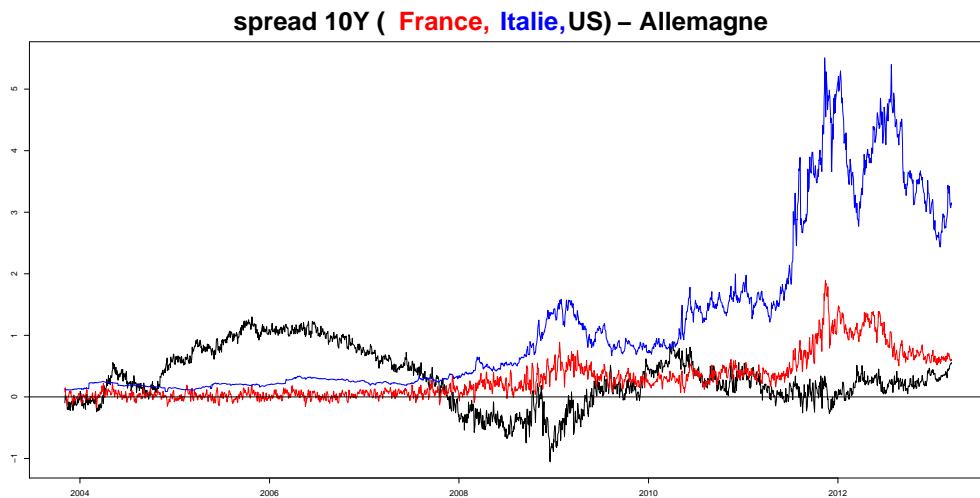


# Modèles de Taux, Surface de Volatilité et Introduction au Risque de Crédit

Alexis Fauth<sup>†</sup>

Université Lille I  
Master 2 Mathématiques et Finance  
Spécialités Mathématiques du Risque & Finance Computationnelle

2014/2015



---

membre associé de l'université Paris I - Panthéon Sorbonne, laboratoire SAMM.  
alexis.fauth@gmail.com, <http://samm.univ-paris1.fr/-Alexis-Fauth->

# Table des matières

<b>1 Présentation des marchés financiers</b>	<b>6</b>
1.1 Marché financier . . . . .	6
1.1.1 Organisé . . . . .	6
1.1.2 De gré à fré . . . . .	6
1.1.3 Acteurs . . . . .	6
1.1.4 Coûts de transactions . . . . .	7
1.2 Produits financiers . . . . .	7
1.2.1 Action . . . . .	7
1.2.2 Indice . . . . .	7
1.3 Matière Première . . . . .	9
1.3.1 Taux de change . . . . .	9
1.3.2 Tracker . . . . .	9
1.3.3 Contrats forward et future . . . . .	10
1.3.4 Option . . . . .	10
1.3.5 Obligation . . . . .	11
1.3.6 Dérivé de crédit . . . . .	12
1.4 Comité de Bâle . . . . .	12
1.5 Risques . . . . .	14
1.5.1 Risque de taux d'intérêt . . . . .	14
1.5.2 Risque de crédit . . . . .	15
1.5.3 Risque d'inflation . . . . .	15
1.5.4 Risque de taux de change . . . . .	16
1.5.5 Risque de taxe . . . . .	16
<b>2 Taux d'intérêt</b>	<b>16</b>
2.1 Zoologie . . . . .	16
2.1.1 Obligation zéro-coupon . . . . .	16
2.1.2 Structure par termes . . . . .	17
2.1.3 Taux d'intérêt forward . . . . .	17
2.1.4 Taux d'intérêt court terme . . . . .	18
2.1.5 Obligation à coupons fixe . . . . .	19
2.1.6 Duration et convexité . . . . .	19
2.1.7 Swap de taux d'intérêt . . . . .	20
2.1.8 Taux Libor . . . . .	22
2.1.9 Convention de marché . . . . .	24
2.1.10 Intérêts courus, clean et dirty price . . . . .	24

2.1.11	Cap et floor . . . . .	25
2.1.12	Swaption . . . . .	26
2.2	Modèle de taux court . . . . .	27
2.2.1	Valorisation d'une obligation . . . . .	27
2.2.2	Modèle de Vasicek . . . . .	30
2.2.3	Modèle de Cox-Ingersoll-Ross . . . . .	33
2.2.4	Généralisation, modèle de taux court à un facteur . . . . .	36
2.2.5	Modèle de Hull and White . . . . .	37
2.2.6	Quelques modèles . . . . .	40
2.3	Modèles multifactoriels . . . . .	41
2.3.1	Fondations . . . . .	41
2.3.2	Modèle de Vasicek à deux facteurs . . . . .	43
2.3.3	Quelques modèles . . . . .	47
2.4	Pricing de dérivés de taux . . . . .	50
2.4.1	Mesure forward . . . . .	50
2.4.2	Prix d'une option sur une obligation . . . . .	51
2.4.3	Pricing de cap et floor . . . . .	54
2.4.4	Pricing de swaption . . . . .	56
2.4.5	Pricing dans le cas à deux facteurs . . . . .	60
2.5	Heath-Jarrow-Morton Framework . . . . .	62
2.5.1	Condition de non-arbitrage . . . . .	63
2.5.2	Formulation des taux court avec HJM . . . . .	64
2.5.3	Taux forward de Vasicek . . . . .	65
2.6	Courbe des taux . . . . .	65
2.6.1	Interpolation . . . . .	66
2.6.2	Bootstraping . . . . .	66
2.6.3	Familles de polynôme exponentiel . . . . .	67
2.6.4	Méthodes utilisées par pays . . . . .	69
<b>3</b>	<b>Valorisation sur action</b>	<b>70</b>
3.1	Modèle à volatilité constante . . . . .	70
3.1.1	Modèle de Black and Scholes . . . . .	70
3.1.2	Grecques et hedging . . . . .	73
3.2	Surface de volatilité . . . . .	79
3.2.1	Définition et problème de calibration . . . . .	79
3.2.2	Volatilité implicite forward . . . . .	80
3.3	Modèle à volatilité locale . . . . .	83
3.3.1	Diffusion de Dupire . . . . .	83

3.3.2	Volatilité implicite dans le modèle de Dupire . . . . .	85
3.3.3	Problème de calibration . . . . .	87
3.3.4	Modèle shifted-lognormal . . . . .	90
3.3.5	Modèle CEV . . . . .	91
3.4	Volatilité stochastique . . . . .	92
3.4.1	Lien avec la volatilité implicite . . . . .	95
3.4.2	Modèle de Heston . . . . .	98
3.4.3	Modèle de Wu et Zhang . . . . .	102
3.4.4	Modèle SABR . . . . .	105
3.5	Dérivées sur volatilité . . . . .	108
3.5.1	Swap de variance . . . . .	108
3.5.2	Option sur volatilité . . . . .	110
3.5.3	VIX . . . . .	112
3.6	Calcul de Malliavin . . . . .	113
3.6.1	Eléments de théorie . . . . .	113
3.6.1.1	Chaos de Wiener . . . . .	114
3.6.1.2	Intégrale multiple stochastique . . . . .	116
3.6.1.3	Dérivée de Malliavin . . . . .	119
3.6.1.4	Intégrale de Skorohod . . . . .	122
3.6.2	Application à la finance . . . . .	125
3.6.2.1	Calcul des sensibilités . . . . .	126
3.6.2.2	Sensibilité pour option asiatique . . . . .	130
3.6.2.3	Raffinement du hedging . . . . .	132
<b>4</b>	<b>Introduction au risque de crédit</b>	<b>132</b>
4.1	Processus ponctuel . . . . .	132
4.2	Obligation en cas de défaut . . . . .	134
4.3	Modèle de firme . . . . .	137
4.3.1	Modèle de Merton . . . . .	138
4.3.2	Modèle de Black et Cox . . . . .	138
4.4	Credit Default Swap . . . . .	139
4.5	Credit Valuation Adjustment . . . . .	141
<b>5</b>	<b>Mesure de risque</b>	<b>144</b>
5.1	Value At Risk . . . . .	145
5.2	Expected Shortfall . . . . .	148
5.3	Backtesting . . . . .	148
5.4	Drawdown . . . . .	149

<b>6 Appendix</b>	<b>155</b>
6.1 Equation différentielle . . . . .	155
6.1.1 Equation du premier ordre . . . . .	155
6.1.2 Equation d'ordre $n$ . . . . .	155
6.1.3 Equation de Riccati . . . . .	156
6.2 Calcul stochastique et applications . . . . .	157
6.2.1 Théorème de Girsanov . . . . .	157
6.2.2 Changement de numéraire . . . . .	160
6.2.3 Théorème de Feynman-Kac . . . . .	161
6.3 Pricing d'option par transformée de Fourier . . . . .	162
6.4 Approximation d'intégrale . . . . .	164
6.5 Optimisation non linéaire . . . . .	167
6.5.1 Descente de gradient . . . . .	167
6.5.2 Algorithme génétique . . . . .	172
<b>7 Références</b>	<b>173</b>

# 1 Présentation des marchés financiers

## 1.1 Marché financier

### 1.1.1 Organisé

Un marché *organisé*, ou *bourse*, est une plateforme permettant aux acheteurs et aux vendeurs de produits financiers d'effectuer des transactions, d'exécuter des ordres de bourses, via un intermédiaire qui transmet l'ordre à un membre officiel de la bourse. La *chambre de compensation* est un pilier central dans la bourse, elle intervient comme contrepartie entre acheteurs et vendeurs en garantissant la bonne fin des opérations. Il ne s'agit ni plus ni moins que de l'acheteur de tous les vendeurs et du vendeur de tous les acheteurs. L'établissement [Clearnet](#) s'occupe du bon fonctionnement de l'ensemble des bourses européennes (Euronext).

### 1.1.2 De gré à fré

Un marché de *gré à gré* ou *Other-The-Counter* (OTC) est un marché sur lequel les transactions se font directement entre un acheteur et un vendeur. La principale différence avec la bourse est donc qu'il n'existe pas de chambre de compensation permettant d'éviter le risque de contrepartie. Les raisons de l'existence de cette place sont multiples. Notons déjà qu'une partie des sociétés qui y sont présentes le sont car elles sont trop petites, ne remplissent pas les critères pour être cotées sur les marchés organisés. De plus, comme les contrats sont établis directement entre les deux parties, cela permet d'acheter (ou vendre) un produit correspondant bien plus à ses besoins, comme par exemple pour couvrir un risque de change. Par construction, ce marché est moins transparent que la bourse.

### 1.1.3 Acteurs

(Presque) n'importe qui peut intervenir sur les marchés financiers, nous pouvons tout de même classifier les intervenants en 3 grandes catégories, les *hedgers*, les *spéculateurs* et les *market makers*. Pour les deux premières catégories, sans grandes surprises, il s'agit de gestionnaires intervenant respectivement pour se couvrir d'un certain risque et, pour tirer un profit suite à une transaction. Les market makers sont présents pour assurer la liquidité du marché. Ils offrent une contrepartie à tout acheteur ou vendeur sur certains titres souvent illiquides. Ces agents sont en général rémunérés par les bourses qui souhaitent offrir à leur client le plus de liquidité possible.

#### 1.1.4 Coûts de transactions

Si l'on achète ou vend un produit financier, quel qu'il soit, il faut rémunérer les intermédiaires, personne ne travaille gratuitement ! On ne peut pas donner ici de pourcentage précis de coût, cela dépend du volume, du montant du produit financier que l'on traite et également de votre broker, plus on achète, plus le pourcentage va être avantageux. Il faut donc bien se renseigner sur les montants demandés par votre broker et les prendre en compte dans votre stratégie.

### 1.2 Produits financiers

#### 1.2.1 Action

Une action (*share*, ou *stock*) est un *titre de propriété* d'une entreprise conférant à son détenteur (l'actionnaire) un droit de vote dans la politique de gestion de l'entreprise ainsi que des dividendes. Il s'agit donc d'une fraction du capital social de l'entreprise émise afin de se financer sur les marchés. La valeur initiale de l'action est déterminée par la société, ensuite, elle évolue sur les marchés en fonction de l'offre et de la demande.

Il existe deux types d'action, *ordinaire* ou *préférentielle*. Le détenteur d'une action préférentielle a un droit de priorité par rapport au détenteur d'une action ordinaire, une société n'a pas le droit de verser des dividendes aux actionnaires ordinaires tant que les privilégiés n'ont pas été rémunérés. La deuxième priorité intervient en cas de liquidation de l'entreprise, le détenteur bénéficiera d'un remboursement prioritaire pour la valeur de chacune de ses actions.

Enormément de facteurs rentrent en compte dans la formation des prix, en premier lieu les fondamentaux de l'entreprise, notons les variations de taux de changes, les annonces économiques/macroéconomiques, les décisions gouvernementales ainsi que les mouvements purement spéculatifs. Un simple ordre passé sur le marché, disons de vente, quelle qu'en soit la raison, s'il est relativement important, peut entraîner à lui seul un effet de panique et conduire à une vente de la part des autres détenteurs de la même action, on parle alors de *price impact*.

#### 1.2.2 Indice

Un *indice* est un indicateur proposé par une bourse pour évaluer les performances d'un pays, d'un secteur.

Citons quelques exemples parisiens. Le **CAC 40** est le principal indice de la bourse de Paris (Euronext Paris), initialement, l'acronyme CAC signifiait 'Compa-

gnie des Agents de Change', aujourd'hui 'Cotation Assistée en Continu'. La valeur de l'indice est une pondération de chacun des 40 titres de sociétés en fonction de leur capitalisation flottante. Sa composition est mise à jour tous les trimestres. Quand une société n'est plus cotée, elle est remplacée par une des valeurs du CAC Next 20 en fonction de la liquidité de son titre, capitalisation boursière, volumes échangés, etc. Enfin, notons que sa cotation est réactualisée toutes les 15 secondes. Le CAC Next 20 regroupe les vingt valeurs dont l'importance suit celle des valeurs composant le CAC 40. Le SBF 120 (Société des Bourses Françaises) est composé du CAC 40 et des 80 valeurs les plus liquides à Paris parmi les 200 premières capitalisations boursières françaises. Il existe également le SBF 250 dont on comprend facilement la composition. Notons tout de même que le SBF 80 est quant à lui les 80 valeurs suivant le CAC 40, donc hors CAC 40.

Parmi les indices internationaux nous pouvons également noter le [S&P 500](#) (Standard & Poors, filiale de McGraw-Hill), composé des 500 plus grandes capitalisations boursières des bourses américaines. Comme précédemment, il existe des variantes, S&P 100, 400 et 600. Et toujours comme précédemment, il est pondéré par la capitalisation boursière de chacune des ses composantes. Pour une plus grande diversité sur les marchés américains on pourra regarder le Russell 1000, 2000, 3000 ou le Wilshire 5000.

Pour voyager un peu dans différents pays, notons que l'indice principal en Allemagne est le DAX (30 valeurs), à Londres le FTSE 100 ('Footsie'), Hong-Kong le Hang Seng Index, Taiwan le TSEC et le SSE Composite Index pour Shangai.

Les indices sont donc donnés par,

$$I(t) = C \sum_{i=1}^N \alpha_i x_{i,t} \quad (1.1)$$

où  $N$  est le nombre de valeurs composant l'indice,  $x_{i,t}$  la valeur de l'action  $i$  à l'instant  $t$ ,  $\alpha_i$  le nombre d'actions  $i$  en circulation et  $C$  est une constante de normalisation à une valeur de référence (par exemple 100) à une date bien précise.

On peut conclure cette courte présentation en notant que le DJIA (Dow Jones Industrial Average), indice des 30 plus grosses valeurs du NYSE et du NASDAQ, tout comme le Nikkéi 225 à Tokyo ne sont pas pondérés par la capitalisation boursière de leurs composantes, mais par la valeur des actions. Le DJIA est le plus vieil indice boursier au monde, 1896, il était calculé à la main toutes les heures, on peut donc comprendre en partie pourquoi il est pondéré ainsi.

## 1.3 Matière Première

Les *commodities*, c'est à dire les matières premières, prennent elles aussi une importante place dans les marchés organisés, tant pour les besoins de gestion des risques que pour leur qualité de valeur refuge. Pour le secteur de l'énergie on pourra citer le baril de pétrole brut, le gaz naturel ou l'électricité ; pour les matériaux précieux ou industriels, l'or, l'argent, le platinium, le cuivre ou le palladium ; pour les grains, le maïs, le soja, l'avoine, le riz ou le blé ; les viandes comme les bovins vivants, bovins d'engraissement, le porc maigre, ou la poitrine de porc ; et autres comme le bois, le coton, le caoutchouc, le jus d'orange, le sucre, le café, le cacao, le lait, etc.

### 1.3.1 Taux de change

Nous aurions pu mettre cette section dans le chapitre sur les marchés financiers puisque l'ensemble des *devises* qui sont échangées les unes contre les autres à un certain *taux de change* sont cotées sur le *Forex*, contraction de Foreign Exchange. Ce marché mondial est le deuxième marché financier en terme de volume échangé derrière celui des taux d'intérêts, il est ouvert 24h/24h du fait des décalages horaires entre les bourses européenne, américaine, australienne et japonaise, tout de même fermées le week-end ; fermant à 22h GMT le vendredi à New-York et ouvrant à 22h GMT avec la bourse australienne.

Deux types de devises se distinguent, celle fixée par un Etat, on parle alors d'exotique comme le yuan chinois, et celle flottante, valorisée par l'offre et la demande sur les marchés, comme l'euro ou le dollar américain.

Un taux de change est toujours exprimé de la manière suivante  $X/Y=z$ , en d'autres termes, 1 X vaut  $z$  Y. La devise de gauche est la *devise de base* et celle de droite la *devise de contrepartie*. Les taux de changes varient jusqu'à la quatrième décimale, ce que l'on appelle le *pip*.

### 1.3.2 Tracker

Un *tracker* a pour but de *répliquer* un produit financier bien précis. Il peut porter sur un indice actions, obligataire ou encore de matières premières. Si l'on prend l'indice CAC 40, il n'est bien sûr pas possible d'acheter du CAC 40, on pourrait quand même avoir en portefeuille toutes les composantes du CAC 40 en prenant garde de bien réactualiser son portefeuille suite aux sorties et rentrées dans le CAC, d'avoir le même montant que la pondération, cela peut s'avérer bien compliqué. L'achat d'un tracker sur le CAC 40, par exemple le Lyxor ETF CAC 40

permet de simplifier tout cela en achetant un seul produit, au passage cela diminue forcément les coûts de commissions, les *fees*. Le **SPDR S&P 500**(Standard & Poor's Depositary Receipts), communément appelé Spyder S&P 500, est naturellement le tracker sur le S&P 500 et le SPDR Gold Trust qui réplique la performance de l'or physique sont les deux trackers les plus échangés au monde avec des encours respectifs de 76.5 et 77.5 milliards de dollars au 08/11. Les gérants d'un fond indiciel (Exchange traded funds) sont tenus d'avoir en portefeuille toutes les composantes de l'indice en les pondérant, et pour le cas de l'or, de stocker l'équivalent en or dans leurs coffres.

### 1.3.3 Contrats forward et future

Les contrats *forward* et *future* sont des *contrats à termes*, ils s'engagent fermement à acheter ou à vendre à une échéance définie  $t = T$ , à un prix fixé, une certaine quantité d'un actif financier le *sous-jacent*, à la date  $t = 0$ . La différence entre un contrat forward et un contrat future est que le premier est OTC tandis que le second s'échange sur les marchés organisés. Le 'sûr-mesure' proposé par le marché de gré à gré s'expose naturellement au risque de liquidité. Le sous-jacent du contrat peut être physique, or, pétrole, etc., une action, un indice, taux d'intérêt, devises, obligations. Il s'agit initialement d'un produit destiné à la gestion des risques.

### 1.3.4 Option

Les options les plus connues sont le *call* et le *put européens*, donnant respectivement le droit d'acheter et le droit de vendre un actif sous-jacent à une date déterminée à l'avance, moyennant le versement d'une prime. La différence entre ces deux types de produits financiers avec les forward ou future est qu'ils confèrent le droit, et non l'obligation d'exercer son contrat à maturité. Ces deux premiers exemples sont dit *vanille*. Les options *exotiques* sont plus complexes et la liste complète serait trop longue à citer ici. Rapidement, les options *américaines* qui donnent le droit d'exercer son option à n'importe quelle date entre le moment d'achat et la maturité ; l'option *bermuda* qui donne le droit d'exercer son contrat à plusieurs dates bien fixées ; une option *asiatique* paye la valeur moyenne du sous jacent durant un certain intervalle de temps ; option *lookback* paye n'importe quel prix atteint par l'action au cours de la période, on choisira naturellement celui qui maximise le gain ; option *as-you-like-this*, donne le droit de choisir si elle deviendra un call ou un put.

### 1.3.5 Obligation

Une *obligation* est un *titre de créance négociable* émis par une société ou une collectivité publique, principalement échangée sur les marchés de gré à gré. Les paramètres rentrant dans le calcul d'une obligation sont la date d'émission, la date d'échéance, le taux d'intérêt, la devise dans laquelle elle est émise et la périodicité du coupon. Les modalités de remboursement et le mode de rémunération des prêteurs sont fixés contractuellement, la rémunération pouvant être fixe ou variable, indexée alors sur un taux d'intérêt et non sur le résultat de l'entreprise. La valeur nominale permet le calcul des intérêts (coupon) que l'emprunteur s'engage à verser à des échéances fixées à l'avance. Le marché obligataire est une des principales raisons justifiant l'importance des agences de rating, dont les trois principales sont S&P, Moody's et Fitch. Le système de notation de S&P va du fameux triple A : AAA pour une forte capacité à rembourser au C pour peu d'espoir de recouvrement et D pour défaut de paiement. On peut classifier ces produits en 4 catégories bien distinctes :

- Un Etat dans sa propre devise, *emprunt d'Etat*.
- Un Etat dans une autre devise que la sienne, *obligation souveraine*.
- Une entreprise du secteur public, un organisme public, une collectivité locale, *obligation du secteur public*.
- Une entreprise privée, une association, ou tout autre personne morale, *obligation corporate*.

Nous présentons maintenant un peu plus en détail le cas des obligations d'[Etat Américaines](#). Les obligations à plus courte échéance sont les Treasury Bill (T-Bill) avec une maturité allant de 1 mois à 1 an. De la même manière que les obligations zéro-coupons, ils ne versent pas d'intérêts avant l'échéance. Les maturités habituelles sont 1 mois, 3 mois ou 6 mois. Les T-Bills sont reconnus comme constituant les obligations du trésor les moins risquées aux États-Unis par les investisseurs.

Les Treasury Notes (T-Notes) sont à échéances moyennes 2, 5 et 10 ans et leur rémunération est assurée par un coupon payé tous les six mois au souscripteur. La T-Note à échéance de 10 ans est devenue la valeur la plus fréquemment citée dans les études des performances du marché obligataire américain.

Enfin, les obligations avec la plus grande maturité sont les Treasury Bond (T-Bond), variant entre 10 et 30 ans ils sont également rémunérés par un coupon payé tous les six mois. Le revenu que les détenteurs perçoivent possède en outre l'avantage de n'être imposé qu'au niveau fédéral. Les équivalents [Français](#) des obligations à courte, moyenne et longue maturité sont respectivement les BTF

(Bons du Trésor à taux Fixe et à intérêt précompté), BTAN (Bons du Trésor à intérêts ANnuels) et les OAT (Obligations Assimilables du Trésor).

### 1.3.6 Dérivé de crédit

Les dérivés de crédit sont traitées sur les marchés OTC et par définition, écrites sur des produits sensibles au crédit (obligations, etc.). Elles peuvent être utilisées en vue de transférer un risque de contrepartie vers un autre pour contrôler son exposition. Comme pour les dérivés 'classiques', il existe deux type de dérivé de crédit, celle dite vanille, et celle dite exotique. Pour les vanilles nous pouvons par exemple citer les crédit défaut swap (CDS) single name (i.e. ne protégeant que sur un seul produit) et les collateralized debt obligation (CDO) sur indices. Pour les exotiques, nous avons par exemple les CDS sur zéro-coupons et les CDO<sup>2</sup>. Nous reviendrons plus en détail sur les CDS à la fin de ce cours.

## 1.4 Comité de Bâle

La création du comité de Bâle se fait en 1974 sous l'impulsion des régulateurs financiers, plus précisément les directeurs des banques centrales des principales puissances économiques. Le comité se réunit à la banque des règlements internationaux (BIS) à Bâle, Suisse, souvent appelée la banque centrale des banques centrales. Le but affiché est d'avoir un framework unifié en terme de régularisation pour éviter d'une part les risques d'une trop forte exposition aux marchés émergents et, d'autre part, avoir un certain pacte de non-agression entre les différentes banques centrales. Le premier rapport important se fera en 1988 par la publication des accords de Bâle 1 (Basel I) [Ba88]. Cette première publication ne concerne uniquement les produits de risques de crédits ainsi que la méthode à utiliser pour calculer le capital minimum requis en fonction de l'exposition de la banque. Le fameux ratio de Cooke y est écrit ainsi que le ratio asset to capital multiple. Ce dernier doit être inférieur à 20%,

$$\frac{\text{Total des actifs}}{\text{Capital}} \leq 20\%.$$

Le ratio de Cooke, également connu comme étant le ratio de risque capital est la somme des actifs détenus par l'entité divisée par leurs facteurs de risques respectifs, il doit être au moins supérieur à 8%

$$\frac{\text{Capital}}{\text{facteurs de risque des actifs}} \geq 8\%.$$

Les actifs non risqués sont par exemple l'or, les obligations émises par les plus grandes puissances économiques (et les plus saines !), leurs poids sont fixés à 0. Les banques des pays les puissants économiquement ont un risque de 20%, etc. L'une des raisons de l'essor du risque management provient du comité de Bâle lui même.

En effet, en 1996, l'expansion des activités de trading au sein des firmes ont poussé le comité à publier un amendement [Ba96] proposant aux banques de soit, utiliser un modèle de contrôle de risque proposé par le rapport (méthode de mesure standardisé) soit, de développer en interne un modèle. Bien entendu, ce modèle doit se plier à certaines règles. Les méthodes standardisées sont différentes selon la classe d'actif regardé, taux d'intérêt, equity, commodities ou Forex. Les obligations ont un risque prédéfini selon la maturité et le pays, les ETF sur indice un risque de 2%, une certaine méthodologie pour la couverture en delta et en gamma est proposée, etc. Pour que le modèle interne soit éligible, il faut naturellement qu'une équipe de risk management soit mise en place. Un back test robuste doit être implémenté, au minimum sur un an. Les stress tests doivent être suffisamment complets pour couvrir des événements extrêmes, ils doivent être à la fois quantitatifs et qualitatifs, incorporer les risques marché ainsi que de liquidité. Des calculs de Var avec un indice de confiance de 99% doivent être fait quotidiennement, etc. Il est possible d'avoir quelque chose de bien moins contraignant en faisant soit même son modèle, tout en respectant les recommandations du rapport que d'utiliser l'approche standardisée, c'est pourquoi les équipes de risk management ont considérablement gonflé depuis lors.

Le comité de Bâle publiera un nouveau rapport en 2004 [Ba05] pour répondre aux imperfections du premier rapport ainsi que pour s'adapter à l'évolution des marchés. Bâle II repose sur trois piliers. Le premier étant l'exigence de fonds propres, le deuxième pilier les procédures de surveillance et enfin, le troisième un effort de transparence et de discipline. En ce qui concerne le pilier I, la méthode de calcul des risques a été modifiée. Par exemple, pour le risque de crédit, nous avons l'approche standardisée, l'approche IRBA (Foundation Internal Ratings Based Approach) et l'IRBA avancé. Conformément à ces approches, les banques doivent en plus utiliser les rating fournis par les agences de rating. Sous IRBA avancé, les calculs de LGD (Loss Given Default) et EAD (Exposure At Defaut) sont calculés en interne.

La crise depuis 2007 a mis en avant la non-adéquation de la réforme Bâle II en cas d'événements extrêmes. Bâle III est en fait composé de deux amendements [Ba11] et [Ba13] le premier portant sur le capital et l'autre, portant sur la liquidité.

Hormis les exigences accrues, nous pouvons noter l'introduction d'un ratio de levier devant être supérieur à 6%,

$$\frac{\text{Capitaux propres}}{\text{Dettes}} > 6\%.$$

En plus des provisions en cas de risque de défaut d'une contrepartie, les banques doivent pouvoir couvrir le risque de perte mark-to-market sur le risque de contrepartie attendu d'un dérivé OTC. Il s'agit des pertes CVA (Credit Value Adjustments).

Le deuxième amendement porte sur le risque de liquidité. Le but annoncé du ratio de couverture en liquidité (LCR) est de permettre une résistance à court terme du risque de liquidité des banques. Il s'agit d'avoir un stock d'actifs non gelé et suffisamment liquide pour pouvoir être converti instantanément en cash pour couvrir une perte nette de trésorerie durant 30 jours. Le NSFR (taux net de financement stable) est le rapport entre le montant de financement stable par rapport au montant de financement stable requis sur un 1 an, il doit être supérieur à 1.

## 1.5 Risques

Tout comme les actions, l'achat d'une obligation d'Etat ou d'entreprise comporte des risques. Nous pouvons lister certains de ces différents risques :

- Risque de taux d'intérêt
- Risque de crédits
- Risque d'inflation
- Risque de crédit
- Risque de taxe
- Risque de régulation.

### 1.5.1 Risque de taux d'intérêt

Le risque de taux d'intérêt est certainement le plus évident et l'un des plus importants. Par définition, une obligation est indexée sur les taux d'intérêts, les cash flows peuvent donc être plus moins ou intéressants selon le niveau du taux sous-jacent et, plus particulièrement, sont sensibles aux variations instantanées des taux. Les produits typiques de fixed income bougent dans la direction opposée des variations de taux, si le taux diminue, le prix du produit de fixed income va monter, et inversement. Ce risque apparaît dans le cas où l'investisseur ne compte pas aller

à terme du contrat, une variation peut donc amener à une perte de capital. Ce risque de taux d'intérêt est mesuré par la duration.

### 1.5.2 Risque de crédit

Dans les modèles de valorisations que nous allons étudier, il est sous entendu que les obligations achetées arriveront à terme et payeront le principal. Cela n'est bien entendu pas le cas à chaque fois. En effet, il est tout à fait possible que l'émetteur de l'obligation ne puisse pas rembourser en partie ou la totalité de sa dette. Ce risque est appelé risque de crédit et n'existe que pour ce type d'instruments financiers, pas pour les marchés actions ou Forex par exemple. Si l'émetteur de la dette ne peut pas rembourser le principal, nous dirons qu'il fait défaut. Bien entendu, ce type de risque apparaît minime dans le cas d'obligations émises par le gouvernement américain ou allemand, et, bien plus important pour certains pays de la zone Euro, de pays émergents ou d'entreprises. Néanmoins, la probabilité de défaut de Lehman Brother était notée extrêmement faible, alors que finalement, ce n'était pas le cas. Même pour des produits, à priori, à faible risque de crédit, il est important de prendre en compte cette notion. En effet, la plupart des institutions financières comme des institutions d'Etat possèdent des produits dit toxiques, même si en soit, elles ne sont pas sujettes à défaut, il existe un risque de contagion et donc, une institution sans risque le deviendra.

### 1.5.3 Risque d'inflation

Le risque d'inflation n'est pas un risque de perte en soit, c'est un risque de voir son cash flow futur diminuer de valeur. L'inflation est définie en terme de pourcentage de variation sur un même panier de biens. L'indice de référence est le CPI, Consumer Price Index. Pour comprendre ce risque, il suffit de réaliser que si l'inflation augmente de manière importante, le paiement du coupon, même s'il n'a pas changé de valeur par rapport à ce qui est prévu, ne vaudra plus rien. Bien sûr, ce risque augmente pour des obligations à long-terme, il est peu probable de voir une augmentation de l'inflation significative sur 1 mois. Pour se couvrir des risques sur inflation, il existe des produit dérivés sur inflation, plus précisément, le sous-jacent est le CPI. Egalement, nous pouvons nous prémunir en achetant des produits comme des matières premières qui elles, sont sensibles à l'inflation.

#### 1.5.4 Risque de taux de change

Une obligation émise en monnaie étrangère, va avoir, comme pour n'importe quel autre produit financier, un risque de change. L'obligation va émettre des cash flow dans une certaine devise, qu'il va falloir transcrire dans la monnaie domestique. Supposons que nous achetons une obligation cotée en Yen, alors si le Yen se déprécie par rapport à l'euro, le montant gagné va être moindre, voir, négatif en cas de forte dépréciation. A l'inverse, si le Yen s'apprécie par rapport à l'euro, le montant reçu va être supérieur. Bien entendu, en plus d'être exposé au risque de change, nous sommes également exposés au risque de variation du taux étranger.

#### 1.5.5 Risque de taxe

Pour la plupart des produits de fixed income, ils sont taxables par rapport au taux en vigueur dans le pays considéré. De plus, ils sont sujets à des variations de régulation, d'imposition, menées par le gouvernement étranger. On peut parler de risque de taxe tout comme de risque de politique. Par exemple, en juin 1980, Battery Park City Authority (l'autorité supervisant le développement de Battery Park City, quartier situé à l'extrême sud-ouest de l'île de Manhattan à New York, États-Unis) a émis une dette de 97.315\$ pour lancer des constructions immobilières. Au moment de l'émission, les autorités légales ont considéré que ces notes seraient exemptées d'impôt fédéral. En novembre 1980, l'IRS (Internal Revenue Service), l'agence du gouvernement des États-Unis qui collecte l'impôt s'oppose à l'exonération de taxe sur les coupons, entraînant une baisse de leur valeur.

## 2 Taux d'intérêt

### 2.1 Zoologie

#### 2.1.1 Obligation zéro-coupon

Un dollar d'aujourd'hui vaudra toujours plus qu'un dollar de demain. Un zéro-coupon de maturité  $T$ , est un titre versant 1 dollar à la date  $T$ . La valeur faciale (ou nominal) de l'obligation est  $N = 1\$$ . Nous noterons  $B(t, T)$  le prix du zéro-coupon de maturité  $T$  à l'instant  $t$ , et donc,  $B(T, T) = 1$ . Nous supposerons toute la suite que le prix de l'obligation  $B(t, T)$  est processus positif strictement et adapté sur un espace de probabilité filtré  $(\Omega, \mathcal{F}, \mathbb{P})$ .

### 2.1.2 Structure par termes

Le rendement à maturité (yield-to-maturity) d'un zéro-coupon de maturité  $T$  est le processus adapté vérifiant,

$$R(t, T) = -\frac{1}{T-t} \ln B(t, T), \quad 0 \leq t \leq T. \quad (2.1)$$

La structure par terme des taux, également appelée la courbe des taux (yield curve) est la fonction qui donne les différents taux de la date  $t$  en fonction de leur maturité  $\theta$ ,  $\theta \rightarrow R(t, \theta)$ .

Connaissant le rendement à maturité, le prix de l'obligation est déterminé de manière unique par,

$$B(t, T) = e^{-R(t, T)(T-t)}, \quad 0 \leq t \leq T. \quad (2.2)$$

En pratique, la structure par termes des taux d'intérêts est déduite du prix de plusieurs instruments financiers très liquides, comme les T-bills, T-bond, swap et future.

### 2.1.3 Taux d'intérêt forward

Un contrat forward sur taux d'intérêt est le contrat effectué à l'instant  $t$ , d'achat à  $T_1 (>t)$  d'une obligation zéro-coupon de maturité  $T_2 (>T_1)$ . Par un simple argument de non-arbitrage, nous voyons que le prix du contrat forward est  $B(T_1, T_2) = \frac{B(t, T_2)}{B(t, T_1)}$ . Le taux d'intérêt forward à l'instant  $t$ ,  $f(t, T_1, T_2)$ , doit donc intuitivement vérifier

$$\frac{B(t, T_2)}{B(t, T_1)} = e^{-f(t, T_1, T_2)(T_2-T_1)}, \quad t \leq T_1 \leq T_2. \quad (2.3)$$

En d'autres termes, il est donc donné par

$$f(t, T_1, T_2) = -\frac{1}{T_2 - T_1} \ln \frac{B(t, T_2)}{B(t, T_1)}. \quad (2.4)$$

Il s'agit du taux forward continu. Non continu, il s'écrit naturellement,

$$f(t, T_1, T_2) = \frac{1}{T_2 - T_1} \left( \frac{B(t, T_1)}{B(t, T_2)} - 1 \right), \quad (2.5)$$

Le taux d'intérêt forward instantané est la limite du taux forward lorsque  $T_1 = T$  et  $T_2 = T + \Delta T$ ,

$$F(t, T) = -\lim_{\Delta t \rightarrow 0} \frac{\ln B(t, T + \Delta T) - \ln B(t, T)}{\Delta T} = -\frac{1}{B(t, T)} \frac{\partial B}{\partial T}(t, T) \quad (2.6)$$

Il s'agit donc du rendement marginal obtenu en conservant un court instant supplémentaire l'obligation. En intégrant (2.6), nous obtenons le prix de l'obligation en fonction du taux forward instantané,

$$B(t, T) = \exp \left( - \int_t^T F(t, u) du \right) \quad (2.7)$$

Enfin, en utilisant (2.2) et (2.6), nous avons,

$$F(t, T) = \frac{\partial}{\partial T} (R(t, T)(T - t)) = R(t, T) + (T - t) \frac{\partial}{\partial T} R(t, T). \quad (2.8)$$

Soit,

$$R(t, T) = \frac{1}{T - t} \int_t^T F(t, u) du. \quad (2.9)$$

Nous voyons donc que le prix de l'obligation ainsi que son rendement peuvent être entièrement déterminés par le taux forward instantané.

#### 2.1.4 Taux d'intérêt court terme

Le taux d'intérêt spot, i.e. le taux d'intérêt instantané, est défini mathématiquement comme la limite du taux continu moyen quand la maturité tend vers 0,

$$r_t = \lim_{T \downarrow t} R(t, T) = R(t, t) = F(t, t). \quad (2.10)$$

Nous pouvons réécrire le prix de l'obligation en fonction du taux spot,

$$\begin{aligned} B(0, T) &= \exp \left( - \int_0^T F(0, u) du \right) \\ &= \exp \left( - \int_0^{\Delta t} F(0, u) du - \int_{\Delta t}^T F(\Delta t, u) du \right) \\ &= \exp \left( - \sum_{i=1}^{[T/\Delta t]} \int_{(i-1)\Delta t}^{\Delta t} F((i-1)\Delta t, u) du \right) \end{aligned}$$

en passant à la limite  $\Delta t \rightarrow 0$ ,

$$\begin{aligned} B(0, T) &= \exp \left( - \sum_{i=1}^{[T/\Delta t]} \int_{(i-1)\Delta t}^{\Delta t} r_{i\Delta t} du \right) \\ &= \exp \left( - \int_0^T r_u du \right). \end{aligned}$$

Ainsi,

$$B(t, T) = \exp \left( - \int_t^T r_u du \right). \quad (2.11)$$

Sous forme différentielle, nous avons,

$$\frac{dB(t, T)}{B(t, T)} = r_t dt, \quad (2.12)$$

avec comme condition terminale  $B(T, T) = 1$ .

### 2.1.5 Obligation à coupons fixe

Une obligation à coupons fixes (coupon-bearing bond), de maturité  $T_n$ , est un actif financier qui paye à son détenteur les montants  $c_1, \dots, c_n$  aux dates  $T_1 < \dots < T_n$ . Le prix de ce coupon est donné par la somme des cash flows actualisés,

$$B_n(t, T_n) = \sum_{i=1}^n B(t, T_i) c_i + B(t, T_n) N, \quad 0 \leq t \leq T_1. \quad (2.13)$$

$N$  est le nominal de l'obligation. La séquence  $\{T_1, T_2, \dots, T_n\}$  est appelée le tenor.

### 2.1.6 Duration et convexité

Rappelons que pour un zéro-coupon  $B(t, T)$ , le rendement  $R(t, T)$  est donné par,

$$B(t, T) = e^{-R(t, T)(T-t)}, \quad 0 \leq t \leq T.$$

Soit maintenant  $B_n$  le prix d'un coupon avec pour maturités  $T_1 < \dots < T_n$ , de coupons  $c_1, \dots, c_n$  et de nominal  $N$ . Nous supposons que  $N$  est compris dans les  $c_i$ , le prix est alors,

$$B_n(t) = \sum_{i=1}^n B(t, T_i) c_i.$$

Nous nous plaçons en  $t = 0$ , en notant  $y$  le rendement à maturité (YTM) et  $y(0, T_i) = y_i$ ,  $B_n(0) = B_n$ , nous pouvons réécrire le prix de l'obligation pour tout  $t \leq T_1$ ,

$$B_n = \sum_{i=1}^n c_i e^{-y_i T_i}.$$

La duration de Macaulay d'un coupon est définie par,

$$D_{Mac} := \frac{\sum_{i=1}^n T_i c_i e^{y_i T_i}}{B_n}. \quad (2.14)$$

Il s'agit donc de la moyenne des coupons aux dates  $T_1, \dots, T_n$ . En d'autre terme, la duration de Macaulay d'une obligation correspond à la période à l'issue de laquelle sa rentabilité n'est pas affectée par les variations de taux. Nous pouvons également noter que (2.14) nous donne la sensibilité relative du prix de l'obligation par rapport au YTM,

$$\frac{dB_n}{dy} = \frac{d}{dy} \left( \sum_{i=1}^n c_i e^{y_i T_i} \right) = -D_{Mac} B_n.$$

Nous définissons maintenant la convexité d'une obligation par,

$$C := \frac{dB_n^2}{dy^2} = \sum_{i=1}^n c_i e^{-y_i T_i} T_i^2.$$

Il s'agit de la sensibilité de la duration par rapport à une variation du rendement  $y$ .

### 2.1.7 Swap de taux d'intérêt

Un swap de taux d'intérêt permet d'échanger un taux variable (taux flottant) contre un taux fixe  $\kappa$  sur une période donnée.

La structure d'un swap de taux est donnée par la suite d'intervalles  $[T_1, T_2], \dots, [T_{n-1}, T_n]$  tel que  $T_i - T_{i-1} = \delta_i$ ,  $\delta_i$  une constante connue à l'avance,  $T_n$  la maturité du swap, un taux fixe,  $\kappa$  et, le nominal  $N$ . Nous notons  $\delta_i$ ,  $i = 2, \dots, n$ , et non  $\delta$ , en effet, bien que les dates de versement arrivent à intervalles réguliers, tous les 6 mois par exemple, nous n'avons pas exactement le même nombre de jours entre le  $i$ -ème 6mois et le  $i + 1$ -ème 6mois, simplement du fait du nombre de jours dans un mois (mois de février) ainsi que du nombre de jours fériés.

A l'instant  $T_i$ , le détenteur du contrat paye  $\kappa\delta_i N$  et, reçoit le taux flottant  $f(t, T_{i-1}, T_i)\delta_i N$ . Ce taux flottant est généralement le Libor dollar pour le dollar américain et l'Euribor pour l'euro, il s'agit des deux contrats les plus liquides sur le marché des taux. Le cashflow à l'instant  $T_i$  est donc,

$$(f(t, T_{i-1}, T_i) - \kappa)\delta_i N, \quad 0 \leq t \leq T_{i-1}.$$

Ainsi, en utilisant (2.5), nous trouvons,

$$N(B(t, T_{i-1}) - B(t, T_i) - \kappa\delta_i B(t, T_i)). \quad (2.15)$$

En faisant la somme cumulée de  $T_0$  à  $T_n$ , la valeur totale du swap est,

$$S_t = N(B(t, T_0) - B(t, T_n) - \kappa\delta_i \sum_{i=1}^n B(t, T_i)), \quad 0 \leq t \leq T_0. \quad (2.16)$$

Il reste à savoir comment déterminer la valeur 'juste' (fair value) du taux fixe  $\kappa$ . En supposant que tous les coupons  $B(t, T_i)$  soient déterminés par le marché, la valeur de  $\kappa$  est donnée par l'équilibre de (2.16),  $S_t = 0$ . Alors,

$$R_{\text{swap}}(t) = \frac{B(t, T_0) - B(t, T_n)}{\sum_{i=1}^n \delta_i B(t, T_i)}, \quad 0 \leq t \leq T_0.$$

$R_{\text{swap}}(t)$  est appelé le taux swap forward. Il est possible d'exprimer le payoff d'un swap ainsi que la valeur d'équilibre de  $\kappa$ ,  $R_{\text{swap}}$ , cela pourra servir pour pricer les swaptions. En revenant à (2.15) et en utilisant encore (2.5), nous avons comme cashflow du swap à l'instant  $t$ ,

$$N\delta_i B(t, T_i)(f(t, T_{i-1}, T_i) - \kappa).$$

La somme cumulée est,

$$S_t = N \sum_{i=1}^n \delta_i B(t, T_i)(f(t, T_{i-1}, T_i) - \kappa), \quad (2.17)$$

la valeur d'équilibre du taux fixe  $\kappa$  est alors,

$$R_{\text{swap}(t)} = \sum_{i=1}^n w_i(t) f(t, T_{i-1}, T_i),$$

avec comme poids  $w_i(t) = B(t, T_i) / \sum_{j=1}^n B(t, T_j)$ . Ces poids sont aléatoires, mais les résultats empiriques tendent à voire la variabilité de  $w_i$  petite comparée aux variations de  $f(t, T_{i-1}, T_i)$ .

### 2.1.8 Taux Libor

Rappelons que le taux forward  $f(t, T, S)$ ,  $0 \leq t \leq T \leq S$  est donné par la relation,

$$f(t, T, S) = -\frac{1}{T-S} \ln \frac{B(t, S)}{B(t, T)}. \quad (2.18)$$

Sur le marché Libor, les forward  $L(t, T, S)$ , pour un prêt sur  $[T, S]$ , sont donnés à la différence de (2.18) par,

$$1 + (S - T)L(t, T, S) = \frac{B(t, T)}{B(t, S)}.$$

Ainsi,

$$L(t, T, S) = \frac{1}{S - T} \left( \frac{B(t, T)}{B(t, S)} - 1 \right), \quad 0 \leq t \leq T < S. \quad (2.19)$$

Il s'agit en fait de la limite du taux forward instantané,

$$\begin{aligned} F(t, T) &= -\frac{1}{B(t, T)} \frac{\partial B(t, T)}{\partial T} \\ &= -\lim_{S \downarrow T} \frac{B(t, S) - B(t, T)}{(S - T)B(t, S)} \\ &= \lim_{S \downarrow T} L(t, T, S). \end{aligned}$$

Il convient tout particulièrement de s'attarder quelques secondes sur le taux Libor. Il s'agit du London interbank offered rate, soit, taux interbancaire pratiqué à Londres. C'est un taux fondamental, c'est le taux auquel les banques se prêtent entre elles, c'est pour cela qu'il est si important. Nous étudierons certains produits dérivés directement sur les taux Libor d'ailleurs. Nous présentons sur la figure (1) la courbe de taux forward du 19/08/13.

Remarquons que le taux Libor peut s'exprimer en fonction du prix forward  $Y_t = X_t/N_t$  avec  $N_t = B(t, S)/(S - T)$  jouant le rôle de numéraire et  $X_t = B(t, T) - B(t, S)$ . Le taux Libor sous la mesure forward  $\mathbb{Q}$  est une martingale avec  $\mathbb{Q}$  définie par,

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{1}{B(0, S)} \exp \left( - \int_0^S r_u du \right).$$

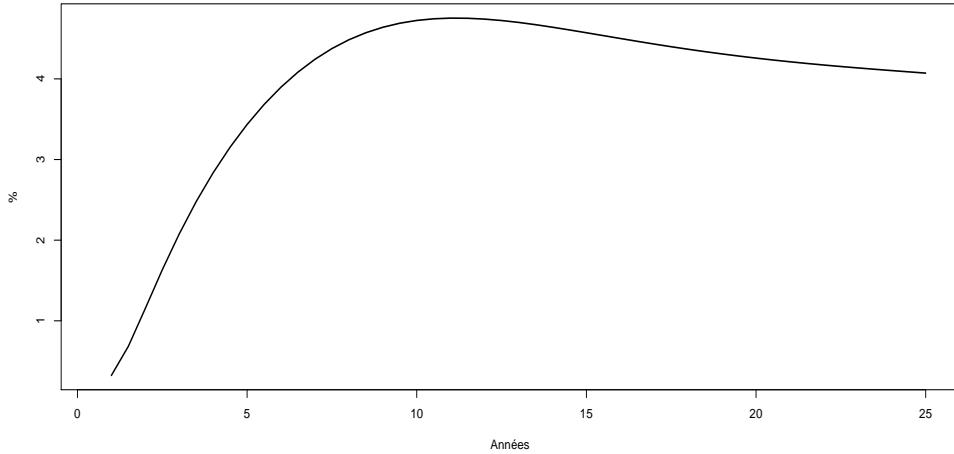


FIGURE 1 – Taux forward LIBOR  $T \mapsto F(t, T)$  du 19/08/13, maturité de 1 à 25 ans. Source : Bank of England.

Le swap de taux sur le Libor satisfait la même relation que précédemment (2.17) avec bien entendu le taux forward  $f(t, T_1, T_n)$  remplacé par le taux Libor  $L(t, T_k, T_{k+1})$ ,

$$\sum_{i=2}^n (T_i - T_{i-1}) B(t, T_i) (L(t, T_{i-1}, T_i) - S(t, T_1, T_n)) = 0,$$

où nous avons supposé que  $N = 1$ . Nous avons donc,

$$\begin{aligned}
 S(t, T_1, T_n) &= \frac{1}{B(t, T_1, T_n)} \sum_{i=2}^n (T_i - T_{i-1}) B(t, T_i) L(t, T_{i-1}, T_i) \\
 &= \frac{1}{B(t, T_1, T_n)} \sum_{i=2}^n B(t, T_i) \left( \frac{B(t, T_{i-1})}{B(t, T_i)} - 1 \right) \\
 &= \frac{1}{B(t, T_1, T_n)} \sum_{i=2}^n (B(t, T_{i-1}) - B(t, T_i)) \\
 &= \frac{B(t, T_1) - B(t, T_n)}{B(t, T_1, T_n)}. \tag{2.20}
 \end{aligned}$$

Notons pour conclure cette section que le swap de taux LIBOR peut être vu

comme  $Y_t = X_t/N_t$  avec le numéraire  $N_t = B(t, T_1, n)$  et  $X_t = B(t, T_1) - B(T_n)$ , sous la mesure forward  $\mathbb{Q}$  tel que,

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{B(T_1, T_1, T_n)}{B(0, T_1, T_n)} \exp\left(-\int_0^{T_1} r_u du\right).$$

### 2.1.9 Convention de marché

Si  $t$  et  $T$  sont deux dates exprimées en jour, mois, année, il n'est pas clair en quoi  $T - t$  doit être. Le marché évalue la fraction d'année entre  $t$  et  $T$ . Il existe une convention de calcul de jours. Connaissant le nombre de jour entre les deux dates  $t$  et  $T$ , nous avons directement,

$$T - t = \frac{\# \text{ jours entre } t \text{ et } T}{365}.$$

Insistons bien sur le fait que nous comptons tous les jours de l'année, et non seulement les jours ouvrés, sinon, nous aurions dû diviser par 261 et exclure les week-ends et jours fériés du calcul de l'intervalle. Pour ne pas à avoir à compter le nombre de jour, en notant  $T_1 = (d_1, m_1, y_1)$  et  $T_2 = (d_2, m_2, y_2)$  tel que  $T_1 < T_2$ ,

$$T_2 - T_1 = \frac{\min(d_2, 30) - \min(d_1, 30)}{360} + \frac{m_2 - m_1}{12} + y_2 - y_1.$$

Prenons par exemple  $T_1 = 25$  juillet 2013 et  $T_2$  le 30 janvier 2014, alors,

$$T_2 - T_1 = \frac{30 - 25}{360} + \frac{1 - 7}{12} + 2014 - 2013 = 0.51.$$

Il s'agit ici de la convention 30E/360. Il en existe beaucoup d'autre.

### 2.1.10 Intérêts courus, clean et dirty price

Les intérêts courus sont la fraction de l'intérêt annuel écoulé sur un titre. A l'instant  $t \in (T_{i-1}, T_i]$ , les intérêts courus sont,

$$AI(i, t) := c_i \frac{t - T_{i-1}}{T_i - T_{i-1}}$$

Le clean price est le prix de l'obligation sans les intérêts courus,

$$p_{\text{clean}}(t) := p(t) - AI(i, t), \quad T_{i-1} < t \leq T_i.$$

Autrement dit, le clean price est le prix auquel on achète le coupon à l'instant  $t \in (T_{i-1}, T_i]$ . Le dirty price est le prix incluant les intérêts courus,

$$p_{\text{dirty}}(t) := p_{\text{clean}}(t) + AI(i, t), \quad T_{i-1} < t \leq T_i.$$

### 2.1.11 Cap et floor

Un caplet verse à l'instant  $T + \delta_i$  la différence positive entre le taux à l'instant  $T$  et  $T + \delta$  et le strike  $\kappa$ . Son payoff est,

$$\text{Cpl}(t, T, T + \delta) = \delta_i(f(T, T, T + \delta_i) - \kappa)^+, \quad 0 \leq t < T. \quad (2.21)$$

Un cap est une bandelette de caplet. Pour le tenor  $T_0 < \dots < T_n$ , le payoff d'un cap à l'instant  $t < T_0$ ,

$$\text{Cp}_t = \sum_{i=2}^n \text{Cpl}(t, T_{i-1}, T_i).$$

Le cap garanti à son détenteur une protection contre la hausse des taux d'intérêts. Le taux à payer ne sera jamais supérieur au strike  $\kappa$ . On peut montrer que les cashflow (2.21) à l'instant  $T_i$  sont proportionnels aux cashflow d'un put à l'instant  $T_{i-1}$  sur une obligation de maturité  $T_i$  et de strike  $1/(1 + \delta_i \kappa)$  soit,

$$(1 + \delta_i \kappa) \left( \frac{1}{1 + \delta_i \kappa} - P(T_{i-1}, T_i) \right)^+.$$

Ce lien est très important puisqu'il permet de pricer des caps en utilisant les formules de pricing d'obligation que nous développerons plus tard. De plus, nous voyons bien qu'un cap est un pari à la baisse des taux comme évoqué précédemment.

Un floor est l'inverse d'un cap. Comme pour un cap, un floor est une bandelette de floorlets, de cashflow à l'instant  $t$  donné par,

$$\text{Flt}(t, T_{i-1}, T_i) = \delta_i(\kappa - F(T_{i-1}, T_{i-1}, T_i))^+, \quad 0 \leq t \leq T_0.$$

Le prix du floor est alors donné par,

$$\text{Fl}_t = \sum_{i=1}^n \text{Flt}(t, T_{i-1}, T_i), \quad 0 \leq t \leq T_0.$$

Il existe un équivalent à la relation de parité call-put pour les caps et les floors,

$$\text{Cp}_t - \text{Fl}_t = S_t.$$

où  $S_t$  est la valeur du swap à l'instant  $t$  de taux  $\kappa$ , de nominal 1 et de même discrétisation  $T_0 < T_1 \cdots < T_n$  que le cap et le floor.

Enfin, on dira que le cap/floor est à la monnaie (ATM) si,

$$\kappa = R_{\text{swap}}(0) = \frac{B(0, T_0) - B(0, T_n)}{\sum_{i=1}^n \delta_i B(0, T_i)}.$$

le taux swap forward. On dira que le cap (floor) est à l'intérieur de la monnaie si  $\kappa < R_{\text{swap}}(0)$  ( $\kappa > R_{\text{swap}}(0)$ ) et à l'extérieur de la monnaie si  $\kappa > R_{\text{swap}}(0)$  ( $\kappa < R_{\text{swap}}(0)$ )

### 2.1.12 Swaption

Une swaption est une option sur swap donnant le droit de vendre, put swaption ou d'acheter, call swaption, à une maturité fixée  $T$ , à un prix  $\kappa$  fixé à l'avance, le strike.

Rappelons que la valeur d'un swap de taux fixe  $\kappa$  commençant en  $T_0$  est,

$$\Pi_p(T_0, \kappa) = N \sum_{i=1}^n B(T_0, T_i) \delta_i (f(T_0, T_{i-1}, T_i) - \kappa).$$

Ainsi, le payoff d'un call swaption de strike  $\kappa$  est,

$$N \left( \sum_{i=1}^n B(T_0, T_i) \delta_i (f(T_0, T_{i-1}, T_i) - \kappa) \right)^+ \quad (2.22)$$

Notons que comme  $\Pi_p(T_0, R_{\text{swap}}(T_0)) = 0$ , nous pouvons réécrire le payoff (2.22) du call swaption,

$$N \delta_i (R_{\text{swap}}(T_0) - \kappa)^+ \sum_{i=1}^n B(T_0, T_i),$$

et du put swaption,

$$N \delta_i (\kappa - R_{\text{swap}}(T_0))^+ \sum_{i=1}^n B(T_0, T_i).$$

A l'instant  $t \leq T_0$ , le call (put) swaption de strike  $K$  est dit respectivement ATM, ITM et OTM (At, In, Out The Money) si,

$$\kappa = R_{\text{swap}}(t), \quad \kappa > (<)R_{\text{swap}}(t), \quad \kappa > (<)R_{\text{swap}}(t).$$

## 2.2 Modèle de taux court

### 2.2.1 Valorisation d'une obligation

Nous allons proposer deux méthodes pour obtenir la formule de valorisation des zéros-coupons sous forme d'une espérance conditionnelle et l'EDP des taux. La deuxième reprend l'approche fondatrice de Vasicek, [Va77], libre au lecteur de choisir celle avec laquelle il se sent le plus à l'aise.

