On Autocorrelation Based on Hermann Weyl's Discrepancy Norm for Time Series Analysis

Jean-Luc Bouchot, Johannes Himmelbauer and Bernhard Moser

***Abstract*— Hermann Weyl's concept of a discrepancy measure is discussed in the context of time series analysis. A concept for autocorrelation based on this discrepancy notion is introduced. It is shown that in particular for high frequent signals as they, for example, are typically encountered in a financial context, the introduced autocorrelation concept stands out by a better discriminative power than its classical counterpart. While the computational complexity of this novel autocorrelation is of quadratic order in terms of the number of given time steps an approximation based on L_p-norms is introduced which can be computed by convolution, and therefore reduces the order of complexity to that of its classical counterpart. It is shown that the proposed approximation can be tuned to be arbitrarily close to the original discrepancy based version, and that it shows similar desirable behavior.**

I. Introduction

Autocorrelation is a widely used concept for analyzing time series. In this paper we address properties of the standard autocorrelation which are intrinsically tied to the point-wisely construction of the standard correlation measure. It will be argued that due to this construction principle especially for high frequent signals the standard autocorrelation fails as even small shifts in the data might cause a break down of the correlation measure. This behaviour becomes the more apparent the more the signal shows chaotic characteristics with high frequencies. To overcome this problem we introduce an alteranative concept based on a construction principle that relies on the evaluation of partial sums rather than the aggregation of point-wisely computed similarity measures.

First of all let us start with recalling the standard concepts. Later on as an alternative concept Hermann Weyl's discrepancy measure is introduced in this context.

In signal processing, the cross-correlation is similar in principle to the convolution of two functions and is defined by $(f \odot g)(t) = \int f(\tau)g(\tau + t)d\tau$. Therefore, cross-correlation of two signals can be understood as similarity between these signals which is constructed by a sliding dot product. It is commonly used to identify a shorter, known feature in a long duration signal. Applications can be found in various fields of science and engineering like pattern recognition, single particle analysis, electron tomographic averaging, cryptanalysis, neurophysiology and financial information systems. The cross-correlation of a signal with itself is also referred to as autocorrelation, which at least has a peak at a lag of zero. In statistics similar concepts are used to measure the correlation between random variables or sequences of random variables. For example Pearson's correlation coefficient is given by

$$\mathrm{corr}(f,g) = \frac{\sum_i (f_i - \bar{f})(g_i - \bar{g})}{\sqrt{\sum_i f_i + \bar{f}}\sqrt{\sum_i g_i - \bar{g}}},$$

where $\bar{f}$ denotes the arithmetic mean and is useful for measuring linear or affine correlation. A value close to zero indicates there is little linear correlation between the variables, but does not rule out significant nonlinear correlation. Rank correlation coefficients like those proposed by Spearman or Kendall are related to a monotonic relationship between the variables. For example Spearman's coefficient tells how well the relationship between two variables can be described using a monotonic function. If there are no repeated data values, a Spearman correlation of $+1$ or -1 indicates that each of the variables can be obtained by a monotone function of the other, see [4], [6], [15] In [10] it was pointed out that the classical cross-correlation concept suffers from some shortcomings due to its pointwise construction, i.e. summation over pointwise products. The construction principle which is characterized by aggregating pointwisely determined similarities or, analogously, distances applies to a wide class of similarity measures. For instance f-divergence and f-information measures based on information theoretical concepts that rely on the evaluation of histograms [1], [13], [16] are also based on this principle. Examples of such measures include mutual information, Kullback-Leibler distance and the Jensen-Rényi divergence measure [3], [7], [13], [14] as well as Bhattacharyya histogram-based coefficient [18], [21].

In particular, such measures cannot satisfy the following properties simultaneously

- **[C1]** a lag at zero entails a global maximum and vice versa (*positive-definiteness*)
- **[C2]** the similarity measure behaves continuously with respect to changes of the lag (*continuity*)
- **[C3]** an increasing lag implies a decreasing similarity and vice versa (*monotonicity*).

At least for non-negative functions $f \geq 0$ the conditions **[C1]-[C3]** are met by an approach based on Hermann Weyl's discrepancy measure which is outlined in the next sections.

Jean-Luc Bouchot is with the Fuzzy Logic Laboratory Linz-Hagenberg, Department of Knowledge-Based Mathematical Systems, University of Linz, Austria phone: +43 732 3343 432; email: jean-luc.bouchot@jku.at

Johannes Himmelbauer and Bernhard Moser are with the Knowledge-Based Technology Area, Software Competence Center Hagenberg, Austria (phone: +43 699 1 3343 833; email: {johannes.himmelbauer, bernhard.moser}@scch.at)

978-1-4244-8126-2/10/$26.00 ©2010 IEEE

In Section II we introduce the discrepancy norm based on Hermann Weyl's concept, and an autocorrelation based on this norm in Section III. Computational aspects are considered in Section IV which also proposes an approximation that can be computed as convolution. Section V shows an application on financial data.

