Lead-Lag Relationship using a Stop-and-Reverse-MinMax Process

STANISLAUS MAIER-PAAPE
Institut für Mathematik, RWTH Aachen, Templergraben 55, D-52056 Aachen, Germany
maier@instmath.rwth-aachen.de

ANDREAS PLETTEN
Institut für Mathematik, RWTH Aachen, Templergraben 55, D-52056 Aachen, Germany
platen@instmath.rwth-aachen.de

April 24, 2015

Abstract
The intermarket analysis, in particular the lead-lag relationship, plays an important role within financial markets. Therefore a mathematical approach to be able to find interrelations between the price development of two different financial underlyings is developed in this paper. Computing the differences of the relative positions of relevant local extrema of two charts, i.e., the local phase shifts of these underlyings, gives us an empirical distribution on the unit circle. With the aid of directional statistics such angular distributions are studied for many pairs of markets. It is shown that there are several very strongly correlated underlyings in the field of foreign exchange, commodities and indexes. In some cases one of the underlyings is significantly ahead with respect to the relevant local extrema, i.e., there is a phase shift unequal to zero between these two underlyings.

Keywords lead-lag relationship, intermarket analysis, local extrema, empirical distribution

1 Introduction

It is well-known that financial markets can be strongly correlated in such a way that their market values show a similar behavior. Knowing the exact connection between two markets would be very helpful for risk-averse investment strategies. In case that two markets are perfectly correlated it would make no difference to invest in either one of them or both together. One simply cannot diversify the risk on both markets. In case it is known that one market leads the other, one is able to use the leading market as an indicator to predict the price development of the other market. Knowing this connection between the two markets can be useful to improve the investment strategy. Therefore we develop a method for quantizing the interrelation of two markets from a different point of view: We want to be able to identify a possible phase shift between two markets if they are correlated.

This subject has been approached in a variety of articles. One approach is to decompose the time series of two markets on a scale-by-scale basis into components with different frequencies using wavelets. The lead-lag relationship is studied by comparing the components of one selected level of the wavelet transformation for two markets, see e.g. [5, 10, 14, 16, 21, 22]. More on wavelet methods in finance can be found in the book of Gençay, Selçuk and Whitcher [13].
Other methods working with correlation, auto-correlation and similar quantities can be found e.g. in [4, 6, 7, 9, 12, 13, 24]. Didier, Love and Pería [8] studied the comovements during 2007–2008 crisis. A different but related topic is the lead-lag relationship between news, e.g. on Twitter, and stock prices, see e.g. [5, 18].

For the intermarket analysis from a point of view of the technical analysis see e.g. the book of Murphy [20] and also of Ruggiero [23].

However, to the best of the author’s knowledge, the approaches found in the literature do not follow a geometric approach, e.g. they do not take local extreme values of the time series into account. Decomposing the time series using wavelets permits to write the time series as the sum of wavelike components with different frequency spectrum. Using these components for comparison of different markets will therefore compare only parts of the original time series. The problem is that these components can be hidden in the original time series such that a possible lag observed between the components of the same level does not necessarily mean that this lag can be observed in the time series itself, e.g. by comparing reversal points. Therefore it is not clear how to interpret the results with regards to an application.

Since we want to be able to receive results giving us an observable lead-lag relationship of two time series, we prefer a geometrical approach. For this reason we need significant points to be able to uniquely identify a lead or lag if any. Very important situations are reversal points and thus the points in time of relevant local extreme values which represents the moment of reversal. A possible lead or lag can then directly be seen by comparing the local extrema of both charts. Such an ansatz could be used for trading these financial products and offers a deep insight into the lead-lag relationship between two markets because an empirical distribution over all local phase shifts can be identified. Additionally the results are not hidden in just one single value like cross-correlation.

The paper is organized as follows: The search for the relevant local extreme values is far from being unique. Therefore we discuss in Section 2 the approach to find these extreme values for a given pair of markets which we want to compare. Using these values we can compute local phase shifts of both markets which gives us a corresponding empirical distribution. To analyze the results we introduce the directional statistics in Section 3. Now we can apply our approach to historical data, e.g. for foreign exchange, commodities and indexes, which we do in Section 4. In Section 5 we give some conclusions.

2 Method for intermarket analysis

Suppose we want to compare two financial underlyings namely market A and market B for lead and lag. First we take one chart for each underlying with the same bar size, e.g. a 60 min chart, depending on our interest. Now we want to decide whether these two charts are correlated and show lead or lag. Of course if both underlyings are fully uncorrelated we cannot compare them. Therefore let us assume that there is a connection between these two charts.

