My journey through finance and stochastics

Marek Musiela[1]

Abstract

This year, *Finance and Stochastics* celebrates its 25th anniversary. The journal provides a platform for the community of researchers on which they can publish their ideas and results.

Publication is an outcome of research which may be conducted for a number of years before it reaches the required maturity. I find this research process to be very important. Unfortunately, it is almost impossible to decode it from reading the research publications. This special issue of *Finance and Stochastics* gives me an opportunity to focus on it. I am grateful I can present my personal memory of this process. Understanding why questions are asked and how the answers are found is critical.

Keywords Financial markets · Valuation and risk management of financial assets · Quantitative finance · Stochastic calculus · Mathematical models · Interaction between academia and industry · Research and education

Mathematics Subject Classification (2010) 91-03 · 01A99 · 91G99

JEL Classification B26 · B23 · N01 · N20

1 Introduction

When reflecting on my scientific and business experience with finance, I am often surprised by how much of it was underpinned by pure chance. I wonder if this is because I have interacted with a large number of people working in this field, both from academia and industry, whom I met largely by chance? In any case, this interaction proved critical for me. It was a driving force of my journey through finance and stochastics. It was the main reason why I worked on many diverse projects. Some results of my research are published in scientific journals. Others were only presented at academic and industry conferences. However, the *why and how* was often hidden from the audience. Herein, I focus on these aspects of my research.

✉ M. Musiela
  mgmusiela@gmail.com

[1] 8 Chuter Street, McMahons Point, 2060 NSW, Sydney, Australia

2 Discovering finance

It must have been late November or early December of 1985 when I received a phone call from the head of the School of Mathematics and Statistics at the University of New South Wales, asking me if I would volunteer to supervise a student over summer on a so-called vacation scholarship. Given that I had joined the school only in September of that year, I felt he made me an offer I could not refuse. The student already had a subject, she wanted to learn about options. At that time, I had no idea what options were. I borrowed a book from the library and learned what a call option was, how one would define the option payoff and what the price could be. The derivation of the option price used arguments of the second order Taylor expansion. I was not completely happy with it. Luckily, the student was happy with the project.

In March 1986, I kept teaching courses on probability and stochastic processes and came back to my research on estimation, filtering and control of stochastic processes. I was working in this field from the very beginning of my scientific career. I was inspired by work of A. N. Shiryaev and his collaborators, and also by the French School of stochastic calculus. In 1980, I moved from Wrocław in Poland to Grenoble in France to continue working with A. Le Breton with whom I had collaborated for some time. There, I had an easy access to the scientific journals published in English, which was not so easy back then in Poland. One of them, which I consulted regularly, was *Stochastic Processes and their Applications*. I have published some of my research there.

I remember reading the volume in which the paper by M. Harrison and S. Pliska [2] was published in 1981. I did not understand what the paper was about. I did not know anything about options. Why should I? I lived in Grenoble, the capital of the French Alps, and not in a financial centre, where people knew or wanted to learn about options.

Forward 6 years in Sydney, some time during 1986, and after the project with the student who wanted to learn about options, I rediscovered the paper by Harrison and Pliska. At that time, I already knew something about options and how they were used in practice. When I was reading this paper again, after all these years, I could not believe the fit between stochastic calculus and finance. It felt like martingale theory, Itô's formula, Girsanov's theorem were all developed to be applied in finance. At this very moment I got hooked. I knew that for many years to come, I would work in this emerging new field on the crossroads between stochastics and finance.

When I think of the road I took to get to work in finance and stochastics, I only see random events, not clear road signs. Of course, I did not learn stochastic calculus because it would later become useful in finance. Also, I did not learn finance because I knew I could apply stochastic calculus there. I just was in the right place at the right time. I did not plan to educate myself for a career in finance. I was probably not the only one who had such an experience. There were a handful of pioneers who ended up in the same place, also by chance.

Very soon it became clear that there was a big push to understand how stochastic calculus was related to and could be applied in modern finance. In the city of Sydney, with a small but relatively developed financial market, the appetite was big. The Sydney Futures Exchange, which opened in 1979, was the first exchange outside of the

US to offer a financial futures contract on a 90 day bank bill. Banks were interested in the development of new products. Software companies started to develop pricing and risk management tools. The demand for knowledge was huge, but the supply of people who had it was very limited. I started to work with banks and software companies on the pricing and risk management of exotic options. Back then, pricing a simple barrier option was a significant challenge. Only few people knew how to use the reflection principle to price a barrier option in closed form.

Thanks to the contacts with the industry, it became clear to me that a course in financial mathematics was needed. People I interacted with wanted to learn the concepts and techniques that were relevant to their jobs. It was very gratifying to see that such an abstract field like stochastic calculus became so central for many people working in the finance industry.

3 Teaching a course and working on a book

I was teaching courses on probability and stochastic processes as part of the Masters of Statistics program. I had developed a new course on financial mathematics and offered it as part of the Masters. The course was also offered to people who did not study for this degree but had sufficient background in mathematics to be able to deal with the challenging syllabus. In particular, one needed to know some probability and stochastic processes theory. I was teaching the course in the evening so people working in the industry could attend. In the first year, which must have been in 1987 or 1988, over 60 people enrolled. For a city like Sydney, with a relatively small financial market, such an enrolment was phenomenal.

Every year we had a significant flow of students who wanted to study financial mathematics. Enrolment numbers decreased but stabilised at around 25 to 30 students. Such numbers were big, when you compare with the enrolment in other graduate courses. The demand persisted for at least a decade or more. When I was leaving Sydney for London, the course was still going strong. Clearly, a new field of finance was created which needed people who could understand the language of modern finance. Very similar developments took place in France, where education in financial mathematics was offered initially at Paris 6 by Nicole El Karoui. In parallel, I was teaching a course I developed for UNSW in Sydney at ENSIMAG in Grenoble. I worked in ENSIMAG before moving to Sydney in 1985 and I was quite happy to help. Every Australian summer, in January and February, and up until 2000, when I took a job in London, I would go to Grenoble to teach financial mathematics to engineers.

Courses in financial mathematics were offered in many places. I have no knowledge of what was happening in the US in the early stage of the development of the field. I know about Poland. The reason is quite simple. I had an ARC research grant on estimation, filtering and control of linear systems driven by stable noises. This must have been around 1987. The grant gave me an opportunity to recruit a research fellow to work on the project. I managed to attract Marek Rutkowski, who came to Sydney to work with me. Those who ever applied for scientific grants know very well that you spend plenty of time writing grant proposals. You apply and might not get

it. You reapply and you have no idea what will happen. Your institution expects you to keep applying. In the end, you may be successful but it may turn out that at this stage you are no longer interested in the research you proposed. This was the case of my grant. It came when I was already working full steam in financial mathematics. The grant proposal was really interesting but, in all honesty, I cannot say I dedicated to it sufficient time of my own. Marek did all the work, and in the meantime he was also learning financial mathematics. When the grant finished, Marek returned to Poland and started to teach at Warsaw University of Technology a course on financial mathematics.

Marek came to Sydney several times on different contracts before settling here permanently. We worked jointly on a number of research projects. I write more about it later. The one project I want to mention now is the book we wrote together — *Martingale Methods in Financial Modelling*, published by Springer in 1997.

I had no intention to write a book. My focus was on teaching courses and working with the industry. The years went on, Marek was teaching in Warsaw, I was teaching in Sydney and Grenoble. The notes we used for the courses developed over time. A one-term course became two- and then three-term. One could be offered to undergraduates and two to postgraduate students.

During one of his multiple visits to Sydney, Marek proposed that we write a book based on the lecture notes we had developed. I agreed but it was Marek who took the responsibility of putting it all together. It was not a simple task. It took about two years before we had something presentable to a publisher. It is worth pointing out that back then, not much was available in a book format addressing the topic the way we wanted. Our book covered many subjects, interest rate modelling included, and it had an appendix covering major tools from probability and stochastic calculus. In that sense, it was quite self-contained. One could learn modern finance from it without the need to first digest the difficult field of stochastic calculus.

