CONSIDERING TIME AS THE RANDOM VARIABLE: A NEW POINT OF VIEW FOR STUDYING FINANCIAL TIME SERIES

GILLES ZUMBACH
Olsen & Associates, Research Institute for Applied Economics, Seefeldstr. 233, CH-8008 Zürich, Switzerland

1. Introduction

In finance the decade 1990 to 2000 is characterized by a move from low frequency data, typically weekly or daily, to high frequency intraday tick-by-tick data. This was made possible mainly by the availability of cheap and powerful computers and by communication networks (in turn, this technological improvement is changing high frequency data into a commodity!). High frequency data means a real quantitative jump: when going from daily to intraday data, the amount of information increases by a factor of 100 to 10 000 (for the FX rate DEM/USD). By tapping this added information, financial calculations can be improved, for example in risk assessment, portfolio optimization or forecasting.

The first big difficulty when moving into the intra-day realm is seasonalities. Clearly, because markets open and close, we can expect strong seasonalities with daily and weekly periods. These seasonalities are a real nuisance, because they are strong and predictable, and will hide completely other finer structures of the process. Since it is precisely these finer properties that we are interested in, seasonalities must be treated first.

Discounting for seasonality is not new in time series analysis, for example the regularly published unemployment rates are deseasonalized for yearly variations. But most of these classical techniques rely on the assumption that the time series is regularly spaced in time. Moreover, they are not very flexible, for example holidays can not be taken into account. For these reasons, we have devised a method based on a change of time scale that will be presented first. Throughout this paper, it should be kept in mind that our ultimate goal is to use high frequency data to forecast prices and build trading models. This implies hunting for some subtle dependencies, and it explains why we are quite picky about 'trivial seasonalities'.

Then, after treating the seasonalities, we will analyse the first hitting time, namely the random time interval before a fixed return is reached. This will shed light on extreme events and their scaling properties. The examples and figures will be taken from the FX market but the same technique and analysis can be used for other markets as well.

2. Notations and Definitions

Tick-by-tick data contains a time stamp t and bid and ask prices p_{bid}, p_{ask}. We will consider mainly the logarithmic middle price x as the primary useful time series

$$x = \frac{1}{2} \left\{ \ln(p_{\text{bid}}) + \ln(p_{\text{ask}}) \right\} . \tag{1}$$

Time series are denoted with a simple letter, like x. The value at time t of a time series x is denoted by $x(t)$. If the time series depends on a parameter p, it is denoted $x[p]$.

The return in a time interval Δt is $r[\Delta t](t) = x(t) - x(t - \Delta t)$. Because of the ln in the definition of x, this corresponds to a relative return (up to second order correction). The absolute value of the return $|r[\Delta t]|$ is often taken as an estimator of the volatility at time scale Δt. Because of the strong tail of the price distribution, and because of the possible noise in the data sources, this is a more robust estimator of the volatility than the usual square return.

3. Seasonality

A time series is called *seasonal* if it exhibits a deterministic periodic pattern. There are essentially two different ways to detect seasonal patterns:

- Intra-day and intra-week statistics.
  An intra-day statistics for x is an average, conditional to the time in the day $< x \mid t \bmod 24 >$. We will denote symbolically the intra day time by $t \bmod 24$, meaning the elapsed time since last midnight. Similarly, an intra-week statistics corresponds to $< x \mid t \bmod 168 >$ (168 is the number of hours in a week).
- lagged autocorrelation function.
  The lagged autocorrelation function examines if there is a linear dependence between the current and past values of a variable.

$$\rho[\tau] \quad = \quad \frac{\sum [x(t-\tau) - \overline{x}][x(t) - \overline{x}]}{\sqrt{\sum [x(t-\tau) - \overline{x}]^2 \sum [x(t) - \overline{x}]^2}}$$

The lagged autocorrelation function peaks at time intervals corresponding to the periods of seasonal patterns.

Note that both of these statistics are sensitive to seasonalities, although not in identical ways. The autocorrelation assumes no a-priori periodicity. It is also sensitive to dependences where a 'trigger' event appears at a random time, followed by an 'echo' event at a fixed time interval τ later. On the other hand, an intra-day statistics shows the timing of the seasonality, for example the opening of a market.

Turning to empirical data, the lagged autocorrelation for the return and volatility is shown in Fig. 1. The return does not seem to exhibit any seasonalities, but the volatility clearly shows daily and weekly peaks, i.e. there is *seasonal heteroskedasticity.* We also observe a small eight-hourly seasonality of the volatility. Similarly, we have detected strong seasonalities in the volatility, the tick frequency and the spread for both FX and IR rates. The same volatility data are used to build an intra-week conditional average shown in Fig. 2. Again, both the daily and weekly seasonalities clearly appear.

From the forecasting point of view, these seasonalities are a real nuisance. What is the point of making a sophisticated forecaster, if the main outcome is to predict a substantial increase of volatility each morning, corresponding to the opening of the market! Moreover, these very strong seasonality signals completely hide other properties of the processes. For example, it is known that financial data exhibit a slow decay of the volatility correlation, corresponding to volatility clustering. This property of the volatility has prompted the development of the ARCH process, and all its descendants, like GARCH, HARCH, ... This volatility clustering should be visible on the lagged autocorrelation graph, but clearly the seasonalities are dominant.

