Mathematics and Finance

Walter Schachermayer*

April 12, 2016

Abstract

This article consists of two parts. The first briefly discusses the history and the basic ideas of option pricing. Based on this background, in the second part we critically analyze the role of academic research in Mathematical Finance relating to the emergence of the 2007-2008 financial crisis.

1 Introduction

Mathematical Finance serves as a prime example of a flourishing application of mathematical theory. It became an important tool for several tasks in the financial industry and this "mathematization" of the financial business seems to be irreversible. Therefore in many curricula of mathematics departments, but also in business schools, mathematical finance is now regularly taught.

In this survey we want to summarize how these ideas developed, starting from the seminal work of Louis Bachelier [2] who defended his thesis "Théorie de la spéculation" in 1900 in Paris. Henri Poincaré was a member of the jury and wrote a very positive report. Bachelier used probabilistic arguments, thus introducing Brownian motion for the first time as a mathematical model, in order to develop a rational theory of option pricing.

This theme subsequently remained dormant for almost 70 years until it was taken up again by the eminent economist Paul Samuelson. In the sequel Fisher Black, Robert Merton, and Myron Scholes applied a slightly modified version of Bachelier's model and the resulting "Black-Scholes formula" for

*Fakultät für Mathematik, Universität Wien, Oskar-Morgenstern-Platz 1, A-1090 Wien, walter.schachermayer@univie.ac.at and the Institute for Theoretical Studies, ETH Zurich. Partially supported by the Austrian Science Fund (FWF) under grant P25815, the Vienna Science and Technology Fund (WWTF) under grant MA14-008 and by Dr. Max Rössler, the Walter Haefner Foundation and the ETH Zurich Foundation.

the price of a European option quickly became very influential in the world of finance. We shall sketch this development.

In the second part we want to give a critical view of the success or failure of these mathematical insights in the real world. We shall argue that the probabilistic approach turned out to be highly successful with respect to the original goal of pricing options on a liquidly traded risky asset, e.g. a share of a large company. On the other hand, the probabilistic approach was subsequently applied to many other tasks, such as credit risk, risk management, "real options" etc. We shall analyze to which extent mathematical models were involved in the financial crisis of 2007-2008. It is sometimes claimed in the public discourse that "nobody warned about the misuse of mathematical models". We shall see that such claims are not justified.

2 Louis Bachelier and Black-Scholes

We outline the remarkable work of L. Bachelier (1870 – 1946) by following the more extensive presentation [18] which I gave at the summer school 2000 in St. Flour.

It is important to note that the young Louis Bachelier did not attend any of the grandes écoles in Paris, apparently for economic reasons. In order to make a living he worked as a subordinate clerk at the Bourse de Paris where he was exposed on a daily basis to the erratic movements of prices of financial securities.

L. Bachelier was interested in designing a rational theory for the prices of term contracts. The two forms which were traded at the Bourse de Paris at that time also play a basic role today: forward contracts and options. We shall focus on the mathematically more interesting of these two derivatives, namely options.

Definition 2.1. *A European call (resp. put) option on an underlying security S consists of the* right (but not the obligation) *to buy (resp. to sell) a fixed quantity of the underlying security S, at a fixed price K and a fixed time T in the future.*

The underlying security S, usually called the *stock*, can be a share of a company, a foreign currency, gold etc. In the case of Bachelier the underlying securities were *"rentes"*, a form of perpetual bonds which were very common in France in the nineteenth century (compare [18]). The nominal value was 100 francs and it would pay 3 francs of interest every year. But the nominal capital was never paid. While the specifics of these assets are not relevant, it is worthwhile to note the following features (the terminology below will be explained later):

- the underlying asset S (the "rentes" in the concrete case of Bachelier) were liquidly traded.
- the value of the asset would typically not deviate too much from its nominal value of 100 francs.

In addition, they had the following properties.

- low volatility of the underlying asset.
- short term to maturity of the option (maximum: 2 months).
- approximately "at the money" options.

We mention these features explicitly as it is important in many applications to keep in mind for which purposes a mathematical model was originally intended, in particular, if it is later also applied to quite different situations.

Fixing the letter K for the strike price of the option, one arrives — after a moment's reflection — at the usual "hockey-stick" shape for the pay-off function of a call option at time T. We draw the value of the option as a function of the price S_T of the underlying asset S at time T.

Let $\widehat{C}$ denote the upcounted (from time $t = 0$ to time $t = T$) price C of the option. We shall not elaborate on the rather boring aspects of upcounting and discounting and assume that the riskless rate of interest equals zero so that $C = \widehat{C}$.

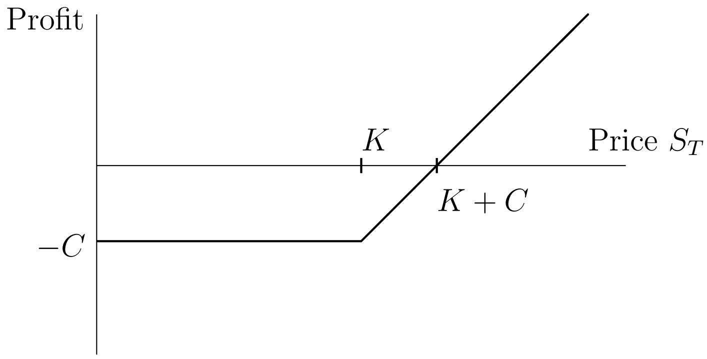

Figure 1: Pay-off function of a call option at time T.