Since we prefer a geometrical ansatz we need the points in time of relevant local extreme values. If each maximum occurs for both charts at the exact same time and the same holds true for the minimal values we can say that both underlyings run perfectly synchronous. If the maximum of chart B occurs shortly after the maximum of chart A, we observe that market B has a lag compared to market A.
LEAD-LAG RELATIONSHIP USING A SAR-MINMAX PROCESS

Such a comparison could easily be done by hand in a very intuitive way. If one compares two markets and gets a feeling for lead-lag relationship, e.g. assume market $A$ leads $B$, one directly benefits from this knowledge because right after a reversal point in market $A$ would most likely occur a reversal point in market $B$. This can be very useful for several strategies (for position entries and also for exits).

Of course doing an extensive study by hand would be very time consuming and not objective. For an automatic approach we first need an appropriate method to identify local extrema for both time series. The MinMax algorithm introduced by Maier-Paape [17] is a method which yields such a series of alternating relevant local extrema (called MinMax process) and will therefore be used in the following. This method uses a so called SAR (stop and reverse) process to identify up and down movements. If an up movement is detected the MinMax algorithm searches for a maximum and fixes this local maximal value if the movement phase reverses to a down movement. Minimal values are searched during down movement phases. The underlying SAR process could be the MACD (moving average convergence/divergence) indicator of [1] which, simplified speaking, indicates an up movement if the MACD series lies above its signal line and a down movement when its vice versa. See [17] for the details.

The SAR process controls the frequency of detected local extreme values and, in general, is controlled by some parameters (default for MACD are 12, 26 and 9). In this paper we will always use the MACD as SAR process. Instead of adjusting several parameters separately we use just a common factor, called timescale, that scales the three default parameters at the same time. Increasing the timescale leads to less extreme values while decreasing timescale leads to more extreme values, i.e. a finer resolution.

Note that the MACD series can oscillate quickly around the signal line which leads to many small and insignificant local extreme values. To avoid this problem we require for a change of the direction of the SAR process that the distance of MACD and its signal line needs to exceed some minimal threshold of $\delta = 0.3 \cdot \text{ATR}(100)$, where ATR means the average true range, see [17, Subsection 2.1] for the details.

From now on we use this MinMax algorithm because this is a very flexible tool to identify local extreme values. As far as we know this method is the only one which identifies local extreme values exactly and is continuously adjustable. Since a financial time series always has some noise there is no unique objective choice for relevant local extrema of a financial time series. Therefore this process needs to be parameter dependent to adjust the resolution of the minima and maxima.

One question is how to choose the “right” parameter. This will be discussed at the end of this section. For the moment let us assume we already found “good” parameters for market $A$. The MinMax process then yields consecutive minima and maxima denoted by $(t_i, X_i)_{i=1,...,N}$ with points in time $t_1 \leq ... \leq t_N$ and consecutive price values $X_i$. To be able to compare these points, we measure the time in seconds since 1st January 1970. For this wavelike time series we can compute the mean wavelength by

$$\lambda := \frac{1}{N - 1} \sum_{i=1}^{N-1} 2(t_{i+1} - t_i) = 2 \frac{t_N - t_1}{N - 1}. \quad (1)$$

Note that $\lambda$ depends on the parameters used in the MinMax algorithm since the minima and maxima depend on the used parameters.
Figure 1: Moving average of wavelengths over $N - 1 = 49$ cycles for S&P 500 on a 60 min chart.

Fixing these parameters for the second market gives us the extreme values $(\tilde{t}_i, \tilde{X}_i)_{i=1,\ldots,N}$ with mean wavelength $\tilde{\lambda}$. Of course it makes no sense to compare both markets using these extreme values for very different wavelengths $\lambda$ and $\tilde{\lambda}$. Therefore we fit the parameters of the MinMax process for market $B$ so that $\tilde{\lambda} = \lambda$ holds true.

**Remark 1.** Note that in general we can not expect a constant but only a time dependent wavelength which can vary a lot, see Figure 1, where the moving average of wavelengths over $N - 1 = 49$ cycles is shown, i.e. $\frac{1}{99} \sum_{s=1}^{80} (2\tilde{t}_{i+1} - \tilde{t}_i)$ where $s$ is the current time index. Therefore matching the mean wavelength for both markets means just matching the level of refinement and not the position of the extreme values themselves.

Since we are interested in the lead-lag relationship between market $A$ and $B$ we only need to find the relationship of points in time of the extrema by finding the relative positions of $(\tilde{t}_j)_{j=1,\ldots,N}$ within $(t_j)_{j=1,\ldots,N}$. In this case we call market $A$ the primary market and market $B$ the secondary market. The overall procedure is as follows:

1. Fix the desired mean wavelength $\lambda^* > 0$.

2. Find all local extreme values $(t_i, X_i)_{i=1,\ldots,N}$ and $(\tilde{t}_j, \tilde{X}_j)_{j=1,\ldots,N}$ for the primary and the secondary market, respectively, such that the mean wavelengths $\bar{\lambda}$ for both markets on the full data base matches $\lambda^*$, i.e. that we have $\frac{2\tilde{t}_j - \tilde{t}_i}{N-1} \approx \lambda^* \approx \frac{2t_N - t_1}{N-1}$.

