The Physics of Finance

James Owen Weatherall, 2013

The first person renown for taking a mathematical interest in games of chance was the Italian Renaissance man **Gerolamo Cardano**.
A notable advance towards a theory of probability was made by Chevalier de Méré, a French writer, with the help of Blaise Pascal and Pierre de Fermat.

Jacob Bernoulli, a Swiss mathematician, figured out how to think about the relationship between probabilities and the frequency of events. Bernoulli showed that if the probability of getting heads is 50%, then the probability that the percentage of heads you actually got would differ from 50% by any given amount got smaller and smaller the more times you flipped the coin. You were more likely to get 50% heads if you flipped the coin 100 times than if you flipped it just twice. It has since been proven that if the chance of getting heads when you flip a coin is 50%, and you flip a coin an infinite number of times, then it is essentially certain that half of the times will be heads. This result is known as the law of large numbers, and underwrites one of the most important interpretations of probability.
If something is normally distributed and your sample is large enough, the sample's average value tends to converge to a particular number. This rule can be thought of as the law of large numbers for probability distributions, a generalisation of the principle discovered by Bernoulli, linking probabilities to the long-term frequencies with which events occur.
In practice: if something is governed by certain probability distributions, as men's heights are governed by a normal distribution, then once you have a large enough sample, new instances aren't going to affect the average value very much. Once you have measured many men's heights, measuring one more man won't change the average height very much.
Not all probability distributions satisfy the law of large numbers however (ex: with Cauchy distributions, the average value is highly unstable, as a new instance could dramatically affect the average; as a matter of fact, even the long-term average is unpredictable).

What the field of finance understood is that you need to start with the simplest theory that works, get as far as you can, and then ask where the theory has gone wrong. In this case, once you have established that stock market prices are random, at least in some sense, the next step is to assume that they are random in the simplest possible way: that they just follow a "random walk". This is what **Louis Bachelier** did (beginning of 20C); he provided the first model for how market prices change with time, the **random walk model**, which states that stock prices are normally distributed.

A normal distribution, aka a Gaussian distribution, is represented by a bell curve.

Maury Osborne (worked at the Naval Research Lab) began to look at stock prices and noticed that they behaved like a collection of particles moving randomly in a fluid; they were exhibiting Brownian motion (1959 paper entitled "Brownian Motion in the Stock Market). He then pointed out that Bachelier's model could not be right, since it would mean that stock prices could become negative after a sufficiently long period of time, and so he complicated the model ever so slightly by suggesting that it's the market rates of return which follow a random walk: rates of return are normally distributed, not prices. Since price and rate of return (average percentage by which the price changes each instant) are related by a logarithm, his model implied that the probability distribution of stock prices would be given by log-normal distributions.

Osborne later devised an "extended Brownian motion model" taking volume changes into account.

He also noticed observed that a great preponderance of ordinary investors placed their orders at whole-number prices (ex. 10, or 11).

But he ultimately concluded that there was no profit to be had on speculating on the unrelieved bedlam of financial markets.

Then **Benoit Mandelbrot** (affiliated with IBM) discovered the shortcomings of the Bachelier-Osborne approach; he saw a different pattern than Osborne's in the randomness of prices, although not radically different. Still, in his view the world would be more wildly random, by considering that markets were Levy-stable distributed (1965 article).
→ Mandelbrot revealed that financial markets were governed by fat-tail distributions.

The volatility of Levy-stable distributions being infinite, most standard statistical tools didn't apply for analysing such distributions.

The differences only become important in the context of extreme events. As on a typical day, there aren't going to be any extreme event (according to either theory), you usually won't notice much of a difference between the two models. For this reason, when it came for economists to put the randomness of stock market prices to work by using statistics to predict derivative prices or to calculate the amount of risk in a portfolio, they had to pick between the simple theory that gave good results the vast majority of the time and the more cumbersome one that better accounted for certain extreme events. In doing so they could solve problems that otherwise couldn't be solved, and get a solution that it quite close to correct.

Mandelbrot also introduced the concept of fractals.