Formellement, nous allons travailler sur l'espace de probabilité filtré  $(\Omega, \mathcal{F}, (\mathcal{F}_t, 0 \leq t \leq T), \mathbb{P})$ , la filtration étant celle du mouvement Brownien standard. Comme pour les modèles actions, nous allons supposer par hypothèse d'absence d'opportunité d'arbitrage l'existence d'une probabilité risque-neutre  $\mathbb{Q}$  telle que sous  $\mathbb{Q}$ , les prix des zéros-coupons actualisés  $B(t, T)/S_t^0$  pour tout  $0 \leq t \leq T$  sont des martingales. Le numéraire  $(S_t^0)_{t \in \mathbb{R}_+}$  est pris classiquement comme le processus de cash,

$$S_t^0 = \exp \left( - \int_t^T r_s ds \right), \quad t \leq T.$$

Ainsi, utilisant le fait que  $B(T, T) = 1$  et la propriété de martingale nous pouvons écrire,

$$\begin{aligned} \frac{B(t, T)}{S_t^0} &= \mathbb{E}_{\mathbb{Q}} \left( \frac{B(T, T)}{S_T^0} \middle| \mathcal{F}_t \right) \\ &= \mathbb{E}_{\mathbb{Q}} \left( \frac{1}{S_T^0} \middle| \mathcal{F}_t \right), \end{aligned}$$

Il est clair que  $S_t^0$  est  $\mathcal{F}_t$ -mesurable, alors,

$$\begin{aligned} B(t, T) &= S_t^0 \mathbb{E}_{\mathbb{Q}} \left( \frac{B(T, T)}{S_T^0} \middle| \mathcal{F}_t \right) \\ &= \mathbb{E}_{\mathbb{Q}} \left( \frac{S_t^0}{S_T^0} \middle| \mathcal{F}_t \right) \\ &= \mathbb{E}_{\mathbb{Q}} \left( \exp \left( - \int_t^T r_s ds \right) \middle| \mathcal{F}_t \right). \end{aligned} \tag{2.23}$$

Nous avons une première méthode de pricing des zéros-coupons que nous appellerons l'approche par martingale. Comme pour le cas du modèle de Black & Scholes, nous pouvons aussi obtenir une EDP reliant les prix, donc des zéros-coupons cette fois ci. Supposons pour le moment une dynamique générale du taux spot sous la probabilité historique  $\mathbb{P}$ ,

$$dr_t = \mu_r(t, r)dt + \sigma_r(r, t)dW_t \quad (2.24)$$

avec  $(W_t)_{t \in \mathbb{R}_+}$  un mouvement Brownien standard et  $\mu_r$  et  $\sigma_r$  deux fonctions vérifiant les conditions de Cauchy-Lipschitz.  $\mu_r$  est classiquement le drift du processus  $(r_t)_{t \in \mathbb{R}_+}$  et  $\sigma_r$  sa volatilité. Le prix zéro-coupon à l'instant  $t$  est une fonction de  $t$ ,  $r$  et de  $T$ ,  $B(t, T) = \mathcal{B}(t, r, T)$ . Le prix  $\mathcal{B}$  est une fonction réelle, supposée suffisamment régulière et de classe  $C^2$  afin de pouvoir appliquer le lemme d'Itô,

$$\begin{aligned} dB(t, T) &= \left( \frac{\partial \mathcal{B}}{\partial t} + \mu_r \frac{\partial \mathcal{B}}{\partial r} + \frac{1}{2} \sigma_r^2 \frac{\partial^2 \mathcal{B}}{\partial r^2} \right) (t, r, T) dt + \sigma_r \frac{\partial \mathcal{B}}{\partial r} (t, r, T) dW_t \\ &= B(t, T) \mu_B(t, r, T) dt + B(t, T) \sigma_B(t, r, T) dW_t \end{aligned} \quad (2.25)$$

avec,

$$\begin{aligned} \mu_B(t, r, T) &= \frac{1}{\mathcal{B}(t, r, T)} \left( \frac{\partial \mathcal{B}}{\partial t} + \mu_r \frac{\partial \mathcal{B}}{\partial r} + \frac{1}{2} \sigma_r^2 \frac{\partial^2 \mathcal{B}}{\partial r^2} \right) (t, r, T) \\ \sigma_B(t, r, T) &= \frac{1}{\mathcal{B}(t, r, T)} \sigma_r(t, r) \frac{\partial \mathcal{B}}{\partial r} (t, r, T). \end{aligned} \quad (2.26)$$

Comme précédemment, nous allons travailler sous probabilité risque neutre  $\mathbb{Q}$ . D'après le théorème de Girsanov, nous supposons alors l'existence de la prime de risque, fonction du temps et du taux spot,  $(\lambda(t, r))_{(t,r) \in \mathbb{R}_+ \times \mathbb{R}}$ , tel que,

$$\tilde{W}_t = W_t + \int_0^t \lambda(s, r) ds$$

soit un mouvement brownien sous  $\mathbb{Q}$ . L'équation de diffusion de  $r$  sous  $\mathbb{Q}$  est alors donnée par,

$$dr_t = (\mu_r(t, r) - \lambda(t, r)\sigma_r(t, r))dt + \sigma_r(t, r)d\tilde{W}_t.$$

La dynamique du prix zéro-coupons sous la mesure risque neutre  $\mathbb{Q}$  est alors,

$$dB(t, T) = B(t, T)(\mu_B(t, r, T) - \lambda(t, r)\sigma_B(t, r, T))dt + B(t, T)\sigma_B(t, r, T)d\tilde{W}_t. \quad (2.27)$$

Notons que le prix actualisé  $B(t, T)/S_t^0$  est une martingale sous  $\mathbb{Q}$ , le terme de tendance doit donc être égal à  $r_t$ , soit,

$$\mu_B(t, r, T) - \lambda(t, r)\sigma_B(t, r, T) = r_t. \quad (2.28)$$

En remplaçant avec les valeurs de  $\mu$  et  $\sigma$  trouvées en (2.26) nous obtenons,

$$\begin{aligned} \left( \frac{\partial \mathcal{B}}{\partial t} + \mu_r \frac{\partial \mathcal{B}}{\partial r} + \frac{1}{2} \sigma_r^2 \frac{\partial^2 \mathcal{B}}{\partial r^2} \right) - \lambda \sigma_r \frac{\partial \mathcal{B}}{\partial r} &= r \mathcal{B} \\ \frac{\partial \mathcal{B}}{\partial t} + (\mu_r - \sigma_r \lambda) \frac{\partial \mathcal{B}}{\partial r} + \frac{1}{2} \sigma_r^2 \frac{\partial^2 \mathcal{B}}{\partial r^2} - r \mathcal{B} &= 0. \end{aligned} \quad (2.29)$$

La condition terminale de l'EDP étant la même que précédemment,  $\mathcal{B}(T, r, T) = 1$ .

Notons qu'il est possible de retrouver l'EDP des taux sans faire intervenir explicitement la probabilité risque-neutre  $\mathbb{Q}$ . Les taux d'intérêts n'étant pas un titre négocié à proprement parler, nous ne pouvons pas le couvrir avec le titre, comme avec le sous-jacent d'une option sur equity. Nous allons donc supposer que nous achetons un zéro-coupon de valeur  $V_1$  et de maturité  $T_1$  et que nous shortons un autre zéro-coupon de valeur  $V_2$  de maturité  $T_2$ . La valeur du portefeuille est alors,

$$V = V_1 - V_2.$$

En conservant la dynamique (2.25), nous avons,

$$dV = (V_1 \mu_B(t, r, T_1) - V_2 \mu_B(t, r, T_2))dt + (V_1 \sigma_B(t, r, T_1) - V_2 \sigma_B(t, r, T_2))dW_t.$$

L'astuce consiste à prendre des montants bien particuliers pour  $V_1$  et  $V_2$  de telle sorte que l'aléa  $B$  disparaisse. Ce choix est,

$$V_1 = \frac{\sigma_B(t, r, T_2)}{\sigma_B(t, r, T_2) - \sigma_B(t, r, T_1)} V \quad \text{and} \quad V_2 = \frac{\sigma_B(t, r, T_1)}{\sigma_B(t, r, T_2) - \sigma_B(t, r, T_1)} V.$$

La diffusion de la valeur du portefeuille est alors,

$$dV = V \frac{\mu_B(t, r, T_1) \sigma_B(t, r, T_2) - \mu_B(t, r, T_2) \sigma_B(t, r, T_1)}{\sigma_B(t, r, T_2) - \sigma_B(t, r, T_1)} dt. \quad (2.30)$$

Le portefeuille est maintenant sans risque, il doit donc rapporter la valeur du taux d'intérêt,  $dV/V = r_t dt$ . En identifiant le terme de (2.30) avec  $r_t$  pour qu'il n'y ait pas d'opportunité d'arbitrage, nous avons,

$$\frac{\mu_B(t, r, T_1) - r_t}{\sigma_B(t, r, T_1)} = \frac{\mu_B(t, r, T_2) - r_t}{\sigma_B(t, r, T_2)}.$$

L'équation est valide quelle que soit la maturité  $T_i$  et nous noterons la quantité  $\lambda$ ,

$$\frac{\mu_B(t, r) - r}{\sigma_B(t, r)} = \lambda(t, r).$$

Pour conclure, nous n'avons qu'à remplacer les valeurs de  $\mu$  et  $\sigma$  des équations (2.26) pour obtenir l'EDP des taux,

$$\frac{\partial \mathcal{B}}{\partial t} + (\mu_r - \sigma_r \lambda) \frac{\partial \mathcal{B}}{\partial r} + \frac{1}{2} \sigma_r^2 \frac{\partial^2 \mathcal{B}}{\partial r^2} - r \mathcal{B} = 0.$$

Une application directe du théorème de Feyman-Kac (voir Appendix, 6.2.3) avec la condition terminale de notre prix  $\Phi(X_T) = \mathcal{B}(T, r, T) = 1$ , nous donne directement,

$$B(t, T) = \mathbb{E}_{\mathbb{Q}} \left( \exp \left( - \int_t^T r_s ds \right) \middle| \mathcal{F}_t \right).$$

Nous avons deux méthodes de pricing de zéro coupon avec taux spot, soit en résolvant l'EDP des taux, (2.29), soit, par le calcul de l'espérance conditionnelle sous  $\mathbb{Q}$  (2.23).

### 2.2.2 Modèle de Vasicek

La première diffusion que nous allons considérer, certainement la plus simple, est celle du processus d'Ornstein-Uhlenbeck. Le taux spot a la dynamique,

$$dr_t = \kappa(\theta - r_t)dt + \sigma dW_t.$$

où  $(W_t)_{t \in \mathbb{R}_+}$  est un mouvement Brownien standard. Il s'agit d'un processus ayant la caractéristique d'être 'mean-reverting', en d'autre terme, il oscille autour de sa valeur moyenne, ce qui est le cas général des taux spot, n'ayant pas de tendance bien définie à long terme.  $\kappa$ ,  $\theta$  et  $\sigma$  sont des constantes positives.  $\theta$  correspond à la valeur d'équilibre, la valeur moyenne de  $r$ ,  $\kappa$  est la force de retour à la moyenne, la vitesse à laquelle le processus va revenir à sa valeur d'équilibre. Prenons un exemple, supposons que  $r_t$  soit au dessus de sa valeur moyenne  $t$ , alors, le terme  $\theta - r_t$  va être négatif, soit une tendance négative, ce qui paraît logique.  $\kappa$ , va quant à lui donner plus ou moins d'importance à ce drift, plus il sera élevé, plus la convergence se fera rapidement, et réciproquement. Si  $r_t$  est à sa valeur d'équilibre, le drift est nul, il n'y a aucune raison de vouloir voir bouger le prix.  $\sigma$  est la volatilité, il va donner plus ou moins d'importance au bruit  $W$ . On présente

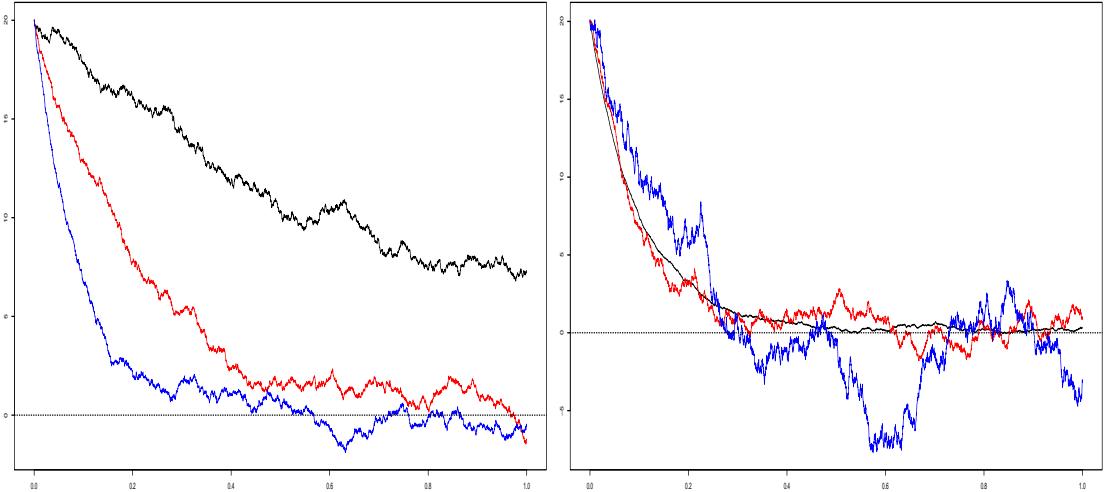


FIGURE 2 – 2\*3 trajectoires d'un processus d'Ornstein-Uhlenbeck avec même valeur d'équilibre  $\theta = 0$ . A gauche, 3 forces de retour à la moyenne différente et a droite, 3 volatilités différentes.

quelques réalisations de ce processus sur la figure (2.2.2).

Pour travailler sous  $\mathbb{Q}$ , de manière à conserver la même dynamique, nous prenons une prime de risque constante,  $\lambda$  tel que  $dW_t = d\tilde{W}_t - \lambda dt$ , alors,

$$\begin{aligned} dr_t &= \kappa(\theta - r_t)dt + \sigma(d\tilde{W}_t - \lambda dt) \\ &= \kappa(\theta - \sigma\lambda/\kappa - r_t)dt + \sigma d\tilde{W}_t \\ &= \kappa(\tilde{\theta} - r_t)dt + \sigma d\tilde{W}_t \end{aligned} \tag{2.31}$$

avec  $\tilde{\theta} = \theta - \sigma\lambda/\kappa$ .

**Proposition 2.1.** *La solution de (2.31) est,*

$$r_t = \tilde{\theta} + (r_0 - \tilde{\theta})e^{-\kappa t} + \sigma \int_0^t e^{-\kappa(t-s)} d\tilde{W}_s \tag{2.32}$$

**Preuve 2.1.** Posons  $f(x, t) = e^{\kappa t}r_t$ . Il est clair que la fonction est  $C^2$ , nous pouvons appliquer le lemme d'Itô,

$$\begin{aligned} df(x, t) &= \kappa e^{\kappa t}r_t dt + e^{\kappa t}dr_t \\ &= \kappa \tilde{\theta} e^{\kappa t} dt + \sigma e^{\kappa t} d\tilde{W}_t. \end{aligned}$$

En intégrant,

$$\begin{aligned} e^{\kappa t} r_t &= r_0 + \tilde{\theta}(1 - e^{\kappa t}) + \sigma \int_0^t e^{\kappa s} d\tilde{W}_s \\ r_t &= \tilde{\theta} + (r_0 - \tilde{\theta})e^{-\kappa t} + \sigma \int_0^t e^{-\kappa(t-s)} d\tilde{W}_s. \end{aligned} \quad (2.33)$$

On remarque que  $r_t$  est alors un processus gaussien sous  $\mathbb{Q}$  (en supposant que  $r_0$  est gaussien) de moyenne  $\tilde{\theta} + (r_0 - \tilde{\theta})e^{-\kappa t}$  et de variance  $\frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa t})$ . De plus, on peut montrer qu'il est stationnaire.

Pour valoriser un zéro-coupon, nous avons besoin de calculer  $\int_t^T r_s ds$ . Un moyen simple est d'intégrer (2.31)

$$\begin{aligned} \int_t^T dr_s &= \int_t^T \kappa(\tilde{\theta} - r_s)ds + \int_t^T \sigma d\tilde{W}_s \\ r_T - r_t &= -\kappa \int_t^T r_s ds + \kappa \tilde{\theta}(T - t) + \sigma \int_t^T d\tilde{W}_s \\ \kappa \int_t^T r_s ds &= r_t - r_T + \kappa \tilde{\theta}(T - t) + \sigma \int_t^T d\tilde{W}_s \end{aligned}$$

En remplaçant  $r_T$  par sa solution (2.33) nous obtenons

$$\int_t^T r_s ds = (r_t - \tilde{\theta}) \frac{1 - e^{-\kappa(T-t)}}{\kappa} + \tilde{\theta}(T - t) + \int_t^T \sigma \frac{1 - e^{-\kappa(T-t)}}{\kappa} dW_s \quad (2.34)$$

Il est clair que  $\int_t^T r_s ds$  est également gaussienne sous  $\mathbb{Q}$  conditionnellement à  $\mathcal{F}_t$ , on est donc invité à utiliser la transformée de Laplace. On sait que pour toute variable  $X \sim \mathcal{N}(\mu, \sigma^2)$

$$\mathbb{E}(e^{zX}) = \exp\left(z\mu + \frac{1}{2}z^2\sigma^2\right).$$

Les deux premiers moments conditionnels sont,

$$\begin{aligned}
 \mathbb{E} \left( - \int_t^T r_s ds \middle| \mathcal{F}_t \right) &= -\tilde{\theta}(T-t) + (\tilde{\theta} - r_t) \frac{1 - e^{-\kappa(T-t)}}{\kappa} \\
 \mathbb{V}\text{ar} \left( - \int_t^T r_s ds \middle| \mathcal{F}_t \right) &= \mathbb{E} \left[ \left( \sigma \int_t^T \frac{1 - e^{-\kappa(T-s)}}{\kappa} d\tilde{B}_s \right)^2 \middle| \mathcal{F}_t \right] \\
 &= \sigma^2 \mathbb{E} \left[ \int_t^T \left( \frac{1 - e^{-\kappa(T-s)}}{\kappa} \right)^2 ds \middle| \mathcal{F}_t \right] \\
 &= \sigma^2 \int_t^T \left( \frac{1 - e^{-\kappa(T-s)}}{\kappa} \right)^2 ds \\
 &= -\frac{\sigma^2}{2\kappa^3} (1 - e^{-\kappa(T-t)})^2 + \frac{\sigma^2}{\kappa^2} \left( T - t - \frac{1 - e^{-\kappa(T-t)}}{\kappa} \right). \tag{2.35}
 \end{aligned}$$

Sous la dynamique d'Ornstein-Uhlenbeck (modèle de Vasicek), le prix d'un zéro-coupon de maturité  $T$  est,

$$\begin{aligned}
 B(t, T) &= \exp \left\{ \mathbb{E} \left( - \int_t^T r_s ds \middle| \mathcal{F}_t \right) + \frac{1}{2} \mathbb{V}\text{ar} \left( - \int_t^T r_s ds \middle| \mathcal{F}_t \right) \right\} \\
 &= \exp \left\{ -R_\infty(T-t) + (R_\infty - r_t) \frac{1 - e^{-\kappa(T-t)}}{\kappa} - \frac{\sigma^2}{4\kappa^3} (1 - e^{-\kappa(T-t)})^2 \right\}, \tag{2.36}
 \end{aligned}$$

où  $R_\infty = \tilde{\theta} - \frac{\sigma^2}{2\kappa^2}$ . La courbe des taux à l'instant  $t$  est quant à elle,

$$\begin{aligned}
 R(t, T) &= -\frac{1}{T-t} \ln B(t, T) \\
 &= -R_\infty + (R_\infty - r_t) \frac{1 - e^{-\kappa(T-t)}}{\kappa(T-t)} - \frac{\sigma^2}{4\kappa^3(T-t)} (1 - e^{-\kappa(T-t)})^2. \tag{2.37}
 \end{aligned}$$

### 2.2.3 Modèle de Cox-Ingersoll-Ross

Nous avons vu sur la figure (2.2.2) que le processus d'Ornstein Uhlenbeck pouvait très bien passer sous le zéro, or, cela n'est pas vraiment réaliste dans la vraie vie (quoi que...). Le taux spot sous probabilité historique dans le modèle de Cox-Ingersoll-Ross (CIR), [CoInRo85] est donné par,

$$dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t.$$

Il est possible de montrer que si  $2\kappa\theta > \sigma^2$ , alors, la solution n'atteint pas 0, dans ce cas, le drift de l'équation est suffisamment important pour contraindre la diffusion d'atteindre 0.

Sous probabilité risque neutre, en prenant  $\lambda_t = \lambda\sqrt{r_t}$  comme prime de risque, soit,  $d\tilde{W}_t = dW_t + \lambda\sqrt{r_t}dt$ , nous obtenons,

$$dr_t = \tilde{\kappa}(\tilde{\theta} - r_t)dt + \sigma\sqrt{r_t}d\tilde{W}_t$$

avec  $\tilde{\kappa} = \kappa + \sigma\lambda$  et  $\tilde{\theta} = \frac{\kappa\theta}{\kappa+\sigma\lambda}$ . En écrivant l'EDP des taux pour cette diffusion, nous obtenons

$$\frac{\partial \mathcal{B}}{\partial t} + \tilde{\kappa}(\tilde{\theta} - r)\frac{\partial \mathcal{B}}{\partial r} + \frac{1}{2}\sigma^2 r\frac{\partial \mathcal{B}^2}{\partial r^2} - r\mathcal{B} = 0.$$

Pour pouvoir résoudre cette EDP, on pose l'ansatz,

$$\mathcal{B}(t, T, r) = A(T-t)e^{-C(T-t)r}. \quad (2.38)$$

Ainsi, en notant  $\tau = T - t$  et en remplaçant dans l'EDP précédente,

$$-A'(\tau) + A(\tau)C'(\tau)r - \tilde{\kappa}(\tilde{\theta} - r)A(\tau)C(\tau) + \frac{1}{2}\sigma^2 rA(\tau)C^2(\tau) - rA(\tau) = 0.$$

En égalisant à zéro les coefficients du polynôme en  $r$  ( $r_t$  est donné par le marché, il n'y a donc pas lieu de vouloir modifier sa valeur et de plus nous avons supposé que les fonctions  $A$  et  $C$  ne dépendent que de  $\tau$ ), nous obtenons,

$$\begin{cases} C'(\tau) + \tilde{\kappa}C(\tau) + \frac{1}{2}\sigma^2 C^2(\tau) - 1 = 0 \\ -A'(\tau) - \tilde{\kappa}\tilde{\theta}A(\tau)C(\tau) = 0 \end{cases} \quad (2.39)$$

La première des deux équations est une équation de Riccati, il est donc tout à fait possible de lui trouver une solution, que l'on réinjecte dans l'équation 2 pour déterminer  $A$ . Notons également que nous avons pour condition terminale  $\mathcal{B}(T, r, T) = 1$ , soit,  $A(0) = 1$  et  $C(0) = 0$ . En notant  $\rho = \sqrt{\tilde{\kappa}^2 + 2\sigma^2}$ , la solution générale est,

$$A(\tau) = \left( \frac{2\rho e^{\tau(\tilde{\kappa}+\rho)/2}}{(\tilde{\kappa}+\rho)(e^{\rho\tau}-1)+2\rho} \right)^{\frac{2\tilde{\kappa}\tilde{\theta}}{\sigma^2}}, \quad C(\tau) = \frac{2(e^{\rho\tau}-1)}{(\tilde{\kappa}+\rho)(e^{\rho\tau}-1)+2\rho}.$$

**Preuve 2.2.** La seconde équation de (2.39) dépend des fonctions  $A$  et  $C$ , on ne peut donc pas la résoudre pour le moment. La première équation est quant à elle uniquement écrite en fonction de  $C$ , il s'agit d'une équation de Riccati, nous pouvons donc la résoudre. Nous avons donc

$$C' = 1 - \tilde{\kappa}C - \frac{1}{2}\sigma^2C^2.$$

Nous appliquons la méthode classique de ce type d'équation (voir appendix). Nous obtenons

$$u'' + \tilde{\kappa}u' - \frac{1}{2}\sigma^2u = 0.$$

Le polynôme caractéristique associé est

$$z^2 + \tilde{\kappa}z - \frac{1}{2}\sigma^2 = 0.$$

avec pour racines

$$r_{\pm} = \frac{-\tilde{\kappa} \pm \sqrt{\tilde{\kappa}^2 + 2\sigma^2}}{2}.$$

La solution de  $u$  est donc du type

$$u = \lambda_1 e^{r_1 \tau} + \lambda_2 e^{r_2 \tau}$$

et en déduit utilisant le fait que  $C = -u'/((-1/(2\sigma^2))u)$

$$C(\tau) = -\frac{\lambda_1 r_1 e^{r_1 \tau} + \lambda_2 r_2 e^{r_2 \tau}}{-\frac{1}{2}\sigma^2 (\lambda_1 e^{r_1 \tau} + \lambda_2 e^{r_2 \tau})}.$$

Utilisant la condition terminale  $C(0) = 0$ , en déduit que

$$\frac{\lambda_1}{\lambda_2} = -\frac{r_2}{r_1}.$$

En injectant cela

$$\begin{aligned} C(\tau) &= \frac{2}{\sigma^2} \left( \frac{-r_2 e^{r_1 \tau} + r_2 e^{r_2 \tau}}{\frac{r_2}{r_1} e^{r_1 \tau} + e^{r_2 \tau}} \right) \\ &= 2 \frac{r_2}{\sigma^2} \left( \frac{e^{\rho \tau} - 1}{\frac{r_2}{r_1} + e^{\rho \tau}} \right) \end{aligned}$$

où nous avons posé  $\rho = \sqrt{\tilde{\kappa}^2 + 2\sigma^2}$ . Enfin, après quelques simplifications, nous avons le résultat voulu

$$C(\tau) = \frac{2(e^{\rho\tau} - 1)}{(\tilde{\kappa} + \rho)(e^{\rho\tau} - 1) + 2\rho}.$$

Pour déterminer  $A$ , il faut un peu travailler pour arriver à une forme du type  $f'(\tau)/f(\tau) + cst$  où  $f$  est une fonction que je laisse le lecteur déterminer.

La courbe des taux dans le modèle CIR, (2.1), est alors,

$$\begin{aligned} R(t, T) &= -\frac{1}{T-t} \ln \left( A(T-t)e^{-C(t,T)r} \right) \\ &= \frac{C(T-t)}{T-t} r_t - \frac{\ln A(T-t)}{T-t}. \end{aligned}$$

Les deux principaux attraits de ce modèle sont l'existence d'une formule fermée, et à la différence du modèle de Vasicek, des taux positifs.

#### 2.2.4 Généralisation, modèle de taux court à un facteur

Il existe une littérature abondante sur les modèles de taux court à un facteur. Les modèles de Vasicek et de Cox-Ingersoll-Ross ne sont qu'une infime partie. La plupart de ces modèles peuvent s'écrire sous la forme,

$$dr_t = (\alpha + \beta r_t)dt + \rho r_t^\gamma dW_t,$$

où les paramètres  $\alpha$ ,  $\beta$ ,  $\rho$  et  $\gamma$  sont des constantes. Nous pouvons citer quelques uns des principaux modèles à un facteur :

Dothan	$dr_t = \rho r_t dW_t$
Brennan-Schwartz	$dr_t = (\alpha + \beta r_t)dt + \rho r_t dW_t$
Cox-Ingersoll-Ross variable rate	$dr_t = \rho r_t^{3/2} dW_t$
Constant elasticity of variance	$dr_t = \beta r_t dt + \rho r_t^\gamma dW_t$

Les études empiriques montrent généralement que les modèles avec une sensibilité entre la variation de la volatilité et le niveau du taux  $r$ , i.e.,  $\gamma$  'grand' ( $\geq 1$ ), capturent mieux la dynamique de  $r$ . Formellement, la volatilité des variations non supposées est  $\rho^2 r^{2\gamma}$ . Le paramètre  $\rho$  est un facteur d'échelle, tandis que  $\gamma$  contrôle le degré d'influence du niveau du taux sur la volatilité. Cet effet est dit *effet de niveau*.

La relation entre le niveau de  $r$  et sa volatilité est plus importante que la caractéristique de retour à la moyenne dans les modèles de taux court. De plus, la prise en compte de cette caractéristique augmente la complexité des modèles dans l'analyse de la structure par termes. Comme son importance est moindre par rapport à celle de la variance, cette prise en compte n'est pas toujours facilement justifiable. Hormis la possibilité de voir un taux négatif, le modèle de Vasicek propose un  $\gamma = 0$ , soit, une mauvaise prise en compte de l'effet de niveau. La volatilité des variations est indépendante du niveau de  $r$ .

Les modèles à un facteur forment une famille assez limitée et ne peuvent pas permettre de valoriser correctement n'importe quelles obligations.

### 2.2.5 Modèle de Hull and White

Les processus que nous avons vus précédemment sont invariants en temps. Les paramètres sont supposés constants. Pour plus de flexibilité, Hull et White, [HuWh90], proposent une diffusion du taux plus générale,

$$dr_t = (\theta(t) - \alpha(t)r_t)dt + \sigma(t)r_t^\beta dW_t.$$

Ce modèle est également appelé modèle de Vasicek généralisé. les fonctions  $\theta(t)$ ,  $\alpha(t)$  et  $\sigma(t)$  varient en fonction du temps et peuvent être utilisées pour calibrer exactement le modèle aux prix du marché. Le prix à payer pour cette calibration exacte est que le prix des obligations et de leurs options n'est plus obtenable analytiquement.

Il est toujours tentant de laisser plus de flexibilité aux paramètres, donc d'écrire  $\alpha(t)$  et  $\sigma(t)$  au lieu de  $\alpha$  et  $\sigma$ . Néanmoins, une conséquence importante est que la volatilité de la courbe des taux devient en général non stationnaire et évolue de manière totalement imprévisible. Hull et White eux-mêmes dans un papier parlent d'un risque de sur-paramétrisation du modèle si on laisse trop de liberté. Le modèle de Hull et White est en fait,

$$dr_t = (\theta(t) - \alpha r_t)dt + \sigma r_t^\beta dW_t,$$

où  $\alpha$ ,  $\sigma$  et  $\beta \geq 0$  sont des constantes, et  $\theta(t)$  une fonction déterministe inconnue. Nous considérerons ici le cas où  $\beta = 0$ , ce qui nous permettra d'avoir une solution simplement. Nous commençons par déterminer le prix du zéro-coupon avec une approche par EDP.

L'EDP des taux prend la forme,

$$\frac{\partial \mathcal{B}}{\partial t} + (\phi(t) - \alpha r) \frac{\partial \mathcal{B}}{\partial r} + \frac{\sigma^2}{2} \frac{\partial \mathcal{B}^2}{\partial r^2} - r \mathcal{B} = 0,$$

avec  $\phi(t) = \theta(t) - \sigma \lambda$  (en d'autres termes nous avons pris  $\lambda$  tel que  $dB = d\tilde{B} - \lambda dt$ ). Comme précédemment, nous supposerons que le prix du zéro-coupon est de la forme,

$$\mathcal{B}(t, T, r) = A(T-t)e^{-C(T-t)r}.$$

En injectant cette équation dans l'EDP des taux, nous obtenons comme système,

$$\begin{cases} \frac{dC}{dt} - \alpha C + 1 = 0 \\ \frac{dA}{dt} - \phi(t)AC + \frac{\sigma^2}{2}AC^2 = 0 \end{cases}$$

avec comme condition terminale à  $t = T$ ,  $A(T, T) = 1$  et  $C(T, T) = 0$ . Il s'agit d'un système d'équations différentielles somme toute classique, et nous trouvons après quelques calculs,

$$\begin{aligned} C(t, T) &= \frac{1}{\alpha} \left( 1 - e^{-\alpha(T-t)} \right) \\ A(t, T) &= \exp \left\{ \frac{\sigma^2}{2} \int_t^T C^2(u, T) du - \int_t^T \phi(u)C(u, T) du \right\}. \end{aligned} \quad (2.40)$$

Il ne nous reste alors qu'à déterminer l'expression de  $\phi(t)$ . En utilisant la deuxième équation de (2.40) et en utilisant la formulation du prix d'un zéro coupon comme fonction exponentielle, nous obtenons,

$$\int_t^T \phi(u)C(u, T) du = \frac{\sigma^2}{2} \int_t^T C^2(u, T) du - \ln \mathcal{B}(r, t, T) - rC(t, T). \quad (2.41)$$

En différentiant le terme de droite,

$$\begin{aligned} \frac{\partial}{\partial T} \int_t^T \phi(u)C(u, T) du &= \phi(u)C(u, T)|_{u=T} + \int_t^T \phi(u) \frac{\partial}{\partial T} C(u, T) du \\ &= \int_t^T \phi(u) e^{-\alpha(T-u)} du, \end{aligned}$$

et en égalisant avec la dérivée du terme de gauche,

$$\int_t^T \phi(u) e^{-\alpha(T-u)} du = \frac{\sigma^2}{\alpha} \int_t^T (1 - e^{-\alpha(T-u)}) e^{-\alpha(T-u)} du - \frac{\partial}{\partial T} \ln \mathcal{B}(r, t, T) - r e^{-\alpha(T-t)}.$$

Il ne reste plus qu'à supprimer le terme  $e^{-\alpha(T-u)}$ . On injecte l'équation (2.41) en la multipliant par  $\alpha$ . Après quelques calculs, on obtient,

$$\int_t^T \phi(u) du = \frac{\sigma^2}{2\alpha} \int_t^T (1 - e^{-2\alpha(T-u)}) du - \frac{\partial}{\partial T} \ln \mathcal{B}(r, t, T) - \alpha \ln \mathcal{B}(r, t, T) - r.$$

En différentiant une nouvelle fois l'équation, on arrive au résultat voulu,

$$\phi(T) = \frac{\sigma^2}{2\alpha} (1 - e^{-2\alpha(T-t)}) - \frac{\partial^2}{\partial T^2} \ln \mathcal{B}^2(r, t, T) - \alpha \frac{\partial}{\partial T} \ln \mathcal{B}(r, t, T)$$

La volatilité du prix de l'obligation, donnée par (2.26), dans le cadre du modèle de Hull and White est,

$$\sigma_B(t, T) = \frac{\sigma}{\beta} \frac{\partial \mathcal{B}}{\partial r} = -\sigma C(t, T) = -\frac{\sigma}{\alpha} (1 - e^{-\alpha(T-t)}).$$

Pour déterminer la diffusion du prix de l'obligation, nous n'avons plus qu'à appliquer le lemme d'Itô à (2.38),

$$\frac{d\mathcal{B}(r, t, T)}{\mathcal{B}(r, t, T)} = r dt - \frac{\sigma}{\alpha} (1 - e^{-\alpha(T-t)}) dW_t.$$

Il est également possible de trouver la solution par martingale.

Rappelons que la diffusion de Hull and White est,

$$dr_t = (\theta(t) - \alpha r_t) dt + \sigma dW_t \quad (2.42)$$

avec  $(W_t)_{t \in \mathbb{R}_+}$  un processus de Wiener. En utilisant la même astuce que pour le modèle de Vasicek, et donc, le processus d'Ornstein Uhlenbeck, nous pouvons trouver la solution de (2.42),

$$dr_t e^{\alpha t} = \theta(t) e^{\alpha t} dt + \sigma e^{\alpha t} dW_t$$

Ainsi,

$$\begin{aligned}
 \int_t^T r(u)du &= r_t \frac{1 - e^{-\alpha(T-t)}}{\alpha} + \int_t^T \int_t^u \theta(s) e^{-\alpha(u-s)} ds du \\
 &\quad + \int_t^T \int_t^u \sigma e^{-\alpha(u-s)} dW_s du \\
 &= r_t \frac{1 - e^{-\alpha(T-t)}}{\alpha} + \int_t^T \theta(s) \frac{1 - e^{-\alpha(T-s)}}{\alpha} ds \\
 &\quad + \int_t^T \frac{\sigma}{\alpha} (1 - e^{-\alpha(T-s)}) dW_s.
 \end{aligned}$$

Il est clair que la moyenne et variance sont données par,

$$\begin{aligned}
 \mathbb{E} \left( \int_t^T r_u du \right) &= r_t \frac{1 - e^{-\alpha(T-t)}}{\alpha} + \int_t^T \theta(s) \frac{1 - e^{-\alpha(T-s)}}{\alpha} ds, \\
 \mathbb{V} \left( e^{-\int_t^T r_u du} \right) &= \int_t^T \frac{\sigma^2}{\alpha^2} (1 - e^{-\alpha(T-s)})^2 ds
 \end{aligned}$$

La valorisation du zéro coupon est alors simplement,

$$\begin{aligned}
 B(t, T) &= \mathbb{E} \left( e^{-\int_t^T r_u du} \right) \\
 &= \exp \left( -\mathbb{E} \left( e^{-\int_t^T r_u du} \right) + \frac{1}{2} \mathbb{V} \left( e^{-\int_t^T r_u du} \right) \right) \\
 &= \exp \left( r_t \frac{1 - e^{-\alpha(T-t)}}{\alpha} + \int_t^T \theta(s) \frac{1 - e^{-\alpha(T-s)}}{\alpha} ds \right. \\
 &\quad \left. + \frac{1}{2} \int_t^T \frac{\sigma^2}{\alpha^2} (1 - e^{-\alpha(T-s)})^2 ds \right).
 \end{aligned}$$

### 2.2.6 Quelques modèles

Pour conclure cette section, comme nous n'avons bien évidemment pas donné tous les modèles existants, nous présentons ici quelques unes des diffusions les plus utilisées. Notons qu'il n'y a pas de consensus sur les modèles ATS, chaque modèle a ses avantages/inconvénients. Le modèle de Black & Scholes par exemple, bien que les hypothèses de bases soient erronées, le modèle reste largement utilisé par les praticiens. Plusieurs raisons, il est largement connu et l'on sait que la plupart des agents l'utilisent, nous avons donc un effet moutonnier non négligeable. La formule de robustesse est très intéressante, bien que le modèle soit erroné, l'intervalle de confiance reste correct et donne donc une bonne indication au trader.

Notons enfin un point qui peut paraître secondaire mais finalement non, le modèle de B&S a une formule fermée pour la pricing d'option, ce qui est loin d'être le cas des principaux modèles concurrents, plus fidèles à la réalité des marchés.

- Vasicek (77) :  $\mathbb{R}$

$$dr_t = \kappa(\theta - r_t)dt + \sigma dW_t.$$

- Longstaff (89) :  $\mathbb{R}_+$

$$dr_t = \kappa(\theta - \beta\sqrt{r_t})dt + \sigma\sqrt{r_t}dW_t$$

- Cox-Ingersoll-Ross (85) :  $\mathbb{R}_+$

$$dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t.$$

- Dothan (78) :  $\mathbb{R}_+$

$$dr_t = \beta r_t dt + \sigma r_t dW_t.$$

- Black-Derman-Toy (90) :  $\mathbb{R}_+$

$$dr_t = \beta(t)r_t dt + \sigma(t)dW_t.$$

- Black - Karasinki (91) :  $\mathbb{R}_+$

$$d\log r_t = \kappa(t)(\theta(t) - \log r_t)dt + \sigma(t)dW_t.$$

- Ho-Lee (86) :  $\mathbb{R}$

$$dr_t = \mu(t)dt + \sigma dW_t.$$

- Hull and White (90) :  $\mathbb{R}$

$$dr_t = \kappa(t)(\theta(t) - r_t)dt + \sigma(t)dW_t.$$

- Hull and White (90) :  $\mathbb{R}_+$

$$dr_t = \kappa(t)(\theta(t) - r_t)dt + \sigma(t)\sqrt{r_t}dW_t.$$

## 2.3 Modèles multifactoriels

### 2.3.1 Fondations

Le problème des modèles monofactoriels, c'est-à-dire n'ayant qu'une seule diffusion pour  $r_t$  est qu'il n'est pas possible de prendre les corrélations entre des

## Modèle de Taux

---

	1w	1m	6m	12m	5ans	30ans
1w	1					
1m	0,976	1				
6m	0,960	0,961	1			
12m	0,953	0,952	0,999	1		
5ans	0,935	0,907	0,943	0,946	1	
30ans	0,772	0,789	0,797	0,795	0,802	1

TABLE 1 – Matrice de Corrélation d’obligations françaises à différentes maturités.  
Données journalières du 02/01/12 au 31/07/13.

obligations à différentes maturités. Il est clair que dans chaque cas, nous avons la corrélation d’un zéro coupon de maturité  $T$  et  $S$  égale à 1. Sur le tableau (1), la corrélation entre des obligations françaises à différentes maturités.

De plus, les modèles à un facteur ont fait l’objet de plusieurs critiques. Les modèles n’arrivent pas à reproduire la réalité empirique du marché. D’un point de vue théorique, il semble absurde de vouloir ‘tout’ expliquer par la simple diffusion du taux court  $r_t$ . Les surfaces de volatilité des taux forward sont difficilement reproductibles sans avoir un modèle très compliqué pour le taux court.

Afin de pallier à ces différents problèmes, l’usage de la seule variable aléatoire  $r_t$  comme exogène a été de plus en plus abandonné. Les modèles multi-facteurs ont été introduits. Le choix de ces facteurs est nécessairement fondamental. Les conditions de non-arbitrage doivent être également vérifiées. La condition de non-arbitrage est certainement l’une des hypothèses en finance les plus importantes, et qui reste vérifiée au fil des années. A l’inverse, l’hypothèse de normalité des marchés de L. Bachelier apparaît aujourd’hui totalement incohérente avec les faits stylisés. L’hypothèse d’infinité divisibilité des actifs est bien entendu invalidée par la réalité, tout comme le caractère continue des données, spécialement à haute fréquence et pour les données non liquides. Néanmoins, l’hypothèse de non-arbitrage, c’est-à-dire qu’il est impossible de faire de l’argent de manière certaine avec une richesse nulle, reste vérifiée. La plupart des grandes banques et des fonds d’investissement ont tout une batterie d’arbitragistes, leur métier consiste donc à repérer toute opportunité d’arbitrage à des temps microscopiques. Il est donc, pour nous, impossible de pouvoir en bénéficier.

La plus simple extension du modèle à un facteur est de considérer le taux court comme fonction de deux diffusions. Plus généralement, nous pouvons étendre le

cas à un modèle  $d$ -dimensionnel

$$r_t = \sum_{i=1}^n x_t^i$$

où chaque facteur est du type

$$dx_t^i = \mu_t^i dt + \sigma_t^i dW_t^i.$$

Dans le cas où les  $W^i$  sont des mouvements Brownien indépendants, nous avons simplement comme formule de pricing de zéro-coupon

$$\begin{aligned} B(t, T) &= \mathbb{E} \left( e^{-\int_t^T r_s ds} \middle| \mathcal{F}_t \right) \\ &= \mathbb{E} \left( e^{-\int_t^T \sum_{i=1}^n x_s^i ds} \middle| \mathcal{F}_t \right) \\ &= \prod_{i=1}^n \mathbb{E} \left( e^{-\int_t^T x_s^i ds} \middle| \mathcal{F}_t \right). \end{aligned}$$

### 2.3.2 Modèle de Vasicek à deux facteurs

Nous considérons le cas à deux facteurs, i.e.

$$r_t = x_t + y_t$$

avec

$$\begin{aligned} dx_t &= -\kappa_x x_t dt + \sigma_x dW_t^x \\ dy_t &= -\kappa_y y_t dt + \sigma_y dW_t^y, \end{aligned} \tag{2.43}$$

où  $(W^x, W^y)$  est un mouvement Brownien bi-dimensionnel tel que

$$\langle dW_t^x dW_t^y \rangle = \rho dt$$

avec  $\rho \in [-1, 1]$ . Nous supposerons enfin que  $\kappa_x, \kappa_y, \sigma_x$  et  $\sigma_y$  sont des constantes positives. Les processus  $(x_t)_{t \in \mathbb{R}_+}$  et  $(y_t)_{t \in \mathbb{R}_+}$  sont deux processus de Ornstein Ulhenbeck.

En intégrant les équations (2.43), nous avons pour tout  $s < t$

$$\begin{aligned} r_t &= x_s e^{-\kappa_x(t-s)} + y_s e^{-\kappa_y(t-s)} \\ &\quad + \sigma_x \int_s^t e^{-\kappa_x(t-u)} dW_u^x + \sigma_y \int_s^t e^{-\kappa_y(t-u)} dW_u^y. \end{aligned}$$

Il est clair que  $r_t$ , conditionnellement à son historique, est gaussien de moyenne et variance

$$\begin{aligned}\mathbb{E}(r_t | \mathcal{F}_s) &= x_s e^{-\kappa_x(t-s)} + y_s e^{-\kappa_y(t-s)} \\ \mathbb{V}(r_t | \mathcal{F}_s) &= \frac{\sigma_x^2}{2\kappa_x} (1 - e^{-2\kappa_x(t-s)}) + \frac{(\sigma_y)^2}{2\kappa_y} (1 - e^{-2\kappa_y(t-s)}) \\ &\quad + 2\rho \frac{\sigma_x \sigma_y}{\kappa_x + \kappa_y} (1 - e^{-(\kappa_x + \kappa_y)(t-s)}).\end{aligned}\quad (2.44)$$

Comme pour le cas des modèles à un facteur, le prix d'un zéro-coupon  $B(t, T)$  est donné sous la probabilité risque-neutre par

$$B(t, T) = \mathbb{E} \left( e^{-\int_t^T r_s ds} \middle| \mathcal{F}_t \right).$$

Nous introduisons la quantité suivante :

$$I(t, T) := \int_t^T (x_u + y_u) du. \quad (2.45)$$

**Lemme 2.1.** Pour tout  $t \leq T$ , la quantité  $I(t, T)$  donnée par (2.45) est gaussien conditionnellement à  $\mathcal{F}_t$  de moyenne  $M(t, T)$  et de variance  $V(t, T)$  respectivement données par

$$M(t, T) = \frac{1 - e^{-\kappa_x(T-t)}}{\kappa_x} x_t + \frac{1 - e^{-\kappa_y(T-t)}}{\kappa_y} y_t$$

et

$$\begin{aligned}V(t, T) &= \frac{\sigma_x^2}{\kappa_x^2} \left[ T - t + \frac{2}{\kappa_x} e^{-\kappa_x(T-t)} - \frac{1}{2\kappa_x} e^{-2\kappa_x(T-t)} - \frac{3}{2\kappa_x} \right] \\ &\quad + \frac{\sigma_y^2}{\kappa_y^2} \left[ T - t + \frac{2}{\kappa_y} e^{-\kappa_y(T-t)} - \frac{1}{2\kappa_y} e^{-2\kappa_y(T-t)} - \frac{3}{2\kappa_y} \right] \\ &\quad + 2\rho \frac{\sigma_x \sigma_y}{\kappa_x \kappa_y} \left[ T - t + \frac{e^{-\kappa_x(T-t)} - 1}{\kappa_x} + \frac{e^{-\kappa_y(T-t)} - 1}{\kappa_y} - \frac{e^{-(\kappa_x + \kappa_y)(T-t)} - 1}{\kappa_x + \kappa_y} \right]\end{aligned}$$

**Preuve 2.3.** Il est possible de reprendre les calculs fait pour le cas du modèle à un facteur (2.32). Nous proposons ici une autre preuve du calcul de  $\int_t^T r_s ds$  et donc des quantités  $M$  et  $V$ . Par application de la formule d'intégration par partie stochastique, nous avons

$$\begin{aligned}\int_t^T x_u du &= Tx_T - tx_t - \int_t^T u dx_u \\ &= \int_t^T (T-u) dx_u + (T-t)x_t.\end{aligned}\quad (2.46)$$

Par définition de  $x$ , la partie droite de (2.46) peut être réécrite par :

$$\int_t^T (T-t)dx_u = -\kappa_x \int_t^T (T-t)x_u du + \sigma_x \int_t^T (T-t)dW_u^x,$$

avec

$$\int_t^T (T-t)dx_u = x_t \int_t^T (T-u)e^{-\kappa_x(u-t)} du + \sigma_x \int_t^T (T-u) \int_t^u e^{-\kappa_x(u-s)} dW_s^x du.$$

En calculant les deux intégrales,

$$-\kappa_x x_t \int_t^T (T-u)e^{-\kappa_x(u-t)} du = -x_t(T-t) - \frac{e^{-\kappa_x(T-t)} - 1}{\kappa_x} x_t$$

et,

$$\begin{aligned} & -\kappa_x \sigma_x \int_t^T (T-u) \int_t^u e^{-\kappa_x(u-s)} dW_s^x du \\ &= -\kappa_x \sigma_x \int_t^T \left( \int_t^u e^{\kappa_x s} dW_s^x \right) du \left( \int_t^u (T-v)e^{-\kappa_x v} dv \right) \\ &= -\kappa_x \sigma_x \left[ \left( \int_t^T e^{\kappa_x u} dW_u^x \right) \left( \int_t^T (T-v)e^{-\kappa_x v} dv \right) \right. \\ &\quad \left. - \int_t^T ((T-v)e^{-\kappa_x v}) e^{\kappa_x u} dW_u^x \right] \\ &= -\kappa_x \sigma_x \int_t^T \left( \int_u^T (T-v)e^{-\kappa_x v} dv \right) e^{\kappa_x u} dW_u^x \\ &= -\sigma \int_t^T \left( (T-u) + \frac{e^{-\kappa_x(T-u)} - 1}{\kappa_x} \right) dW_u^x, \end{aligned}$$

où pour la dernière étape nous avons utilisé le fait que

$$\int_u^T (T-u)e^{-\kappa_x v} dv = \frac{(T-u)e^{-\kappa_x u}}{\kappa_x} + \frac{e^{-\kappa_x T} - e^{-\kappa_x u}}{\kappa_x^2}.$$

Ainsi

$$\int_t^T x_u du = \frac{1 - e^{-\kappa_x(T-t)}}{\kappa_x} x_t + \frac{\sigma_x}{\kappa_x} \int_t^T \left( 1 - e^{-\kappa_x(T-u)} \right) dW_u^x.$$

Nous obtenons un résultat tout à fait symétrique pour  $y$

$$\int_t^T y_u du = \frac{1 - e^{-\kappa_y(T-t)}}{\kappa_y} y_t + \frac{\sigma_y}{\kappa_x} \int_t^T \left( 1 - e^{-\kappa_y(T-u)} \right) dW_u^y.$$

La variance conditionnelle est

$$\begin{aligned}\mathbb{V}ar(I(t, T) | \mathcal{F}_t) &= \mathbb{V}ar\left[\frac{\sigma_x}{\kappa_x} \int_t^T (1 - e^{-\kappa_x(T-u)}) dW_u^x + \frac{\sigma_y}{\kappa_y} \int_t^T (1 - e^{-\kappa_y(T-u)}) dW_u^y\right] \\ &= \frac{\sigma_x^2}{\kappa_x^2} \int_t^T (1 - e^{-\kappa_x(T-u)})^2 du + \frac{\sigma_y^2}{\kappa_y^2} \int_t^T (1 - e^{-\kappa_y(T-u)})^2 du \\ &\quad + 2\rho \frac{\sigma_x \sigma_y}{\kappa_x \kappa_y} \int_t^T (1 - e^{-\kappa_x(T-u)}) (1 - e^{-\kappa_y(T-u)}) du\end{aligned}$$

**Théorème 2.1.** Le prix à l'instant  $t$  d'un zéro-coupon de maturité  $T$  est

$$B(t, T) = \exp \left\{ -\frac{1 - e^{-\kappa_x(T-t)}}{\kappa_x} x_t - \frac{1 - e^{-\kappa_y(T-t)}}{\kappa_y} y_t + \frac{1}{2} V(t, T) \right\}.$$

**Preuve 2.4.** Pour toute variable aléatoire réelle  $Z$  de moyenne  $m_Z$  et de variance  $\sigma_Z^2$ , nous avons  $\mathbb{E}e^Z = e^{m_Z + \frac{1}{2}\sigma_Z^2}$ .  $r_t$  étant conditionnellement gaussien par rapport à  $\mathcal{F}_t$ , le lemme (2.1) nous permet de conclure.

**Remarque 2.1.** Pour que le modèle à deux facteurs puisse refléter au mieux les données du marché après calibration, nous introduisons une fonction déterministe  $(\varphi(t))_{t \in \mathbb{R}_+}$  telle que

$$r_t = x_t + y_t + \varphi(t). \quad (2.47)$$

Il est possible de montrer alors que

$$\varphi(t) = F^M(0, T) + \frac{\sigma_x^2}{2\kappa_x^2} (1 - e^{\kappa_x T})^2 + \frac{\sigma_y^2}{2\kappa_y^2} (1 - e^{\kappa_y T})^2 + \rho \frac{\sigma_x \sigma_y}{\kappa_x \kappa_y} (1 - e^{\kappa_x T})(1 - e^{\kappa_y T}).$$

**Preuve 2.5.** voir [BrMe01].

Notons enfin qu'il est possible après quelques calculs de montrer que

$$\begin{aligned}\frac{\mathbb{C}or(dF(t, T_1), dF(t, T_2))}{dt} &= \frac{\sigma_x^2 e^{-\kappa_x(T_1+T_2-2t)} + \sigma_y^2 e^{-\kappa_y(T_1+T_2-2t)}}{\sigma_F(t, T_1) \sigma_F(t, T_2)} \\ &\quad + \frac{\rho \sigma_x \sigma_y (e^{-\kappa_x(T_1+T_2-2t)} + e^{-\kappa_y(T_1+T_2-2t)})}{\sigma_F(t, T_1) \sigma_F(t, T_2)},\end{aligned}$$

qui est en général différent de 1.

### 2.3.3 Quelques modèles

Comme précédemment pour le cas mono-factoriel, nous présentons quelques uns des modèles les plus usuels.

- **Richard, 1978, [Ri78].**

Le modèle de Richard propose d'expliquer la structure par termes des taux d'intérêts par deux facteurs, le taux court observé  $(q_t)_{t \in \mathbb{R}_+}$  et le taux d'inflation instantané espéré  $(\pi_t)_{t \in \mathbb{R}_+}$ . Les équations de diffusion sont

$$\begin{aligned} dq_t &= \mu_q dt + \sigma_q dW_t^q \\ d\pi_t &= \mu_\pi dt + \sigma_\pi dW_t^\pi \end{aligned}$$

avec  $(W_t^q)_{t \in \mathbb{R}_+}$  et  $(W_t^\pi)_{t \in \mathbb{R}_+}$  deux mouvements brownien indépendants.

- **Brennan et Schwartz, 1979, 1982, [BrSc79],[BrSc82].**

L'idée du modèle de Brennan et Schwartz, est de supposer que le taux long  $(\ell_t)_{t \in \mathbb{R}_+}$  contient de l'information sur la valeur future du taux court  $(r_t)_{t \in \mathbb{R}_+}$ . La structure par termes des taux est ainsi expliquée par la diffusion jointe

$$\begin{aligned} dr_t &= \mu_r() dt + \sigma_r() dW_t^r \\ d\ell_t &= \mu_\ell() dt + \sigma_\ell() dW_t^\ell, \end{aligned}$$

avec  $(W_t^r)_{t \in \mathbb{R}_+}$  et  $(W_t^\ell)_{t \in \mathbb{R}_+}$  deux mouvements Brownien corrélés et  $\mu_r, \mu_\ell, \sigma_r$  et  $\sigma_\ell$  des fonction réelles. Après quelques calculs similaires à ceux des modèles à un facteur, nous obtenons

$$\frac{\partial B}{\partial t} + (\mu_r - \lambda_r \sigma_r) \frac{\partial B}{\partial r} + (\mu_\ell - \lambda_\ell \sigma_\ell) \frac{\partial B}{\partial \ell} + \frac{\sigma_r^2}{2} \frac{\partial^2 B}{\partial r^2} + \frac{\sigma_\ell^2}{2} \frac{\partial^2 B}{\partial \ell^2} + \rho \sigma_r \sigma_\ell \frac{\partial^2 B}{\partial r \partial \ell} - r B = 0 \quad (2.48)$$

avec toujours comme condition terminale  $B(T, T) = 1$ . Les primes de risque des taux courts et des taux longs, respectivement  $\lambda_r$  et  $\lambda_\ell$  vérifient

$$\mu_B - r_t B(t, T) = \lambda_r(t, r, \ell) \sigma_{B,r} + \lambda_\ell(t, r, \ell) \sigma_{B,\ell}.$$

avec

$$\begin{aligned}\mu_B &= \frac{\partial B}{\partial t} + \mu_r \frac{\partial B}{\partial r} + \mu_\ell \frac{\partial B}{\partial \ell} + \frac{\sigma_r^2}{2} \frac{\partial^2 B}{\partial r^2} + \frac{\sigma_\ell^2}{2} \frac{\partial^2 B}{\partial \ell^2} + \rho \sigma_r \sigma_\ell \frac{\sigma_r^2}{2} \frac{\partial^2 B}{\partial r \partial \ell} \\ \sigma_{B,r} &= -\sigma_r \frac{\partial B}{\partial r} \\ \sigma_{B,\ell} &= -\sigma_\ell \frac{\partial B}{\partial \ell}.\end{aligned}$$

Le problème de ce modèle, est que l'équation (2.48) n'admet pas de solution explicite en tant que tel, elle doit être résolue numériquement. Brennan et Schwartz, [BrSc82], ont par la suite choisi une forme spécifique pour le drift et la volatilité

$$\begin{aligned}dr_t &= (a_1 + b_1(\ell_t - r_t))dt + \sigma_1 r_t dW_t^1 \\ d\ell_t &= \ell_t(a_2 + b_2 r_t + c_2 \ell_t)dt + \sigma_2 \ell_t dW_t^2\end{aligned}$$

Le taux court est donc directement fonction du spread de taux long-taux court. Néanmoins, ce modèle a été invalidé par Hogan [Ho93] et Duffie et al. [DuMaYo94], qui ont prouvé qu'il n'y avait pas de solution réelle à ce système d'EDS et, qu'il était possible d'exhiber un arbitrage.