Inspiration for writing the first chapter of our book came from the way the first chapter was written by A. N. Shiryaev in his book *Probability*, published by Springer in 1980. There, the most important and also difficult concepts of probability theory were presented on a finite probability space, where all mathematical difficulties are removed and the ideas and techniques become accessible at the undergraduate level. Marek and I did exactly the same. We used finite probability spaces and introduced concepts from finance and probability to explain the idea of pricing under the assumption of no arbitrage. Mathematical difficulties are removed so that one could focus on the concepts and ideas behind the theory, by now accessible at the undergraduate level. In the chapters that followed, the models became more advanced.

The book was relatively well received and up until today is considered to be one of the best texts in financial mathematics. I personally think the reason for it is the combination of rigorous mathematics with focus on the practical problems we encountered when working with the industry. We were not trying to produce the most general results possible. We were focused on the ideas and techniques that could be useful in practical applications. For example, we priced a number of stock and FX options under a lognormality assumption for the underlying asset because such a model framework was used in practice at that time.

Beyond the classical models used for pricing of options in equity and foreign exchange markets, the book also presented various approaches to deal with interest

rate risk. This was quite important because at that time, modelling interest rate risk was a domain of very active research, published in scientific journals but not in a book format. Because both of us were involved in research on interest rate modelling, it was quite natural for us to include this new topic.

4 HJM approach to interest rate risk

One day, soon after I started teaching a course on financial mathematics at UNSW, a gentleman showed up in my office. He introduced himself very politely, and said he heard about me from some people in the city. Clearly, he must have been talking with someone who attended my course or simply heard about me from someone else. He came with a well-formulated question which required some knowledge of stochastic calculus. This is how my collaboration with Alan Brace started. Alan worked as a quant with a group of traders taking positions in the interest rates futures at the Sydney Futures Exchange.

More or less at the same time, an early draft of the D. Heath, R. Jarrow and A. Morton paper [3] on interest rate modelling was already circulating. It was well before it was published in 1992 in *Econometrica*. If I am not mistaken, it took 5 years to publish the paper, so it must have been around 1987. I had a preprint from a software company who in turn got it from a well-connected place in New York. The software company wanted to implement the HJM model and asked me to help them. There were not many people in Sydney at that time who could understand and implement the HJM framework. My role was to understand the model, the software company was to take care of its implementation. Banks were also interested in the HJM model.

It turned out that the simplest way to understand and implement the HJM model was to assume that the volatility of the instantaneous forward rates is deterministic. Instantaneous forward rates are the fundamental building blocks of this framework. From them, one could recalculate the zero-coupon rates and zero-coupon bond prices. When the instantaneous forward rates have deterministic volatility under the historical measure, they become Gaussian under the risk-neutral measure. The price processes of the zero-coupon bonds then follow lognormal processes. This is the reason why we initially worked under the assumption of a deterministic volatility. There were many nice consequences of this assumption. One could price many options, but also forwards, swaps or futures on stocks and FX, in closed form. For example, a call option on a stock was still given by the Black–Scholes formula even under the assumption of a stochastic interest rate. One only needs to interpret differently the model inputs. In particular, the volatility now had to be seen as the aggregate volatility of the forward prices on the stock. The implied volatility was not the volatility in the classical Black–Scholes model any more. For longer maturity options, it incorporated the risk of changes in the short term interest rate.

The only really disappointing feature of Gaussian HJM was the way it was pricing LIBOR options. There was a real disconnect between the approach adopted by the market and the caplet formula (call option formula on a forward rate) derived within a Gaussian HJM framework.

From the very beginning, I was not that happy with the way the HJM framework was set up. Being a mathematician, working with stochastic processes and using tools

from stochastic calculus, I had a habit to analyse the processes I worked with from the perspective of their dynamics. In the case of the Black–Scholes model, it was clear that I had to deal with a Markov process with positive values. The transition probabilities defined the joint distribution along the paths, and the infinitesimal generator of the process built a link between the stochastic and analytic descriptions of the movement. The Feynman–Kac formula connected expectations with the solutions to the relevant PDEs. All this was possible because it was clear what was the state space in which evolutions of prices were taking place. But this was not the case for the HJM model.

As in the Black–Scholes model, where the underlying was a price process, I wanted to understand what was the underlying analogue in the HJM framework. Intuitively, I expected a yield curve or a forward curve to represent the underlying. However, within the HJM model, on every transition from one day to the next, they both became shorter by a day and hence the state space was changing with each transition. Luckily, there was a simple fix to the HJM parametrisation, which showed that in fact the HJM model could be viewed as an infinite-dimensional evolution equation; see Musiela and Sondermann [13] and Vargiolu [20]. After this fix, when reduced to the case of Gaussian HJM, it became an infinite-dimensional Ornstein–Uhlenbeck process with a drift containing a first derivative operator. The associated semigroup generates the shift of the curve to the left. I explained all this in the paper "Stochastic PDEs and term structure models", which was presented at many conferences but was never published, except for the proceedings from a conference in La Baule; see Musiela [9]. For the first time, I presented the paper at the Stochastic Processes and their Applications conference in Amsterdam probably in 1990 or 1991. There, it attracted the attention of several well-known mathematicians who had already published seminal results on martingale theory and pricing in the absence of arbitrage. My talk focused on the case of a deterministic volatility. The evolution of the forward curve under the risk-neutral distribution was studied in some detail. Many results were very intuitive and comparable with the analogous finite-dimensional systems, where the first derivative operator is replaced by a matrix.

Publication in a peer-reviewed journal proved to be more difficult. After reading the reviews from reputable journals, I gave up on publishing this paper. Clearly, it did not fit anywhere — it was too mathematical for some and incomprehensible for others. Luckily, this paper was picked up later by many researchers and the reparametrisation of the HJM model, dubbed as the Musiela parametrisation, was studied in detail by many mathematicians. Cases of random volatility were studied and understood well. Thankfully, no real damage was done by giving up on the publication of this paper. The practical implications of this research were significant. Conditions under which the evolution takes place in a finite-dimensional manifold help to construct a more robust yield curve, which in turn leads to better risk management.

The HJM dynamics also helped to understand how to construct a yield curve from a finite number of points, given by the prices of securities traded in the market. Clearly, the HJM framework was very useful in a number of situations. However, it could not easily replicate the market method of pricing caps and swaptions. One of the reasons was the fact that the HJM framework had an implicit assumption of trading a continuum of discount bonds. This had rung alarm bells for me for all approaches to term structure modelling, those either based on the specification of the

short-rate dynamics or on the HJM framework. Both of them were leading to a continuum of discount bonds. Recall that at that time, we only had a theory of arbitrage-free pricing for models with a finite number of assets.

What does it mean to trade a continuum of assets? How do you define self-financing portfolios? All these questions had to be addressed before we could believe the theory. What I am trying to say is that from a purely mathematical point of view, there were some questions that needed to be answered in order to fully connect the HJM framework with the theory of pricing in the absence of arbitrage that involved only a finite number of assets. This comment is more for the connoisseurs of the mathematical formalism, because it has more to do with the Kolmogorov measure extension theorem, measure-valued self-financing strategies, etc., than with the practicalities and the use of models in applications. For me personally, it was important to understand the theory because I could see how people in the industry were happy to twist it and cut corners to get what they wanted from the business perspective. This disconnect and the pressure that came with it existed for decades to come, and I suspect it is present until today, but perhaps it is focused now on different questions. The reason behind is the push to convert theoretical developments and knowledge into something that generates positive monetary outcomes. This is why I think it is critical to make sure that the mathematics of the models used for pricing and risk management is correct.

5 The market model of interest rate dynamics, also known as BGM

I was working with Alan Brace on yield curve related projects. We had already implemented the Gaussian HJM, but the traders did not like the pricing of caplets that the model was producing. Instead, they were putting forward LIBOR rates into the Black–Scholes formula, for pricing options on LIBOR, and multiplied the result by the discount bond for the settlement date. This procedure was adopted globally by all desks trading caps and swaptions. However, I was not aware of any scientific justification for doing so. It seemed traders considered all forward LIBOR rates as individual assets to price options on them. This was obviously a problem because there are clear links between the LIBOR rates, well described by term structure models. Some people in the academic community held the view that the market procedure was wrong and could lead to arbitrage. I did not share this view. My first reaction was to try to understand why they did what they did. I just did not believe they all could be wrong. Indeed, they competed against each other and some of them would probably find arbitrage if there was one. This, in turn, would probably destabilise the market and lead to a better valuation method, for example based on the Gaussian HJM. I concluded that the market may not know why what they do is scientifically correct, but definitively it is not wrong to the point to admit arbitrage. This explanation was much more plausible for me and proved correct later on. Motivated by the experience of other people, I was looking for ways to explain the market practice within an arbitrage-free term structure model.