The normal and Cauchy distributions are both examples of **Lévy-stable distributions**, but Lévy showed that there is a spectrum of randomness ranging between the two. Wildness can be captured by a number, usually called alpha, which characterises the tails of a Lévy-stable distribution: normal distributions have an alpha of 2; Cauchy distributions have an alpha of 1 ("drunken firing squad model"). The lower the number, the more wildly random the process (and the fatter the tail. Distributions that have alpha of 1 or less don't satisfy the law of large numbers – in fact, it isn't possible to even define the average value for a quantity that wild. Distributions with alpha between 1 and 2, meanwhile, have average values, but they don't have a well-defined average variability – what statisticians call volatility or variance – which means it can be very hard to calculate an average value from empirical data, even when the average exists.

Edward Thorp (mathematics instructor at MIT) showed that physics and mathematics could be used to profit from financial markets. Building on the work of Bachelier and Osborne, and on his own experience with gambling systems, Thorp invented the modern hedge fund, by applying the new field of information theory.

Claude Shannon (worked at Bell Labs on an encrypted telephone system in WW2) founded the entirely new science of information theory (with a landmark paper that he published in 1948), which is the mathematics (and electrical engineering) behind the digital revolution – it undergirds computer science, modern telecommunications, cryptography, and code-breaking. The basic object of study is data: bits (a term Shannon coined). The basic principle is that the amount of information gained depends on how unpredictable that information/signal was; by connecting information with probability, Shannon discovered a way to assign a number to a message that measures the amount of information it contains, which in turn was the first major step in building a mathematical theory of information.

Edward Thorp met Claude Shannon at MIT. When Thorp mentioned the money management problem to Shannon (in a gambling situation, how much to bet in function of the advantage/partial information), Shannon directed Thorp to a paper written by one of his colleagues at Bell Labs named John Kelly Jr., which provided the essential connection between information theory and gambling: the equation now called the Kelly criterion or Kelly bet size: the percentage of money to bet on any given outcome is advantage/payout.

Thorp ushered in a switch in money management to quantitative, computer-driven methods. Upon publishing Beat The Market in 1967, he came up with a way to calculate a warrant's "true" price (he saw a relationship able to give him an "advantage", one between warrants and stock prices – considering the latter log-normally distributed) and invented a strategy now called delta hedging (ex: shorting warrants while buying the underlying stocks), and he created a hedge fund with a partner.

He met Warren Buffet, the money manager of one of his potential clients (the Dean of the graduate school where he had a job).

Fischer Black is one of the most famous figures in the history of finance. His most important contribution, the Black-Scholes model of option pricing (sometimes called Black-Scholes-Merton) in 1973, remains the standard by which all other derivatives models are measured. He made quantitative finance, with its deep roots in physics, an essential part of investment banking.

Before that, he met **Jack Treynor**, who had developed a new way of understanding the relationship between risk, probability and expected value, now known as the Capital Asset Pricing Model (CAPM). This model linked risk and return, via a cost-benefit analysis of risk premiums. CAPM was an equilibrium theory, insofar as it described economic value as the natural balance between risk and reward. CAPM formed the foundation for virtually all of the work Fischer Black would go on to do.

To start with, Fischer Black found a fundamental relationship that would ultimately give rise to the Black-Scholes equation; his approach was to find a portfolio consisting of stocks and options that was risk-free, and then argue by CAPM reasoning that his portfolio should be expected to earn the risk-free rate of return. Black's strategy of building a risk-free asset from stocks and options is now called dynamic hedging.

Fischer Black and Myron Scholes later worked on a model of option pricing together (1970), and their result was confirmed by Robert Merton (although he had taken an entirely different starting point). This gave investment banks the tool they needed to manufacture options: they could sell options and buy other assets in such a way that at least theoretically they didn't carry any risk.

However, both Thorp's and the Black-Scholes-Merton options models were based on Osborne's random walk hypothesis, which amounted to assuming that rates of return are normally distributed. So they did not effectively account for extreme events, such as the Black Monday stock market crash of 1987, during which world financial markets fell more than 20% literally overnight.

Michael Greenbaum and Clay Struve, working at a company called O'Connor and Associates, internally improved the Black-Scholes model, to take into account things like sudden jumps in prices (and thereby dealt with a price fluctuation called the "volatility smile"). But their findings were not made public.

Mandelbrot left finance at the end of the 1960s, but he returned in the 1990s, as the Black-Scholes model was contested (apart from its inadequacy relative to extreme events, discrepancies in pricing options began to appear, before the crash but mostly after, such as its inability to cope for the "volatility smile"). His claim, that probability distributions describing market returns have fat tails, was gaining traction.

Emmanuel Derman later discovered a way of modifying the Black-Scholes model to account for the volatility smile, and thereby, large shifts in prices (1994).

