PRE-PRINT
UD 2014-02

The Geometry of Interest Rate Risk

Michele Maio*
and
Jorrit-Jaap de Jong†

uGly Duckling B.V.

Burgemeester le Fevre de Montignylaan 30,
3055LG Rotterdam, the Netherlands

June 2, 2015

Abstract

In this paper we consider the process of interest rate risk management. The yield curve construction is revisited and emphasis is given to aspects such as input instruments, bootstrap and interpolation. For various financial products we present new formulas that are crucial to define sensitivities to changes in the instruments and/or in the curve rates. Such sensitivities are exploited for hedging purposes. We construct the risk space, which eventually turns out to be a curve property, and show how to hedge any product or any portfolio of products in terms of the original curve instruments.

Keywords: Yield curve, hedging, interest rate risk management.

*`m.maio@uglyduckling.nl`
†`j.jong@uglyduckling.nl`

Contents

1 Introduction

In this paper we will describe a framework for managing interest rate risk of linear fixed-income products. The aim is to find a way to hedge interest rate risk. In order to achieve that, we reconsider well-known practices from the business side and put them in a wider more mathematical framework. This is done by looking at the present value of one basis point (PV01) or equivalently by the instrument value of one basis point (IV01). These quantities allow us to define hedging strategies and work out the positions vector that one should take to hedge the risk of small movements in the interest rates of market instruments.

Many quantities of interest require the knowledge of the term structure of interest rates at values that are different from the input data, namely the yield curve is needed. Examples of such quantities are the discount factors for generic maturities or any product's PV01 which requires the derivatives of the yield curve with respect to its zero nodes.

Two approaches are normally used to solve this issue: interpolation and parametric fitting methods. When using interpolation one is guaranteed that the curve will pass through all the input nodes by construction. This is not in general the case for parametric methods, where typically one is given a curve that depends on a number of parameters whose values must be chosen appropriately. This is what happens for example in the Nelson-Siegel-Svensson models [1, 2] where the parameters are fixed by minimizing the sum of the squares of the distances between the observed input nodes and their corresponding values expected from the curve. Other parametric models are those developed by Wiseman [3, 4] by Bjork and Christensen [5], by James and Webber [6], and extensions thereof. In this paper we will use interpolation only.

The yield curve will turn out to be crucial in constructing the hedging positions for any product. In fact, the whole risk space will be determined by the properties of the yield curve. On one side, those positions clearly must depend on the details of the product that we want to hedge, and on the other side they are fixed by the yield curve. Moreover, they are independent on the particular hedging strategy. A strategy based on the curve PV01 will generally give the same positions as a strategy based on the curve IV01. This will be easy to see once we realize that PV01 and IV01 are just two different representations of the same concept but in two different basis, with a Jacobian matrix relating them. Intuitively, this behavior is the counterpart of the bootstrap procedure which relates instruments rates to zero rates.

To construct the framework three steps are needed.
First, we construct a yield curve. The various methods of curve construction all share a bootstrap approach, but vary in interpolation technique. These techniques are not novel and an overview of the most common approaches can be found in Bolder and Stréliski's report [7]. A recent contribution to the literature is the method described by Hagan and West [8, 9]. These authors also give a list of good features that a yield curve should have, that we can summarize with the following criteria: i) the curve should price back the market (or at least deviations from market prices should be small), so that market data are recovered; ii) forward rates should be continuous in order to avoid arbitrage, but also positive and stable in order to price instruments correctly; iii) the curve should be local, so that a small change in one of the input nodes modifies the curve values only in the neighborhood of that node and not far away; iv) the hedges should be local, in such a way that if one of the curve instruments is hedged then the hedge is assigned to that instrument only.
Second, given the yield curve, we can compute the value of any linear fixed-income product, namely a product that depends only on rates[1]. Using the present value formula as a sum of discounted cash flows, we can compute the sensitivity to changes in the zero rates of the curve (PV01) or to the instrument rates (IV01). These quantities are not completely independent of each other and will be crucial for hedging. We will see that they are related by a change-of-basis transformation and that they are the starting point to determine hedging positions. Consistently, the hedging positions that can be determined for any product will be the same, independently from whether one uses the product's PV01 or IV01. Ultimately this happens because for any product the risk space is determined by the yield curve, or more fundamentally by the instruments used to bootstrap the yield curve. Our approach to compute PV01 and IV01 is fully analytic: the relevant derivatives are computed exactly, however for the PV01 and IV01 we will use a first-order approximation. This is a perturbative approach.
Thirdly, and finally, we can use these sensitivities to hedge (or replicate) any position.

[1] This would is not the case for the so-called non linear products, which depend on rates, volatilities and in some cases other quantities such as correlation.

For example, as referred to by Whittaker [10], the interest rate risk of a swap portfolio can be managed in this way. To check whether the hedge is local, the preferred way is to use exactly the set $\mathcal{I}$ as hedging instruments. Then, for any curve instrument we can compute the change in the present value due to a small change in any of the curve rates. We should find that this change in the instrument present value is zero for those changes in the curve rates that do not involve the rate of that particular instrument. Equivalently, changes in a given curve rate only triggers changes in the present value of the instrument used to construct that curve rate.

As pointed out by Reitano [11], there are problems with this approach, mainly because the replication is only a first-order approximation. Hence, all higher-order factors are ignored. In practice this means that instruments with high-convexity will not be properly hedged. This is indeed true, however when a portfolio is closely monitored and the hedge is updated frequently, this first-order framework can be used.

Since the credit crunch crisis, it is common practice to use the multi-curve approach for the valuation of fixed-income products. This approach is described by Bianchetti [12] (see also [13, 14]) and is considered the market standard to accurately include basis risk. In this paper the multi-curve framework is not considered, mainly for simplicity's reasons, however this framework can be extended to include multi-curve and we hope to do that in the future.

The plan of this paper is as follows.

In Section 2 we define the set up and fix our notation. In particular, we describe the role of bootstrap and interpolation for the yield curve construction, and we define the PV01 and IV01. Section 3 explains how the zero rates, that are the node points of the yield curve, are derived from various instruments' rates. In particular we will compute the zero rates related to cash and swaps. We will also compute the Jacobian matrix of derivatives to change from zero to instrument rates.

Section 4 contains a summary of the various interpolation methods that we have used to check our approach to measure risk in this paper, namely linear, cubic splines (Bessel-Hermite and monotone-preserving), and forward monotone convex spline. The methods that we use here are all partially treated in the Hagan-West papers [8, 9].

In Section 5 we describe how the PV01 and the IV01 are used for hedging purposes.

In section 6 we present a helicopter view of the quantities introduced in this paper, how they are linked to each other and how they are used for risk management purposes.

Finally, in Section 7 we summarize our results, briefly discuss open issues, present possible directions for future investigation, and throw our conclusions.

2 Set up

In this section we will explain our set up and fix our notation. Let us start by recalling what the term structure of interest rates, or simply yield curve, is. A yield curve is constructed out of a discrete set of instruments $\mathcal{I}$ which allows us to derive another discrete set $\mathcal{N}$ of input nodes for the curve. The set $\mathcal{I}$ will contain several instruments of various maturities, e.g.[2]

$$\mathcal{I} = \{\text{cash}_{2\text{w}}, \text{cash}_{1\text{m}}, \text{cash}_{3\text{m}}, \text{cash}_{6\text{m}}, \text{swap}_{1\text{y}}, \text{swap}_{5\text{y}}, \text{swap}_{10\text{y}}, \text{swap}_{20\text{y}}\}\,. \tag{2.1}$$

[2] We will consider only cash and swap in this paper, but in principle any fixed-income instrument can be used. The choice of the instruments used to construct the curve will be important when we move the discussion to hedging.

Instrument rates and maturities are in one-to-one correspondence with the zero nodes of the yield curve. The exact relationship, that will be explained in section 3, allows us to construct the set of zero nodes:

$$\mathcal{N} = \{(t_1, r_1),\, (t_2, r_2),\, \ldots, (t_n, r_n)\}\,, \tag{2.2}$$

which are the curve inputs. In general, there exists a *bootstrap* algorithm $\mathcal{B}$ that allows us to compute the zero rate of an instrument given its price and its principal value in a way that is consistent with the zero rates of the other instruments with shorter maturities (see e.g. section 4.5 of Hull's book 9$^{\text{th}}$ ed. [15] for a review, or the Deaves-Parlar paper [16] for a generalized approach to bootstrap):

$$\mathcal{B} : \mathcal{I} \longrightarrow \mathcal{N} \tag{2.3}$$

A yield curve is a function γ that maps the discrete set of zero rates into the real numbers

$$\gamma \equiv \gamma_{\mathcal{N}}^{\text{int}} : \mathcal{N} \longrightarrow \mathbb{R}\,. \tag{2.4}$$

The suffix denotes that the curve construction is strongly dependent on interpolation. If one uses parametric fitting methods (e.g. Nelson-Siegel-Svensson) to construct the curve, then it will not necessarily pass through all the nodes in $\mathcal{N}$. However, if interpolation is used, then γ will go through all the node points by definition and consequently the market is priced back. Equivalently, if the market instruments are priced using the yield curve then the input prices are recovered. In any case, the curve can also be extrapolated outside the interval $[t_1, t_n]$ (e.g. linear extrapolation, or any other model-dependent prescription), even if the interpretation of the values that one obtains might be subject to discussions.

The yield curve is needed to compute the present value (PV) of any product via the discount factor. We use the continuously-compounded convention for interest rates throughout this paper

$$D(t) = e^{-r(t)\cdot t}\,, \tag{2.5}$$

but other choices are equally allowed. One such example is the the n-periodic compounding

$$D(t) = \left(\frac{1}{1+\frac{r}{n}}\right)^{-nt} \tag{2.6}$$

which converges to the continuous compounding in the $n \to \infty$ limit.

Typically, a product promises intermediate cash flows in the future during its lifetime. The present values of these cash flows are given by their discounted amounts. Hence, to each product we can associate a set of discount factors

$$\mathcal{D} = \{D_1, D_2, \ldots, D_m\}\,, \tag{2.7}$$

where $D_i = D(\tilde{t}_i)$ and $\tilde{t}_i$, $i = 1\ldots, m$, correspond to the time of the m future cash flows. Hence, to get the PV of a product we have to follow the steps as in the sequence (2.8):

$$\mathcal{I} \xrightarrow{\ \mathcal{B}\ } \mathcal{N} \xrightarrow{\ \gamma\ } \mathbb{R}^{\gamma} \xrightarrow{\ D\ } \mathcal{D} \xrightarrow{\ PV\ } \mathbb{R}^{PV}\,. \tag{2.8}$$

Here, $\mathbb{R}^{\gamma}$ is the range of the curve γ which is normally $\mathbb{R}^{+}$, $\mathbb{R}^{PV}$ denotes the range of the present value PV, typically $\mathbb{R}$, and the discount function D represents the convention chosen for discounting.