#### – Schaefer et Schwartz, 1984, [ScSc84]

Schaefer et Schwartz [ScSc84], proposent cette fois de ne plus modéliser directement le taux court  $(r_t)_{t \in \mathbb{R}_+}$  mais le spread de taux court - taux long  $(s_t)_{t \in \mathbb{R}_+}$  et, le taux long  $(\ell_t)_{t \in \mathbb{R}_+}$

$$\begin{aligned}ds_t &= m(\mu - s_t)dt + \sigma_1 dW_t^1 \\ d\ell_t &= \beta(s_t, \ell_t, t)dt + \sigma_2 \sqrt{\ell_t} dW_t^2.\end{aligned}$$

Le spread est un processus d'Ornstein-Uhlenbeck, i.e. mean-reverting, le taux long a une variance proportionnelle à ses propres variations. Les auteurs n'arrivent pas à obtenir de formule fermée pour le pricing de zéro-coupon mais uniquement une approximation.

#### – Longstaff et Schwartz, 1992,[LoSc92].

Longstaff et Schwartz proposent d'expliquer la courbe des taux par le taux court  $(r_t)_{t \in \mathbb{R}_+}$  et la variation du taux court  $(v_t)_{t \in \mathbb{R}_+}$ . Le prix de l'obligation

prend la forme

$$dB(t, T) = B(t, T)((\mu X_t + \theta Y_t)dt + \sigma \sqrt{Y} dW_t)$$

avec  $(X_t)_{t \in \mathbb{R}_+}$  et  $(Y_t)_{t \in \mathbb{R}_+}$  deux facteurs économiques correspondants respectivement à la valeur du rendement espéré et, à la corrélation par rapport aux variations de prix. La dynamique choisie est

$$\begin{aligned} dX_t &= (a - bX_t)dt + c\sqrt{X_t}dW_t^X \\ dY_t &= (d - eY_t)dt + f\sqrt{Y_t}dW_t^Y, \end{aligned}$$

où  $(W_t^X)_{t \in \mathbb{R}_+}$  et  $(W_t^Y)_{t \in \mathbb{R}_+}$  sont deux mouvements Brownien non corrélés et l'ensemble des constantes sont prises positives strictement. Les auteurs montrent qu'il est possible d'exprimer les deux quantités  $X_t$  et  $Y_t$  en fonction de  $r_t$  et  $v_t$

$$\begin{aligned} r_t &= \alpha X_t + \beta Y_t \\ v_t &= \alpha^2 X_t + \beta^2 Y_t \end{aligned}$$

avec  $\alpha = \mu c^2$  et  $\beta = (\theta - \sigma^2)f^2$ , ainsi,  $r_t$  et  $v_t$  sont positifs. Les variables économiques sont alors données par

$$\begin{aligned} X_t &= \frac{\beta r_t - v_t}{\alpha(\beta - \alpha)} \\ Y_t &= \frac{v_t - \alpha r_t}{\beta(\beta - \alpha)}. \end{aligned}$$

Au prix de quelques calculs, il est cette fois possible de trouver une solution explicite pour la valorisation d'un zéro-coupon ainsi qu'une option sur zéro-coupon. Néanmoins, la solution finale comporte énormément de paramètres, donc, compliqué à estimer de manière 'efficace'.

– **Chen, 1996, [Ch96].**

Chen propose un modèle à trois facteurs, le taux court  $(r_t)_{t \in \mathbb{R}_+}$ , sa moyenne stochastique  $(\theta_t)_{t \in \mathbb{R}_+}$  et sa volatilité stochastique  $(v_t)_{t \in \mathbb{R}_+}$ .

$$\begin{aligned} dr_t &= \kappa(\bar{\theta} - \theta_t)dt + \sqrt{\sigma(t)}\sqrt{r_t}dW_t^r \\ d\theta_t &= \nu(\bar{\theta} - \theta_t)dt + \xi\sqrt{\theta}dW_t^\theta \\ d\sigma_t &= \mu(\bar{\sigma} - \sigma_t)dt + \eta\sqrt{\sigma}dW_t^\sigma, \end{aligned}$$

chacun des trois processus est donc un processus de retour à la moyenne. Il est extrêmement difficile d'obtenir une solution pour le prix des obligations ainsi que pour les produits dérivés. Il est néanmoins possible de simuler le modèle en le discrétilisant.

## 2.4 Pricing de dérivés de taux

### 2.4.1 Mesure forward

Rappelons qu'en absence d'arbitrage, le processus

$$e^{-\int_0^t r_s ds} B(t, T), \quad t \leq T$$

est une  $\mathcal{F}_t$ -martingale sous  $\mathbb{Q}$  et donc, la suite de variables  $(B(t, T_i))_{t \in [0, T_i]}$  peut être prise comme le numéraire dans la définition

$$\frac{d\mathbb{Q}_i}{d\mathbb{Q}} = \frac{1}{B(0, T_i)} e^{-\int_0^{T_i} r_s ds}$$

de la mesure forward  $\mathbb{Q}_i$ . Nous avons alors, pour toute fonction  $\phi$  'suffisamment' régulière  $\phi$  et pour toute variable aléatoire réelle  $X$ ,

$$\mathbb{E} \left( \phi(X) e^{-\int_t^{T_i} r_s ds} \middle| \mathcal{F}_t \right) = B(t, T_i) \mathbb{E}_{\mathbb{Q}_i}(\phi(X) | \mathcal{F}_t), \quad 0 \leq t \leq T. \quad (2.49)$$

Pour pouvoir utiliser l'équation (2.49), il pourrait être utile d'écrire les dynamiques de  $r_t$ ,  $f(t, T, S)$  et  $B(t, T)$  sous la mesure forward  $\mathbb{Q}_i$ .

Supposons simplement que la dynamique du prix de l'obligation soit du type,

$$\frac{dB(t, T_i)}{B(t, T_i)} = r_t dt + \sigma_i(t) dW_t^i, \quad 0 \leq t \leq T_i, \quad i = 1, \dots, n$$

avec  $(W_t^i)_{t \in \mathbb{R}_+}$  un mouvement Brownien standard sous  $\mathbb{Q}$  pour tout  $i = 1, \dots, n$ , et  $(r_t)_{t \in \mathbb{R}_+}$  et  $(\sigma(t))_{t \in \mathbb{R}_+}$  deux processus adaptés à la filtration  $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$  générée par  $(W_t)_{t \in \mathbb{R}_+}$ . Notons que cette diffusion est en accord avec les résultats précédents de diffusion sous la probabilité risque-neutre (2.27), (2.28).

D'après le théorème de Girsanov, nous savons que

$$\tilde{W}_t^i := W_t - \int_0^t \sigma_i(s) ds, \quad 0 \leq t \leq T_i,$$

est un mouvement Brownien standard sous la mesure forward  $\mathbb{Q}_i$  pour tout le tenor  $i = 1, \dots, n$ . La dynamique du prix de l'obligation sous  $\mathbb{Q}_i$  est

$$\frac{dB(t, T_i)}{B(t, T_i)} = (r_t + \sigma_i^2(t))dt + \sigma_i(t)d\tilde{W}_t^i, \quad 0 \leq t \leq T_i, \quad i = 1, \dots, n.$$

Dans sa forme générale, le taux court  $r_t$  est supposé Markovien et solution de l'EDS

$$dr_t = \mu(t, r_t)dt + \sigma(t, r_t)dW_t.$$

Sous  $\mathbb{Q}_i$ , nous avons directement

$$dr_t = \mu(t, r_t)dt + \sigma(t, r_t)\sigma_i(t)dt + \sigma(t, r_t)d\tilde{W}_t^i.$$

Nous prenons (une fois de plus) l'exemple du modèle de Vasicek

$$dr_t = \kappa(\theta - r_t)dt + \sigma dW_t^i, \quad 0 \leq t \leq T_i.$$

D'après (2.57), (2.60) et (2.63), nous avons

$$\begin{aligned} \frac{dB(t, T_i)}{B(t, T_i)} &= r_t dt + \sigma \int_t^{T_i} e^{-\kappa(T_i-u)} du dW_t^i \\ &= r_t dt - \frac{\sigma}{\kappa}(1 - e^{-\kappa(T_i-t)})dW_t^i, \quad 0 \leq t \leq T_i. \end{aligned}$$

Ainsi

$$\sigma_i(t) = -\frac{\sigma}{\kappa}(1 - e^{-\kappa(T_i-t)}), \quad 0 \leq t \leq T_i.$$

La diffusion de  $(r_t)_{t \in \mathbb{R}^+}$  sous  $\mathbb{Q}_i$  est

$$dr_t = \kappa(\theta - r_t)dt - \frac{\sigma^2}{\kappa}(1 - e^{-\kappa(T_i-t)})dt + \sigma d\tilde{W}_t^i$$

et la diffusion du prix de l'obligation sous  $\mathbb{Q}_i$  est

$$\frac{dB(t, T_i)}{B(t, T_i)} = r_t dt + \frac{\sigma^2}{\kappa^2}(1 - e^{-\kappa(T_i-t)})^2 dt - \frac{\sigma}{\kappa}(1 - e^{-\kappa(T_i-t)})d\tilde{W}_t^i.$$

#### 2.4.2 Prix d'une option sur une obligation

Nous donnons ici une forme générale pour valoriser des options sur obligations.

**Proposition 2.2.** *Le prix d'un call sur  $B(T_i, T_j)$  en  $t \leq T_i$ , de payoff  $(B(T_i, T_j) - \kappa)_+$  est donné par,*

$$\begin{aligned} & \mathbb{E} \left( e^{-\int_t^{T_i} r_s ds} (B(T_i, T_j) - \kappa)_+ | \mathcal{F}_t \right) \\ &= B(t, T_j) \phi \left( \frac{v}{2} + \frac{1}{v} \log \frac{B(t, T_j)}{\kappa B(t, T_i)} \right) - B(t, T_i) \kappa \phi \left( -\frac{v}{2} + \frac{1}{v} \log \frac{B(t, T_j)}{\kappa B(t, T_i)} \right) \end{aligned}$$

avec

$$v^2(t, T_i, T_j) = \int_t^{T_i} |\sigma_j(s) - \sigma_i(s)|^2 ds$$

et  $\phi$  la densité d'une loi normale centrée réduite.

**Preuve 2.6.** *D'après (2.49)*

$$\mathbb{E} \left( e^{-\int_t^{T_i} r_s ds} (B(T_i, T_j) - \kappa)_+ | \mathcal{F}_t \right) = B(t, T_i) \mathbb{E}_i ((B(T_i, T_j) - \kappa)_+ | \mathcal{F}_t). \quad (2.50)$$

Ensuite, en utilisant le lemme d'Itô sur la fonction  $f(x) = x^{-1}$  nous avons

$$\begin{aligned} d \frac{1}{B(t, T_i)} &= -\frac{1}{B(t, T_i)^2} dB(t, T_i) + \frac{1}{B(t, T_i)^3} \langle dB(t, T_i) \rangle \\ &= -\frac{1}{B(t, T_i)} (r_t dt + \sigma_i(t) dW_t) + \frac{\sigma_i^2(t)}{B(t, T_i)} dt \\ &= -\frac{1}{B(t, T_i)} (r_t dt + \sigma_i(t) (dW_t^i + \sigma_i(t))) + \frac{\sigma_i^2(t)}{B(t, T_i)} dt \\ &= -\frac{1}{B(t, T_i)} (r_t dt + \sigma_i(t) dW_t^i) \end{aligned}$$

Ainsi,

$$\begin{aligned} d \left( \frac{B(t, T_j)}{B(t, T_i)} \right) &= \frac{1}{B(t, T_i)} dB(t, T_j) + B(t, T_j) d \frac{1}{B(t, T_i)} + \langle dB(t, T_j), d \frac{1}{B(t, T_i)} \rangle \\ &= \frac{B(t, T_j)}{B(t, T_i)} (\sigma_j dW_t - \sigma_i dW_t^i) - \frac{B(t, T_j)}{B(t, T_i)} \sigma_j \sigma_i dt \\ &= \frac{B(t, T_j)}{B(t, T_i)} (\sigma_j - \sigma_i) dW_t^i. \end{aligned}$$

En appliquant une nouvelle fois le lemme d'Itô nous avons

$$\frac{B(t, T_j)}{B(t, T_i)} = \frac{B(0, T_j)}{B(0, T_i)} \exp \left( \int_0^{T_i} (\sigma_j(s) - \sigma_i(s)) dW_t^i - \frac{1}{2} \int_0^{T_i} (\sigma_j(s) - \sigma_i(s))^2 ds \right).$$

Et donc

$$\frac{B(u, T_j)}{B(u, T_i)} = \frac{B(t, T_j)}{B(t, T_i)} \exp \left( \int_0^{T_i} (\sigma_j(s) - \sigma_i(s)) dW_s^i - \frac{1}{2} \int_0^{T_i} (\sigma_j(s) - \sigma_i(s))^2 ds \right)$$

pour tout  $u \geq t$ , et en particulier pour  $u = T_i$  nous obtenons

$$B(T_i, T_j) = \frac{B(t, T_j)}{B(t, T_i)} \exp \left( \int_0^{T_i} (\sigma_j(s) - \sigma_i(s)) dW_s^i - \frac{1}{2} \int_0^{T_i} (\sigma_j(s) - \sigma_i(s))^2 ds \right). \quad (2.51)$$

Injectons maintenant (2.51) dans (2.50)

$$\begin{aligned} & \mathbb{E} \left( e^{-\int_t^{T_i} r_s ds} (B(T_i, T_j) - \kappa)_+ | \mathcal{F}_t \right) \\ &= B(t, T_i) \mathbb{E}_i \left( \left( \frac{B(t, T_j)}{B(t, T_i)} e^{\int_0^{T_i} (\sigma_j(s) - \sigma_i(s)) dW_s^i - \frac{1}{2} \int_0^{T_i} (\sigma_j(s) - \sigma_i(s))^2 ds} - \kappa \right)_+ | \mathcal{F}_t \right) \\ &= B(t, T_i) \mathbb{E}_i \left( \left( \frac{B(t, T_j)}{B(t, T_i)} e^{X(t) - \frac{1}{2} v^2(t, T_i, T_j)} - \kappa \right)_+ | \mathcal{F}_t \right) \\ &= B(t, T_i) \mathbb{E}_i \left( \frac{B(t, T_j)}{B(t, T_i)} e^{X(t) - \frac{1}{2} v^2(t, T_i, T_j)} \mathbf{1}_D | \mathcal{F}_t \right) - B(t, T_i) \kappa \mathbb{P}_i(D | \mathcal{F}_t) \\ &= I_1 - I_2. \end{aligned}$$

avec  $X(t)$  une variable aléatoire gaussienne centrée et de variance  $v^2(t, T_i, T_j)$  et  $D$  est donnée par

$$D = \left\{ \frac{B(t, T_j)}{B(t, T_i)} e^{X(t) - \frac{1}{2} v^2(t, T_i, T_j)} > \kappa \right\}$$

Commençons par calculer  $I_2$ .

$$\begin{aligned} I_2 &= B(t, T_i) \kappa \mathbb{P}_i(D | \mathcal{F}_t) \\ &= B(t, T_i) \kappa \mathbb{P}_i \left( \frac{B(t, T_j)}{B(t, T_i)} e^{X(t) - \frac{1}{2} v^2(t, T_i, T_j)} > \kappa | \mathcal{F}_t \right) \\ &= B(t, T_i) \kappa \mathbb{P}_i \left( X(t) < -\frac{v}{2} + \frac{1}{v} \log \frac{B(t, T_j)}{\kappa B(t, T_i)} | \mathcal{F}_t \right) \\ &= B(t, T_i) \kappa \phi \left( -\frac{v}{2} + \frac{1}{v} \log \frac{B(t, T_j)}{\kappa B(t, T_i)} \right) \end{aligned}$$

Pour calculer  $I_1$  nous avons besoin d'introduire la densité de probabilité  $L_{T_i}$  définie par :

$$\begin{aligned} L_{T_i} &= \exp \left( \int_0^{T_i} (\sigma_j(s) - \sigma_i(s)) dW_s^i - \frac{1}{2} \int_0^{T_i} (\sigma_j(s) - \sigma_i(s))^2 ds \right) \\ &= \exp \left( X - \frac{1}{2} v^2(t, T_i, T_j) \right) \end{aligned}$$

Alors si  $\tilde{\mathbb{P}}_i$  désigne la probabilité de densité  $L_{T_i}$  par rapport à  $\mathbb{P}_i$ , nous avons  $\tilde{X}(t) = X(t) - \int_0^t (\sigma_j(s) - \sigma_i(s)) ds$  et donc,

$$\begin{aligned} I_1 &= B(t, T_i) \mathbb{E}_i \left( \frac{B(t, T_j)}{B(t, T_i)} e^{X(t) - \frac{1}{2} v^2(t, T_i, T_j)} \mathbf{1}_D | \mathcal{F}_t \right) \\ &= B(t, T_j) \tilde{\mathbb{P}}_i (\tilde{D} | \mathcal{F}_t) \\ &= B(t, T_j) \phi \left( \frac{v}{2} + \frac{1}{v} \log \frac{B(t, T_j)}{\kappa B(t, T_i)} \right). \end{aligned}$$

#### 2.4.3 Pricing de cap et floor

Rappelons qu'un caplet est une option d'achat sur le forward  $f(T, T, T + \delta)$  de strike  $\kappa$ ,  $Cpl_t = \delta(f(T, T, T + \delta) - \kappa)_+$ . Si nous prenons le cas du Libor  $L(T, T, T + \delta)$

$$Cpl_t = \delta(L(T, T, T + \delta) - \kappa)_+.$$

Le modèle BGM [BrGaMu97], propose de diffuser le Libor de tenor  $T_1 < \dots < T_n$  par un Brownien géométrique sous la mesure  $T_{i+1}$ -forward, i.e.

$$\frac{dL(t, T_i, T_{i+1})}{L(t, T_i, T_{i+1})} = \gamma_i(t) dW_t^{i+1}, \quad 0 \leq t \leq T_i,$$

où  $\gamma_i(t)$  est une fonction déterministe pour tout  $i$ . La solution est donc

$$L(u, T_i, T_{i+1}) = L(t, T_i, T_{i+1}) \exp \left( \int_t^u \gamma_i(s) dW_s^{i+1} - \frac{1}{2} \int_t^u \gamma_i(s)^2 ds \right)$$

pour tout  $u \geq t$ , et en particulier pour  $u = T_i$ , nous avons

$$L(T_i, T_i, T_{i+1}) = L(t, T_i, T_{i+1}) \exp \left( \int_t^{T_i} \gamma_i(s) dW_s^{i+1} - \frac{1}{2} \int_t^{T_i} \gamma_i(s)^2 ds \right).$$

Ainsi, le taux Libor  $L(T_i, T_i, T_{i+1})$  étant un mouvement Brownien géométrique sous  $\mathbb{Q}_{i+1}$ , les caplets peuvent pricer pour tout  $t \in [0, T_i]$  par la formule de Black & Scholes.

**Proposition 2.3.** *Le caplet sur  $L(T_i, T_i, T_{i+1})$  est valorisé à l'instant  $t \in [0, T_i]$  par*

$$\begin{aligned} & \mathbb{E} \left[ (T_{i+1} - T_i) e^{-\int_t^{T_{i+1}} (L(T_i, T_i, T_{i+1}) - \kappa)_+ | \mathcal{F}_t} \right] \\ &= (T_{i+1} - T_i) B(t, T_{i+1}) \{ L(t, T_i, T_{i+1}) \phi(d_+) - \kappa \phi(d_-) \} \end{aligned} \quad (2.52)$$

où

$$d_{\pm} = \frac{\log(L(t, T_i, T_{i+1})/\kappa) \pm \sigma_i^2(t)(T_i - t)/2}{\sigma_i(t)\sqrt{T_i - t}},$$

et

$$\sigma_i^2(t) = \frac{1}{T_i - t} \int_t^{T_i} \gamma_i(s) ds.$$

**Preuve 2.7.** *Comme précédemment, sous la mesure  $T_{i+1}$ -forward, nous avons,*

$$\begin{aligned} & \mathbb{E} \left[ (T_{i+1} - T_i) e^{-\int_t^{T_{i+1}} (L(T_i, T_i, T_{i+1}) - \kappa)_+ | \mathcal{F}_t} \right] \\ &= (T_{i+1} - T_i) B(t, T_{i+1}) \mathbb{E}_{i+1} [(L(T_i, T_i, T_{i+1}) - \kappa)_+ | \mathcal{F}_t] \\ &= (T_{i+1} - T_i) B(t, T_{i+1}) Bl(\kappa, L(t, T_i, T_{i+1}), \sigma_i(t), 0, T_i - t) \end{aligned}$$

pour tout  $t \leq T_i$  et  $Bl$  la fonction de Black & Scholes

$$Bl(\kappa, x, \sigma, r, \tau) = x \phi(d_+) - \kappa e^{-r\tau} \phi(d_-)$$

et

$$\sigma_i^2(t) = \frac{1}{T_i - t} \int_t^{T_i} \gamma_i^2(s) ds.$$

De la même manière que pour le cas d'un call ou d'un put, nous pouvons valoriser un floorlet en calculant à l'aide la formule de Black & Scholes pour tout  $t \in [0, T_i]$

$$\begin{aligned} & \mathbb{E} \left[ (T_{i+1} - T_i) e^{-\int_t^{T_{i+1}} (\kappa - f(T_i, T_i, T_{i+1}))_+ | \mathcal{F}_t} \right] \\ &= (T_{i+1} - T_i) B(t, T_{i+1}) \mathbb{E}_{i+1} [(f(T_i, T_i, T_{i+1}) - \kappa)_+ | \mathcal{F}_t]. \end{aligned}$$

Considérons maintenant un cap de tenor  $T_1 < \dots < T_n$  de payoff actualisé

$$\sum_{k=1}^{n-1} (T_{k+1} - T_k) e^{-\int_t^{T_{k+1}} r_s ds} (f(T_k, T_k, T_{k+1}))_+.$$

Le pricing d'un cap se déduit simplement du résultat obtenu pour les caplets (2.52)

$$\begin{aligned} \text{Cp}_t &= \mathbb{E} \left[ \sum_{i=1}^{n-1} (T_{i+1} - T_i) e^{-\int_t^{T_{i+1}}} (\kappa - f(T_i, T_i, T_{i+1}))_+ | \mathcal{F}_t \right] \\ &= \sum_{i=1}^{n-1} (T_{i+1} - T_i) \mathbb{E} \left[ e^{-\int_t^{T_{i+1}}} (\kappa - f(T_i, T_i, T_{i+1}))_+ | \mathcal{F}_t \right] \\ &= \sum_{i=1}^{n-1} (T_{i+1} - T_i) B(t, T_{i+1}) \mathbb{E}_{i+1} [(\kappa - f(T_i, T_i, T_{i+1}))_+ | \mathcal{F}_t]. \end{aligned}$$

Dans le cadre du modèle BGM, nous obtenons donc pour tout  $t \in [0, T_1]$

$$\text{Cp}_t = \sum_{i=1}^{n-1} (T_{i+1} - T_i) B(t, T_{i+1}) \text{Bl}(\kappa, L(t, T_i, T_{i+1}), \sigma_i(t), 0, T_i - t).$$

#### 2.4.4 Pricing de swaption

Pour pouvoir pricer une swaption, nous avons besoin d'introduire la mesure forward-swap. Nous considérons le numéraire annualisé

$$B(t, T_i, T_j) = \sum_{k=i}^{j-1} (T_{k+1} - T_k) B(t, T_{k+1}), \quad 0 \leq t \leq T_i.$$

**Proposition 2.4.** *Le numéraire annualisé actualisé*

$$e^{-\int_0^t r_s ds} B(t, T_i, T_j) = e^{-\int_0^t r_s ds} \sum_{k=i}^{j-1} (T_{k+1} - T_k) B(t, T_{k+1}), \quad 0 \leq t \leq T_i$$

est une martingale sous  $\mathbb{Q}$ .

La mesure forward-swap  $\mathbb{Q}_{i,j}$  est définie par

$$\frac{d\mathbb{Q}_{i,j}}{d\mathbb{Q}} = e^{-\int_0^{T_i} r_s ds} \frac{B(T_i, T_i, T_j)}{B(0, T_i, T_j)}$$

pour tout  $1 \leq i < j \leq n$ . Nous avons d'après (2.4) pour tout  $0 \leq t \leq T_i$

$$\begin{aligned} \mathbb{E} \left( \frac{d\mathbb{Q}_{i,j}}{d\mathbb{Q}} \middle| \mathcal{F}_t \right) &= \frac{1}{B(0, T_i, T_j)} \mathbb{E} \left( e^{-\int_0^{T_i} r_s ds} B(T_i, T_i, T_j) \middle| \mathcal{F}_t \right) \\ &= \frac{B(t, T_i, T_j)}{B(0, T_i, T_j)} e^{-\int_0^t r_s ds}. \end{aligned}$$

Egalement,

$$\frac{d\mathbb{Q}_{i,j}|\mathcal{F}_t}{d\mathbb{Q}|\mathcal{F}_t} = e^{-\int_0^{T_i} r_s ds} \frac{B(T_i, T_i, T_j)}{B(t, T_i, T_j)}, \quad 0 \leq t \leq T_{i+1}$$

Nous pouvons aussi montrer que le processus

$$v_k^{i,j}(t) := \frac{B(t, T_k)}{B(t, T_i, T_j)}$$

est une  $\mathcal{F}_t$ -martingale sous  $\mathbb{Q}_{i,j}$ . Ainsi, le taux swap

$$S(t, T_i, T_j) = \frac{B(t, T_i) - B(t, T_j)}{B(t, T_i, T_j)} = v_i^{i,j}(t) - v_j^{i,j}(t), \quad 0 \leq t \leq T_i$$

est une martingale sous  $\mathbb{Q}_{i,j}$ .

En utilisant la mesure forward-swap, nous avons la formule de valorisation pour tout payoff de la forme  $B(T_i, T_i, T_j)g(T_i)$

$$\begin{aligned} \mathbb{E} \left( e^{-\int_t^{T_i} r_s ds} B(T_i, T_i, T_j)g(T_i) \middle| \mathcal{F}_t \right) &= B(t, T_i, T_j) \mathbb{E} \left( g(T_i) \frac{d\mathbb{Q}_{i,j}|\mathcal{F}_t}{d\mathbb{Q}|\mathcal{F}_t} \middle| \mathcal{F}_t \right) \\ &= B(t, T_i, T_j) \mathbb{E}_{i,j}(g|\mathcal{F}_t). \end{aligned} \quad (2.53)$$

Rappelons qu'une swaption sur le taux forward  $f(T_i, T_k, T_{k+1})$  est un contrat de payoff

$$\left( \sum_{k=i}^{j-1} (T_{k+1} - T_k) e^{-\int_{T_i}^{T_{k+1}} r_s ds} (f(T_i, T_k, T_{k+1}) - \kappa) \right)_+,$$

à l'instant  $T_i$ . La valorisation pour tout  $t \in [0, T_i]$  sous la mesure risque-neutre est donc

$$\mathbb{E} \left[ e^{-\int_t^{T_i} r_s ds} \left( \sum_{k=i}^{j-1} (T_{k+1} - T_k) e^{-\int_{T_i}^{T_{k+1}} r_s ds} (f(T_i, T_k, T_{k+1}) - \kappa) \right)_+ \middle| \mathcal{F}_t \right],$$

soit, en d'autre terme,

$$\mathbb{E} \left[ e^{-\int_t^{T_i} r_s ds} \left( \sum_{k=i}^{j-1} (T_{k+1} - T_k) B(T_i, T_{k+1}) (f(T_i, T_k, T_{k+1}) - \kappa) \right)_+ \middle| \mathcal{F}_t \right]. \quad (2.54)$$

Comme pour les cap et floor, nous allons considérer une swaption sur le taux Libor  $L(t, T, S)$ . Cela va nous permettre de réécrire le payoff (2.54) en fonction du taux forward swap  $R_{\text{swap}}(T_i, T_i, T_j)$ .

**Lemme 2.2.** *Le payoff d'une swaption sur le taux Libor  $L(t, T, S)$  peut s'écrire sous la forme*

$$\begin{aligned} & \left( \sum_{k=1}^{j-1} (T_{k+1} - T_k) B(T_i, T_{k+1}) (L(T_i, T_k, T_{k+1}) - \kappa) \right)_+ \\ & = B(T_i, T_i, T_j) (R_{swap}(T_i, T_i, T_j) - \kappa)_+. \end{aligned}$$

**Preuve 2.8.** *Le taux forward swap  $R_{swap}$  est donné par la relation d'équilibre*

$$\sum_{k=1}^{j-1} (T_{k+1} - T_k) B(t, T_{k+1}) (L(t, T_k, T_{k+1}) - R_{swap}(t, T_i, T_j)) = 0.$$

*Nous avons ainsi*

$$\begin{aligned} \sum_{k=1}^{j-1} (T_{k+1} - T_k) B(t, T_{k+1}) L(t, T_k, T_{k+1}) & = R_{swap}(t, T_i, T_j) \sum_{k=1}^{j-1} (T_{k+1} - T_k) B(t, T_{k+1}) \\ & = R_{swap}(t, T_i, T_j) B(t, T_i, T_j). \end{aligned}$$

*Donc*

$$\sum_{k=1}^{j-1} (T_{k+1} - T_k) B(t, T_{k+1}) (L(t, T_k, T_{k+1}) - \kappa) = B(t, T_i, T_j) (R_{swap}(t, T_i, T_j) - \kappa),$$

*en particulier pour  $t = T_i$*

$$\begin{aligned} & \left( \sum_{k=1}^{j-1} (T_{k+1} - T_k) B(T_i, T_{k+1}) (L(T_i, T_k, T_{k+1}) - \kappa) \right)_+ \\ & = B(T_i, T_i, T_j) (R_{swap}(T_i, T_i, T_j) - \kappa)_+. \end{aligned}$$

**Proposition 2.5.** *Le prix de la swaption (2.54) peut se réécrire comme un call européen sur le taux swap forward  $R_{swap}$  sous la mesure forward-swap  $\mathbb{Q}_{i,j}$*

$$B(t, T_i, T_j) \mathbb{E}_{i,j} ((R_{swap}(T_i, T_i, T_j) - \kappa)_+, 0 \leq t \leq T_i).$$

**Preuve 2.9.** *En appliquant l'équation (2.53) et le lemme (2.2) nous obtenons*

$$\begin{aligned} & \mathbb{E} \left[ e^{-\int_t^{T_i} r_s ds} \left( \sum_{k=i}^{j-1} (T_{k+1} - T_k) B(T_i, T_{k+1}) (L(T_i, T_k, T_{k+1}) - \kappa) \right)_+ \middle| \mathcal{F}_t \right] \\ & = \mathbb{E} \left[ e^{-\int_t^{T_i} r_s ds} B(T_i, T_i, T_j) (R_{swap}(T_i, T_i, T_j) - \kappa)_+ \middle| \mathcal{F}_t \right] \\ & = \frac{1}{B(t, T_i, T_j)} \mathbb{E} \left[ \frac{d\mathbb{Q}_{i,j}|\mathcal{F}_t}{d\mathbb{Q}|\mathcal{F}_\approx} (R_{swap}(T_i, T_i, T_j) - \kappa)_+ \middle| \mathcal{F}_t \right] \\ & = B(t, T_i, T_j) \mathbb{E}_{i,j} [(R_{swap}(T_i, T_i, T_j) - \kappa)_+ | \mathcal{F}_t]. \end{aligned}$$

Fort des différents résultats que nous avons obtenus, nous pouvons donner une formule de pricing de swaption dans le cas où le taux swap forward suit un mouvement Brownien géométrique sous la mesure swap-forward.

**Proposition 2.6.** *Supposons que le taux swap forward est modélisé par un mouvement Brownien géométrique sous la mesure  $\mathbb{Q}_{i,j}$*

$$dR_{swap}(t, T_i, T_j) = R_{swap}(t, T_i, T_j)\sigma(t)dW_t^{i,j}$$

avec  $(\sigma(t))_{t \in \mathbb{R}_+}$  une fonction déterministe. Nous pouvons alors valoriser une swaption de payoff

$$B(t, T_i, T_j)(R_{swap}(T_i, T_i, T_j) - \kappa)_+$$

par la formule de Black & Scholes

$$\begin{aligned} & \mathbb{E} \left[ e^{-\int_t^{T_i} r_s ds} B(T_i, T_i, T_j)(R_{swap}(T_i, T_i, T_j) - \kappa)_+ \middle| \mathcal{F}_t \right] \\ &= (B(t, T_i) - B(t, T_j))\phi(d_+) - \kappa\phi(d_-) \sum_{k=1}^{j-2} (T_{k+1} - T_k)B(t, T_{k+1}), \end{aligned}$$

avec

$$\begin{aligned} d_+ &= \frac{\log \left( \frac{R_{swap}(t, T_i, T_j)}{\kappa} \right) + \frac{\sigma_i^2(t)(T_i - t)}{2}}{\sigma_i^2(t)\sqrt{T_i - t}} \\ d_- &= \frac{\log \left( \frac{R_{swap}(t, T_i, T_j)}{\kappa} \right) - \frac{\sigma_i^2(t)(T_i - t)}{2}}{\sigma_i^2(t)\sqrt{T_i - t}} \end{aligned}$$

et

$$\sigma_i^2 = \frac{1}{T_i - t} \int_t^{T_i} \sigma^2(s)ds$$

**Preuve 2.10.** *Nous avions supposé que le taux  $R_{swap}$  était diffusé par un Brownien géométrique de variance  $(\sigma(t))_{t \in \mathbb{R}_+}$  sous  $\mathbb{Q}_{i,j}$ , nous avons*

$$\begin{aligned} & \mathbb{E} \left[ e^{-\int_t^{T_i} r_s ds} B(T_i, T_i, T_j)(R_{swap}(T_i, T_i, T_j) - \kappa)_+ \middle| \mathcal{F}_t \right] \\ &= B(t, T_i, T_j) \mathbb{E}_{i,j} [(R_{swap}(T_i, T_i, T_j) - \kappa)_+ | \mathcal{F}_t] \\ &= B(t, T_i, T_j) Bl(\kappa, R_{swap}(T_i, T_i, T_j), \sigma_i(t), 0, T_i - 1) \\ &= B(t, T_i, T_j) [R_{swap}(T_i, T_i, T_j)\phi(d_+) - \kappa\phi(d_-)] \\ &= (B(t, T_i) - B(t, T_j))\phi(d_+) - \kappa B(T_i, T_i, T_j)\phi(d_-) \\ &= (B(t, T_i) - B(t, T_j))\phi(d_+) - \kappa\phi(d_-) \sum_{k=1}^{j-2} (T_{k+1} - T_k)B(t, T_{k+1}) \end{aligned}$$

### 2.4.5 Pricing dans le cas à deux facteurs

Nous reprenons le modèle (2.47), i.e.

$$r_t = x_t + y_t$$

avec

$$\begin{aligned} dx_t &= -\kappa_x dt + \sigma_x dW_t^x \\ dy_t &= -\kappa_y dt + \sigma_y dW_t^y, \end{aligned}$$

avec  $\kappa_x, \kappa_y, \sigma_x$  et  $\sigma_y$  des constantes positives et  $(W^x, W^y)$  un mouvement Brownien bi-dimensionnelle de corrélation  $\rho$ . Sous la mesure  $T$ -forward, les processus  $(x_t)_{t \in \mathbb{R}_+}$  et  $(y_t)_{t \in \mathbb{R}_+}$  évoluent selon

$$\begin{aligned} dx_t &= \left( -\kappa_x x_t - \frac{\sigma_x^2}{\kappa_x} (1 - e^{-\kappa_x(T-t)}) - \rho \frac{\sigma_x \sigma_y}{\kappa_y} (1 - e^{-\kappa_y(T-t)}) \right) dt + \sigma_x d\tilde{W}_t^x \\ dy_t &= \left( -\kappa_y y_t - \frac{\sigma_y^2}{\kappa_y} (1 - e^{-\kappa_y(T-t)}) - \rho \frac{\sigma_x \sigma_y}{\kappa_x} (1 - e^{-\kappa_x(T-t)}) \right) dt + \sigma_y d\tilde{W}_t^y, \end{aligned} \quad (2.55)$$

avec  $(\tilde{W}_t^x)_{t \geq 0}$  et  $(\tilde{W}_t^y)_{t \geq 0}$  deux mouvements browniens sous la mesure  $T$ -forward tels que  $\langle d\tilde{W}_t^x d\tilde{W}_t^y \rangle = \rho dt$ . Les solutions de (2.55) sont

$$\begin{aligned} x_t &= x_s e^{-\kappa_x(t-s)} - M_x(s, t) + \sigma_x \int_s^t e^{-\kappa_x(t-u)} d\tilde{W}_u^x \\ y_t &= y_s e^{-\kappa_y(t-s)} - M_y(s, t) + \sigma_y \int_s^t e^{-\kappa_y(t-u)} d\tilde{W}_u^y, \end{aligned}$$

pour tout  $s \leq t \leq T$  avec

$$\begin{aligned} M_x(s, t) &= \left( \frac{\sigma_x^2}{\kappa_x^2} + \rho \frac{\sigma_x \sigma_y}{\kappa_x \kappa_y} \right) (1 - e^{-\kappa_x(t-s)}) - \frac{\sigma_x^2}{2\kappa_x^2} (e^{-\kappa_x(T-t)} - e^{-\kappa_x(T+t-2s)}) \\ &\quad - \rho \frac{\sigma_x \sigma_y}{\kappa_y (\kappa_x + \kappa_y)} \left( e^{-\kappa_y(T-t)} - e^{-\kappa_y T - \kappa_x t + (\kappa_x + \kappa_y)s} \right) \\ &= \left( \frac{\sigma_y^2}{\kappa_y^2} + \rho \frac{\sigma_x \sigma_y}{\kappa_x \kappa_y} \right) (1 - e^{-\kappa_y(t-s)}) - \frac{\sigma_y^2}{2\kappa_y^2} (e^{-\kappa_y(T-t)} - e^{-\kappa_y(T+t-2s)}) \\ &\quad - \rho \frac{\sigma_x \sigma_y}{\kappa_x (\kappa_x + \kappa_y)} \left( e^{-\kappa_x(T-t)} - e^{-\kappa_x T - \kappa_y t + (\kappa_x + \kappa_y)s} \right) \end{aligned}$$

La distribution est donc une nouvelle fois conditionnellement gaussienne par rapport à  $\mathcal{F}_t$ , en particulier, sous la mesure  $T$ -forward

$$\begin{aligned}\mathbb{E}(r_t | \mathcal{F}_s) &= x_s e^{-\kappa_x(t-s)} - M_x(s, t) + y_s e^{-\kappa_y(t-s)} - M_y(s, t) \\ \mathbb{V}\text{ar}(r_t | \mathcal{F}_s) &= \frac{\sigma_x^2}{2\kappa_x} (1 - e^{-2\kappa_x(t-s)}) + \frac{\sigma_y^2}{2\kappa_y} (1 - e^{-2\kappa_y(t-s)}) \\ &\quad + 2\rho \frac{\sigma_x \sigma_y}{\kappa_x + \kappa_y} (1 - e^{-(\kappa_x + \kappa_y)(t-s)}).\end{aligned}$$

**Preuve 2.11.** *Similaire à la valorisation d'un zéro-coupon avec le modèle de Vasicek à deux facteurs.*

**Théorème 2.2.** *Le prix d'un call de maturité  $T$  et de strike  $\kappa$  sur zéro-coupon de maturité  $S$  dans le cadre du modèle à deux facteurs gaussien est donné à l'instant  $t \leq S$  par*

$$\begin{aligned}\mathbb{E} \left[ e^{\int_t^T r_s ds} ((B(T_i, T_{i+1}) - \kappa)_+) \right] &= B(t, T_{i+1}) \phi \left( \frac{1}{v} \log \frac{B(t, T_{i+1})}{\kappa B(t, T_i)} + \frac{1}{2} v \right) \\ &\quad - B(t, T_i) \kappa \phi \left( \frac{1}{v} \log \frac{B(t, T_{i+1})}{\kappa B(t, T_i)} - \frac{1}{2} v \right)\end{aligned}$$

avec

$$\begin{aligned}v^2 &= \frac{\sigma_x^2}{2\kappa_x^3} (1 - e^{-\kappa_x(T_{i+1}-T_i)})^2 (1 - e^{-2\kappa_x(T_i-t)}) \\ &\quad + \frac{\sigma_y^2}{2\kappa_y^3} (1 - e^{-\kappa_y(T_{i+1}-T_i)})^2 (1 - e^{-2\kappa_y(T_i-t)}) \\ &\quad + 2\rho \frac{\sigma_x \sigma_y}{\kappa_x \kappa_y (\kappa_x + \kappa_y)} (1 - e^{-\kappa_x(T_{i+1}-T_i)}) (1 - e^{-\kappa_y(T_{i+1}-T_i)}) \\ &\quad \times (1 - e^{-(\kappa_x + \kappa_y)(T_{i+1}-T_i)}).\end{aligned}$$

**Preuve 2.12.** *Par analogie avec les calculs faits dans le cas à un facteur et en remarquant que  $B(t, T_i)$  suit une loi log normale, théorème 2.1, de paramètres*

$$\begin{aligned}\mathbb{E}(\log B(t, T_{i+1}) | \mathcal{F}_t) &= -\frac{1 - e^{-\kappa_x(T_{i+1}-T_i)}}{\kappa_x} \mathbb{E}(x_t | \mathcal{F}_t) - \frac{1 - e^{-\kappa_y(T_{i+1}-T_i)}}{\kappa_y} \mathbb{E}(y_t | \mathcal{F}_t) \\ &\quad + \frac{1}{2} V(t, T_{i+1})\end{aligned}$$

et

$$\begin{aligned}
 \mathbb{V}(\log B(t, S) | \mathcal{F}_t) &= \frac{(1 - e^{-\kappa_x(T_{i+1}-T_i)})^2}{2\kappa_x^2} \mathbb{V}(x_t | \mathcal{F}_t) \\
 &\quad + \frac{(1 - e^{-\kappa_y(T_{i+1}-T_i)})^2}{2\kappa_y^2} \mathbb{V}(y_t | \mathcal{F}_t) \\
 &\quad + \frac{1 - e^{-\kappa_x(T_{i+1}-T_i)}}{\kappa_x} \frac{1 - e^{-\kappa_y(T_{i+1}-T_i)}}{\kappa_y} \mathbb{C}(x_t, y_t | \mathcal{F}_t) \\
 &= \frac{\sigma_x^2}{\kappa_x^3} (1 - e^{-\kappa_x(T_{i+1}-T_i)})^2 (1 - e^{-2\kappa_x(T_{i+1}-T_i)}) \\
 &\quad + \frac{\sigma_y^2}{\kappa_y^3} (1 - e^{-\kappa_y(T_{i+1}-T_i)})^2 (1 - e^{-2\kappa_y(T_i-t)}) \\
 &\quad + 2\rho \frac{\sigma_x \sigma_y}{\kappa_x \kappa_y (\kappa_x + \kappa_y)} (1 - e^{-\kappa_x(T_{i+1}-T_i)}) (1 - e^{-\kappa_y(T_{i+1}-T_i)}) \\
 &\quad \times (1 - e^{-(\kappa_x + \kappa_y)(T_{i+1}-T_i)}).
 \end{aligned}$$

## 2.5 Heath-Jarrow-Morton Framework

Le prix de l'obligation  $B(t, T)$ , le rendement  $R(t, T)$  et le taux forward  $F(t, T)$ , nous donnent tous la même information sur la courbe des taux. Le framework de Heath-Jarrow-Morton (HJM), [HeJaMo90], [HeJaMo91] et [HeJaMo92] tente de construire une famille de processus stochastiques à temps continu pour la structure par termes, cohérente avec les observations initiales. La variable d'état du modèle choisi est le taux instantané  $F(t, T)$ . La diffusion générale choisie par les auteurs est,

$$F(t, T) = F(0, T) + \int_0^t \alpha_F(u, T) du + \int_0^t \sigma_F(u, T) dW_u,$$

et sous forme différentielle,

$$dF(t, T) = \alpha_F(t, T) dt + \sigma_F(t, T) dW_t, \quad 0 \leq t \leq T. \quad (2.56)$$

$F(0, T)$  est le taux forward initial,  $\alpha_F(t, T)$  la tendance du taux forward instantané et  $\sigma_F(t, T)$  sa volatilité. Nous supposerons les deux processus  $\alpha_F$  et  $\sigma_F$  adaptés à leur filtration naturelle.

La diffusion (2.56) ne donne pas nécessairement un non-arbitrage, et, qu'il existe donc une probabilité risque neutre. Nous allons donner une condition dans la prochaine section.

### 2.5.1 Condition de non-arbitrage

Sous la mesure risque-neutre, la tendance du prix de l'obligation est  $r_t$ , (2.27), (2.28). La dynamique du prix prend alors la forme,

$$\frac{dB(t, T)}{B(t, T)} = r_t dt + \sigma_B(t, T) dW_t, \quad 0 \leq t \leq T, \quad (2.57)$$

avec  $\sigma_B$  un processus adapté. Nous pouvons également rajouter comme condition terminale  $\sigma_B(T, T) = 0$ . Il s'agit de l'effet Pull-to-Par, plus l'obligation va arriver à maturité, plus le prix va se stabiliser vers le par (sa valeur faciale), et donc, la volatilité va disparaître. Rappelons que le prix forward est donné par

$$df(t, T_1, t_2) = \frac{d \ln B(t, T_1) - d \ln B(t, T_2)}{T_2 - T_1}, \quad (2.58)$$

par le lemme d'Itô, nous obtenons,

$$d \ln B(t, T_i) = \left( r_t - \frac{\sigma_B(t, T_i)^2}{2} \right) dt + \sigma_B(t, T_i) dW_t, \quad i = 1, 2.$$

Ainsi, nous pouvons réécrire l'équation (2.58),

$$df(t, T_1, t_2) = \frac{\sigma_B(t, T_2)^2 - \sigma_B(t, T_1)^2}{2(T_2 - T_1)} dt + \frac{\sigma_B(t, T_1) - \sigma_B(t, T_2)}{T_2 - T_1} dW_t.$$

Posons maintenant  $T_1 = T$  et  $T_2 = T + \Delta$ , en passant à la limite,

$$\lim_{\Delta t \rightarrow 0} \frac{\sigma_B(t, T) - \sigma_B(t, T + \Delta t)}{\Delta t} = -\frac{\partial \sigma_B}{\partial T}(t, T)$$

et

$$\lim_{\Delta t \rightarrow 0} \frac{\sigma_B(t, T + \Delta t)^2 - \sigma_B(t, T)^2}{2\Delta t} = \sigma_B(t, T) \frac{\partial \sigma_B}{\partial T}(t, T)$$

$(W_t)_{t \in \mathbb{R}_+}$  étant un processus de Wiener, nous pouvons changer le bruit sans perte de généralité. On peut alors écrire la diffusion du taux forward instantané,

$$dF(t, T) = \sigma_B(t, T) \frac{\partial \sigma_B}{\partial T}(t, T) dt + \frac{\partial \sigma_B}{\partial T}(t, T) dW_t, \quad (2.59)$$

En égalisant les équations (2.56) et (2.59), nous devons avoir,

$$\sigma_F(t, T) = \frac{\partial \sigma_B}{\partial T}(t, T), \quad \alpha_F(t, T) = \sigma_B(t, T) \frac{\partial \sigma_B}{\partial T}(t, T). \quad (2.60)$$

La condition de non-arbitrage dans le framework HJM est alors,

$$\alpha_F(t, T) = \sigma_F(t, T) \int_t^T \sigma_F(t, u) du. \quad (2.61)$$

Les modèles de taux d'intérêt HJM sont donc entièrement spécifiés par le sous-jacent et, la volatilité. Le sous-jacent est la structure par terme, la volatilité décrivant quant à elle comment cette structure par terme évolue au cours du temps.

### 2.5.2 Formulation des taux court avec HJM

Les taux cours à l'instant  $t$  sont égaux aux forward de maturité  $t$ ,

$$r_t = F(t, t) = F(0, t) + \int_0^t dF(u, t) du.$$

En utilisant (2.59),

$$r_t = F(0, t) + \int_0^t \sigma_B(u, t) \frac{\partial \sigma_B}{\partial t}(u, t) du + \int_0^t \frac{\partial \sigma_B}{\partial t}(u, t) dW_u.$$

En différentiant,

$$\begin{aligned} dr_t &= \frac{\partial F}{\partial t}(0, t) dt + \left\{ \int_0^t \sigma_B(u, t) \frac{\partial^2 \sigma_B}{\partial t^2}(u, t) du + \left( \frac{\partial \sigma_B}{\partial t}(u, t) \right)^2 du \right\} dt \\ &\quad + \left\{ \int_0^t \frac{\partial^2 \sigma_B}{\partial t^2}(u, t) dW_u \right\} dt + \frac{\partial \sigma_B}{\partial t}(u, t)|_{u=t} dW_u. \end{aligned} \quad (2.62)$$

Considérons la forme suivante pour la volatilité,

$$\sigma_B(t, T) = \sigma(T - t),$$

$\sigma$  une constante. Avec la méthodologie HJM, la diffusion du taux instantané (2.62) est,

$$dr_t = \left( \frac{\partial F}{\partial t}(0, t) + \sigma^2 t \right) dt + \sigma dW_t.$$

On retrouve le modèle en temps continu de Ho-Lee.

### 2.5.3 Taux forward de Vasicek

La valeur d'un zéro coupon dans le modèle de Vasicek est donnée par (2.36). Le taux forward est alors (après quelques calculs),

$$\begin{aligned} f(t, T, S) &= -\frac{\log B(t, S) - \log B(t, T)}{S - T} \\ &= -R_\infty \\ &\quad - \frac{1}{S - T} \left[ \left( \frac{r_t}{\kappa} + \frac{R_\infty}{\kappa} - \frac{\sigma^2}{2\kappa^3} \right) \left( e^{-\kappa(S-t)} - e^{-\kappa(T-t)} \right) - \frac{\sigma^2}{4\kappa^3} \left( e^{-2\kappa(S-t)} - e^{-2\kappa(T-t)} \right) \right]. \end{aligned}$$

Le taux instantané est,

$$\begin{aligned} F(t, T) &= -\frac{\partial \log B(t, T)}{\partial T} \\ &= r_t e^{-\kappa(T-t)} + \theta \left( 1 - e^{-\kappa(T-t)} \right) - \frac{\sigma^2}{2\kappa^2} \left( 1 - e^{-\kappa(T-t)} \right)^2. \end{aligned}$$

En supposant  $\theta = 0$ , nous avons la forme différentielle,

$$dF(t, T) = \sigma^2 e^{-\kappa(T-t)} \int_t^T e^{-\kappa(t-s)} ds dt + \sigma e^{-b(T-t)} dW_t.$$

Nous pouvons alors retrouver la condition de HJM,

$$\alpha(t, T) = \sigma^2 e^{-\kappa(T-t)} \int_t^T e^{-\kappa(s-t)} ds = \sigma(t, T) \int_t^T \sigma(t, s) ds,$$

où

$$\sigma(t, T) = \sigma e^{-\kappa(T-t)}. \tag{2.63}$$

## 2.6 Courbe des taux

Il existe plusieurs moyens de reconstruire la courbe des taux. Commençons déjà par expliquer pourquoi, dans le Fixed Income, il est important de pouvoir reconstruire la courbe des taux. Remarquons déjà que nous ne disposons pas de zéro-coupon à toutes les maturités, par exemple 1 ans, 2 mois et 3 jours. Nous n'avons qu'un échantillon de points, que nous souhaitons donc relier pour avoir une courbe continue. Cela permet de construire des portefeuilles robustes en termes de

Risk management.

Les méthodes les plus connues sont les méthodes d'interpolations linéaires et cubiques, de bootstrapping et, les méthodes paramétriques de type Svensson. Nous présentons rapidement ces différentes méthodes.

### 2.6.1 Interpolation

Les méthodes d'interpolations consistent simplement à retrouver le taux  $R(t, \tau)$  connaissant  $R(t, t_{i-1})$  et  $R(t, t_i)$  pour tout  $t_{i-1} < \tau < t_i$ . Le plus simple étant naturellement l'interpolation linéaire

$$R(t, \tau) = \frac{\tau - t_{i-1}}{t_i - t_{i-1}} R(t, t_{i-1}) + \frac{t_i - \tau}{t_i - t_{i-1}} R(t, t_i)$$

Le problème majeur résultant de cette méthode est que le taux forward déduit n'est plus différentiable en tout point.

L'interpolation par splines cubiques prend la forme

$$R(t, \tau) = a_i + b_i(\tau - t_i) + c_i(\tau - t_i)^2 + d_i(\tau - t_i)^3, \quad t_i \leq \tau \leq t_{i+1}.$$

Pour que l'interpolation soit cohérente, nous devons instaurer quelques contraintes. La fonction doit rencontrer tous les points connus, i.e.  $a_i = R(t, t_i)$  pour tout  $i < n-1$  et  $a_{n-1} + b_{n-1}(t_{i+1} - t_i) + c_{n-1}(t_{i+1} - t_i)^2 + d_{n-1}(t_{i+1} - t_i)^3 = R(t, t_n) = a_n$ , également, la fonction doit être continue et différentiable en tout point. Après quelques calculs, il est possible de trouver une forme explicite pour les paramètres  $(a_i, b_i, c_i, d_i)$ . Malheureusement, bien que les méthodes par splines soient supposées moins sujettes au sur-apprentissage et aux fortes oscillations entre deux points, les conditions que nous devons imposer en  $t_1$  et  $t_n$  peuvent compromettre la qualité du résultat. Il est également possible de calibrer les paramètres  $(a_i, b_i, c_i, d_i)$  par moindres carrés pénalisés pour s'assurer d'une bonne stabilité du résultat. Pour plus de détails sur ces méthodes, nous renvoyons à [HaWe08].

### 2.6.2 Bootstrapping

Le principe du bootstrap pour reconstruire la courbe des taux consiste à partir des taux à petites maturités, obtenir les taux de maturités supérieures. L'idée

repose simplement qu'une obligation versant différents coupons fixes est donnée par

$$\begin{aligned} B_n(t, T_n) &= \sum_{i=1}^n B(t, T_i) c_i + B(t, T_n) \\ &= \sum_{i=1}^n c_i e^{-(T_i-t)R(t, T_i)} + e^{-(T_n-t)R(t, T_n)}, \quad t \leq T_1. \end{aligned}$$

Ainsi,

$$R(t, T_n) = -\frac{1}{T_n - t} \log \frac{B_n(t, T_n) - \sum_{i=1}^{n-1} c_i e^{-(T_i-t)R(t, T_i)}}{1 + c_n}.$$

Pour les taux courts, sur le marché européen, nous prenons en général les taux Euribor (1w à 3m) et l'Eonia (Euro OverNight Index Average) pour le taux au jour le jour. Pour les taux moyens nous pouvons prendre indifféremment des futures ou forward. Il semble que les futures étant plus liquides et cotées sur le marché et non de gré à gré, ils conviennent mieux car plus proches des valeurs efficientes. Néanmoins, déduire les taux forward en fonction des taux futures n'est pas facilement réalisable car il a ce que l'on appelle un biais de convexité entre ces deux instruments, les forwards ont une convexité supérieure. Pour des raisons de liquidité, nous utilisons les taux swap pour des maturités supérieures à 2 ans, bien que les maturités à 6 et 9 ans sont généralement peu liquides et peuvent considérablement diminuer la qualité du résultat.

### 2.6.3 Familles de polynôme exponentiel

La matrice de variance covariance  $\Sigma$  a pour entrées  $x(1), \dots, x(N)$ , les incrément de la courbe forward. En diagonalisant la matrice  $\Sigma$  (non singulière), nous retrouvons les valeurs propres et les vecteurs propres. On constate généralement que les trois premières valeurs propres expliquent à elles seules plus de 95% de la variation totale. Empiriquement, nous pouvons interpréter ces trois facteurs comme suit :

- Le premier facteur a une forme relativement plate, il donne le niveau moyen de la courbe.
- Le second donne les inversions de la courbe autour d'un point, il s'agit d'un facteur de pente.
- Le troisième facteur est un facteur de courbure.

Cela nous permet d'avoir une idée pour proposer une forme paramétrique de la courbe des taux.

Les polynômes exponentiels sont des fonctions du type

$$p_1(x)e^{-\alpha_1 x} + \cdots + p_n(x)e^{-\alpha_1 x},$$

où les  $p_i$ ,  $i = 1, \dots, n$  sont des polynômes de degré  $n_i$ .