The graph displayed in Figure 1 appears explicitly in Bachelier's thesis. It gives the profit or loss of the option at time $t = T$ when we shall know the price S_T of the underlying asset S. But we have to determine the price C of the option which we have to pay at time $t = 0$. We note in passing that the special form of the above payoff function is not really relevant. Its only crucial feature is that it is not linear.

Louis Bachelier now passes to the central topic, *Probabilities in Operations on the Exchange*. Somewhat ironically, he had already addressed the basic difficulty of introducing probability in the context of the stock exchange in the introduction to the thesis in a very sceptical way: "The calculus of probabilities, doubtless, could never be applied to fluctuations in security quotations, and the dynamics of the Exchange will never be an exact science."

Nevertheless he now proceeds to model the price process of securities by a probability distribution distinguishing "two kinds of probabilities":

> "1. The probability which might be called "mathematical", which can be determined *a priori* and which is studied in games of chance.
> 2. The probability dependent on future events and, consequently impossible to predict in a mathematical manner.

This last is the probability that the speculator tries to predict."

My personal interpretation of this — somewhat confusing — definition is the following: sitting daily at the stock exchange and watching the movement of prices, Bachelier got the same impression that we get today when observing price movements in financial markets, e.g., on the internet. The development of the charts of prices of stocks, indices etc. on the screen or on the blackboard resembles a "game of chance". On the other hand, the second kind of probability seems to refer to the expectations of a speculator who has a personal opinion on the development of prices. Bachelier continues:

"His (the speculator's) inductions are absolutely personal, since his counterpart in a transaction necessarily has the opposite opinion."

This insight leads Bachelier to the remarkable conclusion, which in today's terminology is called the "efficient market hypothesis":

"It seems that the market, the aggregate of speculators, at a given instant can believe in neither a market rise nor a market fall since, for each quoted price, there are as many buyers as sellers."

He then makes clear that this principle should be understood in terms of "true prices", i.e., discounted prices. Finally he ends up with his famous dictum:

"In sum, the consideration of true prices permits the statement *of this fundamental* principle:

The mathematical expectation of the speculator is zero."

This is a truly fundamental principle and the reader's admiration for Bachelier's pathbreaking work will increase even more when continuing to the subsequent paragraph of Bachelier's thesis:

"It is necessary to evaluate the generality of this principle carefully: It means that the market, at a given instant, considers not only currently negotiable transactions, but even those which will be based on a subsequent fluctuation in prices as having a zero expectation.

For example, I buy a bond with the intention of selling it when it will have appreciated by 50 centimes. The expectation of this complex transaction is zero exactly as if I intended to sell my bond on the liquiditation date, or at any time whatever."

In my opinion, in these two paragraphs, the basic ideas underlying the concepts of martingales, stopping times, trading strategies, and Doob's optional sampling theorem already appear in a very intuitive way. It also sets the basic theme of the modern approach to option pricing which is based on the notion of a martingale.

Let us look at the implications of the *fundamental principle*: In order to draw conclusions from it, Bachelier had to determine the probability distribution of the random variable S_T (the price of the underlying security at expiration time T) or, more generally, of the entire stochastic process $(S_t)_{0\leq t\leq T}$. It is important to note that Bachelier had the approach of considering this object as a *process*, i.e., by thinking of the *pathwise behavior* of the random trajectories $(S_t(\omega))_{0\leq t\leq T}$; this was very natural for him, as he was constantly exposed to observing the behavior of the prices, as t "varies in continuous time".

To determine the law of the process S, Bachelier assumes that, for $0 \leq t \leq T$, the probability $p_{x,t}dx$, that the price S of the underlying security, starting at time t_0 from S_{t_0}, lies at time $t_0 + t$ in the infinitesimal interval $(S_{t_0} + x, S_{t_0} + x + dx)$ is *symmetric around* $x = 0$ *and homogeneous in time* t_0 *as well as in space.*

Bachelier notices that this creates a problem, as it gives positive probabilities to negative values of the underlying security, which is absurd. But one should keep in mind the proportions mentioned above: a typical yearly standard deviation σ of the prices of the underlying stock S considered by L. Bachelier was of the order of 2.4 %. Hence the region where the bond price becomes negative after one year is roughly 40 standard deviations away from the mean; anticipating that Bachelier uses the normal distribution, this effect is — in his words — "considered completely negligible". This was certainly justified as the horizons for the options were only fractions of a year. On the other hand, we should be warned when considering Bachelier's results asymptotically for $t \to \infty$ (or $\sigma \to \infty$ which roughly amounts to the same), as in these circumstances the effect of assigning positive probabilities to negative values of S_t is not "completely negligible" any more. But this was not Bachelier's concern. As J.M. Keynes phrased so nicely: in the long run we all are dead.

After these specifications, Bachelier argues that *"by the principle of joint probabilities"* (apparently he means the independence of the increments), we obtain

$$p_{z,t_1+t_2} = \int_{-\infty}^{+\infty} p_{x,t_1} p_{z-x,t_2} dx. \tag{1}$$

In other words, he obtains what we call today the Chapman-Kolmogoroff equation. Then he observes that "this equation is confirmed by the function"

$$p_{x,z} = \frac{1}{\sigma\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2\sigma^2 t}\right), \tag{2}$$

concluding that "evidently the probability is governed by the Gaussian law already famous in the calculus of probabilities".