James Doyne Farmer and **Norman Packard**, experts in nonlinear forecasting (seeking to identify predictive patterns in apparently random phenomena), set out to predict the behaviour of financial markets directly (not that of derivatives). In Packard's words, it involved identifying the order at the *edge of chaos* (=the small windows of time in which there was enough structure in a chaotic process to predict where a system would go next).

The ideas at the heart of their work were first developed by **Edward Lorenz**. Working in meteorology, his own computer simulations (one of the very first computer simulations in service of a scientific problem) revealed (by mistake) that extremely small differences in initial conditions could alter the end results dramatically. Despite the fact that Lorenz' system was entirely deterministic, wholly governed by the laws he programmed, extremely small differences in the state of a system at a given time would quickly explode into large differences later on. In 1960s, Lorenz called his theory sensitive dependence on initial conditions, which captured the essence of chaotic behaviour (the word *chaos* came later, with the work of James Yorke and Tien-Yien Li who published a paper called "Period Three Implies Chaos" in 1975).

But their work at the Prediction Company (founded in 1991) was not directly using chaos theory; instead, their experience in that field gave them:

- an excellent understanding of how complex systems work
- the ability to use computers for mathematics
- an appreciation of how regular patterns (patterns with real predictive power) could be masked by the appearance of randomness
- the means to apply the right statistical measures to identify truly predictive patterns
- the means to test data against their models of market behaviour
- the means to figure out when those models were no longer doing their job

They were at ease with the statistical properties of fat-tailed distributions and wild randomness, which are characteristic of complex systems in physics as well as financial markets, so that they could easily apply some of Mandelbrot's ideas for risk management in ways that people with more traditional economics training could not.

They were trying to extract small amounts of information from a great deal of noise – a search for regularities / patterns.

One strategy they used was statistical arbitrage, which works by betting that certain statistical properties of stocks will tend to return even if they disappear briefly (ex: applied for pairs trading: when 2 usually closely-correlated stocks go out of synch, bet on the return of that correlation). Although this had been done before, they used a novel approach, based on genetic algorithms (assess the results of a wide variety of experimental configurations, select the most successful solutions and recombine them in novel ways to produce a second generation of solutions, which are allowed to compete again; and then proceed iteratively) and optimisation algorithms (different algorithms tailored to different tasks).

Another idea they had was to use several models at once (each based on a different set of simplified assumptions about the statistical properties of various assets) and see which trades were ranking high on most models.

This kind of modeling, where one uses algorithmic methods to identify optimal strategies, is often called "black box modeling" in the financial industry. Black Box models are very different from models like Black-Scholes and its predecessors, whose inner workings are not only transparent but often provide deep insights into why the models (should) work. Black Box models are much more opaque, and as a result they are often scarier, especially to people who don't understand where they come from or why they should be trusted. By definition it is impossible to pinpoint why they work, or to fully predict when they are going to fail. This means that black-box modellers don't have the luxury of being able to guess when the assumptions that have gone into their models are going to be wrong. In place of this sort of theoretical backing, the reliability of black-box models has to be constantly tested by statistical methods. But at the end of the day, data outclass theory; any model has to be tested by the same kind of statistical methods, and theoretical backing can give a false confidence otherwise.

The Prediction Company was one of the very first companies to build an entire business model based on those Black Box models; it was a whole new way of thinking about trading. The company later made a deal with O'Connor and Associates, to fund themselves.

The efficient market hypothesis entails that sophisticated investors can adopt trading strategies exploiting patterns when these arise, temporarily, until the market self-corrects and the pattern washes out, thereby playing a role in making markets random, and efficient (by correcting small-scale deviations from the perfect theoretical efficiency).

(But the efficient market hypothesis is still flawed, as speculative bubbles and crashes still appear.)

Didier Sornette worked in identifying patterns endemic to the structures of complex systems and using these patterns to predict critical phenomena – ruptures, quakes, crashes – in more than a dozen fields (material science, geophysics, decision theory, neuroscience, financial markets, etc.). He understood that ruptures require a high degree of structure and a capacity for coordinated action: they require a mechanism for system-wide feedback and amplification, something to transform an otherwise small event into a large one (following the pattern of fractals) → <u>self-organised criticality</u> (Per Bak, 1996).