Therefore, the leading signal originates in the seasonalities, and we want to remove them. In order to do so, the main idea is to **change the time scale**. In fact, this is an old and common practice in finance; for example, daily sampled data have 5 days per week, Saturday and Sunday being eliminated. Another common practice when studying hourly intra-day stock data is to divide the 'market day' regularly from the opening to the closing of the stock exchange, in effect discarding the closed period from the time scale. Implicitly, these two procedures define new time scales adapted to common financial processes, and this kind of off-handedness with time caused some commotion to physicists moving to finance. As the FX market is open 24 hours a day, we have to be a bit more sophisticated, but the main idea is similar, namely to expand the period with high activity and contract the period with low activity. Clearly, this corresponds to defining a proper financial time scale for each financial object. As we start mangling with

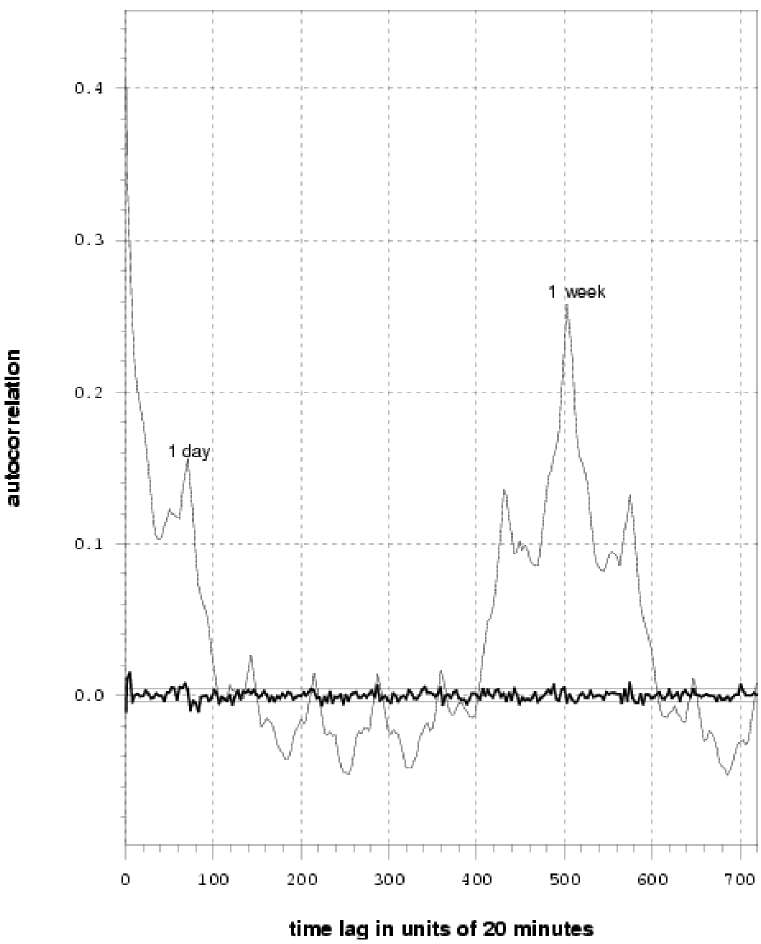

Figure 1. Autocorrelation of 20 minute returns (bold curve) and absolute values of returns (thin curve), computed on the physical time scale (uniform time, no special treatment of weekends and Holidays). FX rate: USD-DEM; sampling period: from 2 June 1986 to 1 June 1993. The confidence limits represent the 95% confidence interval of a Gaussian random walk.

time, and defining time scales other than the usual physical time, several questions arise. Eventually, we will reach the fundamental question: what is time, and which time scale should we use in finance? The next section is a small 'divertimento' around this theme, without any pretention to answer the question, but more as an invitation for further thought.

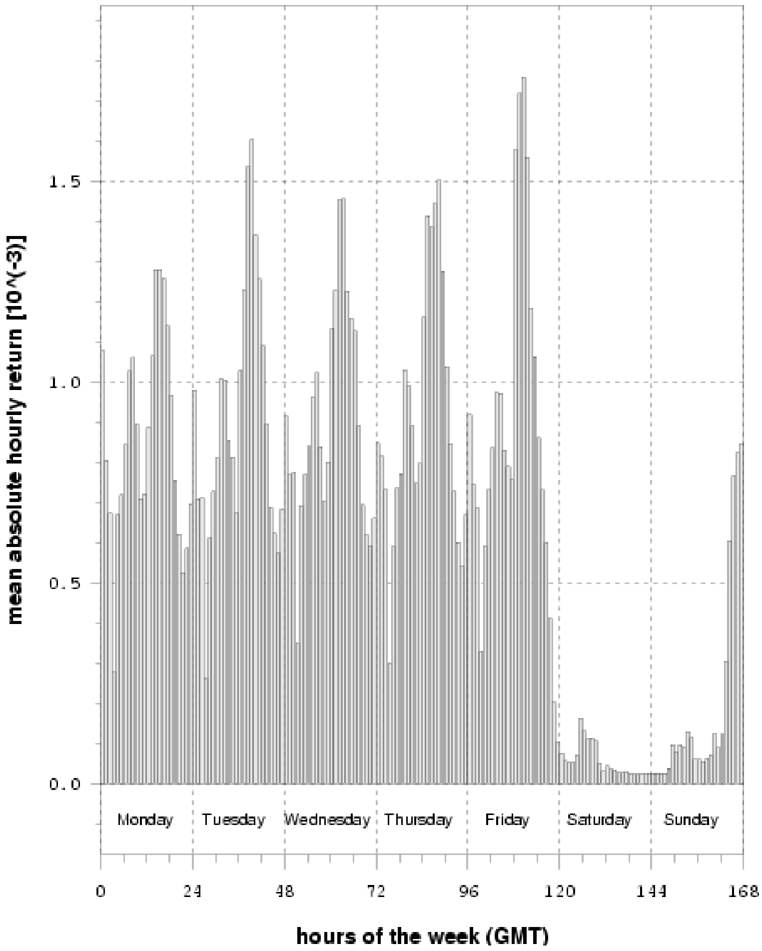

Figure 2. The Intra-weekly histogram of mean absolute hourly returns. The hourly intra-week grid is in Greenwich Mean Time (GMT). No adjustment has been made for daylight saving time. FX rate: USD-DEM; sampling period: from 3 Feb 1986 to 26 Sep 1993 (398 full weeks).

