THE FAST WALSH-TRANSFORMATION

By

T. FREY and T. NAGYPÁL

Department of Mathematics of the Faculty of Electrical Engineering,
Technical University, Budapest

Received June 20, 1978

1. It is a very important and not yet satisfactorily solved problem to classify a process into a finite set of subclasses on the basis of the time-evolution of some observable parameters. The selected finite set of subclasses represents either some essential differences. between the physical system components of the process: or some essential variation in the structure of interconnections, between the system components.

The classification has to be based on the functional description of the time-evolution of the observable parameters. A theoretically complete treatment of this problem can be established as follows: let us construct a complete metric functional space, containing all the possible evolution functions of the measured parameters, and the probability measures over this space, characterizing the distribution of the evolution functions of the cumulated class as well as the one of the individual subclasses. However, in this case a functional space of infinite dimensional must be worked in, where there is no effective numerical method. Even therefore we propose to transpose the entire analysis into a finite dimensional space, where the possible evolution functions may be characterized by a finite set of highly relevant parameters. It is well known that this problem can be solved by the help of projecting the evolution functions onto a predefined finite dimensional space. The most effective choice of the latter is the space, spanned by the orthonormal set of functions, defined by the *Karhunen—Loeve* expansion. However, this set is not easy to manage numerically, moreover it cannot be done if the probability degree is not known at least up to the second moments. In general we have no preliminary knowledge of it. Even therefore the usual practice in this case is based on a finite dimensional *Hilbert*-space, spanned by an optionally chosen orthonormal system, hoping that the projection of the evolution functions onto this space yet contains enough information to assure an efficient partition into the set of subclasses.

In the light of the above heuristics, any orthonormal system is competitive, and even therefore we may choose the one needing the least numerical effort. It will be demonstrated in the following that this is assured by the Walsh-system.

2. The orthonormal Walsh-system [1] is defined over (0,1) as follows:

$$w_0(t) \equiv 1, \quad 0 \leq t \leq 1;$$

$$w_{2^n}(t) = \operatorname{sign}(\sin 2^{n+1} t\pi), \quad 0 \leq t \leq 1, \quad n = 0, 1, 2 \ldots$$

$$w_k(t) = \prod_{j=j_1(k)}^{j=j_2(k)} w_{2^j}(t), \text{ if } k = 2^{j_1} + \ldots + 2^{j_2}, \quad 0 < t < 1;$$
$$k = 0, 1, 2, \ldots; 0 \leq j_1 < \ldots < j_2.$$

This system is known [2] to be complete in L^2 (0, 1), to have, however, no good pointwise convergence properties. But in the case where only a well defined part sequence of the Walsh-expansion is considered, the pointwise convergence behaviour will be extremely good.

Theorem 1. Let $f(t) \in C[0, 1]$, and denote $w_k(t; f)$ the k-th cut of the Walsh-expansion of $f(t)$:

$$w_k(t, f) = \sum_{j=0}^{k} \varphi_j w_j(t); \; \varphi_j = \int_0^1 w_j(z) \cdot f(z) dz.$$

Then $w_{2^n-1}(t, f)$ converges uniformly to f in (0, 1).

Proof: the convergence properties of the expansion are determined by the kernel:

$$W_k(t, z) = \sum_{j=0}^{k} w_j(t) \cdot w_j(z).$$

Let k be an integer between 2^{n-1} and 2^n, i.e. $2^{n-1} \leqslant k < 2^n$. Then, by the definition of $w_j(t)$, the kernel is an integer constant on each cell $C_n(l, m) = (l-1) \cdot 2^{-n+1} < t < l \cdot 2^{-n+1} < z < m \cdot 2^{-n+1}$; $l, m = 1, 2, 3, \ldots$ We shall prove by induction that if $k = 2^n - 1$, then

$$W_{2^n-1}(t, z) = \begin{cases} 2^n, \text{ on } C_n(l, 1); \; l = 1, 2, 3, \ldots \\ 0, \text{ on } C_n(l, m); \; l \neq m \end{cases}$$

This assertion is evidently true for $n = 0$ and $n = 1$. Let us suppose that our assertion is true for $n = N$, and consider the case $n = N + 1$. Now

$$W_{2^{N+1}-1}(t, z) = W_{2^N-1}(t, z) + \sum_{j=2^N}^{2^{N+1}-1} w_j(t) \cdot w_j(z).$$