(metaphor: an inflated balloon would be much more sensitive to the same events that would otherwise have no impact)

He and others developed a detailed model for how fractures and cracks percolate through a material, accounting for how degrees of organisation and coordination could serve to amplify fractures, to turn small causes into large effects.

In 1991, he had the idea that a rupture would be preceded by smaller events, following a very specific, accelerating pattern defined as <u>log-periodic</u> because the time between the smaller events decreases in a particular way (related to the logarithm of the time). He initially tested his theory on the pressure tanks of rockets and then earthquakes, by looking for log-periodic patterns in the data in order to predict the time of a rupture. He then identified such a pattern in the stock market, and managed to benefit from the worldwide crash on 27 October 1997 – explained by historians as a reverberation effect.

Critical phenomena often have what physicists call <u>universal properties</u>; despite the profound differences in their microscopic details, under certain circumstances they exhibit the exact same larger-scale behaviours. Certain universal laws seem to apply at a statistical level; they govern coordination between parts, irrespective of what the parts happen to be.

In 1994, Didier Sornette co-authored a paper with **Jean-Philippe Bouchaud** on how to price options even if the underlying stock does not follow the kind of random walk assumed by Black and Scholes. This effectively extended the theory of options pricing to more sophisticated models of price changes, including those with fat-tail distributions (O'Connor and Associates had already done work along those lines but only internally).

In 1996, Sornette, Bouchaud and Anders Johansen published a paper, extending Sornette's earlier work on predicting earthquakes and ruptures to predict market crashes.

The central idea behind the market-crash-as-critical-event hypothesis involved <u>collective action</u>, or <u>herding behaviour</u>. This is hardly surprising, as the suggestion that market crashes have something to do with mob psychology is old: in 1841, Charles Mackay wrote a book on, among other things, economic bubbles (market conditions under which prices become entirely divorced from the value of the things being traded), which he called *Extraordinary Popular Delusions and the Madness of Crowds* – ex. tulip mania in the Netherlands, in 1635-1637. The question is then why under some circumstances herding seems to take over. How does something like tulip mania strike? When do the normal mental brakes that would keep someone from spending his entire life savings on a tulip bulb give out? Sornette doesn't have an answer to this question, although he has developed some models that predict which circumstances will lead herding effects to become particularly strong. What Sornette can do is identify when herding effects have taken over; this amounts to identifying when a speculative bubble has taken hold in a particular market and to predicting the probability that the bubble will pop before a certain fixed time (the critical point).

Sornette also predicted the crashes in 1998, following the collapse in the Russian ruble (he claimed that even though the largely unanticipated Russian debt default may have triggered the market turmoil that summer, the crash showed the log-periodic precursors characteristic of herding behaviour, meaning that it would have occurred even without the ruble collapse), 2000 (dot-com crash), 2008.

Extreme financial events happen for the same reasons as more mundane events; big market drawdowns, at their core, are just smaller drawdowns that didn't stop. Which entails that there is no way to predict catastrophes. Indeed, self-organisation, one of the principal parts of the theory of critical phenomena, is usually associated with just the kind of fat-tail distributions that make predicting extreme events so difficult. The three physicists who first introduced the notion of self-organisation, Per Bak, Chao Tang, and Kurt Wiesenfeld, took their discovery as evidence that extreme events are, in principle, indistinguishable from more moderate events. The moral, they thought, was that predicting such events was a hopeless endeavour. This concern is at the heart of Nassim Taleb's argument against modeling in finance. In *The Black Swan*, Taleb introduces the concept of *Black Swans*, events that are far from standard, normal distribution expectations – unpredictable – but with a substantial impact. Taleb takes it to be a consequence of Mandelbrot's arguments that these kinds of extreme events, the events with the most dramatic consequences, occur much more frequently that any model can account for. To trust a mathematical model in a wildly random system like a financial market is foolish, because the models exclude the most important phenomena: the catastrophic crashes. Sornette introduced a new term for extreme events: he calls them *dragon kings*. They bear dramatic consequences, but unlike Black Swans, you can hear them coming, issuing warnings in the form of log-periodic precursors, oscillations in some form of data that occur only when the system is in the special state where a massive catastrophe can occur. These precursors arise only when the right combination of positive feedback and amplifying processes is in place, along with the self-organised criticality necessary to lead to a rupture.