One could also wonder what happens if one of the input zero nodes is changed by a small amount, typically one basis point[3]. In this case we can use the one-variable Taylor series that approximates the present value with respect to changes in the interest rate:

$$PV(r+\delta r) = \sum_{n=0}^{\infty} \frac{1}{n!}\frac{d^n PV}{dr^n}(\delta r)^n = PV(r) + \frac{dPV}{dr}\delta r + \frac{1}{2}\frac{d^2 PV}{dr^2}(\delta r)^2 + \ldots \tag{2.9}$$

In this paper we will focus on the changes in the present value due to chances in the interest rate (changes in the term structure) for linear products, that is the value of a product without optionality. As a result the first order approximation will suffice, for which equation (2.9) reduces to:

$$\delta PV(r) = \frac{dPV}{dr}\delta r + O(\delta r^2) \tag{2.10}$$

where $\delta PV(r) = PV(r+\delta r) - PV(r)$. When $\delta r = 1$ basis point, then the l.h.s. gives back the standard PV01. Clearly, in the limit when $\delta r \to 0$, δPV vanishes as well, but the derivative on the r.h.s. does not.

Actually the *exact* counterpart of formula (2.10) is

$$\delta PV(r) \equiv \int_r^{r+\delta r} \frac{d}{dr'} PV(r')\, dr' \,. \tag{2.11}$$

From this we can see that if we know analytically the expression for the derivative of the present value with respect to the rate, then we can re-construct the PV01 *exactly* by integrating the derivative on the interval $[r, r+\delta r]$ with $\delta r = 1$bp.

We will compute the derivative $\frac{dPV}{dr}$, which depends on the definition of the financial product, the formula used for the discount factors, the interpolation method and the instruments used to construct the yield curve. In particular, we consider a yield curve constructed out of n market instruments. Each instrument will result in fixing a node point in the yield curve. It will be shown that the gradient vector ∇PV w.r.t. the node points spans the complete risk space for linear fixed-income products (and it is the best first-order approximation for more complex products).

Defining r_i as the *zero rate* at node point i, from eq. (2.10) we get:

$$\delta PV_i = (\nabla PV)_i \cdot \delta r_i \,, \tag{2.12}$$

where ∇PV is the gradient vector with components $(\nabla PV)_i = \frac{\partial PV}{\partial r_i}$, and the total risk for the product -given all the curve nodes- is:

$$\delta PV(r) = \nabla PV \cdot \Delta r = \sum_i \frac{\partial PV}{\partial r_i}\delta r_i \,. \tag{2.13}$$

Here Δr is the shift vector, $\Delta r = (\delta r_1, \delta r_2, \ldots, \delta r_n)$. If δr_i is one basis point and all the others are zero, i.e. $\delta r_j = 0.0001 \cdot \delta_{i,j}$, then each component in the sum gives the present value of one basis point, a.k.a. $PV01_i$, with respect to zero rate i:

$$PV01_i = \frac{\partial PV}{\partial r_i}\delta r_i \,. \tag{2.14}$$

[3] By definition of basis point, $1\text{bp} = 10^{-4}$.

When the PV01 is implemented numerically or with a first-order Taylor approximation, one should choose the bump $\delta r_i = 0.0001$. The exact counterpart of this is

$$PV01_i \equiv \int_{r_i}^{r_i+1bp} dr'_i \, \frac{\partial}{\partial r'_i} PV(r'_i)\,. \tag{2.15}$$

PV01 measure is a widely-used metric for interest rate risk. In addition, if all the rate changes in the vector Δr are set to 0.0001 at the same time, then the result is a parallel shift of the complete curve of one basis point. This is also often called DV01 (dollar value of one basis point) or confusingly PV01 (which conflicts with our earlier definition).

Similarly to the zero rate r_i, we can also consider the *instrument rate* x_i at node i, which is the rate specified by the instrument that is used to construct the corresponding zero rate in the yield curve, and define the $IV01_i$ (instrument value of one basis point) at i as

$$IV01_i = PV(r(x_i + 1bp)) - PV(r(x_i))\,. \tag{2.16}$$

In order to compute this, since $\delta x_1 = 1$bp is small enough to be in the linear regime, we can use first-order Taylor approximation as

$$IV01_i = \frac{\partial PV}{\partial x_i}\delta x_i\,. \tag{2.17}$$

However, there is an *exact* counterpart in this case too. By extension of the argument that led to the exact relations (2.11) and (2.15), *mutatis mutanda*, the same reasoning holds for the calculation of the IV01, with the only difference that now the relevant derivative is $\frac{\partial PV}{\partial x_i}$:

$$IV01_i \equiv \int_{r_i}^{r_i+1bp} dx'_i \, \frac{\partial}{\partial x'_i} PV(x'_i)\,. \tag{2.18}$$

The present value of a product is by definition the discounted sum of its future cash flows:

$$PV = \sum_{cf} A_{cf} D(t_{cf})\,. \tag{2.19}$$

Here, starting from $t = 0$ as today's date, the index cf runs over all the future cash flows, A_{cf} is the amount exchanged at time t_{cf} and $D(t_{cf}) = \exp\left(-r(t_{cf})t_{cf}\right)$ is the corresponding discount factor, where $r(t_{cf})$ is the interest rate obtained from the knowledge of the yield curve. In order to compute these derivatives, the best way is to use the chain rule as follows:

$$\frac{\partial PV}{\partial r_i} = \sum_{cf} \frac{\partial PV}{\partial D_{cf}} \frac{\partial D_{cf}}{\partial r} \frac{\partial r}{\partial r_i} \tag{2.20}$$

for the PV01$_i$, and

$$\frac{\partial PV}{\partial x_i} = \sum_j \frac{\partial PV}{\partial r_j} \frac{\partial r_j}{\partial x_i} = \sum_{cf} \frac{\partial PV}{\partial D_{cf}} \frac{\partial D_{cf}}{\partial r} \sum_j \frac{\partial r}{\partial r_j} \frac{\partial r_j}{\partial x_i} \tag{2.21}$$

for the IV01$_i$. Here $r(t) \equiv \gamma(t)$ is the yield curve. The key difference between PV01 and IV01 is the presence of the Jacobian $\frac{\partial r_j}{\partial x_i}$. We will see later that the Jacobian is in general not diagonal, but lower triangular matrix, as a consequence of the bootstrap procedure. Formulas (2.20) and (2.21) are the mirror of the sequence (2.8). In fact:

- $\frac{\partial r_j}{\partial x_i}$ represents the bootstrap on the set $\mathcal{I}$;
- $\frac{\partial r}{\partial r_j}$ represents interpolation from the set $\mathcal{N}$;
- $\frac{\partial D_{cf}}{\partial r}$ depends on the choice of discounting convention;
- $\frac{\partial PV}{\partial D_{cf}}$ is related to the product only via its PV.

Many quantities in the formulas above require the knowledge of the term structure of interest rates at values that are different from the input data, namely the yield curve is needed. Examples of such quantities are the discount factors for the cash flows $D(t_{cf})$ or the derivatives $\frac{\partial r}{\partial r_i}$. As discussed earlier, we will generate these values by using interpolation rather than parametric fitting methods.

3 Zero rates from cash and swap instruments

In this section we describe how to derive the zero rates that are then used to construct the yield curve via bootstrap and interpolation from various instrument rates.

A zero rate is the interest rate matured by a zero-coupon bond. Zero rates are constructed out of marked data, and in particular from instruments traded in the market. Here we will focus on the relation between zero rates and cash rates as well as swap rates[4].

Typically cash rates are used to construct the short-term structure of interest rates, while swap rates are responsible of the long-term yield curve.

Later we will move to hedging and we will need to compute the quantities PV01 and the IV01. These two quantities differ by the presence of the Jacobian matrix[5]

$$\mathcal{J} \equiv \mathcal{J}_x(r_1, r_2, \dots, r_n) = \frac{\partial(x_1, x_2, \dots, x_n)}{\partial(r_1, r_2, \dots, r_n)} \qquad \Longleftrightarrow \qquad \mathcal{J}_{kl} = \frac{\partial x_k}{\partial r_l} \tag{3.2}$$

between zero rates and instrument rates in the latter. Note that $\mathcal{J}$ is a curve property, since it depends only on the instruments used to construct the curve:

$$\mathcal{J} \equiv \mathcal{J}(\mathcal{I}) \,. \tag{3.3}$$

Actually, as already pointed out in (2.20) and (2.21), we will mostly need the inverse of such a matrix, namely $\mathcal{J}^{-1}$:

$$\mathcal{J}^{-1} = \frac{\partial(r_1, r_2, \dots, r_n)}{\partial(x_1, x_2, \dots, x_n)} \qquad \Longleftrightarrow \qquad \left(\mathcal{J}^{-1}\right)_{lk} = \frac{\partial r_l}{\partial x_k} \tag{3.4}$$

In order to compute (3.4) we will first compute the actual Jacobian (3.2) and then take its inverse, rather than directly computing the entries $\frac{\partial r_l}{\partial x_k}$. In fact, it turns out that this

[4]With the word swap we actually mean plain vanilla interest rate swap.

[5]Recall that the notation for the Jacobian matrix is:

$$\mathcal{J}_x(r_1, r_2, \dots, r_n) \equiv \frac{\partial(x_1, x_2, \dots, x_n)}{\partial(r_1, r_2, \dots, r_n)} = \begin{pmatrix} \frac{\partial x_1}{\partial r_1} & \frac{\partial x_1}{\partial r_2} & \cdots & \frac{\partial x_1}{\partial r_n} \\ \frac{\partial x_2}{\partial r_1} & \frac{\partial x_2}{\partial r_2} & \cdots & \frac{\partial x_2}{\partial r_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial x_n}{\partial r_1} & \frac{\partial x_n}{\partial r_2} & \cdots & \frac{\partial x_n}{\partial r_n} \end{pmatrix} . \tag{3.1}$$

approach is both easier and provides a natural way to deal with the first node. Moreover, it is pretty straightforward to derive formulas for the instrument rates x_l in terms of the zero rates $\{r_k\}$ and in addition, even though the actual values depend on the specific interpolation scheme, these formulas are interpolation-independent. We will also see that generically both $\mathcal{J}$ and $\mathcal{J}^{-1}$ are lower triangular matrices and that this is the case because the zero rate at position i is sensitive to changes in the previous instruments at positions up to i, as consequence of the bootstrap procedure.

The first node of the curve is special because of the *settlement delay*. Once entering a contract, there is always a settlement delay, that could in principle be zero, but in reality is non-zero[6]. As a consequence, almost any contract will be likely to start at a future date[7]. We will include this effect in our discussion and denote the real starting time by s ($s \geq 0$).

We will soon see that the Jacobian will be expressed in terms of derivatives of the type $\frac{\partial r(t)}{\partial r_i}$. This quantity strongly depends on the interpolation method used to compute the yield curve $r(t)$. In this section we will focus on the Jacobian, while in section 4 we will discuss interpolation.

3.1 Cash

Cash instruments are simply deposits that promise a pre-agreed interest over a predetermined time on an initial invested amount. Consider a set of cash instruments with different maturities τ (e.g. $\tau = 2$ weeks, $\tau = 1$ month, $\tau = 3$ months, $\tau = 6$ months, etc.). They all promise to pay back a simply-compounded interest at time $t = \tau + s$, where s is the settlement delay and t will be the location of the node used to construct the yield curve.

