Shifts between two periodic time series

1st Piotr Marian Zelewski
Dundas, Ontario, Canada
pzelewski@cogeco.ca

2nd Valery P Voevudsky
Toronto, Ontario, Canada
v.voevudsky@gmail.com

***Abstract*—This paper presents results on the discrete mathematics of time delay (a.k.a. cross-correlation) [4][7] between two time series f, g indexed by a discrete ordered set $\{0, 1, \ldots, m-1\} = [0, m)$, with the time delay modeled by circular permutations of $[0, m)$. Vectors f, g can represent measurements of a quantity (i.e. price) changes at two distinct points x, y in the phase space, in which case interpolating the lagging signal at x by shifting the leading signal from y or even measuring the change in the lead/lag relationship is of interest. Let Z/mZ denotes the cyclic group of order m with generator σ acting freely and transitively on $[0, m)$ by the 'circular shift' $\sigma(t) = t+1 \mod m$ identyfying Z/mZ with $[0, m)$ by $\sigma^t \to \sigma^t(0) = t \in [0, m)$. We equip $[0, m)$ with a Z/mZ invariant measure giving each of its elements the equal weight. The vector space of complex valued vectors $Map([0, m), C)$ has left action $f \to f^s$ of Z/mZ given by the circular permutation of f given by the equation $f_i^s = f_{i+s \mod m}$. This action preserves the hermitian product $< f, g >= \sum_{t\in[0,m)} f(t)\overline{g(t)}$ and the corresponding norm $||$. Given a pair $g(t), f(t) \in Map([0, m), C)$ we define the energy function $d(f, g)(s) = |g - f^s|^2$ and the crosscorrelation function [4] $\psi(f, g)(s) = Re < g, f^s >$. We define $\mathbf{S}(f, g) = \arg \max_{t\in Z/mZ}(\psi) \subset Z/mZ$. Since $\arg \min_{t\in Z/mZ}(d(t)) = \arg \max_{t\in Z/mZ}(\psi(t))$, hence $s \in \mathbf{S}(f, g)$ is the circular transformation of $[0, m)$ for which f is the best approximation to g^s in $||$-norm. In section 1 we establish basic results for how perturbations of f, g affect $\mathbf{S}(f, g)$. In Section 2 we write out an equation relating $\psi(f^*, g^*)$ to $\psi(f, g)$ where f^*, g^* denote restrictions of f, g to a sub-interval $[0, m') \subset [0, m)$ for $m' < m$. In section 3 we derive the well know equation $\psi(f, g) = f * g^\tau$ where $*$ denotes the convolution on $Map([0, m), C)$ (induced by addition in $[0, m)$) and τ the involution on $Map([0, m), C)$ induced by the minus operator on $[0, m)$. Using this, we show that the associativity and commutativity of the convolution operator [6] implies, for generic, nonnegative f, g, that $\mathbf{S}(f, g)$ stays unchanged when the mesh $[0, m)$ gets refined to $[0, ml)$ by adding an equal number l of points between each consecutive pair $\{k, k+1\}$ and $g, f \in Map([0, m), C)$ are extended to $[0, ml)$ using the customary 'as of the last available measurement' rule. We also re-proof the well known fact [5] that DFT O(mlog(m))-computes $\mathbf{S}(f, g)$. The last section summarizes the case when f^*, g^* are sampled from f, g that are l-periodic, nonnegative and containing sufficiently many fundamental l-periods $\hat{f}, \hat{g}$. This occurs when f^*, g^* approximate signals emitted by a high frequency cyclical process. We show that $\mathbf{S}(\hat{f}, \hat{g}) = \{0\}$ implies $\{0\} \subset \mathbf{S}(f^*, g^*)$ while $\mathbf{S}(\hat{f}, \hat{g}) = \{p\}$ with $p \neq 0$ implies $\mathbf{S}(f^*, g^*) \cap \{p, p-l\} \neq \emptyset$. Thus, asymptotically, none of these expressions contain the size m of the of data windows as depenedency.**

***Index Terms*—time series, cross-correlation, time delay, convolution, FFT, DFT, best fit**

I. Best fit and cross-correlation

A. Definitions

Let $T_m = Z/mZ$ denotes the cyclic group of order m with generator σ acting freely and transitively on $[0, m) = \{0, 1, \ldots, m-1\}$ by the 'circular shift' $\sigma(t) = t+1 \mod m$ identyfying T_m with $[0, m)$ by $\sigma^t \to \sigma^t(0) = t \in [0, m)$. Let $\mathbb{C}([0, m))$ denote the vector space of functions on $[0, m)$ with values in the field of complex numbers $\mathbb{C}$. Clearly, any $f \in \mathbb{C}([0, m))$ is a complex m-dimensional vector $(f_0, f_1, \ldots, f_{m-1})$. We equip $[0, m)$ with a T_m invariant measure giving each of its elements the equal weight. Then $\mathbb{C}([0, m))$ has left action $f \to f^s$ of T_m given by the circular permutations of f: $f_i^s = f_{i+s \mod m}$. This action preserves the hermitian product $< f, g >= \sum_{t\in[0,m)} f(t)\overline{g(t)}$ and the corresponding norm $||$.