II. Hermann Weyl's Discrepancy Measure

For a summable sequence of real values $f = (f_i)_{i\in\mathbb{Z}}$, $\sum_{i\in\mathbb{Z}} |f_i| < \infty$, Weyl's discrepancy concept, see [17], leads to the definition

$$\|f\|_D := \sup\left\{\left|\sum_{i=n_1}^{n_2} f(i)\right| : n_1, n_2 \in \mathbb{Z}\right\}. \quad (1)$$

Applications of Weyl's concept of discrepancy can be found in the field of numerical integration, especially for Monte Carlo methods in high dimensions [12], in computational geometry [5] and in pattern recognition [11]. There are also applications in image processing in the context of pixel classification [2], see also [8],[9], [10] for extensions to higher dimensions and applications to texture analysis and tracking.

What is special about this discrepancy concept is its dependency from the ordering, see also [11] [10], for example

$$\begin{aligned}\|(1,1,-1)\|_D &= 2\\ \|(1,-1,1)\|_D &= 1.\end{aligned}$$

In this paper we are studying its behavior in the context of time-shifted signals, therefore for the interval $\mathcal{I} \subset \mathbb{Z}$ let us define

$$F_{\mathcal{I}} := \{f : \mathcal{I} \subseteq \mathbb{Z} : \sum_{i\in\mathcal{I}} |f(i)| < \infty\},$$

$$\|f\|_D^{\mathcal{I}} := \|f(.)1_{\mathcal{I}}(.)\|_D \quad (2)$$

and

$$\alpha_{[f]}^{\mathcal{I}}(t) := \|f(.) - f(.-t)\|_D^{\mathcal{I}}. \quad (3)$$

Definition (2) refers to the restriction of f to a subinterval domain $\mathcal{I}$, where $1_{\mathcal{I}}$ denotes the indicator function. Definition (3) measures the discrepancy of a time shifted signal $f(.-t)$ at lag t with its reference f at lag zero. See Figure 2 which shows the behavior of the discrepancy in terms of the difference with a time-shifted signal, $f(.) - f(.-t)$ for the simple but illustrative example depicted in Figure 1. It is interesting to observe that $\alpha_{[f]}^{\mathcal{I}}$ is monotonically increasing on $t \geq 0$ whereas Pearson's autocorrelation $(f \odot f)(t) = \sum_i f(i)f(i+t)$, as well as Spearman's and Kendall's rank correlation coefficient, show a non-monotonic behavior, see Figure 3.

With the above definitions we obtain the following properties, its proofs can be found in the Appendix, VI-A:

- $\|.\|_D^{\mathcal{I}}$ is a norm on $F_{\mathcal{I}}$

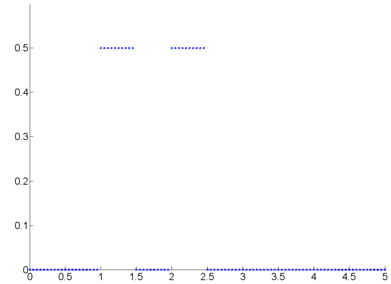

Fig. 1. Illustrative example s with two steps

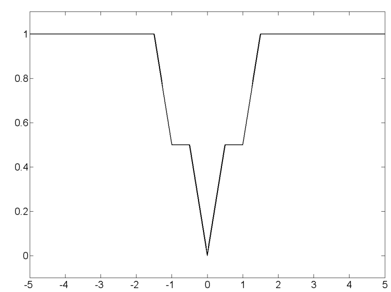

Fig. 2. $\|s(.) - s(.-t)\|_D$ as function of t for example s in Figure 1, which is monotonically increasing for $t \geq 0$

- with $f \in F_{\mathcal{I}}$ and $\mathcal{I}_k = \{n \in \mathbb{Z} : n \leq k\}$:

$$\begin{aligned}\|f\|_D^{\mathcal{I}} &= \max\{0, \sup_{k\in\mathcal{I}} \sum_{i\in\mathcal{I}\cap\mathcal{I}_k} f(i)\}\\ &- \min\{0, \inf_{k\in\mathcal{I}} \sum_{i\in\mathcal{I}\cap\mathcal{I}_k} f(i)\} \quad (4)\end{aligned}$$

- $\alpha_{[f]}^{\mathcal{I}}(0) = 0$ for $f \in F_{\mathcal{I}}$
- Let $f \in F_{\mathbb{Z}}$, $f \geq 0$, then $\alpha_{[f]}^{\mathbb{Z}}(t) = \alpha_{[f]}^{\mathbb{Z}}(-t)$
- Let $f \in F_{\mathbb{Z}}$, $f \geq 0$, then $\alpha_{[f]}^{\mathbb{Z}}$ is monotonically increasing on $\mathbb{N} \cup \{0\}$

Equation (4) allows us to compute the discrepancy of a sequence of length n with $O(n)$ operations instead of $O(n^2)$ number of operations resulting from the original definition (1).