4. Some Reflections about Time

With such a big question to ponder at, a first lead is to reach for the dictionary in order to search for a definition of time, and to read what linguists have to say about it. First, I invite the reader to pause and think about his/her own definition of time, and then to get his/her favorite dictionary and read the definition.

A FEW DEFINITIONS OF "TIME"

You will find below short excerpts from a random sampling of dictionaries. A common point about all dictionaries is that they give a very long definition for time, typically 5 to 10 times longer than the average definitions. For this reason, I kept only the main headings.

- **The Oxford English Dictionary**
  - A limited stretch or space of continued existence, as the interval between two successive events or acts, or the period through which an action, condition, or state continues.
  - A particular period indicated or characterized in some way.
  - ... (52 definitions, 5 pages)
- **Oxford Advanced Learner's Dictionary**
  - All the years of the past, present and future (*The world exists in space and time*).
  - Passing of these taken as a whole (*Time has not been kind to her looks, i.e. She is no longer as beautiful as she was*).
  - Indefinite period in the future.
  - Portion or measure of time.
  - ... (1.5 page)
- **Webster Dictionary**
  - The system of those sequential relations that any event has to any other, as past, present, or future; indefinite and continuous duration regarded as that in which events succeed one another.
  - Duration regarded as belonging to the present life as distinct from the life to come or from eternity; finite duration.
  - A system or method of measuring or reckoning the passage of time.
  - ... (1 page)
- **Micro Robert** [1]
  - Indefinite medium in which lives in their changing states, events and phenomena seem to take place according to a definite order.

[1] The original French definition is:

- Milieu indéfini où paraissent se dérouler selon un ordre les existences dans leur changement, les événement et les phénomènes.
- Durée, portion limitée de cette durée.
- Point repérable dans une succession par référence à un ‘avant’ et un ‘après’.

- Duration, limited portion of it.
- A reference point in a succession according to a 'before' and an 'after'.
- ... (0.5 page)

What appears from these definitions is that the word 'time' is used for two concepts. One is for a point along the 'absolute' time axis, as in *today at 1 o'clock*. The other meaning is for a time interval, as in *in 1 hour*. The german language has two different words for these two different concepts: *die Uhr* (a time point) and *die Stunde* (a time interval). Some English authors also advocate the use of two different words, for example 'epoch' for a time point. Another reflection about these definitions is that each one of us certainly has quite a deep understanding of what time is, and a strong feel for the ineluctable passing of time. But if there exist people who do not, I wonder if the above definitions will be of any help!

A PHYSICIST'S VIEW OF TIME

Physicists are certainly big users of the time concept. Time appears in all basic equations: Newton, Schröndiger, heat, relativity, etc... In order to have the equivalent of the above dictionary definitions but for physicist, I consulted the 'Encyclopedia of Physics' (Lerner and Trigg, 1991), which contains an article on time writen by P.C.W. Davies. This article about time is in a remarkable contrast to the rest of the encyclopedia: it does not contain any equations or greek letters, it does not provide for a definition but assumes the knowledge of time, and it is more about open questions than a beautiful answer provided by science. As it is also quite a long article, I reproduce below only its very beginning and end. The missing parts contains a discussion of fundamental questions and topics such as: is there a beginning and end to time, is there a minimum granularity to time, the arrow of time as provided by the increasing entropy, time is measured (created ?) by events therefore if there is nothing, is there still a time, etc...

Encyclopedia of Physics

"Time is the most fundamental aspect of our experience. Its properties have intrigued and baffled philosophers, theologians and scientists for centuries. Question about the nature of time, its global and microscopic structure, continue to fill books and provoke experimental research. ...

Psychological Time

One fundamental aspect of time has not appeared in the physical description already given. The human mind divides time into past, present and future. Everyone is aware of a flow, or flux, of time toward the future. This *flow appears to sweep the past out of existence and bring the future into being. Such temporal activity is not a feature of time in physics, at least as understood at present. Its appearance in the human mind is sometimes dismissed as purely psychological, and sometimes conjectured to indicate a great unsolved problem in physics. In either case it is a mystery."*

Again, if someone does not know what time is, I am afraid that this article will be of no more help than the dictionaries! Being more down to earth, the every day physical time corresponds essentially to

– The variable t appearing in various equations. However, it is not so obvious why, say, the t in the heat equation is the same as the t in the Schrödinger equation and why a wrist-watch still measures the same t! Indeed, it is quite non trivial that the same concept is used to index the advance of events in the quantum, classic and relativistic realm.
– Practically, the measure of time passing by corresponds to counting the beats of some *regular* pendulum. Moreover, pendulums with widely different periods can be related simply by counting, with ranges from an atomic clock to the rotation of celestial bodies. The location of the clocks can also differ, with a distant pendulum being in the next room, on a satelite or on a remote star. Essentially, a pendulum is needed to reckon the advance of time (which can be followed by the question, relevant in finance: if there is 'nothing', is there still a time?).

The uniqueness of the time concept across realms and scales is what makes it so pervasive in physics. Turning now to finance, it is not clear why the same physical time concept must be valid in this field. Indeed, there is no regular 'financial' pendulum, and an economic market is not driven by physical laws. In physics, one is able to construct and reproduce a pendulum, and all copies will measure essentially the same time. In finance, each event is unique and non reproducible. Yet we still want a scale that measures the passing of events.

Another concept of the time scale is the *psychological time*. Each one of us has his own psychological time scale, where important events take a relatively larger span — say for example your wedding — and unimportant days are shrunk to zero — like a dull eventless November day five years ago. Then, you can have a 'before' and 'after' your wedding, for example. In this way, each individual constructs his/her personal time scale from his/her own appreciation and recollection of the lived events. The physical time scale is used to relate different psychological time scales, for example my wedding to the birth of your first kid. Clearly, there is a large degree of arbitrariness in such psychological time scales.

Returning to finance, the idea is to introduce a time scale similar to psychological time, in which elapsed time is measured by the (subjective) importance of the events. In this sense, each market will create its own time scale.