L'un des premiers modèles est celui de Nelson-Siegel [NeSi87], la courbe des taux forward est

$$f(t, t+m) = \beta_0 + \beta_1 \exp\left(-\frac{m}{\tau_1}\right) + \beta_2 \frac{m}{\tau_1} \exp\left(-\frac{m}{\tau_1}\right)$$

Les paramètres  $\beta_0$  et  $\tau_1$  sont supposés positifs. Le taux forward a donc trois composants. Le premier correspond au niveau moyen, le second est un terme exponentiel décroissant si  $\beta_1 < 0$  (croissant si  $\beta_1 > 0$ ) jusque zéro comme une fonction du temps de livraison. Le troisième terme nous donne la courbure, voûtée dans le cas où  $\beta_2 > 0$ , en U si négatif. Svensson [Sv94], propose de rajouter un 4ème terme pour laisser plus de flexibilité lors de la calibration

$$f(t, T) = \beta_0 + \beta_1 \exp\left(-\frac{T}{\tau_1}\right) + \beta_2 \frac{T}{\tau_1} \exp\left(-\frac{T}{\tau_1}\right) + \beta_3 \frac{T}{\tau_2} \exp\left(-\frac{T}{\tau_2}\right)$$

Le modèle comprend donc non plus 4 paramètres à estimer mais 6 et la courbure est donnée par  $\beta_2 \frac{m}{\tau_1} \exp\left(-\frac{m}{\tau_1}\right) + \beta_3 \frac{m}{\tau_2} \exp\left(-\frac{m}{\tau_2}\right)$ . La courbe des taux est alors

$$R(t, T) = \beta_0 + \beta_1 \frac{1 - e^{-\frac{T}{\tau_1}}}{\frac{T}{\tau_1}} + \beta_2 \left( \frac{1 - e^{-\frac{T}{\tau_1}}}{\frac{T}{\tau_1}} - e^{-\frac{T}{\tau_1}} \right) + \beta_3 \left( \frac{1 - e^{-\frac{T}{\tau_2}}}{\frac{T}{\tau_2}} - e^{-\frac{T}{\tau_2}} \right).$$

La différence à retenir entre les deux modèles est donc que le premier, celui de Nelson-Neigel a 'une courbure' (one-hump) tandis que celui de Svensson a 'deux courbures' (two-hump). Nous présentons sur la figure (3) l'influence du premier et second/troisième terme, i.e. respectivement  $\beta_1 \exp\left(-\frac{T}{\tau_1}\right)$  et  $\beta_2 \frac{T}{\tau_1} \exp\left(-\frac{T}{\tau_1}\right)$ . Sur la figure (4), nous présentons la calibration des deux modèles sur les données réelles des taux gouvernementaux américains le 09/03/13.

En fait, cette famille n'est pas simplement présente dans la théorie des modèles de taux pour reconstruire la courbe de taux, elle est beaucoup plus large que cela. Les modèles que nous avons vus précédemment, nous avons supposé que la solution du zéro-coupon était du type

$$B(t, T) = A(T - t)e^{-C(T-t)r_t},$$

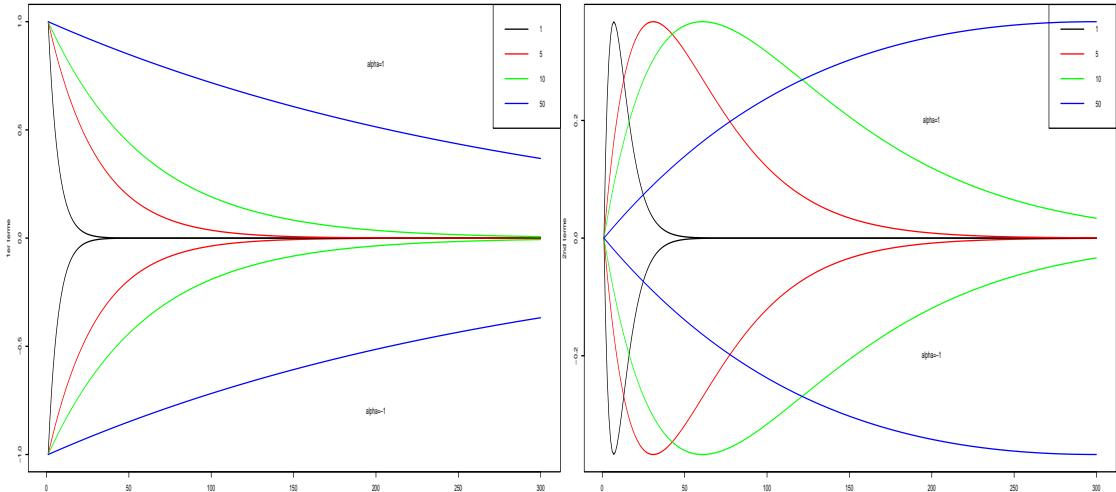


FIGURE 3 – A gauche, influence du paramètre  $\tau_1$  sur le premier terme, à droite, influence de  $\tau_1$  sur le second terme de l'équation de Svensson.

cela ne sort pas d'un chapeau. Il est possible de montrer que le type de diffusion que nous avons étudié (pour les modèles à un facteur) ont toutes une solution de cette forme. Cette forme et donc, cette classe de modèle, est dites de structure par terme affine. En effet, le YTM est alors du type

$$R(t, T) = -\frac{1}{T-t} (\ln(A(T-t)) - C(T-t)r_t),$$

soit, une fonction affine. La famille de modèle polynôme exponentiel propose donc des modèles à structure par terme polynomiale exponentiel. Cela étant relativement 'long' à expliquer correctement, mais néanmoins très intéressant, nous renvoyons au lecteur à [Fi00] où [Pr12]. Cette théorie est également relativement récente, comme nous l'avons notifiée tout au long de ce cours, il est loin d'exister un consensus sur quel modèle utiliser, à la différence des produits sur equity avec les modèles de Black & Scholes et Heston qui se taillent une bonne part du lion pour les modèles de pricing à temps continu, ou GARCH pour les modèles discrets.

#### 2.6.4 Méthodes utilisées par pays

Toutes les banques centrales n'estiment pas leur courbe de taux ni par la même méthode, ni avec le même spectre, c'est-à-dire l'éventail de maturités prisent en

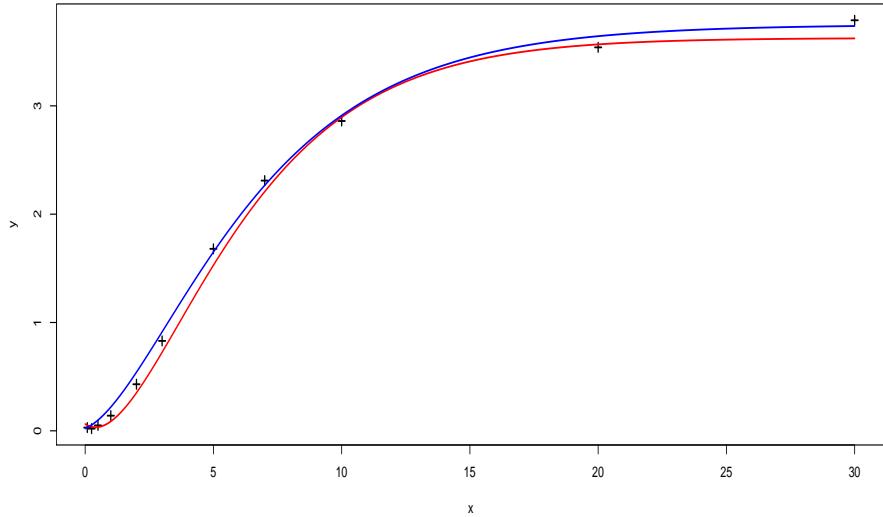


FIGURE 4 – En bleue, équation de Svensson, en rouge celle de Nelson-Siegel, les points noirs correspondent aux taux gouvernementaux américains de tenor 1m, 3m, 6m, 1Y, 2Y, 3Y, 5Y, 7Y, 10Y, 20Y, 30Y en date du 09/03/13.

compte. Nous présentons les différents modèles utilisés selon les pays sur le tableau (2). Cela a au moins le mérite de nous montrer qu'il n'existe pas encore de consensus sur ce type de problème (et sur les modèles de taux tout court d'ailleurs). Même si certains choix sont discutables, il n'y a pas véritablement un pays qui a plus raison qu'un autre.

### 3 Valorisation sur action

Commençons par un rapide rappel sur le modèle de Black and Scholes (B&S).

#### 3.1 Modèle à volatilité constante

##### 3.1.1 Modèle de Black and Scholes

Notons  $r$  le taux sans risque constant sur le marché,  $\mu$  la tendance de l'actif  $S$  et  $\sigma$  sa volatilité,  $\mathbb{P}$  la probabilité historique et  $\mathbb{Q}$  la probabilité risque-neutre. La

Banque Centrale	Méthode	Maturité min	Maturité max
Belgique	Svensson ou Nelson-Siegel	quelques jours	16 ans
Canada	Merrill Lynch Exponential Spline	1 mois	30 ans
Finlande	Nelson-Siegel	1 jour	12 ans
France	Svensson ou Nelson-Siegel	1 mois	10 ans
Allemagne	Svensson	3 mois	10 ans
Italie	Nelson-Siegel	1 mois	30 ans
Japon	Splines	1 jour	10 ans
Norvège	Svensson	1 mois	10 ans
Espagne	Svensson	1 jour	10 ans
Suède	Spline ou Svensson	1 jour	10 ans
Suisse	Svensson	1 jour	30 ans
Royaume-Uni	VRP*	1 semaine	30 ans
Etats-Unis	Spline multiples	1 ans	10 ans

TABLE 2 – Méthodologie utilisée par pays. Source : BIS, 2005. \* Variable Roughness penalty (VRP), il s'agit d'une méthode de splines proposée par Waggoner [Wa97].

diffusion de B&S est

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

où  $W_t$  est un  $\mathbb{P}$  - mouvement brownien standard. Pour passer en univers risque neutre, nous introduisons la quantité  $\lambda$  tel que  $\tilde{W}_t = W_t - \lambda t$  soit un  $\mathbb{Q}$  - mouvement brownien standard (Formule Cameron-Martin 6.1). Ainsi

$$dS_t = (\mu - \lambda\sigma)S_t dt + \sigma S_t d\tilde{W}_t.$$

Comme nous sommes en univers risque-neutre, les rendements doivent avoir une espérance égale au taux sans risque, la prime de risque  $\lambda$  est donc choisie comme

$$\lambda = \frac{\mu - r}{\sigma},$$

soit, le ratio de Sharpe. Après avoir appliqué le lemme d'Itô à la fonction  $\ln(S_t)$  nous trouvons la solution de la diffusion

$$S_T = S_t e^{\left(r - \frac{\sigma^2}{2}\right)(T-t) + \sigma \tilde{W}_{T-t}}.$$

Le prix d'un call de maturité  $T$  et de strike  $K$  sur  $S_t$  est donné par

$$\begin{aligned} C(t, S_t) &= \mathbb{E}^{\mathbb{Q}}[e^{-r(T-t)}(S_T - K)_+ | \mathcal{F}_t] \\ &= \mathbb{E}^{\mathbb{Q}}[e^{-r(T-t)}S_T \mathbb{1}_{S_T > K} | \mathcal{F}_t] - Ke^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[\mathbb{1}_{S_T > K} | \mathcal{F}_t] \\ &= S_t \mathbb{E}^{\mathbb{Q}}[e^{\sigma \tilde{W}_{T-t} - \frac{\sigma^2}{2}(T-t)} \mathbb{1}_{S_T > K}] - Ke^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[\mathbb{1}_{S_T > K}] \end{aligned}$$

Pour calculer le premier terme, nous introduisons la densité

$$L = e^{\int_0^{T-t} \sigma d\tilde{W}_s - \frac{1}{2} \int_0^{T-t} \sigma^2 ds}$$

Alors si  $\mathbb{Q}'$  est la probabilité de densité  $L$  par rapport à  $\mathbb{Q}$ ,  $\tilde{W}'_t = \tilde{W}_t - \sigma t$  est mouvement brownien sous  $\mathbb{Q}'$ . Ainsi,

$$\begin{aligned} C(t, S_t) &= S_t \mathbb{E}^{\mathbb{Q}}[L \mathbb{1}_{S_T > K}] - Ke^{-r(T-t)}\mathbb{Q}(S_T > K) \\ &= S_t \mathbb{Q}' \left( S_t e^{\left(r + \frac{\sigma^2}{2}\right)(T-t) + \sigma \tilde{W}'_{T-t}} > K \right) \\ &\quad - Ke^{-r(T-t)}\mathbb{Q} \left( S_t e^{\left(r - \frac{\sigma^2}{2}\right)(T-t) + \sigma \tilde{W}_{T-t}} > K \right) \end{aligned}$$

On obtient donc

$$C(t, S_t) = S_t N(d_+) - Ke^{-r(T-t)} N(d_-) \quad (3.1)$$

avec  $N$  la fonction de répartition de la loi normale centrée réduite et

$$d_{\pm} = \frac{\log \left( \frac{S_t}{K} \right) + \left( r \pm \frac{\sigma^2}{2} \right) (T-t)}{\sigma \sqrt{T-t}}.$$

En plus de la formule fermée du prix d'un call obtenue par martingale, nous avons également la possibilité de valoriser ce produit avec l'EDP de B&S.

Notons  $V_t$  un portefeuille constitué d'un call de maturité  $T$ , de strike  $K$  sur  $S_t$  et une quantité  $\Delta$  de l'actif  $S_t$  qui servira à couvrir (hedger) le portefeuille. En appliquant le lemme d'Itô à ce portefeuille

$$\begin{aligned} V_t &= C(t, S_t) - \Delta S_t \\ dV_t &= \frac{\partial C}{\partial t} dt + \frac{\partial C}{\partial S} dS_t + \frac{1}{2} \frac{\partial^2 C}{\partial S^2} d\langle S_t, S_t \rangle - \Delta dS_t \\ &= \frac{\partial C}{\partial t} dt + rS_t \frac{\partial C}{\partial S} dt + \sigma \frac{\partial C}{\partial S} dW_t + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 C}{\partial S^2} dt - \Delta dS_t. \end{aligned} \quad (3.2)$$

Le problème ici est que nous avons un terme d'aléa  $dW_t$ , pour pouvoir le supprimer nous choisissons  $\Delta = \frac{\partial C}{\partial S}$  (on est alors delta-hedgé). Le portefeuille devient

$$dV_t = \frac{\partial C}{\partial t} dt + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} dt. \quad (3.3)$$

Comme nous sommes en univers risque-neutre, le portefeuille ne devrait pas pouvoir faire plus qu'un investissement dans le taux sans risque, c'est-à-dire le portefeuille

$$dV_t = rV_t dt. \quad (3.4)$$

En égalisant (3.3) et (3.4), nous obtenons l'EDP de Black and Scholes

$$\frac{\partial C}{\partial t} + rS \frac{\partial C}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} = rC. \quad (3.5)$$

### 3.1.2 Grecques et hedging

Comme nous l'avons vu sur les équations (3.2) et (3.3), la stratégie de base consiste à ce delta-hedger. Il est possible d'aller plus loin. Notons tout de même que théoriquement nous devrions nous delta-hedger de manière constante. Rebalancer le portefeuille constamment est impossible, ne serait-ce que les coûts de transactions nous absorberaient tout profit éventuel et engendreraient des pertes colossal très rapidement. Il existe des travaux sur ce point, on parle de rebalancement optimal voir e.g. [MaPa99], [Fu11] et [GoLa14] pour le lecteur intéressé. Nous commençons par présenter les grecques (pour le cas d'un call toujours), c'est à dire différentes sensibilités du call par rapport aux variations de ses paramètres. Pour pouvoir calculer ces quantités nous utilisons la formule de B&S (3.1), supposons pour simplifier les calculs que  $t = 0$  et remarquons que

$$n(d_-) = \frac{Se^{rT}}{K} n(d_+), \quad d_+ = d_- - \sigma\sqrt{T}$$

où  $n = N'$ .

Le delta est la sensibilité du call par rapport à une petite variation du sous-

jacent :

$$\begin{aligned}
 \Delta &= \frac{\partial C}{\partial S} \\
 &= N(d_+) + S_t \frac{\partial N(d_+)}{\partial d_+} \frac{\partial d_+}{\partial S} + K e^{-r(T-t)} \frac{\partial N(d_-)}{\partial d_-} \frac{\partial d_-}{\partial S} \\
 &= N(d_+) + S_t \frac{\partial N(d_+)}{\partial d_+} \left( \frac{\partial d_+}{\partial S} - \frac{\partial d_-}{\partial S} \right) \\
 &= N(d_+) + S_t \frac{\partial N(d_+)}{\partial d_+} \left( \frac{\partial(-\sigma\sqrt{T})}{\partial S} \right) \\
 &= N(d_+).
 \end{aligned}$$

Sur la figure (5) nous présentons la forme du delta en fonction de  $T - t$  est de  $\log(S/K)$  pour  $r = 0.02$  et  $\sigma = 0.3$ .

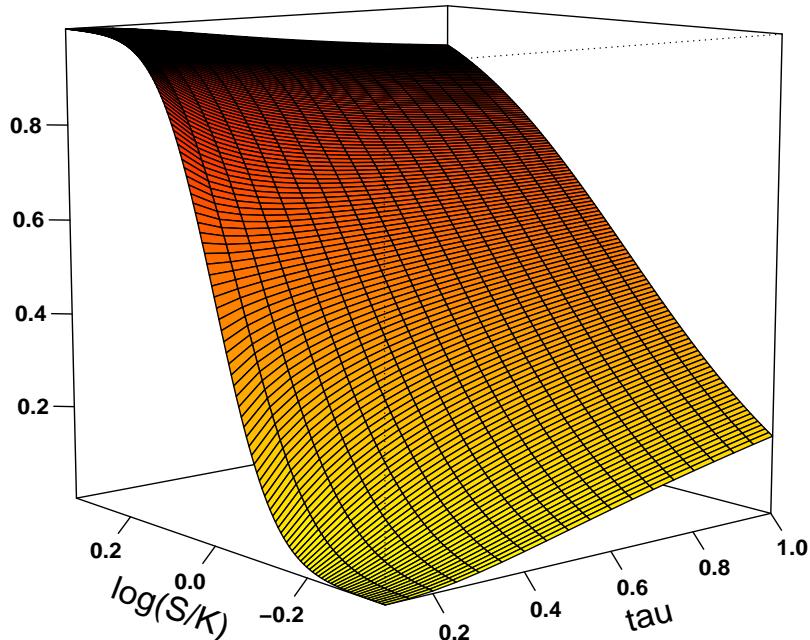
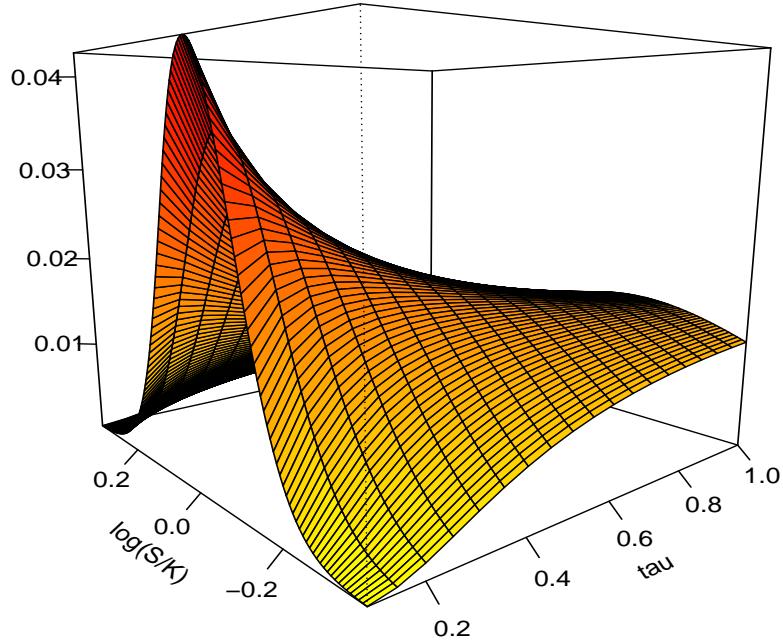


FIGURE 5 – Graphe du delta d'un call européen en fonction de  $T - t$  et  $\log(S/K)$

Nous déduisons directement que la valeur du gamma (figure (6)), c'est-à-dire


 FIGURE 6 – Graphe du gamma d'un call européen en fonction de  $T - t$  et  $\log(S/K)$ 

la dérivée de l'option par rapport à son sous-jacent est

$$\Gamma = \frac{\partial^2 C}{\partial S^2} = \frac{n(d_+)}{S\sigma\sqrt{T}}.$$

Nous pouvons interpréter la valeur du gamma comme suit. Si sa valeur est faible, cela implique que le delta varie peu et donc que nous n'avons pas besoin de rebalancer notre portefeuille pour être delta-neutre. En revanche, s'il est élevé, le delta est très sensible aux variations du sous-jacent, il conviendra de réajuster sa position régulièrement. Pour faire le lien avec la physique, la dérivée seconde correspond à l'accélération. Pour pouvoir utiliser le gamma dans notre portefeuille, nous pouvons constater qu'en effectuant un développement de Taylor sur notre portefeuille, en supposant qu'il n'est assujetti qu'aux variations en temps et en sous-jacent

$$dV = \frac{\partial V}{\partial S} \Delta S + \frac{\partial V}{\partial t} \Delta t + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} \Delta S^2 + \frac{1}{2} \frac{\partial^2 V}{\partial t^2} \Delta t^2 + \frac{1}{2} \frac{\partial^2 V}{\partial S \partial t} \Delta S \Delta t \dots \quad (3.6)$$

Si nous négligeons les termes  $(\Delta t)^x$  pour  $x > 1$ . Si le portefeuille est gamma neutre, nous avons alors

$$dV = \Theta \Delta t + \frac{1}{2} \Gamma \Delta S^2. \quad (3.7)$$

où  $\Theta = \frac{\partial V}{\partial t}$  est le theta de l'option, la variation par rapport au temps, figure (7).

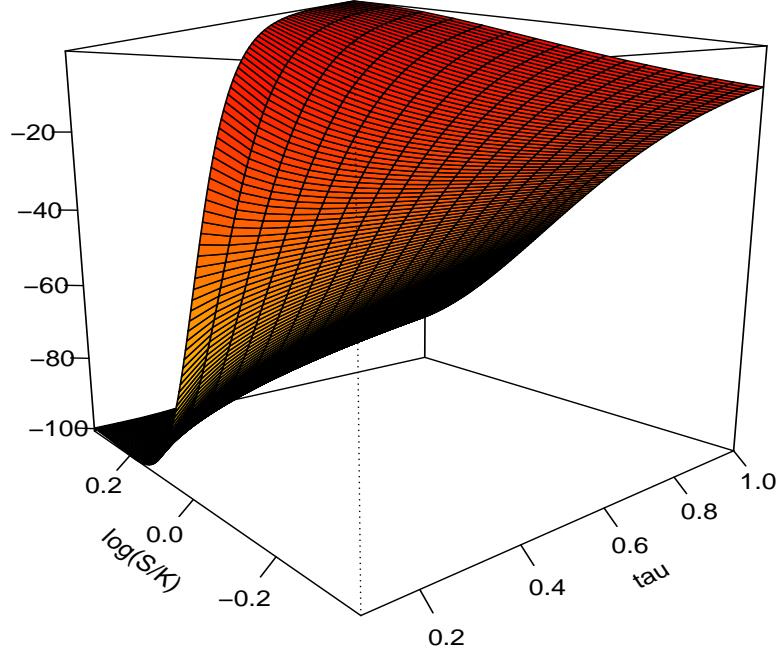


FIGURE 7 – Graphe du theta d'un call européen en fonction de  $T - t$  et  $\log(S/K)$

Par les mêmes astuces que précédemment nous avons

$$\Theta = \frac{\partial C}{\partial T} = \frac{S\sigma n(d_+)}{2\sqrt{T}} + rKe^{-rT}N(d_-).$$

Pour avoir une position gamma neutre, il faut avoir en portefeuille une quantité d'options  $\pi$  annulant le gamma du portefeuille. Si le portefeuille initial a un gamma de  $\Gamma_V$  et l'option  $\Gamma$ , le gamma total du portefeuille sera de

$$\pi\Gamma + \Gamma_V.$$

Pour être gamma neutre il faudra alors avoir une quantité  $\pi^* = -\Gamma_V/\Gamma$  d'option en portefeuille. Bien sûr, cela va modifier la quantité de sous-jacent sous gestion pour

rester delta-neutre car l'ajout d'une autre quantité va modifier le delta. D'après l'équation (3.7) nous pouvons également remarquer que le theta peut se substituer au gamma dans la gestion du portefeuille.

Si nous supposons maintenant que la volatilité est non constante, nous pouvons introduire la sensibilité de l'option par rapport à la volatilité, le vega, figure (8)

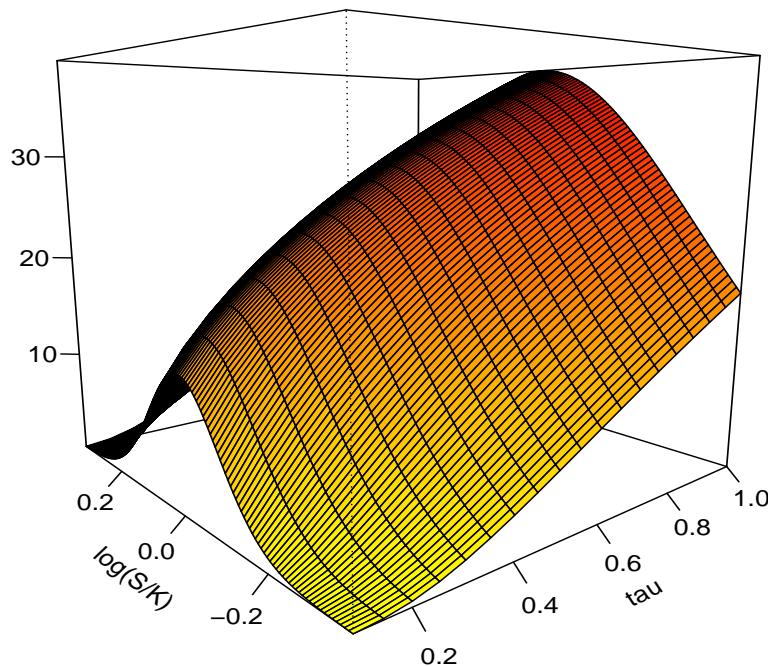


FIGURE 8 – Graphe du vega d'un call européen en fonction de  $T - t$  et  $\log(S/K)$

$$\mathcal{V} = \frac{\partial C}{\partial \sigma} = S n(d_+) \sqrt{T}. \quad (3.8)$$

Bien entendu, plus le niveau de la volatilité est important, plus la variation de l'option et donc du portefeuille sera importante. Nous devons donc nous couvrir également contre ce risque. Cela peut ce faire de la même manière que pour le gamma,  $\pi^* = -\mathcal{V}_V/\mathcal{V}$ . Néanmoins, il est très compliqué du fait de la liquidité et des fonds nécessaires de maintenir une position à la fois delta, vega et gamma hedgé. Il faut avoir plusieurs types d'options en portefeuille pour avoir ce résultat.

Notons que si nous négligeons toujours les termes  $\Delta t^x$  pour  $x > 1$ , l'équation (3.6) devient

$$dV \approx \frac{\partial V}{\partial S} \Delta S + \frac{\partial V}{\partial t} \Delta t + \frac{\partial V}{\partial \sigma} \Delta \sigma + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} \Delta S^2 + \frac{1}{2} \frac{\partial^2 V}{\partial \sigma^2} \Delta \sigma^2.$$

La dernière sensibilité courante est le  $\rho$ , la sensibilité de l'option par rapport au taux d'intérêt.

$$\rho = \frac{\partial C}{\partial r} = T K e^{-rT} N(d_+).$$

Dans un monde pas trop chaotique, les taux d'intérêts varient peu, la priorité n'est donc pas sur cette couverture.

Nous allons maintenant présenter ce que l'on appelle la robustesse de Black & Scholes [KaJeSh98]. La vraie diffusion du sous-jacent sous  $\mathbb{Q}$  est

$$\frac{dS_t}{S_t} = r dt + \sigma_t dW_t$$

où  $\sigma_t$  peut être stochastique. En supposant que l'investisseur utilise la formule de Black & Scholes pour valoriser un call, nous avons en appliquant la formule d'Itô

$$\begin{aligned} dC_{BS}(t, S_t) &= \left( \frac{\partial C_{BS}}{\partial t} + r S_t \frac{\partial C_{BS}}{\partial S} + \frac{1}{2} \sigma_t^2 S_t^2 \frac{\partial^2 C_{BS}}{\partial S^2} \right) dt \\ &\quad + \sigma_t S_t \frac{\partial C_{BS}}{\partial S} dW_t. \end{aligned}$$

D'un autre côté rappelons que l'EDP de Black & Scholes (3.5), est

$$\frac{\partial C_{BS}}{\partial t} + r S \frac{\partial C_{BS}}{\partial S} + \frac{1}{2} \sigma_{BS}^2 S^2 \frac{\partial^2 C_{BS}}{\partial S^2} = r C_{BS}.$$

En combinant les deux équations nous avons

$$\begin{aligned} dC_{CB}(t, S_t) &= r C_{BS} dt + \sigma_t S_t \frac{\partial C_{BS}}{\partial S} + \frac{1}{2} (\sigma_{BS}^2 - \sigma_t^2) \frac{\partial^2 C_{BS}}{\partial S^2} \\ &= r C_{BS} dt + \sigma_t S_t \Delta + \frac{1}{2} (\sigma_{BS}^2 - \sigma_t^2) S_t^2 \Gamma. \end{aligned}$$

La dynamique du portefeuille  $V_t$  de couverture  $\Delta$  de l'agent est

$$dV_t = rV_t + \Delta(dS_t - rS_t dt), \quad V_0 = C_{BS}(0, S_0).$$

L'erreur de couverture de l'agent est donc

$$\begin{aligned}\varepsilon_T &= V_T - C_{BS}(T, S_T) \\ &= \frac{1}{2} \int_0^T e^{r(T-t)} (\sigma_{BS}^2 - \sigma_t^2) S_t^2 \Gamma dt.\end{aligned}$$

Le gamma de l'option mesure donc l'exposition de l'agent à une mauvaise estimation de la volatilité. Si la volatilité choisie est supérieure à la vraie,  $\sigma_{BS} \geq \sigma_t$  alors le portefeuille de couverture Black & Scholes domine le vrai prix. Inversement si  $\sigma_{BS} \leq \sigma_t$ . Le modèle de Black & Scholes permet donc d'obtenir des bornes du vrai prix de l'option.

## 3.2 Surface de volatilité

### 3.2.1 Définition et problème de calibration

Dans le modèle de B&S, la volatilité  $\sigma$  est supposée constante or, si nous cherchons la valeur de  $\sigma$  telle que le prix de B&S soit le même que le prix du marché pour différentes maturités et différents strike, nous constatons que la valeur retournée n'est pas constante. Notons  $C^{Mkt}(t, S_t, T, K)$  le prix du call observé sur le marché, alors la volatilité implicite,  $I(t, S_t, T, K)$  est donnée par :

$$C^{BS}(t, S_t, T, K, I(t, S_t, T, K)) = C^{Mkt}(t, S_t, T, K).$$

De plus, elle est unique. Elle est donc obtenue en inversant la valeur du call. Il n'est pas possible d'obtenir une formule explicite pour la déterminer, on va donc utiliser un algorithme d'optimisation. Posons par exemple le problème comme un problème de recherche de zéro d'une fonction :

$$f(I) = C^{BS} - C^{Mkt} = 0.$$

L'un des algorithmes de recherche le plus simple est la méthode de Newton, l'idée est de faire le développement limite d'une fonction  $g$  de  $x$  en  $x_k$  :

$$g(x) = g(x_k) + (x - x_k)g'(x_k)$$

en égalisant en zéro, on trouve alors

$$x = x_{k+1} = x_k - \frac{g(x_k)}{g'(x_k)}.$$

Pour notre problème :

$$I_{k+1} = I_k - \frac{C^{Mkt} - C^{BS}(I_k)}{\frac{\partial C^{BS}(I_k)}{\partial I_k}}. \quad (3.9)$$

Le dénominateur est le vega de l'option (3.8).

Il est important de noter quelques points sur la volatilité implicite. Notons déjà qu'elle est toujours supérieure à la volatilité du sous-jacent et qu'elle diffère selon les maturités et strikes sur un même sous-jacent.

Les 'mots clés' à savoir lorsque l'on parle de volatilité implicite sont le skew et le smile. Le phénomène de skew correspond à la décroissance de la volatilité implicite par rapport au strike. Le phénomène de smile correspond quant à lui à la remontée de la volatilité implicite pour des strikes très lointains.

Pourquoi la volatilité historique, définie comme les accroissements au carré du sous-jacent, est-elle différente de la volatilité implicite ? Sur le marché, la couverture d'une option est plus chère que dans le modèle de Black & Scholes, frais de transaction, bid-ask spread, market impact et, bien sûr la volatilité supposée constante ainsi qu'une loi gaussienne sous-jacente donc, peu de 'gros' mouvements. Le phénomène de skew, i.e. la première partie de la volatilité par rapport au strike, s'explique par le fait que les agents surestiment le prix par rapport au modèle de Black & Scholes pour se prémunir d'un mouvement contraire lorsque l'on est dans la monnaie (ITM). Le smile s'explique quant à lui pour compenser la faible liquidité des options loin de la monnaie à strike très lointain (OTM). Nous présentons sur la figure (9) la volatilité implicite du call sur l'indice SPY le 9 septembre 2013.

### 3.2.2 Volatilité implicite forward

Nous avons vu comment à partir des prix des call et put européens nous déduisions la volatilité implicite, mais cela ne nous indiquent pas comment est-ce qu'elle évolue. La valorisation d'option européenne ne requiert pas de connaître la volatilité implicite future, en revanche, pour des options path-dependant, i.e. que le payoff dépend de toute la trajectoire du sous-jacent et non uniquement de sa valeur terminale, comme les options barrière ou forward-start. On cherche donc à savoir à quoi correspondra la volatilité implicite en fonction du niveau du prix

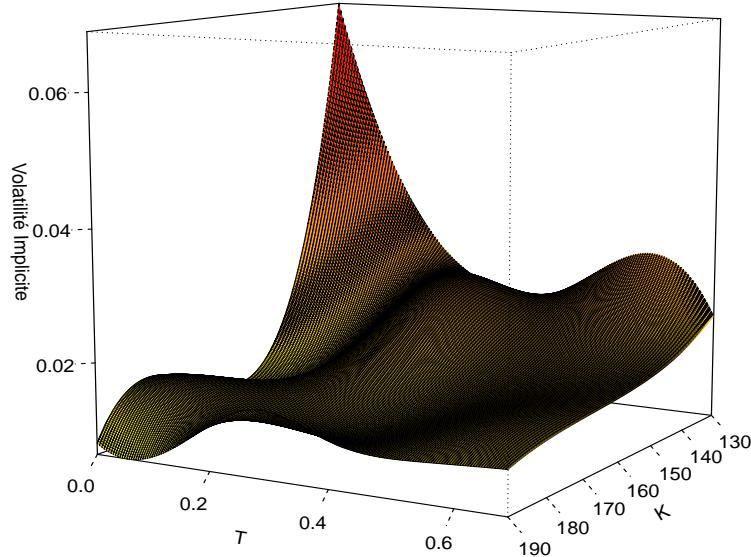


FIGURE 9 – Surface de volatilité du SPY, le 9 septembre 2013,  $\sim 166\$$ , interpolation par spline.

futur. Tous les paramètres du modèle n'influent pas le niveau de la volatilité à court-terme, cela pourrait être plus simple de considérer ce problème comme un problème invariant. En finance on parle plutôt de problème 'sticky' (épineux) et comme dit précédemment, ils portent sur le modèle et non sur le sous-jacent lui-même. Les deux sticky rules les plus connues sont :

- Sticky strike rule.

Pour un certain niveau de skew sur la volatilité implicite, il est possible que même si le sous-jacent bouge, les options cotées sur celui ci avec des niveaux de strike inchangés gardent leur volatilité initiale. C'est ce que l'on appelle la règle sticky strike. Bien entendu, en cas de grosse variation de  $S_t$ , par exemple durant des périodes de crise, la volatilité implicite sera modifiée.

On a

$$I(t + \delta t, K) = I(t, K)$$

- Sticky delta/moneyness rule.

La règle de moneyness correspond à la dépendance entre la volatilité et la moneyness de l'option  $m = K/S$ ,

$$I(t + \delta t, m) = I(t, m)$$

Pour pouvoir déterminer concrètement la volatilité implicite future, nous avons besoin de travailler sur des options à paiement différé, l'exemple le plus courant est celui d'une option forward-start. Un option forward-start est une option contractée à l'instant  $t$ , recevant une proportion du sous-jacent à l'instant  $T_1 \geq t$  à la maturité  $T_2 > T_1$ . Le payoff d'une telle option est

$$\phi(S_{T_2}) = (S_{T_2} - KS_{T_1})_+.$$

Le pricing est donc

$$C(t, T_2) = e^{-r(T_1-t)} \mathbb{E}[(S_{T_2} - KS_{T_1}) | \mathcal{F}_t].$$

Supposons que le sous-jacent soit diffusé par un mouvement brownien géométrique de volatilité constante sous la probabilité risque-neutre, alors

$$\begin{aligned} C(t, T_2) &= e^{-r(T_1-t)} \mathbb{E}[\mathbb{E}[(S_{T_2} - KS_{T_1}) | \mathcal{F}_{T_1}] | \mathcal{F}_t] \\ &= e^{-r(T_1-t)} \mathbb{E}[S_{T_1} N(d_+) - Ke^{-r(T_2-T_1)} N(d_-) | \mathcal{F}_t] \\ &= e^{-r(T_1-t)} N(d_+) - Ke^{-r(T_2-t)} N(d_-), \end{aligned}$$

avec

$$d_{\pm} = -\frac{\ln(Ke^{-r(T_2-T_1)})}{\sigma_{BS}\sqrt{T_2-T_1}} \pm \frac{\sigma_{BS}\sqrt{T_2-T_1}}{2},$$

où nous avons utilisé le fait que  $\mathbb{E}[\mathbb{E}[X | \mathcal{F} | \mathcal{G}] = \mathbb{E}[X | \mathcal{G}]$  pour toute v.a.r.  $X$  et les filtrations telles que  $\mathcal{G} \subset \mathcal{F}$  et que le résultat de Black & Scholes est déterministe.

La dynamique de la volatilité implicite est déduite de tel produit. En fait, on va plutôt utiliser les options cliques. Il s'agit d'une somme d'option forward-start, le payoff est alors simplement

$$\phi(S_n) = \sum_{i=1}^n (S_{T_{i+1}} - KS_{T_i})_+$$

où le tenor  $T_1, \dots, T_n$  est fixé à l'avance. La valorisation se déduit de la formule donnée pour les options forward-start.

### 3.3 Modèle à volatilité locale

Pour l'aspect historique, nous pouvons citer Breeden and Litzenberger, [BrLi78], faisant apparaître la relation entre la densité du stock et le prix des options. Ensuite, Dupire [Du94] et Derman et Kani [DeKa94] ont montré que sous la probabilité risque neutre, il existe une unique diffusion consistante avec cette distribution. Dans [DeKa94], la volatilité locale est construite de manière discrète avec des méthodes d'arbres. Ici, nous nous intéresserons seulement au temps continu avec la formule de Dupire. Avant de donner la formulation pour les caplets, nous donnons les équations pour le prix de call sur stock. L'unique coefficient correspondant  $\sigma(T, K)$  consistant avec le prix des options européennes est la fonction de volatilité implicite. Les modèles à volatilité locale proposent une méthodologie simplifiée pour valoriser des produits exotiques à partir de produits vanilles.

#### 3.3.1 Diffusion de Dupire

La diffusion de driftless du stock  $S_t$  est donnée pour tout  $t \geq 0$  par

$$dS_t = \sigma(t, S_t) S_t dW_t,$$

avec  $\sigma$  une fonction déterministe positive. Le prix d'un call de maturité  $T$  et de strike  $K$  en  $t = 0$  est

$$C(T, K) = \int_K^\infty (s - K) \varphi_T(s) ds \quad (3.10)$$

avec  $\varphi_T$  la densité risque neutre de l'actif. Avant de continuer, rappelons l'équation de Fokker Planck (ou Kolmogorov forward).

**Proposition 3.1.** (*Equation de Fokker Planck*) Soit  $(X_t)_{t \in \mathbb{R}_+}$  un processus de diffusion

$$dX_t = \mu(t, X_t) dt + \sigma(t, X_t) dW_t$$

avec  $(W_t)_{t \in \mathbb{R}_+}$  un mouvement Brownien standard. Alors

$$\frac{d}{dt} g(t, x) = -\frac{\partial}{\partial x} (\mu(t, x) g(t, x)) + \frac{1}{2} \frac{\partial^2}{\partial x^2} (\sigma^2(t, x) g(t, x))$$

où  $g$  est la densité de probabilité de  $X$ ,  $g(t, x) = \mathbb{P}(X_t \in dx)$  et pour condition au bord,

$$g(t, x) dx|_{t=0} = \delta(x - X_0).$$

La densité  $\varphi$  de (3.10) doit vérifier cette équation, ainsi

$$\frac{1}{2} \frac{\partial^2(\sigma^2 s^2 \varphi_T)}{\partial s^2} = \frac{\partial \varphi_T}{\partial t}. \quad (3.11)$$

En dérivant deux fois l'équation (3.10) par rapport au strike  $K$  nous obtenons

$$\varphi_T(K) = \frac{\partial^2 C(T, K)}{\partial K^2}.$$

En injectant cela dans l'équation (3.11),

$$\frac{1}{2} \frac{\partial^2(\sigma^2 K^2 \varphi_T)}{\partial K^2} = \frac{\partial}{\partial T} \left( \frac{\partial^2 C}{\partial K^2} \right).$$

En intégrant deux fois par rapport à  $K$

$$\frac{1}{2} \sigma^2 K^2 \varphi_T = \frac{\partial C}{\partial T}.$$

En remplaçant à nouveau  $\varphi_T$  par la dérivée seconde du call par rapport à  $K$

$$\frac{1}{2} \sigma^2 K^2 \frac{\partial^2 C}{\partial K^2} = \frac{\partial C}{\partial T}.$$

Nous obtenons la volatilité locale définie par

$$\sigma(T, K) = \sqrt{\frac{2 \frac{\partial C}{\partial T}}{K^2 \frac{\partial^2 C}{\partial K^2}}}. \quad (3.12)$$

Quel est l'intérêt de l'EDP de Dupire par rapport à celle de Black & Scholes ? Rappelons que cette dernière est donnée pour un call  $C$  sur  $S$  de même diffusion que précédemment, de strike  $K$  et de maturité  $T$  par

$$\frac{\partial C}{\partial t} + rS \frac{\partial C}{\partial S} + \frac{1}{2} \sigma(t, S)^2 S^2 \frac{\partial^2 C}{\partial S^2} = rC.$$

La volatilité de Black & Scholes est donc

$$\sigma(t, S) = \sqrt{\frac{rC - \frac{\partial C}{\partial t} - rS \frac{\partial C}{\partial S}}{S^2 \frac{\partial^2 C}{\partial S^2}}}. \quad (3.13)$$

Les dérivées sont par rapport à  $t$  et  $S$ , or, ces deux quantités sont fixées, nous ne pouvons pas calculer ou approximer ces dérivées. L'EDP de Dupire 3.12 tient seulement compte des dérivées par rapport à la maturité et au strike. Sur le marché, chaque call/put suffisamment liquide est vendu à des strikes et maturités différentes, il est donc possible en collectant ces données de calculer les dérivées et donc, d'obtenir la volatilité implicite.

Nous allons maintenant regarder le lien entre volatilité locale et volatilité instantanée.

### 3.3.2 Volatilité implicite dans le modèle de Dupire

Le prix d'un call (non actualisé) est donnée par

$$C = \mathbb{E}[(S_T - K)_+],$$

en dérivant par rapport au strike  $K$  nous obtenons

$$\frac{\partial C}{\partial K} = -\mathbb{E}[\mathbf{1}_{S_T > K}].$$

En dérivant une nouvelle fois par rapport à  $K$

$$\frac{\partial^2 C}{\partial K^2} = \mathbb{E}[\delta(S_T - K)]. \quad (3.14)$$

avec  $\delta$  la fonction de Dirac.

Appliquons maintenant le lemme d'Itô à la fonction  $(S_T - K)_+$  avec  $dS_t = \sqrt{v_t} S_t dW_t$

$$d(S_T - K)_+ = \mathbf{1}_{S_T > K} dS_T + \frac{1}{2} v_T S_T^2 \delta(S_T - K) dT.$$

En passant l'opérateur espérance, nous avons

$$d\mathbb{E}[(S_T - K)_+] = d\mathbb{E}\left[\frac{1}{2} v_T S_T^2 \delta(S_T - K)\right] dT$$

avec en utilisant (3.14)

$$\begin{aligned} \mathbb{E}\left[\frac{1}{2} v_T S_T^2 \delta(S_T - K)\right] &= \mathbb{E}[v_t | S_T = K] \frac{1}{2} K^2 \mathbb{E}[\delta(S_T - K)] \\ &= \mathbb{E}[v_t | S_T = K] \frac{1}{2} K^2 \frac{\partial^2 C}{\partial K^2}. \end{aligned} \quad (3.15)$$

En mixant la formulation de la volatilité locale (3.13) avec le résultat précédent (3.15) nous obtenons la représentation suivante :

$$\sigma^2(T, K) = \mathbb{E}[v_t | S_T = K]. \quad (3.16)$$

Dans ce cas, la volatilité locale est donc l'espérance de la volatilité instantanée conditionnellement au fait que le prix final est égal au strike.

Considérons maintenant le cas où la diffusion de l'actif  $S$  dans l'univers risque neutre est

$$\frac{\partial S_t}{S_t} = rdt + \sigma(t, S_t)dW_t,$$

où  $r$  est le taux sans risque. En appliquant à nouveau l'équation de Fokker-Planck pour cette diffusion nous avons

$$\frac{\partial \varphi_T}{\partial T} = -\frac{\partial r S \varphi}{\partial S} + \frac{\partial^2 \sigma^2 S^2 \varphi_T}{\partial S^2}.$$

Avec le même raisonnement que précédemment (3.11), nous obtenons en évaluant en  $S_T = K$

$$\sigma^2(T, K) = 2 \frac{\frac{\partial C}{\partial T} + rK \frac{\partial C}{\partial K}}{K^2 \frac{\partial^2 C}{\partial K^2}}.$$

Il est possible de lier la volatilité locale avec la volatilité implicite. Les prix  $C$  sont les 'vrais' prix du marché de volatilité locale  $\sigma$ , ils doivent donc correspondre au prix des calls marché calculés avec la volatilité implicite, soit :  $C(T, K) = C^{BS}(T, K, I(T, K))$  où  $I$  est la volatilité implicite. En appliquant la formule de dérivation en chaîne, nous obtenons

$$\sigma^2(T, K) = 2 \frac{\frac{\partial C^{BS}}{\partial T} + \frac{\partial C^{BS}}{\partial \sigma} \frac{\partial I}{\partial T} + rK \left( \frac{\partial C^{BS}}{\partial K} + \frac{\partial C^{BS}}{\partial \sigma} \frac{\partial I}{\partial K} \right)}{K^2 \left( \frac{\partial^2 C^{BS}}{\partial K^2} + 2 \frac{\partial^2 C^{BS}}{\partial K \partial \sigma} \frac{\partial I}{\partial K} + \frac{\partial^2 C^{BS}}{\partial \sigma^2} \left( \frac{\partial I}{\partial K} \right)^2 + \frac{\partial C^{BS}}{\partial \sigma} \frac{\partial^2 I}{\partial K^2} \right)}. \quad (3.17)$$

Il est possible d'aller un peu plus loin en supposant que la valeur du strike  $K$  n'influe pas la volatilité implicite, en d'autre terme nous supposons que

$$\frac{\partial I}{\partial K} = 0.$$

En injectant les dérivées du call de Black and Scholes calculées dans la section précédente avec en plus

$$\frac{\partial C}{\partial K} = -e^{-rT} N(d_-), \quad \frac{\partial^2 C}{\partial K^2} = \frac{e^{-rT} n(d_-)}{K \sigma \sqrt{T}}, \quad (3.18)$$

l'expression de la volatilité locale précédente se réduit alors à

$$\begin{aligned} \sigma^2(T) &= 2 \frac{\frac{SIn(d_+)}{2\sqrt{T}} + rKe^{-rT}N(d-) + Sn(d_+) \sqrt{T} \frac{\partial I}{\partial T} + -rKe^{-rT}N(d_-)}{K^2 \frac{e^{-rT} n(d_-)}{KI\sqrt{T}}} \\ &= I(I)^2 + 2TI(T) \frac{\partial I}{\partial T}. \end{aligned} \quad (3.19)$$

La volatilité implicite dans le cas où elle n'est pas influencée par le strike est alors égale à la moyenne quadratique de la volatilité locale sur toute la vie de l'option

$$I^2(T) = \frac{1}{T} \int_0^T \sigma^2(s) ds. \quad (3.20)$$

Nous avons maintenant quelques formules, mais tout cela reste assez théorique et pas forcément trivial de les mettre en oeuvre. En ce qui concerne le calcul de la volatilité implicite, nous pouvons appliquer la formule (3.9), ce n'est pas forcément optimal, mais c'est un bon début. Pour pouvoir estimer la volatilité locale, c'est une autre histoire.

### 3.3.3 Problème de calibration

N'ayant qu'un ensemble fini d'observations  $(T_i, K_j)_{0 \leq i < n, 0 \leq j < m}$  il est évident que même si une solution au problème existe, elle ne sera pas unique puisqu'un ensemble de courbes vont pouvoir passer par les prix du marché. De plus, empiriquement, la moindre petite variation des prix va complètement modifier la volatilité implicite trouvée, la solution n'est absolument pas stable. Nous avons deux approches pour pouvoir 'estimer' notre volatilité locale, soit par régularisation, soit en supposant une forme paramétrique de la volatilité implicite. Nous commençons par la première.

On parle de calibration lorsque l'on cherche numériquement les valeurs des paramètres d'un modèle en fonction des prix actuels et non passés. Le terme d'estimation est réservé à la recherche des paramètres en fonction de l'historique des

données, par exemple par maximum de vraisemblance. On parlera donc de processus forward-looking pour la calibration et de processus backward-looking pour l'estimation.

Le problème de calibration est donc un problème inverse. En effet, dans le cadre le plus simple du modèle de Black & Scholes nous utilisons la formule de pricing pour déterminer le paramètre de volatilité en fonction des prix observés plutôt que de déterminer le prix des options en fonction de la volatilité. Dans l'univers risque neutre, le drift disparaît au profit du taux sans risque, disponible sur le marché, le seul paramètre restant à trouver est bien la volatilité. Néanmoins, pour un modèle plus sophistiqué comme pour notre volatilité locale au sens de Dupire, cela est bien plus compliqué. Les problèmes inverses sont souvent plus difficiles à résoudre car ils sont bien souvent 'mal posés' (ill-posed).

Rappelons qu'au sens de Hadamard, un problème est bien posé (well-posed) si ces conditions sont remplies :

- Le problème admet une solution
- Le problème admet une solution et, elle est unique
- La solution dépend continûment des données.

Pour obtenir un problème bien posé, nous avons besoin de le régulariser. L'une des méthodes les plus connues est la régularisation de Tikhonov.

Notons  $C^{Mkt}(K, T)$  les prix des options observés sur le marché,  $C(\theta, T, K)$  les prix du modèle à volatilité locale et  $\theta \in \Theta \subset \mathcal{H}$  le vecteur de paramètres à estimer dans le sous-espace convexe  $\Theta$ ,  $\mathcal{H}$  étant un espace de Hilbert. Le problème de calibration s'écrit comme un problème de moindre carré non-linéaire

$$\tilde{\theta} = \arg \min_{\theta \in \Theta} \|C^{Mkt}(K, T) - C(\theta, T, K)\|^2,$$

cela nous permet de s'assurer de l'existence d'une solution,  $\|\cdot\|$  étant une norme sur  $\mathcal{H}$ . Supposons maintenant que nous ayons un à priori sur  $\theta$  disons  $\theta_0$ , nous introduisons une pénalisation (de Tikhonov) sur la fonctionnelle pour régulariser le problème

$$J_\alpha(\theta) = \|C^{Mkt}(K, T) - C(\theta, T, K)\|^2 + \alpha \|\sigma\|_{\mathcal{H}^1}^2, \quad (3.21)$$

$\alpha > 0$  va influencer le niveau de régularité pour assurer la stabilité. Il s'agit donc de faire un compromis entre précision et stabilité. Plus  $\alpha$  sera choisi grand, plus la solution sera stable mais loin de la réalité. A l'inverse, plus nous choisirons un

$\alpha$  proche de zéro, plus les paramètres trouvés (dont la volatilité locale) reproduiront rigoureusement le marché mais ils seront très instables rendant le modèle inutilisable. Pour le terme de pénalisation, on peut montrer que

$$\|\sigma\|_{\mathcal{H}^1}^2 = \int_{t_0}^T \int_0^\infty (\sigma^2(t, S) + \left(\frac{\partial \sigma}{\partial t}(t, S)\right)^2 + \left(\frac{\partial \sigma}{\partial S}(t, S)\right)^2) dt dS_t.$$

Le problème (3.21) peut être résolu numériquement par un algorithme de descente (voir appendix). Remarquons un point faible de cette procédure, la régularisation,  $\alpha \|\sigma\|_{\mathcal{H}^\infty}^2$ , est la même quelque soit la région dans laquelle on se trouve, que la volatilité influe sur les prix ou non. Il est préférable de ne pas chercher à minimiser l'erreur quadratique mais plutôt l'erreur quadratique pondérée par l'inverse du bid-ask spread. Cette pondération s'interprète comme un contrôle du résultat pour des produits peu liquides avec un spread élevé. Pour plus de détails sur ces problèmes de calibrations nous renvoyons à [Cr03] et [LaOs97].

L'autre approche que nous présentons consiste à calculer la volatilité locale à partir de la volatilité implicite avec l'équation reliant ces deux quantités (3.17). Plutôt que de prendre la 'vraie' volatilité implicite, J. Gatheral [Ga04] propose une forme paramétrique de la volatilité implicite, comme cela nous n'avons pas de problème de grille de discréétisation et de bruit. Il s'agit du modèle SVI (Stochastic Volatility Inspired). Inspiré des modèles à volatilité stochastique car il vérifie certaines des propriétés de volatilité implicite asymptotique vérifiées par les modèles à volatilité stochastiques comme celui de Heston [Lee04]. La forme paramétrique de la volatilité implicite est

$$I^2(T, k) = \frac{1}{T} \left( a + b \left( \rho(k - m) + \sqrt{(k - m)^2 + \sigma^2} \right) \right) \quad (3.22)$$

avec  $k$  le log-strike,  $k = \log \frac{K}{S} - rT$ ,  $a \geq 0, b \geq 0, \sigma \geq 0, \rho \in [-1, 1]$  et  $m \in \mathbb{R}$ . Les différents paramètres ont les influences suivantes :

$a$  détermine le niveau de la volatilité implicite

$b$  détermine les pentes à droite et gauche de la volatilité implicite

$\rho$  détermine l'asymétrie de la volatilité implicite

$m$  détermine le moment auquel on observe la volatilité implicite remonte (smile après skew)

$\sigma$  détermine la convexité à la monnaie.

Après avoir estimé les paramètres par rapport aux prix marchés  $C^{Mkt}(T_i, K_j)_{i,j}$  nous obtenons une surface de volatilité implicite lisse, non bruitée. C'est celle ci que nous allons injecter dans l'équation de la volatilité locale (3.17) qui d'ailleurs devient plus précisément que (3.19), après quelques calculs fastidieux, dans le cas où la volatilité implicite est influencée par le strike

$$\sigma^2(T, K) = \frac{\frac{I}{T} + 2\frac{\partial I}{\partial T} + 2rK\frac{\partial I}{\partial K}}{Kp2 \left( \frac{1}{K^2 IT} + 2\frac{d_+}{KI\sqrt{T}}\frac{\partial I}{\partial K} + \frac{d_+ d_-}{I} \left( \frac{\partial I}{\partial K} \right)^2 + \frac{\partial^2 I}{\partial K^2} \right)}.$$

Connaissant une version plus 'propre' de la volatilité implicite avec le modèle SVI (3.22), à tous les points  $(T_j, K_j)$  avec  $j$  notre indice de grille de discréétisation, nous pouvons alors estimer la valeur de la volatilité locale. Pour plus de détails sur la calibration nous renvoyons par exemple à [Ga04] et [GaJa12] pour un modèle plus élaboré.

Des méthodes pour estimer la volatilité locale par splines ont également été proposées. Nous ne rentrerons pas dans ces détails pour ce cours et nous renvoyons e.g. à [JaSuHo99] pour le lecteur intéressé.