In the diagonal cells $C_{N+1}(l, l)$ the factors $w_j(t)$ and $w_j(z)$ have the same sign for all $j = 2^N, 2^N + 1, \ldots, 2^{N+1} - 1$. Therefore by the induction hypothesis we have here

$$W_{2^{N+1}-1}(t, z) = 2^N + \sum (\pm 1) \cdot (\pm 1) = 2^N + 2^N = 2^{N+1}$$

For any off-diagonal cell $C_{N+1}(l, m)$ with $|l - m| > 1$ and for any j, $2^N < j \leqslant 2^{N+1} - 1$ we have

$$w_j(t) = w_{2^N}(t) \cdot \prod_{p=p_1(j)}^{p_2(j)} w_{2^{p_1(j)}}(t) \cdot \ldots \cdot w_{2^{p_2(j)}}(t),$$

$$w_j(z) = w_{2^N}(z) \cdot \prod_{p=p_1(j)}^{p_2(j)} w_{2^{p_1(j)}}(z) \cdot \ldots \cdot w_{2^{p_2(j)}}(z),$$

with $0 \leqslant p_1(j) < \ldots < p_2(j) < N + 1$. Therefore

$$\sum_{j=2^N}^{2^{N+1}-1} w_j(t) \cdot w_j(z) =$$

$$= w_{2^N}(t)\, w_{2^N}(z) \cdot \sum_{j=2^N}^{2^{N+1}-1} \left[\prod_{p=p_1(j)}^{p_2(j)} w_{2^{p_1(j)}}(t) \cdot w_{2^{p_1(j)}}(z) \cdot \ldots \cdot w_{2^{p_2(j)}}(z) \right] =$$

$$= w_{2^N}(t) \cdot w_{2^N}(z) \cdot W_{2^N-1}(t, z),$$

because the terms in the sum are the same as in $W_{2^N-1}(t, z)$, surely each combination of the exponents occurs exactly once in both. But $(l - m) > 1$ assures that (t, z) is lying in some off-diagonal cell $C_N\left(\left[\frac{l}{2}\right], \left[\frac{m}{2}\right]\right)$ with respect to N, and therefore

$$W_{2^{N+1}-1}(t, z) = W_{2^N-1}(t, z) \cdot [1 + w_{2^N}(t) \cdot w_{2^N}(z)] = 0 \cdot [l + (\pm l)] = 0.$$

Finally for $(l - m) = 1$, we have either $l = (m - 1)$ or $l = m + 1$. In the former case, (t, z) is lying either in an off-diagonal cell $C_N\left(\left[\frac{l}{2}\right], \left[\frac{m}{2}\right]\right)$ – namely if and only if $l = 2s$ – or in a diagonal cell, namely if and only if $l = 2s + 1$. In the case where $l = 2s$,

$$W_{2^{N+1}-1}(t, z) = 0 \cdot [1 + (-1)] = 0 \cdot 0 = 0,$$

however, for $l = 2s + 1$, hence for (t, z) is lying in a diagonal cell with respect to N:

$$W_{2^{N+1}-1}(t, z) = W_{2^N-1}(t, z) \cdot [1 + (\pm 1) \cdot (\mp 1)] = 2^N \cdot (1 - 1) = 0.$$

In the letter case, for $l = m + 1$, and for $l = 2s$, (t, τ) is lying in a diagonal cell with respect to N, and therefore

$$W_{2^{N+1}-1}(t, z) = 2^N \cdot (1 - 1) = 0,$$

and otherwise $l = m + 1, l = 2s + 1$, the point (t, z) is lying in an off-diagonal cell with respect to N, hence

$$W_{2^{N+1}-1}(t, z) = 0 \cdot [1 + (\pm 1)] = 0.$$

With the above we have fully proven the induction hypothesis over the kernel.