3. Find $j_1, j_2 \in \{1, \ldots, N\}$ such that $\tilde{t}_{j_1} = \min\{\tilde{t}_j : \tilde{t}_j \geq t_1\}$ and $\tilde{t}_{j_2} = \max\{\tilde{t}_j : \tilde{t}_j < t_N\}$. For each $j \in \{j_1, \ldots, j_2\}$ do the following:
   a) Find $i \in \{1, \ldots, N - 1\}$ such that $t_i \leq \tilde{t}_j < t_{i+1}$.
   b) Define the phase shift of extreme value $(\tilde{t}_j, \tilde{X}_j)$ regarding the extreme values $(t_i, X_i)$ and $(t_{i+1}, X_{i+1})$. Here we use the linear relative distance between the corresponding extrema values measured as an angle. We set

   \[
   \alpha_j^\lambda := \frac{\tilde{t}_j - s_j}{t_{i+1} - t_i} \cdot \pi \in [-\pi, \pi), \tag{2}
   \]

   where

   \[
   s_j := \begin{cases} t_i, & \text{if } X_i \text{ and } \tilde{X}_j \text{ are both maxima or both minima,} \\ t_{i+1}, & \text{if } X_{i+1} \text{ and } \tilde{X}_j \text{ are both maxima or both minima.} \end{cases} \tag{3}
   \]
LEAD-LAG RELATIONSHIP USING A SAR-MINMAX PROCESS

Figure 2: Computation of $\alpha$ in (2).

Figure 2 shows some examples for the position of a maximum of the secondary market relative to some extreme values of the primary market.

4. We end up with the empirical circular distribution $(\alpha^*_j)_{j=j_1,\ldots,j_2} \subset (-\pi, \pi)$ depending on the mean wavelength $\lambda^*$.

Negative $\alpha$ resemble a front-running (lead) of the secondary market, positive $\alpha$ resemble a time lag of the secondary market. The result can be interpreted on the unit sphere $S^1 = \{(\sin \alpha, \cos \alpha) \in \mathbb{R}^2 : \alpha \in [-\pi, \pi]\}$ and gives us all observations of local phase shifts between two markets.

Remark 2. This approach is independent of the openings of the stock exchange for market A and market B. Since we measure the points in time $t_i$ and $\tilde{t}_j$ in seconds since 1st January 1970 we just put these values into (2) and the machinery works straight forward.

Remark 3. The above method has only one parameter, namely the mean wavelength $\lambda^*$, see step 1. Therefore we can compute different distributions for different wavelengths. It turns out that the results in most cases do not depend on the wavelength. Therefore we compute $(\alpha^*_i)_{i=1,\ldots,n(\lambda)}$ for many values of the mean wavelength $\lambda$. For each $\lambda$ we can generate a histogram or rather a bar plot and at the end we can compute the average of all bars including standard deviation.

Remark 4. Note that the extreme values cannot be determined in real time. There is always at least a small time lag. Therefore we can also identify such an empirical distribution if we use the point in time when the extreme value is confirmed by the MinMax algorithm instead of the point in time of the extreme value itself.

3 Directional statistics

Since we work with circular distributions, the mean and variance must be computed in an appropriate way, see e.g. [11, 19]. This can be used to identify a possible phase shift. We
introduce the basic statistical quantities in Subsection 3.1. For a deeper analysis we list some interesting statistical tests in Subsection 3.2 and give an approximation of the lead or lag in Subsection 3.3.

3.1 Basic quantities

Now we will discuss how to calculate estimators, e.g. for the mean angular direction. Details on computations for a general distribution with a $2\pi$ periodic probability density function $f$ can be found in [11, Section 3.2].

The first step is to identify the angles by vectors on the unit sphere $S^1$. Let $(\alpha_j)_{j=1,...,N} \subset [-\pi,\pi)$ be the outcomes of a discrete distribution for the phase shift of two markets of interest. We can identify each angle $\alpha_j$ with a point on the unit sphere

$$r_j := \begin{pmatrix} \sin \alpha_j \\ \cos \alpha_j \end{pmatrix} \in S^1$$

for $j = 1, ..., N$. In this two-dimensional space we can compute the mean resultant vector which is defined by

$$\hat{r} := \frac{1}{N} \sum_{j=1}^{N} r_j.$$

Note that for the length of $\hat{r}$ we have $\|\hat{r}\|_2 \leq 1$ because it is a convex combination of vectors in $S^1$. If $\hat{r} \neq 0$ choose the mean angular direction $\hat{\alpha} \in [-\pi,\pi)$ such that

$$\begin{pmatrix} \sin \hat{\alpha} \\ \cos \hat{\alpha} \end{pmatrix} = \frac{1}{\|\hat{r}\|_2} \hat{r}. \tag{4}$$