Summing up, Bachelier derived from some basic principles the transition law of Brownian Motion and its relation to the Chapman-Kolmogoroff equation.

Bachelier then gives an *"Alternative Determination of the Law of Probability"*. He approximates the continuous time model $(S_t)_{t\geq 0}$ by a random walk, i.e., a process which during a time interval Δt moves up or down with probability $\frac{1}{2}$ by Δx. He clearly works out that Δx must be of the order $(\Delta t)^{\frac{1}{2}}$ and — using only Stirling's formula — he obtains the convergence of the one-dimensional marginal distributions of the random walk to those of Brownian motion.

Suming up, Bachelier arrives at the model for the stock price process

$$S_t = S_0 + \sigma W_t, \qquad 0 \leq t \leq T, \tag{3}$$

where, in modern terminology, $(W_t)_{0\leq t\leq T}$ denotes standard Brownian motion and the constant $\sigma > 0$ is the "volatility", which Bachelier has called the "coefficient de nervosité du marché".

Having fixed the model, Bachelier is now able to determine the price C of an option appearing in Figure 1. Indeed, the probability distribution in this picture is now given by a Gaussian distribution with mean S_0 (the current price of the underlying S) and variance $\sigma^2 T$. The "fundamental principle" (the mathematical expectation of the speculator is zero) states that the integral of the function depicted in Figure 1 with respect to this probability distribution equals zero. This yields the equation

$$-C + \int_{K-S_0}^{\infty} (x - (K - S_0))\, f(x) dx = 0, \tag{4}$$

where

$$f(x) = \frac{1}{\sigma\sqrt{2\pi T}} e^{-\frac{x^2}{2\sigma^2 T}}, \tag{5}$$

which clearly determines the relation between the premium C of the option and the strike price K. In other words, equation (4) determines the price for the option and therefore solves the basic problem considered by Bachelier.

It is straightforward to derive from (4) an "option pricing formula" by calculating the integral in (4): denoting by $\phi(x)$ the standard normal density function, i.e., $\phi(x)$ equals (5) for $\sigma^2 T = 1$, by $\Phi(x)$ the corresponding distribution function, and using the relation $\phi'(x) = -x\phi(x)$, an elementary calculation reveals that

$$\begin{aligned} C &= \int_{\frac{K-S_0}{\sigma\sqrt{T}}}^{\infty} \left(x\sigma\sqrt{T} - (S_0 - F) \right) \phi(x) dx \qquad (6) \\ &= (S_0 - K)\Phi\left(\frac{S_0 - K}{\sigma\sqrt{T}}\right) + \sigma\sqrt{T}\phi\left(\frac{S_0 - K}{\sigma\sqrt{T}}\right), \end{aligned}$$

which is a very explicit and tractable formula. Note that the only delicate parameter is σ while all the other quantities are given.

Finally in Bachelier's thesis a section follows, which is not directly needed for the subsequent applications in finance, but which — retrospectively — is of utmost mathematical importance: *"Radiation of probability"*. Consider the discrete random walk model and suppose that the grid in space is given by

$$\ldots, x_{n-2}, x_{n-1}, x_n, x_{n+1}, x_{n+2}, \ldots$$

having the same distance

$$\Delta x = x_n - x_{n-1},$$

for all n, and such that at time t these points have probabilities

$$\ldots, p_{n-2}^t, p_{n-1}^t, p_n^t, p_{n+1}^t, p_{n+2}^t, \ldots$$

for the random walk under consideration. What are the probabilities

$$\ldots, p_{n-2}^{t+\Delta t}, p_{n-1}^{t+\Delta t}, p_n^{t+\Delta t}, p_{n+1}^{t+\Delta t}, p_{n+2}^{t+\Delta t}, \ldots$$

of these points at time $t + \Delta t$? A moment's reflection reveals the rule which is so nicely described by Bachelier in the subsequent phrases:

"Each price x during an element of time radiates towards its neighboring price an amount of probability proportional to the difference of their probabilities.

I say proportional because it is necessary to account for the relation of Δx to Δt.

The above law can, by analogy with certain physical theories, be called the law of radiation or diffusion of probability."

Passing formally to the continuous limit and denoting by $P_{x,t}$ the distribution function associated to the density function (2)

$$P_{x,t} = \int_{-\infty}^{x} p_{z,t} dz \tag{7}$$

Bachelier deduces in an intuitive and purely formal way the relation

$$\frac{\mathrm{d}P}{\mathrm{d}t} = \frac{1}{c^2}\frac{\mathrm{d}p}{\mathrm{d}x} = \frac{1}{c^2}\frac{\mathrm{d}^2P}{\mathrm{d}x^2} \tag{8}$$

where $c > 0$ is a constant. Of course, the heat equation was known to Bachelier: he claims that *"this is a Fourier equation"*.