Our theorem is now easy to prove. If namely, $f(t) \in C[0, 1]$ is valid, then we can give each fixed $\epsilon > 0$ bound $\Delta(\varepsilon)$ for the displacement $|t_2 - t_1|$ with

$$|f(t_2) - f(t_1)| < \varepsilon \text{ if } |t_2 - t_1| < \Delta\varepsilon.$$

In addition, because of orthonormality, we have for each fixed x, with $\varphi(t) \equiv f(x)$ the relation

$$f(x) = \int_0^1 \varphi(\tau) W_{2^n-1}(x, \tau)\, d\tau.$$

With the above:

$$|f(x) - w_{2^n-1}(x; f)| = \left| \int_0^1 \varphi(\tau)\, W_{2^n-1}(x, \tau)\, d\tau - \int_0^1 f(\tau)\, W_{2^n-1}(x, \tau)\, d\tau \right| =$$

$$= \left| \int_0^1 [\varphi(\tau) - f(\tau)]\, W_{2^n-1}(x, \tau)\, d\tau \right| \leqslant \int_0^1 |\varphi(\tau) - f(\tau)|\, W_{2^n-1}(x, \tau)\, d\tau =$$

$$= \int_{x-2^{-n}}^{x+2^{-n}} |\varphi(\tau) - f(\tau)|$$

$$W_{2^n-1}(x, \tau)\, d\tau + \int_0^{x-2^{-n}} |\varphi(\tau) - f(\tau)|\, W_{2^n-1}(x, \tau)\, d\tau +$$

$$+ \int_{x+2^{-n}}^1 |\varphi(\tau) - f(\tau)|\, W_{2^n-1}(x, \tau)\, d\tau \leq \varepsilon \cdot 2^{-n} \cdot 2^n + C \cdot 0 = \varepsilon,$$

if $2^{-n} < \Delta(\varepsilon)$, which proves our assertion, because ε is arbitrarily small.

Let us remark that the proof demonstrates, with respect to the speed of convergency, that $|f(t) - w_{2^n-1}(t; f)| \leqslant \omega(2^{-n}; f)$, where ω denotes the modulus of continuity of f.

3. Concerning the fast version of the Walsh expansion, i.e., the fast calculation of the value of such an expansion, let us first consider the expansion

$$f(t) = \sum_{j=0}^{2^n-1} \varphi_j w_j(t),$$

which defines a piecewise constant step function, with

$$f(t) \equiv f(t_k) \quad \text{if} \quad t \in [(k-1)\, 2^{-n+1},\ k \cdot 2^{-n+1}],$$

$$t_k = \left(k - \frac{1}{2}\right) \cdot 2^{-n+1}; \quad k = 1, 2, \ldots 2^{n-1},$$

and the expansion coefficient

$$\varphi_j = \int_0^1 f(t) \cdot w_j(t)\, dt = 2^{-n+1} \sum_{k=1}^{2^n-1} f(t_k) \cdot w_j(t_k).$$

Theorem 2. Both $f(t)$ and φ_j are linear combinations of the φ_k and the $f(t_k)$, resp., with coefficients ± 1. The coefficient of φ_j in the evaluation of $f(t_k)$ is the same as that of $f(t_j)$ in the evaluation of φ_k, disregarding the factor 2^{-n+1}.

Proof: The assertion about the coefficients is evident, because $|w_j(t_k)| \equiv \equiv 1$, for all j and k. We must only demonstrate that the signs are the same, namely that

$$w_j(t_k) \equiv w_k(t_j); \quad j, k = 0, 1, \ldots, 2^{-n+1}.$$

The assertion is trivially true for $n = 0$ and $n = 1$, and also for $n = 2$, because for the latter

$$w_0(t_1) = w_0(t_2) = w_0(t_3) = 1 = w_1(t_0) = w_2(t_0) = w_3(t_0);$$

$$w_1(t_2) = w_1(t_3) = -1 = w_2(t_1) = w_3(t_1);$$

$$w_2(t_3) = w_3(t_2) = 1.$$

Let us now suppose that the assertion is true for $n = N$, and consider the case $n = N+1$. For the subscript $j < 2^N$, $w_j(t)$ is the same function for $n = N$ and $n = N + 1$, but the base point $t_k^{(N+1)}$ coincides with no base point $t_k^{(N)}$. However, for any even k and odd $k + 1$, the base points $t_k^{(N+1)}$ and $t_{k+1}^{(N+1)}$ embrace the base point $t_{k/2}^{N}$ and both are lying on the interval $(t_{k/2}^{(N)} - 2^{-N-1}, t_{k/2}^{(N)} + 2^{-N-1})$; thus $w_j^{(N)}(t_k^{(N+1)}) = w_j^{(N)}(t_{k+1}^{(N+1)}) = w_j^{(N)}(t_{k/2}^{(N)})$.