For cash, the relation between the zero rate at time t and cash rate with maturity τ is:

$$D(s) = (1 + x \cdot \tau) D(t) \,, \tag{3.5}$$

where x is the cash rate of the cash instrument with maturity τ, $\tau = t - s$, and $D(s)$ and $D(t)$ are the discount factors at times s and t[8]. Formula (3.5) can be inverted to give the cash rate as a function of the node zero rates:

$$x = \frac{1}{\tau} \left(\frac{D(s)}{D(t)} - 1 \right) . \tag{3.6}$$

Observe that $D(s)$ is typically function of the first node only and its value depends on the interpolation method, while $D(t) = D(\tau + s)$ is independent of the interpolation since $t = \tau + s$ defines a node point.

For more cash instruments, there will be more cash rates $x_1, x_2, \ldots, x_n$, with maturities respectively $0 \leq s \leq \tau_1 \leq \tau_2 \leq \cdots \leq \tau_n$. Before going to the general case, let us consider a simpler example. Let us take $s = 2$ days as settlement delay and $n = 3$ cash

[6] E.g. in the case of the Eurozone the settlement delay is two days.

[7] Exceptions to this case are e.g. all the overnight products, which start on the spot date.

[8] The relation $\tau = t - s$ is actually valid only in our idealized mathematical world where there is only one measure of times. In practice, conventions appear all over the places. The yield curve carries its own convention to measure times, while instruments follow other conventions for day counting, business days, and holidays, that are typically different from the curve one. In the real world, s and t are measured with the curve convention, while τ with the instrument convention. The relation $\tau = t - s$ is only approximately true.

instruments with tenors $\tau_1 = 2$ weeks, $\tau_2 = 1$ month, $\tau_3 = 3$ months. Formula (3.5) gives now explicitly:

$$D(s) = (1 + x_{2w} \cdot \tau_1)\, D(t_1) \tag{3.7a}$$

$$D(s) = (1 + x_{1m} \cdot \tau_2)\, D(t_2) \tag{3.7b}$$

$$D(s) = (1 + x_{3m} \cdot \tau_3)\, D(t_3) \tag{3.7c}$$

where $t_i = \tau_i + s$. Without doing any calculation, we can immediately see that

$$\frac{\partial x_{2w}}{\partial r_1} \neq 0 \qquad \frac{\partial x_{2w}}{\partial r_2} = 0 \qquad \frac{\partial x_{2w}}{\partial r_3} = 0 \tag{3.8a}$$

$$\frac{\partial x_{1m}}{\partial r_1} \neq 0 \qquad \frac{\partial x_{1m}}{\partial r_2} \neq 0 \qquad \frac{\partial x_{1m}}{\partial r_3} = 0 \tag{3.8b}$$

$$\frac{\partial x_{3m}}{\partial r_1} \neq 0 \qquad \frac{\partial x_{3m}}{\partial r_2} = 0 \qquad \frac{\partial x_{3m}}{\partial r_3} \neq 0\,. \tag{3.8c}$$

This structure is typical for cash instruments only, where some of the lower-diagonal entries are zero. The derivative with respect to the first node is non-zero due to the discount factor $D(s)$, which is usually fixed in terms of r_1[9], while the derivative with respect to the node i is non-zero only for the cash instrument responsible for that node, since $D(t_i)$ is fixed by r_i only. It is straightforward to compute these derivatives:

$$\frac{\partial x_j}{\partial r_i} = \frac{1}{(t_j - s)} \frac{D(s)}{D(t_j)} \left(-s \frac{\partial r(s)}{\partial r_i} + t_j \frac{\partial r(t_j)}{\partial r_i} \right)\,. \tag{3.9}$$

One can rewrite this in various ways, using (3.5) and the definition of the discrete forward rate:

$$\begin{aligned} \frac{\partial x_j}{\partial r_i} &= \frac{1}{(t_j - s)} (1 + x \cdot \tau) \cdot \frac{\partial}{\partial r_i} \left(-s r(s) + t_j r(t_j) \right) \\ &= (1 + x \cdot \tau) \cdot \frac{\partial}{\partial r_i} \left(\frac{t_j r(t_j) - s r(s)}{t_j - s} \right) \\ &= (1 + x \cdot \tau) \cdot \frac{\partial}{\partial r_i} f^d(s, t_j)\,, \end{aligned}$$

where $f^d(s, t_j)$ is the discrete forward rate between times s and t_j.

A few remarks are now in order. First. Notice that already for cash we see a lower triangular matrix for the quantity $\mathcal{J}$ as defined in (3.2):

$$\mathcal{J} \text{ is a lower triangular matrix}\,. \tag{3.10}$$

In addition, some entries in the lower triangle are zero. The non-zero entries are on the diagonal and in the first column. This is true if only cash instruments are considered[10]. This happens because the zero rate at position i is sensitive only to its cash instrument at i and to the initial cash instrument (via the discount factor $D(s)$). This structure will appear again later when we will start adding swap instruments to the picture, but those zeroes in the lower triangle will generically disappear. The zeroes in the triangle are typical of cash instruments.

[9] This is true in many cases, e.g. when flat extrapolation is used.

[10] Otherwise the positions of the vanishing and the non-vanishing entries will generally be different.

Second. Later we will need to compute the matrix of derivatives $\frac{\partial r_i}{\partial x_j}$. This matrix will simply be given by the inverse matrix of (3.10). If we denote the elements of the Jacobian matrix (3.10) as

$$\mathcal{J}_{kl} = \frac{\partial x_k}{\partial r_l}\,, \tag{3.11}$$

where for each cash k the entries are given by (3.9), then we have

$$\frac{\partial r_i}{\partial x_j} = \left(\mathcal{J}^{-1}\right)_{ij}\,. \tag{3.12}$$

3.2 Swap

For plain vanilla interest rate swaps, or swaps in short, the reasoning is similar, but the formulas are more complicated. The defining relation for the swap with total maturity τ is:

$$D(s) - D(t) = x \sum_{k=1}^{N_{cf}} \alpha_k D(t_k)\,. \tag{3.13}$$

The notation here is as follows. The l.h.s. represents the floating side of the swap, where s is the settlement delay and t is the end date[11] using the conventions for the floating leg of the swap. The r.h.s. is the fixed side of the swap, x is the swap rate, N_{cf} is the number of fixed cash flows, t_j is the time when the j^{th} cash flow is paid, using the conventions for the fixed leg of the swap, and α_j is the tenor of each fixed cash flow[12]. If we invert this relation we obtain

$$x = \frac{D(s) - D(t)}{\sum_{k=1}^{N_{cf}} \alpha_k D(t_k)} \tag{3.14}$$

for the swap rate and

$$\begin{aligned}\frac{\partial x}{\partial r_i} \quad = \quad & \frac{1}{\left(\sum_{k=1}^{N_{cf}} \alpha_k D(t_k)\right)^2} \left[\left(-sD(s)\frac{\partial r(s)}{\partial r_i} + tD(t)\frac{\partial r(t)}{\partial r_i}\right) \cdot \sum_{k=1}^{N_{cf}} \alpha_k D(t_k)\right. \\ & \left. + \Big(D(s) - D(t)\Big) \cdot \sum_{k=1}^{N_{cf}} \alpha_k D(t_k) t_k \frac{\partial r(t_k)}{\partial r_i}\right] \end{aligned} \tag{3.15}$$

for the derivative of the swap rate w.r.t. a given zero rate r_i. Observe that the quantity $\sum_{k=1}^{N_{cf}} \alpha_k D(t_k)$ appearing in the equation above is the PV01 (or better the IV01) of the swap, obtained by deriving (3.13) by x.

In general, in the curve building we have more than just one swap instrument, and hence there will be more swap rates x_1, x_2, ..., x_n, with tenors respectively $0 \leq s \leq \tau_1 \leq \tau_2 \leq \cdots \leq \tau_n$. As an example, for $n = 3$ we could have $\tau_1 = 1$ year, $\tau_2 = 3$ years, $\tau_3 = 5$ years. As we have already done for cash, in this case we have

$$x_{1y} = \frac{D(s) - D(t_1)}{\sum_{k=1}^{N_{cf}} \alpha_k D(t_k)} \tag{3.16a}$$

[11] In the mathematical world, $t = \tau + s$ and later $t_{N_{cf}} = \tau + s$. However, the floating leg has generally different conventions for measuring times than the fixed leg, hence it might happen that the end date on the floating and fixed side are not equal.

[12] It is computed as the time between the $(j-1)^{\text{th}}$ and the j^{th} fixed cash flow and it depends on the index j because of the day count convention.

$$x_{3y} = \frac{D(s) - D(t_2)}{\sum_{k=1}^{N_{cf}} \alpha_k D(t_k)} \tag{3.16b}$$

$$x_{5y} = \frac{D(s) - D(t_3)}{\sum_{k=1}^{N_{cf}} \alpha_k D(t_k)} \tag{3.16c}$$

where $t_i = \tau_i + s$. From these expressions it easy to see that

$$\frac{\partial x_{1y}}{\partial r_1} \neq 0 \qquad \frac{\partial x_{1y}}{\partial r_2} = 0 \qquad \frac{\partial x_{1y}}{\partial r_3} = 0 \tag{3.17a}$$

$$\frac{\partial x_{3y}}{\partial r_1} \neq 0 \qquad \frac{\partial x_{3y}}{\partial r_2} \neq 0 \qquad \frac{\partial x_{3y}}{\partial r_3} = 0 \tag{3.17b}$$

$$\frac{\partial x_{5y}}{\partial r_1} \neq 0 \qquad \frac{\partial x_{5y}}{\partial r_2} \neq 0 \qquad \frac{\partial x_{5y}}{\partial r_3} \neq 0\,. \tag{3.17c}$$

Hence we recover the lower triangular structure as in (3.10), but now all the entries are generically non-zero[13] because of the intermediate cash flows coming from the fixed leg. For instance, this generically happens for the entry $\frac{\partial x_{5y}}{\partial r_2}$, where the r_2-dependence is in the denominator of x_{5y}.

Extending this $n = 3$-result to the general situation is straightforward. We will have a Jacobian matrix of the form

$$\mathcal{J}_{kl} = \frac{\partial x_k}{\partial r_l}\,, \tag{3.18}$$

where for each swap k the entries are given by (3.15) and whose inverse is

$$\frac{\partial r_i}{\partial x_j} = \left(\mathcal{J}^{-1}\right)_{ij}\,. \tag{3.19}$$

Moreover, both matrices are lower triangular because the zero rate at position i is sensitive to all the swap instruments up to i (through the settlement discount $D(s)$ and all the fixed-leg cash flows).