Hence Bachelier in 1900 very explicitly discovered the fundamental relation between Brownian motion and the heat equation; this fact was rediscovered five years later by A. Einstein [8] and resulted in a goldmine of mathematical investigation through the work of Kolmogoroff, Kakutani, Feynman, Kac, and many others up to recent research. It is worth noting that H. Poincaré in his (very positive) report on Bachelier's thesis saw the seminal importance of this idea when he wrote "On peut regretter que M. Bachelier n'ait pas developpé d'avantage cette partie de sa thèse" (One may regret that M. Bachelier did not further develop this part of his thesis.)g

But unfortunately the thesis of Bachelier obtained only a "mention bien". Apparently the two other jury members did not have the same positive opinion as H. Poincaré towards this unusual student who was working at the stock exchange. But in order to make an academic career a "mention très bien" was an absolute must, just as it is today in France. Louis Bachelier subsequently had a rather difficult life and his work was not well received in France. On the other hand, A. Kolmogoroff or K. Itô did appreciate his writings.

We focused on the early work by L. Bachelier as his contribution is less known to a wider public than the "Black-Scholes option pricing formula".

After Bachelier's pioneering work, it remained silent around the theme of option pricing for many decades. This is in sharp contrast to the progress made during this period in the mathematical theory of stochastic processes and their applications in physics and biology.

Eventually in 1965 the eminent economist P. Samuelson rediscovered Bachelier's thesis in the library of Harvard University, following a request of the statistician J. Savage. Samuelson was immediately fascinated by Bachelier's work and started a line of research on option pricing and related topics which at this time had much more repercussions than Bachelier's thesis. Samuelson [17] proposed a multiplicative version of Bachelier's model defined by the stochastic differential equation

$$\frac{dS_t}{S_t} = \sigma dW_t + \mu dt, \qquad 0 \leq t \leq T, \tag{9}$$

where $(W_t)_{0 \leq t \leq T}$ denotes a standard Brownian motion and $\sigma \in \mathbb{R}_+, \mu \in \mathbb{R}$ are constants.

Given the initial value S_0 of the stock, Itô's formula yields the solution

$$S_t = S_0 \exp\Big(\sigma W_t + \big(\mu - \frac{\sigma^2}{2}\big)t\Big), \qquad 0 \leq t \leq T. \tag{10}$$

The SDE (9) states that the *relative* increments $\frac{dS_t}{S_t}$ of the price process are driven by a Brownian motion with drift. Today, the model (9) is usually called the *Black-Scholes model*.

In 1973, the papers by F. Black and M. Scholes [3] and R. Merton [15] appeared. Departing from the *no arbitrage principle* and using the concept of *dynamic trading* these authors derived the – by now famous – *Black-Scholes formula* for the price of a call option. Maintaining for convenience the above hypothesis that the riskless rate of interest equals zero (which presently happens to be close to the actual economic situation), one obtains the formula

$$C = S_0 \Phi(d_1) - K\Phi(d_2), \tag{11}$$

with

$$d_1 = \frac{\ln(\frac{S_0}{K}) + \frac{\sigma^2}{2}T}{\sigma\sqrt{T}}, \; d_2 = \frac{\ln(\frac{S_0}{K}) - \frac{\sigma^2}{2}T}{\sigma\sqrt{T}}. \tag{12}$$

This formula looks quite different from Bachelier's result (6) above. However, for moderate values of T and σ, as was the case in Bachelier's original application, the difference of the numerical values of (6) and (12) is remarkably small. In [19] the difference for typical data used by Bachelier has been estimated to be of the order of $10^{-8}S_0$. In a way, this is not too surprising as the difference between Bachelier's model (3) and the Black-Scholes model (9) is analoguous to the difference between linear growth and exponential growth. In the short run this difference is remarkably small.

We do not give the derivation of the Black-Scholes formula here as it may be found in many textbooks (e.g. [6]). It is remarkable that the solution (11) for the option pricing formula is eventually obtained by applying precisely Bachelier's *fundamental principle*, i.e., by choosing $\mu = 0$ in (9) and calculating the expectation of the payoff of the option under the law of S_T.

3 Mathematics and the Financial Crisis

The Black-Scholes formula and the related concepts of *hedging* and *replication* of derivate securities had enormous impact on the paradigms of financial markets. In particular, the use of *stochastic models* became ubiquitous in the financial industry. In this section we shall have a critical look at the effects of this probabilistic approach to the real world.

3.1 Value at Risk

Let us start with the concept of *value at risk*. The CEO of J.P. Morgan, Dennis Weatherstone, asked the bank's quants in the wake of the 1987 crash to come up with a short daily summary of the market risk facing the bank. He wanted one single number every day at 4:15 pm which indicates the risk exposure of the entire bank. By that time the quants, i.e. the quantitative financial analysts, disposed of mathematical models for "market risk", such as the above considered price movements of stocks, options etc. Stochastic models were used to calculate the distribution of total profits or losses from these sources during a fixed period, e.g., the consecutive 10 business days. The *"value at risk"* was then defined as the 1 % quantile of this distribution, i.e. the smallest number $M \in \mathbb{R}$ such that the probability of the total loss being bounded by M, is at least 99 %. This was the famous "4:15 number".

It was quickly noticed by the quants that the above models of Bachelier and Black-Scholes are not well suited for the estimation of extreme events. After all, they are based on the Gaussian distribution which is derived from the central limit theorem.