Option prices within the Gaussian HJM framework were given by the expectation of the payoffs discounted with the savings account. Under the risk-neutral measure,

such calculations were often rather tedious. A remarkable trick was invented that consisted of changing the probability measure under which the expectation of the payoff, without discounting, would need to be calculated and the result would be multiplied by the zero-coupon bond for the option maturity. This new measure was dubbed the forward measure.

The credit for it is often attributed to H. Geman; however, N. El Karoui should also be mentioned for her input. Working with Alan in Sydney, we came up independently with the same idea. Moreover, about the same time, F. Jamshidian explained how one can adjust the short rate in order to eliminate the discounting. His idea is in fact a particular case of a measure change. I recall my meetings in New York with F. Jamshidian during which I explained the forward measure trick, and in Paris with N. El Karoui, where we discussed this technique in detail.

The forward measure transformation worked under an arbitrary volatility of the HJM framework. Forward prices are martingales under the forward measure. Of course, discounted spot prices are martingales under the spot risk-neutral measure. These properties became important building blocks in the development of the market model of interest rate dynamics.

For me, it was not only New York and Paris that were very important places during the early stage of development of the conceptual framework of arbitrage-free pricing. Bonn played a very important role as well. In fact, I started to visit Bonn before I began working in finance. I worked with N. Christopeit on various questions in stochastic processes theory. Towards the end of our collaboration, during one of my visits to Bonn, we finished a project motivated by finance. I presented the results at the departmental seminar, after which I met D. Sondermann who was in the audience. Over the years, Dieter came twice to Sydney and I visited him in Bonn several times. We published several papers together, all related to the dynamics of interest rates. His research group in Bonn was very much focused on this topic. There was a very natural fit because of the different backgrounds of people working in both groups. We exchanged information on the research progress we made, and the ideas developed almost independently in both teams, until the next exchange, most of the time related to mutual visits.

Both teams worked on the development of a model which would be consistent with the way the market was pricing caplets. In Sydney, we relied on the forward measure and produced a term structure model within the HJM framework which returned the caplet formula used by the market. In Bonn, an equivalent result was obtained, but the main argument used an explicit solution to a PDE. The Sydney paper by A. Brace, D. Gątarek and M. Musiela was published in *Mathematical Finance* with the title "The market model of interest rate dynamics" [1], and the Bonn paper by K. Miltersen, K. Sandmann and D. Sondermann was published in the *Journal of Finance* with the title "Closed form solutions for term structure derivatives with log-normal interest rates" [8]. By choosing our title, we wanted to express the view that we only explained what the market was doing anyway. The model presented in our paper was dubbed very soon the *BGM model*.

You know by now what B and M stand for in BGM. How about G? Well, when the model was defined, I wanted to analyse its properties. One important property of term structure dynamics is ergodicity. There is a general and empirically supported belief

that interest rates are mean-reverting. In mathematical terms, one could say that interest rates are stationary or ergodic. From the perspective of a re-parametrisation of the HJM model, this would correspond to ergodicity of the infinite-dimensional evolution equation. I analysed the ergodicity of the Ornstein–Uhlenbeck process corresponding to the deterministic volatility of the HJM model. However, in the BGM model, the HJM volatility was random. I wanted to study ergodicity in this case as well, but we had so much work to do that we clearly needed help from an expert. There was one available, just next door, working with a colleague, B. Goldys, on a research grant. His name was D. Gątarek and so the G in BGM stands for Gątarek. Dariusz proved ergodicity of the BGM dynamics under the so-called Musiela parametrisation of the HJM framework.

Implementation of any term structure model is a challenge. This is true for models derived from the specification of the short rate as well as for the HJM model, and of course also for the BGM model. There is an extensive literature on this subject. It is important to notice that in each of these models, one could talk about the instantaneous short rate and the savings account it generated. This was a potential problem if one wanted to discretise the tenor structure for the model implementation. This is why we started to work on the next generation of term structure models, namely, models which had only a finite number of discount bonds available for trading. Until then, existence of a riskless asset, the so-called savings account, was assumed in the vast majority of models used for pricing options on interest rates, equity and foreign exchange. It was not clear at all how to remove the riskless asset from the model specification. Note, however, that from the general theory point of view, there was no problem. One could assume trading in a finite number of securities, choose one of them as a numeraire, discount all others with it and account in the units of the numeraire asset. Simpler said than done because at that point in time, changes of units were not generally used and understood. We were still learning to live with the spot and forward risk-neutral measures. The link between them was given by the savings account, which in turn was defined by the instantaneous continuously compounded short rate.

6 Discrete-tenor term structure models

So how do you construct a term structure model with only a finite number of zero-coupon bonds, and how do you eliminate arbitrage from it? Recall how this was done within the HJM framework. The defining quantity was the instantaneous forward rate for a fixed maturity. You assume, for example, Itô dynamics for the forward rate processes. The short rate is the forward rate for the 'next' maturity. The savings account is defined in the usual way. The zero-coupon bond prices are defined in terms of the forward rates. You discount the zero-coupon bonds with the savings account and apply Itô's formula to the ratio. In turn, you ask under what conditions on the drifts in the forward rates you can construct a new measure under which the ratio becomes a martingale. You assume such a form for the drifts in the forwards and you are done. Well, not exactly, because back then we only had models with a finite number of assets in them. Mathematicians fixed this problem later on by developing a theory which allowed trading in a continuum of discount bonds.

Of course, in the discrete-tenor case, there is no savings account; so you cannot discount your bond prices with it. A way to deal with this is to choose one discount bond as a numeraire asset. For example, you can take the one with the longest maturity. The construction works by backward induction. It is more convenient to work with the forward LIBOR rates than with the forward prices. This is, after all, how the market looked at the problem. It assumed lognormal forward LIBOR rates and applied the Black–Scholes formula to price a call on LIBOR. So you postulate dynamics for the 'last' forward LIBOR rate. To define it, you need to use the last two discount bonds. Then, you use the volatility of the 'last' forward LIBOR and the LIBOR itself to define the volatility of the forward price to the last maturity of the penultimate discount bond. This allows you to construct a forward measure for the penultimate maturity and a Brownian motion under this measure by shifting the Brownian motion used in the construction of the 'last' forward LIBOR rate. To construct the penultimate forward LIBOR rate, you only need to choose its volatility. Proceeding in this way, you construct the joint dynamics of all forward LIBOR rates. The details of this construction and much more can be found in my paper with Marek (Rutkowski), "Continuous-time term structure models. Forward measure approach", published in *Finance and Stochastics* in September 1997 [12]. The construction may look simple and intuitive. In fact, this is not how things happened. Earlier we had tried many alternative constructions, but all had some problems. Finally, Marek came up with the one which was published.

When the first version of the paper was completed, I sent it to Farshid (Jamshidian) for comments. A couple of months later, he sent me the first version of his paper "LIBOR and swap market models and measures", which was later published in the same volume of *Finance and Stochastics* [4]. The published versions of both papers differ very significantly from the initial ones. Unfortunately, both were stripped of information on how we collaborated on this topic. Often, published versions become dry and formal and the human dimension of the interaction evaporates. This is quite unfortunate.

The other two papers published in volume 1, issue 4, of *Finance and Stochastics* are also directly related to the development of the LIBOR and swap market models. B. Goldys, a member of the Sydney team, in "A note on pricing interest rate derivatives when forward rates are lognormal", derives pricing formulae for contracts written on zero-coupon bonds for lognormal forward rates. His method is purely probabilistic, in contrast to the earlier results obtained by Miltersen, Sandmann and Sondermann which used the solution to a PDE. S. Rady, associated with the Bonn team, in "Option pricing in the presence of natural boundaries and a quadratic diffusion term", uses a probabilistic change-of-numeraire technique to compute closed-form prices of European options to exchange one asset against another. One can easily claim that *Finance and Stochastics* published almost all important contributions to the development of the LIBOR and swap market models.