4 Interpolation

We have seen in section 3 that the Jacobian $\mathcal{J}$ is computed in terms of the matrix

$$\frac{\partial r(t)}{\partial r_i}\,, \tag{4.1}$$

which is interpolation-dependent. Therefore it is important to understand how it is computed in various interpolation schemes. For all the methods considered here, except Hagan and West's interpolation, we use *flat extrapolation* outside the data points (t_i, r_i), with $i = 1, \ldots, n$:

$$r(t) = r_1 \quad t \leq t_1 \qquad \text{and} \qquad r(t) = r_n \quad t \geq t_n\,. \tag{4.2}$$

This immediately implies that for those methods one has:

$$\frac{\partial r(t)}{\partial r_i} = \delta_{i,1} \quad t \leq t_1 \qquad \text{and} \qquad \frac{\partial r(t)}{\partial r_i} = \delta_{i,n} \quad t \geq t_n\,, \tag{4.3}$$

[13]Some of them could still be zero, but it would be by chance and case dependent.

where $\delta_{i,j}$ denotes the Kronecker delta. This is not the case for Hagan and West's method. For all the methods however the following is true: for any time t within the input range, we can always determine the interval $[t_{\hat{i}}, t_{\hat{i}+1}]$ for a specific index $\hat{i}$ that encloses t, i.e. $t_{\hat{i}} \leq t \leq t_{\hat{i}+1}$[14].

In the following subsections we will give the exact analytic expressions for the derivative (4.1) for some specific interpolation schemes. Observe that, within our framework, actually all the higher-order derivatives are always known *exactly*. In practice, the highest order is limited by the specific interpolation method used to construct the curve. Typically, spline-like methods result in curve which are $\mathcal{C}^1$, sometimes $\mathcal{C}^2$, with vanishing higher derivatives, so that in a Taylor-like approach the sum will stop quite soon.

4.1 Linear Interpolation

In linear interpolation, the curve $r(t)$ is specified by

$$r(t) = \left(\frac{r_{\hat{i}+1} - r_{\hat{i}}}{t_{\hat{i}+1} - t_{\hat{i}}}\right)(t - t_{\hat{i}}) + r_{\hat{i}}\,, \qquad t \in [t_{\hat{i}}, t_{\hat{i}+1}]\,. \tag{4.4}$$

Hence our derivative is

$$\frac{\partial r(t)}{\partial r_i} = \begin{cases} 1 - \left(\frac{t-t_{\hat{i}}}{t_{\hat{i}+1}-t_{\hat{i}}}\right) & \text{if} \quad i = \hat{i} \\ \left(\frac{t-t_{\hat{i}}}{t_{\hat{i}+1}-t_{\hat{i}}}\right) & \text{if} \quad i = \hat{i}+1 \\ 0 & \text{otherwise} \end{cases} . \tag{4.5}$$

4.2 Cubic Splines

In interpolations based on cubic spline the curve $r(t)$ is specified by

$$r(t) = a_x + b_{\hat{i}}(t - t_{\hat{i}}) + c_{\hat{i}}(t - t_{\hat{i}})^2 + d_{\hat{i}}(t - t_{\hat{i}})^3\,, \qquad t \in [t_{\hat{i}}, t_{\hat{i}+1}]\,. \tag{4.6}$$

Hence our derivative is generically:

$$\frac{\partial r(t)}{\partial r_i} = \frac{\partial a_{\hat{i}}}{\partial r_i} + \frac{\partial b_{\hat{i}}}{\partial r_i}(t - t_{\hat{i}}) + \frac{\partial c_{\hat{i}}}{\partial r_i}(t - t_{\hat{i}})^2 + \frac{\partial d_{\hat{i}}}{\partial r_i}(t - t_{\hat{i}})^3\,. \tag{4.7}$$

The choice of the coefficients a, b, c and d defines different methods. For the spline-like methods that we will consider here, it is enough to specify the b's, i.e. the derivatives of the curve at the node points. In fact, the a's are fixed by the nodes, i.e. $a_{\hat{i}} = r_{\hat{i}}$, and for c's and d's we have (see [8]):

$$c_{\hat{i}} = \frac{3m_{\hat{i}} - b_{\hat{i}+1} - 2b_{\hat{i}}}{h_{\hat{i}}} \qquad \text{and} \qquad d_{\hat{i}} = \frac{b_{\hat{i}+1} + b_{\hat{i}} - 2m_{\hat{i}}}{h_{\hat{i}}^2}\,, \tag{4.8}$$

where we have defined

$$h_{\hat{i}} = t_{\hat{i}+1} - t_{\hat{i}} \qquad \text{and} \qquad m_{\hat{i}} = \frac{a_{\hat{i}+1} - a_{\hat{i}}}{h_{\hat{i}}}\,. \tag{4.9}$$

[14]For Hagan and West's interpolation in section 4.3 the interval will be denoted by $[t_{\hat{i}-1}, t_{\hat{i}}]$.

Consequently, we have

$$\frac{\partial a_{\hat{i}}}{\partial r_i} = \delta_{i,\hat{i}}\,, \tag{4.10}$$

and the derivatives

$$\frac{\partial c_{\hat{i}}}{\partial r_i} \quad \text{and} \quad \frac{\partial d_{\hat{i}}}{\partial r_i} \tag{4.11}$$

are fixed in terms of the derivatives of the b's. The only method-dependent quantity is then

$$\frac{\partial b_{\hat{i}}}{\partial r_i}\,. \tag{4.12}$$

We will consider two kinds of cubic splines, namely the Bessel-Hermite and the monotone-preserving splines, and show how the b's and its derivative are computed there.

4.2.1 Bessel-Hermite Cubic Splines

In Bessel-Hermite splines the b-coefficients are fixed by considering three consecutive points, drawing the only parabola passing through all of them and computing the derivative at the node. This procedure defines formulas for these coefficients, that we recall here [8]:

$$\begin{aligned}
b_1 &= \frac{1}{t_3 - t_1}\left[\frac{(t_3 + t_2 - 2t_1)(r_2 - r_1)}{t_2 - t_1} - \frac{(t_2 - t_1)(r_3 - r_2)}{t_3 - t_2}\right] && (4.13)\\
b_i &= \frac{1}{t_{i+1} - t_{i-1}}\left[\frac{(t_{i+1} - t_i)(r_i - r_{i-1})}{t_i - t_{i-1}} + \frac{(t_i - t_{i-1})(r_{i+1} - r_i)}{t_{i+1-t_i}}\right] \quad (1 < i < n) && (4.14)\\
b_n &= -\frac{1}{t_n - t_{n-2}}\left[\frac{(t_n - t_{n-1})(r_{n-1} - r_{n-2})}{t_{n-1} - t_{n-2}} - \frac{(2t_n - t_{n-1} - t_{n-2})(r_n - r_{n-1})}{t_n - t_{n-1}}\right]. && (4.15)
\end{aligned}$$

The calculation of $\frac{\partial b_{\hat{i}}}{\partial r_i}$ is now straightforward. First, for any given value of t to interpolate, one has to determine the lower index $\hat{i}$ that identifies the corresponding $b_{\hat{i}}$, and then the derivative can be carried out as usual.

4.2.2 Monotone-Preserving Cubic Splines

As the name suggests, these splines preserve the monotonicity of the curve by following the trend of the input points. This is achieved in the following way [8]. First, the derivatives at the boundary are set to zero, $b_1 = b_n = 0$. Then, if the curve has a turning point at position $\hat{i}$, its derivative is also set to zero, $b_{\hat{i}} = 0$, otherwise its value is

$$b_{\hat{i}} = \beta_{\hat{i}}\,, \tag{4.16}$$

where $\beta_{\hat{i}}$ is equal to

$$\beta_{\hat{i}} = \frac{3 m_{\hat{i}-1} m_{\hat{i}}}{\max(m_{\hat{i}-1}, m_{\hat{i}}) + 2\min(m_{\hat{i}-1}, m_{\hat{i}})}\,. \tag{4.17}$$

In the original paper [8] the following additional adjustment was included:

$$b_{\hat{i}} = \begin{cases} \min\big(\max(0, b_{\hat{i}}), 3\min(m_{\hat{i}-1}, m_{\hat{i}})\big) & \text{if } m_{\hat{i}} > 0 \\ \max\big(\min(0, b_{\hat{i}}), 3\max(m_{\hat{i}-1}, m_{\hat{i}})\big) & \text{if } m_{\hat{i}} < 0 \end{cases}\,. \tag{4.18}$$

However, it can be shown [18, 19] that the latter adjustment is not necessary, since monotonicity had already been achieved at the previous step. At this point the calculation of the derivative $\frac{\partial b_{\hat{i}}}{\partial r_i}$ is straightforward: if the curve has a turning point at $\hat{i}$, then $\frac{\partial b_{\hat{i}}}{\partial r_i} = 0$, otherwise the curve is locally monotonic at $\hat{i}$ and one has:

$$\frac{\partial b_{\hat{i}}}{\partial r_i} = \frac{\partial \beta_{\hat{i}}}{\partial r_i} \tag{4.19}$$

(see [18, 19] for the relevant expressions).

4.3 Forward Monotone Convex Splines

This method was introduced by Hagan and West [8]. In [9] they also give the algorithm to implement it.

The method starts from the realistic assumption that forward rates are constant in the intervals $[t_{i-1}, t_i]$, with $1 \leq i \leq n$, $t_0 = 0$. Eventually, the yield curve is recovered from the forward function as

$$f(t) = \frac{d(r(t)t)}{dt} \implies \quad r(t) = \frac{1}{t}\left[r_{\hat{i}-1}t_{\hat{i}-1} + f^d_{\hat{i}}(t - t_{\hat{i}-1}) + I_t\right], \qquad t \in [t_{\hat{i}-1}, t_{\hat{i}}], \tag{4.20}$$

where

$$I_t = \int_{t_{\hat{i}-1}}^{t} g(\tau)d\tau \tag{4.21}$$

and $t \in [t_{\hat{i}-1}, t_{\hat{i}}]$ for a specific index $\hat{i} \in [1, n]$. This integral is also defined in four regions.

Finally, since we are interested in the Jacobian $\mathcal{J}$, we need to compute the derivative of the yield curve w.r.t. the input zero rates:

$$\frac{\partial r(t)}{\partial r_i} = \frac{1}{t}\left[\delta_{i,\hat{i}-1}t_{\hat{i}-1} + \left(\frac{\partial}{\partial r_i} f^d_{\hat{i}}\right)(t - t_{\hat{i}-1}) + \frac{\partial}{\partial r_i} I_t\right]. \tag{4.22}$$

The derivative of the integral is done region by region, while the derivative of the forward rates gives a linear-like contribution:

$$\frac{\partial}{\partial r_i} f^d_{\hat{i}} = \delta_{i,\hat{i}-1}\left(1 - \frac{t - t_{i-1}}{t_i - t_{i-1}}\right) + \delta_{i,\hat{i}}\left(\frac{t - t_{i-1}}{t_i - t_{i-1}}\right). \tag{4.23}$$

The relevant formulas for $g(\tau)$, I_t and $\frac{\partial I_t}{\partial r_j}$ can be found in appendix A.

5 Hedging

In this section we will describe how hedging works. There are various approaches to compute hedging-related quantities. One such approach is the so-called *wave* or *scenario* method, which allows one to separate the risks of the yield curve from the instruments. This is sometimes desirable since in principle the curve instruments do not need to be exactly the same as the hedging instruments. However, we will not pursue this approach here, but leave it for the future.