### 3.3.4 Modèle shifted-lognormal

Un exemple de modèle que nous allons étudier est le modèle shifted-lognormal. La diffusion du taux forward  $F^j$  associée à la mesure  $T_j$ -forward est

$$dF_t^j = \beta(t)(F_t^j - \alpha)dW_t, \quad t < T_j,$$

où  $\alpha$  est une constante réelle,  $(\beta(t))_{t \in \mathbb{R}_+}$  une fonction déterministe par rapport au temps et  $(W_t)_{t \in \mathbb{R}_+}$  un mouvement Brownien standard. Pour tout  $t < T \leq T_j$  nous avons

$$F_T^j = \alpha + (F_t^j - \alpha)e^{-\frac{1}{2} \int_t^T \beta^2(u)du + \int_t^T \beta(u)dW_u}.$$

Conditionnellement à  $F_t^j$ , la densité de  $F_T^j$  est pour tout  $t < T \leq T_j$

$$f_{F_T^j|F_t^j}(x) = \frac{1}{(x - \alpha)U(t, T)\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left( \frac{\ln \frac{x-\alpha}{F_t^j-\alpha} + \frac{1}{2}U^2(t, T)}{U(t, T)} \right)^2 \right\},$$

pour  $x > \alpha$  et

$$U(t, T) = \sqrt{\int_t^T \beta^2(u)du}.$$

Remarquons que nous pouvons réécrire le modèle

$$\begin{aligned} F_t^j &= X_t^j + \alpha \\ dX_t^j &= \beta(t)X_t^j dW_t, \end{aligned}$$

et donc

$$\begin{aligned} B(t, T_j)\mathbb{E}_j((F_{T_{j-1}}^j - \kappa)_+ | \mathcal{F}_t) &= B(t, T_j)\mathbb{E}_j((X_{T_{j-1}}^j + \alpha - \kappa)_+ | \mathcal{F}_t) \\ &= B(t, T_j)\mathbb{E}_j((X_{T_{j-1}}^j - (\kappa - \alpha))_+ | \mathcal{F}_t). \end{aligned}$$

$X$  étant un mouvement Brownien géométrique, nous avons simplement pour la valeur d'un caplet

$$\text{Cpl}(t, T_{j-1}, T_j, \delta, \kappa) = \delta_j B(t, T_j) \text{Bl}(\kappa - \alpha, F_t^j - \alpha, U(t, T_{j-1})).$$

La volatilité implicite,  $\hat{\sigma}(K, \alpha)$ , est alors obtenue en résolvant numériquement

$$\delta_j B(t, T_j) \text{Bl}(\kappa, F_t^j, \hat{\sigma}(K, \alpha) \sqrt{T_{j-1} - t}) = \delta_j B(t, T_j) \text{Bl}(\kappa - \alpha, F_t^j - \alpha, U(t, T_{j-1})).$$

### 3.3.5 Modèle CEV

Le modèle CEV pour Constant Elasticity of Variance a été introduit par [Co75] et appliqué au marché des taux d'intérêts par [AnAn00]. La dynamique du taux forward est sous la mesure  $T_j$ -forward

$$dF_t^j = \sigma_j(t)(F_t^j)^\gamma dW_t, \quad 0 < \gamma \leq 1. \quad (3.23)$$

Le cas  $\gamma = 1$  correspond bien entendu au modèle de Black & Scholes, i.e. lognormal, et  $\alpha \rightarrow 0$  au cas gaussien (Bachelier). 0 est la barrière absorbante, si  $F_t = 0$  alors  $F_s = 0$  pour tout  $s \geq t$ . Il est possible de montrer que pour tout  $0 < \gamma \leq 1$  le

processus est une vraie martingale (locale si  $\gamma > 1$ ). La valorisation d'un caplet de maturité  $T_j$  à l'instant  $t$  de strike  $\kappa$  est

$$\text{Cpl}(t, T_j, \delta) = \delta B(t, T_j) \left\{ F_t^j \left[ 1 - \chi^2 \left( 2\kappa^{1-\gamma}; \frac{1}{1-\gamma} + 2; 2u \right) \right] - \kappa \chi^2 \left( 2u; \frac{1}{1-\gamma}; 2k\kappa^{1-\gamma} \right) \right\},$$

où  $\chi^2(\cdot, r, \rho)$  est la fonction de répartition d'un chi-deux décentré de  $r$  degrés de liberté et de paramètre de décentrage  $\rho$  et

$$k = \frac{1}{2v(t, T)(1-\gamma)^2}, \quad u = k(F_t^j)^{2(1-\gamma)}.$$

**Preuve 3.1.** *On peut montrer qu'il s'agit d'un processus de Bessel, voir [AnAn00].*

Les auteurs de ([HaWe08]) ont montré par perturbation singulière que la volatilité implicite d'un caplet de maturité  $T_j$  est donnée par

$$\hat{\sigma}(K, \alpha) \approx \frac{\sigma_j}{u^{1-\gamma}} \left\{ 1 + \frac{1}{24}(1-\gamma)(2-\gamma) \left( \frac{F_0 - \kappa}{u} \right)^2 + \frac{1}{24}(1-\gamma)^2 \frac{\sigma_j^2 T_j}{u^{2-2\gamma}} \right\},$$

avec

$$u = \frac{1}{2}(F_0 + \kappa).$$

Nous présentons sur la figure (10) un exemple de surface de volatilité avec le modèle CEV.

### 3.4 Volatilité stochastique

Avant de discuter des modèles à volatilité stochastique pour les produits de taux d'intérêts nous posons la forme générale des équations régissant le prix du sous-jacent et de sa volatilité

$$\begin{aligned} dS_t &= \mu(t)S_t dt + \sqrt{v_t}S_t dZ_t^1 \\ dv_t &= \alpha(S, v, t)dt + \eta\beta(S, v, t)dZ_t^2 \end{aligned}$$

avec  $\langle dZ_t^1 dZ_t^2 \rangle = \rho dt$ .  $\mu(t)$  est une fonction déterministe. Nous allons essayer de retrouver l'EDP de valorisation pour les modèles à volatilité stochastique. Dans le

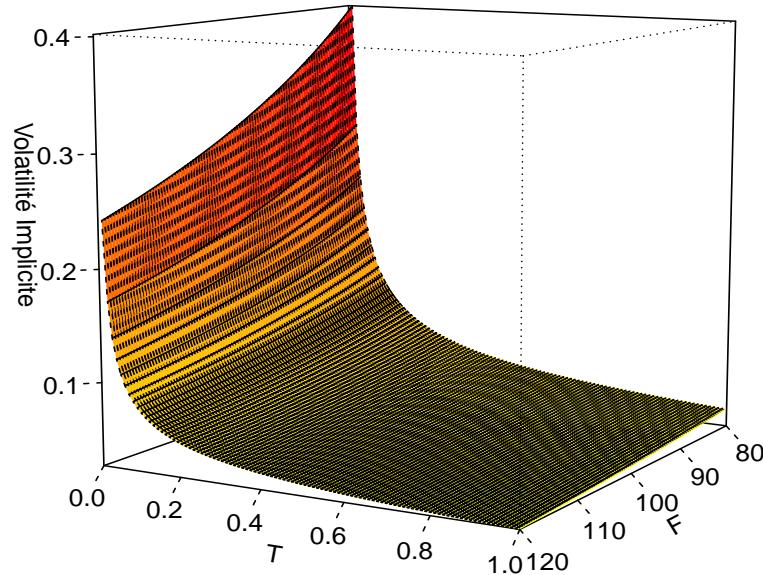


FIGURE 10 – Surface de volatilité simulée à l'aide du modèle CEV avec  $F = 100$ .

cadre du modèle de Black & Scholes, il n'y a qu'une seule source d'aléa, de risque, le prix du sous-jacent et, cela peut être couvert en achetant  $\Delta$  sous-jacent (pour un call). Ici, nous avons deux sources d'aléa, le prix du sous-jacent et sa volatilité, il n'est plus possible de se couvrir à l'aide d'un seul actif. La seconde source de risque a elle aussi besoin d'être couverte. Ainsi, nous construisons un portefeuille  $V$  constitué d'un produit  $C_1$  de sous-jacent  $S$ , une quantité  $-\Delta$  du stock (comme avec le modèle de Black and Scholes) et une quantité  $-\Theta$  d'un autre produit  $C_2$  écrit toujours sur  $S$  qui dépend donc de la volatilité  $v$ . Nous avons alors

$$V = C_1 - \Delta S - \Theta C_2.$$

En considérant un portefeuille auto-finançant et à l'aide de la formule d'Itô multidimensionnelle pour  $V$ , dépendant de  $t$ ,  $v$  et  $S$ , nous obtenons

$$\begin{aligned}
 dV &= dC_1 - \Delta dS - \Theta dC_2 \\
 &= \left\{ \frac{\partial C_1}{\partial t} dt + \frac{\partial C_1}{\partial S} dS + \frac{\partial C_1}{\partial v} dv + \frac{\partial^2 C_1}{\partial S \partial v} d\langle S, v \rangle + \frac{1}{2} \frac{\partial^2 C_1}{\partial S^2} d\langle S, S \rangle + \frac{1}{2} \frac{\partial^2 C_1}{\partial v^2} d\langle v, v \rangle \right\} \\
 &\quad - \Delta dS \\
 &\quad - \Theta \left\{ \frac{\partial C_2}{\partial t} dt + \frac{\partial C_2}{\partial S} dS + \frac{\partial C_2}{\partial v} dv \right. \\
 &\quad \left. + \frac{\partial^2 C_2}{\partial S \partial v} d\langle S, v \rangle + \frac{1}{2} \frac{\partial^2 C_2}{\partial S^2} d\langle S, S \rangle + \frac{1}{2} \frac{\partial^2 C_2}{\partial v^2} d\langle v, v \rangle \right\} \\
 &= \left\{ \frac{\partial C_1}{\partial t} dt + \frac{\partial C_1}{\partial S} dS + \frac{\partial C_1}{\partial v} dv + \eta \beta \rho v S \frac{\partial^2 C_1}{\partial S \partial v} dt \right. \\
 &\quad \left. + \frac{1}{2} v S^2 \frac{\partial^2 C_1}{\partial S^2} dt + \frac{1}{2} v \eta \beta \frac{\partial^2 C_1}{\partial v^2} dt \right\} \\
 &\quad - \Delta dS \\
 &\quad - \Theta \left\{ \frac{\partial C_2}{\partial t} dt + \frac{\partial C_2}{\partial S} dS + \frac{\partial C_2}{\partial v} dv + \eta \beta \rho v S \frac{\partial^2 C_2}{\partial S \partial v} dt \right. \\
 &\quad \left. + \frac{1}{2} v S^2 \frac{\partial^2 C_2}{\partial S^2} dt + \frac{1}{2} v \eta \beta \frac{\partial^2 C_2}{\partial v^2} dt \right\} \\
 &= \left\{ \frac{\partial C_1}{\partial t} + \eta \beta \rho v S \frac{\partial^2 C_1}{\partial S \partial v} + \frac{1}{2} v S^2 \frac{\partial^2 C_1}{\partial S^2} + \frac{1}{2} v \eta \beta \frac{\partial^2 C_1}{\partial v^2} dt \right\} dt \\
 &\quad - \Theta \left\{ \frac{\partial C_2}{\partial t} + \eta \beta \rho v S \frac{\partial^2 C_2}{\partial S \partial v} + \frac{1}{2} v S^2 \frac{\partial^2 C_2}{\partial S^2} + \frac{1}{2} v \eta \beta \frac{\partial^2 C_2}{\partial v^2} dt \right\} dt \\
 &\quad + \left\{ \frac{\partial C_1}{\partial S} - \Theta \frac{\partial C_2}{\partial S} - \Delta \right\} dS \\
 &\quad + \left\{ \frac{\partial C_1}{\partial v} - \Theta \frac{\partial C_2}{\partial v} \right\} dv.
 \end{aligned}$$

Pour éliminer les sources d'aléas nous mettons à 0 les termes en  $dS$  et  $dv$ , soit,

$$\frac{\partial C_1}{\partial S} - \Theta \frac{\partial C_2}{\partial S} - \Delta = 0 \tag{3.24}$$

et

$$\frac{\partial C_1}{\partial v} - \Theta \frac{\partial C_2}{\partial v} = 0 \tag{3.25}$$

Ainsi,

$$dV_t = \left\{ \frac{\partial C_1}{\partial t} + \eta\beta\rho v S \frac{\partial^2 C_1}{\partial S \partial v} + \frac{1}{2} v S^2 \frac{\partial^2 C_1}{\partial S^2} + \frac{1}{2} v \eta \beta \frac{\partial^2 C_1}{\partial v^2} \right\} dt - \Theta \left\{ \frac{\partial C_2}{\partial t} + \eta\beta\rho v S \frac{\partial^2 C_2}{\partial S \partial v} + \frac{1}{2} v S^2 \frac{\partial^2 C_2}{\partial S^2} + \frac{1}{2} v \eta \beta \frac{\partial^2 C_2}{\partial v^2} \right\} dt \quad (3.26)$$

Le portefeuille étant maintenant sans risque, son rendement devrait être égal à celui du taux sans risque  $r$ ,

$$\begin{aligned} dV_t &= rV_t dt \\ &= r(C_1 - \Delta S - \Theta C_2) dt. \end{aligned} \quad (3.27)$$

En égalisant les équations (3.26) et (3.27), et en utilisant les équations (3.24) et (3.25), nous avons

$$\begin{aligned} &\frac{1}{\frac{\partial C_1}{\partial v}} \left( \frac{\partial C_1}{\partial t} + \eta\beta\rho v S \frac{\partial^2 C_1}{\partial S \partial v} + \frac{1}{2} v S^2 \frac{\partial^2 C_1}{\partial S^2} + \frac{1}{2} v \eta \beta \frac{\partial^2 C_1}{\partial v^2} + rS \frac{\partial C_1}{\partial S} - rC_1 \right) \\ &= \frac{1}{\frac{\partial C_2}{\partial v}} \left( \frac{\partial C_2}{\partial t} + \eta\beta\rho v S \frac{\partial^2 C_2}{\partial S \partial v} + \frac{1}{2} v S^2 \frac{\partial^2 C_2}{\partial S^2} + \frac{1}{2} v \eta \beta \frac{\partial^2 C_2}{\partial v^2} + rS \frac{\partial C_2}{\partial S} - rC_2 \right). \end{aligned}$$

Le terme de gauche dépend uniquement de  $C_1$ , le terme de droite de  $C_2$ , sans perte de généralité, nous pouvons alors supposer qu'ils sont égaux à une certaine fonction  $f$  des variables  $S$ ,  $v$  et  $t$ . Ainsi

$$\frac{\partial C}{\partial t} + \eta\beta\rho v S \frac{\partial^2 C}{\partial S \partial v} + \frac{1}{2} v S^2 \frac{\partial^2 C}{\partial S^2} + \frac{1}{2} v \eta^2 \beta \frac{\partial^2 C}{\partial v^2} + rS \frac{\partial C}{\partial S} - rC = f \frac{\partial C}{\partial v}. \quad (3.28)$$

Pour des considérations économiques que nous ne détaillons pas ici, la fonction  $f$  est choisie tel que  $f(S, v, t) = (-\alpha(S, v, t) - \phi(S, v, t)\sqrt{v_t})$  où  $\phi(S, v, t)$  est appelé le prix du risque de volatilité.

### 3.4.1 Lien avec la volatilité implicite

Considérons un modèle à volatilité stochastique où les processus de prix et de volatilité sont non corrélés. Le pricing d'une option de payoff  $\phi(S_T, \sigma_T)$  à la forme générique

$$V_{SV}(t, T) = e^{-r(T-t)} \int \int \phi(x, y) d\mathbb{P}(S_T = x, \sigma_T = y),$$

où  $SV$  ou pour valorisation sous un modèle à volatilité stochastique. En appliquant simplement la formule de Bayes nous avons

$$\begin{aligned} V_{SV}(t, T) &= e^{-r(T-t)} \int \int \phi(x, y) d\mathbb{P}(S_T = x | \sigma_T = y) \mathbb{P}(\sigma_T = y) \\ &= \int \left\{ e^{-r(T-t)} \int f_{S_T | \sigma_T=y}(x) \phi(x, y) dx \right\} d\mathbb{P}(\sigma_T = y) \\ &= \int f_{\sigma_T}(y) V_{BS}(t, T, \sigma_T) dy. \end{aligned}$$

La dernière équation se déduit de fait que si la volatilité est fixée alors on se retrouve dans le cadre de Black and Scholes. Le modèle de Black and Scholes est une variable déterministe, la seule variable aléatoire est  $\sigma_T$ , la dernière équation est en fait l'espérance du modèle de Black and Scholes sous  $\sigma_T$

$$V_{SV}(t, T) = \mathbb{E}[V_{BS}(\sigma_T)]. \quad (3.29)$$

Valoriser une option avec un modèle à volatilité stochastique correspond donc à faire une moyenne des prix de Black and Scholes pour toutes les valeurs possibles de  $\sigma_T$ . Ce résultat est connu sous le terme de formule de mélange. Nous allons maintenant en déduire une approximation de la volatilité implicite dans le cadre d'un modèle à volatilité stochastique ainsi qu'une correspondance entre la sensibilité de la volatilité implicite par rapport au sous-jacent et par rapport au strike.

Notons  $V_{SV} = V_{SV}(\sigma_T)$ . En faisant un développement limité de  $V_{SV}(\sigma_T)$  au voisinage de  $\bar{\sigma}_T = \mathbb{E}[\sigma_T]$  nous avons

$$\begin{aligned} V_{SV} &= \int \left[ V_{BS}(\bar{\sigma}_T) + \frac{\partial V_{BS}(\bar{\sigma}_T)}{\partial \sigma_T} (\sigma_T - \bar{\sigma}_T) + \frac{1}{2} \frac{\partial^2 V_{BS}(\bar{\sigma}_T)}{\partial \sigma_T^2} (\sigma_T - \bar{\sigma}_T)^2 + \dots \right] f(\sigma_T) d\sigma_T \\ &\approx V_{BS}(\bar{\sigma}_T) + \frac{\partial V_{BS}(\bar{\sigma}_T)}{\partial \sigma_T} (\mathbb{E}[\sigma_T] - \bar{\sigma}_T) + \frac{1}{2} \frac{\partial^2 V_{BS}(\bar{\sigma}_T)}{\partial \sigma_T^2} \mathbb{E}[(\sigma_T - \bar{\sigma}_T)^2] \\ &= V_{BS}(\bar{\sigma}_T) + 0 + \frac{1}{2} \frac{\partial^2 V_{BS}(\bar{\sigma}_T)}{\partial \sigma_T^2} Var(\sigma_T). \end{aligned} \quad (3.30)$$

Si nous supposons maintenant que valoriser une option avec un modèle à volatilité stochastique correspond à la même chose que valoriser une option avec le modèle de Black and Scholes mais en prenant en compte la volatilité implicite plutôt qu'une volatilité constante, nous avons

$$\begin{aligned} V_{SV} &= V_{BS}(\Sigma) \\ &\approx V_{BS}(\bar{\sigma}_T) + \frac{\partial V_{BS}(\bar{\sigma}_T)}{\partial \sigma_T} (\Sigma - \bar{\sigma}_T). \end{aligned} \quad (3.31)$$

où nous avons supposé que  $\frac{\partial V_{BS}(\bar{\sigma}_T)}{\partial \Sigma} = \frac{\partial V_{BS}(\bar{\sigma}_T)}{\partial \sigma_T}$ . Il ne reste plus qu'à égaliser les équations (3.30) et (3.31) pour obtenir une approximation de la volatilité implicite

$$\Sigma \approx \bar{\sigma}_T + \frac{\frac{1}{2} \frac{\partial^2 V_{BS}(\bar{\sigma}_T)}{\partial \sigma_T^2} Var(\sigma_T)}{\frac{\partial V_{BS}(\bar{\sigma}_T)}{\partial \sigma_T}}.$$

Notons qu'il est possible de raffiner l'approximation en faisant un développement plus lointain de (3.30), incorporant skewness et kurtosis par exemple. Dans le cas simple où le produit est une option européenne, ou certaines exotiques, nous pouvons calculer explicitement les dérivées du produit.

Pour montrer la correspondance entre la sensibilité de la volatilité implicite par rapport au sous-jacent et par rapport au strike, rappelons que la formule de Black and Scholes d'un call européen est homogène en  $S$  et  $K$ ,  $C_{BS}(t, T, S, K, \sigma) = SC_{BS}(t, T, 1, K/S, \sigma)$ . Si nous revenons à notre formule de mélange (3.29), un call sous volatilité stochastique correspond à la moyenne des call de Black and Sholes pour différentes volatilités. Si nous n'avons que deux volatilités à disposition, la plus faible  $\sigma_L$  et la plus haute,  $\sigma_H$ , (3.29) devient

$$\begin{aligned} C_{SV} &= \frac{1}{2} [C_{BS}(S, K, \sigma_L) + \sigma_{BS}(S, K, \sigma_H)] \\ &= \frac{1}{2} S \left[ C_{BS}\left(1, \frac{K}{S}, \sigma_L\right) + C_{BS}\left(1, \frac{K}{S}, \sigma_H\right) \right] \\ &= S f\left(\frac{K}{S}\right) \end{aligned}$$

pour  $f$  une fonction ne dépendant que de  $K/S$ . Nous avons ainsi

$$C_{SV} = S f\left(\frac{K}{S}\right) = S C_{BS}\left(1, \frac{K}{S}, \Sigma\right)$$

et donc

$$\Sigma = g\left(\frac{K}{S}\right)$$

pour une fonction  $g$  de  $K/S$ . Ainsi

$$\begin{aligned} \frac{\partial \Sigma}{\partial S} &= -\frac{K}{S^2} g' \\ \frac{\partial \Sigma}{\partial K} &= \frac{1}{S} g' \\ S \frac{\partial \Sigma}{\partial S} + K \frac{\partial \Sigma}{\partial K} &= 0. \end{aligned}$$

Si nous sommes proche de la monnaie ( $K \approx S$ ), nous avons alors

$$\frac{\partial \Sigma}{\partial S} \approx \frac{\partial \Sigma}{\partial K}.$$

### 3.4.2 Modèle de Heston

Dans le modèle de Heston, les équations de diffusions sous la probabilité risque neutre sont donc

$$\begin{aligned} dS_t &= \mu_t S_t dt + \sqrt{v_t} S_t dZ_t^1 \\ dv_t &= \kappa(\theta - v_t) dt + \sigma \sqrt{v_t} dZ_t^2, \end{aligned}$$

avec  $\langle dZ^1, dZ^2 \rangle = \rho dt$ .

En reprenant (3.28),

$$\begin{aligned} \frac{\partial C}{\partial t} + \sigma \rho v S \frac{\partial^2 C}{\partial S \partial v} + \frac{1}{2} v S^2 \frac{\partial^2 C}{\partial S^2} + \frac{1}{2} v \sigma^2 \frac{\partial^2 C}{\partial v^2} + r S \frac{\partial C}{\partial S} - r C &= \\ (\kappa(v - \theta) - \phi) \frac{\partial C}{\partial v}. \end{aligned} \quad (3.32)$$

Il paraît justifié de choisir le prix marché du risque de volatilité comme fonction de la volatilité. Plus précisément, nous choisissons un risque proportionnel à la variance  $v$ ,  $\phi = \lambda \sqrt{v}$  avec  $\lambda$  une constante. Définissons les paramètres ajustés au risque  $\kappa'$  et  $\theta'$  tel que  $\kappa' = \kappa - \lambda$  et  $\kappa' \theta' = \kappa \theta$ . L'équation (3.32) devient alors

$$\begin{aligned} \frac{\partial C}{\partial t} + \sigma \rho v S \frac{\partial^2 C}{\partial S \partial v} + \frac{1}{2} v S^2 \frac{\partial^2 C}{\partial S^2} + \frac{1}{2} v \sigma^2 \frac{\partial^2 C}{\partial v^2} + r S \frac{\partial C}{\partial S} - r C &= \\ = \kappa'(v - \theta') \frac{\partial C}{\partial v}. \end{aligned} \quad (3.33)$$

Posons

$$x_t = \ln \left( \frac{F(t, T)}{K} \right)$$

avec  $F(t, T)$  le prix forward de  $S_t$ ,  $F(t, T) = S_t e^{\int_t^T r_s ds}$ . La valeur d'un call européen  $C$  est alors

$$-\frac{\partial C}{\partial \tau} + \frac{1}{2} v C_{11} + \frac{1}{2} v C_1 + \frac{1}{2} \sigma^2 v C_{22} + \rho \sigma v(t) C_{12} - \kappa(v(t) - \theta) C_2 = 0, \quad (3.34)$$

où  $\tau = T - t$  et les paramètres  $\kappa$  et  $\theta$  sont risque-ajustés (nous avons retiré les primes pour simplifier). Nous allons maintenant chercher une solution à la 'Black & Scholes'. Plus particulièrement, nous posons

$$C(x, v, \tau) = e^x P_1(x, v, \tau) - P_0(x, v, \tau). \quad (3.35)$$

Substituant (3.35) dans (3.34), nous obtenons,

$$-\frac{\partial P_j}{\partial \tau} + \frac{1}{2}v \frac{\partial P_j}{\partial x^2} + \left(\frac{1}{2} - j\right)v \frac{\partial P_j}{\partial v^2} + \frac{1}{2}\sigma^2 v \frac{\partial^2 P_j}{\partial v^2} + \rho\sigma v \frac{\partial^2 P_j}{\partial x \partial v} + (a - b_j) \frac{\partial P_j}{\partial v} = 0 \quad (3.36)$$

pour  $j = 0, 1$ ,  $a = \kappa\theta$ ,  $b = \kappa - j\rho\sigma$  et, comme condition terminale

$$P_j(x, v, 0) = \mathbf{1}_{x>0} \quad (3.37)$$

Pour pouvoir résoudre l'équation (3.36) avec la condition terminale (3.37), nous utilisons la transformée de Fourier. La transformée de Fourier (ou fonction caractéristique) de  $P$  est définie par

$$\tilde{P}(k, v, \tau) = \int_{-\infty}^{\infty} e^{-ikx} dx P(x, v, \tau)$$

son inverse étant

$$P(x, v, \tau) = \int_{-\infty}^{\infty} e^{ikx} \tilde{P}(k, v, \tau) \frac{dk}{2\pi}. \quad (3.38)$$

En substituant à nouveau dans (3.36)

$$-\frac{\partial \tilde{P}_j}{\partial \tau} - \frac{1}{2}k^2 v \tilde{P}_j - \left(\frac{1}{2} - j\right) ikv \tilde{P}_j + \frac{1}{2}\sigma^2 v \frac{\partial^2 \tilde{P}_j}{\partial v^2} + \rho\sigma ikv \frac{\partial \tilde{P}_j}{\partial v} + (a - b_j) \frac{\partial \tilde{P}_j}{\partial v} = 0.$$

Pour simplifier l'expression, nous introduisons les notations

$$\begin{aligned} \alpha &= -\frac{k^2}{2} - \frac{ik}{2} + ijk \\ \beta &= \kappa - \rho\sigma j - \rho\sigma ik \\ \gamma &= \frac{\sigma^2}{2}. \end{aligned}$$

Nous obtenons ainsi

$$v \left( \alpha \tilde{P}_j - \beta \frac{\partial \tilde{P}_j}{\partial v} \right) + a \frac{\partial \tilde{P}_j}{\partial v} - \frac{\partial \tilde{P}_j}{\partial \tau} = 0. \quad (3.39)$$

Posons, une nouvelle fois, l'ansatz suivante

$$\begin{aligned}\tilde{P}_j(k, v, \tau) &= \exp\{C_j(k, \tau)\theta + D_j(k, \tau)v\} \tilde{P}(k, v, 0) \\ &= \frac{1}{ik} \exp\{C_j(k, \tau)\theta + D_j(k, \tau)v\},\end{aligned}$$

où les fonctions  $C_j$  et  $D_j$  sont ce que l'on recherche. Les dérivées sont

$$\begin{aligned}\frac{\partial \tilde{P}_j}{\partial \tau} &= \left( \theta \frac{\partial C_j}{\partial \tau} + v \frac{\partial D_j}{\partial \tau} \right) \tilde{P}_j \\ \frac{\partial \tilde{P}_j}{\partial v} &= D_j \tilde{P}_j \\ \frac{\partial^2 \tilde{P}_j}{\partial v^2} &= D_j^2 \tilde{P}_j.\end{aligned}$$

L'équation (3.39) devient ainsi

$$v \left( \alpha P_j - \beta D_j \tilde{P}_j + \gamma D_j^2 \tilde{P}_j + \frac{\partial D_j}{\partial \tau} \tilde{P}_j \right) + a D_j \tilde{P}_j - \theta \frac{\partial C_j}{\partial \tau}$$

On annule maintenant les termes en  $v$  et les termes ne dépendant pas de  $v$ , on obtient le système

$$\begin{aligned}\frac{\partial C_j}{\partial \tau} &= \frac{a}{\theta} D_j \\ &= \kappa D_j \\ \frac{\partial D_j}{\partial \tau} &= \alpha - \beta D_j + \gamma D_j^2 \\ &= \gamma(D_j - r_+)(D_j - r_-)\end{aligned}\tag{3.40}$$

avec

$$r_{\pm} = \frac{\beta \pm d}{\sigma^2}, \quad d = \sqrt{\beta^2 - 4\alpha\gamma}$$

Il ne reste plus qu'à intégrer le système (3.40) avec les conditions terminales  $C_j(k, 0) = 0$  et  $D_j(k, 0) = 0$  (correspondant à (3.37)). Après calculs

$$\begin{aligned}C(k, \tau) &= \kappa \left[ r_- \tau - \frac{2}{\sigma^2} \ln \left( \frac{1 - ge^{-d\tau}}{1 - g} \right) \right] \\ D(k, \tau) &= r_- \frac{1 - e^{-d\tau}}{1 - ge^{-d\tau}}\end{aligned}$$

où  $g = r_-/r_+$ . Il ne reste plus qu'à inverser la transformée de Fourier pour trouver le résultat voulu. En utilisant (3.38), les pseudo-probabilités  $P_j, j = 1, 2$  sont données par

$$P_j(x, v, \tau) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re \left\{ \frac{\exp \{C_j(k, \tau)\theta + D_j(k, \tau)v + ikx\}}{ik} \right\} dk,$$

Equation que l'on peut estimer par des méthodes numériques classiques comme la quadrature de Gauss - Laguerre pour résoudre le problème de l'intégrale infinie.

Nous allons maintenant essayer d'extraire la volatilité implicite du modèle de Heston. Notons  $x_t$  le log-strike à l'instant  $t$ ,  $x_t = \log(S_t/K)$  et supposons que le drift du sous-jacent est nul  $\mu = 0$ , le modèle de Heston devient alors

$$\begin{aligned} dx_t &= -\frac{v_t}{2} dt + \sqrt{v_t} dZ_t \\ dv_t &= -\kappa(v_t - \theta)dt + \rho\eta\sqrt{v_t} dZ_t + \sqrt{1 - \rho^2}\eta\sqrt{v_t} dW_t \end{aligned}$$

où nous avons posé  $Z_t^1 = Z_t$  et  $Z_t^2 = \rho Z_t + \sqrt{1 - \rho^2}W_t$  avec  $W_t$  un mouvement brownien indépendant de  $Z_t$ . En appliquant l'opérateur espérance conditionnelle à  $v_t$ , nous avons en notant  $\sigma_t = \mathbb{E}(v_t|x_T)$

$$d\sigma_t = \kappa(\theta - \sigma_t)dt + \rho\eta \left( d\mathbb{E}(x_t|x_T) + \frac{1}{2}\sigma_t \right) + \sqrt{1 - \rho^2}\eta\mathbb{E}(\sqrt{v_t}dW_t|x_T). \quad (3.41)$$

Nous assumons que la dernière espérance est égale à zéro et que

$$\mathbb{E}(x_t|x_T) = x_T \frac{\int_0^t \mathbb{E}(v_s)ds}{\int_0^T \mathbb{E}(v_s)ds} = x_T \frac{s_t}{s_T}$$

(la dernière égalité est juste une notation). En injectant cela dans (3.41)

$$d\sigma_t = \kappa(\theta - \sigma_t)dt + \rho\eta \frac{x_T}{s_T} ds_t - \frac{1}{2}\rho\eta\sigma_t.$$

Dans le cas particulier où nous conditionnons par  $x_T = 0$ , soit  $K = S_T$ ,  $\sigma_t$  est la volatilité locale comme il a été montré (3.16). La volatilité locale de Heston vérifie donc

$$\begin{aligned} d\sigma_t &= \kappa(\theta - \sigma_t)dt + \frac{1}{2}\rho\eta\sigma_t. \\ &= \kappa'(\theta' - \sigma_t)dt \end{aligned}$$

avec  $\kappa' = \kappa - \rho\eta/2$  et  $\theta' = \kappa\theta/\kappa$ . La solution de ce processus est

$$\sigma_t = e^{-\kappa't}(\sigma_0 - \theta') + \theta'$$

Ainsi, dans le cas particulier où la volatilité implicite ne dépend pas du strike (3.20) nous en déduisons la formule suivante

$$\begin{aligned} I_{Heston}^2(t) &= \frac{1}{T} \int_0^T (e^{-\kappa't}(\sigma_0 - \theta') + \theta') dt \\ &= \theta' + \frac{\sigma_0 - \theta'}{\kappa'} (1 - e^{-\kappa'}). \end{aligned}$$

### 3.4.3 Modèle de Wu et Zhang

Le papier de Wu et Zang, [WuZh02], prolonge le modèle de Heston pour le marché Libor. Les équations de diffusions d'un forward sous la probabilité risque neutre sont

$$\begin{aligned} df_t^j &= f_t^j \sqrt{V_t} \gamma_t^j \cdot (d\mathbf{Z}_t - \sqrt{V_t} \sigma_{j+1} dt) \\ dV_t &= \kappa(\theta - V_t) + \epsilon \sqrt{V_t} dW_t, \end{aligned} \tag{3.42}$$

où  $f_t^j = f(t, T_j, T_{j+1})$ . La volatilité (stochastique) suit donc un processus d'Ornstein Ulhenbeck,  $\epsilon$  est habituellement appelé le paramètre volvol, il contrôle la courbure du smile.  $\cdot$  est le produit scalaire,  $\mathbf{Z}$  est un vecteur de  $\mathbb{Q}$ -mouvements Brownien standard indépendants et  $(W_t)_{t \in \mathbb{R}_+}$  un mouvement Brownien standard tel que

$$\frac{(\gamma_t^j \cdot \mathbf{Z}_t) dW_t}{\|\gamma_t^j\|} = \rho_t^j dt.$$

Les fonctions  $(\gamma_t^j)_{t \in \mathbb{R}_+}$  donnent le niveau de courbure du smile. Elles sont données par

$$\gamma_t^j = \frac{1 + (T_{j+1} - T_j) f_t^j}{(T_{j+1} - T_j) f_t^j} (\sigma(t, T_{j+1}) - \sigma(t, T_j)).$$

Comme précédemment pour les mesures forward nous introduisons les mouvements browniens  $\mathbf{Z}_t^{j+1}$  et  $W_t^{j+1}$  sous la mesure  $T_{j+1}$ -forward tel que

$$\begin{aligned} d\mathbf{Z}_t^{j+1} &= d\mathbf{Z}_t - \sqrt{V_t} \sigma^{j+1}(t) dt \\ dW_t^{j+1} &= dW_t + \xi_t^j \sqrt{V_t} dt, \end{aligned}$$

où

$$\xi_t^j = \sum_{k=1}^j \frac{(T_{j+1} - T_j) f_t^k \rho_t^k ||\gamma_t^k||}{1 + (T_{k+1} - T_k) f_t^k} \quad (3.43)$$

Les équations de diffusions (3.42) sous la mesure  $T_{j+1}$ -forward sont alors

$$\begin{aligned} df_t^j &= f_t^j \sqrt{V_t} \gamma_j \cdot d\mathbf{Z}_t^{j+1} \\ dV_t &= (\kappa\theta - (\kappa + \epsilon\xi_t^j)V_t)dt + \epsilon\sqrt{V_t}dW_t^{j+1}. \end{aligned}$$

Malheureusement,  $\xi_t^j$  dépend de la dynamique des taux forward  $(f_t^j)_{t \geq 0}$ , il n'est pas possible d'obtenir des formules 'à la Heston'. Utilisant la propriété de martingale de  $(f_t^j)_{t \geq 0}$  sous la mesure  $T_{j+1}$ -forward, nous avons  $\mathbb{E}_{j+1}(f_t^j | \mathcal{F}_0) = f_0^j$  et nous approximons (3.43) par

$$\xi_t^j \approx \sum_{k=1}^j \frac{(T_{j+1} - T_j) f_t^k \rho_t^k ||\gamma_t^k||}{1 + (T_{k+1} - T_k) f_t^k},$$

nous avons donc 'freezé' les coefficients. Posons maintenant  $\tilde{\xi}_t^j = 1 + \frac{\epsilon}{\kappa}\xi_t^j$ , alors,

$$dV_t = \kappa(\theta - \tilde{\xi}_t^j V_t)dt + \epsilon\sqrt{V_t}dW_t^{j+1}.$$

Maintenant que les équations régissant la dynamique des taux forward et de sa volatilité sont posées sous la mesure forward neutre, nous pouvons nous intéresser à la valorisation d'un caplet en  $t = 0$  de maturité  $T_{j+1}$  et commençant en  $T_j$  de strike  $K$

$$\begin{aligned} \text{Cpl}(0, T_j, T_{j+1}) &= B(0, T_{j+1})(T_{j+1} - T_j)\mathbb{E}_{j+1}[(f_{T_j}^j - K)_+ | \mathcal{F}_0(t)] \\ &= B(0, T_{j+1})(T_{j+1} - T_j)f_0^j \mathbb{E}_{j+1}\left[\left(\frac{f_{T_j}^j}{f_0^j} - \frac{K}{f_0^j}\right)_+ | \mathcal{F}_t^0\right] \\ &= B(0, T_{j+1})(T_{j+1} - T_j)f_0^j \mathbb{E}_{j+1}[e^{\ln(f_{T_j}^j/f_0^j)} \mathbf{1}_{f_t^j > K} | \mathcal{F}_t^0] \\ &\quad + B(0, T_{j+1})(T_{j+1} - T_j)K \mathbb{E}_{j+1}[\mathbf{1}_{f_t^j > K} | \mathcal{F}_t^0]. \end{aligned}$$

Les deux espérances peuvent être évaluées en utilisant la fonction génératrice des moments du taux forward

$$\phi(X_t, V_t, t, z) = \mathbb{E}[e^{zX_t} | \mathcal{F}_t], \quad z \in \mathbb{C},$$

avec  $X = \ln(f_{T_j}^j/f_0^j)$ . Par définition de la fonction caractéristique (voir e.g. [DuPaSi00]), nous obtenons

$$\begin{aligned}\mathbb{E}_{j+1}[e^{\ln(f_{T_j}^j/f_0^j)} \mathbf{1}_{f_t^j > K} | \mathcal{F}_t^0] &= \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{\Im(e^{-iu \ln(K/f_0^j)} \phi_T(iu))}{u} du \\ \mathbb{E}_{j+1}[e^{\ln(f_j(T_j)/f_j(0))} \mathbf{1}_{f_t^j > K} | \mathcal{F}_t^0] &= \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{\Im(e^{-iu \ln(K/f_0^j)} \phi_T(1+iu))}{u} du,\end{aligned}$$

où  $\phi_T(z) = \phi(0, V_0, 0, z)$  et  $\Im(x)$  dénote la partie imaginaire de  $x$ . Les deux espérances peuvent ainsi être approximées par des méthodes numériques usuelles. Il est tout de même possible d'obtenir une formule fermée. Par simple application du théorème de Feyman-Kac multi - dimensionnel,

$$\begin{aligned}\frac{\partial \phi}{\partial t} + \kappa(\theta - \tilde{\xi}_j V) \frac{\partial \phi}{\partial V} - \frac{1}{2} \|\gamma_j(t)^2\| V \frac{\partial \phi}{\partial x} + \frac{1}{2} \epsilon^2 V \frac{\partial^2 \phi}{\partial V^2} \\ + \epsilon \rho_j V \|\gamma^j\| \frac{\partial^2 \phi}{\partial V \partial x} + \frac{1}{2} \|\gamma_t^j\|^2 V \frac{\partial^2 \phi}{\partial x^2} = 0\end{aligned}\quad (3.44)$$

avec comme condition terminale

$$\phi(x, V, T, z) = e^{zx}. \quad (3.45)$$

De la même manière que Heston, nous posons l'ansatz,

$$\phi(x, V, t, z) = \hat{\phi}(x, V, \tau, z) = e^{A(\tau, z) + B(\tau, z)V + zx}$$

avec  $\tau = T_{j+1} - t$ . En substituant dans les équations (3.44) et (3.45)

$$\begin{aligned}\frac{dA}{d\tau} &= \kappa \theta B \\ \frac{dB}{d\tau} &= \frac{1}{2} \epsilon^2 B^2 + (\rho \epsilon \|\gamma_j\| - \kappa \tilde{\xi}_j) B + \frac{1}{2} \|\gamma^j\|^2 (z^2 - z)\end{aligned}$$

avec comme conditions initiales

$$A(0, z) = 0, \quad B(0, z) = 0.$$

On reconnaît alors la forme des équations différentielles de Riccati. Dans le cas où les fonctions  $\rho_t^j$  et  $\gamma_t^j$  sont constantes par morceaux, après résolution (voir [WuZh02]), nous obtenons pour tout  $\tau_j \leq \tau < \tau_{j+1}$  les équations de récurrences

$$\begin{aligned}A(\tau, z) &= A(\tau_j, z) + \frac{\kappa \theta}{\epsilon^2} \left[ (a + d)(\tau - \tau^j) - 2 \ln \left( \frac{1 - g_j e^{d(\tau - \tau^j)}}{1 - g_j} \right) \right] \\ B(\tau, z) &= B(\tau^j, z) + \frac{(a + d - \epsilon^2 B(\tau_j, z))(1 - e^{d(\tau - \tau^j)})}{\epsilon^2 (1 - g_j e^{d(\tau - \tau^j)})},\end{aligned}$$

où

$$a = \kappa \tilde{\xi}_t^j - \rho_t \epsilon ||\gamma_t^j|| z, \quad d = \sqrt{a^2 - ||\gamma_t^j||^2 \epsilon^2 (z^2 - z)}, \quad g_j = \frac{a + d - \epsilon^2 B(\tau^j, z)}{a - d - \epsilon^2 B(\tau^j, z)}.$$

### 3.4.4 Modèle SABR

Nous passons au modèle de Hagan et al., [HaKuLeWo02], SABR, pour "Stochastique, Alpha, Beta, Rho. Il s'agit d'un modèle à volatilité stochastique caractérisant la dynamique d'un forward sous la probabilité risque neutre par

$$\begin{aligned} dF_{t,j} &= V_t F_j^\beta(t) dZ_{t,j}^j \\ dV_t &= \epsilon V_t dW_t^j \\ V_0 &= \alpha \end{aligned}$$

avec  $(Z_{t,j}^j)_{t \in \mathbb{R}_+}$  et  $(W_t^j)_{t \in \mathbb{R}_+}$  deux mouvements Brownien standard vérifiant

$$\langle dZ_{t,j}^j dW_t^j \rangle = \rho dt,$$

et  $\beta \in (0, 1]$ ,  $\epsilon$  et  $\alpha$  deux constantes réelles positives et  $\rho \in [-1, 1]$ . Le taux forward  $F_{t,j}$  est donc type CEV avec une volatilité stochastique suivant un mouvement Brownien géométrique sans drift. Il s'agit d'un des modèles les plus populaires actuellement pour sa simplicité à mettre en oeuvre et ses approximations pour les volatilités implicites. Depuis, des résultats plus généraux ont été montrés, nous renvoyons par exemple à [La05] ou [La08] pour quelque chose de plus complet. Notons que les prémisses de cette théorie ont été instiguées par [BeBuFl04], abordable par un étudiant un peu motivé.

Les auteurs de [HaKuLeWo02] démontrent par perturbation singulière que le prix en  $t = 0$  d'un caplet est

$$\text{Cpl}(0, T_{j-1}, T_j, \kappa) = (T_j - T_{j-1}) B(0, T_j) (F_{0,j} \phi(d_+) - K \phi(d_-))$$

où

$$d_\pm = \frac{\ln(F_{0,j}/K) \pm \frac{1}{2} \sigma^{\text{imp}}(\kappa, F_{0,j})^2 T_{j-1}}{\sigma^{\text{imp}}(\kappa, F_{0,j})^2 \sqrt{T_{j-1}}}$$

et  $\sigma^{\text{imp}}(\kappa, F)$  est la volatilité implicite vérifiant l'approximation

$$\begin{aligned} \sigma^{\text{imp}}(\kappa, F) &= \frac{\alpha}{(F\kappa)^{\frac{1-\beta}{2}} \left[ 1 + \frac{(1-\beta)^2}{24} \ln^2 \left( \frac{F}{\kappa} \right) + \frac{(1-\beta)^4}{1920} \ln^4 \left( \frac{F}{\kappa} \right) + \dots \right] x(z)} \\ &\cdot \left\{ 1 + \left[ \frac{(1-\beta)^2 \alpha^2}{24(F\kappa)^{1-\beta}} + \frac{\rho \beta \epsilon \alpha}{4(F\kappa)^{\frac{1-\beta}{2}}} + \epsilon^2 \frac{2-3\rho^2}{24} \right] T_{j-1} + \dots \right\}, \end{aligned} \quad (3.46)$$

avec

$$z = \frac{\epsilon}{\alpha} (F\kappa)^{\frac{1-\beta}{2}} \ln \left( \frac{F}{\kappa} \right)$$

et

$$x(z) = \ln \left( \frac{\sqrt{1 - 2\rho z + z^2} + z - \rho}{1 - \rho} \right).$$

Dans le cas particulier où nous sommes à la monnaie (ATM), i.e.  $\kappa = F = F_{0,j}$ , nous avons

$$\begin{aligned} \sigma^{\text{ATM}} &= \sigma^{\text{imp}}(F_{0,j}, F_{0,j}) \\ &= \frac{\alpha}{F_{0,j}^{1-\beta}} \left[ 1 + \left( \frac{(1-\beta)^2 \alpha^2}{24 F_{0,j}^{2-2\beta}} + \frac{\rho\beta\epsilon\alpha}{4F_{0,j}^{1-\beta}} + \epsilon^2 \frac{2-3\rho^2}{24} \right) T_{j-1} + \dots \right] \end{aligned} \quad (3.47)$$

Le terme principal de l'équation (3.47) est  $\frac{\alpha}{F_{0,j}^{1-\beta}}$ , les paramètres  $\alpha$  et  $\beta$  vont influer sur la courbure comme sur le niveau de la volatilité implicite ATM.

Pour comprendre le rôle de chacun des paramètres, les auteurs de [HaKuLeWo02] proposent une nouvelle approximation de (3.46),

$$\begin{aligned} \sigma^{\text{imp}}(K, F_{0,j}) &= \frac{\alpha}{F_{0,j}^{1-\beta}} \left\{ 1 - \frac{1}{2}(1-\beta-\rho\lambda) \ln \frac{\kappa}{F_{0,j}} \right. \\ &\quad \left. + \frac{1}{12} [(1-\beta)^2 + (2-3\rho^2)\lambda^2 \ln^2 \frac{\kappa}{F_{0,j}}] + \dots \right\} \end{aligned} \quad (3.48)$$

avec

$$\lambda = \frac{\epsilon F_{0,j}^{1-\beta}}{\alpha}$$

Le premier terme de (3.48) correspond à celui de la volatilité implicite ATM, en d'autres termes, celui approximant au plus la volatilité implicite ATM. Le terme  $\ln \kappa / F_{0,j}$  nous donne le niveau de la volatilité implicite par rapport au strike. Il est lui-même amplifié par deux termes,  $-(1-\beta)/2$ , que l'on appelle habituellement le 'beta skew', et  $\rho\lambda/2$ , le 'vanna skew'. La partie associée à  $\ln^2 \frac{\kappa}{F_{0,j}}$  nous donne la convexité de la volatilité implicite. Le premier terme amplifiant son importance est appelé naturellement le beta skew au carré, le second est 'volga' (volatility-gamma). Nous présentons la figure (11) une surface de volatilité simulée à l'aide du modèle, ainsi que sur (3.4.4) l'influence des paramètres  $\beta$  et  $\alpha$  sur le smile de volatilité en fonction du strike.

## Modèle de Taux

---

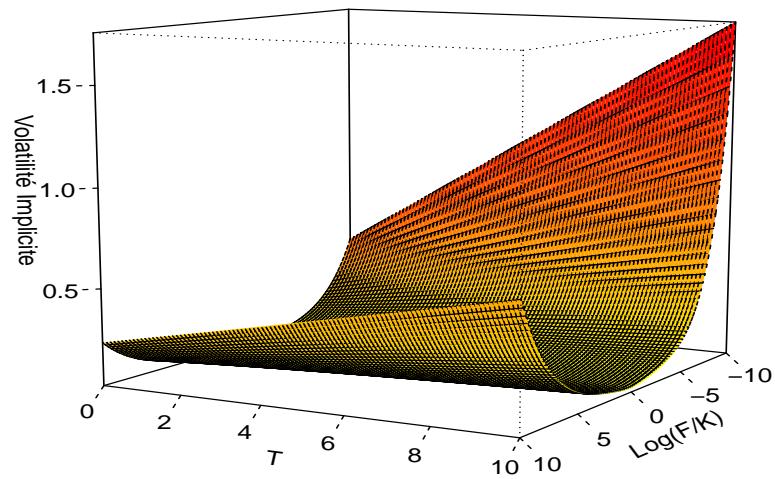


FIGURE 11 – Surface de volatilité simulée à l'aide du modèle SABR.

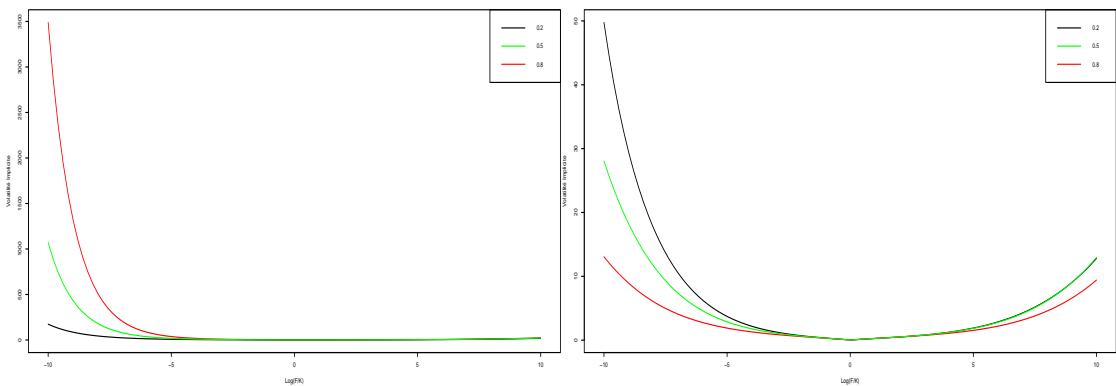


FIGURE 12 – A gauche, influence du paramètre  $\beta$  sur le smile, à droite, influence du paramètre  $\alpha$ .

### 3.5 Dérivées sur volatilité

Il existe des produits dérivés sur des prix d'equity, taux, forex, commodities, inflation, etc., et également, sur la volatilité de ces produits. En effet, il est possible d'acheter de la volatilité sur les marchés.

#### 3.5.1 Swap de variance

Il est possible de faire un pari sur le niveau de la volatilité dans  $X$  jours, ou de se hedger, selon le point de vue. L'exemple basique est le swap de variance. Le payoff d'un tel produit est

$$SV = \frac{NA}{n} \sum_{i=1}^n \left( \log \frac{S_i}{S_{i-1}} \right)^2 - N\sigma_K^2,$$

où  $N$  est le nominal,  $n$  le nombre de jours avant la maturité  $T$ ,  $A = 251$  le nombre de jours dans une année,  $\sigma_K^2$  le strike de variance.

Dans le cas où nous travaillons en temps continu, le swap de variance devient

$$SV = N \left( \frac{1}{T-t} \int_t^T \sigma_s^2 ds - \sigma_K^2 \right)$$

Comme pour les swap de taux vu précédemment où nous échangions un taux fixe contre un taux variable, nous échangeons cette fois une volatilité fixe  $\sigma_T$  contre une volatilité flottante  $\log^2(S_i/S_{i-1})$ . Nous avons deux possibilités pour valoriser ce produit, soit nous calculons simplement la valeur de la volatilité intégrée, soit nous la répliquons par un portefeuille d'option.

Avec le modèle de Heston, la volatilité est donnée par

$$d\sigma_t = -\kappa(\sigma_t - \theta)dt + \eta\sqrt{\sigma_t}dW_t.$$

La variance intégrée est donc

$$\frac{1}{T-t} \mathbb{E} \left[ \int_t^T \sigma_t dt \right] = \frac{1 - e^{-\kappa(T-t)}}{\kappa(T-t)} (\sigma_t - \theta) + \theta$$

pour une preuve de ce résultat on peut regarder les résultats sur le processus CIR vu précédemment.

Plutôt que de travailler sur la volatilité en elle-même, nous pouvons essayer de la répliquer par un portefeuille d'option cela peut par exemple nous permettre de hedger le portefeuille par des techniques classiques. On suppose le modèle général sous  $\mathbb{Q}$

$$\frac{dS_t}{S_t} = rdt + \sigma_t dW_t.$$

En appliquant le lemme d'Itô à  $\ln(S_t)$  nous avons

$$d\ln(S_t) = (r - \frac{1}{2}\sigma_t^2)dt + \sigma_t dW_t$$

et donc

$$\begin{aligned} \frac{1}{2}\sigma_t^2 dt &= rdt + \sigma_t dW_t - d\ln(S_t) \\ \frac{1}{2}\sigma_t^2 dt &= \frac{dS_t}{S_t} - d\ln(S_t) \\ \frac{1}{2} \int_t^T \sigma_s^2 ds &= \int_t^T \frac{dS_s}{S_s} - \int_t^T d\ln(S_s) \\ \frac{1}{2} \int_t^T \sigma_s^2 ds &= \int_t^T \frac{dS_s}{S_s} - \ln \frac{S_T}{S_t}. \end{aligned}$$

La valeur d'un investissement en  $\sigma_t$  est ainsi

$$e^{-r(T-t)} \frac{1}{2} \int_t^T \mathbb{E}(\sigma_s | \mathcal{F}_t) ds = e^{-r(T-t)} \mathbb{E} \left( \int_t^T \frac{dS_s}{S_s} \middle| \mathcal{F}_t \right) - e^{-r(T-t)} \mathbb{E} \left( \frac{S_T}{S_t} \middle| \mathcal{F}_t \right).$$

$W_t$  étant un processus de Wiener, pour le premier terme nous avons

$$e^{-r(T-t)} \mathbb{E} \left( \int_t^T \frac{dS_s}{S_s} \middle| \mathcal{F}_t \right) = e^{-r(T-t)} r(T-t). \quad (3.49)$$

Pour répliquer le second terme, nous avons besoin d'introduire le résultat suivant. Tout payoff  $\phi(S_T)$  avec  $S_T > 0$  peut être décomposé en une somme infinie de call et de put à différent strikes

$$\begin{aligned} \phi(S_T) &= \phi(F) + \phi'(F)(S_T - F) + \int_0^F \phi''(K)(K - S_T)_+ dK \\ &\quad + \int_F^\infty \phi''(K)(S_T - K)_+ dK \end{aligned} \quad (3.50)$$

où  $F = S_t e^{r(T-t)}$  le forward. Pour prouver ce résultat rappelons que

$$\max(x - K, 0)' = \mathbb{1}_{x-K}, \quad \mathbb{1}'_{x-K} = \delta(x - K).$$

Pour toute fonction  $f : \mathbb{R}_+^* \mapsto \mathbb{R}_+$

$$\begin{aligned} f(x) &= \int_0^\infty f(K)\delta(K-x)dK \\ &= \int_0^{x^*} f(K)\frac{\partial^2}{\partial K^2}(K-x)_+dK + \int_{x^*}^\infty f(K)\frac{\partial^2}{\partial K^2}(K-x)_+dK, \end{aligned}$$

où nous avons simplement appliquer le théorème de Chasles pour les intégrales, les deux dérivées secondes sont égales. En faisant deux intégrations par parties avec  $x^* = F$  et  $x = S_T$  on trouve le résultat voulu. En multipliant (3.50) par  $e^{-r(T-t)}$  et en passant sous l'espérance risque neutre on obtient pour toute valorisation d'option de payoff  $\phi(S_T)$

$$\begin{aligned} e^{-r(T-t)}\mathbb{E}(\phi(S_T)|\mathcal{F}_t) &= e^{-r(T-t)}\phi(F) + \int_0^F \phi''(K)\text{Put}(T, K)dK \\ &\quad + \int_F^\infty \phi''(K)\text{Call}(T, K)dK. \end{aligned} \tag{3.51}$$

En appliquant (3.51) au payoff  $\phi(x) = \log\left(\frac{x}{S_t}\right)$  nous avons

$$\begin{aligned} e^{-r(T-t)}\mathbb{E}\left(\ln\left(\frac{S_T}{S_t}\right)|\mathcal{F}_t\right) &= e^{-r(T-t)}r(T-t) - \int_0^F \frac{\text{Put}(T, K)}{K^2}dK \\ &\quad - \int_F^\infty \frac{\text{Call}(T, K)}{K^2}dK. \end{aligned} \tag{3.52}$$

La variance intégrée peut donc s'écrire comme la somme de (3.49) et (3.52), soit

$$\begin{aligned} \frac{e^{-r(T-t)}}{T-t} \int_t^T \mathbb{E}(\sigma_s^2|\mathcal{F}_t)ds &= \frac{2}{T-t} \int_0^F \frac{\text{Put}(T, K)}{K^2}dK \\ &\quad + \frac{2}{T-t} \int_F^\infty \frac{\text{Call}(T, K)}{K^2}dK. \end{aligned} \tag{3.53}$$

### 3.5.2 Option sur volatilité

Pour proposer un framework généraliste d'option ayant pour sous-jacent la volatilité d'un actif nous avons besoin de faire un calcul auparavant. Supposons pour simplifier que le taux sans risque est égal à zéro et que les processus de prix et de volatilité sont indépendants, (3.50) devient alors

$$\mathbb{E}(\phi(S_T)|\mathcal{F}_t) = \phi(F) + \int_0^F \phi''(K)\text{Put}_{BS}(K, T)dK + \int_F^\infty \phi''(K)\text{Call}_{BS}(K, T)dK.$$

Considérons maintenant le cas particulier du payoff  $\phi(S_T) = S_T^p$ , il s'agit d'une option power. Une telle option peut donc être répliquée par

$$\mathbb{E}(S_T^p | \mathcal{F}_t) = F^p + p(p-1) \int_0^F K^{p-2} \text{Put}(K, T) dK + p(p-1) \int_F^\infty K^{p-2} \text{Call}(K, T) dK.$$

Pour simplifier cette expression nous allons faire le changement de variable  $k = \log(K/F)$ , la parité call/put devient

$$\text{Put}(k) = e^{-k} \text{Call}(-k).$$

Ainsi,

$$\begin{aligned} \mathbb{E}(S_T^p | \mathcal{F}_t) &= F^p \left( 1 + p(p-1) \int_{-\infty}^0 e^{pk} \text{Put}(k, T) dk + p(p-1) \int_0^\infty e^{pk} \text{Call}(k, T) dk \right) \\ &= F^p \left( 1 + 2p(p-1) \int_0^\infty k e^{\frac{k}{2}} \cosh\left(p - \frac{1}{2}\right) \text{Call}(k, T) dk \right). \end{aligned} \quad (3.54)$$

Considérons maintenant la diffusion du log-prix générale avec toujours les taux d'intérêts égaux à zéro

$$\begin{aligned} x_T &= \log\left(\frac{S_T}{S_0}\right) \\ &= \int_0^T \sigma_t dW_t - \frac{1}{2} \langle x \rangle_T. \end{aligned}$$

Le processus  $x_T$  est une loi normale de moyenne  $-1/2\langle x \rangle_T$  et de variance  $\langle x \rangle_T$ , sa transformée de Laplace est donc

$$\mathbb{E}(e^{px_T}) = e^{-\frac{1}{2}p\langle x \rangle_T + \frac{1}{2}p^2\langle x \rangle_T}.$$

Ainsi

$$\begin{aligned} \mathbb{E}(e^{px_T}) &= \mathbb{E}\left(e^{(p^2/2 - p/2)\langle x \rangle_T}\right) \\ &= \mathbb{E}\left(e^{\lambda\langle x \rangle_T}\right) \end{aligned}$$

où  $\lambda = p^2/2 - p/2$ . En d'autres termes

$$\mathbb{E}\left(e^{\lambda\langle x \rangle_T}\right) = \mathbb{E}\left(e^{p(\lambda)x_T}\right)$$

avec  $p(\lambda) = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 2\lambda}$ . En utilisant la formule (3.54)

$$\begin{aligned} \mathbb{E}\left(e^{\lambda\langle x \rangle_T}\right) &= \mathbb{E}\left(e^{p(\lambda)x_T}\right) \\ &= \mathbb{E}(S_T^p(\lambda)) \\ &= F^p \left( 1 + 2p(\lambda)(p(\lambda) - 1) \int_0^\infty k e^{\frac{k}{2}} \cosh\left(p(\lambda) - \frac{1}{2}\right) \text{Call}(k, T) dk \right) \end{aligned}$$

ce que l'on peut simplifier en remarquant que  $p(\lambda)(p(\lambda) - 1) = 2\lambda$ . Pour conclure et avoir une formule de valorisation générale nous n'avons plus qu'à remarquer que

$$\mathbb{E}(\langle x \rangle_T) = \frac{\partial}{\partial \lambda} \mathbb{E}(e^{\lambda \langle x \rangle_T}) \Big|_{\lambda=0}.$$

Ainsi,

$$\begin{aligned} \mathbb{E}(\langle x \rangle_T) &= \frac{\partial}{\partial \lambda} \left( 1 + 4\lambda \int_0^\infty k e^{k/2} \cosh \left( p(\lambda) - \frac{1}{2} \right) \text{Call}(k, T) dk \right) \Big|_{\lambda=0} \\ &= 4 \int_0^\infty k e^{k/2} \cosh \left( p(0) - \frac{1}{2} \right) \text{Call}(k, T) dk \\ &= 2 \int_0^\infty (1 + e^k) \text{Call}(k, T) dk \\ &= 2 \left( \int_{-\infty}^0 \text{Put}(k, T) dk + \int_0^\infty \text{Call}(k, T) dk \right). \end{aligned} \tag{3.55}$$

On retrouve ainsi le même résultat que (3.53).