Of course $\hat{r}$ could be zero and thus no unique mean angular direction would exist. This is the case, e.g., if the angles are uniformly distributed all around $S^1$. If $\hat{r} \neq 0$ choose the mean angular direction $\hat{\alpha} \in [-\pi,\pi)$ such that

$$\begin{pmatrix} \sin \hat{\alpha} \\ \cos \hat{\alpha} \end{pmatrix} = \frac{1}{\|\hat{r}\|_2} \hat{r}. \tag{4}$$

Of course $\hat{r}$ could be zero and thus no unique mean angular direction would exist. This is the case, e.g., if the angles are uniformly distributed all around $S^1$. If this is the case for the phase shifts between two markets then there is no connection between them and the analysis of the results would already be finished. Since we are interested in at least slightly correlated markets we do not expect this behavior.

Nevertheless even in the case where $\|\hat{r}\|_2 > 0$, the length of $\hat{r}$ could be small. This happens if the outcomes of the distribution have a large variance. In contrast a length of $\hat{r}$ near 1 indicates a small variance and a high concentration of the outcomes near to its mean angular direction. Therefore we need to consider the circular variance (cf. [11, Section 2.3.1, Equation (2.11)]) which can be defined by

$$\hat{S} := 1 - \|\hat{r}\|_2 \in [0,1].$$

To be able to also measure the skewness and the peakedness we define the circular skewness by

$$\hat{b} := \frac{1}{N} \sum_{j=1}^{N} \sin(2(\alpha_j - \hat{\alpha})) \in [-1,1]$$
LEAD-LAG RELATIONSHIP USING A SAR-MINMAX PROCESS

0.1
0.2
0.3
0
−π
2
±π
+ π
2
0.1
0.2
0.3
0
−π
2
±π
+ π
2
0.1
0.2
0.3
0
−π
2
±π
+ π
2

Figure 3: First: Example of a possible distribution of phase shifts and a hat function on $S^1$; Second: Corresponding weighted version of the example distribution from first plot; Third: Plot of the probability density functions of von Mises distributions mean location parameter $\mu = 0$ and concentration parameters $\kappa = 50$; Fourth: Same as third but with $\kappa = 1$.

and the circular kurtosis by

$$\hat{k} := \frac{1}{N} \sum_{j=1}^{N} \cos(2(\alpha_j - \hat{\alpha})) \in [-1, 1].$$

Since we are interested in the possible lead or lag between two markets we want to reduce the influence of outliers which are far away from the mean angular direction. For this reason we use a hat function on $S^1$ to weight the empirical distribution with the hat near the position of the highest peak of the distribution. Then all reasonable data near the peak get high weights and thus more influence in our statistics, while less important data, i.e. the outliers, obtain small weights. We expect that the peaks of the distributions are near zero up to some lead or lag, i.e. the two markets are positive correlated. Therefore we use the hat function which has its hat (maximum) at zero and is zero (minimum) at $\pm \pi$. The first two plots of Figure 3 show an example for an observed distribution and its weighted counterpart, respectively. From the weighted distribution we can compute the weighted mean angular direction $\hat{\alpha}_{(w)}$ as in [1].

3.2 Statistical tests

Most of the statistical tests require an underlying von Mises distribution, see e.g. [11, Section 3.3.6], which is often used as an analogon to normal distribution on the unit sphere. The distribution we get for our application is not exactly a von Mises distribution but has a similar shape, see Figure 3. In this figure the distribution of phase shifts has a similar shape to two superposed von Mises distribution, one with a large and one with a small concentration parameter $\kappa$. Thus it is possible, that the phase shifts correspond to a von Mises distribution plus noise, e.g. white noise. Nevertheless we use the following statistical
tests in order to be able to classify the results even if they are designed for von Mises distributions.

Since we do not know the underlying distribution for the phase shifts we only get some realizations. Computing the quantities in Section 3.1 using the formulas by putting in our observations will give us the estimators which will be denoted by \( \hat{\alpha} \), \( \hat{\alpha}^{(w)} \), \( \hat{S} \), \( \hat{b} \) and \( \hat{k} \), respectively.

Next we want to verify the quality of our mean angular direction. Therefore we compute the \((1 - \delta)\%\)-confidence intervals for the population mean, such that \( L_1 := \hat{\alpha} - d \) and \( L_2 := \hat{\alpha} + d \) are the lower and upper confidence limits of the mean angular direction, respectively, see [25, Section 26.7]. For the weighted mean \( \hat{\alpha}^{(w)} \) we denote the confidence interval by \( d^{(w)} \). We always use \( \delta = 5\% \).