The market uses term structure models to price more exotic options, relatively to the prices of caps and swaptions. Lognormality of the forward LIBOR rates was important for the consistency with the method of pricing caps used in the market. The discrete-tenor term structure models are important from the practical perspective. Simulation algorithms are written for the finite-dimensional systems of stochastic

differential equations that discrete-tenor models generate, and not on the basis of continuous-tenor term structure models like HJM or BGM.

Earlier versions of the published papers were presented at a number of conferences. I should like to mention one particularly important event which took place in Cambridge, UK, at the Isaac Newton Institute for Mathematical Sciences. From January 1st, 1995 to June 30th, 1995, there was a programme dedicated to Financial Mathematics. From June 12th to June 16th, there was a meeting with special emphasis on the term structure of interest rates organised by D. Duffie. It was a perfect venue to present our research and meet people.

During my numerous business trips and because of many academic and industry conferences I attended, I met a very large number of brilliant individuals working in quantitative finance. All of them helped me progress in this difficult field by sharing their views and knowledge. A big *thank you* goes to them all. Two conferences, in particular, spring to my mind, because I met there people with whom I worked on many research projects later on. One took place in Hong Kong, where I met J.-M. Lasry and P.-L. Lions, the other at Princeton, where I met T. Zariphopoulou.

7 From academia to industry

As I already mentioned, I was involved in a number of consulting projects from a very early stage of my work in finance. A step change, however, took place after meeting Jean-Michel Lasry and Pierre-Louis Lions in Hong Kong. Jean-Michel was looking after the quantitative research for Paribas, a small French investment bank based in London. Pierre-Louis was based in Paris but worked closely with Jean-Michel. Jean-Michel asked if I could start working with Paribas on some research projects, for example on the implementation of the market models. There was a small problem, however — I was based in Sydney and they worked between London and Paris. The only solution we came up with was for me to travel to London for two weeks every six weeks.

Thanks to my contacts with the industry, I discovered that very significant research ideas were developed by the people working within the finance industry. I was working at the UNSW in Sydney and my understanding of how knowledge spreads was based on the belief that things were first discovered at universities and later utilised in a commercial environment. This may be true for many disciplines, but it definitely was not the case for finance. When I was talking with the desk quants, sitting on the trading desks, or with the researchers working at a small distance from the trading, I was very surprised to see books on stochastic calculus and probability they were consulting when working on various problems. There, it was a community capable to understand complex mathematics and use it in a commercial environment. Never before did I see how close the latest scientific discoveries and ideas came to their practical applications. This was truly inspirational and fascinating.

I have already mentioned Farshid (Jamshidian), whom I met during one of my trips to New York. On the same trip, I also met with F. Black. I recall a conversation we had in his office after I explained to him what I was working on. He asked if I believed it was better to conduct research at a university or in a bank. For me, it was

obvious then that a university was a much better place. However, just a couple of years later, I had changed my mind. When you work as a consultant to the industry, you have no authority to implement what you consider to be the best practice. Your best ideas and efforts are also exposed to risk generated by internal politics and power struggles. I had many such negative experiences in the past. I concluded that one had to become a full-time member of a management team in order to be able to push the best practice of growing business thanks to the quantitative ideas. It became clear that the choice was to continue to have fun and play with mathematics, no responsibility attached, or to take the risk and the responsibility that comes with it. I continued to play for some time.

The years went by and it became obvious I was spending too much time on planes. I decided to move back to Europe. Part of my family was there and I was doing a lot of work in London. Still I was not in the mood of entirely leaving the academic life. I was offered a position at the University of Geneva and the plan was to move there during the northern summer of 2000. The academic year in Sydney ended in December 1999 and I had 6 months before starting in Geneva. Jean-Michel suggested that I go to London for this period.

In January 2000 I left Sydney for the US, where I was visiting Madison before going to New York to meet the research team of Paribas there. It turned out that for technical reasons, I had to stay in New York until Paribas arranged a visa for me to work in the UK. This took 6 weeks or a bit more.

At the same time, a major consolidation of French investment banking was under way. To cut a long story short, BNP decided to take over Paribas and Société Générale in one go. The bid for Paribas was successful and the one for Société Générale failed. A new company, called BNP Paribas, emerged as a result of this takeover (dubbed *merger* by the senior management of BNP). Such events typically generate a lot of uncertainty and movement of people. It turned out that the BNP and Paribas merger was particularly successful with a very limited disruption to the business and a minimal loss of people. However, during the merger, Jean-Michel decided to move back to Paris and I agreed, for the newly created BNP Paribas, to take on the role of the head of research, based in London, for interest rates, credit and foreign exchange. Stephane Tyc, who was the head of research for equities and commodities, was based in Paris.

This is essentially how I ended up at BNP Paribas, a sequence of random events that led me to renounce the offer from the University of Geneva and accept the offer from BNP Paribas. I did not plan to leave academia and join the industry. This was just a natural consequence of a sequence of unpredictable events. As part of the deal, I was to teach a course in Geneva until they could find a replacement, which of course I did.

Thanks to my previous frequent trips to Paribas in London, I knew many people in the business and all people in the London and New York research teams. I was also familiar with the job requirements. I agreed to take on this responsibility because I understood the challenge and considered the environment as the best I had seen. Nevertheless, I cannot say it was easy to make the transition from academia to industry. It took several years before I felt really comfortable in my new role. Again, I was very lucky because many people both from the business and the research sides helped me a lot in this transition. My big *thank you* goes to them all.

When I was still in New York waiting for the working visa for the UK, one of the business heads gave me some handwritten notes by Pat Hagan on the derivation of the approximate volatility formula for the so-called SABR model. It turned out that Pat worked for Paribas New York, but left before I had arrived. Later on I found out that Bruno Dupire worked for Paribas in London and had left when Jean-Michel arrived. Obviously, I was coming to a place with a history of extremely strong research. This could be seen in the way the business treated research. Trading, sales and research were part of what was called the front office.

The approximate implied volatility formula represented well the market smile. Very soon, it was adopted as the market standard for quoting caps and swaption volatilities. But from the perspective of term structure models, this presented another difficulty. Ideally, one would need to replace the lognormality of LIBOR rates of the discrete-tenor market model by the distribution generated by SABR dynamics. This made implementation of such a model very inefficient. Other methods and models were developed to deal with this issue.

8 Research in the industry

Paribas had a culture of using the latest quantitative ideas in support of business development. This continued to be the case for BNP Paribas. We had negotiated a budget for consultants, from which we also supported academic conferences.

Consultants played many different roles. Some worked on long-term research projects with teams supporting various business lines. Others were used on one-off projects because of their expertise in a particular area. Below, I talk about longer-term projects I was personally involved in.

Pierre-Louis worked with Jean-Michel before I joined Paribas. He kindly agreed to continue consulting with BNP Paribas after Jean-Michel left for Paris. It was a real privilege to work with such a brilliant individual. Again, I was very lucky. I could interact with someone who would almost instantly have answers to many mathematical questions I had, or at least had an idea how to approach them. The challenge I had was to formulate a problem that would be sufficiently important from the practical perspective and at the same time intriguing and mathematically nontrivial so that Pierre-Louis would be keen and excited to think about. I keep very fond memories of many discussions we had related to the challenges for the pricing and risk management of options generated for institutions like BNP Paribas, and how one could try to deal with them using a mathematical formalism. Below I talk about one of them in greater detail.

8.1 SABR model

Pat (Hagan) had already developed a formula to approximate the implied volatility, using methods which required the diffusion in question to be elliptic, which is not the case for the SABR dynamics. The question was if the approximation would still work for such dynamics or if it could fail in certain regimes. It turned out that the additional correction terms, missing in the formula, were small and would not change anything

from the practical perspective. This was good news, because the formula was already used in the market to quote cap and swaption volatilities.

The SABR model itself presented many mathematical challenges. To deal with them, we had conducted an extensive mathematical analysis to see if we did not face regimes of parameter values for which there could be problems with the model behaviour.