### 3.5.3 VIX

L'indice VIX est la volatilité de l'indice S&P500. Aujourd'hui (les méthodes de calculs ont changé au fil des ans), il est déterminé par

$$\text{VIX}_T^2 = \frac{2e^{rT}}{T} \sum_{i=1}^n \frac{\delta K_i}{K_i^2} Q_i(K_i) - \frac{1}{T} \left( \frac{F}{K_0} - 1 \right)^2 \tag{3.56}$$

où  $Q_i$  est le prix de l'option, call ou put, hors de la monnaie de strike  $K_i$ ,  $F$  est le forward,  $K_0 = \min\{\sup F\}$  et  $\delta K_i$  la différence entre deux strikes consécutifs. A l'origine, le VIX était calculé pour reproduire la volatilité implicite à la monnaie de maturité 1 mois. Il s'agissait d'une moyenne sur 8 options, des calls et des puts proches de l'ATM. Le nouveau calcul prend en compte plus de maturité pour pouvoir représenter la skew, également une moyenne de bandelette d'option pondérée, enfin, initialement, le VIX était calculé sur le S&P 100 et non sur le 500, la vision du marché est donc plus large et consistante.

Il est possible de retrouver l'équation (3.56) à l'aide de la formule (3.53) et donc

(3.55), on suppose toujours le taux sans risque égal à zéro

$$\begin{aligned}
 \frac{\text{VIX}_T^2}{2} &= \int_0^F \frac{\text{Put}(K, T)}{K^2} dK + \int_F^\infty \frac{\text{Call}(K, T)}{K^2} dK \\
 &= \int_0^F \frac{\text{Put}(K, T)}{K^2} dK + \int_F^\infty \frac{\text{Call}(K, T)}{K^2} dK + \int_{F_0}^K \frac{\text{Put}(K, T) - \text{Call}(K, T)}{K^2} dK \\
 &= \int_0^\infty \frac{Q(K, T)}{K^2} dK + \int_{F_0}^F \frac{(K - F)}{k2} dK \\
 &= \int_0^\infty \frac{Q(K, T)}{K^2} dK - \frac{1}{2} \frac{(K_0 - F)^2}{2}.
 \end{aligned}$$

Il ne reste alors plus qu'à discréteriser pour retomber sur (3.56).

Il n'y a pas que le VIX comme indice de volatilité, on peut par exemple penser au Russell 2000 Daily Volatility, Corn Volatility Index, DJIA Volatility et au NVX pour le Nasdaq.

## 3.6 Calcul de Malliavin

Nous faisons ici une introduction au calcul de Malliavin, cela nous permettra entre autre de calculer des grecques pour des options à payoff compliquées. Le calcul de Malliavin est une théorie de calcul infiniment différentiel sur l'espace de Wiener. Il ne s'agit certainement pas de la chose la plus simple, nous ne ferons donc pas toutes les preuves. Pour aider à la compréhension, nous rajoutons des exemples de calculs corrigés. Pour plus de détails sur la partie théorique, nous renvoyons à [Nu06], [Ok97] ou [Wa84].

### 3.6.1 Eléments de théorie

**Définition 3.1.** (*Processus gaussien isonormal*). Soit  $H$  un espace de Hilbert. Un processus stochastique  $W = \{W(h), h \in H\}$  sur l'espace de probabilité  $(\Omega, \mathcal{F}, P)$  est un processus gaussien isonormal si  $W$  est gaussien et vérifie pour tout  $h, g \in H$   $\mathbb{E}(W(h)W(g)) = \langle h, g \rangle_H$ .

Remarquons que la fonction  $h \rightarrow W(h)$  est linéaire, en effet, pour tout  $\lambda, \mu \in \mathbb{R}$

et  $h, g \in H$

$$\begin{aligned}\mathbb{E}[(W(\lambda h + \mu g) - \lambda W(h) - \mu W(g))^2] &= \|\lambda h + \mu g\|_H^2 + \lambda^2 \|h\|_H^2 + \mu^2 \|g\|_H^2 \\ &\quad - 2\lambda \langle \lambda h + \mu g, h \rangle_H - 2\mu \langle \lambda h + \mu g, g \rangle_H \\ &\quad + 2\lambda \mu \langle h, g \rangle_H \\ &= 0 \quad p.s..\end{aligned}$$

Ainsi, la fonction donne une isométrie entre  $H$  est un sous-espace  $\mathcal{H}_1 \subset \mathbb{L}^2(\Omega, \mathcal{F}, P)$  contenant les variables aléatoires isonormales.

Un exemple bien connu de processus gaussien isonormal est l'intégrale de Wiener. Avec  $B_t = (B_t^1, \dots, B_t^d)$  un mouvement brownien  $d$ -dimensionnel sur  $(\Omega, \mathcal{F}, P)$  et pour tout  $h \in H = \mathbb{L}^2(\mathbb{R}_+, \mathbb{R}^d)$ , l'intégrale de Wiener

$$W(h) = \sum_{i=1}^d \int_{\mathbb{R}_+} h_i(s) dB_s^i$$

est un processus gaussien isonormal. Il est clair que  $W(h)$  est gaussien par définition de l'intégrale stochastique et en simplifiant avec  $d = 1$ ,

$$\mathbb{E}(W(h)W(g)) = \int g(s)h(s)ds = \langle g, h \rangle_H.$$

### 3.6.1.1 Chaos de Wiener

Nous allons maintenant montrer l'un des résultat fondamental de cette théorie, la décomposition en chaos de Wiener de toute fonction de  $\mathbb{L}^2$ . Pour ce faire nous avons besoin d'introduire les polynômes d'Hermite. Il s'agit de polynômes orthogonaux pour la mesure  $\mu$  de densité  $\frac{d\mu}{dx} = \frac{e^{-x^2/2}}{\sqrt{\pi}}$ . Pour tout  $x \in \mathbb{R}$  et  $\lambda > 0$

$$\begin{aligned}H_0(x, \lambda) &= 1 \\ H_n(x, \lambda) &= \frac{(-\lambda)^n}{n!} e^{\frac{1}{2\lambda}x^2} \frac{d^n}{dx^n}(e^{-\frac{1}{2\lambda}x^2}), \quad n \geq 1.\end{aligned}$$

Remarquons que nous avons le développement en série entière

$$\exp\left(tx - \frac{t^2\lambda}{2}\right) = \sum_{n=0}^{\infty} t^n H_n(x, \lambda).$$

Pour tout  $n \geq 1$  nous définissons le chaos de Wiener d'ordre  $n$ ,  $\mathcal{H}_n$ , comme la fermeture du sous-espace linéaire de  $\mathcal{L}^2(\Omega, \mathcal{F}, P)$  générée par

$$\{H_n(W(h)), h \in H, \|h\|_H = 1\}$$

avec  $\mathcal{H}_0$  un ensemble de constante. Le lemme suivant nous permet de montrer que les ensembles  $\mathcal{H}_n$  sont orthogonaux entre eux pour tout  $n \neq m$

**Lemme 3.1.** *Soit  $X$  et  $Y$  deux variables gaussiennes centrées réduites, alors pour tout  $n, m \geq 0$*

$$\mathbb{E}(H_n(X)H_m(Y)) = \begin{cases} 0, & n \neq m \\ \frac{1}{n!} (\mathbb{E}(XY))^n, & n = m \end{cases}$$

**Preuve 3.2.** Pour tout  $s, t \in \mathbb{R}$ ,

$$\mathbb{E}\left(\exp\left(sX - \frac{s^2}{2}\right)\exp\left(tY - \frac{t^2}{2}\right)\right) = \exp(st\mathbb{E}(XY)).$$

En prenant les dérivées  $\frac{\partial^{n+m}}{\partial s^n \partial t^m}$  en  $s = t = 0$  de chaque côté de l'équation précédente nous avons

$$\mathbb{E}(n!H_n(X)m!H_m(Y)) = \begin{cases} 0, & n \neq m \\ n! (\mathbb{E}(XY))^n, & n = m \end{cases}$$

Nous pouvons maintenant montrer la décomposition de  $\mathbb{L}^2(\mathcal{G})$  en chaos de Wiener où  $\mathcal{G}$  est la  $\sigma$ -algèbre générée par  $W$ .

**Théorème 3.1.** *L'espace  $\mathbb{L}^2(\mathcal{G})$  peut être décomposé en une somme infinie de sous-espaces orthogonaux*

$$\mathbb{L}^2(\Omega, \mathcal{G}, P) = \bigoplus_{n=0}^{\infty} \mathcal{H}_n.$$

**Preuve 3.3.** Nous avons déjà montré l'orthogonalité des sous-espaces  $\mathcal{H}_n$ , il ne nous reste plus qu'à montrer que si  $X \in \mathbb{L}^2(\Omega, \mathcal{G}, P)$  tel que  $\mathbb{E}(XH_n(W(h))) = 0$  alors  $X = 0$ . Notons déjà qu'il est possible d'exprimer  $x^n$  par une combinaison linéaire de polynômes d'Hermite d'ordres inférieurs ou égaux à  $n$ , ainsi,  $\mathbb{E}(XW^n(h)) = 0$ . Le développement en série entière de la fonction exponentielle nous permet d'écrire  $\mathbb{E}(W \exp(W(h))) = 0$  et la linéarité de  $W$

$$\mathbb{E}\left(X \exp\left(\sum_{i=1}^m t_i W(h_i)\right)\right) = 0$$

avec pour  $m \geq 1$ ,  $t_1, \dots, t_m \in \mathbb{R}$  et  $h_1, \dots, h_m \in H$  tel que  $\|h_i\|_H = 1$ . Cette équation peut se voir comme la transformée de Laplace de la mesure

$$\nu(B) = \mathbb{E}(X \mathbf{1}_B(W(h_1), \dots, W(h_m))), \quad B \in \mathcal{B}(\mathbb{R}^m).$$

Si la transformée de  $\nu$  est zéro, alors la mesure doit être elle-même-égale à zéro pour tout  $G \in \mathcal{G}$ , soit,  $\mathbb{E}(X \mathbf{1}_G) = 0$  et donc  $X = 0$ .

Nous avons un moyen de décomposer l'espace  $\mathbb{L}^2(\mathcal{G})$ . Le prochain but est de déterminer un moyen de décomposer une variable aléatoire en une somme de variables aléatoires orthogonales.

### 3.6.1.2 Intégrale multiple stochastique

Nous allons définir l'intégrale multiple stochastique d'une fonction  $f \in \mathbb{L}^2(T^n)$ ,

$$\int_{T^n} f(t_1, t_2, \dots, t_n) dB(t_1) dB(t_2) \cdots dB(t_n).$$

Nous définissons l'espace  $\mathcal{E}_n$  contenant les fonctions élémentaires sur  $T^n$

$$f(t_1, \dots, t_n) = \sum_{i_1, \dots, i_n=1}^k a_{i_1, \dots, i_n} \mathbf{1}_{[\tau_{i_1}-1, \tau_{i_1}[} \times \cdots \times \mathbf{1}_{[\tau_{i_n}-1, \tau_{i_n}[}(t_1, \dots, t_n)$$

avec  $a = \tau_0 < \tau_1 < \cdots < \tau_k = b$  tel que  $a_{i_1 \dots i_n} = 0$  si  $i_p = i_q$ . Pour tout  $f \in \mathcal{E}_n$  nous définissons

$$I_n(f) = \sum_{i_1, \dots, i_n=1}^k a_{i_1, \dots, i_n} \xi_{i_1} \cdots \xi_{i_n}$$

où  $\mathcal{E}_{i_p} = B(\tau_{i_p}) - B(\tau_{i_p} - 1)$  pour  $p = 1, \dots, n$ . Rappelons que la symétrisation  $\tilde{f}$  de  $f$  est

$$\tilde{f}(t_1, \dots, t_n) = \frac{1}{n!} \sum_{\sigma \in \mathcal{S}_n} f(t_{\sigma(1)}, \dots, t_{\sigma(n)})$$

où  $\mathcal{S}_n$  est l'ensemble des permutations possibles sur  $\{1, \dots, n\}$ . Par exemple

$$\begin{aligned} f(x_1, x_2) &= \cos x_1 - x_2^3 \\ \tilde{f}(x_1, x_2) &= \frac{1}{2}(\cos x_1 + \cos x_2 - x_1^3 - x_2^3). \end{aligned}$$

**Lemme 3.2.** Si  $f \in \mathcal{E}_n$ , alors  $I_n(f) = I_n(\tilde{f})$

**Preuve 3.4.** Par linéarité de  $f$ , nous pouvons montrer le résultat simplement pour la fonction

$$f(t_1, \dots, t_m) = \mathbf{1}_{[t_1^{(1)}, t_1^{(2)}] \times \dots \times [t_n^{(1)}, t_n^{(2)}]}$$

où les intervalles  $[t_i^{(1)}, t_i^{(2)}] \subset T$  sont disjoints. Nous avons

$$\begin{aligned} I_n(f) &= \prod_{i=1}^n \left( B(t_i^{(2)}) - B(t_i^{(1)}) \right) \\ I_n(\tilde{f}) &= \frac{1}{n!} \sum_{\sigma \in \mathcal{S}_n} \prod_{i=1}^n \left( B(t_{\sigma(i)}^{(2)}) - B(t_{\sigma(i)}^{(1)}) \right). \end{aligned}$$

En observant que pour toutes les permutations  $\sigma \in \mathcal{S}_n$ ,  $\prod_{i=1}^n \left( B(t_i^{(2)}) - B(t_i^{(1)}) \right) = \prod_{i=1}^n \left( B(t_{\sigma(i)}^{(2)}) - B(t_{\sigma(i)}^{(1)}) \right)$ , nous obtenons le résultat voulu.

**Lemme 3.3.** Pour tout  $f \in \mathcal{E}_n$  et  $g \in \mathcal{E}_m$ ,  $n, m \geq 1$ , alors  $\mathbb{E}(I_n(f)) = 0$  et

$$\mathbb{E}(I_n(f)I_m(g)) = \begin{cases} 0 & \text{si } n \neq m \\ n! \langle \tilde{f}, \tilde{g} \rangle_{\mathbb{L}^2(T^n)} & \text{si } n = m. \end{cases}$$

**Preuve 3.5.** La preuve pour l'espérance est triviale. Considérons donc  $f \in \mathcal{E}_n$  et  $g \in \mathcal{E}_m$  partageant la même partition  $[t_1^{(1)}, t_1^{(2)}] \times \dots \times [t_n^{(1)}, t_n^{(2)}]$  avec pour coefficients respectifs  $a_{i_1, \dots, i_m}$  et  $b_{i_1, \dots, i_m}$ . Nous avons

$$I_n(f) = n! \sum_{i_1 \leq \dots \leq i_n} a_{i_1 \dots i_n} \xi_{i_1} \dots \xi_{i_n}$$

le même résultat s'obtient pour  $g$ . Ainsi

$$\mathbb{E}(I_n(f)I_m(g)) = n!m! \sum_{i_1 \leq \dots \leq i_n} \sum_{j_1 \leq \dots \leq j_m} a_{i_1 \dots i_n} b_{i_1 \dots i_n} \mathbb{E}(\xi_{i_1} \dots \xi_{i_n} \xi_{j_1} \dots \xi_{j_m})$$

où

$$\mathbb{E}(\xi_{i_1} \dots \xi_{i_n} \xi_{j_1} \dots \xi_{j_m}) = \begin{cases} \prod_{p=1}^n (\tau_{i_p} - \tau_{i_p-1}) & \text{si } n = m, i_1 = j_1, \dots, i_n = j_n. \\ 0 & \text{sinon} \end{cases}$$

Ainsi, si  $n \neq m$ ,  $\mathbb{E}(I_n(f)I_m(g)) = 0$  et si  $n = m$

$$\begin{aligned} \mathbb{E}(I_n(f)I_m(g)) &= (n!)^2 \sum_{i_1 \leq \dots \leq i_n} a_{i_1 \dots i_n} b_{i_1 \dots i_n} \prod_{p=1}^n (\tau_{i_p} - \tau_{i_p-1}) \\ &= n! \int_{T^n} f(t_1, \dots, t_n) g(t_1, \dots, t_n) dt_1 \dots dt_n. \end{aligned}$$

Notons qu'il est possible d'étendre ces deux derniers résultats pour le cas où  $f \in \mathbb{L}^2(T^n)$  et  $g \in \mathbb{L}^2(T^m)$ . Nous avons également le résultat suivant.

**Théorème 3.2.** *Soit  $f \in \mathbb{L}^2(T^n)$ ,  $n \geq 1$ , alors*

$$I_n(f) = n! \int_a^b \cdots \int_a^{t_{n-2}} \left( \int_a^{t_{n-1}} \tilde{f}(t_1, \dots, t_n) dB(t_n) \right) dB(t_{n-1}) \cdots dB(t_1).$$

**Preuve 3.6.** *Considérons le cas simple et suffisant où  $f$  est donné par*

$$f(t_1, \dots, t_n) = \mathbb{1}_{[t_1^{(1)}, t_1^{(2)}] \times \dots \times [t_n^{(1)}, t_n^{(2)}]}(t_1, \dots, t_n)$$

où  $t_n^1 < t_n^2 \leq t_{n-1}^1 \leq \dots \leq t_2^{(2)} \leq t_1^{(1)} < t_1^{(2)}$ . Alors

$$I_n(f) = \prod_{i=1}^n (B(t_i^{(2)}) - B(t_i^{(1)})).$$

D'un autre côté  $\tilde{f} = \frac{1}{n!} f$  si  $t_n < t_{n-1} < \dots < t_1$ , ainsi

$$\int_a^{t_{n-1}} \tilde{f}(t_1, \dots, t_n) dB(t_n) = \frac{1}{n!} \mathbb{1}_{[t_1^{(1)}, t_1^{(2)}] \times \dots \times [t_{n-1}^{(1)}, t_{n-1}^{(2)}]}(t_1, \dots, t_{n-1}) (B(t_n^{(2)}) - B(t_n^{(1)})).$$

Notons qu'il s'agit clairement d'une variable  $\mathcal{F}_{t_{n-1}^{(1)}}$  mesurable. Nous obtenons alors par itérations

$$\int_a^b \cdots \int_a^{t_{n-2}} \left( \int_a^{t_{n-1}} \tilde{f}(t_1, \dots, t_n) dB(t_n) \right) dB(t_{n-1}) \cdots dB(t_1) = \frac{1}{n!} \prod_{i=1}^n (B(t_i^{(2)}) - B(t_i^{(1)})).$$

Il est possible de relier les polynômes d'Hermite avec l'intégrale multiple stochastique.

**Proposition 3.2.** *Soit  $H_m(x)$  le  $m$ -ième polynôme d'Hermite,  $f \in H = \mathbb{L}^2(T)$ , alors*

$$I_n(f^{\oplus n}) = n! H_n(W(f), \|f\|_{\mathbb{L}^2(T)}^2)$$

**Preuve 3.7.** *Nous allons montrer la proposition par récurrence. Le cas  $m = 1$  se déduit du fait que  $\mathcal{E}_1$  est dense dans  $\mathbb{L}^2(T)$ . Supposons à l'ordre  $m$  et montrons à l'ordre  $m + 1$ . Notons  $f^{\oplus m+1} = f(t_1) \cdots f(t_m)$ , par (3.2)*

$$I_{m+1}(h^{\oplus m+1}) = (m+1)! I_m \int_a^b f(t_1) X_{t_1} dB(t_1)$$

avec

$$X_t = \int_a^t \cdots \left( \int_a^{t_n} f(t_2) \cdots f(t_{n+1}) dB(t_{n-1}) \right) \cdots dB(t_2).$$

**Théorème 3.3.** *Toute variable aléatoire  $F \in \mathbb{L}^2(\mathcal{G})$  symétrique peut être décomposée en une série unique d'intégrale multiple stochastique*

$$F = \sum_{n=0}^{\infty} I_n(f_n).$$

Passons à quelques exemples pour clarifier les choses sur la décomposition en chaos de Wiener. Dans tous les cas,  $(W_t)_{t \in [0, T]}$  est un mouvement brownien.

- $F = W(t)$ .

$W$  est un processus gaussien isonormal, il s'agit en fait d'une intégrale stochastique par rapport au mouvement brownien,

$$W(t) = \int_0^T \mathbf{1}_{[0,t]}(s) dW_s.$$

On en déduit directement que  $f_0 = 0$ ,  $f_1 = \mathbf{1}_{[0,t]}$  et  $f_n = 0, n \geq 2$ .

- $F = \int_0^T g(s) dW_s$  où  $g$  est une fonction déterministe.

On a de manière évidente  $f_0 = 0$ ,  $f_1 = g$  et  $f_n = 0, n \geq 2$ .

- $F = W^2(t)$ .

Il y a une petite astuce cette fois, il faut remarquer que

$$\int_0^t W_s dW_s = \frac{1}{2} W^2(t) - \frac{1}{2} t.$$

On a donc

$$\begin{aligned} W^2(t) &= t + 2 \int_0^t \int_0^s dW_u dW_s \\ &= t + \int_0^T \int_0^s \mathbf{1}_{[0,t]}(u) \mathbf{1}_{[0,t]}(s) dW_u dW_s \\ &= t + I_2(f_2). \end{aligned}$$

Ainsi,  $f_0 = t$ ,  $f_1 = 0$ ,  $f_2 = \mathbf{1}_{[0,t]}(u) \mathbf{1}_{[0,t]}(s)$ ,  $f_n = 0, n \geq 3$ .

### 3.6.1.3 Dérivée de Malliavin

Nous passons maintenant à l'un des objets qui va nous être utile pour déterminer les sensibilités d'options. Introduisons l'espace  $C_p^\infty(\mathbb{R}^n)$  des fonctions  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $C^\infty$  et avec des dérivées partielles à croissances polynomiales et l'ensemble  $\mathcal{S}$  des fonctions régulières

$$\mathcal{S} = \left\{ F = f(W(h_1), \dots, W(h_n)), f \in C_p^\infty(\mathbb{R}^n), n \geq 1 \right\}.$$

Nous noterons  $\mathcal{S}_b$  pour l'espace des fonctions  $f \in C_b^\infty(\mathbb{R}^n)$  à dérivées bornées. Nous avons donc  $\mathcal{S}_b \subset \mathcal{S}$ .

**Définition 3.2.** Pour tout  $F \in \mathcal{S}$ , nous définissons la dérivée de  $F$ ,  $DF$  par la variable aléatoire

$$D_t F = \sum_{i=1}^n \partial_i f(W(h_1), \dots, W(h_n)) h_i. \quad (3.57)$$

La dérivée de Malliavin peut également s'écrire comme une dérivée directionnelle

$$\langle D_t F, h \rangle = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} ((f(W(h_1)) + \varepsilon \langle h_1, h \rangle_H, \dots, f(W(h_n)) + \varepsilon \langle h_n, h \rangle_H) - F).$$

Pour bien comprendre les choses prenons comme exemple un mouvement brownien défini sur  $[0, T]$ ,  $(B_t)_{t \in [0, T]}$ , et l'espace de Hilbert  $H = \mathbb{L}^2([0, T], \mathcal{B}([0, T]), \lambda)$  où  $\lambda$  est la mesure de Lebesgue. Pour tout  $h \in H$  posons  $F = W(h(t))$ , alors

$$\begin{aligned} D_t F &= D_t W(h(t)) \\ &= h(t). \end{aligned}$$

Il est sous entendu que nous avons posé  $f(W(h(t))) = W(h(t))$ , soit  $f(x) = x$  qui a pour dérivée  $f'(x) = 1$ . En utilisant (3.57), on obtient le résultat.

L'un des éléments importants lorsque l'on souhaite introduire une théorie de dérivation est la formule d'intégration par partie.

**Lemme 3.4.** Soit  $F \in \mathcal{S}$  et  $h \in H$ , alors

$$\mathbb{E}[\langle D_t F, h \rangle_H] = \mathbb{E}[F W(h)]$$

**Preuve 3.8.** Considérons le cas  $\|h\|_H = 1$  et une base orthonormale de  $H$ ,  $\{e_1, \dots, e_n\}$  tel que  $h = e_1$  et  $F = f(W(e_1), \dots, W(e_n))$  où  $f \in C_p^\infty(\mathbb{R}^n)$ . En notant  $n(x)$  la densité d'une gaussienne multivariée sur  $\mathbb{R}^n$ ,

$$n(x) = (2\pi)^{-n/2} \exp\left(-\frac{1}{2} \sum_{i=1}^n x_i^2\right).$$

Alors,

$$\begin{aligned}\mathbb{E}[\langle D_t F, h \rangle_H] &= \int_{\mathbb{R}^n} \frac{\partial f(x)}{\partial x_1} n(x) dx \\ &= \int_{\mathbb{R}^n} f(x) n(x) x_1 dx \\ &= \mathbb{E}[FW(e_1)] \\ &= \mathbb{E}[FW(h)].\end{aligned}$$

En utilisant la propriété (la règle du produit)

$$D_t(FG) = D_tFG + FD_tG,$$

on en déduit une seconde version de la formule d'intégration par partie,

$$\mathbb{E}[G\langle D_t F, h \rangle_H] = -\mathbb{E}[F\langle D_t G, h \rangle_H] + \mathbb{E}[FGW(h)].$$

En plus de la formule d'intégration par partie, nous avons la formule de dérivation en chaîne. Introduisons d'abord l'espace complet de  $\mathcal{S}$ ,  $\mathcal{D}^{1,p}$  muni de la semi-norme

$$\|F\|_{1,p} = \{\mathbb{E}(|F|^p) + \mathbb{E}(\|D_t F\|_H^p)\}^{1/p}$$

pour tout  $p \geq 1$  et notons  $\mathbb{D}^{1,p}$  sa fermeture.

**Lemme 3.5.** Soit  $g$  une fonction  $\mathcal{C}^1$  de  $\mathbb{R}^d$  dans  $\mathbb{R}$  à dérivées partielles bornées  $F = (F^1, \dots, F^d)$  un vecteur aléatoire tel que  $F^i \in \mathbb{D}^{1,p}$  pour tout  $i$ . Alors  $g(F) \in \mathbb{D}^{1,p}$  et

$$D_t(g(F)) = \sum_{i=1}^d \partial_i g(F) D_t F^i. \quad (3.58)$$

De la même manière que pour la variable aléatoire  $F$ , nous avons une décomposition en chaos de Wiener de la dérivée de Malliavin.

**Proposition 3.3.** Soit  $F$  une variable aléatoire de carré intégrable admettant la décomposition de Wiener  $F = \sum_{n=0}^{\infty} I_n(f_n)$  avec  $f_n \in \mathbb{L}^2(T^n)$ . Alors  $F \in \mathbb{D}^{1,2}$  si et seulement si

$$\sum_{n=1}^{\infty} nn! \|f_n\|_{\mathbb{L}^2(T^n)}^2 < \infty,$$

et dans ce cas nous avons

$$D_t F = \sum_{n=1}^{\infty} n I_{n-1}(f_n(\cdot, t)). \quad (3.59)$$

Passons maintenant à quelques exemples de calcul de dérivées. Dans tous les cas,  $(W_t)_{t \in [0, T]}$  est un mouvement brownien.

- $F = W(t)$

Remarquons que

$$\begin{aligned} W(t) &= \int_0^t dW_s = \int \mathbb{1}_{[0,t]}(s) dW_s \\ &= W(\mathbb{1}_{[0,t]}). \end{aligned}$$

On a donc

$$D_t W(t) = \mathbb{1}_{[0,t]}(s)$$

- $F = \int_0^T s^2 dW_s$

On remarque directement que  $h(t) = t^2$ , donc  $D_t F = t^2$ .

- $F = \int_0^T \int_0^{t_2} \ln(t_1 + t_2) dW_{t_1} dW_{t_2}$ .

Il s'agit d'une intégrale double, on est donc amené à appliquer la formule (3.59). On a  $\int_0^T \int_0^{t_2} \ln(t_1 + t_2) dW_{t_1} dW_{t_2} = I_2(\ln(t_1 + t_2))$ , ainsi,

$$\begin{aligned} D_t F &= 2 \frac{1}{2} I_1(\ln(\cdot + t)) \\ &= \int_0^T \ln(t_1 + t) dW_{t_1}. \end{aligned}$$

- $F = \exp(W(t) + W(s)^2)$ .

$$D_t F = \exp(W(t) + W(s)^2) \mathbb{1}_{[0,t]} + 2W(s) \exp(W(t) + W(s)^2) \mathbb{1}_{[0,s]}.$$

### 3.6.1.4 Intégrale de Skorohod

Nous allons introduire l'opérateur dual de la dérivée de Malliavin, appelé l'opérateur de divergence et dans le cas gaussien, l'intégrale de Skorohod.

L'opérateur de divergence  $\delta$  et donc l'adjoint de  $D$ , vérifiant la relation de dualité pour tout  $F \in \mathbb{D}^{1,2}$

$$\mathbb{E}[F\delta(u)] = \mathbb{E}[\langle DF, u \rangle_H] \tag{3.60}$$

où  $u$  appartient au domaine de  $\delta$ ,  $Dom(\delta)$  tel que

$$Dom(\delta) = \{u \in \mathbb{L}^2(\Omega, H); |\mathbb{E}[\langle DF, u \rangle_H]| \leq c_u \|F\|_{\mathbb{L}^2(\Omega)}\}.$$

Nous avons également une représentation différente de l'opérateur de divergence. En prenant toujours  $F \in \mathbb{D}^{1,2}$  et  $u \in Dom(\delta)$  tel que  $Fu \in \mathbb{L}^2(\Omega, H)$  et  $F\delta(u) - \langle DF, u \rangle_H \in \mathbb{L}^2(\Omega)$  alors  $Fu \in Dom(\delta)$  et

$$\delta(Fu) = F\delta(u) - \langle DF, u \rangle_H. \quad (3.61)$$

La preuve de ce résultat se déduit directement de (3.60). En effet, en considérant une variable  $G \in \mathcal{S}$ ,

$$\begin{aligned} \mathbb{E}[\langle Fu, DG \rangle_H] &= \mathbb{E}[\langle u, D(FG) - GDF \rangle_H] \\ &= \mathbb{E}[(F\delta(u) - \langle u, DF \rangle_H)G]. \end{aligned}$$

Nous avons les propriétés suivantes :

- Si  $u \in Dom(\delta)$ , alors  $\mathbb{E}[\delta(u)] = 0$ .
- $\delta$  est opérateur linéaire et fermé de  $Dom(\delta)$ .
- Si  $u \in Dom(\delta)$ ,

$$\delta(u) = \sum_{j=1}^n F_j W(h_j) - \sum_{j=1}^n \langle DF_j, h_j \rangle_H.$$

Comme dit précédemment, dans le cas gaussien, nous parlons d'intégrale de Skorohod pour l'opérateur de divergence  $\delta(u)$ . Nous allons nous restreindre maintenant à  $W(h) = \int_{[a,b]} h(s) dB_s$  pour tout  $h \in H = \mathbb{L}^2(T)$  avec donc  $T = [a, b]$ .

Remarquons que dans le cas particulier où  $u \in \mathbb{L}^2(T \times \Omega)$  est un processus adapté alors  $u \in Dom(\delta)$  et  $\delta$  coïncide avec l'intégrale d'Itô par rapport au mouvement brownien

$$\delta(u) = \int_a^b u_s dB_s.$$

En effet,  $u$  est un processus élémentaire adapté pouvant s'écrire sous la forme

$$u_t = \sum_{j=1}^n F_j \mathbf{1}_{]t_j, t_{j+1}]}(t),$$

avec  $F_j \in \mathbb{L}^2(\Omega, \mathcal{F}_{t_j})$  pour  $a \leq t_1 < \dots < t_{n+1} \leq b$ . Alors en appliquant la formule de l'opérateur de divergence (3.61), nous obtenons,

$$\begin{aligned}\delta(F_j \mathbf{1}_{[t_j, t_{j+1}]}) &= F_j \delta(\mathbf{1}_{[t_j, t_{j+1}]}) - \int_t D_t F_j \mathbf{1}_{[t_j, t_{j+1}]}(t) dt \\ &= F_j(B_{t_{j+1}} - B(t_j)).\end{aligned}$$

On conclut par la linéarité de  $\delta$  en passant à la limite.

Il est bien entendu possible de lier l'intégrale de Skorohod avec l'intégrale multiple. En supposant que le processus stochastique  $u$  admet la décomposition

$$u(t) = \sum_{n=0}^{\infty} I_n(f_n(\cdot, t)),$$

son intégrale de Skorohod vérifie

$$\delta(u) = \sum_{n=0}^{\infty} I_{n+1}(\tilde{f}_n) \tag{3.62}$$

où  $\tilde{f}$  est la symétrisation de  $f$  comme précédemment.

Pour conclure cette section, donnons la dérivée de l'intégrale de Skorohod. Pour tout  $u \in \mathbb{D}^{1,2}(H)$  tel que  $D_t u_s \in Dom(\delta)$  et  $\delta(D_t u_s) \in \mathbb{L}^2(T \times \Omega)$  alors

$$D_t(\delta(u)) = u_t + \delta(D_t u) \tag{3.63}$$

Encore une fois, faisons quelques exemples.

- $F = W(t)$

Il est clair que  $W(t)$  est  $\mathcal{F}_t$  - adapté, l'intégrale de Skorohod coïncide donc avec l'intégrale d'Itô et

$$\begin{aligned}\delta(W(t)) &= \int_0^T W(t) \delta W(t) \\ &= \int_0^T W(t) dW(t) \\ &= \frac{1}{2} W^2(T) - \frac{1}{2} T.\end{aligned}$$

-  $F = \int_0^T W^2(t_0) \delta W(t).$

Nous avons

$$\int_0^T W^2(t_0) \delta W(t) = \int_0^t (t_0 + I_2(f_2)) \delta W(t)$$

où  $f_2(t_1, t_2, t) = \mathbb{1}_{[0,t_0]}(t_1)\mathbb{1}_{[0,t_0]}(t_2)$ . Il faut maintenant déterminer  $\tilde{f}$  pour pouvoir utiliser (3.62).

$$\begin{aligned} \tilde{f}_2(t_1, t_2, t) &= \frac{1}{3}[f_2(t_1, t_2, t) + f_2(t, t_2, t_1) + f_2(t_1, t, t_2)] \\ &= \frac{1}{3}[\mathbb{1}_{[0,t_0]}t_1\mathbb{1}_{[0,t_0]}t_2 + \mathbb{1}_{[0,t_0]}t\mathbb{1}_{[0,t_0]}t_2 + \mathbb{1}_{[0,t_0]}t_1\mathbb{1}_{[0,t_0]}t] \\ &= \frac{1}{3}[\mathbb{1}_{t_1,t_2 < t_0} + \mathbb{1}_{t,t_2 < t_0} + \mathbb{1}_{t_1,t < t_0}] \\ &= \mathbb{1}_{t,t_1,t_2 < t_0} + \frac{1}{3}\mathbb{1}_{t_1,t_2 < t_0 < t} + \frac{1}{3}\mathbb{1}_{t,t_2 < t_0 < t_1} + \frac{1}{3}\mathbb{1}_{t,t_1 < t_0 < t_2}. \end{aligned}$$

Ainsi,

$$\begin{aligned} \int_0^T W^2(t_0) \delta W(t) &= t_0 W(t) + \int_0^T I_2(f_2) \delta W(t) \\ &= t_0 W(T) + I_3(\tilde{f}_2) \\ &= t_0 W(T) + 6 \int_0^T \int_0^{t_3} \int_0^{t_2} (\mathbb{1}_{[0,t_0]}(t_1, t_2, t_3))^3 dW(t_1) dW(t_2) dW(t_3) \\ &\quad + 6 \int_0^T \int_0^{t_3} \int_0^{t_2} \frac{1}{3} \mathbb{1}_{(t_1,t_2 < t_0 < t_3)} dW(t_1) dW(t_2) dW(t_3) \\ &= t_0 W(T) + t_0^{2/3} h_3 \left( \frac{W(t_0)}{\sqrt{t_0}} \right) + 2 \int_0^T \int_0^{t_3} \int_0^{t_2} dW(t_1) dW(t_2) dW(t_3) \\ &= t_0 W(T) + W^3(t_0) - 3W(t_0) + (W^2(t_0) - t_0)(W(T) - W(t_0)) \\ &= W^2(t_0)W(T) - 2t_0 W(t_0). \end{aligned}$$

### 3.6.2 Application à la finance

Considérons la diffusion générale d'un actif financier  $(S_t)_{t \geq 0}$ ,

$$dS_t = \mu(X_t)dt + \sigma(X_t)dW_t \quad (3.64)$$

où  $(W_t)_{t \geq 0}$  est un mouvement brownien et  $\mu$  et  $\sigma$  sont des fonctions adaptées vérifiant de 'bonnes propriétés'. Par bonnes, nous sous-entendons qu'il n'y a pas que les conditions usuelles à vérifier sur les deux processus mais un peu plus afin de pouvoir déterminer la dérivée de Malliavin de  $S_t$ . Pour tout  $r \leq t$ , nous avons

$$D_t S_t = \sigma(X(r)) + \int_r^t \partial \sigma(X(s)) D_r(X(s)) dW_s + \int_r^t \partial \mu(X(s)) ds. \quad (3.65)$$

Procédons étape par étape pour montrer ce résultat. Réécrivons déjà (3.64)

$$S_t = \int_0^t \mu(X_s)ds + \int_0^t \sigma(X_s)dW_s$$

Pour calculer les dérivées des processus  $\mu$  et  $\sigma$ , nous appliquons la formule de dérivation en chaîne (3.58),

$$\begin{aligned} D_r(\mu(X_s)) &= \partial\mu(X(s))D_r(X(s))\mathbf{1}_{r \leq s} \\ D_r(\sigma(X_s)) &= \partial\sigma(X(s))D_r(X(s))\mathbf{1}_{r \leq s}. \end{aligned}$$

Pour l'intégrale de  $\mu$ , il n'y a pas de problème, par linéarité on a directement

$$D_r \left( \int_0^t \mu(X(s))ds \right) = \int_r^t D_r(\mu(X_s))ds.$$

Pour l'intégrale de  $\sigma$ , nous avons une intégrale d'Itô et donc, de Skorohod, nous appliquons la formule de dérivation de l'intégrale de Skorohod (3.63)

$$D_r \left( \int_0^t \sigma(X(s))dW_s \right) = \sigma(X_s) + \int_r^t D_r(\sigma(X_s))dW_s.$$

En réunissant tous les résultats, on obtient le résultat voulu.

### 3.6.2.1 Calcul des sensibilités

Il s'agit de l'un des principaux points d'application du calcul de Malliavin à la finance quantitative. Son utilisation va nous permettre de calculer 'simplement' les grecques (ou sensibilités) de n'importe quelle option. En fait, nous n'obtenons pas forcément une formule fermée quelque soit la forme du payoff, mais cela permet de minimiser les erreurs d'approximation du Monte Carlo du fait que nous n'avons plus à estimer la dérivée de différentes espérances, l'erreur d'approximation de l'espérance se répercute donc dans l'approximation de la dérivée. Avec le calcul de Malliavin, nous pouvons ôter l'étape de l'approximation de la dérivée, le problème se réduit donc à un problème d'approximation d'espérance. Nous renvoyant à [FoLaLeLiTo99] et [FoLaLeLi01] pour plus de détails que ce que nous allons présenter.

La proposition suivante joue un rôle clé dans ce calcul.

**Proposition 3.4.** Soit  $F$  et  $G$  deux variables aléatoires tel que  $F \in \mathbb{D}^{1,2}$  et  $\langle DF, u \rangle \neq 0$  et  $Gu\langle DF, u \rangle^{-1} \in \text{Dom}(\delta)$ . Alors pour toute fonction  $f$  différentiable à dérivées bornées,

$$\mathbb{E}[f'(F)G] = \mathbb{E}[f(F)H(F, G)], \quad (3.66)$$

où

$$H(F, G) = \delta \left( \frac{Gu}{\langle DF, u \rangle} \right).$$

**Preuve 3.9.** Avec la formule de dérivation en chaîne (3.58), on a

$$\langle Df(F), f \rangle = f'(F)\langle DF, u \rangle,$$

alors en utilisant (3.60)

$$\begin{aligned} \mathbb{E}[f'(F)G] &= \mathbb{E}[\langle Df(F), u \rangle \cdot \langle DF, u \rangle^{-1}G] \\ &= \mathbb{E}[\langle Df(F), u(\langle DF, u \rangle)^{-1} \rangle \cdot G] \\ &= \mathbb{E}[f(F)\delta(Gu\langle DF, u \rangle)^{-1}]. \end{aligned}$$

On peut maintenant passer au calcul des grecques. Pour les sensibilités du premier ordre sur une option  $V$  de payoff  $\phi$ , nous avons en supposant que nous sommes sous la probabilité risque neutre et que le modèle dépend d'un paramètre  $\alpha$

$$\begin{aligned} \frac{\partial V}{\partial \alpha} &= \frac{\partial}{\partial \alpha} \mathbb{E}[\phi(S_T^\alpha)] \\ &= \mathbb{E} \left[ \frac{\partial \phi(S_t^\alpha)}{\partial \alpha} \right] \\ &= \mathbb{E} \left[ \phi'(S_T^\alpha) \frac{\partial S_t^\alpha}{\partial \alpha} \right]. \end{aligned}$$

En appliquant le résultatat (3.66) on peut alors écrire pour tout payoff  $\phi$  dérivable

$$\frac{\partial V}{\partial \alpha} = \mathbb{E} \left[ \phi'(S_T^\alpha) \frac{\partial S_t^\alpha}{\partial \alpha} \right] = \mathbb{E} \left[ \phi(S_T^\alpha) \delta \left( \frac{S_T^\alpha}{\int_0^T D_t S_T dt} \right) \right].$$

Comme nous travaillons sous la probabilité risque neutre, nous avons dans le cadre du modèle de Black and Scholes

$$dS_t = rS_t dt + \sigma S_t dW_t \quad (3.67)$$

où  $r$  est le taux sans risque et  $\sigma$  la volatilité constante. Pour calculer la dérivée de Malliavin de ce processus, nous reprenons le résultat obtenu précédemment (3.65) et en déduisons que

$$D_r S_T = \sigma e^{\int_r^T (r - \frac{1}{2}\sigma^2) ds + \int_r^T \sigma dW_s}$$

et que

$$D_t S_T = \sigma S_T.$$

En effet,

$$\begin{aligned} D_t S_T &= D_t \left( S_0 \exp \left( \left( r - \frac{\sigma^2}{2} \right) T \right) + \sigma W_T \right) \\ &= S_0 \exp \left( r - \frac{\sigma^2}{2} \right) T D_t (\exp(\sigma W_T)) \\ &= S_0 \exp \left( r - \frac{\sigma^2}{2} \right) T \exp(\sigma W_T) D_t \int_0^T \sigma dW_u \\ &= \sigma S_T \mathbf{1}_{t < T}. \end{aligned}$$

Choisissons maintenant le cas particulier du delta  $\Delta$ , c'est-à-dire la dérivée du produit par rapport au sous-jacent à l'instant  $t$ . Pour simplifier on va supposer que  $t = 0$ , la solution de (3.67) est

$$S_T = S_0 e^{\int_0^T (r - \frac{1}{2}\sigma^2) ds + \int_0^T \sigma dW_s}.$$

Ainsi,

$$\begin{aligned} \Delta &= \frac{\partial}{\partial S_0} \mathbb{E} [\phi(S_T^{S_0})] = e^{-rT} \mathbb{E} \left[ \frac{\partial \phi(S_T^{S_0})}{\partial S_0} \right] \\ &= e^{-rT} \mathbb{E} \left[ \phi'(S_T^{S_0}) \frac{\partial S_T}{\partial S_0} \right] = \frac{e^{-rT}}{S_0} \mathbb{E} [\phi'(S_T^{S_0}) S_T] \\ &= \frac{e^{-rT}}{S_0} \mathbb{E} \left[ \phi(S_T^{S_0}) H \left( S_T, \frac{\partial S_T}{\partial S_0} \right) \right] \\ &= \frac{e^{-rT}}{S_0} \mathbb{E} \left[ \phi(S_T^{S_0}) \delta \left( \frac{\frac{\partial S_T}{\partial S_0}}{\int_r^T D_r S_T dt} \right) \right] \\ &= \frac{e^{-rT}}{S_0} \mathbb{E} \left[ \phi(S_T^{S_0}) \delta \left( \frac{S_T}{\sigma S_T} \right) \right] \\ &= \frac{e^{-rT}}{\sigma T S_0} \mathbb{E} [\phi(S_T^{S_0}) W_T] \end{aligned}$$

où sur la dernière égalité nous avons utilisé le fait que l'intégrale de Skorohod porte sur un processus adapté, il s'agit donc d'une intégrale d'Itô. Maintenant on se pose la question du comment faire pour les dérivées secondes, ou troisièmes si on veut plus de raffinement. Le gamma, la dérivée seconde d'une option, est l'une des sensibilités les plus connues et utilisées. En dehors de se couvrir par rapport au gamma comme nous l'avons expliqué dans la section sur les grecques, elle permet aussi d'adapter sa stratégie de delta-hedging, la plus fondamentale en pratique. En effet, on ne peut pas se delta hedger correctement en temps continu, le delta bouge sans cesse, avoir un delta nul tout le temps est illusoire, c'est possible d'un point de vue pratique mais les coûts de transactions pour maintenir cette position impliqueraient une stratégie perdant de l'argent. Le gamma peut être compris comme la réponse à de grande variation du sous-jacent, on peut donc en déduire grâce au gamma les moments auxquels il est plus 'important' de rebalancer son portefeuille pour être delta-neutre.

Pour le calcul de gamma,

$$\begin{aligned}\Gamma &= e^{-rT} \frac{\partial^2}{\partial S_0^2} \mathbb{E} [\phi(S_T^{S_0})] \\ &= \frac{e^{-rT}}{S_0^2} \mathbb{E}[\phi''(S_T^{S_0}) S_T^2].\end{aligned}$$

On applique une nouvelle fois le résultat (3.66) avec  $G = S_T^2$ ,  $F = S_T$  et  $u = 1$ , l'astuce est simplement de poser  $\phi'' = (\phi')'$ , l'intégrale de Skorohod est

$$\begin{aligned}\delta \left( \frac{S_T^2}{\int_0^T D_t S_T dt} \right) &= \delta \left( \frac{S_T}{\sigma T} \right) \\ &= S_T \left( \frac{W_T}{\sigma T} \right)\end{aligned}$$

ainsi,

$$\mathbb{E}[\phi''(S_T^{S_0}) S_T^2] = \mathbb{E}[\phi'(S_T^{S_0}) S_T \left( \frac{W_T}{\sigma T} - 1 \right)].$$

On reprend encore une fois le résultat (3.66), avec  $G = S_T \left( \frac{W_T}{\sigma T} - 1 \right)$ ,  $F = S_T$  et  $u = 1$ , l'intégrale de Skorohod est

$$\begin{aligned}\delta \left( \frac{S_T \left( \frac{W_T}{\sigma T} - 1 \right)}{\int_0^T D_t S_T dt} \right) &= \delta \left( \frac{W_T}{\sigma^2 T^2} - \frac{1}{\sigma T} \right) \\ &= \frac{W_T^2}{\sigma^2 T^2} - \frac{1}{\sigma^2 T} - \frac{W_T}{\sigma T},\end{aligned}$$

et donc,

$$\Gamma = \frac{e^{-rT}}{S_0^2 \sigma T} \mathbb{E} \left[ \phi(S_T) \left( \frac{W_T^2}{\sigma T} - \frac{1}{\sigma} - W_T \right) \right].$$

Pour le vega, nous avons

$$\begin{aligned} \vartheta &= \frac{\partial}{\partial \sigma} e^{-rT} \mathbb{E}[\phi(S_T^\sigma)] \\ &= e^{-rT} \mathbb{E} \left[ \phi(S_T) \left( \frac{W_T^2}{\sigma T} - \frac{1}{\sigma} - W_T \right) \right]. \end{aligned}$$

Le rho,

$$\begin{aligned} \rho &= \frac{\partial}{\partial r} e^{-rT} \mathbb{E}[\phi(S_T^r)] \\ &= e^{-rT} \mathbb{E} \left[ \phi(S_T) \left( \frac{W_T}{\sigma} - T \right) \right] \end{aligned}$$

### 3.6.2.2 Sensibilité pour option asiatique

Dans la section précédente nous avons considéré uniquement des payoff opérant sur la valeur terminale de l'actif, cela ne couvre pas tous les cas de figure. Prenons par exemple les options asiatiques, il s'agit d'option ne prenant en compte la moyenne du sous-jacent au court de la durée de vie de l'option. Le payoff est du type pour toute fonction  $\phi$

$$V = e^{-rT} \mathbb{E} \left[ \phi \left( \frac{1}{T} \int_0^T S_t dt \right) \right].$$

Dans ce cas, le delta d'une option est

$$\Delta = \frac{2e^{-rT}}{\sigma^2 S_0} \mathbb{E} \left[ \phi \left( \frac{1}{T} \int_0^T S_t dt \right) \left( \frac{S_T - S_0}{\int_0^T S_t dt} - (r - \sigma^2/2) \right) \right].$$

La preuve est purement calculatoire. Nous assumons comme toujours que  $\phi$  est une fonction continue différentiable à dérivées bornées. Nous avons,

$$\begin{aligned} \Delta &= \frac{\partial}{\partial S_0} e^{-rT} \mathbb{E} \left[ \phi \left( \frac{1}{T} \int_0^T S_t dt \right) \right] \\ &= e^{-rT} \mathbb{E} \left[ \phi' \left( \frac{1}{T} \int_0^T S_t dt \right) \frac{1}{T} \int_0^T \frac{\partial S_t}{\partial S_0} dt \right] = \frac{e^{-rT}}{S_0} \mathbb{E} \left[ \phi' \left( \frac{1}{T} \int_0^T S_t dt \right) \frac{1}{T} \int_0^T S_t dt \right]. \end{aligned}$$

En posant  $F = \frac{1}{T} \int_0^T S_s ds$ ,  $G = \frac{1}{T} \int_0^T S_s ds$ , nous avons,

$$\Delta = \frac{e^{-rT}}{S_0} \mathbb{E} \left[ \phi \left( \frac{1}{T} \int_0^T S_t dt \right) \delta \left( \frac{\frac{1}{T} S_t \int_0^T S_s ds}{\int_0^T S_u D_u \left( \frac{1}{T} \int_0^T S_s ds du \right)} \right) \right].$$

Alors,

$$D_u \left( \frac{1}{T} \int_0^T S_s ds \right) = \frac{\sigma}{T} \int_u^T S_s ds$$

et

$$\begin{aligned} \frac{\sigma}{T} \int_0^T S_u \left( \int_u^T S_s ds \right) du &= \frac{\sigma}{T} \int_0^T S_s \left( \int_0^s S_u du \right) ds \\ &= \frac{\sigma}{T} \int_0^T \frac{1}{2} \left( \int_0^s S_u du \right)^2 ds \\ &= \frac{\sigma}{T} \frac{1}{2} \left( \int_0^T S_u du \right)^2. \end{aligned}$$

On peut maintenant calculer l'intégrale de Skorohod,

$$\begin{aligned} \delta \left( \frac{\frac{1}{T} S_t \int_0^T S_s ds}{\int_0^T S_u D_u \left( \frac{1}{T} \int_0^T S_s ds du \right)} \right) &= \delta \left( \frac{\int_0^T S_u d.S_t}{\frac{\sigma}{2} \left( \int_0^T S_u du \right)^2} \right) \\ &= \frac{2}{\sigma} \delta \left( \frac{S_t}{\int_0^T S_u du} \right) \\ &= \frac{2}{\sigma} \left( \frac{\int_0^T S_t dW_t}{\int_0^T S_t dt} + \frac{\sigma}{2} \right) \\ &= \frac{2}{\sigma} \left( \frac{S_T - S_0}{\sigma \int_0^T S_t dt} - \frac{r}{\sigma} + \frac{\sigma}{2} \right) \\ &= \frac{2(S_T - S_0)}{\sigma^2 \int_0^T S_t dt} - \frac{2r}{\sigma^2} + 1. \end{aligned}$$

Ainsi,

$$\Delta = \frac{e^{-rT}}{\sigma^2 S_0} \mathbb{E} \left[ \phi \left( \frac{1}{T} \int_0^T S_t dt \right) \left( \frac{S_T - S_0}{\sigma^2 \int_0^T S_t dt} - r + \frac{\sigma^2}{2} \right) \right].$$

Il est également possible de trouver des résultats pour les options dépendant du maximum ou du minimum, comme les options barrières où lookback, voir e.g. [GoKo03] et [BeDoKo03]. Egalement, pour des modèles à diffusion plus complexe comme à volatilité stochastique, [AlEw07] ou [DeGaRi05] entre autre.

### 3.6.2.3 Raffinement du hedging

## 4 Introduction au risque de crédit

Le risque de défaut est de plus en plus pris en compte dans les modèles de pricing. Quand un dérivé sur un produit OTC (over-the-counter) est pricé, nous devrions naturellement prendre en compte le risque de défaut de contrepartie. A l'instant où l'entité en face du détenteur du produit fait défaut, le détenteur de l'actif financier ne recevra pas la totalité de ses paiements comme initialement prévu, mais une fraction. La problématique de cette partie va donc consister à introduire ce risque de défaut dans nos modèles de valorisation. Nous allons considérer des obligations pour lesquelles il est possible d'avoir un défaut de l'émetteur au temps aléatoire  $\tau$ .

### 4.1 Processus ponctuel

Soit  $\tau$  la durée de vie aléatoire et  $\mathbb{P}(\tau \geq t), t > 0$  la probabilité de survivre au moins jusqu'à l'instant  $t$ . Le premier point à introduire est la notion de processus ponctuel.