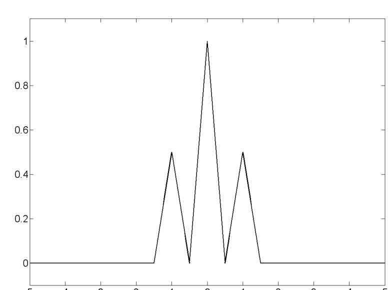

Fig. 3. Autocorrelation for example s in figure 1, which is not monotonic on $t \geq 0$

III. Autocorrelation Based on Hermann Weyl's Discrepancy Norm

In this section a correlation coefficient based on the discrepancy norm is proposed. The construction imitates the classical Pearson's formula, which can be expressed by means of an inner product

$$\mathrm{corr}(f,g) = \frac{\langle f,g\rangle}{\|f\|\|g\|}.$$

For any Hilbert space with inner product $\langle .,.\rangle$ the parallelogram law can be derived $\|f-g\|^2 + \|f+g\|^2 = 2(\|f\|^2 + \|g\|^2)$, which together with $\|f+g\|^2 = \|f\|^2 + \|g\|^2 + 2\langle f,g\rangle$, induces a representation of the inner product by means of the norm

$$\langle f,g\rangle = (\|f+g\|^2 - \|f-g\|^2)/4. \qquad (5)$$

Note that the discrepancy norm does not fulfill the parallelogram law, consider, e.g., $f = (1,1,-1)$ and $g = (1,-1,1)$, then $\|f-g\|_D^2 + \|f+g\|_D^2 = \|(2,0,0)\|_D^2 + \|(0,2,2)\|_D^2 = 20 \neq 2(\|f\|_D^2 + \|g\|_D^2) = 10$. Anyway, formula (5) yields a reasonable correlation coefficient for not vanishing $f, g \in F_{\mathcal{I}}$

$$\mathrm{corr}_D(f,g) := \frac{1}{4}\frac{\|f+g\|_D^2 - \|f-g\|_D^2}{\|f\|_D^2\|g\|_D^2} \in [-1,1]. \qquad (6)$$

The range of $\mathrm{corr}_D(.,.) \in [-1,1]$ directly follows from the triangle inequality of the norm, so does $\mathrm{corr}_D(f,\lambda f) = 1$ for $\lambda \in \mathbb{R}\backslash\{0\}$.

Observe that for $f \geq 0$ we have $\|f\|_D = \|f\|_1 = \sum_f(i)$, therefore, for $g \geq f \geq 0$, $f \neq 0$ we obtain

$$\begin{aligned}\mathrm{corr}_D(f,g) &= \frac{(\sum_i f(i)+g(i))^2 - (f(i)-g(i))^2}{4\sum_f f(i)\sum_i g(i)} \\ &= 1.\end{aligned}$$

If we correlate f with itself for a nonnegative $f \geq 0$ note that

$$\begin{aligned}&\mathrm{corr}_D(f(.), f(.-t)) \\ =\ & \frac{1}{4}\frac{\|f(.)+f(.-t)\|_D^2 - \|f(.)-f(.-t)\|_D^2}{\|f(.)\|_D^2 + \|f(.-t)\|_D^2} \\ =\ & \frac{4\|f(.)\|_D^2 - \|f(.)-f(.-t)\|_D^2}{4\|f(.)\|_D^2}\end{aligned}$$

which means that, due to the monotonicity of the discrepancy norm of the difference of time-shifted non-negative functions as a function of the time-shift, the autocorrelation based on the discrepancy norm

$$\mathrm{autocorr}_D[f](t) := \mathrm{corr}_D(f(.), f(.-t)) \qquad (7)$$

is monotonically decreasing for $t \geq 0$.

IV. Approximation and Computation by Convolution

In this section the discrepancy norm is approximated by a formula that can be computed by convolution which reduces the computational complexity to $O(n\log(n))$ with n the number of discrete support points. In this section we restrict to finite support intervals $\mathcal{I}$, $n = |\mathcal{I}| > 1$. For convenience without loss of generality, we restrict to intervals $\mathcal{I} = \{1,\ldots,n\}$ starting at 1. Further on, for convenience let us set

$$\begin{aligned}F_k(f) &:= \sum_{i=0}^{k} f_i \\ \chi_p[f](k) &:= e^{p\sum_{i=0}^{k} f_i} = e^{pF_k(f)}\end{aligned}$$