We will assume that the fundamental vector is a vertical column and hence its transpose is a horizontal row. Moreover, the relevant vectors will be denoted by lower-case Greek letters, while matrices by upper-case Greek letter (except for the Jacobian, which is still denoted by $\mathcal{J}$). Due to the amount of notation used in this section, let us list the main quantities first:

- $\mathcal{J}$ is the Jacobian introduced earlier
- Ψ will denote the matrix whose columns are the IV01 of the curve instruments $\mathcal{I}$
- Ψ_i, with $i = 1, \ldots, n$ will denote the columns of Ψ
- ψ will denote the IV01 of an arbitrary product that we want to hedge
- Π will denote the matrix whose columns are the PV01 of the curve instruments $\mathcal{I}$
- Π_i, with $i = 1, \ldots, n$ will denote the columns of Π
- ξ will denote the PV01 of an arbitrary product that we want to hedge
- ω will denote the vector with the hedging position

We have already defined the PV01 and the IV01 in (2.14) and (2.17). Let us recall them here for convenience:

$$PV01_i = \frac{\partial PV}{\partial r_i}\delta r_i\,, \qquad IV01_i = \frac{\partial PV}{\partial x_i}\delta x_i\,. \tag{5.1}$$

These are vector components (not summed over), corresponding to the nodes in the yield curve. The PV01 measures the sensitivity of any product's present value to changes in the zero rates by one basis point, while the IV01 measures a similar sensitivity when the underlying instrument rate changes by the same amount. We have already seen that they are connected by the Jacobian $\mathcal{J}$

$$\mathcal{J}_{kl} = \frac{\partial x_k}{\partial r_l} \qquad \Longleftrightarrow \qquad \left(\mathcal{J}^{-1}\right)_{kl} = \frac{\partial r_k}{\partial x_l} \tag{5.2}$$

via the relation

$$PV01_i = \sum_j IV01_j \mathcal{J}_{ji} \qquad \Longleftrightarrow \qquad IV01_i = \sum_j PV01_j \left(\mathcal{J}^{-1}\right)_{ji} \tag{5.3}$$

where the sum in j runs over the node points of the yield curve, or in matrix notation

$$PV01 = \mathcal{J}^T \cdot IV01 \qquad \Longleftrightarrow \qquad IV01 = \mathcal{J}^{-1T} \cdot PV01\,, \tag{5.4}$$

where T denotes transposition.

In hedging, we want to replicate a product (or a portfolio of products) using the instruments that are available in the market, in such a way that fluctuations on the present value of our original product due to fluctuations in the underlying rates are matched by the same fluctuation in the portfolio. In this way one can make a portfolio immune to small changes in the yield curve.

In this framework, the hedging construction is achieved in the following way, by looking both at the curve and at the product to be hedged.

Given a yield curve constructed out of n instruments, we can define an n-by-n matrix Ψ whose *columns* are the IV01's of each instrument. Since the zero rate at a generic position i is sensitive to all the previous instruments up to and including i, due to the bootstrap procedure, then the matrix Ψ will be upper triangular:

$$\Psi \equiv \Psi(\mathcal{I}) = \left(\begin{array}{c|c|c|c} \vdots & \vdots & \vdots & \vdots \\ \Psi_1 & \Psi_2 & \ldots & \Psi_n \\ \vdots & \vdots & \vdots & \vdots \end{array}\right) = \left(\begin{array}{c|c|c|c} \psi_{11} & \psi_{12} & \ldots & \psi_{1n} \\ 0 & \psi_{22} & \ldots & \psi_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \psi_{nn} \end{array}\right), \tag{5.5}$$

where Ψ_i are the column vectors defined as the IV01 of the i^{th} instrument, with $\psi_{ij} = (\Psi_j)_i$. Also note that Ψ is a curve property since it depends only on the set of instruments $\mathcal{N}$. Moreover, the diagonal entries of Ψ are strictly non-zero, hence this matrix is always invertible.

Now consider a product to be hedged. We can work out its IV01 as explained in section 3. It is a *vector* whose entries depend on its complexity. Let's call it ψ:

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{pmatrix}, \qquad \psi_i \in \mathbb{R}\,. \tag{5.6}$$

The hedging strategy is now to write the vector ψ as a linear combination of the column vectors of the matrix Ψ with some suitable vector of coefficients $\omega = (\omega_1, \omega_2, \ldots, \omega_n)^T$, where T denotes transposition:

$$\psi = \omega_1 \cdot \Psi_1 + \omega_2 \cdot \Psi_2 + \cdots + \omega_n \cdot \Psi_n\,. \tag{5.7}$$

More formally[15]:

$$\psi_i = \sum_{j=1}^{n} \Psi_{ij}\omega_j\,, \qquad \omega_j \in \mathbb{R} \qquad \Longleftrightarrow \qquad \psi = \Psi \cdot \omega\,, \tag{5.8}$$

in components and matrix notation respectively. Since Ψ is invertible, for any given product that needs to be hedged it is always possible to find a solution for ω. The solution for the hedging vector is unique and given by

$$\omega = \Psi^{-1} \cdot \psi\,. \tag{5.9}$$

These values for the ω_i's tell us how to replicate the initial product. In particular, if we choose the product ψ to be exactly one of the curve instruments in $\mathcal{I}$, then we expect all the components of the ω vector to vanish, except for the component corresponding to the original product which should be equal to one.

In detail, in order to insure ourselves against moves in the present value of a product due to changes in the underlying curve rates, we can build a portfolio with the same instruments in $\mathcal{I}$, each in the right abundance, such that the replicated portfolio will move in the opposite direction:

$$\delta\zeta \sim \bigoplus_{j=1}^{n} \omega_j \iota_j\,, \qquad \iota_j \in \mathcal{I}\,. \tag{5.10}$$

We will refer to this fact by saying that for each product *the curve instruments spans the complete risk space*. We will be more formal in the next subsection.

Ψ and $\mathcal{J}$ are both curve properties. Once we know them, we can define another matrix Π as the product

$$\Pi = \mathcal{J}^T \cdot \Psi\,, \tag{5.11}$$

[15] We define the positions ω with a plus sign. That will allow us to replicate the original product or portfolio of products. For hedging purposes one must use opposite positions, obtained by changing the sign of ω.

which is called the *PV01 hedging matrix in the IV01 hedging strategy* for reasons that will be clear in a moment. This matrix is also a curve property:

$$\Pi \equiv \Pi(\mathcal{I})\,. \tag{5.12}$$

Since the Jacobian $\mathcal{J}$ is lower triangular and Ψ is upper triangular, Π will also be upper triangular. We can now compute the PV01 of our original product as follows, using (5.4) and (5.8). If we denote it by ξ

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}\,, \qquad \xi_i \in \mathbb{R}\,, \tag{5.13}$$

then we have:

$$\xi = \Pi \cdot \omega\,, \tag{5.14}$$

in terms of the replicated portfolio. This justifies its name. The figure below summarizes the process.

$$\begin{tikzcd} \psi \arrow[r, "\Psi^{-1}"] \arrow[d, "\mathcal{J}^T"'] & \omega \arrow[dl, "\Pi=\mathcal{J}^T\cdot\Psi"] \\ \xi & \end{tikzcd}$$

Similarly, we could repeat the same construction and start from the PV01 instead. We can define the matrix $\hat{\Pi}$ whose *columns* are the PV01's of each instrument:

$$\hat{\Pi} \equiv \hat{\Pi}(\mathcal{I}) = \left(\begin{array}{c|c|c|c} \vdots & \vdots & \vdots & \vdots \\ \hat{\Pi}_1 & \hat{\Pi}_2 & \dots & \hat{\Pi}_n \\ \vdots & \vdots & \vdots & \vdots \end{array}\right) = \left(\begin{array}{c|c|c|c} \xi_{11} & \xi_{12} & \dots & \xi_{1n} \\ 0 & \xi_{22} & \dots & \xi_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \xi_{nn} \end{array}\right)\,, \tag{5.15}$$

where $\hat{\Pi}_i$ is the column vector containing the PV01 of the i^{th} instrument, with $\xi_{ij} = \left(\hat{\Pi}_j\right)_i$. Considerations similar to the previous case hold here as well. In particular, $\hat{\Pi}$ is upper triangular and we can write the product's PV01 as a linear combination of the columns of $\hat{\Pi}$

$$\xi = \sum_{j=1}^{n} \hat{\Pi}_{ij}\hat{\omega}_j \qquad \Longleftrightarrow \qquad \xi = \hat{\Pi}\cdot\hat{\omega} \tag{5.16}$$

with coefficients $\hat{\omega}_j$. Hence

$$\hat{\omega} = \hat{\Pi}^{-1}\cdot\xi\,. \tag{5.17}$$

The process diagram is now the following:

$$\begin{tikzcd} \xi \arrow[r, "\hat{\Pi}^{-1}"] & \hat{\omega} \arrow[dl, "\hat{\Psi}=\left(\mathcal{J}^T\right)^{-1}\cdot\hat{\Pi}"] \\ \psi \arrow[u, "\mathcal{J}^T"] & \end{tikzcd}$$

where $\hat{\Psi} = \left(\mathcal{J}^T\right)^{-1} \cdot \hat{\Pi}$ is the *IV01 hedging matrix in the PV01 hedging strategy*, since the product's IV01 is recoved as

$$\psi = \left(\mathcal{J}^T\right)^{-1} \cdot \xi = \left(\mathcal{J}^T\right)^{-1} \cdot \hat{\Pi} \cdot \hat{\omega} \equiv \hat{\Psi} \cdot \hat{\omega} \, . \tag{5.18}$$

By gluing the two diagrams together one gets:

$$\begin{tikzcd} & \psi \arrow[r, "\Psi^{-1}"] \arrow[d, "\mathcal{J}^T"] & \omega \arrow[dl, "\Pi"] \\ \hat{\omega} \arrow[ur, "\hat{\Psi}"] & \xi \arrow[l, "\hat{\Pi}^{-1}"] & \end{tikzcd} \tag{5.19}$$

This diagram summarizes all the relations between the matrices and vectors that we have introduced so far. From the diagram it follows e.g. that

$$\xi = \Pi\omega = \mathcal{J}^T \hat{\Psi}\hat{\omega} \tag{5.20}$$

which we will solve in a moment.

Recall that each column vector of $\hat{\Pi}$ is equal to the matrix product of $\mathcal{J}^T$ with the corresponding column of Ψ. This tells us that these two matrices are related by

$$\hat{\Pi} = \mathcal{J}^T \cdot \Psi \, , \tag{5.21}$$

which is exactly Π. Using this observation, together with the diagram (5.19), it is now easy to show that the position vectors are also equal:

$$\hat{\omega} = \hat{\Pi}^{-1} \cdot \xi = \Psi^{-1} \left(\mathcal{J}^T\right)^{-1} \cdot \xi = \Psi^{-1} \left(\mathcal{J}^T\right)^{-1} \cdot \mathcal{J}^T \cdot \psi = \Psi^{-1} \left(\mathcal{J}^T\right)^{-1} \cdot \mathcal{J}^T \cdot \Psi \cdot \omega \, ,$$

i.e.

$$\hat{\omega} = \omega \, . \tag{5.22}$$

Similarly, the sets of matrices in the two cases are the same:

$$\hat{\Pi} = \Pi \qquad \text{and} \qquad \hat{\Psi} = \Psi \, . \tag{5.23}$$

This implies that, independently of the hedging strategy (either using PV01 or IV01), all the relevant quantities are invariant (for the position vector we have $\omega = \hat{\omega}$, for the curve PV01 matrix we have $\Pi = \hat{\Pi}$, for the curve IV01 matrix we have $\Psi = \hat{\Psi}$).