**Définition 4.1** (Processus Ponctuel). *Soit  $t$  le temps physique et  $\{t_i\}_{i \in \{1, 2, \dots\}}$  une suite croissante aléatoire de temps d'arrivée,  $0 \leq t_i \leq t_{i+1}, \forall i$ . Alors la séquence est un processus ponctuel sur le segment  $[0, \infty)$ .*

**Définition 4.2** (Processus de Comptage). *Soit  $(N_t)_{t \geq 0}$  un processus stochastique à valeurs réelles tel que  $N_t = \sum_{i \geq 1} \mathbf{1}_{\{t_i \leq t\}}$ . On dit que  $N$  est un processus de comptage si p.s.,  $N(0) = 0$ ,  $N$  est continu à droite et  $N$  est croissant.*

L'exemple certainement le plus connu est le processus de Poisson, qui en plus des propriétés énoncées vérifie,

- pour tout  $t, s \geq 0$ ,  $N_{t+s} - N_t$  suit une loi de Poisson de paramètre  $\lambda s$  (stationnarité)
- pour tout  $t, s \geq 0$ , les variables aléatoires  $N_{t+s} - N_s$  indépendant de la sigma-algèbre  $\sigma(N_u, u \leq s)$ , (accroissement indépendant)

La probabilité conditionnelle que l'émetteur ne fasse pas défaut jusqu'à l'instant  $T$  sachant qu'il n'est pas encore en défaut en  $t$  est

$$\begin{aligned}\mathbb{P}(\tau > T | \tau > t) &= \frac{\mathbb{P}(\tau > T \text{ et } \tau > t)}{\mathbb{P}(\tau > t)} \\ &= \frac{\mathbb{P}(\tau > T)}{\mathbb{P}(\tau > t)}, \quad 0 \leq t \leq T.\end{aligned}$$

Nous définissons maintenant l'intensité conditionnelle par

$$\lambda(t|\mathcal{F}_t) = \frac{\mathbb{P}(\tau < t + dt | \tau > t)}{dt}$$

La  $\mathcal{F}_t$ -intensité caractérise l'évolution du processus  $N_t$  conditionnellement à sa filtration naturelle. On peut l'interpréter comme la probabilité conditionnelle de faire défaut au prochain instant. Nous avons également

$$\begin{aligned}\lambda(t|\mathcal{F}_t) &= \frac{1}{\mathbb{P}(\tau > t)} \frac{\mathbb{P}(\tau > T \text{ et } \tau > t)}{dt} \\ &= \frac{1}{\mathbb{P}(\tau > t)} \frac{\mathbb{P}(\tau > t) - \mathbb{P}(\tau > t + dt)}{dt} \\ &= -\frac{d}{dt} \ln \mathbb{P}(\tau > t).\end{aligned}$$

Ainsi, nous obtenons une équation différentielle de condition initiale  $\mathbb{P}(\tau > 0) = 1$ . La solution est

$$\mathbb{P}(\tau > t) = e^{-\int_0^t \lambda_u du}, \quad t \in \mathbb{R}_+,$$

où par abus de notation nous avons noté  $\lambda_s$  pour  $\lambda(s|\mathcal{F}_s)$  avec  $\mathcal{F}_s$  la filtration naturelle engendrée par  $\lambda$ . Nous en déduisons directement

$$\frac{\mathbb{P}(\tau > T)}{\mathbb{P}(\tau > t)} = e^{-\int_t^T \lambda_u du}, \quad 0 \leq t \leq T. \quad (4.1)$$

Posons maintenant  $T = t + \delta t$ , il vient, pour  $dt$  tendant vers 0,

$$\begin{aligned}\mathbb{P}(\tau > t + dt | \tau > t) &= e^{-\lambda_t dt} \\ &\approx 1 - \lambda_t dt.\end{aligned}$$

La probabilité de défaut sur  $[t, t + \delta t]$  est donc

$$\begin{aligned}\mathbb{P}(\tau < t + dt | \tau > t) &= 1 - e^{-\lambda_t dt} \\ &\approx \lambda_t dt.\end{aligned}$$

## 4.2 Obligation en cas de défaut

Nous supposerons maintenant que  $(\lambda_t)_{t \in \mathbb{R}_+}$  est un processus aléatoire réel adapté à une certaine filtration. Le prix d'une obligation de maturité  $T$  et d'instant de défaut  $\tau$  est donné à l'instant  $t$  par

$$\begin{aligned} B(t, T) = & \mathbb{E} \left[ \mathbf{1}_{\{\tau > T\}} \exp \left( - \int_t^T r_s ds \right) \middle| \mathcal{G}_t \right] \\ & + \mathbb{E} \left[ \text{Rec} \mathbf{1}_{\{\tau \leq T\}} \exp \left( - \int_t^T r_s ds \right) \middle| \mathcal{G}_t \right], \quad 0 \leq t \leq T, \end{aligned}$$

avec  $\text{Rec} \in [0, 1]$  le taux de recouvrement et  $\mathcal{G}_t = \mathcal{F}_t \vee \sigma\{\{\tau < u\}, 0 \leq u \leq t\}$  la plus petite sigma-algèbre contenant l'historique des taux ainsi que la connaissance de la probabilité de défaut ou non de la contrepartie. Nous allons calculer le premier terme. Soit  $F$  une variable aléatoire adaptée à  $\mathcal{F}_T$ , nous avons

$$\begin{aligned} \mathbb{E}[F \mathbf{1}_{\{\tau > T\}} | \mathcal{G}_t] &= \mathbb{E}[F \mathbf{1}_{\{\tau > T\}} \mathbf{1}_{\{\tau > t\}} | \mathcal{G}_t] \\ &= \mathbf{1}_{\{\tau > t\}} \mathbb{E}[F \mathbf{1}_{\{\tau > T\}} | \mathcal{F}_t \vee \{\tau > t\}] \\ &= \mathbf{1}_{\{\tau > t\}} \frac{\mathbb{E}[F \mathbf{1}_{\{\tau > T\}} | \mathcal{F}_t]}{\mathbb{P}(\tau > t | \mathcal{F}_t)} \\ &= \mathbf{1}_{\{\tau > t\}} \frac{\mathbb{E}[\mathbb{E}[F \mathbf{1}_{\{\tau > T\}} | \mathcal{F}_T] | \mathcal{F}_t]}{\mathbb{P}(\tau > t | \mathcal{F}_t)}, \quad \mathcal{F}_t \subset \mathcal{F}_T \\ &= \mathbf{1}_{\{\tau > t\}} \frac{\mathbb{E}[F \mathbb{E}[\mathbf{1}_{\{\tau > T\}} | \mathcal{F}_T] | \mathcal{F}_t]}{\mathbb{P}(\tau > t | \mathcal{F}_t)} \\ &= \mathbf{1}_{\{\tau > t\}} \frac{\mathbb{E}[F \mathbb{P}(\tau > T | \mathcal{F}_T) | \mathcal{F}_t]}{\mathbb{P}(\tau > t | \mathcal{F}_t)} \\ &= \mathbf{1}_{\{\tau > t\}} \mathbb{E} \left[ F \frac{\mathbb{P}(\tau > T | \mathcal{F}_T)}{\mathbb{P}(\tau > t | \mathcal{F}_t)} \middle| \mathcal{F}_t \right]. \end{aligned}$$

Le choix  $F = \exp \left( - \int_t^T r_s ds \right)$  avec (4.1) nous permet de conclure

$$\mathbb{E} \left[ \mathbf{1}_{\{\tau > T\}} \exp \left( - \int_t^T r_s ds \right) \middle| \mathcal{G}_t \right] = \mathbf{1}_{\{\tau > T\}} \mathbb{E} \left[ \exp \left( - \int_t^T (r_s + \lambda_s) ds \right) \middle| \mathcal{F}_t \right],$$

ainsi,

$$\begin{aligned} B(t, T) = & \mathbf{1}_{\{\tau > T\}} \mathbb{E} \left[ \exp \left( - \int_t^T (r_s + \lambda_s) ds \right) \middle| \mathcal{F}_t \right] \\ & + \mathbb{E} \left[ \text{Rec} \mathbf{1}_{\{\tau \leq T\}} \exp \left( - \int_t^T r_s ds \right) \middle| \mathcal{G}_t \right], \quad 0 \leq t \leq T. \end{aligned}$$

Dans le cas où  $\text{Rec} = 0$ , c'est-à-dire qu'il n'y a pas de recouvrement possible, nous avons,

$$B(t, T) = \mathbb{1}_{\{\tau > T\}} \mathbb{E} \left[ \exp \left( - \int_t^T (r_s + \lambda_s) ds \right) \middle| \mathcal{F}_t \right], \quad 0 \leq t \leq T. \quad (4.2)$$

L'équation (4.2) est intuitivement correcte. En effet, plus l'intensité  $\lambda$ , c'est-à-dire la probabilité d'avoir un défaut, est importante, plus le prix de l'obligation va baisser.

Nous allons présenter le cas où le taux comme l'intensité sont diffusés par un processus d'Ornstein Ulhenbeck, il s'agit du modèle de Vasicek.

**Proposition 4.1.** Soit  $B(t, T)$  une obligation de maturité  $T$  avec  $(r_t)_{t \in \mathbb{R}_+}$  et  $(\lambda_t)_{t \in \mathbb{R}_+}$  modélisés par des processus d'Ornstein Ulhenbeck

$$\begin{cases} dr_t = -ar_t dt + \sigma dW_t^1 \\ d\lambda_t = -b\lambda_t dt + \eta dW_t^2, \end{cases}$$

avec  $(W_t^1)_{t \in \mathbb{R}_+}$  et  $(W_t^2)_{t \in \mathbb{R}_+}$  deux mouvements browniens standards tel que

$$\langle dW_t^1 dW_t^2 \rangle = \rho dt,$$

alors,

$$\begin{aligned} B(t, T) &= \mathbb{1}_{\{\tau > t\}} \exp(-C(a, t, T)r_t - C(b, t, T)\lambda_t) \\ &\quad \cdot \exp \left( \frac{\sigma^2}{2} \int_t^T C^2(a, s, T) ds + \frac{\eta^2}{2} \int_t^T C^2(b, s, T) e^{-b(T-s)} ds \right) \\ &\quad \cdot \exp \left( \rho \sigma \eta \int_t^T C(a, s, T) C(b, s, T) ds \right), \end{aligned}$$

avec  $C(\kappa, t, T) = \mathbb{E}[\int_t^T r_s ds | \mathcal{F}_t] = -\frac{1}{\kappa}(e^{-\kappa(T-t)} - 1)$ .

**Preuve 4.1.** Nous allons procéder comme pour le pricing par martingale des zéros-coupons. Calculons les quantités  $\int_t^T r_s ds$  et  $\int_t^T \lambda_s ds$ .

$$\begin{aligned} \int_0^t r_s ds &= \frac{1}{a} (\sigma W_t^1 - r_t) \\ &= \frac{\sigma}{a} \left( W_t^1 - \int_0^t e^{-a(t-s)} dW_s^1 \right) \\ &= \frac{\sigma}{a} \int_0^t (1 - e^{-a(t-s)}) dW_s^1. \end{aligned}$$

Ainsi,

$$\begin{aligned}
 \int_t^T r_s ds &= \int_0^T r_s ds - \int_0^t r_s ds \\
 &= \frac{\sigma}{a} \int_0^T (1 - e^{-a(T-s)}) dW_s^1 - \frac{\sigma}{a} \int_0^t (1 - e^{-a(t-s)}) dW_s^1 \\
 &= -\frac{\sigma}{a} \left( \int_0^t (e^{-a(T-s)} - e^{-a(t-s)}) dW_s^1 + \int_t^T (e^{-a(T-s)} - 1) dW_s^1 \right) \\
 &= -\frac{\sigma}{a} (e^{-a(T-t)} - 1) \int_0^t e^{-a(t-s)} dW_s^1 - \frac{\sigma}{a} \int_t^T (e^{-a(T-s)} - 1) dW_s^1 \\
 &= -\frac{1}{a} (e^{-a(T-t)} - 1) r_t - \frac{\sigma}{a} \int_t^T (e^{-a(T-s)} - 1) dW_s^1
 \end{aligned}$$

Nous avons le même résultat pour  $\lambda_t$  naturellement. Il est clair que nous avons

$$\mathbb{E} \left[ \int_t^T r_s ds + \int_t^T \lambda_s ds \mid \mathcal{F}_t \right] = C(a, t, T) r_t + C(b, t, T) \lambda_t,$$

pour  $C(\kappa, t, T)$  une fonction dépendant de  $t, T$  et du drift  $\kappa$ . La variance est quant à elle donnée par

$$\begin{aligned}
 &\mathbb{V}ar \left( \int_t^T r_s ds + \int_t^T \lambda_s ds \mid \mathcal{F}_t \right) \\
 &= \mathbb{V}ar \left( \int_t^T r_s ds \mid \mathcal{F}_t \right) + \mathbb{V}ar \left( \int_t^T \lambda_s ds \mid \mathcal{F}_t \right) + 2\mathbb{C}ov \left( \int_t^T r_s ds, \int_t^T \lambda_s ds \mid \mathcal{F}_t \right) \\
 &= \frac{\sigma^2}{a^2} \int_t^T (e^{-a(T-s)} - 1)^2 ds + 2\rho \frac{\sigma\eta}{ab} \int_t^T (e^{-a(T-s)-1}) (e^{-b(T-s)} - 1) ds \\
 &\quad + \frac{\eta^2}{b^2} \int_t^T (e^{-b(T-s)} - 1)^2 ds \\
 &= \sigma^2 \int_t^T C^2(a, s, T) ds + 2\rho\sigma\eta \int_t^T C(a, s, T) C(b, s, T) ds + \eta^2 \int_t^T C^2(b, s, T) ds.
 \end{aligned}$$

Il est clair que la quantité  $\int_t^T r_s + \int_t^T \lambda_s ds$  est une gaussienne, fort de nos résultats

*précédents, nous avons donc*

$$\begin{aligned}
 B(t, T) &= \mathbb{1}_{\{\tau>t\}} \mathbb{E} \left[ \exp \left( - \int_t^T r_s ds - \int_t^T \lambda_s \right) \middle| \mathcal{F}_t \right] \\
 &\quad \cdot \exp \left[ \frac{1}{2} \mathbb{V}ar \left( - \int_t^T r_s ds - \int_t^T \lambda_s \right) \middle| \mathcal{F}_t \right] \\
 &= \mathbb{1}_{\{\tau>t\}} \exp(-C(a, t, T)r_t - C(b, t, T)\lambda_t) \\
 &\quad \cdot \exp \left( \frac{\sigma^2}{2} \int_t^T C^2(a, s, T)ds + \frac{\eta^2}{2} \int_t^T C^2(b, s, T)e^{-b(T-s)}ds \right) \\
 &\quad \cdot \exp \left( \rho\sigma\eta \int_t^T C(a, s, T)C(b, s, T)ds \right).
 \end{aligned}$$

### 4.3 Modèle de firme

Il s'agit maintenant de poser le problème légèrement différemment, mais finalement les hypothèses restent les mêmes. Nous considérons une firme avec une unique dette de payoff terminal promit  $L$ , en d'autre terme, il s'agit d'un zéro-coupon de maturité  $T$  et de valeur faciale  $L > 0$ . La capacité à rembourser la dette est déterminée pour la valeur totale des actifs de la firme en  $T$ ,  $V_T$ . Nous supposerons que le défaut peut se produire en  $T$  uniquement, i.e.  $\{V_T < L\}$ . Le temps d'arrêt associé est donc

$$\tau = T\mathbb{1}_{V_T < L} + \infty\mathbb{1}_{V_T \geq L}$$

La valeur de recouvrement est donc

$$Rec_T = V_T\mathbb{1}_{\{V_T < L\}} + L\mathbb{1}_{\{V_T \geq L\}}$$

Le payoff d'un tel produit est donc

$$g(T) = \min(V_T, L) = L - \max(L - V_T, 0) = L - (L - V_T)_+.$$

Le dernier terme nous montre que dans ce framework, le problème est équivalent au pricing d'un put écrit sur la valeur de l'entreprise, de strike  $L$  et de maturité  $T$ . Notons maintenant  $D(t, T)$  le prix à l'instant  $t < T$  de l'obligation

$$D(t, T) = LB(t, T) - P(t), \tag{4.3}$$

où  $P$  est le prix d'un put de strike  $L$  et de maturité  $T$ . (4.3) peut également être vu que la valeur de la dette de l'entreprise.

### 4.3.1 Modèle de Merton

Dans le modèle de Merton [Me74], nous supposons que la valeur de la firme est diffusée sous la probabilité risque neutre  $\mathbb{Q}$  par

$$dV_t = (r - \kappa)V_t dt + \sigma dW_t.$$

Le pricing de l'obligation se ramenant au pricing d'un put (4.3), nous avons directement,

$$D(t, T) = V_t e^{(\kappa(T-t))} \phi(d_+) + LB(t, T) \phi(d_-), \quad 0 \leq t < T,$$

où

$$d_{\pm} = \frac{(\kappa - r) \pm \frac{1}{2}\sigma^2(T-t) - \ln(V_t/L)}{\sigma\sqrt{T-t}}.$$

### 4.3.2 Modèle de Black et Cox

Par construction, le modèle de Merton suppose uniquement que le défaut peut intervenir à maturité, or, ce n'est pas le cas. Le modèle de Black et Cox [BlCo76] est un prolongement du modèle de Merton. Le modèle prend en compte plusieurs points : l'engagement du vendeur, la subordination de la dette et les restrictions sur la vente d'actifs. La diffusion de la valeur de la firme reste inchangée

$$dV_t = V_t((r - \kappa)dt + \sigma dW_t).$$

La clause d'engagement permet aux détenteurs de la dette de forcer l'entreprise à déposer le bilan ou se réorganiser en fonction du standard établi. Ce 'standard' est défini dans le modèle de Black et Cox par une barrière  $\bar{v}_t = K e^{-\gamma(T-t)}$  pour  $K$  une constante positive. Tant que la valeur de la firme est au dessus de cette barrière, la firme continue de plein de droit, le cas contraire, les actionnaires prennent le relais et le défaut a lieu. Notons plus généralement

$$v_t = \begin{cases} \bar{v}_t & \text{pour } t < T \\ L & \text{pour } t = T. \end{cases}$$

L'instant de défaut est donc donné par

$$\tau = \inf\{t \in [0, T], V_t \leq v_t\}.$$

Nous avons trois cas, le premier correspondant à l'émetteur ne fait pas défaut et nous touchons  $L$ , le second, l'émetteur passe sous la barrière et nous touchons

$\beta_2 V_\tau$ , enfin, l'émetteur ne peut pas recouvrir sa dette en  $T$  et nous touchons  $\beta_1 V_T$ , avec  $\beta_1, \beta_2$  deux constantes dans  $[0, 1]$ . L'équation de valorisation prend alors la forme

$$\begin{aligned} D(t, T) = & \mathbb{E}(L e^{-r(T-t)} \mathbf{1}_{\{\tau>T, V_T \geq L\}} | \mathcal{F}_t) \\ & + \mathbb{E}(\beta_1 V(T) e^{-r(T-t)} \mathbf{1}_{\{\tau \leq T, V_T < L\}} | \mathcal{F}_t) \\ & + \mathbb{E}(K \beta_2 e^{-\gamma(T-\tau)} e^{-r(\tau-t)} \mathbf{1}_{\{t < \tau < T\}} | \mathcal{F}_t). \end{aligned}$$

Dans le cas particulier où  $\beta_1 = \beta_2 = 1$ , les auteurs de [BlCo76] montrent que  $D(t, T) = I_1(t, T) + I_2(t, T)$  avec

$$\begin{aligned} I_1(t, T) &= LB(t, T)[\phi(d_1) - V_\tau^{2a} \phi(d_2)] \\ I_2(t, T) &= V_t[V_t^{\theta+\zeta} \phi(d_3) - V_t^{\theta-\zeta} \phi(d_4)], \end{aligned}$$

où

$$\begin{aligned} I_1(t, T) &= \frac{\ln(V_t/L) + \nu(T-t)}{\sigma \sqrt{T-t}} \\ I_2(t, T) &= \frac{\ln \bar{v}_t^2 - \ln(Lv_t) + \nu(T-t)}{\sigma \sqrt{T-t}} \\ I_3(t, T) &= \frac{\ln(\bar{v}_t/V_t) + \zeta \sigma^2(T-t)}{\sigma \sqrt{T-t}} \\ I_4(t, T) &= \frac{\ln(\bar{v}_t/V_t) - \zeta \sigma^2(T-t)}{\sigma \sqrt{T-t}}, \end{aligned}$$

et

$$\begin{aligned} \nu &= r - \kappa - \frac{1}{2} \sigma^2 \\ m &= r - \kappa - \gamma - \frac{1}{2} \sigma^2 \\ \zeta &= \frac{1}{\sigma^2} \sqrt{m^2 + 2\sigma^2(r - \gamma)} \\ \theta &= \frac{m}{\sigma^2} + 1 \end{aligned}$$

#### 4.4 Credit Default Swap

Un Crédit Défaut Swap (CDS) est un produit visant à couvrir les risques de crédit. Il s'agit d'un instrument fondamental sur le marché des dérivés de crédit. Les CDS que nous considérerons sont single name, i.e. qu'ils sont écrits sur un seul sous-jacent et, permettent donc de se couvrir du risque de défaut de celui

ci, de manière dynamique. A l'instar des notes publiées par les agences de rating, les CDS jouent également un rôle d'indicateur sur la probabilité de défaut de la firme ou de l'Etat émettant de la dette. L'acheteur du produit de couverture va payer régulièrement un montant fixé à l'avance jusqu'à l'instant défaut, celui ci n'arrivant pas, il payera jusqu'à maturité de produit. Le cas contraire, i.e. l'instant défaut arrivant avant  $T$ , c'est l'émetteur du contrat de couverture que va payer selon des critères définis lors de la signature du contrat, selon l'état de défaut de l'entité, il s'agit de la partie flottante.

Il se compose de deux jambes. La première (premium leg) correspond à l'acheteur de la protection à l'instant  $t$  contre un défaut possible à l'instant  $T_{i+1}$ . L'agent paye un montant fixe  $S_t$  aux instants  $T_{i+1}, \dots, T_n = T$ . La valeur actualisée de la première jambe est donc

$$\begin{aligned} V(t, T) &= \mathbb{E} \left[ \sum_{k=i}^{n-1} e^{-\int_t^{T_{k+1}} r_s ds} S_t (T_{k+1} - T_k) \mathbf{1}_{\{\tau > T_{k+1}\}} \middle| \mathcal{G}_t \right] \\ &= \sum_{k=i}^{n-1} S_t (T_{k+1} - T_k) \mathbb{E} \left[ e^{-\int_t^{T_{k+1}} r_s ds} \mathbf{1}_{\{\tau > T_{k+1}\}} \middle| \mathcal{G}_t \right] \\ &= S_t \sum_{k=i}^{n-1} (T_{k+1} - T_k) B(t, T_{k+1}) \\ &= S_t B(t, T_i, T_n) \end{aligned}$$

où

$$B(t, T_k) = \mathbf{1}_{\{\tau > t\}} \mathbb{E} \left[ e^{-\int_t^{T_k} (r_s + \lambda_s) ds} \middle| \mathcal{F}_t \right], \quad 0 \leq t \leq T$$

est le prix d'une obligation pouvant faire défaut.

Le seconde jambe est appelée la jambe de protection, il s'agit donc de la partie incomptant au vendeur du produit. Le vendeur du contrat paye  $1 - \text{Rec}_{i+1}$  en cas de défaut à l'instant  $T_{i+1}$ . La valeur de cette partie est

$$\mathbb{E} \left[ \sum_{k=i}^{n-1} e^{-\int_t^{T_{k+1}} r_s ds} \mathbf{1}_{(T_k, T_{k+1}]}(\tau) (1 - \text{Rec}_{i+1}) \middle| \mathcal{G}_t \right]$$

où  $\text{Rec}_{k+1}$  est le taux de recouvrement en  $T_{k+1}$ . Dans le cas où ce taux n'est pas aléatoire mais déterminé à l'avance et constant, la valeur de cette protection devient

$$(1 - \text{Rec}) \mathbb{E} \left[ \sum_{k=i}^{n-1} e^{-\int_t^{T_{k+1}} r_s ds} \mathbf{1}_{(T_k, T_{k+1}]}(\tau) \middle| \mathcal{G}_t \right].$$

L'inconnue principale de ce problème est quel montant  $S_t$  payé par l'acheteur est 'juste'. En égalisant les deux jambes, nous avons

$$S_t = \frac{1}{B(t, T_i, T_n)} \mathbb{E} \left[ \sum_{k=i}^{n-1} e^{-\int_t^{T_{k+1}} r_s ds} \mathbb{1}_{(T_k, T_{k+1}]}(\tau) (1 - \text{Rec}_{i+1}) \middle| \mathcal{G}_t \right].$$

Si  $\tau$  prend une valeur dans le tenor  $\{T_1, \dots, T_n\}$  et que le taux de recouvrement est déterministe et constant, la valeur à l'équilibre se réduit à

$$S_t = \frac{1 - \text{Rec}}{B(t, T_i, T_n)} \mathbb{E} \left[ \sum_{k=i}^{n-1} e^{-\int_t^{T_{k+1}} r_s ds} \mathbb{1}_{[t, T]}(\tau) \middle| \mathcal{G}_t \right].$$

## 4.5 Credit Valuation Adjustment

Un CVA (Credit Valuation Adjustment) correspond à la différence entre la valorisation d'un produit pour lequel on suppose qu'il n'y a pas de risque de défaut de la contrepartie contre la valorisation dans le cas d'un défaut possible, la vraie valeur. En notant  $V$  la valeur actuelle d'un produit financier

$$\text{CVA}(t) = V_{\text{risk-free}}(t) - V(t). \quad (4.4)$$

Si  $V$  est une dérivée de maturité  $T$  et de payoff  $\phi$  elle est donnée théoriquement par

$$V(t, T) = \mathbb{E}_{\mathbb{Q}} \left[ \sum_{u=t}^T \phi(u) D(t, u) \middle| \mathcal{F}_t \right]$$

où  $D$  est le facteur d'actualisation et  $\mathbb{Q}$  la probabilité risque-neutre. Il existe plusieurs types de CVA, si nous considérons que c'est la contrepartie qui fait défaut, on parle bien de CVA. S'il s'agit de l'investisseur qui fait défaut, qu'il n'est pas capable de rembourser la contrepartie, on parle de DVA pour Debit Valuation Adjustment. Les CVA et DVA supposent donc qu'un seul défaut est possible avant la date de livraison  $T$ , on parle de CVA unilatéral. Si le défaut peut se produire des deux cotés, on parle de CVA bilatéral.

Nous nous plaçons dans le cas d'un défaut possible de la contrepartie à l'instant  $\tau_c \leq T$ , nous allons déterminer la forme générale d'un CVA unilatéral. Commençons par remarquer que la valeur du produit peut s'écrire

$$\begin{aligned} V^D(t, T) &= V(t, T) \mathbb{1}_{\tau_c > T} \\ &\quad + V(t, \tau_c) \mathbb{1}_{t < \tau_c \leq T} + \mathbb{1}_{t < \tau_c \leq T} [D(t, \tau_c) (\text{Rec}(V(\tau_c, T))_+ + (V(\tau_c, T))_-)]. \end{aligned}$$

Le premier terme correspond à ce que l'on touche si le défaut n'a pas lieu. Le second à la somme que l'on touche avant que le défaut se produise. Le troisième à ce que l'on touche une fois que le défaut est intervenu, Rec est le taux de recouvrement, soit la contrepartie à encore des liquidités pour rembourser une proportion Rec, soit rien du tout. Enfin, le dernier terme avec le min correspond à ce que va devoir payer l'investisseur si la valeur est négative. Si cela peut aider le lecteur remarquons que  $(x)_- = -(-x)_+$ . On passe maintenant sous l'espérance risque neutre où l'on note pour alléger  $\mathbb{E} = \mathbb{E}_{\mathbb{Q}}$

$$\begin{aligned}
 V^D(t, T) &= \mathbb{E} \left\{ V(t, T) \mathbb{1}_{\tau_c > T} + V(t, \tau_c) \mathbb{1}_{t < \tau_c \leq T} \right. \\
 &\quad \left. + \mathbb{1}_{t < \tau_c \leq T} [D(t, \tau_c) (\text{Rec}(V(\tau_c, T))_+ + (V(\tau_c, T))_-)] \middle| \mathcal{G}_t \right\} \\
 &= \mathbb{E} \left\{ V(t, T) \mathbb{1}_{\tau_c > T} + V(t, \tau_c) \mathbb{1}_{t < \tau_c \leq T} \right. \\
 &\quad \left. + \mathbb{1}_{t < \tau_c \leq T} [D(t, \tau_c) (-(\text{Rec}(V(\tau_c, T))_+ + V(\tau_c, T)))] \middle| \mathcal{G}_t \right\} \quad (4.5) \\
 &= \mathbb{E} \left\{ V(t, T) (\mathbb{1}_{\tau_c > T} + \mathbb{1}_{t < \tau_c \leq T}) \right. \\
 &\quad \left. - \mathbb{1}_{t < \tau_c \leq T} D(t, \tau_c) (\text{Rec}(V(\tau_c, T))_+) \middle| \mathcal{G}_t \right\} \\
 &= \mathbb{E}[V(t, T) | \mathcal{F}_t] - \mathbb{E} \left[ \mathbb{1}_{t < \tau_c \leq T} D(t, \tau_c) (\text{Rec}(V(\tau_c, T))_+) \middle| \mathcal{G}_t \right].
 \end{aligned}$$

En comparant les équations (4.4) et (4.5), on en déduit que le CVA unilatéral est

$$\text{UCVA}(t) = \mathbb{E} \left[ \mathbb{1}_{t < \tau_c \leq T} D(t, \tau_c) (\text{Rec}(V(\tau_c, T))_+) \middle| \mathcal{G}_t \right].$$

Il s'agit donc de l'espérance du Loss Given Default (LGD = 1-Rec) de l'exposition actualisée en cas de défaut.

Pour le cas d'un DVA unilatéral, nous notons  $\tau_I$  l'instant de défaut de l'investisseur et déduisons des résultats précédents sur le UCVA que

$$\begin{aligned}
 V^D(t, T) &= V(t, T) \mathbb{1}_{\tau_I > T} \\
 &\quad + V(t, \tau_I) \mathbb{1}_{t < \tau_I \leq T} + \mathbb{1}_{t < \tau_I \leq T} [D(t, \tau_I) ((V(\tau_c, T))_+ + \text{Rec}(V(\tau_I, T))_-)]
 \end{aligned}$$

et donc

$$\text{UDVA}(t) = \mathbb{E} \left[ \mathbf{1}_{t < \tau_I \leq T} D(t, \tau_I) (1 - \text{Rec})(V(\tau_c, T))_- \middle| \mathcal{G}_t \right].$$

Passons maintenant au cas le plus intéressant, le plus réaliste finalement, celui où tout le monde peut faire défaut. Nos avons trois cas possibles :

- La contrepartie et l'investisseur ne font pas défaut avant la maturité.
- La contrepartie fait défaut à l'instant  $\tau_c$ , avant l'investisseur.
- L'investisseur fait défaut à l'instant  $\tau_I$ , avant la contrepartie.

En se servant des résultats sur les CVA et DVA, nous pouvons écrire pour chacun de ces trois cas les correspondances

- 
- $V_{\tau_c > T \wedge \tau_I > T} = V(t, T).$
- 
- $V_{\tau_c \leq T \wedge \tau_c < \tau_I} = V(t, \tau_c) + D(t, \tau_c)(\text{Rec}_c(V(\tau_c, T))_+ + (V(\tau_c, T))_-).$
- 
- $V_{\tau_I \leq T \wedge \tau_I < \tau_c} = V(t, \tau_I) + D(t, \tau_I)((V(\tau_I, T))_+ + \text{Rec}_I(V(\tau_I, T))_-).$

En passant sous l'espérance risque neutre

$$\begin{aligned}
 V^D(t, T) &= \mathbb{E} \left\{ V(t, T) \mathbb{1}_{\tau_c > T \wedge \tau_I > T} \right. \\
 &\quad + \mathbb{1}_{\tau_c \leq T \wedge \tau_c < \tau_I} [V(t, \tau_c) + D(t, \tau_c)(\text{Rec}_c(V(\tau_c, T))_+ + (V(\tau_c, T))_-)] \\
 &\quad \left. + \mathbb{1}_{\tau_I \leq T \wedge \tau_I < \tau_c} [V(t, \tau_I) + D(t, \tau_I)((V(\tau_I, T))_+ + \text{Rec}_I(V(\tau_I, T))_-)] \middle| \mathcal{G}_t \right\} \\
 &= \mathbb{E} \left\{ \mathbb{1}_{\tau_c > T \wedge \tau_I > T} V(t, T) \right. \\
 &\quad + \mathbb{1}_{\tau_c \leq T \wedge \tau_c < \tau_I} [V(t, \tau_c) + (V(\tau_c, T) - (1 - \text{Rec}_c)(V(\tau_c, T))_+ D(t, \tau_c))] \\
 &\quad \left. + \mathbb{1}_{\tau_I \leq T \wedge \tau_I < \tau_c} [V(t, \tau_I) + (V(\tau_I, T) - (1 - \text{Rec}_I)(V(\tau_I, T))_- D(t, \tau_I))] \middle| \mathcal{G}_t \right\} \\
 &= \mathbb{E} \left\{ (\mathbb{1}_{\tau_c > T \wedge \tau_I > T} + \mathbb{1}_{\tau_c \leq T \wedge \tau_c < \tau_I} + \mathbb{1}_{\tau_I \leq T \wedge \tau_I < \tau_c}) V(t, T) \right. \\
 &\quad - \mathbb{1}_{\tau_c \leq T \wedge \tau_c < \tau_I} (1 - \text{Rec}_c)(V(\tau_c, T))_+ D(t, \tau_c) \\
 &\quad \left. - \mathbb{1}_{\tau_I \leq T \wedge \tau_I < \tau_c} (1 - \text{Rec}_I)(V(\tau_I, T))_- D(t, \tau_I) \middle| \mathcal{G}_t \right\} \\
 &= \mathbb{E}(V(t, T) | \mathcal{F}_t) \\
 &\quad - \mathbb{E} [\text{LGD}_c \mathbb{1}_{\tau_c \leq T \wedge \tau_c < \tau_I} (V(\tau_c, T))_+ D(t, \tau_c) | \mathcal{G}_t] \\
 &\quad - \mathbb{E} [\text{LGD}_I \mathbb{1}_{\tau_I \leq T \wedge \tau_I < \tau_c} (V(\tau_I, T))_- D(t, \tau_I) | \mathcal{G}_t].
 \end{aligned}$$

On en déduit qu'un CVA bilatéral est

$$\begin{aligned}
 \text{BCVA}(t) &= \mathbb{E} [\text{LGD}_c \mathbb{1}_{\tau_c \leq T \wedge \tau_c < \tau_I} (V(\tau_c, T))_+ D(t, \tau_c) | \mathcal{G}_t] \\
 &\quad - \mathbb{E} [\text{LGD}_I \mathbb{1}_{\tau_I \leq T \wedge \tau_I < \tau_c} (V(\tau_I, T))_- D(t, \tau_I) | \mathcal{G}_t].
 \end{aligned}$$

## 5 Mesure de risque

Avant de rentrer dans le vif du sujet, nous présentons la notion de mesure de risque monétaire. Il s'agit d'une fonction  $\rho$  vérifiant les deux propriétés suivantes pour tout  $X$  et  $Y$  deux v.a.r., la valeur de deux portefeuilles

- Monotone : Si  $X \leq Y$ , alors  $\rho(X) \geq \rho(Y)$ .
- Invariant : Si  $m \in \mathbb{R}$ , alors  $\rho(X + m) = \rho(X) - m$ .

La monotonie implique que le risque décroît à mesure que le payoff du portefeuille croît. Pour expliquer la propriété d'invariance supposons que  $m$  est le montant correspondant à un investissement sur un taux sans risque, alors le risque du portefeuille  $X + m$  est réduit de  $m$ .

Ce n'est pas tout à fait suffisant pour avoir une 'bonne' mesure de risque, ce que l'on appelle une mesure cohérente du risque, il nous faut rajouter deux propriétés

- Convexe : Pour tout  $0 \leq \lambda \leq 1$

$$\rho(\lambda X + (1 - \lambda)Y) \leq \rho(X) + (1 - \lambda)\rho(Y).$$

- Homogène : Pour tout  $\lambda \geq 0$ ,  $\rho(\lambda X) = \lambda\rho(X)$ .

La propriété de convexité se comprend comme une minimisation du risque par diversification. L'homogénéité est simplement qu'en valeur, le portefeuille présente même les risques.

## 5.1 Value At Risk

Quelle est la perte maximale de mon portefeuille ? C'est en résumé ce à quoi la VaR, pour Value at Risk, tente de répondre. Il s'agit d'une mesure de risque, la plus connue. Concrètement on détermine le niveau de perte d'un actif ou d'un portefeuille pour un certain niveau de probabilité pré-fixé sur une période bien définie. Par exemple, si la Var sur un portefeuille est de 1M\$ sur une semaine à un niveau de confiance de 99%, alors, nous avons une chance sur cent d'avoir une telle perte la semaine à venir. Formellement, en notant  $\alpha$  le risque, où  $1 - \alpha$  le niveau de confiance, la Var d'horizon  $T$  sur l'actif/portefeuille de rendements  $r_T$  est (les deux équations sont équivalentes)

$$\begin{aligned}\mathbb{P}(r_T \leq VaR(\alpha, T)) &= \alpha \\ \mathbb{P}(VaR(\alpha, T) \leq r_T) &= 1 - \alpha.\end{aligned}$$

Il a trois moyens de déterminer la VaR, historiquement, par Monte-Carlo ou paramétriquement.

- VaR historique. Pour déterminer la Var par cette méthode, il suffit de prendre en compte l'historique des rendements  $r_t$ . Supposons par exemple que nous cherchons la VaR à 99% sur  $T = 1000$  jours, nous n'avons qu'à trouver le

quantile correspondant. En triant les données,  $r_{(1)} \leq r_{(2)} \leq \dots \leq r_{(1000)}$ , nous avons  $VaR(0.01, 1000) = r_{(10)}$ .

C'est une méthode non paramétrique, c'est-à-dire que nous n'avons fait aucune hypothèse de distribution sur les rendements et que nous prenons en compte toutes les caractéristiques réelles de la série financière prise en compte. Néanmoins, pour calculer la VaR historique, nous ne prenons qu'un échantillon, la statistique n'est pas forcément représentative de ce qu'il va se passer dans le futur. Prendre un historique 'grand' pour avoir plus d'information et donc avoir l'estimateur convergeant vers la valeur théorique n'est pas forcément une bonne idée, en effet, les variations très lointaines (par exemple d'il y a plus de 30ans) ne sont pas forcément significatives. Pour améliorer ce moyen de calculer la VaR nous pouvons affecter des poids aux observations, plus les données sont récentes, plus on leur donne de l'importance dans le calcul. Egalement, nous pouvons tirer des échantillons bootstrap.

- VaR paramétrique. Avec la méthode précédente, nous n'avions qu'une statistique issue d'un échantillon, qui ne converge pas forcément vers la valeur théorique. Nous allons donc supposer que les données suivent une loi bien particulière. Supposons que la variable  $r_T$  soit de moyenne  $\mu T$  et de variance  $\sigma^2 T$ , alors

$$\begin{aligned} \mathbb{P}(r_T \leq VaR(\alpha, T)) &= \mathbb{P}\left(\frac{r_T - \mu T}{\sigma \sqrt{T}} \leq \frac{VaR(\alpha, T) - \mu T}{\sigma \sqrt{T}}\right) \\ &= \alpha. \end{aligned}$$

Il n'y a plus qu'à inverser la probabilité

$$VaR(\alpha, T) = \mu T + z_\alpha \sigma \sqrt{T}. \quad (5.1)$$

où  $z_\alpha = F^{-1}(\alpha)$  avec  $F$  la fonction de répartition de  $r_T$ . On pourra regarder dans des tables statistiques la valeur de  $F^{-1}$ . Notons que la VaR est principalement fonction de la volatilité puisque les rendements sont en moyenne proches de zéro. De la on peut en déduire une sorte de formule de récurrence en temps, en effet, d'après (5.1)

$$\begin{aligned} VaR(\alpha, T_1) &\approx z_\alpha \sigma \sqrt{T_1} \\ VaR(\alpha, T_2) &\approx z_\alpha \sigma \sqrt{T_2}. \end{aligned}$$

De la on en déduit

$$VaR(\alpha, T_2) = \frac{\sqrt{T_2}}{\sqrt{T_1}} VaR(\alpha, T_1).$$

La connaissance de la VaR à 1 jour nous permet donc de connaître la Var sur 10 jours par exemple. Cela peut simplifier les calculs.

Par le même principe que (5.1), on en déduit le résultat pour un portefeuille de  $N$  actifs. En fait, on peut avoir un résultat aussi simple seulement pour le cas gaussien multivarié de matrice de covariance  $\Sigma$ , c'est pour cela que par abus de langage, on parle de méthode variance/covariance, vrai uniquement dans le cas gaussien, malheureusement trop utilisé. Dans le cas de variables corrélées, il n'existe pas forcément de distribution jointe comme pour le cas gaussien et nous avons besoin de sortir l'artillerie des copules, cela sera vu dans un autre cours.

Cette méthode présente deux risques d'erreur. La première est de considérer une distribution non réaliste, par exemple une distribution gaussienne. La VaR dans ce cas sera sous-estimée car la loi normale ne présente pas de queue épaisse. Le second risque porte sur l'estimateur de la variance/covariance. Comme nous l'avons vu, la VaR est principalement décrite par la variance de la série financière, non-constante, une mauvaise estimation de celle-ci induira une VaR non cohérente.

- VaR Monte-Carlo. Pour résumer, il s'agit d'un mix des deux méthodes précédentes. La VaR historique se confronte au problème des données disponibles et de leurs significativités. La VaR paramétrique suppose quant à elle de connaître la distribution des données et surtout de pouvoir inverser la distribution pour faire le calcul. Si l'on suppose que les données sont gaussiennes, pas de problème, mais cela n'est pas réaliste avec les faits stylisés. Si l'on suppose une distribution plus complexe, nous ne connaissons pas forcément la densité, par exemple avec le modèle de marche aléatoire multifractal (MRW). Supposons que notre choix de distribution porte sur la MRW, pour estimer la VaR dans ce cas, la seule possibilité est de générer  $n$  réalisations, et de déterminer la VaR pour chacune d'entre elles, le résultat étant la moyenne

$$VaR(\alpha, T) = \frac{1}{n} \sum_{i=1}^n Var^i(\alpha, T)$$

où  $Var^i$  correspond à la Var de la  $i - \text{ème}$  réalisation d'une trajectoire. Le point noir de cette méthode est la complexité de mise en oeuvre, le choix de la distribution, l'estimation des paramètres correspondants et, le coût de calcul. Enfin, le choix de la distribution est basé sur des données passées, il n'est pas improbable que la distribution ainsi que les paramètres changent dans le futur. Ce dernier point est valable pour les deux autres méthodes.

La VaR n'est pas une mesure de cohérente, elle ne vérifie pas l'hypothèse de convexité. La diversification des positions ne va pas forcément diminuer le risque. Comme nous l'avons déjà dit, la VaR ne donne aucune information sur le niveau de la perte, seulement la probabilité d'occurrence. Il existe d'autre mesure de risque, plus ou moins intéressante, la plus connue est l'expected shortfall qui prend en compte ces deux derniers points.

## 5.2 Expected Shortfall

L'expected shortfall, également appelé conditional value at risk est une mesure de risque cohérente, elle vérifie les propriétés suffisantes. Notons  $ES(\alpha, T)$  l'expected shortfall de niveau  $\alpha$  sur  $T$ , alors

$$ES(\alpha, T) = \frac{1}{\alpha} \int_0^\alpha Var(x, T) dx.$$

Il s'agit donc d'une moyenne de VaR pour différent niveau  $\alpha$ . Dans le cas où la variable aléatoire considérée est continue, nous pouvons montrer que

$$ES(\alpha, T) = \mathbb{E}(r_T | r_T \geq Var(\alpha, T)).$$

Pour être complet sur cette partie il faudrait introduire des notions de décision dans l'incertain, d'intégrale de Choquet par exemple. Dans tous les cas, cela reste très théorique, il faut backtester les résultats et que le niveau  $\alpha$  est bien souvent imposé, entre autre, par votre futur management.

## 5.3 Backtesting

Comme pour une stratégie de trading, nous avons besoin de backtester notre mesure de risque, est-ce qu'elle est correcte ou non. L'idée est de déterminer notre mesure on line, pour un certain niveau de risque  $\alpha$  est-ce que nous avons bien les dépassements prévus. Graphiquement, si nous traçons les rendements avec les

intervalles des VaR par exemple, il s'agit de voir si les rendements dépassent, comme prévu, ces intervalles. Si nous avons choisi de prendre une distribution gaussienne, nous allons sous-estimer les dépassements. Si nous avons choisi une loi trop 'extrême', nous allons sur-estimer les dépassements, mais finalement, cela n'est pas forcément plus mal.

Remarquons tout de même un gros inconvénient de ce type de mesure. Supposons que tout cela fonctionne bien, nous avons la probabilité exacte de passer sous une certaine perte, mais, de combien va être la perte, cela est une autre question.

## 5.4 Drawdown

Le drawdown représente la perte de richesse par rapport au plus haut niveau de richesse atteint par la stratégie. Si nous notons  $(V_t)_{t \geq 0}$  la valeur du portefeuille, le drawdown est défini par

$$D_t = \max_{0 \leq s \leq t} V_s - V_t.$$

Nous représentons graphiquement cette quantité sur (13).

Il faudrait des drawdowns les plus faibles possible, cela reviendrait à dire que la stratégie ne cesse d'empocher des gains. Un drawdown positif est donc équivalent à une perte et le maximum drawdown peut être vu comme une mesure de risque. Pour pouvoir modéliser cette quantité et donc de pouvoir déterminer des probabilités de perte, il faut dans un premier temps trouver la distribution. Pour ce faire, nous pouvons appliquer ce que l'on appelle un test d'adéquation, il en existe énormément. Nous pouvons différencier les tests paramétriques des tests non paramétriques. La différence est que dans le premier cas nous supposons une distribution, par exemple si nous supposons que les rendements sont gaussiens, nous pouvons utiliser le test de Jarque Bera, Lilliefors ou Shapiro-Wilk entre autres. Nous présentons ici des tests non paramétriques, c'est-à-dire que nous ne supposons aucune distribution en particulier. L'un des plus connu est le test de Kolmogorov-Smirnov. Il s'appuie sur la distance  $D$  définie par,

$$D_n = \sqrt{n} \max_x |F_n(x) - F(x)|$$

où  $F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{x_i \leq x}$  est la fonction de répartition empirique de la série  $(x_1, \dots, x_n)$  et  $F$  est la fonction de répartition à laquelle on souhaite comparer notre échantillon. Sous l'hypothèse  $H_0 : F_n = F$ , la statistique  $D_n$  converge vers la loi de Kolmogorov-Smirnov,

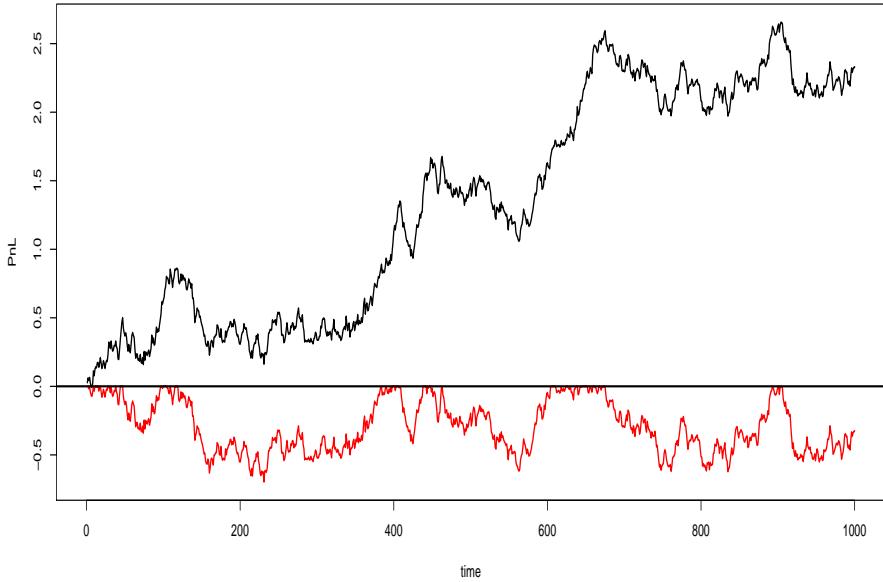


FIGURE 13 – En noir la valeur du portefeuille, en rouge les drawdowns (négatif) correspondants.

$$\mathbb{P}(D_n > c) = 2 \sum_{n=1}^{\infty} (-1)^{n-1} e^{-2n^2 c^2}$$

Par la construction de  $D_n$ , on voit bien que le test est happé par les valeurs centrales et que les queues de distributions ne sont pas bien prises en compte par ce test. Le test suivant est graphique, cela permet de prendre en compte les valeurs extrêmes, mais la contrepartie est que cela reste subjectif.

Soit  $(Y_t, 0 < t \leq n)$  une série d'observations. On souhaite vérifier si elle suit une certaine loi  $\mathcal{L}$  de paramètre  $\theta = (\theta_1, \dots, \theta_d) \in \mathbb{R}^d$ . On génère un échantillon  $(x_1, x_2, \dots, x_n)$  de loi  $\mathcal{L}$  et de même paramètres  $\theta$  que ceux trouvés sur l'échantillon  $(y_1, y_2, \dots, y_n)$ . Le qq-plot est le graphique formé par les couples  $Q_j = (q_{x_j}, q_{y_j})$  où  $q_{u_j}$  est le quantile à l'ordre  $j$  de la variable  $u$ . Si les deux séries ont la même loi sous jacente, alors les  $Q_j$  décrivent une droite. Pour faire simple, il s'agit simplement du graphe des quantiles de l'échantillon observé sur un axe et des quantiles d'une loi de probabilité bien choisie sur l'autre axe. Naturellement, si les deux lois sous

## Modèle de Taux

---

jacentes aux observations sont les mêmes, nous devrions avoir une droite comme graphe.

On présente le qq-plot des log-rendements du S&P 500 sur la figure (5.4) en prenant la loi normale comme loi de référence.

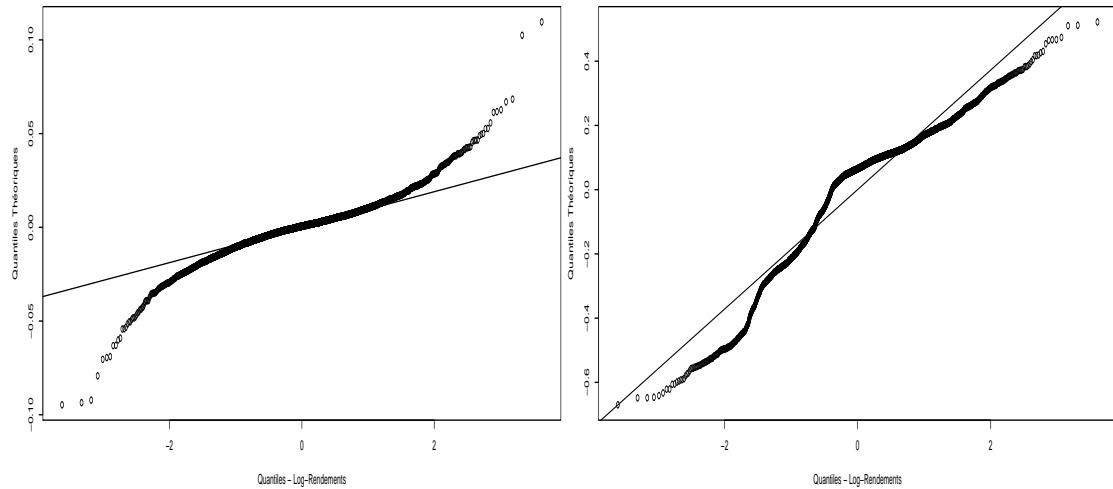


FIGURE 14 – qq-plot des log-rendements du S&P 500 du 13/07/1998 au 16/03/2012, à gauche avec  $\tau = 1$  et à droite avec  $\tau = 251$ , soit les rendements sur des lags d'une année. Le trait plein est la droite d'Henry, si les observations suivaient une loi gaussienne, les quantiles devraient être par dessus cette droite. On constate donc que plus la fréquence choisie est basse, plus la distribution des rendements se rapproche de la loi normale, un 'bon' modèle devra donc avoir cette propriété.

Un autre moyen de regarder la distribution des rendements financiers est simplement de tracer la distribution empirique (5.4),

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{X_i < x}. \quad (5.2)$$

Supposons maintenant pour simplifier que les drawdowns sont gaussien et plus particulièrement qu'ils ont comme diffusion un mouvement brownien géométrique, soit le modèle de Black and Scholes. Par exemple nous pouvons supposer pour simplifier que le portefeuille est constitué d'un seul actif de diffusion de Black and

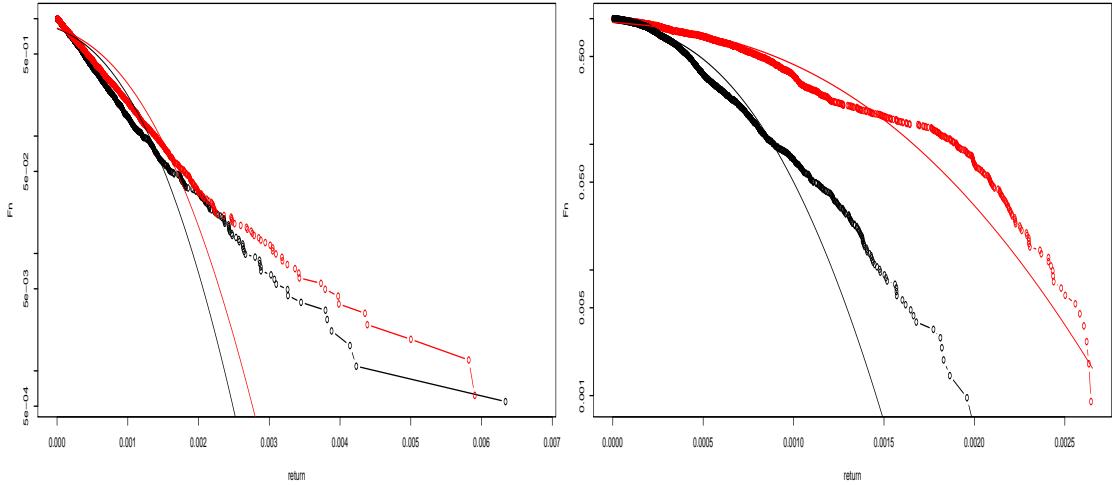


FIGURE 15 – Distribution cumulée empirique,  $\mathbb{P}(\delta_\tau X \geq x)$ , des log-rendements du S&P 500, du 13/07/1998 au 16/03/2012, échelle log, à gauche  $\tau = 1$ , à droite  $\tau = 251$ . En noir les rendements positifs, en rouge les rendements négatifs en valeurs absolues. Trait plein pour le fit gaussien.

Scholes. Pour déterminer la probabilité maximum des drawdowns, nous avons besoin de faire des calculs. Cela repose principalement sur le principe de réflexivité du mouvement brownien.

Notons  $M_T = \max_{0 \leq t \leq T} S_t$  avec

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

où  $(W_t)_{t \in \mathbb{R}}$  un mouvement brownien standard issu de zéro. Notons déjà que si

$$X_T = \max_{0 \leq t \leq T} W_t$$

alors

$$X_T \geq 0 \quad p.s.$$

comme  $W_0 = 0$  par hypothèse. Introduisons maintenant le temps d'arrêt

$$\tau_a = \inf\{t \in \mathbb{R}_+ : W_t = a\}.$$

Il s'agit donc du premier instant où le mouvement brownien passe par  $a > W_0$ .  $W_T$  étant centré,  $-W_T$  à la même loi que  $W_T$ . Par symétrie de la loi nous avons donc

$$\mathbb{P}(W_T - W_{\tau_a} > 0 \mid \tau_a < T) = \mathbb{P}(W_T - W_{\tau_a} < 0 \mid \tau_a < T) = \frac{1}{2}.$$

En multipliant de part et d'autre par  $\mathbb{P}(\tau_a < T)$  (formule de Bayes)

$$2\mathbb{P}(W_T - W_{\tau_a} > 0 \& \tau_a < T) = 2\mathbb{P}(W_T - W_{\tau_a} < 0 \& \tau_a < T) = \mathbb{P}(\tau_a < T).$$

Il est clair que si  $X_T \geq a$  alors l'instant de passage  $\tau_a$  à eu lieu avant  $T$  soit  $\{X_T \geq a\} = \{\tau_a < T\}$  et donc,

$$2\mathbb{P}(W_T - W_{\tau_a} > 0 \& X_T \geq a) = 2\mathbb{P}(W_T - W_{\tau_a} < 0 \& X_T \geq a) = \mathbb{P}(X_T \geq a)$$

Si l'instant de passage à lieu avant  $T$  comme c'est le cas ici, nous avons  $W_{\tau_a} = a$  et nous pouvons conclure

$$\begin{aligned}\mathbb{P}(X_T \geq a) &= 2\mathbb{P}(W_T > a \& X_T \geq a) \\ &= 2\mathbb{P}(W_T > a) \\ &= \mathbb{P}(W_T > a) + \mathbb{P}(-W_T > a) \\ &= \mathbb{P}(|W_T| > a).\end{aligned}$$

La distribution du maximum du mouvement brownien est la même que celle de la valeur absolue du mouvement brownien

$$\begin{aligned}\mathbb{P}(X_T \leq a) &= \mathbb{P}(|W_T| \leq a) \\ &= \frac{1}{\sqrt{2\pi T}} \int_{-a}^a e^{-\frac{1}{2} \frac{x^2}{2T}} \\ &= \frac{2}{\sqrt{2\pi T}} \int_0^a e^{-\frac{1}{2} \frac{x^2}{2T}}, \quad a \in \mathbb{R}_+.\end{aligned}$$

La densité est,

$$\begin{aligned}f_{X_T}(a) &= \frac{d\mathbb{P}(X_T \leq a)}{da} \\ &= \frac{2}{\sqrt{2\pi T}} \mathbf{1}_{[0, \infty)}(a) e^{-\frac{a^2}{2T}}, \quad a \in \mathbb{R}.\end{aligned}$$

Pour notre problématique nous n'avons pas besoin du maximum du mouvement brownien mais du maximum de son exponentiel

$$dS_t = S_t \sigma dW_t$$

(nous faisons le cas sans drift pour commencer). Cela ne change pas grand chose en fait

$$\begin{aligned} M_T &= \max_{0 \leq t \leq T} S_t \\ &= S_0 \max_{0 \leq t \leq T} e^{\sigma W_t} \\ &= S_0 e^{\sigma \max_{0 \leq t \leq T} W_t} = S_0 e^{\sigma X_T} \end{aligned}$$

La densité ne change donc pas, c'est toujours celle du maximum du brownien.