where by definition we set $f_0 = 0$. The t-mean M_t is defined by

$$M_t(f) \;=\; \left(\frac{1}{N+1}\sum_{i=0}^{N}|f_i|^t\right)^{\frac{1}{t}}$$

which for $t = +1$ yields the arithmetic and for $t = -1$ yields the harmonic mean. Then, let us define

$$\Gamma_p(f) = \frac{1}{p}\ln\left(\frac{M_{+1}(\chi_p[f])}{M_{-1}(\chi_p[f])}\right) \qquad (8)$$

which is well defined due to the fact that the t-mean M_t is monotonically increasing with t and, that due to the construction of χ_p we have $0 < \min_{k\in\{0,\cdots,N\}}\chi_p[f](k) \leq M_{-1}(\chi_p[f]) \leq M_1(\chi_p[f])$. Though it can be proved that Γ_p does not satisfy all norm properties, we get similar and approximative results, see Appendix, VI-B.

- Γ_p is positive definite
- approximative homogeneity: $\Gamma_p(\lambda f) = |\lambda| \cdot \Gamma_{|\lambda|p}(f)$
- approximative triangular inequality: $\Gamma_p(f+g) \leq \Gamma_p(f) + \Gamma_p(g) + \frac{2}{p}\ln(n+1)$

Finally we state the main result of this section:

Theorem 1: Γ_p approximates the discrepancy norm in the sense of

$$\Gamma_p(f) \leq ||f||_D < \Gamma_p(f) + \frac{2}{p}\ln(n+1) \qquad (9)$$

and the correlation of two signals f and g can be computed by convolution

$$h \;=\; \frac{1}{p}\ln(F * G) + \frac{1}{p}\ln(F^{-1} * G^{-1}) \qquad (10)$$

where $F = \chi_p[f]$, $G = \chi_{-p}[g]$.

Figure 4 illustrates the approximation behavior of Γ_p for the 2-step function of Figure 1.

V. Application

A typical problem for correlation is to find and locate a known patch of pattern in a longer signal. Such a problem is illustrated in Figure 5, where a similar test patch should be identified in the data by checking the correlation at each position. Figure 6 shows a noisy test patch and its reference which corresponds to the data between 450 till 550 in Figure 5. The resulting correlation measures are shown in Figure 7, which show similar behavior at least for sufficiently large parameter p. Both of them, the discrepancy norm and its approximation show a significant and well shaped minimum at the correct position 450. Figure 8 shows time series of so-called daily asset returns for three different companies.

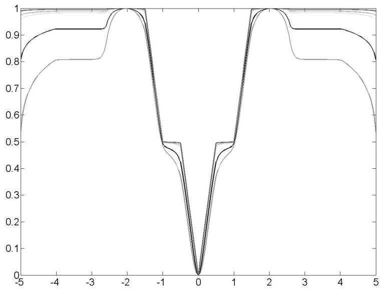

Fig. 4. Illustration of the approximation behavior of Γ_p

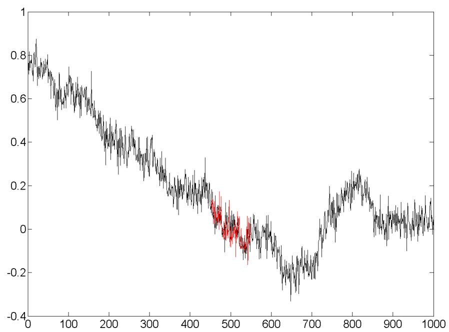

Fig. 5. Signal with reference patch between 450 and 550

The data refer to the period between February 2005 and end of July 2008 with real stock price data. Figure 10 depicts autocorrelation curves according to $\text{autocorr}_D(f)$ where f is supposed to be normalized in the sense that the arithmetic mean is zero and its standard deviation is one. Observe that the discrepancy induced autocorrelation curves in Figure 10 can be distinguished from each other, and, moreover shifts in the lag only cause small changes in the autocorrelation in contrast to Pearson's correlation where a peak appears at lag zero while even very small lag deviations from zero lead to an abrupt change of the autocorrelation. It is interesting to observe that for the classical autocorrelation the time series of Figure 8 can not be discriminated as is demonstrated by Figure 9 in contrast to the autocorrelation based on the discrepancy norm due to Equation (7), see Figure 9.