To test for zero mean which would imply that there is no lead or lag relationship we can perform the one sample test for mean angle, which is similar to the one sample \( t \)-test on a linear scale. Let \( \alpha_0 \in [-\pi, \pi) \) be the mean angular direction for which we want to test and \( \bar{\alpha} \) the mean angular direction of the underlying (unknown) distribution. We test for

\[
H_0 : \bar{\alpha} = \alpha_0, \\
H_1 : \bar{\alpha} \neq \alpha_0
\]

by checking whether \( \alpha_0 \in [L_1, L_2] \) using our estimator \( \hat{\alpha} \) and its 95\% confidence interval, see [25, Section 27.1 (c)]. In our case we will set \( \alpha_0 = 0 \). The result of this test is then given by

\[
h_m := \begin{cases} 0, & \text{if } H_0 \text{ can not be rejected, i.e. } \alpha_0 = 0 \in [L_1, L_2], \\ 1, & \text{otherwise.} \end{cases}
\]

In Remark 3 we noted that we will generate empirical distributions for different mean wavelengths, say \( n \in \mathbb{N} \) different values. To compare all these distributions for the same pair of markets we can use the one-factor ANOVA or Watson-Williams test (multi-sample test). It assesses the question whether the mean directions of two or more groups are identical or not, i.e. it tests for

\[
H_0 : \text{All of } n \text{ groups share a common mean direction, i.e., } \bar{\alpha}^{(1)} = \ldots = \bar{\alpha}^{(n)}. \\
H_1 : \text{Not all groups have a common mean direction},\]  

see [25, Section 27.4 (b)]. The output of this test is a \( p \)-value, i.e. the probability of getting results which are at least as extreme as our observation assuming the null hypothesis is true. Thus a large \( p \)-value indicates that the null hypothesis holds true. We denote this value by \( p_{ww} \in [0, 1] \).

### 3.3 Lead or lag

Using the mean angular direction \( \hat{\alpha} \) and its confidence interval we can roughly approximate the lead or lag. Assume we have a mean wavelength of 100 candles on a 60 min chart. The mean wavelength would then be approximately \( 60 \text{ min} \cdot 100 = 6000 \text{ min} \). This value equates \( 2\pi \). Thus the mean of the lead or lag \( \ell \) can be approximated by

\[
\ell \approx \frac{\hat{\alpha}}{2\pi} \cdot 6000 \text{ min}
\]
and the corresponding confidence interval is approximated by \([\ell - d, \ell + d]\) where

\[
d \approx \frac{d}{2\pi} \cdot 6000 \text{ min}.
\]

Analogously we can compute the lead or lag using the weighted mean angular direction which we denote by \(\hat{\ell}(w)\) and \(d^w\ell\), respectively, i.e. \(\hat{\ell}(w) \approx \frac{\hat{\alpha}(w)}{2\pi} \cdot 6000 \text{ min}\) and \(d^w\ell \approx \frac{d^w}{2\pi} \cdot 6000 \text{ min}\). Note that a positive value for \(\ell\) and \(\hat{\ell}(w)\) means that the primary market leads the secondary and vice versa for a negative value.

To answer the question which market is ahead, if any, we make the following definition.

**Definition 1.** For positive correlated markets, i.e. \(|\hat{\alpha}(w)| \leq \frac{\pi}{2}\), we say one market leads the other if \(\hat{\alpha}(w)\) is significantly different from zero, i.e.

\[
\text{if } \hat{\alpha}(w) - d(w) > 0 \quad \Rightarrow \quad \text{primary market leads},
\]

\[
\text{if } \hat{\alpha}(w) + d(w) < 0 \quad \Rightarrow \quad \text{secondary market leads}.
\]

### 4 Empirical study

Now we study different markets from commodities to foreign exchanges. In Subsection 4.1 we explain the setting and give some details on the choice of parameters. The angular histograms and the statistical results are then shown in Subsection 4.2.

#### 4.1 Settings

In this paper we focus on the 60 min chart. The wavelengths we use to adjust the MinMax process for the primary market, see Remark 3, are of size of 30 candles up to 180 candles. For the Watson-Williams test, see Section 3.2, this leads to \(n = 151\) groups.