Let τ denotes the involution of $[0, m)$ given by $t \to m - t$. Since $m - t = -t \mod m$ we can equivalently define τ as $\tau(t) = -t$ on the abelian group T_m. Then T_m and $Z/2Z$ act on $\mathbb{C}([0, m))$ via $\sigma_m^s(f)(t) = f(\sigma_m^{-s}(t)) = f(t-s \mod m)$, $\tau(f)(t) = f(-t)$ for $f \in \mathbb{C}(T_m)$. To simplify notation we will denote $\sigma_m^s(f)$ as f^s. We have $\tau(f^s)(t) = f^s(-t) = f(-t-s \mod m)$ and $\tau(f)^s(t) = \tau(f)(t-s) = f(s-t)$ from which we get formulae $\tau(f^s) = \tau(f)^{-s}$. It follows that $\mathbb{C}([0, m))$ is acted in the natural way by the dihedral group D_m [3], which is not used in this paper.

Let $\delta_k \in \mathbb{C}([0, m))$ denotes the Dirac delta function on $[0, m)$ equal to 1 at k and 0 elswehere. Clearly δ_i, $0 \leq i \leq m$ is a basis of $\mathbb{C}([0, m))$ and $\delta_k^s = \delta_{k+s \mod m}$, $\tau(\delta_k) = \delta_{-k} = \delta_{m-k}$. Explicitely, given vector $f = (f_0, f_1, \ldots, f_{m-1}) = \sum f_i\delta_i$, we have: $f^s = (f_{m-s \mod m}, f_{m-s+1 \mod m}, \ldots, f_0, f_1, \ldots, f_{m-s-1}) = \sum f_i\delta_{i+s \mod m}$ and $\tau(f) = (f_0, f_{m-1}, f_{m-2}, \ldots, f_1) = \sum f_i\delta_{m-i \mod m}$.

Given $f, g \in \mathbb{C}([0, m))$ consider the best fit function $d(f, g) : T_m \to \mathbb{R}^+$ given by $d(f, g)(t) = ||f - g^t||^2$ and the circular cross-correlation function [4],[7] $\psi(f, g)(t) = Re(< f, g^t >)$ where $Re(z)$ denotes the real part of a complex number z. Since $\|\|$ is T_m invariant, it is easy to see, using the expansion $\|f - g^t\|^2 =< f - g^t, f - g^t >= ||f||^2 + ||g||^2 - 2Re(< f, g^t >)$, that the minimums of $d()$ coincide with the maximums of $\psi()$.

Definition 1.1: Define $\mathbf{S}(f, g) = \arg \max_{t\in T_m}(\psi(f, g)(t)) \subset T_m$ and $gap(f, g)$ as the difference separating the maximum value from all other values $\psi(f, g)(t)$, $t \in T_m$.

Thus $\mathbf{S}(f,g)$ consists of $s \in T_m$ for which g^t is the best approximation of f in $\|\|$ norm or, equivalently, for which the circular cross-correlation function $\psi(f,g)$ attains maximum. Clearly, for any pair $f,g \in \mathbb{C}(T_m)$, $\emptyset \neq \mathbf{S}(f,g) \subset T_m$.

B. Elementary properties of $\mathbf{S}(f,g)$

Theorem 1.2:

For any $f,g,h \in \mathbb{C}([0,m])$:

1) $\mathbf{S}(f,\delta_0)$ (resp. $\mathbf{S}(f,-\delta_0)$ is the set of maximums (resp. minimums) of f
2) $\mathbf{S}(f,f) = T_m$ if and only if f is a constant function
3) for any $c \in \mathbb{C}$: $\mathbf{S}(c,f) = T_m$,
4) for any $s,t \in T_m$ we have $\mathbf{S}(f^s,g^t) = \mathbf{S}(f,g)^{t-s} = t-s+\mathbf{S}(f,g) \mod m$
5) for any $t \in T_m$ we have $\mathbf{S}(f,g) = \mathbf{S}(f^t,g^t)$
6) $\mathbf{S}(g,f) = \tau(\mathbf{S}(f,g)) = m - \mathbf{S}(f,g) \mod m = -\mathbf{S}(f,g) \mod m$
7) for any $c \in \mathbb{C}$: $\mathbf{S}(g+c,f) = \mathbf{S}(g,f)$
8) for $r \in \mathbb{R}$, $r > 0$: $\mathbf{S}(rg,f) = \mathbf{S}(g,f)$
9) $\mathbf{S}(-g,f) = \arg\min_{t\in T_m}(\psi(f,g)(t)) \subset T_m$

Proof: We leave this as excercise to the reader. ∎

Given a set S write $\#S$ for the cardinality of S.

Theorem 1.3:

Given $f \in \mathbb{C}([0,m))$:

1) For any $g \in \mathbb{C}([0,m))$ and for any $\epsilon > 0$ there exist $h_1,h_2 \in \mathbb{C}([0,m))$ with $\|h_1\|^2, \|h_2\|^2 < \epsilon$ such that $\#\mathbf{S}(f+h_1,g+h_2) = 1$
2) For any $g \in \mathbb{C}([0,m))$ there exists $\epsilon > 0$ so that for any $h \in \mathbb{C}([0,m))$ with $\|h\|^2 < \epsilon$ we have $\mathbf{S}(f,g+h)) \subset \mathbf{S}(f,g)$.
3) For any $h,g \in \mathbb{C}([0,m))$ there exists $\epsilon > 0$ so that for any c with $c > \epsilon$ we have $\mathbf{S}(f,cg+h)) \subset \mathbf{S}(f,g)$.