VI. Conclusions

In this paper we have introduced a novel approach to time series autocorrelation based on the so-called Hermann Weyls discrepancy norm. In particular its autocorrelation behavior was studied and tested on real data from finance showing that in contrast to the standard concepts used in statistics and signal processing like Pearsons correlation coefficient this novel concept shows promising discriminative potential. The reason for this better behavior is that the discrepancy norm is less sensitive to shifts in the time domain which in particular is the more important the higher the amount of high frequencies in the signal as this is for example the case for time series of daily asset returns. Moreover, an approximation

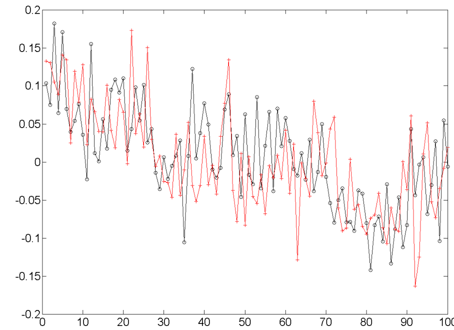

Fig. 6. Comparison of noisy test signal with its reference signal related to the interval $[450, 550]$ of Figure 5

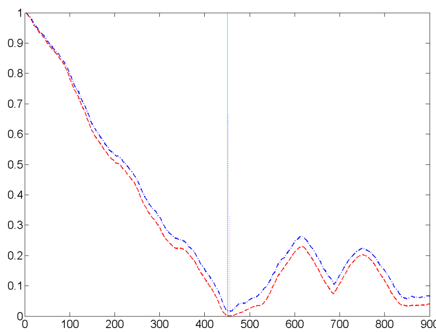

Fig. 7. Discrepancy correlation coefficient and its Γ_p approximation with $p = 128$ for the reference patch of Figure 5; the upper curve refers to the approximation. Both curves show a well formed minimum around the correct position at 450

to this discrepancy correlation coefficient was proposed that allows to efficiently compute auto- and cross-correlation with $O(nlog(n))$ operations which is of the same order than the standard correlation. In future, statistical tests based on these novel concepts should be designed in order to provide a more sensitive and adequate correlation analysis theory to high frequency time series data.

Appendix

A. Properties of the discrepancy norm

In order to show that $\|.\|_D^{\mathcal{I}}$ is a norm on $F_{\mathcal{I}}$ observe that $\|f\|_D^{\mathcal{I}} = 0$ implies $\sup_{n,m\in\mathcal{I}} |\sum_{i=m}^{n} f(i)| = 0$, hence $|f(i)| = 0$ and, therefore $f(i) = 0$ for all $i \in \mathcal{I}$. Homogeneity immediately follows from the definition. For the triangle inequality consider

$$
\begin{aligned}
\|f+g\|_D^{\mathcal{I}} &= sup_{n_1,n_2} |\sum_{i=n_1}^{n_2} f(i) - g(i)| \\
&\leq sup_{n_1,n_2} \{\sum_{i=n_1}^{n_2} f(i)| + \sum_{i=n_1}^{n_2} g(i)|\} \\
&\leq sup_{n_1,n_2} \{\sum_{i=n_1}^{n_2} f(i)|\} + sup_{n_1,n_2} \{\sum_{i=n_1}^{n_2} g(i)|\}.
\end{aligned}
$$

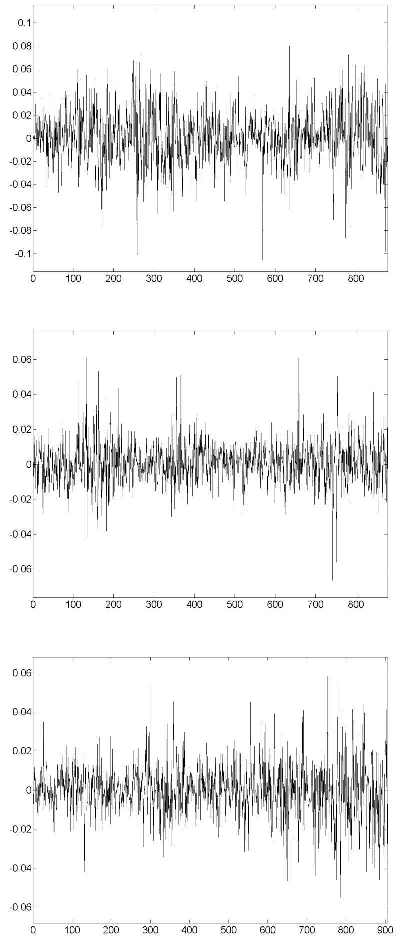

Fig. 8. Daily asset return curves of three different companies related to W&T OFFSHORE INC, MCDONALDS CORP and BMW in the period between February 2005 and end of July 2008

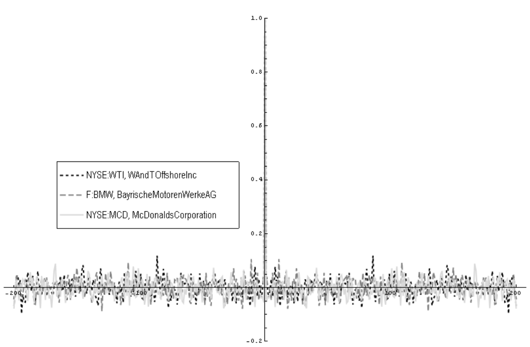

Fig. 9. Pearson's autocorrelation for the time series of the daily asset returns of Figure 8