Since we have given the wavelength in number of candles we proceed as follows to “synchronize” the markets:

1. Choose the desired mean wavelength \(\lambda^*\) in number of candles.
2. Adjust the parameter for the MinMax process on the primary market, such that the wavelength of the primary market in number of candles, ignoring the time when the stock exchange is closed, matches \(\lambda_{\text{candles}}^*\).
3. Calculate the corresponding wavelength \(\lambda^*\) in seconds for the primary market, this time considering the time when stock exchange is closed.
4. Adjust the MinMax process on the secondary market, such that the wavelength of the secondary market in seconds matches \(\lambda^*\), i.e. perform step 2 from Section 2 where the primary market is already fixed.
5. Proceed with steps 3 and 4 from Section 2.

For most computations of the directional statistics the MATLAB library CircStat [2] has been used and all angles are measured in radian.

The markets which we examine including the period of time for the available candle data are listed in Table 1. Note that the start date is not the same for all markets. If we examine a combination of markets with different initial dates we use the smaller period of time for both markets.
Table 1: Examined markets and the period of time of the used candle data of the 60 min chart (terminal date is always May 15, 2014). All historical data are from FIDES.

| market          | underlying                              | initial date   |
|-----------------|-----------------------------------------|----------------|
| Eurex DAX       | DAX Futures                             | December 10, 2003 |
| Eurex BUND      | Euro-Bund Futures                       | December 10, 2003 |
| Eurex DJIST50    | Euro STOXX 50 Index Futures             | December 10, 2003 |
| CME MINI S&P    | E-mini S&P 500 Futures                  | December 12, 2003 |
| CME MINI NSDQ   | E-mini NASDAQ 100 Futures              | September 14, 2004 |
| CME CMX GLD     | Gold Futures                            | July 7, 2005    |
| CME CMX SIL     | Silver Futures                          | July 7, 2005    |
| CME PH CRDE     | Crude Oil Futures                       | November 29, 2004 |
| CME PH NG       | Natural Gas (Henry Hub) Physical Futures| October 18, 2004 |
| CME NYBOT MNRUS2K| Russell 2000 Index Futures              | September 20, 2004 |
| Forex EUR-USD   | EUR-USD                                 | July 17, 2009   |
| Forex JPY-USD   | JPY-USD                                 | July 17, 2009   |
| Forex GBP-USD   | GBP-USD                                 | July 17, 2009   |
| Forex CHF-USD   | CHF-USD                                 | July 17, 2009   |

4.2 Results

Now we look at the results for several futures, indexes and foreign exchanges. The statistical quantities for the phase shift of the extreme values are shown in Table 2 and for the points in time of the confirmation of the extreme values in Table 3. The corresponding empirical distributions are given, according to the following remark, by Figures 4 to 20.

Remark 5. (Notes on figures)
The label of each of the following figures states “A versus B” and each figure shows the following four distributions (in same order):

1. Time of extrema: A as primary and B as secondary market.
2. Time of extrema: B as primary and A as secondary market.
3. Time of extrema confirmed (see Remark 4): A as primary and B as secondary market.
4. Time of extrema confirmed (see Remark 4): B as primary and A as secondary market.

All plots also contain the mean angular direction and the mean angular direction of the weighted distribution (weighted with the hat function, see Figure 3). These directions are the green and red line inside the circle, respectively.

Additionally each bin of the histograms contains information of the single distributions for each wavelength: It shows the largest value of this bin occurred within the 151 single distributions, the smallest value and the bin value of the combined distribution plus and minus the standard deviation.

First we discuss the results for the time of extrema and afterwards the results for the confirmation time of the extrema.
Table 2: Results on 60 min chart (time of extrema). This market leads the other one.
Table 3: Results on 60 min chart (fund of external comment). This market leads the other one.

| Date | Time | Open | High | Low | Close | Volatility | Correlation |
|------|------|------|------|-----|-------|------------|--------------|
| 1    | 696  | 0.68 | 0.73 | 0.70| 0.71  | ±0.027     | 0.944        |
| 1    | 000  | 0.61 | 0.60 | 0.58| 0.60  | ±0.022     | 0.944        |
| 1    | 120  | 0.68 | 0.67 | 0.66| 0.67  | ±0.021     | 0.944        |
| 1    | 240  | 0.66 | 0.67 | 0.66| 0.67  | ±0.028     | 0.944        |
| 1    | 360  | 0.67 | 0.69 | 0.67| 0.68  | ±0.023     | 0.944        |
| 1    | 480  | 0.68 | 0.69 | 0.68| 0.69  | ±0.024     | 0.944        |

Note: The table shows the volatility and correlation for the given market at different time intervals. The volatility values are expressed in percentage points, while the correlation values are on a scale of 0 to 1, with 1 indicating perfect positive correlation.
Time of extrema  First we note that the results are mostly independent of the mean wavelength which we can see from the additional information of each bin, i.e. the minimal and maximal value for this bin and the standard deviation. Next we see a very weak correlation between EUR-USD vs. JPY-USD, Gold vs. EUR-USD, Gold vs. S&P 500, Gold vs. FDAX, Gold vs. Oil, Oil vs. FDAX and Oil vs. EUR-USD. The pairs of markets also have a relatively large standard deviation $\hat{S}$ and small concentration around its mean indicated by the small kurtosis $\hat{k}$.