As is very well known for almost 300 years, this theorem states that a random variable X, in our case the change of a stock price, which is the sum of "many" independent random variable X_n, where each of these random variables has little individual influence on the total effect $X = \sum X_n$, is approximately normally distributed. But in the financial world it happens quite often that a price movement is due to one big event (think, e.g., of 9/11) rather than due to the sum of many small events.

This is well illustrated by the following comment on the use of the Black-Scholes approach by a senior manager of an Austrian bank: "the Black-Scholes theory works very well, in fact surprisingly well, in 99 % or even 99.5 % of the days!" He then continued: "except for the one or two days per year which really matter."

But let us come back to the concept of *value at risk*. By choosing distributions with heavier tails it is not too difficult to correct for the above mentioned shortcomings of the Gaussian models. This was widely done, also by practitioners, in the context of risk management. As a general rule, when choosing a model it is always important to keep the applications in mind. If the purpose is to deal with the day-to-day business of pricing and hedging options, the Black-Scholes model, or even Bachelier's model, is a very efficient tool. However, when it comes to issues like risk management which deal with extreme events, the use of these models is highly misleading. After all, we have to keep in mind that these models were not invented for such purposes as risk management.

A similar fate of misuse happened to the "4:15 number" of Dennis Weatherstone which was originally designed as a very rough but focused information for the senior management of a bank. But this magic number quickly became very popular under the name of *value at risk* and used for to other purposes, notably to the calculation of capital requirements. A risky portfolio of a bank requires sufficient underlying capital as a buffer against potential losses. According to the Basel II regulation this capital requirement is determined by calculating the value at risk of the portfolio and, in order to be on the safe side, eventually multiplying this number by three. Compare [14] for a more detailed discussion.

The use of *value at risk* for regulatory purposes is a prime example of what has become known as "Goodhart's law" which seems to hold true in many contexts: *when a measure becomes a target, it ceases to be a good measure.*

If banks (or traders) get the incentive to design their portfolios in such a way that the "value at risk" is kept low, this may lead to serious misallocations. To sketch the idea we give a somewhat artificial example. Suppose that a bank has a portfolio which causes a sure loss of one million Euros. The bank can decompose this portfolio into 101 sub-portfolios where each of these sub-portfolios makes a loss of one million with probability $\frac{1}{101}$, and zero loss otherwise. While the value at risk of the entire portfolio obviously is one million, each of the sub-portfolios has a value at risk of zero! Hence no capital requirement is necessary for these sub-portfolios. This effect is, of course, in sharp contrast to the basic idea of "diversification": by pooling sub-portfolios into one big portfolio, the risk measure of the sum should be less than or equal to the sum of the risk measures, and not vice versa.

Admittedly, the above example is too blunt to be realistic, but nevertheless it highlights what is happening in practice when *value at risk* is blindly used as a risk measure for risk management purposes. Mathematically speaking, the above effect is due to the fact that the *value at risk map*, which assigns to each random variable X the $1\,\%$ -quantile of its distribution, fails to be sub-additive.

This shortcoming was soon and severely criticized in the academic literature. In 1999, Ph. Artzner, F. Delbaen, J.-M. Eber, and D. Heath [1] proposed a theory of "coherent risk measures" which do not suffer from this defect. According to *Google Scholar*, this paper has been cited more than 7000 times and there has been ample literature on this topic since.

Nevertheless, in practice the concept of *value at risk* still plays a central role for the determination of capital requirements.

3.2 The Gauss copula and CDO's

We now pass to a specific financial product which caused much harm during the financial crisis of 2007/2008, the so-called *collateralized debt obligations*, abbreviated CDOs. The basic idea looks quite appealing. In the banking and insurance business the notion of *risk sharing* plays a central role. If bank A is exposed to the risk of default of loan A and bank B to the risk of default of loan B, it is mutually beneficial if bank A passes over half of the risk of loan A to bank B and vice versa. This practice has existed for centuries and is the reason why, e.g., in the reinsurance business the financial damage of major catastrophes can be absorbed relatively smoothly by distributing the losses over several reinsurance companies.

Turning back to the example of bank A and B there is, however, a slight problem. As bank A negotiates with the obligor of loan A it disposes of better information on the status of this obligor than bank B. Of course, bank B is aware of this asymmetry of information which might work in favor of bank A, and therefore asks bank A for a higher recompensation when accepting half of the risk of loan A.

The original idea of a CDO is to find a mechanism which neutralizes this asymmetry of information. Suppose bank A has one thousand loans $A_1, \ldots, A_{1000}$ in its portfolio and wants to pass over part of the involved risk to other financial institutions or investors. Bank A can pool these loans into one big *special vehicle* and then slice it into *tranches*, e.g. a senior, a mezzanine, and an equity tranche. The tranching might divide the collection of one thousand loans according to the proportion 70 : 20 : 10. When loans fail to perform, the equity tranche is hit first. Only when the losses exceed 10 %, the mezzanine tranche is effected. When the losses exceed 30 %, also the senior tranche has to start to absorb them. The basic idea is that the issuing bank A keeps the equity tranche – which is most effected by the asymmetry of information – in its own portfolio, while it tries to sell the senior tranche and, possibly, the mezzanine tranche to other financial institutions.