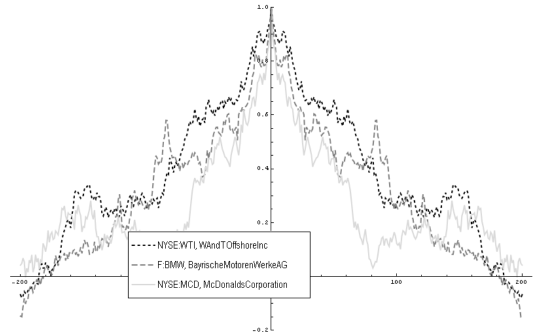

Fig. 10. Discrepancy induced autocorrelation for the time series of the daily asset returns of Figure 8

To prove Equation (4) let us set $\tilde{f}(i) = 0$ for $i \in \mathbb{Z}\backslash\mathcal{I}$, and $\tilde{f}(i) = f(i)$ for $i \in \mathcal{I}$, then we get

$$\begin{aligned}
\|f\|_D^{\mathcal{I}} &= \sup_{n_1,n_2\in\mathbb{Z}} |\sum_{i=n_1}^{n_2} \tilde{f}(i)| \\
&= \sup_{n_1,n_2\in\mathbb{Z}} |\sum_{i=-\infty}^{n_2} \tilde{f}(i) - \sum_{i=-\infty}^{n_1-1} \tilde{f}(i)| \\
&= \sup_{n_2\in\mathbb{Z}} \sum_{i=-\infty}^{n_2} \tilde{f}(i) - \inf_{n_2\in\mathbb{Z}} \sum_{i=-\infty}^{n_1} \tilde{f}(i) \\
&= \max\{0, \sup_{k\in\mathcal{I}} |\sum_{i\in\mathcal{I}\cap\mathcal{I}_k} f(i)|\} - \\
&\quad \min\{0, \inf_{k\in\mathcal{I}} |\sum_{i\in\mathcal{I}\cap\mathcal{I}_k} f(i)|\}.
\end{aligned}$$

The positive-definiteness of $\|f\|_D^{\mathcal{I}}$ immediately implies $\alpha_{[f]}^{\mathcal{I}} = 0$ for $f \in F_{\mathcal{I}}$. Now, let us consider $f \in F_{\mathbb{Z}}$ and let us assume $f \geq 0$, then both the monotonicity as well as the symmetry property follows from the observation

$$\begin{aligned}
\|f(.) - f(.-t)\|_D^{\mathbb{Z}} &= \max\{0, \sup_{n\in\mathbb{Z}} |\sum_{i=n-t+1}^{n} f(i)|\} - \\
&\quad \underbrace{\min\{0, \inf_{n\in\mathbb{Z}} |\sum_{i=n-t+1}^{n} f(i)|\}}_{0} \\
&= \max\{0, \sup_{n\in\mathbb{Z}} |\sum_{i=n-t+1}^{n} f(i)|\}.
\end{aligned}$$

B. Properties of Γ_p

Positive definiteness of Γ_p follows from

$$\begin{aligned}
\Gamma_p(f) = 0 &\Leftrightarrow \\
M_1(\chi_p[f]) = M_{-1}(\chi_p[f]) &\Leftrightarrow \\
\forall k \in \ \{0, \cdots, N\},\ \chi_p[f] = C &\Leftrightarrow \\
\forall k \in \ \{0, \cdots, N\},\ \sum_{i=0}^{k} f_i = C' &\Leftrightarrow \\
\forall i \in \{0, \cdots, N\},\ f_i = 0.&
\end{aligned}$$

The approximative homogeneity formular follows from

$$\Gamma_p(\lambda \cdot f) = \frac{|\lambda|}{|\lambda| p} \ln\left(\frac{\sum_{k=0}^{n} e^{\lambda p F_k}}{n+1} \cdot \frac{\sum_{k=0}^{n} e^{-\lambda p F_k}}{n+1}\right)$$

To show the approximative triangular inequality let us consider

$$\Delta_p(f,g) = \Gamma_p(f) + \Gamma_p(g) - \Gamma_p(f+g)$$

then

$$\begin{aligned} & \Delta_p(f,g) \\ = & \frac{1}{p}\ln\left(\sum_{k=0}^{n} e^{F_k(f)} \cdot \sum_{k=0}^{n} e^{-F_k(f)}\right) \\ & +\frac{1}{p}\ln\left(\sum_{k=0}^{n} e^{F_k(g)} \cdot \sum_{k=0}^{n} e^{-F_k(g)}\right) \\ & -\frac{1}{p}\ln\left(\sum_{k=0}^{n} e^{F_k(f+g)} \cdot \sum_{k=0}^{n} e^{-F_k(f+g)}\right) - \frac{2}{p}\ln(n+1) \\ = & \frac{1}{p}\ln\left(\frac{\sum e^{F_k(f)} \cdot \sum e^{F_k(g)}}{\sum e^{F_k(f+g)}}\right) \\ & +\frac{1}{p}\ln\left(\frac{\sum e^{-F_k(f)} \cdot \sum e^{-F_k(g)}}{\sum e^{-F_k(f+g)}}\right) - \frac{2}{p}\ln(n+1) \end{aligned}$$