By our assertions, for any $j < 2^N$, and any even $2 \cdot k$ we have

$$w_j^{(N+1)}(t_{2k}^{(N+1)}) = w_j^{(N)}(t_k^{(N)}) = w_k^{(N)}(t_j^{(N)}) = w_k^{(N+1)}(t_{2j}^{(N+1)}).$$

To prove our assertion for the considered case, we have only to prove that

$$w_k^{(N+1)}(t_{2j}^{(N+1)}) = w_{2k}^{(N+1)}(t_j^{(N+1)})$$

is valid. However, latter is valid if and only if it is valid for any $k = 2^r$, because by the definition of the *Walsh*-functions, w_k is generally the product of such *Rademacher—Walsh* functions. Let us now suppose that $k = 2^r$, hence $2k = 2^{r+1}$. Now if t_j is lying on such an interval, where $w_{2^{r+1}}(t_j)$ equals 1, then $w_{2^r}(t_{2j})$ has the same value, and for $w_{2^{r+1}}(t_j) = -1$, the same is true for $w_{2^r}(t_{2j})$, because every positiv and negativ interval of w_{2^r} is halved in $w_{2^{r+1}}$, latter having a positive value over the first, and a negative value over the second half. Let us now consider the case $j < 2^N$, and an odd $2k + 1$. Then

$$w_j^{(N+1)}(t_{2k+1}^{(N+1)}) = w_j^{(N)}(t_k^{(N)}) = w_k^{(N)}(t_j^{(N)}) = w_k^{(N+1)}(t_{2j+1}^{(N+1)}),$$

and we have to prove that $w_k^{(N+1)}(t_{2j+1}^{(N+1)}) = (t_j^{(N)}) w_{2k+1}^{(N+1)}$,

$w_k^{(N+1)}(t_{2j}^{(N+1)}) = w_{2k}^{(N+1)}(t_j^{(N+1)})$ is seen to be valid, therefore we have to demonstrate that $w_k^{(N+1)}(t_{2j+1}^{(N+1)})$ differs by the same factor from $w_k^{(N+1)}(t_{2j}^{(N+1)})$ as $w_{2k+1}^{(N+1)}(t_j)$ differs from $w_{2k}^{(N+1)}(t_j^{(N+1)})$. This statement is true, because $w_{2k+1} = = w_{2k} \cdot w_1$, and therefore both have the same value for $t_j^{(N+1)} < 1/2$, and an opposite value for $t_j^{(N+1)} > 1/2$. The latter is, however, impossible, because we are treating the case $j < 2^N$. The same is true for $w_k^{(N+1)}(t_{2j}^{(N+1)})$ and $w_k^{(N+1)}(t_{2j}^{(N+1)})$, because also $k < 2^N$ is true, and therefore $w_k^{(N+1)}(t_{2j}^{(N+1)}) = w_k^{(N)}(t_j^{(N)}) = = w_k^{(N+1)}(t_{2j+1}^{(N+11)})$.

Thereby we have demonstrated our assertion for any k and for $j < 2^N$. For $j \geqslant 2^N$, we have

$$w_j^{(N+1)} = w_{2^N}^{(N+1)} \cdot w_{j-2^N}^{(N+1)}.$$

Moreover, for $j = 2^r$, and for $j - 2^N < 2^N$ we have already proven our assertion, that is therefore also true for its product, because

$$w_j^{(N+1)}(t_k^{(N+1)}) = w_{2^N}^{(N+1)}(t_k^{(N+1)}) \cdot w_{j-2^N}^{(N+1)}(t_k^{(N+1)}) = w_k^{(N+1)}(t_{2^N}^{(N+1)}) \cdot w_k^{(N+1)}(t_{j-2^N}^{(N+1)}) =$$
$$= (-1) \cdot w_k^{(N+1)}(t_{j-2^N}^{(N+1)}) = w_k^{(N+1)}(t_{j-2^N+2^N}^{(N+1)}) = w_k^{(N+1)}(t_j^{(N+1)})$$

namely for $k \neq 0$; $w_k^{(N+1)}(t_{2^N}^{(N+1)}) = -1$, and

$$w_k(t_j) = -w_k\left(\frac{1}{2} + t_j\right) \quad \text{for} \quad t_j < \frac{1}{2}.$$

Our assertion is valid for $k = 0$, because $w_j(t_0) \equiv 1 \equiv w_0(t_j)$.