So far, so good. In fact, similar instruments exist for a long time, e.g., the good old German *"Pfandbriefe"* which were intoduced in Prussia under Frederick the Great. Their business model goes as follows: a bank gives loans to communities which are secured by mortgages on their property. To refinance, the bank then sells bonds (the *"Pfandbriefe"*), which are directly secured by the entity of the underlying mortgages, to private or institutional investors. It is worth noting that there is one essential difference to the concept of CDOs: the issuing bank remains fully liable to the owners of the *Pfandbriefe*. This seems to be an important reason why the *Pfandbriefe* have safely survived so many financial crises. During the past hundred years there was not a single failure of a Pfandbrief-bank.

Back to the CDOs: in order to determine e.g. the price of the senior tranche one tries to estimate the probability distribution of the losses of this tranche. Of course, if one assumes that the defaults of the one thousand loans $A_1, \dots, A_{1000}$ are *independent*, the senior tranche has an extremely low probability of loss, even if each of the individual loans bears a relatively large default risk. But obviously nobody is so extremely naive to suppose independence in this context. Rather we expect some positive correlation of the failures of the individual loans. But how to model this dependence structure precisely?

D. Li [11] proposed in 1999 the *Gaussian copula* to handle this issue. For $0 \leq \rho \leq 1$, denote by $\mathbb{P}_\rho$ the centered Gaussian distribution on $\Omega = \mathbb{R}^{1000}$ defined in the following way. Denoting by $(X_i)_{i=1}^{1000}$ the coordinate projections, we prescribe $\mathbb{E}[X_i^2] = 1$, for each i, and $\mathbb{E}[X_i X_j] = \rho$, for each $i \neq j$. This covariance structure uniquely defines $\mathbb{P}_\rho$.

Now suppose that we know, for each $i = 1, \dots, 1000$, the individual default probability p_i of loan A_i. This is not too problematic as banks have, of course, a long experience dealing with the frequency of defaults of individual loans. For simplicity we suppose that all loans have the same size and either fully pay the loan (with probability $1-p_i$) or default totally (with probability p_i).

Denote by $x_i \in \mathbb{R}$ the $(1-p_i)$-quantile of the standard Gaussian distribution so that $\mathbb{P}[X_i > x_i] = p_i$. We identify the event $\{\textit{loan } A_i \textit{ defaults}\}$ with the event $\{X_i > x_i\}$. Having fixed $\rho \in [0, 1]$ as well as the p_i's we can now calculate all quantities of interest in an obvious and tractable way. For example, the probability that more than 30 % of the loans default and therefore the senior tranche suffers a loss is given by

$$\mathbb{P}_\rho[\#\{i : X_i > x_i\} > 300].$$

The delicate task is to determine the correlation parameter ρ. As regards the senior tranche it is rather obvious that (for realistic choices of p_i) its expected loss is increasing in $\rho \in [0, 1]$. Therefore it seems at first glance a reasonable approach to choose a realistic (i.e. sufficiently big) ρ by calibrating to observed prices on the market. This allows to calculate the price of the senior tranche as well as all the other quantities of relevance. In this way the rating agencies often granted a AAA to such senior tranches and other related products, obtained e.g. by pooling once again the mezzanine tranches of different CDOs into a new CDO (called "CDO-squared"). At least, they did so until 2007.

In 2007 it became very clear that the senior tranches of many CDOs were prone to suffer much bigger losses than predicted by "Li's formula". A Financial Times article in 2009 was entitled "The formula that felled Wall street" [12].

What had gone wrong? The sad fact is that David Li and other people applying the above method had not listened to people working in *extreme value theory*. In this theory it is well known that correlations of random variables tell only very little about the joint probabilities of extreme events. Only in the case of a (centered) *Gaussian* random variable X on $\mathbb{R}^{1000}$ the correlation matrix uniquely determines the law of X. As we shall presently see, it does so by giving rather small probabilities to *joint* extreme events, even if the correlation parameter ρ is close to 1.

This was made crystal clear in the paper [9] by P. Embrechts, A. McNeil, D. Straumann which has circulated since 1998. We give a short outline of the relevant concepts. Instead of one thousand random variables $X_1, \ldots, X_{1000}$ we focus for simplicity on X_1, X_2.

Definition 3.1 ([9]). *Let X_1, X_2 be random variables with distribution functions F_1, F_2. The coefficient of tail dependence is defined as*

$$\lambda := \lim_{\alpha \nearrow 1} \mathbb{P}[X_2 > F_2^{-1}(\alpha) | X_1 > F_1^{-1}(\alpha)],$$

provided the limit exists.

It is straightforward to calculate that for a *Gaussian* random variable (X_1, X_2) with $\rho(X_1, X_2) < 1$, we have $\lambda = 0$. This property is called *asymptotic independence* and has an obvious interpretation: whatever choice of

$\rho \in [0, 1[$ in the Gaussian case is made, the probability of *joint extreme events* becomes small, as $\alpha \nearrow 1$, quicker than the probability of the corresponding individual extreme events. This is in sharp contrast to what happened in the real world of 2007 to the loans pooled in CDOs.

But, of course, the Gaussian copula is not the only way of modeling. There are plenty of other ways to model the joint probability of a vector (X_1, X_2) for given marginal distributions X_1, X_2 and correlation $\rho(X_1, X_2)$. As an example the *Gumbel copula*, for which we obtain a strictly positive value of λ, is thoroughly analyzed in [9]. We note in passing that the word "copula" refers to the rather obvious fact, observed by A. Sklar in 1959, that for the specification of the joint law of (X_1, X_2) for given marginals, there is no loss of generality to normalize the marginal distributions of X_1 and X_2 to be uniform on $[0, 1]$.