Proof: We leave it as excercise to the reader. ∎

Corollary 1.4:

For any pair $f,g \in \mathbb{R}([0,m))$ and $k \in [0,m)$ there exists $\epsilon_1 > 0$ (respectively, $\epsilon_2 < 0$) so that for any $r_1 > \epsilon_1$ (resp. $r_2 < \epsilon_2$) for each $s \in \mathbf{S}(f,g+r_1\delta_k)$, $s+k$ is a maximum (resp. for each $s \in \mathbf{S}(f,g+r_2\delta_k)$, $s+k$ is a minimum) of f.

Proof: We leave it as excercise to the reader. ∎

Thus, if f is a generic function and g is 'sufficiently spiked' up (resp. down) at k then $\mathbf{S}(f,g)$ is contained in the set of shifts from k to $\arg\max_{t\in T_m}(f)$ (resp. $\arg\min_{t\in T_m}(f)$) .

Definition 1.5: A pair $f,g \in \mathbb{C}([0,m))$ is ψ-stable if $\#\mathbf{S}(f,g) = 1$.

With $\|h\|^2$ interpreted as the total energy of h item 1) in the theorem above asserts that a generic pair f,g, is ψ-stable. Further, from 2), if f,g, is ψ-stable then it remains such, with unchanged $\mathbf{S}(f,g)$ after perturbations of either argument by a function h of sufficeintly small energy.Thus, ψ-stable pairs form an open dense subset of $\mathbb{C}([0,m))\times\mathbb{C}([0,m))$. The third assertion of the theorem gives a prescription for increasing the energy of g so that the ψ- stability is preserved for an arbitrary h.

Definition 1.6: A pair $f,g \in \mathbb{C}([0,m))$ is ψ-positive if for $s \in \mathbf{S}(f,g)$: $0 < \psi(f,g)(s)$.

It follows that, for any pair $f,g \in \mathbb{C}([0,m))$ with $f \neq 0$ there exists constant $C > 0$ such that $f+c, g+c$ is ψ-positive for $c > C$. Both notions will be needed in Sections III.B and IV.B.

C. Monotonicity

For the reminder of this section we will restrict to the subset $\mathbb{R}([0,m)) \subset \mathbb{C}([0,m))$ of real valued functions on $[0,m)$ and write $\mathbb{R}([0,m))^+$, $\mathbb{R}([0,m))^{++}$ (respectively $\mathbb{R}([0,m))^-$, $\mathbb{R}([0,m))^{--}$) for the subset of real valued, non-decreasing, strictly increasing (respectively, non increasing, strictly decreasing) functions.

Theorem 1.7:

1) If $f,g \in \mathbb{R}([0,m))^{++}$ or $f,g \in \mathbb{R}([0,m))^{--}$ then $\mathbf{S} = \{0\}$.
2) If $f \in \mathbb{R}([0,m))^{++}$ and $g \in \mathbb{R}([0,m))^{--}$ then $0 \notin \mathbf{S}$.

Proof: For (1), denote $f = (f_0,f_1,\ldots,f_{m-1}), g = (g_0,g_1,\ldots,g_{m-1})$. We will consider the case $f_i < f_{i+1}, g_i < g_{i+1}$. Using induction on m we will show that the cross-correlation function $\psi(s) =< f, s(g) >$ has a maximum at $s = 0$. From Theorem 1 (6) , we loose no generality by setting $f_0 = 0$. First consider the case $m = 2$. Then $< f,g >= f_1g_1 < f_1g_0 =< f,1(g) >$. Recall that $1(g)$ denotes the circular 1-shift of g. Our inductive hypothesis is that the theorem holds for $m = n$. For $m = n+1$, if $s = 1$ then, since for $i > 0$ we have $f_ig_{i-1} < f_ig_i$ hence $< f,1(g) >= f_1g_0 + \ldots + f_ng_{n-1} < f_1g_1 + \ldots f_ng_n =< f,g >$. For $s = k$, $1 < k \leq n$ we have $< f,k(g) >= f_1g_{n-k+2} + \ldots + f_{k-1}g_n + f_kg_0 + \ldots + f_ng_{n-k} < f_1g_0 + \ldots + f_{n-k+1}g_{n-k} + f_{n-k+2}g_{n-k+2} + \ldots + f_ng_n < f_1g_1 + \ldots + f_{n-k+1}g_{n-k+1} + f_{n-k+2}g_{n-k+2} + \ldots + f_ng_n =< f,g >$ where the first inequality is the inductive hypothesis and the second follows from $f_ig_{i-1} < f_ig_i$.

For (2), we lose no generality by assuming that $f_0 = 0, f_i < f_{i+1}$ and $g_i > g_{i+1}$ in which case we have $< f,1(g) >= f_1g_0 + \ldots + f_{m-1}g_{m-2} > f_1g_1 + \ldots + f_{m-1}g_{m-1} =< f,g >$. ∎

Closer examination of this proof shows that the theorem holds under weaker assumption that $f \in \mathbb{R}([0,m))^+$, f is not constant and g is strictly monotonic. We leave this as an excercise to the reader.