5. Other Time Scales for Finance

With high frequency data, the time interval between two ticks is a positive random process. The return is another random process, subordinated to the time process. Even when considering daily data, the return can be seen as subordinated to the deterministic time interval. By changing the time scale, the characteristics of the return process are changed, and possibly become simpler. Similarly, other characteristics of the process, for example the volatility, become better behaved in the transformed time scale. This overall idea is quite old, and seems to be due to Bochner back in the '50s (see e.g. (Feller, 1971) and reference therein), or (Mandelbrot and Taylor, 1967) who introduced a number of financial clocks.

The basic idea for every time scale is similar: contract or expand the physical time in such a way that a measure of the market events becomes more regular. Here are some examples of possible financial time scales:

1. *Business time:* counts only when a given market is open, chopping off nights and week-ends.
   This is the most widely used time scale in finance, for example when considering daily averages but omitting Saturdays and Sundays, or when counting 250 days per year.
2. *Transaction time 1:* add one for each transaction.
3. *Transaction time 2:* add the value of each transaction.
4. *Theta-time:* add a measure of the seasonal volatility.
   Contract or expand the physical time in such a way that the average number of market events, as measured by the average seasonal volatility, is constant.
5. *Tau-time:* add some measure of the momentary recent volatility.
   Contract or expand the physical time in such a way that the number of market events, as measured by the volatility, is constant.

The characteristics of these time scales are quite different. Some time scales are predefined (business time, theta time), others depend on the actual events in a given market (transaction time, tau time). This makes a big difference when forecasting, as it is easy to extend a predefined time scale to the future. In contradistinction, the transaction time scale or the tau time scale must be forecasted. Yet the advantage of the transaction time scale is that it could also deal with conditional heteroskedasticity (i.e. volatility clustering).

Another important point is that the relevant information must be available. For example, both transaction time scales are impossible to construct for the FX market because all information about transactions are unavailable.

The next two sections are devoted to the construction of the theta time scale, whose goal is to remove seasonal heteroskedasticity (Dacorogna *et al*., 1993; Dacorogna *et al*., 1996). For this, a suitably robust measure of the seasonal activity must be constructed first.

6. Measuring the Activity

In order to construct a time scale depending on seasonal volatility, we must first take a good estimator of the volatility of a market. Anticipating the next section, where we will build an additive model for the theta time scale, we need an additive quantity, similar to a tick rate. As volatility is not additive, we must define a new quantity called 'activity' which measures the market activity and is additive. A good estimator would be the transaction volume, but these figures are not available for the FX market. Another candidate is the tick frequency, or tick rate. The problem is that this quantity is very much dependent on the data supplier. Say, for example, that a data provider is well represented in Europe, but much less in Asia. Then, the view of the world according to the tick rate of this data supplier is distorted, giving less activity to Asia than it should. Another problem is the technical compromise made by some data suppliers. In order to limit the bandwidth, the tick-by-tick information can be decimated in periods of high activity.

For these reasons, the volatility is a better measure of the particular market's activity, because it gives a much more objective view of the world. But as mentioned earlier, the volatility is not additive. Therefore, we must scale the volatility in order to define an additive activity. The scaling law for the absolute value of the return is given by (Müller *et al.*, 1990)

$$\langle |r[\Delta t]| \rangle = \left(\frac{\Delta t}{\Delta T} \right)^{1/E} , \tag{2}$$

where $\langle\rangle$ denotes the sample average, $r[\Delta t]$ is the return measured for the time interval Δt, ΔT is a constant, and E is the scaling law exponent. This can be rewritten in the form

$$1 = \frac{\Delta T}{\Delta t} \; \Big(\langle |r[\Delta t]| \rangle \Big)^{E} . \tag{3}$$

We *define* the intra-week activity by

$$a(t') = \frac{\Delta T}{\Delta t} \Big(\langle \, |r[\Delta t]| \; | \, t' \rangle \Big)^{E} , \tag{4}$$

where the average is taken conditionally to the time within the week $t' = t \bmod 168$. In the following section, the time interval Δt is fixed to 1 hour.

7. The Theta Time Scale

The ϑ-time scale is defined infinitesimally by reference to the physical time scale. The ϑ-time increment $d\vartheta$ is proportional to the physical time increment, with the activity as a proportionality constant

$$d\vartheta(t') = a(t')\,dt \tag{5}$$

and $t' = t \bmod 168$. In other words, the activity is the Jacobian of the transformation from theta to the physical time scale. Then, the ϑ-time scale is obtained by integration.

$$\vartheta(t) = \Delta\vartheta(t_0, t) \;\equiv\; \int_{t_0}^{t} a(t')\ \mathrm{d}t'. \tag{6}$$

The time t_0 is an arbitrary zero, where the ϑ-time is rooted.

So far, so good. The next important variation of activity is created by holidays, which are predictable, but not seasonal (at least in the above sense). In order to include Holidays into the time scale, we need to construct a model for the activity. The idea is to model the activity as a superposition of three contributions corresponding to the Asiatic, European and American markets. The total average market activity is modeled by 3 geographical components

$$a(t) = a_0 + \sum_{k=1}^{3} a_k(t) \tag{7}$$

where $a_k(t)$ corresponds to the activity of the three generic markets (East Asia, Europe and America), and a_0 is a basic activity. When opened, each market activity is modeled by a simple polynomial

$$a_k(t'') = w(t'' - t_{open})^2(t'' - t_{close})^2(t'' - s)\left[(t'' - t_{lunch})^2 + d^2\right] \tag{8}$$

with $t'' = t \bmod 24 =$ time in the day. When closed, the market activity is zero $a_k = 0$. Then, the model has to be fitted to the activity measured from the data. There is a total of 17 parameters to be fitted, and the reader can consult Dacorogna *et al*. (1993) for more details. Finally, there is an overall proportionality constant that needs to be fixed. This constant fixes the average 'speed' of time in the new scale, and it is chosen so that in the long term, the ϑ-time flows at the same pace as physical time.