All other combinations of markets illustrated in Table 2 and Figures 4 and 6 to 13 show a large peak near the mean angular direction between 20% up to 53%. This means that the probability is significantly high that extreme values for both markets are shaped in almost the exact time. Of course this leads to smaller standard deviations and higher kurtosis.

Confirmation time of extrema  Since the point in time of confirming an extreme value by the MinMax process is more sensitive to the price development than the very fixed point in time of the extreme value itself we already expect scattered observations. However even here we can see a peak in the mean angular direction of about half of the size of the peak for the time of extrema of the strongly correlated pairs of markets. The values in Table 3 are approximately of the same order as in Table 2.

All together  We see strong correlations for extrema and confirmed extrema between combinations of FDAX, Euro-Bund, Euro STOXX, S&P 500, U.S. Treasury, NASDAQ 100, Russel 2000 and between the foreign exchanges except EUR-USD versus JPY-USD. Additionally Gold and Silver has a strong correlation whereas all other combinations with at least one market from commodities seem to be weakly correlated or even nearly uncorrelated. Thus from the point of view of local extreme values the commodities are separated from other markets.

The lead-lag $\ell(w)$, see Section 3.3, is between 5 min and 10 min for the point in time of the extrema for the indexes and foreign exchanges and also for Gold versus Silver. Note that this is just a fraction of the duration of one single period of the 60 min chart. Even the points in time of the extrema are just the time stamp of a candle and not the exact time of the extreme value itself, i.e. these points in time have an uncertainty of ±30 min. Therefore we cannot view the value $\ell(w)$ as an absolute value but more as a tendency of the lead or lag for the candles in which the extreme values occur.

Remark 6. In most of the cases our investigations of the correlation of two markets yields one market leading and one market following, e.g. DAX Futures leads E-mini S&P 500 Futures, no matter which one is considered primary or secondary market. Note however, that in some cases the leading market is not unique like for instance the Gold Futures versus Silver Futures, or, sometimes our calculation cannot decide which market is leading.

Remark 7. For the currency Swiss franc it is more common to analyze USD-CHF instead of CHF-USD as we do in the above discussion. The reason we focus on CHF-USD is to see the positive correlation to EUR-USD and thus to have a more natural interpretation for lead and lag as in Definition 1.

However, it is also possible to compare (strongly) negative correlated markets as EUR-USD versus USD-CHF. In Figure 21 we see the results for this combination. We expect that
Figure 4: DAX Futures versus E-mini S&P 500 Futures

Figure 5: EUR-USD versus JPY-USD

Figure 6: EUR-USD versus GBP-USD
Figure 7: EUR-USD versus CHF-USD

Figure 8: Euro-Bund Futures versus U.S. Treasury Bond Futures
Figure 9: E-mini S&P 500 Futures versus E-mini NASDAQ 100 Futures

Figure 10: Russell 2000 Index Mini Futures versus E-mini S&P 500 Futures
LEAD-LAG RELATIONSHIP USING A SAR-MINMAX PROCESS

Figure 11: Russell 2000 Index Mini Futures versus E-mini NASDAQ 100 Futures

Figure 12: EURO STOXX 50 Index Futures versus DAX Futures
Figure 13: Gold Futures versus Silver Futures

Figure 14: Gold Futures versus EUR-USD

Figure 15: Gold Futures versus E-mini S&P 500 Futures
Figure 16: Gold Futures versus DAX Futures

Figure 17: Gold Futures versus Crude Oil Futures

Figure 18: Crude Oil Futures versus DAX Futures

Figure 19: Crude Oil Futures versus EUR-USD
the results are the same as for the combination EUR-USD versus CHF-USD but shifted by $\pi$. If we compare Figures 7 and 21 we actually see this connection perfectly. This is also the case for the Japanese yen.

5 Conclusion and outlook

We introduced the notion of lead-lag relationship from a market technical point of view. Using the local extreme values of the markets we get an empirical distribution of their phase shifts on the unit sphere. The directional statistics helps us to illustrate and quantify the results.

We observed many strongly correlated pairs of markets with respect to their extreme values while, of course, there are combinations with a very weak connection. Combinations of indexes show the highest correlation and also a measurable lead or lag. Since we use a geometrical approach based on the actual local extreme values of the chart, i.e. on some kind of reversal points, the results can directly be used for trading strategies.

In future work the authors plan to localize this method to shorter time intervals so that we obtain even more meaningful results for live/real time data.
References

[1] Appel, Gerald: Technical Analysis: Power Tools for Active Investors. Financial Times Prentice Hall, Upper Saddle River, NJ, 2005.