Together, both theorems give a rule of thumb for computing $\mathbf{S}(f,g)$ when $m = 2$.

Corollary 1.8:

If $m = 2$ then:

1) $\mathbf{S} = \{0,1\}$ iff at least one of f,g is a constant
2) $\mathbf{S} = \{0\}$ iff f,g are both strictly monotonic in the same direction
3) $\mathbf{S} = \{1\}$ iff f,g are both strictly monotonic in opposite directions

Suppose that for a function $f \in \mathbb{R}([0,m))$ there exists $t \in T_m$ for which f^t is monotonic. We could call such functions 'shift-monotonic'. Notice that if f is shift-monotonic then, unless it is a constant function, it is either shift-decreasing or shift-increasing and the t for which f^t is monotonic is unique. We notice that the theorem above has a natural reformulation for a shift monotonic pair f, g, since knowing $\mathbf{S}(f^t, g^{t'})$ (with $f^t, g^{t'}$ monotonic) determines $\mathbf{S}(f, g)$, via identity given by 1.2.4 . For $m = 3$ we have the following.

Lemma 1.9: Any $f \in \mathbb{R}([0,3))$ is shift-monotonic.

Proof: We leave ithis as excercise to the reader. ■

Thus, to compute $\mathbf{S}(f, g)$ for $m = 3$, one can assume that f is monotonic and g is shift-monotonic, in which case one can write down a set of rules similiar to Corollary 1.6. We leave details to the reader.

II. Sampling equation

In this section we write out an equation relating $\psi(f^*, g^*)$ to $\psi(f, g)$ where f^*, g^* denote restrictions of f, g to a sub-interval $[0, m') \subset [0, m)$ for $m' < m$.

Definition 2.1:

Given $a, b \in [0, m)$ let $\sigma_{a,b} \in T_m$ denotes the unique circular permutation for which $\sigma(a) = b$. Let $T_{a,b} = \{0, 1, \ldots, \sigma_{a,b}\} \subset T_m$. Denote $[a, b] = a + T_{a,b} = \cup_{t \in T_{a,b}} a^t$, $[a, b) = [a, b] - \{b\}$, $(a, b] = [a, b] - \{a\}$. Then $[a, b] \subset [0, m)$ will be called closed interval (from a to b).

With this definition, an interval is always an ordered, proper and, if closed, a nonempty subset of $[0, m)$. Moreover:

1) For $s \in T_m$: $[a, b]^s = [a + s, b + s]$
2) $\#[a, b) = \begin{cases} b - a \text{ if } a < b \\ 0 \text{ if } a = b \\ m - \#[b, a) = m + (b - a) \text{ if } a > b \end{cases}$
3) $[a, b]^\tau = [-b, -a]$
4) $[0, m) - [a, b) = [b, a)$

If we view $[0, m)$ as the subset $\{\mu_m^0, \mu_m^1, \ldots, \mu_m^{m-1}\}$, with $\mu_m = \exp(2\pi i/m)$, of the unit circle, then 4) reflects that a pair of points $a, b \in [0, m)$ determines two intervals $[a, b)$, $[b, a)$ realized by the two complementary arcs of the circle with clock-wise orientation.

Given $f \in \mathbb{C}([0, m))$ and an interval $i = [a, b) \subset [0, m)$ define $i^*(f)$ to be the restriction of $f \in \mathbb{C}([0, m))$ to the interval $[b, a) = [0, m) - [a, b)$ and set $r = \#[a, b)$. To simplify the notation we will write $i^*(f)$ as f^*.

Theorem 2.2:

For $s \in T_{b,a}$ and $f, g \in \mathbb{C}([0, m))$ we have:

$$(f^*)^s = (\delta_{[b,a-s)} g)^s)^* + (\delta_{[a-s,a)} g)^{s+r})^* \tag{II.1}$$

$$\psi(f^*, g^*)(s) = \psi(f, \delta_{[b,a-s)} g)(s) + \psi(f, \delta_{[a-s,a)} g)(s + r) \tag{II.2}$$

Proof:

Since both sides of the first equation are linear with respect to g and $g = \sum_{t=0}^{t=m-1} g_t \delta_t$ hence it suffices to check it on the basis $g_t = \delta_t$, $t \in [0, m)$ of $\mathbb{C}([0, m))$. We leave this as an excercise to the reader. We note that for $t \in [b, a)$ and $g = \delta_t$, the left hand side of the first equation can be interpreted as an equation of motion $\delta_t \to \delta_t^s$ on $[b, a)$ induced by circular permutation s, with the right hand side beeing the coresponding motion in $[0, m)$. The second equation follows from:

$$< f^*, (g^*)^s >=< f, (\delta_{[b,a-s)} g)^s > + < f, (\delta_{[a-s,a)} g)^{s+r} > \tag{II.3}$$

, which is a consequence of the first equation.

■

The equation II.2 is a starting point for proving Theorem 4.4 [8].

III. Applications of convolution

The results in subsections A and C are well known. Since they seem scatterred accross the literature [4],[7],[5] we present them here for the convienience of the reader.

A. Convolution and cross-convolution

In this section we present an alternative definition of function $\psi(f, g)$.

Let $*$ denotes the (circular) convolution operator [6] on $\mathbb{C}(T_m)$. It is well known that $*$ is commutative, associative and has the identity δ_0.