Now the holidays can be introduced simply by 'closing' the corresponding market, i.e. by putting its activity to zero. Another advantage of the model is the possibility to take into account daylight saving time. This is done simply by taking the fitted time in the model polynomial to be local time. An example of the resulting mapping from physical to ϑ time is given in Fig. 3. The week chosen for the drawing is a week with no market holidays (25.9 to 1.10.95). That is why the ϑ time ends up a little above 168 hours.

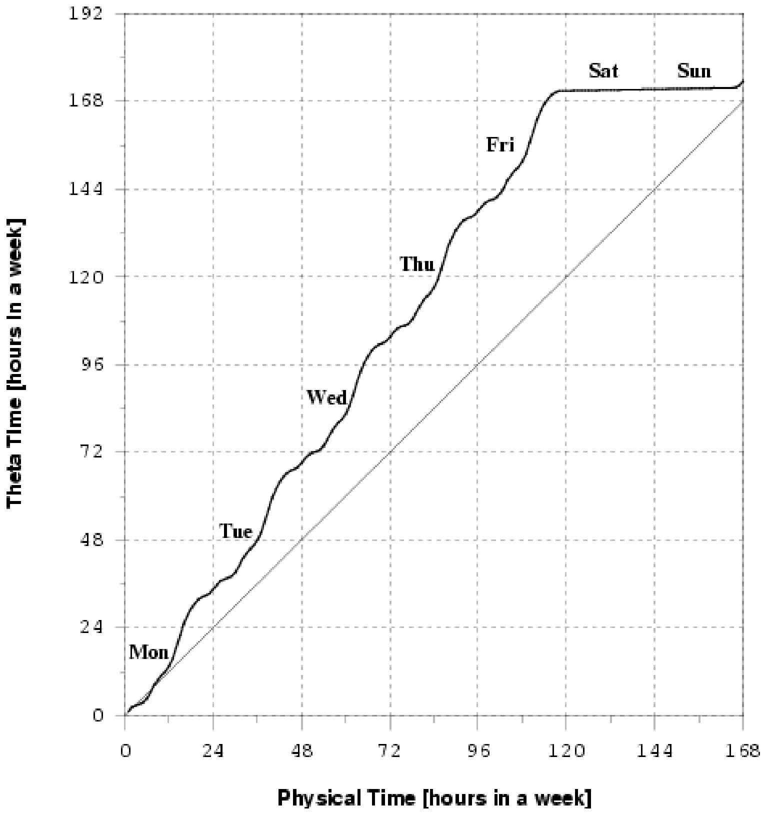

Figure 3. The mapping between physical time and ϑ time for DEM/USD.

One of the motivations for the introduction of a new time scale is that the processes may have simpler characteristics in the transformed time axis. As an heuristic example of what we have accomplished so far, the hourly return is plotted on Fig. 4 against the physical and the ϑ time scales. By an hourly return in ϑ-time, we mean the difference of price, separated by one hour along the scaled time axis. This corresponds to an irregular sampling in physical time, for example the price difference may be taken across the week-end, or separated by only 30 minutes in periods of high activity. Obviously, the process looks simpler in the transformed time axis. In order