We had characterised the existence of moments, which was important for the valuation of CMS-related products. It turned out that the existence of moments was dependent on the correlation parameter. The more negative the correlation, the higher moments would exist.

We showed that convexity is preserved if and only if the correlation between the Brownian motions is negative. Convexity preservation is important because it means that if you start with a convex payoff, its arbitrage-free price, given by the risk-neutral expectation of the payoff, will be a convex function of the underlying asset. From the practical perspective, this leads to a natural interpretation of what it means to be long or short gamma. For one-dimensional diffusions, this convexity preservation is always true, thanks to the maximum principle. However, when models are diffusions but in a higher-dimensional space like in the case of SABR, this is no longer true. Nevertheless, the question remains — will convex payoffs be converted into convex prices by the pricing operator? Again, there was good news because the market was using negative correlation in the calibration of the SABR model, and hence convexity was preserved. Additionally, we characterised the preservation of convexity property for many models with jumps. We analysed multi-dimensional models used in the pricing of options involving many assets like for example average-rate options.

Another problem with the SABR model is that when the power beta parameter is less than one-half, the process will hit zero in finite time. Then the question arises: What to do in such a situation? In order to ensure that the equation has a unique solution, one could for example assume that the process stays at zero once it hits it. Using mathematical language, this means that the process is absorbed at zero. For very low interest rates, which we have experienced now for more than a decade and will probably continue to see for some time, the probability of hitting zero is relatively high. Unfortunately, calibration to the smile requires a very low beta parameter.

For the SABR model absorbed at zero, with a small beta parameter, the distribution of the underlying has a part which is continuous with respect to Lebesgue measure and a Dirac delta part. As a consequence, the sensitivity produced by such a model may be completely wrong. This is because pricing operators generated by diffusions convert even discontinuous payoffs into prices which are smooth, even infinitely many times differentiable. Then, when the underlying asset moves, the hedge is adjusted in a smooth way. But you will not be able to benefit from such a smoothing of the pricing operator if it also contains a Dirac delta.

Selected results of our research into the SABR model were published in the papers "Convexity of solutions of parabolic equations" [5], "Some properties of diffusions with singular coefficients" [6] and "Correlations and bounds for stochastic volatility models" [7].

8.2 Need to understand and explain

Unfortunately, most of the research conducted with Pierre-Louis remains in handwritten notes. The reason was not confidentiality. We were driven by the need to know if all is fine with the models from the mathematical perspective and not so much by the desire to disseminate the results. Time was always limited and there were many questions to answer. Below, I mention only two directions of such research.

The options market provides information about the implied volatility for a finite set of strikes and maturities. On the other hand, pricing models are often continuous-time processes. Information given by the market is used to calibrate the models, so that when we reprice the options using the model, it returns the market quotes. Calibration may be time-consuming and hence the question arises: Can one construct a discrete-time martingale with the marginals at each date given by the smile and use it to price other options. The short answer is yes, provided the smile evolution is captured correctly. This direction of research was picked up later by a number of researchers who focused on minimising a distance and not so much on the smile dynamics.

We also challenged the standard approach to modelling the price. We asked: Should we assume we observe the price, or is it better to assume we only have partial information about it? We could instead assume a price distribution with some properties, say a mean and variance. The question was: If the price of an asset at any fixed time is not a number but a measure, say a normal distribution with a certain mean and variance, what is the concept of option price? It turns out that if wealth is defined as a linear functional over the space of probability distributions and hence is real-valued, such a concept can be developed. The resulting prices become nonlinear in the payoff. This is good and bad news at the same time. I talk about price nonlinearity and its practical consequences later on.

Arbitrage-free pricing theory gives sound mathematical foundations to the options business. A lot of effort is dedicated to the understanding and implementation of the models used in valuation and risk management. The market is very competitive, and there is an expectation from sales and trading that research develops models which help to increase their share of the market. At first appearance, one may think this is true. This view, however, is fundamentally wrong. The price reflects the cost of replication of the risk one takes by selling an option. If risk is not assessed correctly and some of its components are not taken into account, the prices are not competitive, at least for some products. It is relatively easy to correctly assess the risk contained in simple payoffs. But complex exotic products often contain hidden risks that are not easy to identify and capture.

It is not unusual to find that the models you consider to represent the best practice generate prices which are not competitive, at least for some products. This happens when your model prices risks which are not taken into the account by the models used by competitors. In the interest rate area, it is common to see one-factor models producing the best prices, because such models cannot take into account risks present in certain payoffs. The yield curve is subject to risks of parallel shift, rotation and butterfly. A one-factor model cannot capture all of them and at the same time produce a stable sensitivity analysis. Sales and trading will not be satisfied with your latest model, and you will have to explain to your management why the best you can offer

to increase their share of the market is not producing competitive prices, at least for some products. This is why the quantitative culture of an institution is so important. I must say I was lucky, again. I was able to explain, using non-quantitative terms, why we were not competitive on certain products and why we were on some others.

8.3 Multi-period portfolio optimisation and arbitrage-free-based valuation

Beyond the classical theory of pricing options, it is important to develop quantitative methods used in other parts of capital markets. In particular, it is important to understand, from the quantitative perspective, the needs of asset managers, hedge funds and insurance companies. Portfolio optimisation, risk diversification, possible consistency of approaches between the methodology of risk replication and risk-taking are all important questions, which required some attention from research. This is how I embarked on several research projects working with T. Zariphopoulou. Thaleia is an expert in portfolio optimisation. Some of her work on closed-form solutions to nonlinear optimisation problems in incomplete markets gave valuable insights for the valuation of risks that cannot be hedged, in products like weather derivatives. My work on the model development for option pricing had a similar focus. Stay as simple as possible but not too simple with the models, and try to be as explicit as possible with the solutions. It was obvious that the knowledge fit between the two of us was perfect. This is probably why we understood each other very well and, at least in my opinion, produced research which challenged some standard ways of thinking.

In order to build methodological consistency between the arbitrage-free-based valuation and portfolio optimisation and investment, one was tempted to use continuous-time models. In such a framework, on the one hand, it was clear how to eliminate arbitrage from the postulated dynamics, describing asset price movement, and on the other, one could use these dynamics to answer questions related to portfolio optimisation and investment. It is important to mention, however, that while the sell side uses continuous-time models for valuation purposes, the buy side still relies on single-period decision-making frameworks. Some classical continuous-time utility optimisation problems, like the Merton problem, were formulated and solved assuming the same dynamics as the one used in the valuation context. However, these problems and their solutions did not influence the buy side in the same way as the no-arbitrage-based valuation influenced the sell side. In my opinion, the reason is that the Merton problem is in fact a single-period optimisation problem, in which the opportunity set is a result of applying self-financing strategies to the continuous-time asset dynamics and collecting the random variables generated by that procedure. Clearly, such an approach leads to some form of time-inconsistency. Two Merton problems set for two different time horizons, say for 3 and 6 months, are inconsistent from the practical perspective. Something else was needed to deal with this issue.

This is how we came up with the concept of *forward utility*, also called *forward performance*. The name may be somewhat unfortunate, entirely my fault, but it was motivated by the language used on the sell side. In general, utility represents a concept of a non-monetary value. It depends on a time horizon and wealth and it should be adapted to the available information. As an investor, I want to measure a utility of a self-financing strategy and use it to define what is optimal for me.

The stronger form of forward utility assumes that for a fixed level of wealth (discounted with the riskless asset), utility will decrease in time, and for fixed time, it is increasing and concave in wealth. In particular, this means that in discounted units, one prefers one dollar today to one dollar at some later time. It is possible to associate with this assumption the concept of impatience. Note that default considerations are not present here. Even if there is no risk of default, one still prefers the reward earlier than later. Additionally, a forward utility of any self-financing strategy is a supermartingale, and it is a martingale for the optimal strategy. If I use such a concept to measure a utility of a self-financing strategy, I will always be worse off unless on average I preserve the utility of my initial wealth. In essence, I should not invest unless I invest optimally according to this utility criterion.