Theorem 3.1:

1) $\overline{f * g} = \overline{f} * \overline{g}$
2) $f^s = \delta_s * f$
3) $\tau(f * g) = \tau(f) * \tau(g)$
4) $(f * g)^s = (f^s) * g = f * (g^s)$

Proof: For 3), use: $\tau(f * g)(t) = (f * g)(-t) = \sum_{t'} f_{-t-t'} g_s = \sum_{t'} \tau(f)_{t+t'} \tau(g)_{-t'} = \tau(f) * \tau(g)(t)$. For 4), use $(f * g)^s = \delta_s * (f * g) = (\delta_s * f) * g = (f^s) * g$ and thus $(f * g)^s = (g * f)^s = f * (g^s)$. We leave the details to the reader. ■

Theorem 3.2:

$\psi(f, g) = f * \tau(\overline{g})$.

Proof: $(f * \tau(\overline{g}))(t) = \sum_{t'} f_{t'} \overline{\tau(g)}_{-t'+t} = \sum_{t'} f_{t'} \overline{g}_{t'-t} =< f, g^t >$ ■

Corollary 3.3: $\psi(f * f', g * g') = \psi(f, g) * \psi(f', g')$.

Proof: We have: $(f * f') * \tau(\overline{(g * g')}) = (f * f') * (\tau(\overline{g}) * \tau(\overline{g'})) = f * \tau(\overline{g}) * f' * \tau(\overline{g'}) = (f * \tau(\overline{g})) * (f' * \tau(\overline{g'}))$. ■

Definition 3.4: For $f, g \in \mathbb{C}(T_m)$ define $f \hat{*} g = f * \tau(\overline{g})$. The operator $\hat{*}$ will be called cross-convolution.

Corollary 3.5:

1) $\delta_s \hat{*} f = \tau(\overline{f})^s = \tau(\overline{f^{-s}}) = \tau(\delta_{-s} * \overline{f})$
2) $(f \hat{*} g)^s = f^s \hat{*} g = f * \tau(\overline{g})^s = f * \tau(\overline{g^{-s}}) = f \hat{*} g^{-s}$
3) $(f * f') \hat{*} (g * g') = (f \hat{*} g) * (f' \hat{*} g')$
4) $(f \hat{*} f') * (g \hat{*} g') = (f * g) \hat{*} (f' * g')$
5) $\psi(f^s, g) = \psi(f, g)^s$, $\psi(f, g^s) = \psi(f, g)^{-s}$, $\psi(f,^s g^s) = \psi(f, g)$
6) $\psi(f * f', g * g') = \psi(f, g) * \psi(f', g')$

Proof: All these identities can be formally computed from the definitions by using associativity, commutativity of $*$ and Theorem 3.2 . We leave the details to the reader. ■

Definition 3.6: Given $S \subset T_m$, let δ_S denotes the characteristic function of S: $\delta_S(s) = 1$ for $s \in S$ and $\delta_S(t) = 1$ for $t \notin S$.

We can extend the definitions of $*$, $\hat{*}$ to sets: for $S, S' \subset T_m$ define $S * S' = supp(\delta_S * \delta_S)$, $S \hat{*} S' = supp(\delta_S \hat{*} \delta_S)$ where

$supp(f) = \{s : f(s) \neq 0\}$. Clearly $S * S' = \cup_{s\in S, s'\in S'}\{s + s'\}$, $S \hat{*} S' = \cup_{s\in S, s'\in S'}\{s - s'\}$.

B. Mesh refinements

In this section we prove that for generic, nonnegative f, g, $\mathbf{S}(f, g)$ stays essentially unchanged when the mesh $[0, m)$ gets refined to $[0, ml)$ by adding an equal number l of points between each consecutive pair $\{k, k+1\}$ and $g, f \in Map([0, m), C)$ are extended to $[0, ml)$ using the customary 'as of the last available measurement' rule.

Suppose that $m = m_1 m_2$ with $m_1, m_2 > 1$. Then, by means of the canonical injection $[0, m_1) \xrightarrow{\mathrm{i}} [0, m)$ given by $i(k) = m_2 k$, $[0, m]$ can be considered as the m_2-subdivision of $[0, m_1]$. With the identifications from Section I , i is an injective homomorphism of abelian groups $T_{m_1} \subset T_m$ with the subset on the left given by the equation $t = 0 \mod m_2$.

Given $f \in \mathbb{C}([0, m_1))$, consider function $i_*(f) \in \mathbb{C}([0, m))$ given by

$$i_*(f)(t) = \begin{cases} f(t/m_2) \text{ if } t = 0 \mod m_2 \\ 0 \text{ otherwise} \end{cases}.$$

To simplify notation, we will write f_* for $i_*(f)$. It is easy to see that

$$\psi(f_*, g_*) = \psi(f, g)_* \tag{III.1}$$

Observe that if f is a time series defined on mesh $[0, m_1)$ then f_* can be interpreted as this time series on the refined mesh with gaps (zero data points) for $t \in [0, m_1 m_2)$, $t \neq 0 \mod m_2$, while $f_* * \delta_{[0,m_2)}$ as this time series obtained from f_* by filling in the missing values with the 'most recent value' - the usual choice for time series of prices with gaps in the data. For example, if f represents quotes of a price on mesh $[0, m_1)$ then $f * \delta_{[0,m_2)}$ represents the prices on the m_2 subdivison $[0, m_1 m_2)$ of $[0, m_1)$, where the new intermediate price values are the most recent available prices. Recall Definition 1.4

Theorem 3.7: Supose for $f, g \in \mathbb{C}([0, m_1))$ is ψ-stable (cf. Definitin 1.4: $\#\mathbf{S}(f, g) = 1$). Then $\mathbf{S}(f_* * \delta_{[0,m_2)}, g_* * \delta_{[0,m_2)}) = \mathbf{S}(f_*, g_*)$.