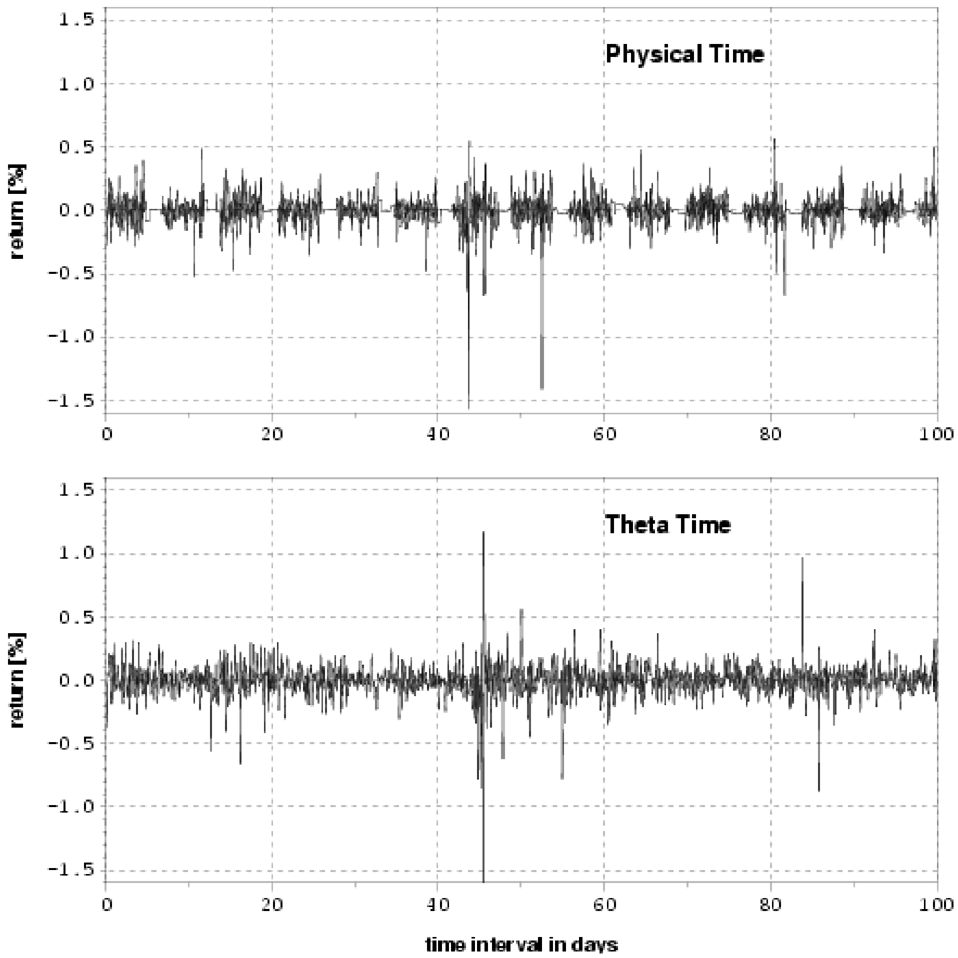

Figure 4. The hourly return in physical and theta time

to be more quantitative, let us compute the lagged autocorrelation of this process. When computed in physical time, this has already been shown in Fig. 1. The same computation but in ϑ-time is displayed in Fig. 5. The structures that were hidden by the strong seasonalities start to unfold. In particular, we see clearly the slow decay of the autocorrelation, corresponding to the clustering of volatility, or the *conditional heteroskedasticity.* The hyperbolic decay of the correlation corresponds to the long memory of the market with respect to volatility. This shows that by working in a time scale in which the process is deseasonalized, we can exhibit some of its finer properties.

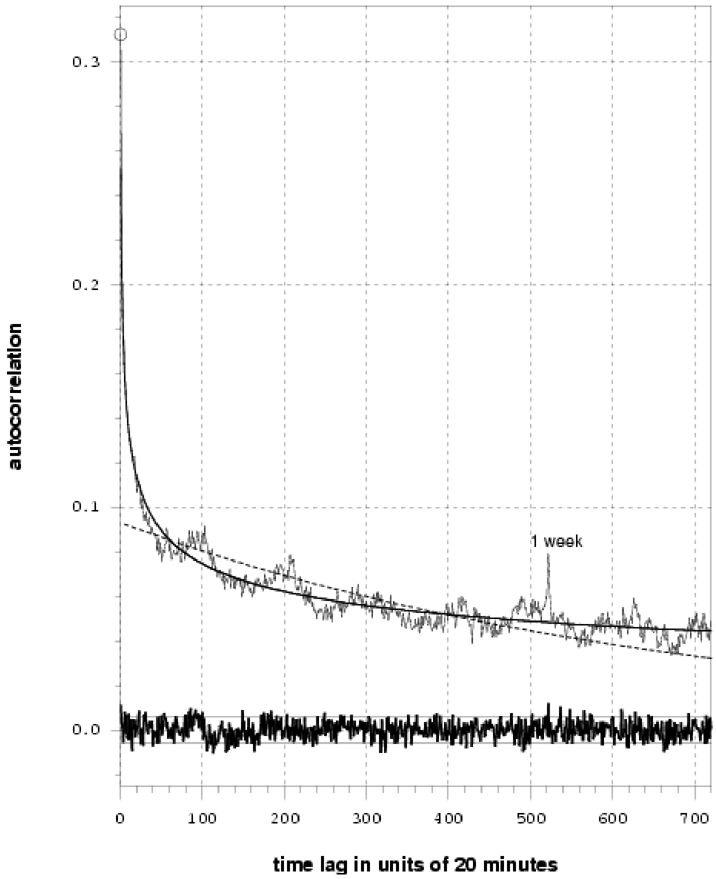

Figure 5. Autocorrelation of the return and its absolute values, in ϑ-time. Autocorrelation of 20 minute return (bold curve) and absolute values of return (thin curve), computed on the ϑ time scale. FX rate: USD-DEM; sampling period: from 5 May 1986 to 4 May 1992. A circle indicates the autocorrelation of absolute returns at lag 1. A hyperbolic function (solid curve) and an exponential function (dotted curve) are shown, both representing the best fit for the absolute price autocorrelation. The confidence limits represent the 95% confidence interval of a Gaussian random walk.

Finally, let us remark that with one time scale, it is possible to deseasonalize at most one quantity. Therefore, even if we had done a perfect job for the activity, the volatility (with your favorite definition of the volatility) will not be completely deseasonalized. Thus, it is not so important to construct a perfect time scale and activity model, but only to take into account the largest contribution to the seasonality.

8. First Hitting Time

Up to this section, our point of view on time was of a continuous parameter indexing the processes. For example, we considered the daily return $r[1 \text{ day}]$, or more generally $r[\Delta t]$ for a given Δt. We may reverse this usual view by considering time as a random variable, subordinated to the price process. In fact, in finance, this is a very natural stand: how long should one wait until one can expect a given return? To be more precise, the first hitting time is defined as the time $\Delta t[r]$ needed to reach — *for the first time* — a given return r. What makes the relationship between $r[\Delta t]$ and $\Delta t[r]$ non trivial is the 'for the first time' in the previous definition.

When studying the return, the probability density $P(r|\Delta t)$ is measured. Similarly, for the first hitting time, $P(\Delta t|r)$ has to be measured, and this is shown in Fig. 6. The time intervals are measured in the theta time scale in order to take care of the seasonalities (a similar picture is obtained when measuring time intervals in physical time). Obviously, the larger the return, the longer we should wait. Note that this graph is on a log-log scale, and if drawn on a linear-linear scale, the curves are completely crushed against the left and lower axes.

In order to compare the curves corresponding to different returns, we have to scale them. Because P is a probability density, we have only one function to choose, namely the scaling relation between Δt and r. Then, the normalization condition on P will enforce the scaling relation for P. The first obvious form is to use an uncorrelated random walk scaling, given by

$$\begin{aligned} \Delta t &\rightarrow z = \Delta t / r^2 \,, \\ P(\Delta t) &\rightarrow r^2 P(z) \,. \end{aligned} \tag{9}$$

The resulting scaled probability densities are shown in Fig. 7, using linear-linear scales. The agreement is fairly good for large values of z but seems to deteriorate for small z. In order to magnify the small z region, the same data are plotted on a log-log scale in Fig. 8. On the same graph, we also plot $P(\Delta t|r)$ for a Gaussian Random Walk (GRW). For this process, $P(z)$ is given by (see e.g. (Borodin and Salminen, 1996))

$$P_{\text{GRW}}(z) = \sqrt{\frac{C}{2\pi}} \, \frac{1}{z^{3/2}} e^{C/(2z)} \tag{10}$$

i.e. this is a Gaussian for the variable $x^2 = 1/z$. The constant C is a free parameter that has to be adjusted, and we draw $P_{\text{GRW}}(z)$ for three possible values of C. Again, the agreement is good for large z. In fact, this agreement is surprisingly good considering that for the first curve ($|r|$ =0.001),