Rajoutons maintenant le drift, nous pouvons par exemple par une astuce considérer le processus

$$\tilde{W}_t = \mu t + W_t.$$

Pour le maximum nous avons

$$\begin{aligned} M_T &= \max_{0 \leq t \leq T} S_t \\ &= S_0 \max_{0 \leq t \leq T} e^{(r - \frac{1}{2}\sigma^2)t + \sigma W_t} \\ &= S_0 \max_{0 \leq t \leq T} e^{\sigma \tilde{W}_t} \\ &= S_0 e^{\max_{0 \leq t \leq T} \tilde{W}_t} \\ &= S_0 e^{\sigma \tilde{X}_T} \end{aligned}$$

$$\mu = -\sigma/2 + r/\sigma.$$

Pour conclure, notons que la principale mesure de performance est le ratio de Sharpe, avec ses défauts. Le ratio de Sortino est également très connu, il s'agit du ratio de Sharpe mais ne considérant que les variations négatives pour le dénominateur, la volatilité. Le ratio de Calmar est quant à lui calculer avec le max drawdowns.

## 6 Appendix

### 6.1 Equation différentielle

Nous avons souvent besoin de savoir résoudre des équations différentielles, d'ordre 1 comme 2, nous allons faire quelques rappels sans rentrer dans la théorie.

#### 6.1.1 Equation du premier ordre

Supposons que nous cherchons la solution de l'équation

$$y' - a(x)y = b(x). \quad (6.1)$$

La première étape consiste à écrire l'équation homogène (EH) associée

$$y' - a(x)y = 0$$

et la résoudre

$$y(x) = \lambda e^{A(x)}$$

où  $A$  est la primitive de  $a$  en  $x$ . Si nous trouvons 'de tête' la solution de (6.1) en prenant en compte le second membre  $b$ , alors il n'y a qu'à faire la somme de la solution de l'équation homogène avec celle-ci. Il s'agit de la méthode de superposition. Si nous n'avons pas d'idée, on peut faire la méthode de variation de la constante, i.e. poser  $\lambda = \lambda(x)$  et faire la dérivée de la solution précédente

$$y'(x) = \lambda'(x)e^{A(x)} + \lambda(x)a(x)e^{A(x)}$$

alors en réinjectant cela dans (6.1)

$$\begin{aligned} \lambda'(x)e^{A(x)} + \lambda(x)a(x)e^{A(x)} - a(x)\lambda(x)e^{A(x)} &= b(x) \\ \lambda'(x)e^{A(x)} &= b(x) \\ \lambda(x) &= \int^x b(t)e^{-A(t)}dt. \end{aligned}$$

Nous en déduisons le résultat final.

#### 6.1.2 Equation d'ordre $n$

Pour les équations du second ordre, c'est-à-dire du type

$$y^{(n)}(x) + a_{n-1}(x)y^{(n-1)}(x) + \cdots + a_1y'(x) + a_0y(x) = b(x)$$

cela est un peu plus compliqué. Nous ne considérons ici que les cas où les coefficients sont constants,  $a_i(x) = a(x)$  pour tout  $i = 0, \dots, n$ . La première étape consiste à considérer l'équation homogène à écrire l'équation caractéristique correspondante

$$z^{(n)} + a_{n-1}z^{(n-1)} + \dots + a_1z + a_0 = 0,$$

et à trouver les racines  $r_m, \dots, r_1$  ( $m \leq n$ ) complexe de multiplicité respective  $\lambda_i$ . L'espace vectoriel solution est alors

$$\mathcal{H} = \left\{ \sum_{k=1}^m P_k(x) e^{\lambda_k x}, \quad Q_k \in \mathbb{C}_{r_k-1}[z] \right\}$$

en d'autre terme,

$$\begin{aligned} y(x) = & R_1(x)e^{x_1 x} + \dots + R_p(x)e^{x_p x} \\ & + (C_{1,1} \cos(\Im(z_1)x) + C_{1,2} \sin(\Im(z_1)x)) e^{\Re(z_1)x} + \dots \\ & + (C_{q,1} \cos(\Im(z_q)x) + C_{q,2} \sin(\Im(z_q)x)) e^{\Re(z_q)x} \end{aligned}$$

où  $x_i$ ,  $i = 1, \dots, p$  sont les racines réelles de multiplicité  $\ell_j$ ,  $z_i$ ,  $i = 1, \dots, q$  les racines complexes de multiplicité  $m_j$  et  $R_i$  des polynômes réels de degré  $\ell_i - 1$  et  $C_{i,1}, C_{i,2}$  des polynômes réels de degré  $m_i - 1$ .

Pour prendre en compte le second membre, deux possibilités, soit il a une 'bonne' forme, soit pas, auquel cas on fait la méthode de variation des constantes. Si bonne forme, i.e. du type  $b(x) = P(x) \cos(\beta x) e^{\alpha x}$  ou  $b(x) = P(x) \sin(\beta x) e^{\alpha x}$  avec  $P$  un polynôme de degré  $d$  alors on va chercher une solution particulière du type  $y(x) = (Q(x) \cos(\beta x) + R(x) \sin(\beta x)) e^{\alpha x}$  où  $Q$  et  $R$  sont des polynômes de degré  $d + m$  avec  $d$  la multiplicité de la racine complexe de l'équation caractéristique (donc égal à 0 si  $\alpha + i\beta$  n'est pas solution).

### 6.1.3 Equation de Riccati

Les équations dites de Riccati sont du type

$$y' = a(x) + b(x)y + c(x)y^2.$$

Avec les coefficients  $a$  et  $c$  différents de zéro. Par les méthodes usuelles, nous ne pouvons pas la résoudre, il faut un peu travailler dessus. La première étape est de faire le changement  $v = c(x)y$ , on obtient alors

$$v' = v^2 + (b(x) + c'(x)/c(x)).$$

En posant ensuite  $v = -u'/u$

$$u'' - (b(x) + c'(x)/c(x))u' + a(x)c(x)u = 0.$$

Cette dernière équation est soluble simplement par les méthodes habituelles. On peut alors retrouver la solution de l'équation initiale  $y = -u'/(c(x)u)$ . Cette équation est entre autre utile pour le modèle CIR dans la valorisation de zéro-coupon ainsi que dans le modèle à volatilité stochastique de Heston.

## 6.2 Calcul stochastique et applications

### 6.2.1 Théorème de Girsanov

Le théorème de Girsanov nous indique qu'un mouvement Brownien  $B_t$  par rapport à la probabilité  $\mathbb{P}$ , de drift  $\lambda t$ , i.e.  $B_t + \lambda t$ , est un mouvement brownien sans drift sous une autre probabilité.

Soit  $L \geq 0$  est une variable aléatoire positive sur l'espace de probabilité  $(\Omega, \mathcal{F}, \mathbb{P})$  tel que  $\mathbb{E}(L) = 1$ . Alors

$$\mathbb{Q} = \mathbb{E}(\mathbf{1}_A L)$$

est une nouvelle probabilité. On dira que  $L$  est la densité de  $\mathbb{Q}$  par rapport à  $\mathbb{P}$  et

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = L.$$

Si l'on considère une variable aléatoire  $X$  définie sur  $(\Omega, \mathcal{F}, \mathbb{Q})$ , nous avons

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}(X) &= \int_{\Omega} X d\mathbb{Q} \\ &= \int_{\Omega} X L d\mathbb{P} \\ &= \mathbb{E}(XL).\end{aligned}$$

En particulier, si  $L_t$  est une  $\mathcal{F}_t$ -martingale sous  $\mathbb{P}$ , i.e.  $\mathbb{E}_{\mathbb{P}}(L_T | \mathcal{F}_t) = L_t$  alors

$$\mathbb{Q}_T(A) = \mathbb{E}_{\mathbb{P}}(\mathbb{E}_{\mathbb{P}}(L_T | \mathcal{F}_t) \mathbf{1}_A) = \mathbb{E}_{\mathbb{P}}(L_t \mathbf{1}_A).$$

**Lemme 6.1.** Soit  $(M_t)_{t \in \mathbb{R}_+}$ , c'est une  $\mathcal{F}_t$ -martingale sous  $\mathbb{Q}$  ssi  $(L_t M_t)_{t \in \mathbb{R}_+}$  est  $\mathcal{F}_t$ -martingale sous  $\mathbb{P}$ .

**Preuve 6.1.** *Supposons que  $M$  soit une  $\mathbb{Q}$ -martingale. Pour tout  $s \leq t \leq T$  et pour tout  $A \in \mathcal{F}_s$  on*

$$\mathbb{E}_{\mathbb{P}}(L_t M_t \mathbf{1}_A) = \mathbb{E}_{\mathbb{Q}}(M_t \mathbf{1}_A) = \mathbb{E}_{\mathbb{Q}}(M_s \mathbf{1}_A) = \mathbb{E}_{\mathbb{P}}(Z_s M_s \mathbf{1}_A),$$

ainsi,  $ZM$  est une  $\mathbb{P}$ -martingale, la preuve de la réciproque étant similaire.

Considérons la suite  $(X_1, \dots, X_n)$  de variables gaussiennes centrées réduites indépendantes sur  $(\Omega, \mathcal{F}, \mathbb{P})$ . Pour tout  $(\mu_1, \dots, \mu_n) \in \mathbb{R}^n$ ,

$$\mathbb{E} \left( \exp \sum_{i=1}^n \mu_i X_i \right) = \prod_{i=1}^n \mathbb{E}(\exp(\mu_i X_i)) = \prod_{i=1}^n \exp \left( \frac{\mu_i^2}{2} \right) = \exp \frac{1}{2} \left( \sum_{i=1}^n \mu_i^2 \right),$$

ainsi,

$$\mathbb{E} \left( \exp \sum_{i=1}^n \left( \mu_i X_i - \frac{\mu_i^2}{2} \right) \right) = 1.$$

Posons

$$Z(\omega) = \exp \sum_{i=1}^n \left( \mu_i X_i(\omega) - \frac{\mu_i^2}{2} \right),$$

alors nous avons une nouvelle probabilité  $\mathbb{Q}$  sur  $(\Omega, \mathcal{F})$ ,  $\mathbb{Q}(d\omega) = Z(\omega) \mathbb{P}(d\omega)$ , soit,

$$\mathbb{Q}(A) = \mathbb{E}_{\mathbb{P}}(Z \mathbf{1}_A) = \int_A Z(\omega) \mathbb{P}(d\omega).$$

La mesure  $\mathbb{Q}$  est bien une probabilité sur  $(\Omega, \mathcal{F}_t)$  puisque  $\mathbb{E}_{\mathbb{P}}(Z) = 1$ . Pour toute la suite, nous posons

$$Z_t^m = \exp \left( mW_t - \frac{m^2 t}{2} \right).$$

**Théorème 6.1.** (*Formule Cameron-Martin*) *Sous la mesure  $\mathbb{Q}_T^m$ , le processus*

$$\tilde{W}_t : t \mapsto W_t - mt, \quad t \leq T$$

*est un mouvement Brownien.*

**Preuve 6.2.** *Remarquons que les filtrations  $\mathcal{F}_t^W$  et  $\mathcal{F}_t^{\tilde{W}}$  sont identiques pour tout  $t \leq T$ . Fixons  $\lambda \in \mathbb{R}$  et posons*

$$L_t^\lambda = \exp(\lambda \tilde{W}_t - \lambda^2 t / 2).$$

Pour tout  $s \leq t \leq T$  et pour  $A \in \mathcal{F}_s^W$ ,

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}(L_t^\lambda \mathbb{1}_A) &= \mathbb{E}_{\mathbb{P}}(L_t^\lambda Z_t \mathbb{1}_A) \\ &= \mathbb{E}_{\mathbb{P}}(\exp(\lambda \tilde{W}_t - \lambda^2 t/2 + mW_t - m^2 t/2) \mathbb{1}_A) \\ &= \mathbb{E}_{\mathbb{P}}(\exp(\lambda W_t - \lambda m t - \lambda^2 t/2 + mW_t - m^2 t/2) \mathbb{1}_A) \\ &= \mathbb{E}_{\mathbb{P}}(\exp((\lambda + m)W_t - (\lambda + m)^2 t/2) \mathbb{1}_A) \\ &= \mathbb{E}_{\mathbb{P}}(\exp((\lambda + m)W_s - (\lambda + m)^2 s/2) \mathbb{1}_A) \\ &= \mathbb{E}_{\mathbb{P}}(L_s^\lambda Z_s \mathbb{1}_A) \\ &= \mathbb{E}_{\mathbb{Q}}(L_s^\lambda \mathbb{1}_A).\end{aligned}$$

Ainsi,

$$\mathbb{E}_{\mathbb{Q}}(L_t^\lambda | \mathcal{F}_s^W) = L_s^\lambda, \quad s \leq t \leq T.$$

$L^\lambda$  est donc une  $(\mathcal{F}_t^W)$ -martingale pour tout  $\lambda \in \mathbb{R}$ . Le théorème de caractérisation de Paul Lévy nous permet de conclure.

D'après le résultat précédent, on sait que l'EDS

$$\begin{aligned}Z_t &:= 1 + \int_0^t \theta(s) Z_s dW_s \\ &= \exp\left(\int_0^t \theta(s) dW_s - \frac{1}{2} \int_0^t \theta^2(s) ds\right)\end{aligned}$$

est une  $(\mathcal{F}_t^W)$ -martingale.

**Théorème 6.2.** (Girsanov) Sous la probabilité  $\mathbb{Q}$ , le processus

$$\tilde{W}_t = W_t - \int_0^T \theta(s) ds, \quad t \leq T$$

est un mouvement brownien.

**Preuve 6.3.** D'après (6.1), on montre que  $S\tilde{W}$  est une martingale sous  $\mathbb{P}$  et par application du lemme,  $\tilde{W}$  est une martingale sous  $\mathbb{Q}$ . De plus, il est clair que nous avons  $\langle \tilde{W}_t \rangle = t$ . On en déduit que sous  $\mathbb{Q}$ ,  $\tilde{W}$  est une martingale continue de crochet  $\langle \tilde{W}_t \rangle = t$ , soit un mouvement Brownien.

### 6.2.2 Changement de numéraire

Nous allons rappeler succinctement l'un des principaux résultats en terme de valorisation de produits dérivés, le changement de numéraire.

Un numéraire est un processus stochastique positif  $(N_t)_{t \in \mathbb{R}_+}$  adapté à la filtration  $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$  qui peut être vu comme une unité de référence. Le prix d'un actif  $S_t$  de numéraire  $N_t$  est en général donné par

$$\tilde{S}_t = \frac{S_t}{N_t}, \quad t \geq 0.$$

L'un des exemples les plus connus et les plus utilisés est l'actualisation

$$N_t = \exp \left( \int_0^t r_s ds \right)$$

où  $r_t$  est le taux d'intérêt sans risque, déterministe ou aléatoire. Dans ce cas,

$$\tilde{S}_t = \frac{S_t}{N_t} = \exp \left( \int_0^t -r_s ds \right) S_t$$

est le prix actualisé de l'actif en  $t = 0$ .

Sur le marché des changes, nous pouvons par exemple prendre  $N_t = R_t$ , la monnaie étrangère. Sur le marché des taux, nous prendrons le numéraire-forward.

L'un des objectifs quand on fait un changement de numéraire est de trouver la probabilité risque neutre  $\mathbb{Q}$  telle que sous cette mesure,  $\tilde{S}_t = S_t/N_t$  soit une martingale.

On associe à un numéraire  $N$  le changement de probabilité suivant : Soit  $\mathbb{Q}$  la probabilité risque-neutre telle que le processus  $(N_t R_t)_{t \in \mathbb{R}_+}$  est une  $\mathbb{Q}$ -martingale. Soit  $\mathbb{Q}^N$  définie par  $\mathbb{Q}^N|_{\mathcal{F}_t} = (N_t R_t) \mathbb{Q}|_{\mathcal{F}_t}$ .

**Proposition 6.1.** *Soit  $(S_t)_{t \in \mathbb{R}_+}$  le prix d'un actif financier et  $N$  un numéraire. Le prix de  $S$  dans le numéraire  $N$ ,  $(S_t/N_t, 0 \leq t \leq T)$  est une  $\mathbb{Q}^N$ -martingale.*

**Preuve 6.4.** *Si  $S$  est un processus de prix, le processus actualisé  $\tilde{S}_t := S_t R_t$  est une  $\mathbb{Q}$ -martingale. Ainsi,  $S_t/M_t$  est une  $\mathbb{Q}^N$ -martingale si  $(S_t/N_t) N_t R_t = R_t S_t$  est une  $\mathbb{Q}$ -martingale.*

Un exemple d'utilisation du changement de numéraire apparaît dans le calcul d'une option avec le modèle Black & Scholes. Le terme  $\mathbb{E}_{\mathbb{Q}}(S_T e^{-rT} \mathbf{1}_A)$  se détermine

très facilement. Sous  $\mathbb{Q}$ , le processus  $N_t = S_t e^{-rt} / S_0$  est une martingale positive d'espérance 1, en posant  $d\mathbb{Q} = N_t d\tilde{\mathbb{Q}}$ , nous avons

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}(S_T e^{-rT} \mathbf{1}_A) &= \mathbb{E}_{\tilde{\mathbb{Q}}}(S_0 \mathbf{1}_A) \\ &= S_0 \tilde{\mathbb{Q}}(S_T \geq A).\end{aligned}$$

Bien entendu, sous  $\mathbb{Q}$  ou  $\tilde{\mathbb{Q}}$ , la dynamique de  $S$  n'est plus la même, nous devons donc la déterminer.

### 6.2.3 Théorème de Feynman-Kac

**Théorème 6.3.** Soit  $\Phi$  une fonction continue et  $u$  une fonction de classe  $C^{1,2}$  à dérivée en  $x$  bornée sur  $[0, T] \times \mathbb{R}$  solution du problème,

$$\begin{cases} \left( \frac{\partial u}{\partial t} + \mu \frac{\partial u}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2 u}{\partial x^2} \right)(t, \rho) - \rho_t u(t, \rho) = 0 \\ u(T, \rho) = \Phi(\rho) \end{cases}$$

Alors

$$M_t = u(t, \rho_t) e^{-\int_0^t \rho_s ds}, t \in [0, T]$$

est une martingale locale. De plus, si

1.  $\frac{\partial M_t \sigma(t, \rho_t)}{\partial \rho} \in \mathcal{L}^2[0, T]$ , ou
2.  $M$  uniformément bornée,

alors  $M$  est une (vraie) martingale et la solution de ce problème est donnée par,

$$u(t, \rho_t) = \mathbb{E}_{\mathbb{Q}} \left[ e^{-\int_t^T \rho_s ds} \Phi(r_T) | \mathcal{F}_t \right],$$

**Preuve 6.5 (Preuve).** Il est clair que  $M$  est de classe  $C^{1,2}$ , alors,

$$\begin{aligned}dM_t &= \left( \frac{\partial u}{\partial t}(t, \rho_t) + \mu \frac{\partial u}{\partial \rho}(t, \rho_t) + \frac{1}{2} \sigma^2 \frac{\partial^2 u}{\partial \rho^2}(t, \rho_t) - \rho_t u(t, \rho_t) \right) e^{-\int_0^t \rho_s ds} dt \\ &\quad + \frac{\partial u}{\partial \rho}(t, \rho_t) e^{-\int_0^t \rho_s ds} \sigma(t, \rho_t) dB_t \\ &= \frac{\partial u}{\partial \rho}(t, \rho_t) e^{-\int_0^t \rho_s ds} \sigma(t, \rho_t) dB_t.\end{aligned}$$

Ainsi,  $M$  est bien une martingale locale. En rajoutant l'une des deux conditions, nous avons une martingale. Ensuite, comme

$$M_T = \Phi(\rho_T) e^{-\int_0^T \rho_u du}$$

nous avons,

$$u(t, \rho_t) e^{-\int_0^t \rho_u du} = \mathbb{E}_{\mathbb{Q}} \left[ \exp(-\int_0^T \rho_u du) \Phi(\rho_T) \middle| \mathcal{F}_t \right].$$

On termine la preuve en multipliant par  $e^{\int_0^t \rho_u du}$ .

### 6.3 Pricing d'option par transformée de Fourier

La transformée de Fourier  $\mathcal{F}f$  d'une fonction intégrable  $f$  de  $\mathbb{R}$  dans  $\mathbb{R}$  est définie pour tout  $k$  par

$$\mathcal{F}f(k) = \int_{-\infty}^{\infty} e^{ikx} f(x) dx.$$

La fonction caractéristique d'une variable aléatoire  $X$  est la transformée de Fourier de sa densité  $p$

$$\phi(k) = \mathbb{E}(\exp(ikX)) = \int_{-\infty}^{\infty} e^{ikx} p(x) dx, \quad k \in \mathbb{R}.$$

Dans l'espace de Fourier, l'opérateur de différentiation devient une multiplication par  $ik$ , pour une fonction  $f$  différentiable, nous avons,

$$\mathcal{F}f^{(m)}(k) = (-ik)^m \mathcal{F}(k), \quad k \in \mathbb{R}.$$

Enfin, la transformée de Fourier inverse est donnée par

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx} \mathcal{F}f(k) dk$$

Il est possible de montrer que toute variable aléatoire  $X$  intégrable au sens de Fourier et de fonction caractéristique  $\phi$ , admet une densité de probabilité faible donnée pour tout  $\xi > 0$  par

$$p(x) = \frac{1}{2\pi} \int_{-\xi - \infty}^{-\xi + \infty} e^{-ikx} \phi(k) dk.$$

En terme de cette probabilité  $p$ , nous pouvons donc écrire pour tout  $\phi$  suffisamment régulier

$$\mathbb{E}(\phi(X)) = \frac{1}{2\pi} \int_{-\xi i - \infty}^{-\xi i + \infty} \phi(k) \left( \int_{-\infty}^{\infty} e^{-iky} \phi(y) dy \right) dk.$$

Si nous prenons le cas particulier de  $\phi(y) = \mathbf{1}_>$  pour  $x \in \mathbb{R}$  nous avons alors

$$\begin{aligned} \mathbb{P}(X > x) &= \frac{1}{2\pi} \int_{-\xi i - \infty}^{-\xi i + \infty} \phi(k) \left( \int_x^{\infty} e^{-iky} dy \right) dk \\ &= \frac{1}{2\pi i} \int_{-\xi i - \infty}^{-\xi i + \infty} \frac{e^{-ikx}}{k} \phi(k) dk. \end{aligned}$$

En faisant tendre  $\xi$  vers  $0^+$ , par application du théorème des résidus de Cauchy nous avons

$$\mathbb{P}(X > x) = \frac{1}{2} + \frac{1}{2\pi} \lim_{\varepsilon \rightarrow 0^+} \int_{|k| > \varepsilon} \frac{e^{-ikx} \phi(k)}{ik} dk.$$

En remarquant que

$$\lim_{\varepsilon \rightarrow 0^+} \int_{|k| > \varepsilon} \frac{e^{-ikx} \phi(k)}{ik} dk = 2 \lim_{\varepsilon \rightarrow 0^+} \int_{\varepsilon}^{\infty} \Re \left( \frac{e^{-ikx} \phi(k)}{ik} \right) dk,$$

nous avons

$$\mathbb{P}(X \leq x) = \frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \Re \left( \frac{e^{-ikx} \phi(k)}{ik} \right) dk \quad (6.2)$$

**Proposition 6.2.** Soit  $(S_t)_{t \in \mathbb{R}_+}$  le prix d'un actif financier. Nous considérons un call européen de maturité  $T$  et de strike  $K$  de sous-jacent  $S$ . En notant  $x_t = \ln(S_t)$ , la valeur du call à l'instant  $t = 0$  est donnée par

$$C(0, S, T, K) = S_0 \Pi_1 - K e^{-rT} \Pi_2$$

où les pseudo-probabilités  $\Pi_1$  et  $\Pi_2$  sont données par

$$\begin{aligned} \Pi_1 &= \frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \Re \left( \frac{e^{-iuk} \phi_T(u-i)}{iu \phi_T(-i)} \right) du \\ \Pi_2 &= \frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \Re \left( \frac{e^{-iuk} \phi_T(u)}{iu} \right) du \end{aligned}$$

où  $\phi_t = \mathbb{E}[\exp(iux_T)]$ .

**Preuve 6.6.** *Revenons en arrière dans le calcul. Nous avons directement*

$$C(0, S, T, K) = \mathbb{E} \left( e^{-rT} e^{x_T} \mathbf{1}_{\{x_T > k\}} \right) - K e^{-rT} \mathbb{P}(x_T > k)$$

où  $k = \ln K$ . Alors, par (6.2),

$$\Pi_2 = \mathbb{P}(x_T > k) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re \left( \frac{e^{-iux} \phi_T(u)}{iu} \right) du$$

Pour déterminer le premier terme, nous introduisons une nouvelle mesure  $\mathbb{Q}$  tel que  $d\mathbb{Q}/d\mathbb{P} = S_T/S_0$ . Ainsi

$$\begin{aligned} \mathbb{E} \left( e^{-rT} e^{x_T} \mathbf{1}_{\{x_T > k\}} \right) &= \int_{t \in \mathbb{R}} e^{-rt} e^{x(T)} \mathbf{1}_{x_T > k} d\mathbb{P}(x) \\ &= \int_{t \in \mathbb{R}} e^{-rt} e^{x(T)} \mathbf{1}_{x_T > k} dS_0 \frac{\mathbb{Q}}{S_T}(x) \\ &= S(0)\mathbb{Q}(x_T > k) \end{aligned}$$

Comme nous l'avons expliqué lors des changements de numéraire, nous avons besoin de déterminer les changements de caractéristiques du processus sous la nouvelle mesure. Ce qui nous intéresse ici va donc être de déterminer la forme de la fonction caractéristique de  $(x_t)_{t \in \mathbb{R}_+}$  sous  $\mathbb{Q}$  en  $t = T$

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}(e^{iux_T}) &= \int_{x \in \mathbb{R}} e^{iux_T} d\mathbb{Q}(x) \\ &= \frac{\mathbb{E}(e^{x_T} e^{iux_T})}{\mathbb{E}(e^{x_T})} = \frac{\phi_T(u-i)}{\phi_T(-i)}. \end{aligned}$$

Toujours par (6.2),

$$\Pi_1 = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re \left( \frac{e^{-iuk} \phi_T(u-i)}{iu \phi_T(-i)} \right) du.$$

Pour plus de précisions sur cette méthode et d'évaluation des options et des approches alternatives, nous renvoyons à [CaMa99].

## 6.4 Approximation d'intégrale

Pour pouvoir appliquer les résultats précédents, nous avons besoin d'approximer des intégrales infinies, ce qui n'est pas intuitivement trivial. Pour des intégrales

définies entre  $a$  et  $b$ , deux réels différents de  $\pm\infty$  nous avons pour une intégrale de Riemann la discrétisation

$$\int_a^b f(x)dx \approx \sum_{i=1}^n f(k_i)$$

où  $a = k_1, \dots, k_n = b$ . La méthode générale d'estimation d'intégrale est la quadrature de Gauss. Pour tout  $a < b \in \mathbb{R}$ ,

$$\int_a^b w(x)f(x)dx \approx \sum_{i=1}^n w'_i f(x_i)$$

où  $w$  est une fonction de pondération,  $w_i$  les coefficients de quadrature et les  $x_i$  sont les noeuds, les racines de polynômes orthogonaux choisies judicieusement. Ces deux coefficients sont choisis en fonction du domaine d'intégration. Bien entendu, la fonction que l'on cherche à intégrer n'a pas toujours la forme en  $w(x)f(x)$ , il suffit de poser  $f(x) = w(x)\frac{f(x)}{w(x)}$ . On présente quelques configurations connues.

- Sur  $[-1,1]$  nous pouvons appliquer la méthode de Gauss-Legendre. Les poids sont simplement  $w(x) = 1$ . Les coefficients  $w_i$  sont données par

$$w_i = \frac{2}{(1 - x_i^2)P'_n(x_i)^2}$$

où les  $P_n$  sont les polynômes de Legendre vérifiant la relation de récurrence

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x), \quad n > 0,$$

avec  $P_0(x) = 1$  et  $P_1(x) = x$ . Les premiers polynômes sont donc

$$\begin{aligned} P_1(x) &= x \\ P_2(x) &= \frac{1}{2}(3x^2 - 1) \\ P_3(x) &= \frac{1}{2}(5x^3 - 3x). \end{aligned}$$

Ainsi, les racines sont respectivement  $\{0\}$ ,  $\{-\sqrt{\frac{1}{3}}, \sqrt{\frac{1}{3}}\}$  et  $\{-\sqrt{\frac{3}{5}}, 0, \sqrt{\frac{3}{5}}\}$ . Après calcul des  $w_i$ , nous avons les approximations suivantes en fonction du choix de  $n$

$$\begin{aligned} (n=1) \quad \int_{-1}^1 f(x)dx &\approx 2f(0) \\ (n=2) \quad \int_{-1}^1 f(x)dx &\approx 1f(-\sqrt{\frac{1}{3}}) + 1f(\sqrt{\frac{1}{3}}) \\ (n=3) \quad \int_{-1}^1 f(x)dx &\approx \frac{5}{9}f(-\frac{3}{5}) + \frac{8}{9}f(0) + \frac{5}{9}f(\sqrt{\frac{3}{5}}). \end{aligned}$$

- Sur  $\mathbb{R}_+$  nous appliquons la quadrature de Gauss-Laguerre. Les poids sont  $w(x) = e^{-x}$ . L'approximation se fera donc sur  $g(x) = e^x f(x)$  pour tenir compte des poids. Les coefficients  $w_i$  sont donnés par

$$w_i = \frac{1}{(n+1)L'_n(x_i)L_{n+1}(x_i)}$$

où les  $L_n$  sont les polynômes de Laguerre vérifiant l'équation de récurrence

$$(n+1)L_{n+1}(x) = (2n+1-x)L_n(x) - nL_{n-1}(x), \quad n > 0$$

avec  $L_0(x) = 1$  et  $L_1(x) = -x + 1$ . Les premiers polynômes sont donc

$$\begin{aligned} L_1(x) &= -x + 1 \\ L_2(x) &= \frac{1}{2}(x^2 - 4x + 2) \\ L_3(x) &= \frac{1}{6}(-x^3 + 9x^2 - 18x + 6). \end{aligned}$$

Le calcul des racines et des  $w_i$  est laissé au lecteur.

- Sur  $\mathbb{R}$  nous appliquons la quadrature de Gauss-Hermite. Les poids sont  $w(x) = e^{-x^2}$ . L'approximation se fera donc sur  $e^{x^2}f(x)$  pour tenir compte des poids. Les coefficients  $w_i$  sont donnés par

$$w_i = \frac{2^{n+1}n!\sqrt{\pi}}{H'_n(x_i)^2}$$

où les  $H_n$  sont les polynômes d'Hermite vérifiant l'équation de récurrence

$$H_{n+1}(x) = xH_n(x) - nH_{n-1}(x), \quad n > 0$$

avec  $H_0(x) = 1$  et  $H_1(x) = x$ . Les premiers polynômes sont donc

$$\begin{aligned} H_1(x) &= x \\ H_2(x) &= x^2 - 1 \\ H_3(x) &= x^2 - 3x. \end{aligned}$$

Le calcul des racines et des  $w_i$  est laissé au lecteur.

## 6.5 Optimisation non linéaire

### 6.5.1 Descente de gradient

**Définition 6.1** (descente de gradient). Soit  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  une fonction  $C^1$  à minimiser. La dérivée partielle de  $f$  s'écrit,

$$\frac{\partial f}{\partial x} = \lim_{\delta x \rightarrow 0} \frac{f(x + \delta x) - f(x)}{\delta x}, \quad x \in \mathbb{R}^d, \quad (6.3)$$

cela nous indique comment un changement de  $x$  affecte  $f$ . Si  $\frac{\partial f}{\partial x_i} < 0$  cela signifie qu'une augmentation infinitésimale du paramètre  $x_i$  ferait baisser la fonction  $f$ . On cherche donc à changer  $x$  dans la direction du vecteur de gradient  $\nabla f$ ,

$$\begin{cases} x_0 \in \mathbb{R}^d \\ x_{k+1} = x_k - \alpha \nabla_k f \end{cases}, \quad (6.4)$$

$\alpha > 0$  est le pas de déplacement, s'il est assez petit on sera assuré d'avoir une réduction de  $f$  à chaque itération, mais si trop petit, la convergence sera très lente vers un minimum local, si trop grand, divergence possible.

On peut proposer une autre interprétation de cet algorithme, si on écrit le développement de Taylor de  $f$  à l'ordre 1 par rapport à  $x_k$  on obtient,

$$f(x) = f(x_k) + \nabla f(x_k)(x - x_k) + \frac{1}{2\alpha} \|x - x_k\|^2. \quad (6.5)$$

Comme on aboutit à une forme quadratique, on peut chercher à minimiser simplement (6.5),

$$\nabla_x(f(x_k) + \nabla f(x_k)(x - x_k) + \frac{1}{2\alpha} \|x - x_k\|^2) = \nabla f(x_k) + \frac{1}{\alpha}(x - x_k). \quad (6.6)$$

On cherche donc à résoudre,

$$\begin{aligned} \nabla f(x_k) + \frac{1}{\alpha}(x - x_k) &= 0 \\ x_{k+1} &= x_k - \alpha \nabla f(x_k). \end{aligned} \quad (6.7)$$

On retrouve bien l'algorithme de descente du gradient.

Il s'agit d'un exemple simple d'un algorithme de descente. Les idées générales d'un *algorithme de descente* sont les suivantes :

- Recherche d'une direction de descente  $d_k$ , direction dans laquelle il est possible de faire décroître  $f$
- Recherche d'un certain  $\alpha$  qui permet une décroissance significative

La forme usuelle est,

$$x_{k+1} = x_k + \alpha d_k. \quad (6.8)$$

Une fois la direction  $d_k$  déterminée, il faut s'occuper de déterminer le pas de déplacement optimal,  $\alpha$ . Comme on cherche à minimiser  $f$ , on peut vouloir minimiser le critère le long de  $d_k$  et donc de déterminer le  $\alpha_k$  comme solution du problème,

$$\alpha_k = \arg \min_{\alpha} f(x_k + \alpha d_k). \quad (6.9)$$

Il s'agit de la règle de Cauchy. La règle suivante s'appelle la règle de Curry,

$$\alpha_k = \inf\{\alpha \geq 0 : f'(x_k + \alpha d_k) = 0, f(x_k + \alpha d_k) < f(x_k)\}. \quad (6.10)$$

Dans ce cas,  $\alpha_k$  est donc choisi parmi les points stationnaires.

Ces deux manières de déterminer le pas optimal  $\alpha_k$  sont dites de *recherche exacte*. Prenons le cas d'une fonction quadratique, i.e.  $\nabla f(x) = Ax + b$ , nous avons simplement,

$$d_k^\top (Ax_k + b + \alpha_k A d_k) = d_k^\top \nabla f(x_k) + \alpha_k d_k^\top A(d_k) = 0, \quad (6.11)$$

si  $A$  est définie positive, et donc que la fonction est convexe, la valeur du pas de déplacement est simplement donnée par,

$$\alpha_k = -\frac{d_k^\top \nabla f(x_k)}{\|d_k\|_A^2}. \quad (6.12)$$

Si la fonction à minimiser n'a pas toutes les 'bonnes propriétés' requises, alors il se peut que les règles (et donc les pas fournis) de Cauchy ou Curry n'existent pas. De plus, la détermination exacte de  $\alpha_k$  (au sens de Cauchy ou Curry) peut s'avérer très coûteuse en temps de calcul, ce qui n'est pas forcément compensé par le gain de temps dans la minimisation de  $f$ . On va naturellement essayer de construire des algorithmes peu gourmands approchant au moins l'une des deux

équations (6.9), (6.10).

Pour essayer d'approcher la condition de Cauchy on peut rechercher  $\alpha_k$  tel que,

$$f(x_k + \alpha_k d_k) \leq f(x_k) + w_1 \alpha_k \nabla f(x_k) d_k^\top, \quad (6.13)$$

où  $w_1 \in ]0, 1[$ . Il s'agit de la condition d'Armijo. Pour arriver à trouver un pas de déplacement qui vérifie l'équation précédente on applique l'algorithme suivant,

---

**Algorithme : Recherche Linéaire approchée - Armijo**

---

Initialisation :  $\alpha_k^1 > 0, \tau \in ]0, 1[$

Tant que (6.13) n'est pas vérifiée i.e.,

$$f(x_k + \alpha_k^i d_k) > f(x_k) + w_1 \alpha_k^i \nabla f(x_k) d_k^\top.$$

(1) On prend

$$\alpha_k^{i+1} \in [\tau \alpha_k^i, (1 - \tau) \alpha_k^i].$$

(2)  $i = i + 1$

---

Le pas déterminé par cet algorithme a la fâcheuse tendance à être trop petit et donné une convergence trop lente. La règle de Goldstein est,

$$f(x_k + \alpha_k d_k) \geq f(x_k) + w'_1 \alpha_k \nabla f(x_k) d_k^\top. \quad (6.14)$$

où  $w'_1 \in [w_1, 1]$ . Cette règle permet de ne pas choisir de  $\alpha_k$  trop petit. Les conditions de Wolfe permettent également d'avoir un pas de déplacement pas trop "petit" et vérifient la condition de décroissance linéaire (6.13),

$$\begin{aligned} f(x_k + \alpha_k d_k) &\leq f(x_k) + w_1 \alpha_k \nabla f(x_k) d_k^\top. \\ \nabla f(x_k + \alpha_k d_k) &\geq w_2 g_k d_k^\top \end{aligned} \quad (6.15)$$

Pour déterminer  $\alpha$  vérifiant les conditions de Wolfe, on peut appliquer l'algorithme de Fletcher-Lemaréchal

Précédemment nous avions montré que  $d_k$  pouvait être égal à  $-\nabla f(x_k)$ , si nous faisons maintenant le développement de Taylor de  $f$  au second ordre (on doit donc avoir une fonction  $f$  de classe  $C^2$ ) nous obtenons,

$$f(x) = f(x_k) + \nabla f(x_k)(x - x_k) + \frac{1}{2}(x - x_k)^\top H f(x_k)(x - x_k), \quad (6.16)$$

**Algorithme : Recherche Linéaire approchée - Fletcher-Lemaréchal**

Initialisation :  $\alpha_k^1 = 0$ , Soient  $\underline{\alpha}^1 = 0$ ,  $\bar{\alpha}^1 = \infty$ ,  $\tau \in ]0, \frac{1}{2}[$  et  $\tau_e > 1$

(1) Si

$$f(x_k + \alpha_k^i d_k) \leq f(x_k) + w_1 \alpha_k^i \nabla f(x_k) d_k^\top.$$

n'est pas vérifiée, alors  $\bar{\alpha}^i = \alpha^i$  et,

$$\alpha^i \in [(1 - \tau)\underline{\alpha}^i + \tau\bar{\alpha}^i, \tau\underline{\alpha}^i + (1 - \tau)\bar{\alpha}^i].$$

(2) Sinon,

(2.1) Si

$$\nabla f(x_k + \alpha_k^i d_k) \geq w_2 g_k d_k^\top.$$

est vérifiée, alors stop.

(2.2) Sinon,  $\underline{\alpha}^i = \alpha^i$ .

(2.3) Si  $\bar{\alpha}^i = +\infty$ , alors prendre  $\alpha^i \in [\tau_e \underline{\alpha}^i, \infty]$ , sinon,

$$\alpha^i \in [(1 - \tau)\underline{\alpha}^i + \tau\bar{\alpha}^i, \tau\underline{\alpha}^i + (1 - \tau)\bar{\alpha}^i].$$


---

où  $Hf(x)$  est le hessien de  $f$  au point  $x$ . Les conditions d'optimalité sont,

$$\nabla f(x_k) + Hf(x_k)(x_{k+1} - x_k) = 0. \quad (6.17)$$

Et donc, la mise à jour est donnée par,

$$x_{k+1} = x_k - Hf(x_k)^{-1} \nabla f(x_k). \quad (6.18)$$

Il s'agit de la méthode de Newton. Le problème ici est le temps de calcul de la matrice hessienne  $Hf$ , on essaye d'approximer  $Hf$ , ou directement  $H^{-1}f$ ,

$$x_{k+1} = x_k - \alpha B_k \nabla f(x_k), \quad (6.19)$$

où  $B_k$  est l'approximation de  $H^{-1}f$  (on parle alors de quasi Newton).

Pour toute la suite on pose  $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$  et  $s_k = x_{k+1} - x_k$ . L'approximation de Broyden est donnée par,

$$B_{k+1} = B_k + \frac{(s_k - B_k y_k)(s_k - B_k y_k)^\top}{(s_k - B_k y_k)^\top y_k}. \quad (6.20)$$

---

**Algorithme : Quasi Newton - Broyden**

---

Initialisation :  $\alpha_0 = 1, H_0 = I$

pour  $k = 1, 2, \dots$

- (1) Calcul de la direction de descente,  $d_k = -B_k \nabla f(x_k)$
- (2) On détermine  $\alpha_k$  en appliquant une recherche linéaire de Wolfe.
- (3) On calcule

$$B_{k+1} = B_k + \frac{(s_k - B_k y_k)(s_k - B_k y_k)^\top}{(s_k - B_k y_k)^\top y_k}.$$

- (4) Stop si  $\|\nabla f(x_k)\| \leq \epsilon$ , alors  $x^* = x_k$ , sinon, retour à l'étape 1.
- 

Le problème ici est que les  $B_k$  ne sont pas forcément définis positifs. La mise à jour de Davidson, Fletcher et Powell permet de pallier à ce problème, elle est donnée par,

$$B_{k+1} = B_k + \frac{s_k s_k^\top}{s_k^\top y_k} - \frac{B_k y_k y_k^\top B_k}{y_k^\top B_k y_k}. \quad (6.21)$$

Nous présentons enfin la méthode de Broyden, Fletcher, Goldfarb et Shanno (BFGS),

---

**Algorithme : Quasi Newton - BFGS**

---

Initialisation :  $\alpha_0 = 1, H_0 = I$

pour  $k = 1, 2, \dots$

- (1) Calcule de la direction de descente,  $d_k = -H_k^{-1} \nabla f(x_k)$
- (2) On détermine  $\alpha_k$  en appliquant une recherche linéaire de Wolfe.
- (3) On calcul

$$B_{k+1} = B_k - \frac{s_k y_k^\top B_k + B_k y_k s_k^\top}{y_k^\top s_k} + \left(1 + \frac{y_k^\top B_k y_k}{y_k^\top y_k}\right) \frac{s_k s_k^\top}{y_k^\top s_k}.$$

- (4) Stop si  $\|\nabla f(x_k)\| \leq \epsilon$ , alors  $x^* = x_k$ , sinon, retour à l'étape 1.
- 

On peut montrer dans ce cas que la convergence est superlinéaire, i.e. que,

$$\lim_{\tau \rightarrow \infty} \frac{\|\mathbf{w}_{k+1} - \mathbf{w}^*\|}{\|\mathbf{w}_k - \mathbf{w}^*\|} = 0. \quad (6.22)$$

Notons qu'il existe une version modifiée de l'algorithme BFGS ne nécessitant pas de calculer explicitement  $Hf$  et donc d'avoir à stocker un trop grand nombre de données, il s'agit de l'algorithme Low Memory BFGS.

### 6.5.2 Algorithme génétique

L'algorithme génétique appartient à la classe des recherches heuristiques. Un algorithme génétique est une fonction d'optimisation pour les cas extrêmes où le maximum (minimum) ne peut pas être déterminé analytiquement. L'idée est de reprendre la théorie de l'évolution : sélection naturelle, croisement et mutation. L'avantage de ce type d'algorithme est qu'il va converger très rapidement dans le voisinage du maximum global, si les paramètres à optimiser sont correctement initialisé, il y a peu de chance que l'on soit bloquer vers un locale. Il existe beaucoup de versions différentes des algorithmes génétiques (il s'agit en fait d'une classe d'algorithme), on présente ici un algorithme appartenant à cette classe pour minimiser une fonction  $f(x)$  avec  $x = (x_1, \dots, x_d)$

### Algorithme Génétique

Paramètres :  $q, p, p'$ .

Initialisation : On génère une population de possibilités  $\mathbf{x}^i = (x_1^i, \dots, x_m^i)$ ,  $i = 1, \dots, P$  avec  $m$  le nombre de paramètres à optimiser et  $P$  la taille de la population. Chacun des  $\mathbf{x}^i$  est appelé un chromosome.

(1) Opérateur de sélection. On sélectionne les  $[P/2]$  chromosomes donnant les meilleurs résultats par rapport à la fonction d'évaluation, pour nous, la fonction d'utilité. Les autres chromosomes sont enlevés de la population.

(2) Opérateur de croisement. Soit  $\mathbf{x}^1 = (x_1^1, \dots, x_m^1)$  et  $\mathbf{x}^2 = (x_1^2, \dots, x_m^2)$  deux chromosomes (en tout  $q\%$  de chromosomes sont croisés), on choisit un entier aléatoirement tel que  $0 \leq r \leq m$  et,

$$\begin{aligned}\mathbf{x}^{1'} &= \{x_j \mid \text{si } j \leq r, \text{ alors } x_j \in \mathbf{x}^1, \text{ sinon } x_j \in \mathbf{x}^2\} \\ \mathbf{x}^{2'} &= \{x_j \mid \text{si } j \leq r, \text{ alors } x_j \in \mathbf{x}^2, \text{ sinon } x_j \in \mathbf{x}^1\}.\end{aligned}$$

(3) Opérateur de mutation. Soit  $\mathbf{x}^1 = (x_1^1, \dots, x_m^1)$  un chromosome (en tout  $p\%$  de chromosomes sont mutés), on choisit de muter  $p' \%$  du chromosome, on tire alors  $mp'/100 = r$  entiers aléatoirement compris entre 0 et  $N$  et,

$$\mathbf{x}^{i''} = \{w_j \mid \text{si } j \neq r, \text{ alors } x_j = x_j^i, \text{ sinon } x_j = \text{random}\}.$$

(4) on régénère  $[P/2]$  nouveaux chromosomes et on répète les étapes 2 à 3 jusqu'à convergence.

---

## 7 Références

### Références

- [AlEw07] E. Alos et C.O Ewald, 2007, *Malliavin Differentiability of the Heston Volatility and Applications to Option Pricing*, preprint, 28pp.
- [AnAn00] L. Andersen, et J. Andreasen, 2000, *Volatility Skews and Extensions of the LIBOR Market Model*, Applied Mathematical Finance 7, 1-32.

- [Ba88] Basel Committee on Banking Supervision, *International convergence of capital measurement and capital standards*, Bank of International Settlements.
- [Ba96] Basel Committee on Banking Supervision, *Amendment to the capital accord to incorporate market risks*, Bank of International Settlements.
- [Ba05] Basel Committee on Banking Supervision, *Basel II : International convergence of capital measurement and capital standards : A revised framework*, Bank of International Settlements.
- [Ba11] Basel Committee on Banking Supervision, *A global regulatory framework for more resilient banks and banking systems - revised version June 2011*, Bank of International Settlements.
- [Ba13] Basel Committee on Banking Supervision, *Basel III : The Liquidity Coverage Ratio and liquidity risk monitoring tools*, Bank of International Settlements.
- [BeBuFl04] H. Berestycki J. Busca, et I. Florent, 2004, *Computing the implied volatility in stochastic volatility models*, Communications on Pure and Applied Mathematics, 57(10), 1352-1373.
- [BeDoKo03] G. Bernis, Gobet et A. Kohatsu-Higa, 2003, *Monte Carlo Evaluation of Greeks for Multidimensional Barrier and Lookback Options*, Math. Finance. 13, 99-113.
- [BlCo76] F. Black et J.C. Cox, 1976, *Valuing corporate securities : Some effects of bond indenture provisions*, J. Finance, 31 :351-367.
- [BrGaMu97] A. Brace, D. Gatarek, et M. Musiela, 1997, *The market model of interest rate dynamics*, Math. Finance 7, no. 2, 127-155.
- [BrSc79] M. J. Brennan et E. S. Schwartz, 1979, *A continuous time approach to the pricing of bonds*, Journal of Banking and Finance, vol. 3, 135-155.
- [BrSc82] M. J. Brennan et E. S. Schwartz, 1982, *An equilibrium model of bond price and a test of market efficiency*, Journal of Financial and Quantitative Analysis, vol. 17, 303-329.
- [BrLi78] D.T. Breeden et R.H. Litzenberger, 1978, *Prices of State-Contingent Claims Implicit in Option Prices*, Journal of Business 51, 621-651.
- [BrMe01] D. Brigo et F. Mercurio, 2001, *Interest rate models/theory and practice*, Springer-Verlag, Berlin, 2001.
- [CaMa99] P. Carr et D.B. Madan, 1999, *Option valuation using the fast Fourier transform*. Journal of Computational Finance, 2(4) :61-73.

- [Ch96] L. Chen, 1996, *Stochastic mean and stochastic volatility : a three factor model of term structure of interest rates and its application to the pricing of interest rate derivatives*, Blackwell publishers.
- [Co75] J.C. Cox, 1975, *Notes on Option Pricing I : Constant Elasticity of Variance Diffusions*, Working paper. Stanford University.
- [CoInRo85] J. C. Cox, J.E. Ingersoll et S.A. Ross, 1985, *A Theory of the Term Structure of Interest Rates*, Econometrica 53 : 385-407
- [Cr03] S. Crépey, 2003, *Calibration of the Local Volatility in a generalized Black-Scholes model using Tikhonov regularization*, SIAM Journal on Mathematical Analysis, Vol. 34 No 5, pp. 1183-1206.
- [DeKa94] E. Derman, I. Kani, 1994, *Riding on a Smile*, Risk, 7(2), pp. 32-39.
- [DeGaRi05] J. Detemple, R. Garcia, M. Rindisbacher, 2005, *Representation formulas for Malliavin derivatives of diffusion processes*, Finance and Stochastics Vol 9, No 3
- [DuMaYo94] D. Duffie, J. Ma, J. Young, 1994, *Black's consol rate conjecture*, working paper, Stanford University.
- [DuPaSi00] D. Duffie, J. Pan et K. Singleton, 2000, *Transform analysis and asset pricing for affine jump-diffusions*, Econometrica, 68(2), 1343-1376.
- [Du94] B. Dupire, 1994, *Pricing with a Smile*, Risk, 18-20.
- [Fi00] D. Filipovic, 2000, *Exponential-Polynomial Families and the Term Structure of Interest Rates*, Bernoulli, 6, 1-27.
- [Fu11] M. Fukasawa, 2011, *Asymptotically efficient discrete hedging*, Stochastic analysis with Financial Applications, Progress in Probability 65 331-346.
- [FoLaLeLiTo99] E. Fournie, J.M. Lasry, J. Lebuchoux, P.L Lions et N. Touzi, 1999, *Applications of Malliavin Calculus to Monte Carlo Methods in Finance*, Finance and Stochastics, 3, 391 - 412.
- [FoLaLeLi01] E. Fournie, J.M. Lasry, J. Lebuchoux, P.L Lions, 2001, *Applications of Malliavin Calculus to Monte Carlo Methods in Finance*, II. Finance and Stochastics. 5, 201-236.
- [Ga06] J. Gatheral, 2003, *The Volatility Surface : A Practitioner's Guide*, John Wiley & Sons.
- [Ga04] J. Gatheral, 2004, *A parsimonious arbitrage-free implied volatility parameterization with application to the valuation of volatility derivatives*, Presentation at Global Derivatives.

- [GaJa12] J. Gatheral and A. Jacquier, 2012, *Arbitrage-free S.V.I. volatility surfaces*. Preprint.
- [GoKo03] E. Gobet et A. Kohatsu-Higa, 2003, *Computation of Greeks for Barrier and Lookback Options using Malliavin Calculus*, Electron. Comm. Probab. 8, 51 - 62.
- [GoLa14] E. Gobet et N. Landon, 2014, *Almost sure optimal hedging strategy*, Annals of Applied Probability, Vol. 24(4), pp.1652-1690.
- [HaKuLeWo02] P. Hagan, D. Kumar, A. Lesniewski, D. Woodward, 2002, *Managing smile risk*, Wilmott Mag. 1, 84-108.
- [HaWe08] P. S. Hagan et G. West, 2008, *Methods for constructing a yield curve*, Wilmott Magazine, 70-81.
- [HaWe08] P. Hagan et E. Woodward, 1999, *Equivalent black volatilities*, Applied Mathematical Finance, 6, pp. 147-157.
- [HeJaMo90] D. Heath, R. Jarrow, et A. Morton, 1990, *Bond Pricing and the Term Structure of Interest Rates : A Discrete Time Approximation*. Journal of Financial and Quantitative Analysis, 25 :419-440.
- [HeJaMo91] D. Heath, R. Jarrow, et A. Morton, 1991, *Contingent Claims Valuation with a Random Evolution of Interest Rates*. Review of Futures Markets, 9 :54-76.
- [HeJaMo92] D. Heath, R. Jarrow, et A. Morton, 1992, *Bond Pricing and the Term Structure of Interest Rates : A New Methodology for Contingent Claims Valuation*. Econometrica, 60(1) :77-105.
- [He93] S.L. Heston, 1993, *A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options*, The Review of Financial Studies, Vol. 6, num. 2, pp. 327-343.
- [Ho93] M. Hogan, 1993, *Problem in certain two factor term structure models*, Annals Appl. Probability, vol 3, no 2, 576-592
- [HuWh90] J. Hull. et A. White, 1990, *Pricing interest-rate derivative securities*. Review of Financial Studies 3 (4) : 573-592.
- [JaSuHo99] N. Jackson, E. Suli, and S. Howison, 1999, *Computation of deterministic volatility surfaces*, Journal of Computational Finance, 2, pp. 5-32.
- [KaJeSh98] N. El Karoui, M. Jeanblanc-Picqué et S. Shreve, 1998, *Robustness of the Black and Scholes formula*, Mathematical Finance, 8(2), p. 93-126.
- [La05] P.H. Labordere, 2005, *A general asymptotic implied volatility for stochastic volatility models*, preprint, 35pp.

- [La08] P.H. Labordère, 2008, *Analysis, geometry, and modeling in finance : Advanced methods in option pricing*, CRC Press, 402pp.
- [Lee04] R. Lee, 2004, *The moment formula for implied volatility at extreme strikes*, Mathematical Finance 14(3) : 469-480.
- [LaOs97] R. Lagnado et S. Osher, 1997, *A Technique for Calibrating Derivative Security Pricing Models : Numerical Solution of an Inverse Problem*, Journal of Computational Finance, 1, 1, pp. 13–25.
- [LoSc92] F. Longstaff et E. Schwartz, 1992, *Interest rate volatility and the term structure : a two factor general equilibrium model*, Journal of Fixed Income Models, 7-14.
- [MaPa99] C. Martini et C. Party, 1999, *Variance optimal hedging in the Black-Scholes model for a given number of transactions*, INRIA Rapport de recherche, no. 3767.
- [Me74] R. Merton, 1974, *On the pricing of corporate debt : the risk structure of interest rates*, J. of Finance, 3 :449-470.
- [NeSi87] C. Nelson et A. Siegel, 1987, *Parsimonious modeling of yield curves*, J. of Business 60, 473-489.
- [Nu06] D. Nualart, 2006, *Malliavin Calculus and Related Topics*, 2nd Edition. Springer-Verlag.
- [Ok97] B.K. Øksendal, 1997, *An Introduction To Malliavin Calculus With Applications To Economics*, lecture notes, Dept. of Mathematics, University of Oslo, 83pp.
- [Pr12] N. Privault, 2012, *An Elementary Introduction to Stochastic Interest Rate Modeling*, Second Edition, Advanced Series on Statistical Science & Applied Probability, 16, World Scientific Publishing Co., Singapore.
- [Ri78] S. Richard, 1978, *An arbitrage model of the term structure of interest rates*, Journal of Financial Economics, vol. 6, 33-47.
- [ScSc84] S. M. Schaefer et E. S. Schwartz, 1984, *A two factor model of the term structure : an approximate analytical solution*, Journal of Financial and Quantitative Analysis, vol. 19, no. 4, 413-424.
- [Sv94] L. E. O. Svensson, 1994, *Estimating and interpreting forward interest rates : Sweden 1992-1994*, IMF Working Paper No. 114.
- [Va77] O. Vasicek, 1977, *An Equilibrium Characterisation of the Term Structure*. Journal of Financial Economics 5 (2) : 177-188.

- [Wa97] D. Waggoner, 1997, *Spline methods for extracting interest rate curves from coupon bond prices*, Federal Reserve Bank of Atlanta, Working Paper series, 97-10.
- [Wa84] S. Watanabe, 1984, *Lectures on Stochastic Differential Equations and Malliavin Calculus*, Springer, New Yor, 113pp.
- [WuZh02] L. Wu et F. Zhang, 2002, *Market Model : from Deterministic to Stochastic Volatility*, Working Paper, Claremont Graduated University and Hong Kong University of Science and Technology, pp 29.