Proof: From 3.5.7, $\psi(f_* * \delta_{[0,m_2)}, g_* * \delta_{[0,m_2)}) = \psi(f_*, g_*) * \psi(\delta_{[0,m_2)}, \delta_{[0,m_2)}) = \psi(f, g)_* * \phi$ where $\phi = \psi(\delta_{[0,m_2)}, \delta_{[0,m_2)})$. It is easy to see that ϕ has support $T = T_{m_2} \hat{*} T_{m_2} = \{-m_2 + 1, \ldots, 0, \ldots, m_2 - 1\}$ and is given by formulae $\phi(u) = \begin{cases} m_2 - |u| \text{ if } u \in T \\ 0 \text{ otherwise} \end{cases}$. Since $\phi^\tau = \phi$ hence we have $\psi(f, g)_* * \phi = \psi(f, g)_* \hat{*} \phi = \psi(\psi(f, g)_*, \phi)$. Denote $\psi_u = \psi(f, g)(u)$. Any $t \in [0, m)$ can be uniquely written as $t = sm_2 + k$ with $s \in [0, m_1)$ and $k \in [0, m2)$, and $(\psi(f_*, g_*), \phi)(sm_2 + k) = \phi(-k)\psi_s + \phi(m_2 - k)\psi_{s+1} = (m_2 - k)\psi_s + k\psi_{s+1}$. The proof is completed using identity:

$$\arg \max_{(s,k)\in[0,m_1)\times[0,m2)} ((m_2 - k)\psi_s + k\psi_{s+1}) = (\arg \max_{s\in[0,m_1)} (\psi_s), 0) \tag{III.2}$$

which holds for an arbitrary sequence $\psi_0, \psi_1, \ldots, \psi_{m_1-1}$ of real numbers with a unique maximum. ∎

Example 3.8: Consider the example of identity functions $f = g = \delta_{[0,m_1)}$. Then $\psi(f, g) = \delta_{[0,m_1)}$, $\#\mathbf{S}(f, g) = m_1$, $\mathbf{S}(f_*, g_*) = i([0, m_1))$ and, since $f_* * \delta_{[0,m_2)} = \delta_{[0,m)}$, hence $\mathbf{S}(f_* * \delta_{[0,m_2)}, g_* * \delta_{[0,m_2)}) = [0, m)$. This shows that the assumption $\#\mathbf{S}(f, g) = 1$ in the theorem above is essential.

Definition 3.9: We will say that pair $f, g \in \mathbb{C}([0, m_1))$ is ψ-convex, if it is ψ-stable and ψ-positive (cf. 1.5, 1.6). This is equivalent to : $\mathbf{S}(f, g) = \{s\}$ with $0 < \psi(f, g)(s)$.

For example, any pair f, g representing prices is ψ-convex.

Lemma 3.10: Suppose $f, g \in \mathbb{C}([0, m_1))$ are ψ-positive. Then $\mathbf{S}(f_*, g_*) = i(\mathbf{S}(f, g))$. If $f, g \in \mathbb{C}([0, m_1))$ is ψ-stable then so is f_*, g_*.

Proof: Follows from the identity III.1. ∎

Notice that for any pair of functions f, g, with $f(t) > 0$ and $g(t) < 0$ for all $t \in [0, m_1)$, the lemma fails and $\mathbf{S}(f_*, g_*) = T_m - i(T_{M_1})$. Thus, the assumption of ψ-positivity is essential.

Putting together the theorem and the lemma gives the following result.

Theorem 3.11: Suppose $f, g \in \mathbb{C}([0, m_1))$ is ψ-convex. Then so is $f_* * \delta_{[0,m_2)}, g_* * \delta_{[0,m_2)} \in \mathbb{C}([0, m))$ and $\mathbf{S}(f_* * \delta_{[0,m_2)}, g_* * \delta_{[0,m_2)}) = i(S(f, g))$.

It can be shown that both, ψ-positivity and stability, are essencial assumptions. Of course, if f, g represent prices then, by perturbing them with arbitrary small energy functions f', g' if necessary, one has $\#\mathbf{S}(f + f', g + g') = 1$ and the theorem applies.

C. Discrete Fourier Transfer and computability

Note that the computation of $S(f, g)$ using the definition given in Section 1 is $O(m^2)$. Corollary 3.13 [4],[5] gives much more efficient method for computation of $S(f, g)$ using the Fast Fourier Transform.