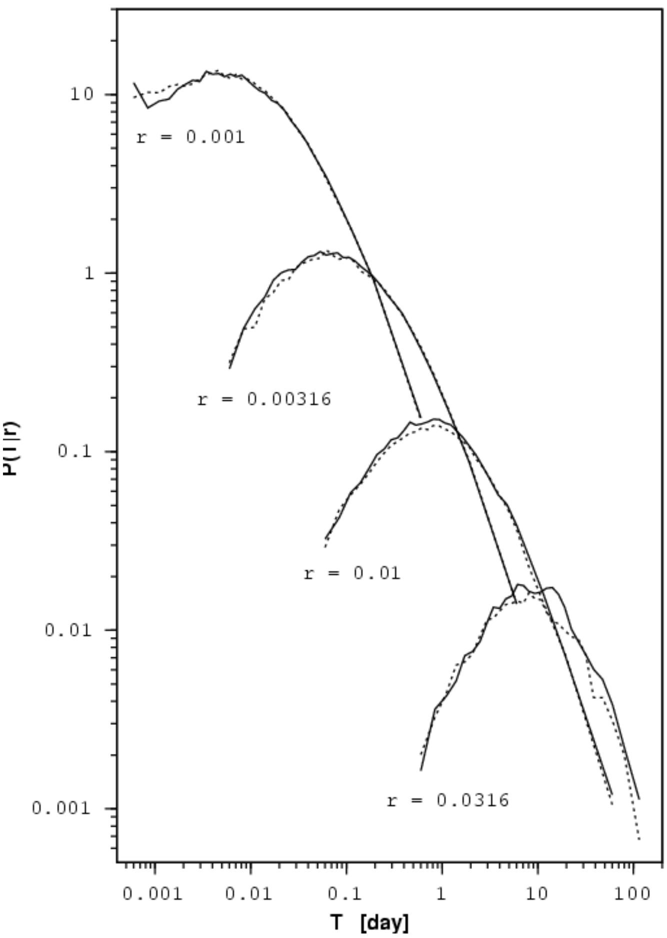

Figure 6. The probability distribution $P(\Delta t|r)$ for the first hitting time, with the time intervals measured in the theta time scale. The returns are $|r|$ =0.001, 0.00316, 0.01, 0.0316. The full (dotted) curves correspond to negative (positive) return. The data are for CHF/USD, from 1.1.87 to 1.1.97.

this corresponds to time intervals going from 1 hour to 1 day. This means that at this time range, for 'normal' events, there are a sufficient number of independent microscopic events to reach the central limit theorem conditions.

Continuing in Fig. 8, the agreement is quite bad for small z. This region corresponds to 'extreme' events, namely the fixed return was reached in a relatively short time interval. For the largest return ($|r|$ =0.0316), this domain corresponds to a time of 1 to 20 days. For such a highly active market as the foreign exchange for CHF/USD, we would expect to have reached the long term scaling limit of the central limit theorem already.

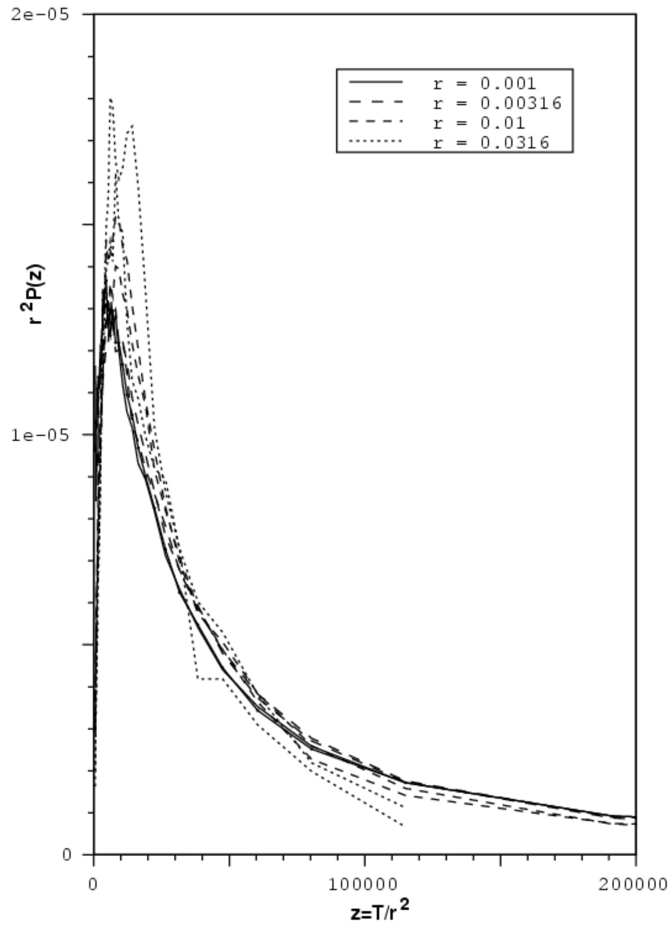

Figure 7. The scaled probability distribution $P(\Delta t|r)$ for the first hitting time, in lin-lin scales.

In the study of the return (given Δt), the large r region is also a domain where scaling seems to be different, yet it is difficult to be more assertive, precisely because this is a domain of rare events and the statistics are quite poor.