We can summarize all the relations involving the yield curve with the following diagram:

$$\begin{tikzcd} \text{curve space} \arrow[d, "\delta"] & \mathcal{I} \arrow[rr, "\mathcal{B}"] \arrow[d, "\delta"] & & \mathcal{N} \arrow[d, "\delta"] \\ \text{risk space} \arrow[d, "\text{hedge}"] & \Psi \arrow[rr, "\mathcal{J}^T"] \arrow[dr, "\text{hedge}"] & & \Pi \arrow[dl, "\text{hedge}"] \\ \text{hedge space} & & \omega & \end{tikzcd} \tag{5.24}$$

which shows all the main features, namely

- the translation from the original instruments and nodes to the risk space is obtained by going to the derivative space and using first-order Taylor approximation[16];
- the bootstrap $\mathcal{B}$ between the instrument set $\mathcal{I}$ and the curve node set $\mathcal{N}$ is translated into the Jacobian matrix $\mathcal{J}^T$ in the risk space;
- the hedging positions are strategy-invariant.

Based on this result, from now on we will use only one notation for the curve IV01 and PV01, as well as for the position vector.

5.1 Portfolios of several products

This reasoning can be easily extended to a portfolio of many products by using linearity. Formally, a portfolio ζ contains m individual products ζ_k, with $k = 1, \dots, m$:

$$\zeta = \{\zeta_1, \dots, \zeta_k, \dots, \zeta_m\}\,, \tag{5.25}$$

Then, by linearity, the total IV01 and PV01 will be just the corresponding sums:

$$\psi = \sum_{k=1}^{m} \psi_k\,, \qquad \xi = \sum_{k=1}^{m} \xi_k\,. \tag{5.26}$$

Using linearity and the fact that the matrices Ψ and Π are curve properties (i.e. they are product-independent), it is easy to see that:

$$\psi = \Psi \cdot \omega \qquad \text{and} \qquad \xi = \Pi \cdot \omega \tag{5.27}$$

where

$$\omega = \Psi^{-1} \cdot \psi = \sum_{k=1}^{m} \omega_k \qquad \text{with} \qquad \omega_k = \Psi^{-1} \cdot \psi_k\,. \tag{5.28}$$

is the hedging vector for the whole portfolio expressed as a sum of the hedging vectors for the single products.

Linearity and the fact that ω_k is fully determined by the individual product ζ_k allow us to compute the complete hedging vector ω for the entire portfolio and hence its full IV01 and PV01. Thus the risk management for portfolios ζ with many products is solved in terms of the risk management for the individual products ζ_k. As a last remark, observe that this generalization is straightforward because of linearity of the portfolio, whose value is linearly determined by the values of the underlying products inside the portfolio itself.

5.2 Risk space

We have seen earlier that the Ψ and Π matrices are triangular. Let us focus on Ψ for the moment. We can define the risk space for Ψ to be the range of Ψ:

$$\mathcal{R}_\Psi \equiv \text{Range}(\Psi)\,. \tag{5.29}$$

[16] Or the exact integral formalism.

The range of Ψ is the span generated by the columns of Ψ. Since Ψ is triangular, it is easy to see that its columns are linearly independent, hence they form a basis for the range and therefore for the risk space (5.29).

However, we could have also started from the matrix Π. In this case we can define the risk space for Π to be the range of Π:

$$\mathcal{R}_\Pi \equiv \mathrm{Range}(\Pi)\,. \tag{5.30}$$

Then, the columns of Π form a basis. Inspection of equation (5.4) shows that

$$\mathcal{R}_\Pi = \mathcal{J}^T\left(\mathcal{R}_\Psi\right) = \mathrm{span}\{\mathcal{J}^T\Psi_1, \mathcal{J}^T\Psi_2, \ldots, \mathcal{J}^T\Psi_n\}\,. \tag{5.31}$$

In fact, for any product,

$$\xi = \Pi\cdot\omega = \sum_i \Pi_i\omega_i = \sum_i \left(\mathcal{J}^T\cdot\Psi_i\right)\omega_i = \mathcal{J}^T\cdot\sum_i \Psi_i\omega_i = \mathcal{J}^T\cdot\psi\,. \tag{5.32}$$

Triangular matrices are not diagonalizable, but one can still find an orthogonal (and orthonormal) basis $\{e_i\}_{i=1}^n$ out of the columns of Ψ. Given the shape of Ψ, we can guess that one orthonormal basis is the canonical basis

$$e_1 = \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},\ e_2 = \begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix},\ \ldots,\ e_n = \begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}. \tag{5.33}$$

We can easily check that this is indeed the case. Alternatively, one can derive the same result by applying the standard Gram–Schmidt algorithm [20] to triangular matrices. So, if we denote by $\Psi_1, \Psi_2, \ldots, \Psi_n$ the columns of Ψ, then an orthonormal basis $\{e_i\}_{i=1}^n$ is constructed recursively as

$$u_i = \Psi_i - \sum_{j=1}^{i-1}\frac{u_j\cdot\Psi_i}{u_j\cdot u_j}u_j \qquad \text{and} \qquad e_i = \frac{u_i}{\left(u_i\cdot u_i\right)^{1/2}}\,, \tag{5.34}$$

for $i = 1, \ldots, n$, and indeed we get back (5.33).

Looking back at (5.31), we can see that the more stringent relation

$$\mathcal{R}_\Pi = \mathcal{R}_\Psi \tag{5.35}$$

holds. One way to see this is to note that these two spaces have a common set of orthonormal basis, namely the canonical basis (5.33), since both Π and Ψ are triangular. Another way is to recall that the two sets of basis for $\mathcal{R}_\Pi$ and $\mathcal{R}_\Psi$ respectively are related by a linear transformation and by definition they span the same space. Moreover, these spaces are equivalent to the whole $\mathbb{R}^n$, since they all have the canonical vectors (5.33) as basis.

One can be even more explicit and write down the component of the vectors ψ and ξ with respect to the canonical basis. Using some manipulations for the indices we find the linear combinations

$$\psi = \sum_i \omega_i\Psi_i = \sum_{i=1}^n \omega_i\sum_{j=1}^i e_j\psi_{ji} = \sum_{i=1}^n e_i\left(\sum_{j=i}^n \psi_{ij}\omega_j\right) \equiv \sum_{i=1}^n e_i c_i^\psi \tag{5.36}$$

with coefficients $c_i^\psi = \sum_{j=i}^n \psi_{ij}\omega_j$ for ψ, and

$$\xi = \sum_i \omega_i \Pi_i = \sum_{i=1}^n \omega_i \sum_{j=1}^i e_j \xi_{ji} = \sum_{i=1}^n e_i \left(\sum_{j=i}^n \xi_{ij}\omega_j \right) \equiv \sum_{i=1}^n e_i c_i^\xi \tag{5.37}$$

with coefficients $c_i^\xi = \sum_{j=i}^n \xi_{ij}\omega_j$ for ξ, where $\xi_{ij} = \sum_k \left(\mathcal{J}^T\right)_{ik} \psi_{kj}$.

As a final remark, observe that for any given product all the other hedge quantities are fixed by the curve via its instruments. This means that the choice of the instrument set $\mathcal{I}$ is crucial for hedging. In fact, in the diagram (5.24) everything starts from $\mathcal{I}$ and the resulting hedge or replication of a product (or a portfolio of products) strongly depends on which instruments are part of $\mathcal{I}$. Only those instruments relevant for measuring value and risk of a product should be used. So $\mathcal{I}$ should be chosen with care, because it dictates the value and risk one will perceive[17].

6 Main points

In this section we would like to summarize, generalize and collect together the main points of the discussion so far.

We start from an obvious result.

Theorem 1 (Bootstrapping prices back market)**.** *Present Values of any instrument using the bootstrapped yield curve is equal to market quoted value.*

Proof. This is true by construction of the yield curve. The set of discrete input data are the market data and interpolation is used to determine a value when the data points are not available. Thus, the yield curve passes through all the market data. Consequently, in the present-value calculation for any instrument where the yield curve is used the computed value will match the market value. □

Observe that the statement of this theorem will not be true if other methods were used to generate the yield curve. In particular, using best-fit-like approaches (such as Nelson-Siegel-Svensson) would give us a smooth curve which will not however pass through the data (market) points.

The next theorem shows the relation between the PV01 and the IV01, which are two sides of the same coin, since they are connected by a change of coordinates.

Theorem 2 (PV01 and IV01)**.** *For a given yield curve γ, the corresponding Jacobian (3.2), which relates zero rates to the underlying instrument rates, also relates PV01 and IV01 for any product ζ. If ξ and ψ are the product's PV01 and IV01 vectors respectively, then*

$$\xi = \mathcal{J}^T \cdot \psi \qquad \Longleftrightarrow \qquad \psi = \mathcal{J}^{-1T} \cdot \xi \,. \tag{6.1}$$

Proof. The theorem was shown to be true in equation (5.4). □

The next theorem is about risk management and replication for any product or portfolio of products.

[17] In principle, many instruments are available for hedging (e.g.in terms of maturities for IRS contracts), however hedging with instruments unrelated to the curve will typically result in bad positions, since in general the yield curve will not go through the desired value, unless it is a node point.

Theorem 3 (Ψ spans complete risk space)**.** *Let ζ be any product or a portfolio of products, and $\mathcal{I} = \{\iota_1, \iota_2, \dots, \iota_n\}$ be the set of instruments used to construct the yield curve γ. Then the column vectors of the matrix Ψ as defined in (5.5) are a basis for the risk of any product and portfolios of products:*

$$\mathcal{R}_\Psi = \mathrm{Range}(\Psi)\,. \tag{6.2}$$

Explicitly, the product's IV01 vector ψ can be written as a linear combination of the form

$$\psi = \Psi \cdot \omega\,, \tag{6.3}$$

for a given ω. The change in the portfolio's value $\zeta \longrightarrow \zeta + \delta\zeta$ due to changes in the instruments rates is then replicated by the instruments in $\mathcal{I}$ together with the vector ω of linear coefficients as

$$\delta\zeta \sim \bigoplus_{j=1}^{n} \omega_j \iota_j\,. \tag{6.4}$$

Similarly, the product's PV01 vector ξ can be written as a linear combination of the form

$$\xi = \Pi \cdot \omega\,, \tag{6.5}$$

where Π is given by

$$\Pi = \mathcal{J}^T \cdot \Psi\,. \tag{6.6}$$

Moreover, the matrices Ψ and Π are curve-dependent only and the vector ω is independent of the hedging strategy. Finally,

$$\mathcal{R}_\Pi = \mathcal{R}_\Psi\,, \tag{6.7}$$

i.e. the risk space based of Π coincides with the one based on Ψ.