Given $g \in \mathbb{C}(T_m)$, denote $g_i = g(\nu_m^i), G_i =< g, \chi_i >, G_i^{-1} =< g, \chi_{n-i} >$ and $F(g)(i) = G_i, F^{-1}(g)(i) = G_i^{-1}$. Then the operator $\mathrm{F} : \mathbb{C}(T_m) \to \mathbb{C}(T_m)$ given by $g \to F(g)$ is the well known Discrete Fourier Transform DFT with inverse $\frac{1}{m}F^{-1}$. For any f, g we have:

1) $\tau(\chi_t) = \overline{\chi_t} = \chi_{-t}$
2) $\mathrm{F}(\tau(f)) = \tau(\mathrm{F}(f)) = \mathrm{F}^{-1}(f)$
3) $\tau(\mathrm{F}(\overline{f})) = \overline{\mathrm{F}}(f)$ and $\tau(\mathrm{F}^{-1}(\overline{f})) = \overline{\mathrm{F}^{-1}}(f)$
4) $\mathrm{F}^{-1}(\tau(f)) = \tau(\mathrm{F}^{-1}(f)) = \mathrm{F}(f)$
5) $\mathrm{F}^{-1}(\mathrm{F}(f)) = \mathrm{F}(\mathrm{F}^{-1}(f)) = mf$
6) $\mathrm{F}(fg) = \mathrm{F}(f) * \mathrm{F}(g)$ and $\mathrm{F}^{-1}(fg) = \mathrm{F}^{-1}(f) * \mathrm{F}^{-1}(g)$
7) $\mathrm{F}(f * g) = \mathrm{F}(f)\mathrm{F}(g)$ and $\mathrm{F}^{-1}(f * g) = \mathrm{F}^{-1}(f)\mathrm{F}^{-1}(g)$
8) $F(\delta_t) = \chi_t$

Theorem 3.12:
$\psi(f, g) = m^{-2} Re(\mathrm{F}^{-1}(\mathrm{F}(f)\overline{\mathrm{F}(g)}))$

Proof:
$\mathrm{F}^{-1}(\mathrm{F}(f)\overline{\mathrm{F}}(f)) = \mathrm{F}^{-1}((\mathrm{F}(f)) * \mathrm{F}^{-1}(\overline{\mathrm{F}}(f)) = m^2 f * \tau(\overline{g}) = f \hat{*} g$. Theorem 3.2 completes the proof. ∎

The key application is the computability of $S(f, g)$.

Corollary 3.13:
$S(f, g) = \arg \max_{t\in T_m}(\hat{\psi})$ where $\hat{\psi} = Re(\mathrm{F}^{-1}(\mathrm{F}(f)\overline{\mathrm{F}(g)}))$.

Corollary 3.14:
For $m = 2^n$, $S(f, g)$ is $O(mlog(m))$ computable.

Proof: It is well known that the Fast Fourier Transform algorithm FFT [2] computing F, F^{-1} is $O(mlog(m))$. Neither the expression for $\hat{\psi}$ nor search for $\arg\max_{t\in T_m}(\hat{\psi})$ increase the complexity. ■

IV. Periodic signals

A. Periodic extensions and restrictions

We have seen in Section 1 that for an arbitrary pair $f, g \in \mathbb{C}([0,m))$ the set $\mathbf{S}(f,g)$ is unstable under perturbations of either argument or the restriction to subintervals of T_m. In contrast, when $f, g \in \mathbb{C}([0,m))$ are sampled from a periodic signal with a common period l then Theorem 4.4 (the main result of this section) states that $\mathbf{S}(f,g)$ is quite stable. This subsection defines criterias for a function $f \in \mathbb{C}([0,m))$ to be a restriction of an l-periodic function, setting a context for Theorem 4.4.

For a pair of positive integers m, l denote $[m/l] = floor(m/l) = max\{k \in \mathbb{Z} : 0 \leq m - kl\}$. Define $\phi : T_m \to T_l$, $\phi(t) = t \mod l$ and the induced homomorphism of $\mathbb{C}$ vector spaces $\phi^* : \mathbb{C}(T_l) \to \mathbb{C}(T_m)$. We have $m = [m/l]l + k$ with $k = m \mod l$.

Definition 4.1: Given a positive integer l, $f \in \mathbb{C}([0,m))$ is an l-periodic extension with $[m/l]$ periods and the reminder $k = \phi(m) = m \mod l$ when there exists $\mathbf{f} \in \mathbb{C}([0,l))$ such that $f = \phi^*(\mathbf{f})$. We will call l an extended period of f.

Let $P_l(m) \subset \mathbb{C}([0,m))$ denotes the set of of l-periodic extensions. Since $P_l(m)$ is closed under pointwise addition and multiplication, hence it is a $\mathbb{C}^*$-sub-algebra of $\mathbb{C}(T_m)$ [1].

Lemma 4.2: The following are equivalent:

1) $f \in \mathbb{C}(T_m)$ is l-periodic extension with $[m/l]$ periods and the reminder k
2) f can be extended to l-periodic function on T_{m+l-k}
3) restriction of f to any interval $[a,b] \subset T_m, 0 \leq a \leq b < m$ with $\#[a,b]$ divisible by l is l-periodic
4) f is a restriction of l-periodic $\mathbf{f} \in \mathbb{C}([0,[m/l]+k))$ to a sub-interval of length m

Proof: We leave this as excercise to the reader. ■

It follows that if l is the shortest extended period of f, then any extended period l' of f is divisible by l. In addition, if l divides m, then any l-periodic extension $f \in \mathbb{C}(T_m)$ is l-periodic with m/l of l-periods.

