In what follows, we only consider the simple case. Path-dependency does not materially modify the reasoning.

---

**8** To model credit this way is more general than with a default time  $\tau$ , in which case .

---

**9** And most other regulations.

10 Many banks use rough closed-form approximations to estimate the future values of transactions for counterparty credit adjustment (CCR). Analytic approximations are other types of proxies, although regression proxies are vastly superior in accuracy when correctly implemented; see [31].

---

11 These notions will be further clarified with concrete code in [Chapter 6](#).

---

12 A process where is -measurable is said to be *adapted*. Self-financing strategies trade adapted amounts. A strategy that trades non-adapted amounts is said to be *forward looking* and does not qualify as a self-financing strategy.

---

13 In theory, in continuous time, and provided that the model parameters are accurate.

---

14 Which essentially means that the same events have zero probability under the two measures.

---

15 In proportional terms, meaning that

where is a standard Brownian motion under .

---

16 Martingale measures are unchanged when numeraires are scaled, so numeraires can always be normalized by their initial value and

---

17 This is the formula for a discrete set of cash-flows; the extension to a continuous set is trivial and not necessary for what follows.

---

18 We will not extend the definition of scenarios any further.

---

[19](#) The model decides what the numeraire is. The product determines when the numeraire is needed and how it intervenes in the payoff. This will be clarified in code.

---

[20](#) This separation between models and products is somewhat muddled in the context of callable products and regulatory calculations, although it is reestablished with proxies.

---

[21](#) Technically, this is not a sufficient condition: linear products are those whose time value today *and in the future* is always zero. Exceptions are rare but they exist, and are sometimes asked in interviews.

---

[22](#) As long as they respect today's market.

---

[23](#) Put prices are given by the *parity*

which is an immediate consequence of where we note that European models must rely on linear models for discounts and forwards to respect today's market.

---

[24](#) In the very rare cases where no such information is available, dynamic models (see next) may be used instead.

---

[25](#) To obtain an arbitrage-free IVS from the interpolation of a limited number of traded option prices is a challenging, specialized exercise outside of our scope. Gatheral's SVI [\[40\]](#) and Andreasen and Huge's LVI [\[41\]](#) are the current best practice solutions. In addition, IVS often interpolate not Black and Scholes's volatility, but the dynamic parameters of more sophisticated models, like SABR [\[36\]](#) or Heston [\[42\]](#), see [\[43\]](#) for details.

---

[26](#) Generalized derivatives are not really needed to demonstrate this result; it can also be demonstrated by more traditional means. However, Schwartz's generalized derivatives allow to demonstrate many financial results, including Dupire's formula and extensions with stochastic volatility, in a natural, terse, and sharp manner. See [\[44\]](#) for a tutorial and applications in finance.

---

27 In particular, the variance of the distribution is related to the at-the-money (ATM) Black-Scholes implied volatility, its skewness is related to the slope of the implied volatility function of the strike around the money, and its kurtosis is related to the curvature (called “smile”) of the implied volatility curve.

---

28 Calibration is the process of tuning the dynamic parameters of the model so that the model matches the IVS prices of European options, hence, marginal distributions.

---

29 Although one-dimensional FDM prices Europeans, one by one or in bulk with the forward FDM technique explained in [44], with speed and accuracy that rivals closed-form formulas, and calibrations are further greatly accelerated with AAD; see [Part III](#).

---

30 Which is not such a drag in this model, the bank account being deterministic, but the method also applies in more sophisticated contexts where it may be beneficial.

---

31 This is a very common application of Euler's theorem to Black and Scholes, although it takes a shortcut as noticed by Rolf Poulsen in [http://web.math.ku.dk/~rolf/BS\\_clarification.pdf](http://web.math.ku.dk/~rolf/BS_clarification.pdf).

---

32 For good reason outside of our scope here; see [46].

---

33 SABR is a strong market practice for interest rate options, where caplets are priced under a forward measure and swaptions under the annuity measure as explained earlier.

---

34 Volatility is a free parameter in Black and Scholes, too, but in SABR, volatility is a process, not a parameter, which is the whole point of stochastic volatility models.

---

35 Which means we solve  $\quad \text{in} \quad .$

---

36 This is the price at  $t$  of the zero-coupon bond of maturity  $T$ , discounted by the numeraire, hence a martingale; the normalization by a constant  $\lambda$  does not affect this result.

---

37 The numeraire  $N_t$  is a finite variation process without a diffusion term; hence, the volatility of the ratio is the volatility of the numerator.

---

38 The formula is exact for caps, approximate but very precise for swaptions.

---

39 However this limitation dissipates in a multifactor context. When this model is written with multiple Brownian motions, the number of free parameters increases linearly with the number of factors. A typical modern implementation in production uses three to five factors, allowing for a very decent calibration to market prices. The Markov dimension increases to the number of factors, but remains low. The main benefit of a multifactor model, however, is to produce a realistic correlation between rates of different maturities: in a one-factor model, with the same Brownian motion driving the entire curve, the correlation between rates is necessarily 100%.

---

40 With the addition of local and/or stochastic volatility, the model is no longer Gaussian, although still linear and Markov, and is no longer called LGM, but Cheyette. Multifactor versions are often called multifactor Cheyette or MFC. A practical, efficient MFC implementation was published in 2005 by Andreasen [8].

---

41 In general, we do not have the same number of times and spots, but we go with that here in the interest of simplicity; in addition, the expectation and standard deviation of  $N_t$  may need to be estimated from the model parameters beforehand.

---

42 Over the counter, or directly between banks and institutional clients, as opposed to traded on an exchange.

---