The methods Sornette has used to identify bubbles and predict when crashes will occur can also be used to identify a situation that Sornette has called an anti-bubble (where stock prices are artificially low), and determine by how much prices are likely to increase. In 1999, based on his observation of log-periodic patterns in the market data, he accurately predicted that the Nikkei would increase by 50% by the end of the year, despite mainstream economist sustaining the contrary.

The Prediction Company, on the one hand, and Sornette, on the other, offer two ways to fill in the gaps in the now-standard Black-Scholes-style reasoning. The Prediction Company's methods might be thought of as local, in the sense that their strategy involved probing the financial data in search for patterns that have some temporary predictive power.

Sornette, conversely, has taken a more global approach, looking for regularities that are associated with the biggest events, the most damaging catastrophes, and trying to use those regularities to make predictions. If Mandelbrot's work can be regarded as a revision to the early accounts of random markets, then Sornette's proposal is a second revision. It is a way of saying that even if markets are wildly random and extreme events occur all the time (he actually believes that catastrophic crashes happen even more than Mandelbrot proposed), at least some extreme events can be anticipated if you know what to look for.

In 1996, Pia Malaney, working in economics and her fiancé Eric Weinstein, in mathematical physics, developed a novel way to solve the index number problem by adapting a tool from mathematical physics known as gauge theory.

The index number problem is concerned with how to take complex pieces of information about the world and turn them into single numbers that can be used to quantify evolutions in time. Index numbers provide a standard to compare economic indicators over time and from place to place. For example, how to best quantify the evolution of the CPI (Consumer Price Index)? How to compare a "reasonable" market basket in 1950 to a "reasonable" market basket in 2010, especially for people with different lifestyles, like a lumberjack and a computer programmer?

Einstein's theory of general relativity makes essential use of the fact that space and time are curved in the sense that parallel transport (the direction of transport is preserved) is path dependent. If you begin with an arrow at one place and then move it around a path that brings it back to the starting point, it might face a different direction. But it will have the same length; **Hermann Weyl** (who was an academic at ETH Zürich, in the same place as Einstein for a year, between 1913 and 1914) came up with an alternative, extended theory in which *length*, too, was path dependent. He called it gauge theory, based on the idea that a mathematical standard would enable comparisons to be made no matter the reference frames the physical quantities are measured in. Weyl's essential innovation was to consider a mathematical theory that accounts for path dependence, for comparing otherwise incomparable quantities.

Even though the theory he put forward was not a success (wasn't consistent with experimental results), gauge theory was resurrected in the 1950s by a pair of researchers called **C. N. Yang** and **Robert Mills**, who took it one step further; if it was possible to construct a theory in which length was path dependent, it would be possible to construct theories in which other quantities were path dependent. They developed a general framework for much more complicated gauge theories than the one Weyl imagined, now known as Yang-Mills theories. These spawned what is sometimes called the gauge revolution; beginning in 1961, fundamental physics got rewritten in terms of gauge theory, a process that only accelerated when Yang, in collaboration with **Jim Simons** (the mathematical physicist turned hedge fund manager who founded Renaissance Technologies in the 1980s), realised a deep connection between Yang-Mills theories and modern geometry later that decade. Gauge theories use geometry to compare apparently incomparable physical quantities, or, here, economic variables.

Gauge theories proved particularly important in physics because they proved to be a natural setting to look for unified theories, where what was being unified was the standard by which different quantities were compared in the theories. By 1973, it appeared that the three fundamental forces of particle physics – electromagnetism, the weak force and the strong force – had been unified into a single gauge-theoretic framework, called the Standard Model of particle physics.

The financial crisis in 2008 prompted Eric Weinstein to call for a new large-scale collaboration between economists and researchers from physics and other fields: an "economic Manhattan Project", to broaden the economists' theoretical framework to account for a wider variety of phenomena. If born to fruition, their ambitious project would translate into a new theory of economics, inspired by ideas developed in physics. But divisions between economists and funding issues are plaguing the initiative.

In a 1965 Supreme Court decision on freedom of speech (*Lamont vs. Post-master General*, found in Sepinuck and Treuthart, 1999), Justice William Brennan coined the expression "marketplace of ideas" to describe how the most important insights might be expected to arise out of a free and transparent public discourse. If this is right, then you would expect that the best new ideas about [economics] would get taken up – even if powerful [economists] rejected them.

The sociologist Donald MacKenzie observed that financial models are as much the engine behind markets as they are a camera capable of describing them, as their implementation generates feedback loops influencing the behaviour of markets.