Example 4.3: Here is one more equivalent definition of $P_l(m)$. Given a natural number m let $\nu_m = e^{2\pi i/m}$ denote the primitive m'th root of unity. The set of characters $\chi_m^k \in \mathbb{C}(T_m)$, $\chi_m^k(t) = \frac{1}{\sqrt[2]{m}}\nu_m^{tk}$, $0 \leq k < m$ is an orthonormal basis of $\mathbb{C}(T_m)$ with respect to the inner product $<>$, $\mathbb{C}(T_m)$ is a $\mathbb{C}^*$ algebra generated by χ_m^1. . If $m = nl$ then it is well known that $\{\chi_m^{0n}, \chi_m^{1n}, \ldots \chi_m^{(l-1)n}\} = \{\chi_{ln}^{nk} : k \in [0,l)\} = \{\chi_l^k : k \in [0,l)\}$ is an orthonrmal basis over $\mathbb{C}$ of $P_l(m)$. In particular, the element χ_m^n generates $P_l(m)$ as $\mathbb{C}^*$ algebra. For an arbitrary m let $m' = nl$ with $m \leq m'$. Let i denotes the natural inclusion $[0,m) \subset [0,m']$. It follows that $P_l(m)$ is generated, as $\mathbb{C}^*$-algebra, by the element $i^*(\chi_{nl}^1)^n$ and $\{i^*(\chi_{nl}^1)^{nk} : k \in [0,l)\} = \{i^*(\chi_l^1)^k : k \in [0,l)\}$ is a basis (not orthonormal, unless $l|m$) over $\mathbb{C}$ of $P_l(m)$. Of course, while all $\chi_{m'}^{kn} = \chi_l^k$ are characters of $T_{m'}$ (group homomorphisms from $T_{m'}$ to $\mathbb{C}^*$), none of $i^*(\chi_l^k)$, except for $k = 0$, are characters of T_m, unless $l|m$.

B. Localization and sampling

In this subsection we will assume that, given $\hat{f}, \hat{g} \in \mathbb{C}(T_l)$, $f_n, g_n \in \mathbb{C}(T_{nl+k})$ are l-periodic extensions containing n fundamental l-periods $\hat{f}, \hat{g} \in \mathbb{C}(T_l)$ with a reminder k. Suppose for each n there is given an interval $[a_n, b_n) \subset T_{nl+k}$ with $\#[a_n, b_n)$ is uniformly bounded by a constant C. Denote $f^* = (1 - \delta_{[a_n,b_n)})f_n$, $g^* = (1 - \delta_{[a_n,b_n)})g_n$. Since $1 - \delta_{[a_n,b_n)} = \delta_{[b_n,a_n)}$, hence f^*, g^* are localizations of f_n, g_n to the interval $[b_n, a_n)$. The following result is a stability theorem for the functor $(\hat{f}, \hat{g}) \to \mathbf{S}(f^*, g^*)$.

Theorem 4.4: Suppose that $\mathbf{S}(\hat{f}, \hat{g}) = \{s\}$ with $0 < \psi(\hat{f}, \hat{g})(s)$ (equivalent to $(\hat{f}, \hat{g})$ being ψ-convex). Then for sufficciently large n:

1) If $s = 0$ then $\{0\} \subset \mathbf{S}(f^*, g^*)$.
2) If $s \neq 0$ then $\mathbf{S}(f^*, g^*) \cap \{s, s-l\} \neq \emptyset$

Moreover, after an arbitrarily small perturbation of f^*, g^*:

1) If $s = 0$ then $\mathbf{S}(f^*, g^*) = \{0\}$.
2) If $s \neq 0$ then $\#(\mathbf{S}(f^*, g^*) \cap \{s, s-l\}) = 1$

These formulaes do not contain the size (given by the number of periods n and the reminder k) of data window or the size of intervals $[a_n, b_n)$ as variable. They hold,in fact, with $[a_n, b_n)$ replased by a union of intervals with sum of cardinalities bounded by C. Authors believe that this indicates a potential for practical applications. The proof of the theorem, while completed using the framework presented here, is too lenghty for this article and can be found in [8].

References

[1] Arveson, W. (1976), An Invitation to C*-Algebra, Springer-Verlag, ISBN 0-387-90176-0.

[2] Brigham, E. Oran (2002). "The Fast Fourier Transform". New York, USA: Prentice-Hall.

[3] Humphreys, John F. (1996). A Course in Group Theory. Oxford University Press. p. 195. ISBN 9780198534594.

[4] https://en.wikipedia.org/wiki/Cross-correlation

[5] Press, William H. and Flannery, Brian P. and Teukolsky, Saul A. and Vetterling, William T., Numerical Recipes in C: The Art of Scientific Computing 1988, 0-521-35465-X, Cambridge University Press, New York, NY, USA. p.433.

[6] Hewitt, Edwin; Ross, Kenneth A. (1979), Abstract harmonic analysis. Vol. I, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 115 (2nd ed.), Berlin, New York: Springer-Verlag, ISBN 978-3-540-09434-0, MR 0551496

[7] Weisstein, Eric W. "Cross-Correlation." From MathWorld–A Wolfram Web Resource.

[8] Zelewski, Piotr and Voevudsky, Valery "Shifts between two periodic time series II", preprint 2018.