This concept is quite different from the classical utility, as formulated in the Merton problem, where utility is defined only for a single point of time in the future. In the forward utility framework, it is defined for an arbitrary future time, and in principle it is not deterministic. It is assumed to be measurable with respect to the flow of information (filtration). Naturally, as time goes by, new information is revealed and one wants to include it in the utility specification. This new utility framework can be compared with the classical recursive utility with an important difference. The forward utility embeds information about investment opportunities and as such is directly linked with the assumptions we are making about the asset dynamics. The concept of recursive utility is more general as it is not linked so strongly with the specification of investment opportunities. As in the Merton problem, the recursive utility may be defined exogenously to the opportunity set. This, however, is not the case for the forward utility, where the existence of an optimal strategy is assumed.

The weaker form of forward utility only assumes that for a fixed level of discounted wealth, the performance (utility) process is a supermartingale and not a decreasing process. Other properties remain the same, namely, the utility of any self-financing strategy is a supermartingale, and a martingale for the optimal strategy.

It is remarkable that under the above assumptions, one can derive a nonlinear stochastic PDE which the forward utility must satisfy. It is even more remarkable that under the strong form of forward utility, namely, when as a function of time it is decreasing for every fixed level of discounted wealth, one can completely characterise all such utilities. In obtaining this result, the complementarity of our skills proved critical. Thaleia has a very strong analysis background. I tend to think in probabilistic terms. It turned out that both skills were fundamentally important.

So how do you solve this problem? Well, you set up your model and calculate the wealth process of any self-financing strategy. Then you put the discounted wealth in place of the discounted initial wealth of the forward utility criteria and apply the Itô–Ventzell formula to get the semimartingale decomposition. By doing so, you compare the value, in utility terms, of holding the riskless asset with the value generated by other self-financing strategies. It turns out that the drift is non-positive for any strategy and you can identify the strategy which sets the drift to zero.

In the case of a decreasing forward utility, you can actually go much further. In fact, you can completely decouple the stochastic and deterministic components of the solution. Setting the drift to zero leads to a nonlinear deterministic PDE for the utility which you need to solve to get in explicit form the optimal strategy and the optimal

wealth it generates. To solve this PDE, you go through nonlinear transformations which link together the fast diffusion equation for the risk tolerance and the ill-posed heat equation which gives you the optimal wealth process in explicit form. You know the solutions to the last equation thanks to Widder's theorem. They are positive space-time harmonic functions. Coupling of the deterministic and stochastic components is done through a martingale and its quadratic variation. The latter aggregates risk premia present in the market and the former traces its non-aggregated evolution. You replace the time argument of the deterministic solutions with the quadratic variation and the wealth argument with its non-aggregated evolution, in order to construct the optimal portfolio and the optimal wealth process.

So what is the intuition behind this criterion? The investor aims to exploit risk premia potentially present in the market. She formulates preferences with respect to these investment opportunities. If there is no premium to exploit, the investor, most naturally, does not invest, i.e., holds the riskless asset. But when risk premia are present, the investor takes a position in the market, but she will be worse off than doing nothing unless she acts optimally. If she does, then on average, conditionally on the information she has at that time, she maintains the average utility of her wealth. The investor makes infinitesimal portfolio decisions at any point in time. Every time the decision is suboptimal, the investor's average utility of her wealth decreases. In this sense, our criterion describes multi-period optimisation.

So what is the difference between the classical formulation of the utility optimisation problem à la Merton and this alternative optimisation criterion? One way to see it is to identify the opportunity sets for both setups. I have argued before that in the Merton problem, one deals with a single-period optimisation over the set of random variables defined by running self-financing strategies until the time at which one calculates the maximal expected utility. Note that the opportunity set over which optimisation takes place for our alternative criteria is different. In fact, it consists of all wealth processes generated by running self-financing strategies. So here, we optimise over all wealth processes and not just over the values of these processes at one future point in time. In this sense, our criterion is consistent through time as it gives a process which is optimal at any point in time — of course, provided the utility at any point in time incorporates the risk premia present in the asset dynamics which we aim to exploit.

A detailed analysis of the decreasing utility case is presented in the paper "Portfolio choice under space-time monotone performance criteria" [18]. The case of a general forward utility in continuous and discrete time is studied in "Portfolio choice under dynamic investment performance criteria" [16]. The general SPDE as well as an HJB equation corresponding to the case when Markovian state variables are introduced are analysed in "Stochastic partial differential equations and portfolio choice" [19].

8.4 Value and its adjustment

When we talk about the value of a financial product, we mean the value derived assuming there is no arbitrage, which in turn means that there exists a martingale measure. Implicit to this assumption is the presence of a market in which one can trade

primary securities. The idea is to synthetically create financial products (contracts) by trading the primary securities, and sell them to the end users. When a product is traded, the profit is booked. To be clear, the bank sells a contract and at the same time puts risk on its balance sheet. The risk is in the future and the profit is booked today. This may look strange to many, but this practice is in line with the Day 1 Profit accounting standard used on the sell side of the capital markets. The institution is left with risk for which it was compensated by the premium it received from the product sale. There are obviously two questions that need answers: What is the price at which it is rational to sell the product, and how to manage the risk that was put on the balance sheet? In each of the business lines, the products are sold one by one and risks are aggregated together and managed in books. Once again, note that products are sold individually, but risks are managed in aggregate. The importance of this is explained below when we talk about the various valuation methods applied to individual products and how this translates into risk management of books.

When there is a unique martingale measure, the obvious way to value a product is to calculate the expected value of the product cash flows under this unique measure. The reason for this is the fact that the price calculated in this way is exactly equal to the cost of running a self-financing and replicating strategy that generates the product payoff. In practice, of course, you would charge an extra margin for delivering such a service. The size of this margin is left to the discretion of the trader who needs to take many considerations into account. Margins will vary depending on the market liquidity, risk position of a book, size of the transaction, credit quality of the counterparty and many other variables. Additional charges, requested by the market-risk department and representing the reserves, come on top. The last global financial crisis exposed the need for the development of model-based ways to adjust the value of financial contracts. However, the valuation procedure remained linear in products. This allows adding and substracting contracts from the portfolio by trading them in the market. On the other hand, the adjustments and reserves, which are relative to the portfolio, need not be linear. One way to think about it is to say that the arbitrage-free pricing theory was applied to capture the expected values and the adjustments were to capture higher order effects of the valuation procedure.

However, in many modelling situations, martingale measures are not unique and the question is which one, if any, should be chosen for the price calculations. Also note that a multiplicity of martingale measures could be just an artefact of thinking about the model. A good example of such a situation is a model with stochastic volatility. If the model assumes trading of the risky and riskless asset only, there will be many martingale measures. However, if we also assume trading in options on the risky asset, the martingale measure may be unique. Unfortunately, such a model enlargement is not always possible.

Alternatively, one could aim to develop a different concept of value altogether. The presence of many martingale measures is related to the fact that not all models generate risks that can be hedged. Of course, when the martingale measure is unique, the price is equal to the cost of hedging and all risks can be hedged away. When there are many martingale measures, not all risks can be perfectly hedged. Therefore, the suggestion was made to focus on super-hedging instead. Using a super-hedging self-financing strategy comes down to generating higher than the contractual cash flows.

It turns out that the infimum over all initial costs of super-hedging strategies is equal to the supremum over the expected values of payoffs calculated under the different martingale measures. From the theoretical point of view, this valuation procedure has many advantages. One, very important to scientists, is that it brings an extra layer of sophisticated and beautiful mathematics to the field of modern finance. From the practical viewpoint, however, it obscures the intuitive idea of value and of its adjustment. The super-hedging-based pricing mechanism is nonlinear in the contract. It is even nonlinear in contract cash flows. This in turn means that it is more complicated to aggregate risks at the portfolio level and distinguish value from its adjustment. The value is related to the Day 1 Profit, the adjustment is related to the margins and reserves. At the risk of oversimplification, one could also say that contracts are valued relatively to the market, but the adjustments are made relatively to the portfolio. Putting the above aside for now, there is an additional issue: the price calculated in this way is not competitive. The reason is not, as explained previously, that certain risks are ignored, but the fact that the super-hedging generates more in payoffs terms than is required, and hence naturally is more expensive.