Proof. This is straightforward from section 5. For one single product ζ, the result follows from equations (5.6), (5.9), (5.10) and (5.13). The statement about the independence of the hedging strategy is a consequence of equation (5.22). The statement about the basis and the one about equality between the two risk spaces derive from the arguments in subsection 5.2. If ζ is a portfolio of products, then one needs to use the analogous equations from subsection 5.1.

□

Observe that once the product ζ and the curve γ have been specified the hedging vector ω, and hence the replication strategy, is fully determined. Also note that what is being replicated is the change in the value of the portfolio, not the portfolio itself. In fact, the portfolio's present values is generically different from the weighted sum of the instruments' present values with weights given by the ω's. Similar statements, *mutatis mutanda*, hold true for the risk space $\mathbb{R}\mathrm{isk}_\Pi$.

We finish with two remarks, one about which instruments should be used to construct the yield curve and for hedging, and one about the use of numerical calculations in this framework.

Remark 1 (Choice of curve instruments)**.** *The choice of the instrument set $\mathcal{I}$ dictates the resulting hedge or replication of a product or a portfolio of products. Only those instruments relevant and available for hedging should be used.*

Proof. The proof is trivial, since the hedge strongly depends on the choice for the curve instruments, as we mentioned at the end of subsection 5.2. □

Remark 2 (No numerical computations)**.** *There are no numerical computations in this framework.*

Proof. Only analytic formulas were used in this paper. □

Observe that the absence of numerical computation is important for many reasons. First of all, numerical computations are intrinsically associated to approximation errors which, despite very small, could lead to significant losses for financial institutions who work with large amounts on a daily basis. Banks' portfolios typically vary from hundreds of billions to a few trillions of Euros ($\approx 10^{11} - 10^{12}$), hence a small numerical error in the hedging strategy, e.g. $O(10^{-9})$, can produce sizable losses.
Secondly, typically a numerical problem could be solved by using various techniques, which perform very well when dealing with a specific aspect, but quite badly with other aspects. As an example, when computing derivatives one can choose to implement it as a right derivative, or as a left derivative, or as a combination of left and right derivative. Clearly this choice is crucial for piece-wise function since it can affect the value of the derivative at the boundary points.
Lastly, when dealing with extremely large or extremely small numbers (e.g. close to the upper or lower limits of the numerical range for a given machine), numerical computations may have stability issues and produce machine-dependent results.
For all these reasons, when possible analytic formulas should be used.

7 Conclusion

In this paper we have considered the problem of evaluating and managing risk for interest rate derivatives, in particular for cash and swap instruments. The crucial ingredient is the construction of the yield curve. It depends on a set $\mathcal{I}$ of input instruments which are uniquely related to the set $\mathcal{N}$ of input zero rates via the bootstrap procedure. The detailed way of how to extract the zero rates from the instruments has been done here for cash and swap instruments only, but it can be generalized further to any other instrument that is part of the curve construction. The starting point is always the formula for the present value of the product, which gives the desired relationship between zero rates and instrument rate.

A crucial ingredient in this framework is the yield curve. The yield curve depends on the specific interpolation method used to generate the values away from the node points. A poor interpolation choice will result in a poor yield curve, where quality criteria such as smoothness, locality, stability and positivity are problematic to have. We have considered only a few interpolation techniques here. The linear method is easy to use and implement, but it is often not accurate enough for our purposes (e.g. it gives serious problems when the derivatives of the curve are needed, since they will be then discontinuous at the node points). Other methods typically suffer of an unnatural wiggly/zig-zag behaviour which is not intrinsic of the initial data, but is a spurious effect of the method. This problem can be corrected by monotonicity constraints, and they have been implemented in some of the splines. Many formulas for basic interpolation uses where already available in the literature.

Interest rate risk management is strictly related to hedging. Both concepts are based on the knowledge of how the value of a particular product changes as a consequence of a change in the yield curve due to small jumps, typically of one basis-point size, in the input rates. This knowledge is mathematically encoded in quantities such as the PV01 and the IV01 of a product, which are defined as the change in the product's price when the input rates change by one basis point. In order to compute them, one can use a numerical approach. However, we have shown that it is possible to compute them analytically, either exactly or at first-order approximation, by focusing on the derivatives of the product's present value with respect to the yield curve zero rates and instrument rates. These derivatives in turn are expressed in terms of the derivatives of the curve with respect to its inputs. The latter derivatives are interpolation-dependent. While many useful interpolation formulas were already known explicitly in the literature, this was not the case for these derivatives. We have given their analytic expression in the main body of the paper as well as in the appendix.

PV01 and IV01 are not independent but related by a change of coordinates, which is mathematically expressed by the Jacobian matrix between instrument rates and zero rates. We have seen that the Jacobian is a triangular matrix. We have also defined the curve PV01 and IV01 matrices and found that they are triangular too. This is a consequence of the bootstrap procedure. We have showed that the whole risk space for any arbitrary product is ultimately determined by the yield curve, since the change of the present value for any product due to changes in the yield curve can be reproduced by a portfolio of instrument consisting exactly of those instruments used to bootstrap the yield curve. In this sense, the yield curve -through its instruments- spans the complete risk space. Consequently, the choice for the instruments used to bootstrap the curve is crucial and only those instruments should be used that are relevant and available for hedging.

Clearly further generalizations of what we have described here could be considered in the future. One obvious extension of this work would be to include many other instruments in the treatment besides cash and swaps. Moreover, one could ask what happens when a multi-curve framework is used instead of the single-curve one. This would be particularly relevant, since the multi-curve is becoming more and more popular after the credit crunch crisis in 2008. Furthermore, many financial institutions have moved to the *wave* or *scenario* method for handling risks. We have used a more standard approach, but we believe that our results could be easily generalizable to this case as well. Finally, in this paper we have considered only linear products, namely those products where convexity effects can be neglected, however it would be interesting to include convexity and higher-order effects[18] into the picture too by working with the exact integral expression and to estimate the error of the approximation in concrete cases. We leave all these point for the future.

[18] Observe that, within our framework, at each order in perturbation theory the derivatives are always known *exactly*. In practice, the highest order that can be reached in the perturbation expansion is limited by the specific interpolation method used to construct the curve. Typically, spline-like methods result in curve which are $\mathcal{C}^1$, sometimes $\mathcal{C}^2$, with vanishing higher derivatives, so the Taylor sum stops quite soon.

Acknowledgements

We would like to thank Patrick Hagan, Julian Wiseman, Stratos Nikolakakis and Sander Scheerders for carefully reading the manuscript and providing us with useful feedback, as well as Laurent Lefort for pointing out some typos.

References

[1] C. Nelson, A. Siegel (1987), *Parsimonious modeling of the Yield Curve*, Journal of Business 60, no 4, pp. 473-489, http://www.math.ku.dk/~rolf/teaching/NelsonSiegel.pdf.

[2] L. Svensson (1995), *Estimating Forward Interest Rates with the Extended Nelson and Siegel Method*, Sveriges Riskbank Quarterly Review 3, http://www.diva-portal.org/smash/get/diva2:342900/FULLTEXT01.pdf.

[3] J. Wiseman (1998) *The Magpie Yield Curve Model*, Technical report, J.P. Morgan Fixed Income Research.

[4] J. Wiseman (2012) *The Magpie Yield Curve Model at SG*, Technical report, Societe Generale Cross Asset Research - Fixed Income.

[5] T. Bjork, B. Christensen (1997), *Interest Rates Dynamics and Consistent Forward Rates Curve*, University of Aarhus, http://www.blackwell-synergy.com/doi/abs/10.1111/1467-9965.00072.

[6] J. James, N. Webber (2000), *Interest Rate Modelling*, Wiley, ISBN: 978-0-471-97523-6.

[7] Bolder, David, and David Stréliski (1999), *Yield Curve Modelling at the Bank of Canada*, Available at SSRN 1082845 http://epe.lac-bac.gc.ca/100/200/301/bankofcanada/technical_reports-ef/no084/tr84.pdf.

[8] P. S. Hagan, G. West (2006), *Interpolation Methods for Curve Construction*, Applied Mathematical Finance, 13: 2, 89 - 129 http://dx.doi.org/10.1080/13504860500396032, http://kalx.net/dsS2011/HagWes2006.pdf.

[9] P. S. Hagan and G. West (2008), *Methods for Constructing a Yield Curve*, Wilmott Magazine, 70 - 81 http://www.math.ku.dk/~rolf/HaganWest.pdf.

[10] Whittaker, J Gregg (1987) *Interest rate swaps: Risk and regulation* Economic review, volume= 72 - 3, pages 3-13, Federal Reserve Bank of Kansas City https://www.kc.frb.org/Publicat/Econrev/EconRevArchive/1987/1q87whit.pdf.

Reitano, Robert R. "Non-parallel yield curve shifts and immunization." The Journal of Portfolio Management 18.3 (1992): 36-43.

[11] Reitano, Robert R (1992) *Non-parallel yield curve shifts and immunization*, The Journal of Portfolio Management volume 18-3, pages 36-43 http://www.robertrreitano.com/files/Non-parallel%20Yield%20Curve%20Shifts%20and%20Immunization.pdf.

[12] Bianchetti, M. (2008). *Two Curves, One Price Pricing & Hedging Interest Rate Derivatives Decoupling Forwarding and Discounting Yield Curves.* arXiv:0905.2770v2 [q-fin.PR] http://arxiv.org/abs/0905.2770.

[13] M. Bianchetti, M. Carlicchi (2008). *Interest Rates After The Credit Crunch: Multiple-Curve Vanilla Derivatives and SABR*, arXiv:1103.2567v2 [q-fin.PR] http://arxiv.org/abs/1103.2567.

[14] C. P. Fries (2013). *Curves and Term Structure Models. Definition, Calibration and Application of Rate Curves and Term Structure Market Models.* preprint, 2013 http://www.fm.mathematik.uni-muenchen.de/download/publications/fries_curves_tsm.pdf.

[15] J. C. Hull (2014), *Options, Futures and Other Derivatives*, 9^{th} edition, Pearson, ISBN-13: 978-0133456318.

[16] R. Deaves, M. Parlar (2000), *A Generalized Bootstrap Method to Determine the Yield Curve*, Applied Mathematical Finance 7, 257- 270 http://www.tandfonline.com/doi/abs/10.1080/.VC049CmSzPY#.VC1SzymSwuE.

[17] J. M. Hyman (1983), *Accurate Monotonicity Preserving Cubic Interpolation*, SIAM J. Sci. Stat. Comput. Vol. 4, No. 4, 645 - 654 http://www.researchgate.net/publication/230676072_Accurate_monotonicity_preserving_cubic_interpolation/file/d912f50c0a8c9585cd.pdf.