and, finally, we get $\Delta_p(f,g) + \frac{2}{p}\ln(N+1) \geq 0$ which yields $\Gamma_p(f+g) \leq \Gamma_p(f) + \Gamma_p(g) + \frac{2}{p}\ln(n+1)$

C. Proof of theorem 1

For convenience we simplify notation by setting $\chi_i = \chi_i[f]$. First of all let us proof the inequality by considerung

$$\begin{aligned} ||f||_D &= \max_{k\in\{0,\cdots,n\}} F_k + \max_{k\in\{0,\cdots,n\}} -F_k \\ &= \ln(\max_{k\in\{0,\cdots,n\}} e^{F_k}) + \ln(\max_{k\in\{0,\cdots,n\}} e^{-F_k}) \end{aligned}$$

which leads to

$$\begin{aligned} \Gamma_p(f) - \|f\|_D &= \ln\left(\frac{\|\chi_1\|_p}{\sqrt[p]{n+1}} \frac{\|\chi_{-1}[f]\|_p}{\sqrt[p]{n+1}}\right) \\ &\quad - \ln\|\chi_1\|_\infty - \ln\|\chi_{-1}\|_\infty \\ &= \ln\left(\frac{\|\chi_1\|_p}{\sqrt[p]{n+1}\|\chi_1\|_\infty} \frac{\|\chi_{-1}\|_p}{\sqrt[p]{n+1}\|\chi_{-1}\|_\infty}\right). \end{aligned}$$

Recall that for $N \in \mathbb{N}$ and a non-vanishing $x \in \mathbb{R}\backslash\{0\}$ we have

$$\|x\|_\infty \leq \|x\|_p \leq \sqrt[p]{N}\|x\|_\infty,$$

hence

$$\frac{1}{\sqrt[p]{N}} \leq \frac{\|x\|_p}{\sqrt[p]{N}\|x\|_\infty} \leq 1.$$

By this, we finally obtain

$$\ln\left(\left(\frac{1}{\sqrt[p]{n+1}}\right)^2\right) \leq \Gamma_p(f) - \|f\|_N \leq 0$$

which is equivalent to

$$\Gamma_p(f) \leq \|f\|_D \leq \Gamma_p(f) + \frac{2}{p}\ln(n+1).$$

In the following we consider the convolution property. Assume that g is defined on an interval with m and f on an interval with n points. The convolution can be seen as follows by introducing the auxiliary function h:

$$h(i) = \Gamma_p(f(i) - g).$$

By

$$\sum_{j=1}^{k} f_{i+j} = F_{k+i}(f) - F_i(f)$$

we otbain

$$\begin{aligned} h(i) &= \frac{1}{p}\ln\left(\sum_{k=1}^{m} e^{p(F_{k+i}(f)-F_i(f)-F_k(g))}\right) \\ &\quad +\frac{1}{p}\ln\left(\sum_{k=1}^{m} e^{-p(F_{k+i}(f)-F_i(f)-F_k(g))}\right) \\ &= \frac{1}{p}\ln\left(\sum_{k=1}^{m} e^{p(F_{k+i}(f)-F_k(g))}\right) \\ &\quad +\frac{1}{p}\ln\left(\sum_{k=1}^{m} e^{-p(F_{k+i}(f)-F_k(g))}\right) \\ &= \frac{1}{p}\ln\left(\sum_{k=1}^{m} e^{pF_{k+i}(f)} \cdot e^{-pF_k(g)}\right) \\ &\quad +\frac{1}{p}\ln\left(\sum_{k=1}^{m} e^{-pF_{k+i}(f)} \cdot e^{pF_k(g)}\right). \end{aligned}$$

Setting $g = 0$ out of the domain proofs that h can be written as a sum of two convolutions.

Acknowledgment

This work was supported in part by the Austrian Science Fund (FWF) under grant no. P21496 N23 and the Austrian COMET programm.

References

[1] A. Hamza, Y. He and H. Krim, An information divergence measure for ISAR image registration, Proc. IEEE Workshop on Statistical Signal Processing, Singapore, pages 130–133, 2001.

[2] P. Bauer, U. Bodenhofer and E. P. Klement, A Fuzzy Algorithm for Pixel Classification based on the Discrepancy norm, Proc. 5th IEEE Int. Conf. on Fuzzy Systems, III, New Orleans, LA, pages 2007–2012, 1996.