Recall that previously the price was given by the initial capital that was required for a hedging (super-hedging) strategy. Now we look at the value from the investment perspective. For a given utility, we compare the value functions which are obtained by maximising expected utility over all self-financing strategies. If we sell an option, we price it at the level which makes us indifferent in terms of the value functions calculated with and without the claim. In essence, we ask by how much our capital should increase for the two value functions to coincide. The same argument is applied when we buy an option. It is easy to see that this valuation method is not linear in the product.

The mathematical finance community working on the indifference-based valuation used continuous-time models. This was also the approach Thaleia (Zariphopoulou) and myself adopted in the beginning. In 2004, we published the paper "An example of indifference prices under exponential preferences" [14]. During one of my visits to Bonn, Dieter, in his wisdom of seeking simplicity, asked how this theory would look in the case of a single-period binomial model. This is how we embarked on the project to explain intuitively what is the meaning of the indifference price and how one can think about risk management in this context. In the same volume of *Finance and Stochastics*, we published the paper "A valuation algorithm for indifference prices in incomplete markets" [15]. Moreover, we wrote the first chapter "The single period binomial model" [17] to the book *Indifference Pricing: Theory and Applications*.

It turns out that in this simple model setup, the indifference price has a very intuitive and beautiful interpretation. Assume a singe-period model with a riskless asset and two binomial assets of which only one is traded. Such a model is incomplete and hence there are many martingale measures. Assume also that you want to price by indifference a claim written on both assets. All calculations can be done explicitly and you end up with a formula. Have a guess on what would it look like.

Of course, all that matters is the joint distribution, under the so-called real (historical) measure, of the traded and non-traded asset at the end of the period and their values in the beginning. From this, you can deduce the conditional distribution of the

non-traded asset given the value of the traded asset. In this way, you have isolated risks that can respectively cannot be hedged. From the actuarial perspective, it is very natural to apply a certainty equivalent to price risk that cannot be hedged. From the investment banking perspective, it is natural to apply the replication method to price risk that can be hedged. This is exactly what the formula tells you to do, and it even gives you an extra hint. Namely, it tells you which martingale measure to take for dealing with risks you cannot hedge.

Of course, for pricing risk you cannot hedge, you use the real (historical) measure. It turns out that the indifference price formula tells you to take the martingale measure for which the conditional distribution of the non-traded asset, given the value of the traded asset, is the same as under the real (historical) joint distribution. Because these conditional distributions are the same under both measures, you can first calculate the certainty equivalent under the conditional distribution, treat the result as a new payoff, and value it computing its expectation under this martingale measure. The algorithm is obvious once you have seen it. To guess it directly from the definition of pricing by indifference is not so easy.

This indifference-based pricing formula is nonlinear in the payoff. Indeed, the certainty equivalent transforms the original payoff into a new payoff given by the log of the expectation of the exponentiated payoff, calculated under the above-mentioned conditional distribution. So the nonlinearity of pricing is directly a consequence of the presence of risk you cannot hedge. The second step in pricing is linear. Note also that in many situations, the payoff transformation might turn it into a new payoff which is smoother. This in itself is good news. However, the nonlinearity remains a problem from the operational perspective. Namely, how can we separate the value from its adjustment, the importance of which was described before? A detailed analysis of the value and of the related risk management is presented in "The single period binomial model" [17].

One of the important results presented there is also the link with the general representation of the indifference price derived under exponential preferences and given by the supremum of the expectations applied to the difference of two terms described before, namely, the payoff and an entropic penalty term. It turns out that the martingale measure used in the indifference price calculation is in fact the minimal entropy martingale measure.

The analysis carried out in [17] addresses also other important issues. The first one is the dependence on the choice of numeraire. The second deals with risk aggregation and risk management under this nonlinear pricing method. To a large extent, [17] can be viewed as an analogue of the first chapter in the book *Martingale Methods in Financial Modelling*, where the effects of linear pricing and risk management based on it are analysed, also in a single-period binomial model setup, in order to focus on ideas and avoid any mathematical complications due to general models. As expected, also in the case of indifference pricing, mathematical considerations are trivial; however, the practical implications of moving away from linear pricing can be exposed. It turns out that one can maintain independence of the numeraire for pricing by indifference provided the risk aversion coefficient in the exponential utility is adjusted accordingly. In practice, this means that exponential utility may become random under different numeraire choices. The same remains true in the

context of optimal investment, which constitutes an integral part of this method of pricing.

Risk aggregation and risk management can be also adapted to satisfy the requirement of separation between value and its adjustment. The linear part of the indifference value is given by the payoff expectation, now calculated under a well-identified martingale measure. The second order correction term, derived from the formula for the indifference price, is related to the expectation of the conditional variance of the claim. Details, including an analysis of the trading strategies and of the residual risk, are included in *Martingale Methods in Financial Modelling* [11].

9 Back to academia

When I left academia and joined the industry in 2000, I had an idea to do this new job for five years or so and move back. Every now and then, I would say to myself to stay for another five years and then go. Not that I had a place to go to or had even started to look for one. The years went by and it became clear that if I would do nothing, I would stay in the industry until retirement. Thaleia knew I was ready to move and mentioned a position not far from London. The job, which I found attractive, was at the Oxford-Man Institute of Quantitative Finance (OMI). After thinking quite a bit about the move, I applied. Thanks to my contacts with academia, I could ask for letters of support from professors who knew me well. My application was successful and I moved to Oxford in 2012. My main mandate was to bring industry research experience and hopefully also some funding. OMI is part of the University of Oxford but is privately funded, mostly from Man Group. AHL, which is part of Man Group, had a research lab in the same building. Interaction between the two research groups was facilitated by the physical presence in the same premises.

AHL is a systematic trading hedge fund. There was an opportunity for me to understand better the way research was done in this part of the finance industry. When I joined OMI, Terry Lyons was its director and I became his deputy. I had two places to learn from, namely, AHL and the Mathematical Institute. At the same time, the finance industry was changing. Appetite for complex transactions was declining after the global financial crisis of 2008. The focus was shifting to trading flow products. This required the development of algorithmic automated solutions to many activities in the industry. In 2015, Terry went back to the Mathematical Institute and Steve Roberts from the Engineering Department became the new director of OMI. This resulted in a big change of research style conducted at OMI.

This period of my professional life was a lot of fun, after all these years in the industry, where people management took far too much of my time away from research. I worked on a number of projects. I mention three below. The first two are continuations of projects I worked on before. The third is new and represents a change in the research focus from derivatives pricing and investment into modelling of the market microstructure.

9.1 Fractional SABR model

The classical SABR model was developed for pricing caps and swaptions in the interest rate market. Later on, it was used in the FX market for pricing options with

very long maturity. It was just a matter of time before it was applied for pricing equity options. However, the equity markets generate an at-the-money volatility skew which is inconsistent with the one generated by SABR. Various attempts were made to modify the volatility specification to solve this problem. It turned out that replacing Brownian motion with a fractional Brownian motion in the specification of the SABR volatility process generated the desired outcomes.

Unfortunately, doing just that leads to a number of mathematical questions which were left unanswered. This alone was not the main reason for me to look again at the SABR framework. There was another reason. Terry is one of the creators of rough paths theory. The fractional Brownian motion used in the modification of SABR had a Hurst exponent of the order of one-tenth. The Brownian motion used in the specification of the volatility process in the classical SABR model is, of course, a fractional Brownian motion with a Hurst exponent of one-half. Trajectories of fractional Brownian motions with a Hurst exponent less than one-half are rough. The classical stochastic calculus does not apply any more and one has to come up with a different way of dealing with the question of how to specify the volatility process for which the stochastic driver is rough. This was the main reason I started working on such a modification of SABR.

As mentioned before, the correlation between the two Brownian motions of the classical SABR model determines many important properties of the model. Replacing Brownian motion with a fractional one introduces a dependence structure into the model which needs to be defined and analysed. There are various ways one can define a fractional Brownian motion using another Brownian motion. That is what some researchers have done, without worrying too much about the consequences for the dependence structure such a construction injects into the model behaviour.

Thanks to the joint self-similarity of bivariate Brownian motion with correlated coordinates, both volatility parameters of the classical SABR model scale as a square root of time. The question is: How will these parameters scale with time in the fractional SABR model, and what property will determine this scaling? It turns out that one needs joint self-similarity of the bivariate process, where in the first coordinate we put the Brownian motion driving the dynamics of the asset value and in the second, we put the fractional Brownian motion defining the volatility process. This ensures that the asset volatility scales as a square root of time and the volatility of the volatility scales as a power of time, which depends on the Hurst exponent. This property is precisely what allows capturing the at-the-money volatility skew observed in the equity market.