[2] Berens, P.: CircStat: A MATLAB Toolbox for Circular Statistics. Journal of Statistical Software, 31(10), 2009.

[3] Bollen, J., H. Mao and X. Zeng: Twitter mood predicts the stock market. Journal of Computational Science, 2(1):1–8, 2011. DOI: 10.1016/j.jocs.2010.12.007

[4] Chan, K.: Imperfect Information and Cross-Autocorrelation Among Stock Prices. The Journal of Finance, 48(4):1211–1230, 1993. DOI: 10.1111/j.1540-6261.1993.tb04752.x

[5] Dajcman, S.: Interdependence Between Some Major European Stock Markets – A Wavelet Lead/Lag Analysis. Prague Economic Papers, 22(1):28–49, 2013.

[6] de Jong, F. and M. W. M. Donders: Intraday Lead-Lag Relationships Between the Futures-, Options and Stock Market. Review of Finance, 1(3):337–359, 1998. DOI: 10.1023/A:1009765322522

[7] de Jong, F. and T. Nijman: High frequency analysis of lead-lag relationships between financial markets. Journal of Empirical Finance, 4(2–3):259–277, 1997. DOI: 10.1016/S0927-5398(97)00009-1

[8] Didier, T., I. Love and M. D. M. Pería: What explains comovement in stock market returns during the 2007–2008 crisis? International Journal of Finance & Economics, 17(2):182–202, 2012. DOI: 10.1002/ijfe.442

[9] Égert, B. and E. Kočenda: Time-varying synchronization of European stock markets. Empirical Economics, 40(2):393–407, 2011. DOI: 10.1007/s00181-010-0341-3

[10] Fiedor, P.: Information-theoretic approach to lead-lag effect on financial markets. The European Physical Journal B, 87:168, 2014. DOI: 10.1140/epjb/e2014-50108-3

[11] Fisher, N. I.: Statistical Analysis of Circular Data. Cambridge University Press, 1996.

[12] Garcia, R. and G. Tsafack: Dependence structure and extreme comovements in international equity and bond markets. Journal of Banking & Finance, 35(8):1954–1970, 2011. DOI: 10.1016/j.jbankfin.2011.01.003

[13] Gençay, R., F. Selçuk and B. Whitcher: An Introduction to Wavelets and Other Filtering Methods in Finance and Economics. Academic Press, New York, 2001.

[14] In, F. and S. Kim: The Hedge Ratio and the Empirical Relationship between the Stock and Futures Markets: A New Approach Using Wavelet Analysis. The Journal of Business, 79(2):799–820, 2006. DOI: 10.1086/499138

[15] Iwaisako, T.: Stock Index Autocorrelation and Cross-autocorrelations of Size-sorted Portfolios in the Japanese Market. Hitotsubashi Journal of Economics, 48(1):95–112, 2007.
[16] Kim, S. and F. In: The relationship between stock returns and inflation: new evidence from wavelet analysis. Journal of Empirical Finance, 12(3):435–444, 2005. DOI: 10.1016/j.jempfin.2004.04.008

[17] Maier-Paape, S.: Automatic One Two Three. Quantitative Finance, 15(2):247–260, 2015. DOI: 10.1080/14697688.2013.814922.

[18] Mao, H., S. Counts and J. Bollen: Predicting Financial Markets: Comparing Survey, News, Twitter and Search Engine Data. http://arxiv.org/abs/1112.1051, 2011.

[19] Mardia, K. V. and P. E. Jupp: Directional Statistics. Wiley series in probability and statistics. Wiley, Hoboken, NJ, 1999.

[20] Murphy, J. J.: Intermarket Analysis: Profiting from Global Market Relationships. John Wiley & Sons, Hoboken, 2004.

[21] Ramsey, J. B. and C. Lampart: The Decomposition of Economic Relationships by Time Scale Using Wavelets: Expenditure and Income. Studies in Nonlinear Dynamics & Econometrics, 3(1):23–42, 1998. DOI: 10.2202/1558-3708.1039.

[22] Ramsey, J. B. and C. Lampart: Decomposition of Economic Relationships by Timescale using Wavelets. Macroeconomic Dynamics, 2(1):49–71, 1998.

[23] Ruggiero, M. A.: Cybernetic Trading Strategies. John Wiley & Sons, Hoboken, 1997.

[24] Stoll, H. R. and R. E. Whaley: The Dynamics of Stock Index and Stock Index Futures Returns. The Journal of Financial and Quantitative Analysis, 25(4):441–468, 1990. DOI: 10.2307/2331010.

[25] Zar, J. H.: Biostatistical Analysis. Pearson, Upper Saddle River, NJ, 5. edition, 2010.