The subsequent highly instructive picture is taken from the paper [9]. For *identical marginal distributions and identical correlation coefficient* ρ the choice of the copula can make a dramatic change to the probability of joint extreme events (the upper right rectangle).

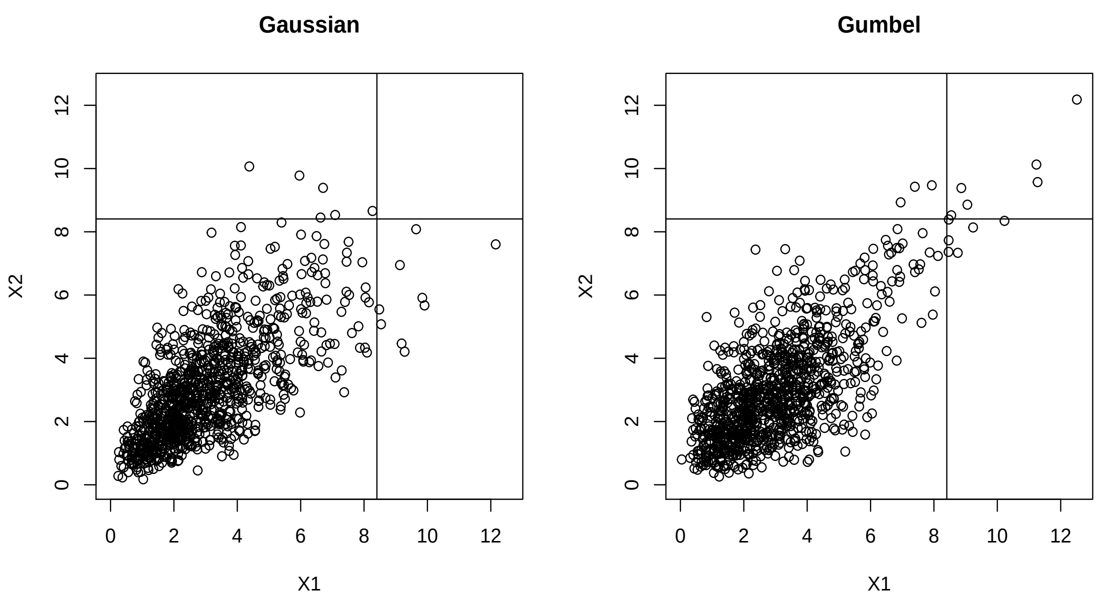

Figure 2: Figures showing 2000 sample points from the named copulas.

As P. Embrechts told me (and as is documented in [7]), he presented the paper [9] on March 27, 1999, at Columbia University. David Li was in the audience and introduced himself during the break. So much for the comments claiming that “nobody has warned.”

To finish this subsection let me mention another highly respected mathematician in the field of Mathematical Finance, L.C.G. Rogers, who also warned very outspokenly and long before 2007 about the misuse of the Gaussian copula. The following quotation on the 2007/2008 crisis [16] dates from 2009:

“ The problem is not that mathematics was *used* by the banking industry, the problem was that it was *abused* by the banking industry. Quants were instructed to build models which fitted the market prices. Now if the market prices were way out of line, the calibrated models would just faithfully reproduce those wacky values, and *the bad prices get reinforced by an overlay of scientific respectability*!”

3.3 An Academic Response to Basel II

In the previous subsections we have looked at two concrete and important examples, value at risk and the Gauss copula. Academic criticism of their misuse arose early and was well argued, but failed to sufficiently influence the practitioners.

Actually, this did not only happen in these two specific examples, as the paper [5] shows very clearly. This paper dates from 2001 and bears the title of this subsection. Written by a number of highly renowned financial economists and mathematicians, among them the above mentioned Ch. Goodhart and P. Embrechts, it was an official response addressed to the Basel Committee for Banking Supervision. It was extremely visible, not only within academia.

Let me quote from the Executive Summary of [5]:

- The proposed regulations fail to consider the fact that risk is endogenous. Value-at-Risk can destabilize an economy and induce crashes when they would not otherwise occur.

- Statistical models used for forecasting risk have been proven to give inconsistent and biased forecasts, notably underestimating the joint downside risk of different assets. The Basel Committee has chosen poor quality measures of risk when better risk measures are available.

- Heavy reliance on credit rating agencies for the standard approach to credit risk is misguided as they have been shown to provide conflicting and inconsistent forecasts of individual clients' creditworthiness. They are unregulated and the quality of their risk estimates is largely unobservable.

- Operational risk modeling is not possible given current databases and technology even if a meaningful definition of this risk were to be provided by Basel. No convincing argument for the need of regulation in this area has yet been made.

- Financial regulation is inherently procyclical. Our view is that this set of proposals will, overall, exacerbate this tendency significantly. In so far as the purpose of financial regulation is to reduce the likelihood of systemic crises, these proposals will actually tend to negate, not promote this useful purpose.

From today's perspective this reads like a clairvoyant description of the key issues of what went wrong in 2007-2008.

Let me try to make some personal comments on these five bullets.