[3] J. F. Bercher, On some entropy functionals derived from Rényi information divergence, Inf. Sci., 178(12), Elsevier Science Inc., New York, NY, USA, pages 2489–2506, 2008.

[4] J. Campbell, A. Lo and C. MacKinlay, The Econometrics of Financial Markets. Princeton University Press, Princeton, NJ., 1996.

[5] B. Chazelle, The Discrepancy Method: Randomness and Complexity, Cambridge University Press, New York, NY, USA, 2000.

[6] M.G. Kendall and A. Stuart, The Advanced Theory of Statistics, 4th Edition, Vol. 2., MacMillan, New York, 1979.

[7] S. Kullback, Information Theory and Statistics, Wiley, New York, NY, USA, 1959.

[8] B. Moser and T. Hoch, Misalignment Measure Based on Hermann Weyl's Discrepancy, In A. Kuijper, B. Heise and L. Muresan, editors, Challenges in the Biosciences: Image Analysis and Pattern Recognition Aspects, Proc. 32nd Workshop of the Austrian Association for Pattern Recognition (AAPR/OAGM), 232, Linz, Austria, pages 187–198, 2008.

[9] B. Moser, T. Kazmar and P. Haslinger, On the potential of Hermann Weyls discrepancy norm for texture analysis. In M. Masoud Mohammadian (editor). Proceedings of the International Conference on Computational Intelligence for Modelling, Control and Automation (CIMCA 2008), pages 187-191, IEEE Computer Society, 2008.

[10] B. Moser, Similarity measure for image and volumetric data based on Hermann Weyls discrepancy. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), volume 99, number 1, DOI 10.1109/TPAMI.2009.50, 2009.

[11] H. Neunzert and B. Wetton, Pattern Recognition Using Measure Space Metrics, University of Kaiserslautern, Department of Mathematics, (Technical Report), 28, 1987.

[12] H. Niederreiter, Random number generation and quasi-Monte Carlo methods, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1992.

[13] J.P.W. Pluim, J.B.A. Maintz and M.A. Viergever, f- Information measures in medical image registration, IEEE Transactions on Medical Imaging, vol. 23, 12, pages 1506-1518, 2004.

[14] A. Renyi, On measures of entropy and information, In Proc. of the 4th Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, pages 547-561, 1961.

[15] C. Spearman, The proof and measurement of association between two things, Amer. J. Psychol., 15, pages 72–101, 1904.

[16] I. Vajda, Theory of Statistical Inference and Information, The Netherlands, Kluwer Academic, Dordrecht, 1989.

[17] H. Weyl, Über die Gleichverteilung von Zahlen mod. Eins, Math. Ann., 77, pages 313–352, 1916.

[18] A. Bhattacharyya, On a measure of divergence between two statistical populations defined by probability distributions, Bull. Calcutta Math, 35, pages 99–109, 1943.

[19] R. Jain, S.N.J. Murthy, P.L.J. Chen and S. Chatterjee, Similarity Measures for Image Databases, Proc. 1995 IEEE International Conference on Fuzzy Systems. International Joint Conference of the Fourth IEEE International Conference on Fuzzy Systems and The Second International Fuzzy Engineering Symposium., 3, pages 1247–1254, 1995.

[20] W. Jiang, G. Er, Q. Dai and J. Gu, Similarity-Based Online Feature Selection in Content-Based Image Retrieval IEEE Transactions on Image Processing , 15(3), pages 702–712, 2006.

[21] M.S. Khalid and M.B. Malik, Biased nature of Bhattacharyya coefficient in correlation of gray-scale objects, Proc. 4th International Symposium on Image and Signal Processing and Analysis, pages 209–214, 2005.

[22] O. Michailovich, Y. Rathi and A. Tannenbaum, Image Segmentation Using Active Contours Driven by the Bhattacharyya Gradient Flow IEEE Transactions on Image Processing, 16(11), pages 2787–2801, 2007.

[23] A. Pérez-Garca and V. Ayala-Ramírez and R. E. Sánchez-Yá ñez and J. G. Avi ña-Cervantes, Monte Carlo Evaluation of the Hausdorff Distance for Shape Matching, In J. F. M. Trinidad, J. A. Carrasco-Ochoa and J. Kittler, editors, CIARP, Springer, Lecture Notes in Computer Science, 4225, pages 686–695, 2006.

[24] J.P.W. Pluim and J.B.A. Maintz and M. A. Viergever, Mutual information-based registration of medical images: a survey, IEEE Transactions on Medical Imaging, volume = 22(8), pages 986–1004, 2003.

[25] B. Zitová and J. Flusser, Image registration methods: a survey, Image Vision Comput., 21(11), pages 977–1000, 2003.