This work was presented at a number of conferences and a paper was submitted for publication to the proceedings from a conference in Jerusalem in 2018 [10].

There are other problems with the specification of a fractional SABR model. There is also a need to analyse the mathematical properties of this new model, extending results obtained for the classical SABR model. However, the biggest challenge of all is the development of an approximate formula for the implied volatility that would generalise the results by Pat Hagan.

9.2 Multi-period investment strategies

In the context of multi-period investment strategies, I worked in the following two directions.

In the first, it turned out that the optimal portfolio generated by a decreasing forward utility is the Markovitz portfolio in which the risk parameter is adjusted in line with the risk tolerance, which is generated by the utility choice and the performance of the optimal strategy to date. This indicates that at least at the continuous-time model level, when making investment decisions on infinitesimally small time horizons, it is optimal to use the Markowitz portfolio. Additionally, the theory tells you how to adjust your tolerance towards risk, consistently with your utility and performance. For forward power utility, the risk tolerance is linear in wealth and hence the ratio is constant through time. This corresponds to using the Markovitz portfolio with a fixed risk parameter at every rebalancing time.

In the second direction, the focus was on the role of information and in particular on the use and value of private information. Enlargement of filtrations is a rapidly developing area in mathematical finance. General results, however, are difficult to obtain. This is why I looked at a very simple case with two managers, where one is more informed than the other. It turns out that the more informed manager will be able to use extra information to his advantage. This work has been presented at a number of conferences, but was never converted into a preprint.

9.3 Modelling the spot FX market

I proposed a continuous-time model for the spot FX market, in which there are market makers and traders. Market makers, also known as liquidity providers, continually quote indicative bid (buy) and ask (sell) prices, relatively to the unobserved fundamental FX price. Traders watch the market and send trade requests to the market makers based on the best bid and ask prices they see in their liquidity pool. Different traders rely on different liquidity pools. Market makers receive streams of trade requests from the traders. They have the right of the last look and may reject or accept a trade.

What would indicate that the unobserved fundamental price has changed?

A market maker should observe unbalanced bid and ask trade requests coming from the traders. A symmetric distribution would indicate that the quoted bid and ask prices correspond well to the unobserved fundamental price. If the requests are skewed towards the bid or the ask, this may indicate that the price has moved and the quotes have to be adjusted accordingly. Therefore, the distribution of trade requests could be used to estimate the unobserved fundamental price.

A trader would observe the evolution of the best bid and the best ask at his liquidity pool and will send trade requests to a market maker. The market maker has the option to accept or reject them. The intensity with which trade orders will reach the market maker depends on how far his bid and ask are from the best bid and the best ask the trader can see in her liquidity pool.

If we assume only one trader, then his requests will go to the market makers publishing the best bid and the best ask prices. However, with many traders acting on

different information, using different market makers and having a better or worse understanding where the fundamental price is, an individual market maker is likely to receive trade requests coming from different traders.

This is essentially the logic of the model. Its mathematical description was presented at a number of conferences and is contained in an unfinished preprint. In the next step, the model should be implemented to provide a simulation platform for the traders and market makers to test various strategies.

10 Final comments

Arbitrage-free pricing theory gave foundations for a new business that capital markets have developed. However, a theory is what it is, a theory. This applies not only to finance but, in general, to any activity in which one uses scientific methods to describe and understand the area of interest. It has always been the case in physics, where mathematics was used to understand our universe in both the micro and the macro scales. This is equally true for engineering, where mathematics is fundamental to understanding how to build things. This is not that surprising because mathematics was developed to help us understand the world we live in. This world was there before we came along and is likely, I hope, to still be there way after we are gone as species.

What is much less obvious to me is why we could use mathematics in the context of finance. Of course, the development of stochastic calculus had originally nothing to do with finance. However, later on it proved to be a theory almost created for it. To me, personally, this is shocking because finance has nothing to do with the laws of physics which are given to us. In contrast, markets represent aggregated actions of individuals. Nevertheless, these individuals generate market movements that at least in the context of risk elimination can be relatively well captured by methods based on stochastic calculus.

The surprise I refer to is not so much that this is possible. It is more about the fact that it reflects some kind of fundamental truth of the mathematical language which transcends not only physics, but even human-generated environments. We may never understand all the secrets of the world we live in, but clearly, we can rely on mathematics to make progress. This applies even to the various social environments we have created.

References

1. Brace, A., Gątarek, D., Musiela, M.: The market model of interest rate dynamics. Math. Finance **7**, 127–155 (1997)

2. Harrison, J.M., Pliska, S.R.: Martingales and stochastic integrals in the theory of continuous trading. Stoch. Process. Appl. **11**, 215–260 (1981)
3. Heath, D., Jarrow, R.A., Morton, A.: Bond pricing and the term structure of interest rates: a new methodology for contingent claim valuation. Econometrica **60**, 77–105 (1992)
4. Jamshidian, F.: LIBOR and swap market models and measures. Finance Stoch. **1**, 293–330 (1997)
5. Lions, P.-L., Musiela, M.: Convexity of solutions of parabolic equations. C. R. Acad. Sci., Sér. 1 Math. **342**, 913–921 (2006)
6. Lions, P.-L., Musiela, M.: Some properties of diffusions with singular coefficients. Commun. Appl. Anal. **10**, 109–126 (2006)
7. Lions, P.-L., Musiela, M.: Correlations and bounds for stochastic volatility models. Ann. Inst. Henri Poincaré C **24**, 1–16 (2007)
8. Miltersen, K.R., Sandmann, K., Sondermann, D.: Closed form solutions for term structure derivatives with log-normal interest rates. J. Finance **52**, 409–430 (1997)
9. Musiela, M.: Stochastic PDEs and term structure models. Journées Internationales de Finance, IGR-AFFI, La Baule, June 1993. (Unpublished)
10. Musiela, M.: Multivariate fractional Brownian motion and generalizations of SABR model. In: Proceedings of the Conference "Options 45 Years After the Publication of the Black–Scholes–Merton Model", Jerusalem 2018. Available online at https://doi.org/10.5281/zenodo.4772004
11. Musiela, M., Rutkowski, M.: Martingale Methods in Financial Modelling. Springer, Berlin (1997)
12. Musiela, M., Rutkowski, M.: Continuous-time term structure models: forward measure approach. Finance Stoch. **1**, 261–291 (1997)
13. Musiela, M., Sondermann, D.: Different dynamical specifications of the term structure of interest rates and their implications. SFB 303 Discussion Paper No. 260, University of Bonn, 1993. Available online at https://www.wiwi.uni-bonn.de/bgsepapers/old%20series/bonsfb/bonsfb260.pdf
14. Musiela, M., Zariphopoulou, T.: An example of indifference prices under exponential preferences. Finance Stoch. **8**, 229–239 (2004)
15. Musiela, M., Zariphopoulou, T.: A valuation algorithm for indifference prices in incomplete markets. Finance Stoch. **8**, 399–414 (2004)
16. Musiela, M., Zariphopoulou, T.: Portfolio choice under dynamic investment performance criteria. Quant. Finance **9**, 161–170 (2009)
17. Musiela, M., Zariphopoulou, T.: The single period binomial model. In: Carmona, R. (ed.) Indifference Pricing: Theory and Applications, pp. 3–43. Princeton University Press, Princeton (2009)
18. Musiela, M., Zariphopoulou, T.: Portfolio choice under space-time monotone performance criteria. SIAM J. Financ. Math. **1**, 326–365 (2010)
19. Musiela, M., Zariphopoulou, T.: Stochastic partial differential equations and portfolio choice. In: Chiarella, C., Novikov, A. (eds.) Contemporary Quantitative Finance. Essays in Honour of Eckhard Platen, pp. 195–215. Springer, Berlin (2010)
20. Vargiolu, T.: Invariant measures for the Musiela equation with deterministic diffusion term. Finance Stoch. **3**, 483–492 (1999)