[18] M. Maio, J. de Jong (2014), *On the Non-differentiability of Hyman Monotonicity Constraint*, pre-print nr UD2014-01 http://uglyduckling.nl/library_files/PRE-PRINT-UD-2014-01.pdf.

[19] M. Maio, J. de Jong (2014), *Curiosities on Monotone Preserving Cubic Spline*, pre-print nr UD2014-03 http://uglyduckling.nl/library_files/PRE-PRINT-UD-2014-03.pdf.

[20] G. B. Arfken, H. J. Webber, F. E. Harris (2012), *Mathematical Methods for Physicists*, 7th ed., Elsevier Academic Press, ISBN: 978-0-12-384654-9.

[21] M. Maio (2014), *Interpolation methods and the Hagan-West paper*, blog post, http://www.uglyduckling.nl/financial-engineering/yield-curve/interpolation-methods-and-the-hagan-west-paper/.

[22] P. F. du Preez, E. Mare (2013), *Interpolating Yield Curve Data in a Manner that Ensures Positive and Continuous Forward Curves*, SAJEMS NS 16 (2013) No 4:395-406 http://sajems.org/index.php/sajems/article/view/388.

Appendix

A Hagan-West: Relevant Formulas

In this appendix we give the collection of relevant formulas for each of the four regions. The method starts from the realistic assumption that forward rates are constant in the intervals $[t_{i-1}, t_i]$, with $1 \leq i \leq n$. Their values are specified as usual by:

$$f_i^d = \frac{r_i t_i - r_{i-1} t_{i-1}}{t_i - t_{i-1}} \qquad (1 \leq i \leq n \, , \, r_0 = t_0 = 0) \, . \tag{A.1}$$

These discrete rates f_i^d belong to the entire interval $[t_{i-1}, t_i]$. Then one defines the instantaneous forward rates at t_i as:

$$f_i = \frac{t_i - t_{i-1}}{t_{i+1} - f_{i-1}} f_{i+1}^d + \frac{t_{i+1} - t_i}{t_{i+1} - f_{i-1}} f_i^d \qquad (i = 1, 2, \ldots, n-1) \tag{A.2a}$$

$$f_0 = f_1^d - \frac{1}{2}(f_1 - f_1^d) \tag{A.2b}$$

$$f_n = f_n^d - \frac{1}{2}(f_{n-1} - f_n^d) \, . \tag{A.2c}$$

where the values $f_0 = f(t_0)$ and $f_n = f(t_n)$ are such that $f'(0) = f'(t_n) = 0$, and interpolates these instantaneous forward rates with an interpolator $f(t)$ whose average on each interval is the discrete rate. In addition $f(t)$ is constructed to be positive and continuous[19]. Continuity is automatic, while positivity is enforced by the substitution:

$$f_0 \longrightarrow \text{collar}(0, f_0, 2f_1^d) \tag{A.3a}$$

$$f_i \longrightarrow \text{collar}(0, f_i, 2\min(f_i^d, f_{i+1}^d)) \qquad (i = 1, 2, \ldots, n-1) \tag{A.3b}$$

$$f_n \longrightarrow \text{collar}(0, f_n, 2f_n^d) \tag{A.3c}$$

where $\text{collar}(a, b, c) = \max(a, \min(b, c))$.

In addition, one can enforce monotonicity and convexity. This is achieved by noting that the interpolant $f(t)$ is defined modulo a smooth function $g(t)$ whose average is zero on each interval. Then $f(t)$ will be given by

$$f(t) = f_i^d + g(t) \, , \tag{A.4}$$

with

$$\int_{t_{i-1}}^{t_i} g(t) dt = \int_0^1 g(x) dx = 0 \, . \tag{A.5}$$

Define: $g_0 = g(t_{i-1})$, $g_1 = g(t_i)$ and $x = \frac{t - t_{i-1}}{t_i - t_{i-1}}$. Inspection of the derivative of g allows us to construct this function in four different regions. We will give explicit expressions below.

Define:

$$x = \frac{t - t_{i-1}}{t_i - t_{i-1}} \, , \qquad t \in [t_{i-1}, t_i] \, . \tag{A.6}$$

For the four regions we have:

[19]Note that the forward function $f(t)$ in *not* constructed to be differentiable. In fact, its derivative can have discontinuities, as already observed in e.g. [21] and [22].

i) if $g_0 > 0$, $-\frac{1}{2}g_0 \geq g_1 \geq -2g_0$ and $g_0 < 0$, $-\frac{1}{2}g_0 \leq g_1 \leq -2g_0$, then:

$$g(t) = g_0(1 - 4x + 3x^2) + g_1(-2x + 3x^2) \tag{A.7}$$

$$I_t = (t_i - t_{i-1}) \left[g_0(x - 2x^2 + x^3) + g_1(-x^2 + x^3)\right] \tag{A.8}$$

$$\frac{\partial I_t}{\partial r_j} = (t_i - t_{i-1}) \left[g_0'(x - 2x^2 + x^3) + g_1'(-x^2 + x^3)\right] \tag{A.9}$$

ii) if $g_0 < 0$, $g_1 > -2g_0$ and $g_0 > 0$, $g_1 < -2g_0$, then:

$$\eta = 1 + 3\frac{g_0}{g_1 - g_0} = \frac{g_1 + 2g_0}{g_1 - g_0} \tag{A.10}$$

$$g(x) = \begin{cases} g_0 & \text{for } 0 \leq x \leq \eta \\ g_0 + (g_1 - g_0)\left(\frac{x-\eta}{1-\eta}\right)^2 & \text{for } \eta < x \leq 1 \end{cases} \tag{A.11}$$

$$I_t = \begin{cases} (t_i - t_{i-1})g_0 x & \text{for } 0 \leq x \leq \eta \\ (t_i - t_{i-1})\left[g_0 x + \frac{1}{3}(g_1 - g_0)\frac{(x-\eta)^3}{(1-\eta)^2)}\right] & \text{for } \eta < x \leq 1 \end{cases} \tag{A.12}$$

$$\frac{\partial I_t}{\partial r_j} = \begin{cases} (t_i - t_{i-1})g_0' x & \text{for } 0 \leq x \leq \eta \\ (t_i - t_{i-1})\Big[g_0' x + \frac{1}{3}(g_1' - g_0')\frac{(x-\eta)^3}{(1-\eta)^2} + \\ \qquad \frac{1}{3}(g_1 - g_0)\eta'(2x + \eta - 3)\frac{(x-\eta)^2}{(1-\eta)^3}\Big] & \text{for } \eta < x \leq 1 \end{cases} \tag{A.13}$$

iii) if $g_0 > 0$, $0 > g_1 > -\frac{1}{2}g_0$ and $g_0 < 0$, $0 < g_1 < -\frac{1}{2}g_0$, then:

$$\eta = 3\frac{g_1}{g_1 - g_0} \tag{A.14}$$

$$g(x) = \begin{cases} g_1 + (g_0 - g_1)\left(\frac{\eta-x}{\eta}\right)^2 & \text{for } 0 < x < \eta \\ g_1 & \text{for } \eta \leq x < 1 \end{cases} \tag{A.15}$$

$$I_t = \begin{cases} (t_i - t_{i-1})\left[g_1 x + \frac{1}{3}(g_0 - g_1)\left(\frac{\eta^3 - (\eta-x)^3}{\eta^2}\right)\right] & \text{for } 0 < x < \eta \\ (t_i - t_{i-1})\left[g_1 x + \frac{1}{3}(g_0 - g_1)\eta\right] & \text{for } \eta \leq x < 1 \end{cases} \tag{A.16}$$

$$\frac{\partial I_t}{\partial r_j} = \begin{cases} (t_i - t_{i-1})\Big[g_1' x + \frac{1}{3}(g_0' - g_1')\left(\frac{\eta^3 - (\eta-x)^3}{\eta^2}\right) + \\ \qquad \frac{1}{3}(g_0 - g_1)\eta'\left(1 - \frac{(\eta-x)^2}{\eta^3}(\eta + 2x)\right)\Big] & \text{for } 0 < x < \eta \\ (t_i - t_{i-1})\left[g_1' x + \frac{1}{3}(g_0' - g_1')\eta + \frac{1}{3}(g_0 - g_1)\eta'\right] & \text{for } \eta \leq x < 1 \end{cases} \tag{A.17}$$

iv) if $g_0 \geq 0$, $g_1 \geq 0$ and $g_0 \leq 0$, $g_1 \leq 0$, then:

$$\eta = \frac{g_1}{g_0 + g_1} \qquad A = -\frac{g_0 g_1}{g_0 + g_1} \tag{A.18}$$

$$g(x) = \begin{cases} A + (g_0 - A)\left(\frac{\eta - x}{\eta}\right)^2 & \text{for } 0 < x < \eta \\ A + (g_1 - A)\left(\frac{x - \eta}{1 - \eta}\right)^2 & \text{for } \eta < x < 1 \end{cases} \tag{A.19}$$

$$I_t = \begin{cases} (t_i - t_{i-1})\left[Ax + \frac{1}{3}(g_0 - A)\frac{\left(\eta^3 - (\eta - x)^3\right)}{\eta^2}\right] & \text{for } 0 < x < \eta \\ (t_i - t_{i-1})\left[Ax + \frac{1}{3}(g_0 - A)\eta + \frac{1}{3}(g_1 - A)\frac{(x-\eta)^3}{(1-\eta)^2}\right] & \text{for } \eta < x < 1 \end{cases} \tag{A.20}$$

$$\frac{\partial I_t}{\partial r_j} = \begin{cases} (t_i - t_{i-1})\left[A'x + \frac{1}{3}(g_0' - A')\frac{\left(\eta^3 - (\eta - x)^3\right)}{\eta^2} + \right. & \\ \qquad \left. \frac{1}{3}(g_0 - A)\eta'\left(1 - \frac{(\eta - x)^2}{\eta^3}(\eta + 2x)\right)\right] & \text{for } 0 < x < \eta \\ & \\ (t_i - t_{i-1})\left[A'x + \frac{1}{3}(g_0' - A')\eta + \frac{1}{3}(g_0 - A)\eta' + \right. & \\ \qquad \frac{1}{3}(g_1' - A')\frac{(x-\eta)^3}{(1-\eta)^2} + & \\ \qquad \left. \frac{1}{3}(g_1 - A)\eta'\frac{(x-\eta)^2}{(1-\eta)^3}(2x + \eta - 3)\right] & \text{for } \eta < x < 1 \end{cases} \tag{A.21}$$

Here a prime typically denotes a derivative w.r.t. r_j, e.g. $\eta' \equiv \frac{d\eta}{dr_j}$. Above we have used the fact that $g_0 \equiv g(t_{i-1}) = f(t_{i-1}) - f_i^d$ and $g_1 \equiv g(t_i) = f(t_i) - f_i^d$, to compute the derivatives $g_0' \equiv \frac{dg_0}{dr_j}$ and $g_1' \equiv \frac{dg_1}{dr_j}$ in terms of the derivative of f and f^d. This calculation is in principle straightforward but cumbersome because of all the constraints (i.e. positivity, monotonicity and convexity) that have been enforced on the forward curve. We will not give the formulas for g_0' and g_1' here.