Bullet 1: We have seen that for Bachelier as well as Black, Scholes, and Merton it was perfectly legitimate to model the risk involved in the price movements of a stock as *exogenous* and given by a stochastic model which is independent of the behavior of the agent. But the picture changes when all the agents believe in such a model or – making things worse – are forced by regulation to apply them. Value at risk plays an important negative role in this context.

Bullet 2 accurately summarizes what we have discussed in the above subsections 3.1 and 3.2.

Bullet 3 was strikingly confirmed by the crisis when the rating agencies, who did the above sketched ratings for the CDOs etc, turned out to have made very poor judgments of default probabilities. In addition, they may have been influenced by conflicts of interest.

Bullet 4: This is the only point which did not lead astray in 2007-2008. While the Basel II regulation of capital requirements for "operational risk", e.g. legal risks, IT failures etc, did not do much harm, it is important to note that is also did not do any good during the crisis.

Bullet 5 addresses the most fundamental issue, the procyclicality of financial regulation. While this is an inherent problem of regulation one should, of course, try to design the rules in a way to mitigate this effect. The prediction that, to the contrary, Basel II exacerbates the procyclicality has materialized in 2007-2008 in a dramatic way.

The final sentence of the introduction of the *Academic Response to Basel II* [5] could not have been more outspoken: “Reconsider before it is too late!” As we know today, the Basel Committee did not follow this urgent advice from academia.

The bottom line of these facts is that academia has not succeeded to influence financial practitioners sufficiently. Despite this rather sad story I do believe that academic research has to continue to try to thoroughly understand the problems at hand and to make itself understood in practice. To quote Sigmund Freud [10]: “The voice of the intellect is soft. But it does not rest before it has made itself understood. (Die Stimme des Intellekts ist leise, aber sie ruht nicht, ehe sie sich Gehör verschafft hat.)”

[1] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath, (1999), *Coherent measures of risk*, Mathematical finance, Vol. 9, No. 3, pp. 203–228.

[2] L. Bachelier, (1900), *Théorie de la Speculation.* Ann. Sci. Ecole Norm. Sup., Vol. 17, pp. 21–86. [english translation with excellent comments: M. Davis and A. Etheridge, *Louis Bacheliers Theory of Speculation: The Origins of Modern Finance*, Princeton University Press, 2006]

[3] F. Black, M. Scholes, (1973), *The pricing of options and corporate liabilities.* Journal of Political Economy, Vol. 81, pp. 637–659.

[4] J.-M. Courtault, Y. Kabanov, B. Bru, P. Crépel, I. Lebon, A. Le Marchand, (2000), *Louis Bachelier: On the Centenary of "Theorie de la Spéculation".* Mathematical Finance, Vol. 10, No. 3, pp. 341–353.

[5] J. Danielsson, P. Embrechts, C. Goodhart, C. Keating, F. Muennich, O. Renault, H.S. Shin, (2001), *An academic response to Basel II*, LSE Financial Markets Group. Special Paper Number 130.

[6] F. Delbaen, W. Schachermayer, (2006), *The Mathematics of Arbitrage*, Springer Finance.

[7] C. Donnelly, P. Embrechts, (2010), *The devil is in the tails: actuarial mathematics and the subprime mortgage crisis*, Astin Bulletin, Vol. 40, No. 1, pp. 1–33.

[8] A. Einstein, (1905), *Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen*, Annalen der Physik, Vol. 322, No. 8, pp. 549–560.

[9] P. Embrechts, A. McNeil, D. Straumann, (2002), *Correlation and dependence in risk management: properties and pitfalls*, Risk management: value at risk and beyond, pp. 176–223.

[10] S. Freud, (1913), *Totem und Tabu*, Internationaler Psychoanalytischer Verlag, Wien.

[11] D.X. Li, (1999), *On default correlation: A copula function approach*, Journal of Fixed Income, Vol. 9, No. 4, pp. 43-54.

[12] S. Jones, (2009), *Of couples and copulas: the formula that felled Wall Street*, April 24 2009, Financial Times.

[13] P. Jorion,(2007), *Value at risk: the new benchmark for managing financial risk*, Vol. 3, McGraw-Hill New York.

[14] A. McNeil, R. Frey, P. Embrechts, (2015), *Quantitative risk management: Concepts, techniques and tools*, Princeton university press.

[15] R.C. Merton, (1973), *Theory of rational option pricing.*, Bell J. Econom. Manag. Sci., Vol. 4, pp. 141–183.

[16] L.C.G. Rogers, (2009), *Financial mathematics and the credit crisis.* Document in response to questions posed by Lord Drayson, UK Science and Innovation Minister.

[17] P.A. Samuelson, (1965), *Proof that properly anticipated prices fluctuate randomly*, Industrial Management Review, Vol. 6, pp. 41–50.

[18] W. Schachermayer, Introduction to the Mathematics of Financial Markets. In: S. Albeverio, W. Schachermayer, M. Talagrand: Lecture Notes in Mathematics 1816 - Lectures on Probability Theory and Statistics, Saint-Flour summer school 2000 (Pierre Bernard, editor), *Springer Verlag, Heidelberg* (2003), pp. 111–177.

[19] W. Schachermayer, J. Teichmann, (2008), *How close are the Option Pricing Formulas of Bachelier and Black-Merton-Scholes?* Mathematical Finance, Vol. 18 (2008), No. 1, pp. 155-170.