The situation may be improved by using another scaling, say $z = \Delta t / r^{\gamma}$ and adjusting γ. For a value of $\gamma = 1.8$ for example, the agreement in the large z region becomes slightly better, but it deteriorates in the small z region. Without an objective a-priori criterion, it is difficult to pick a value for γ. Notice that a lower value of γ would be in line with the scaling of $\langle |r[\Delta t]| \rangle$, which exibits a scaling exponent with respect to Δt larger than $1/2$. Yet, because of the different small z shapes of the scaled curves, the situation can be improved locally at the most, but cannot be globally cured.

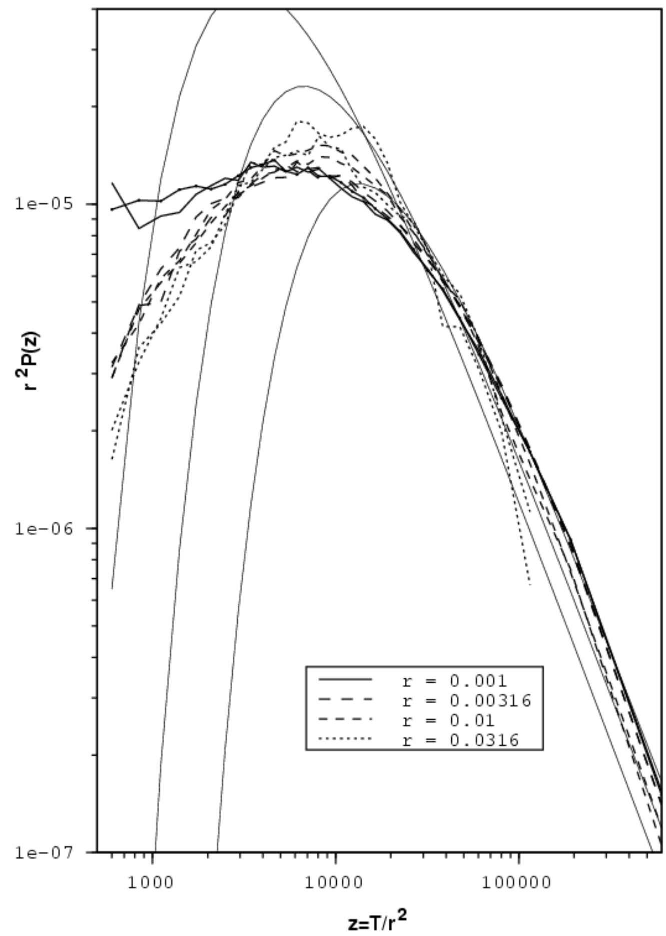

Figure 8. The scaled probability distribution $P(\Delta t|r)$ for the first hitting time, in log-log scales. The parameters for $P_{\mathrm{GRW}}(z)$ are $C = 1\ 10^4, 2\ 10^4, 4\ 10^4$

What makes Fig. 8 particularly interesting is precisely that it clearly shows the domain in which *scaling does not hold*. In short, we do not observe scaling for small z, corresponding to extreme events.

It should also be noted that $P_{\mathrm{GRW}}(z)$ has no first moment due to the slow decay $P(z) \to z^{-3/2}$ for large z, namely the first moment $\langle z \rangle$, or equivalently $\langle \Delta t \rangle$, diverges. Therefore, simple scaling on $\langle \Delta t \rangle$ for example, cannot be computed.

Another interesting point is the comparison with the probability density $P_{\mathrm{GRW}}(\Delta t|r)$ of a Gaussian random walk. For the largest value of $r = 0.0316$, the maximum of the probability corresponds to a time interval of ~ 10 days. Yet, the observed probability distribution is still clearly quite far from the

Gaussian one in the small z region. Let us estimate the size of the domain where we do not observe scaling. The maximum of P_{GRW} is at $z_{\max} = C/3$. The corresponding cummulative distribution function is

$$\mathrm{cdf}_{\mathrm{GRW}}(z) = \int_0^z dz' \, P_{\mathrm{GRW}}(z') = 1 - \Phi(\sqrt{C/z}) \tag{11}$$

which at $z_{\max}$ has a value of cdf $= \sim 8\%$. This gives an estimate of how the small z region corresponds to rare events. In fact, these events are not very rare, they occur once every 10 events; this is better shown in Fig. 7.

9. Conclusion

- In finance, the physical time scale is not necessarily the relevant time scale, because financial markets are not driven by physical laws. Given the freedom to change time scales, we can invent several different ones.
- With high frequency data (intraday), a time scale transformation provides us with an elegant solution to seasonalities. This allows access to finer properties of the process, for example, the conditional heteroskedasticity.
- The first hitting time, conditional to the return, provides another characterization of financial processes. It clearly shows that the processes are not a Gaussian random walk in the short time/large events region. Moreover, we do not observe scaling in this region of large events.

References

Borodin A. N. and Salminen P., 1996, *Handbook of Brownian Motion - Facts and Formulae*, Birkhäuser.

Dacorogna M. M., Gauvreau C. L., Müller U. A., Olsen R. B., and Pictet O. V., 1996, *Changing time scale for short-term forecasting in financial markets*, Journal of Forecasting, **15**(3), 203–227.

Dacorogna M. M., Müller U. A., Nagler R. J., Olsen R. B., and Pictet O. V., 1993, *A geographical model for the daily and weekly seasonal volatility in the FX market*, Journal of International Money and Finance, **12**(4), 413–438.

Feller W., 1971, *An Introduction to Probability Theory and Its Applications*, volume II of *Wiley Series in Probability and Mathematical Statistics*, John Wiley, New York, 2nd edition.

Lerner R. and Trigg G., editors, 1991, *Encyclopedia of Physics*, VCH Publisher, Inc., 2nd edition.

Mandelbrot B. B. and Taylor H. M., 1967, *On the distribution of stock prices differences*, Operations Research, **15**, 1057–1062.

Müller U. A., Dacorogna M. M., Olsen R. B., Pictet O. V., Schwarz M., and Morgenegg C., 1990, *Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis*, Journal of Banking and Finance, **14**, 1189–1208.