Our proof is now complete.

Theorem 2 demonstrates that the evaluation of the φ_j and that of the $f(t_k)$ — s may be done identically except for factor 2^{-N}.

Theorem 3. In the computation pattern

a_{11}	$a_{21} = a_{11} + a_{12}$	$a_{31} = a_{21} + a_{23}$	$a_{41} = a_{31} + a_{35}$	$a_{n,1} =$
a_{12}	$a_{22} = a_{11} - a_{12}$	$a_{32} = a_{22} + a_{24}$	$a_{42} = a_{32} + a_{36}$	$a_{n-1,1} + a_{n-1,2^{n-1}}$
a_{13}	$a_{23} = a_{13} + a_{14}$	$a_{33} = a_{21} - a_{23}$	$a_{43} = a_{33} + a_{37}$	.
a_{14}	$a_{24} = a_{13} - a_{14}$	$a_{34} = a_{22} - a_{24}$	$a_{44} = a_{34} + a_{38}$	.
a_{15}	$a_{25} = a_{15} + a_{16}$	$a_{35} = a_{25} + a_{27}$	$a_{45} = a_{31} - a_{35}$	.
a_{16}	$a_{26} = a_{15} - a_{16}$	$a_{36} = a_{26} + a_{28}$	$a_{46} = a_{32} - a_{36}$	.
a_{17}	$a_{27} = a_{17} + a_{18}$	$a_{37} = a_{25} - a_{27}$	$a_{47} = a_{33} - a_{37}$	.
a_{18}	$a_{28} = a_{17} - a_{18}$	$a_{38} = a_{26} - a_{28}$	$a_{48} = a_{34} - a_{38}$	.
.	.	.	.	
.	.	.	.	
.	.	.	.	
$a_{1,2^n}$	$a_{2,2^n} = a_{1,2^n-1} - a_{1,2^n}$	$a_{3,2^n} = a_{2,2^n-2} - a_{2,2^n}$	$a_{4,2^n} = a_{3,2^n-4} - a_{3,2^n}$	let

$a_{1,i} = f(t_{i^*(i)})$, where i^* is the number in dyadic form, given by the reserved dyadic representation of $i - 1$. Then $a_{n,j}$ gives the value of φ_j in the expansion of $f(t)$.

Proof: A similar pattern, accomplished with some factors of the form $e^{ik\frac{\pi}{r}}$, is known to give the values of the *Fourier* polynomial in the fast version.

It is easy to prove that the same ideas can be used here with some simplification, according to the above given scheme. In the Fourier version we have partitioned the polynomial into an even and an odd part, and the latter into a product with factors z and another even polynomial. The former contains the coefficient with even subscripts, the latter with odd subscripts. Giving the subscripts in dyadic form, and reverting them, the formers are less than the latters. In our case the Walsh-functions with even subscripts do not, the functions with odd subscripts do contain the factor w_1 and therefore the latter part can be factorized into w_1 multiplied by Walsh functions with even subscripts. The factor w_1 has the value $+1$ for $t_k < \frac{1}{2}$, and -1 for $t_k < \frac{1}{2}$. In the second step both parts are partitioned in the Fourier case into two subparts, the formers containing a polynomial with exponents divisible with 4, the latters are factorizable into z^2 and into another polynomial with exponents divisible by 4. Accordingly both parts can be partitioned into two subparts, the formers with indices divisible by 4 — (they do not contain the factor w_2), the latters factorizing into the factor w_2 and a subpart, which does not contain w_2. The factor w_2 has the value $+ 1$ for $t_k < \frac{1}{4}$ and for $\frac{1}{2} < t_k < \frac{3}{4}$, and the value -1 in the intervals $\frac{1}{4} < t_k < \frac{1}{2}$ and $\frac{3}{4} < t_k < 1$. Continuing the process in this way we get terms of the form $w_{2N-1}(\varphi_i + \varphi_j w_{2N})$ at the end, where φ_i has a factor $+1$ anywhere, whereas the factor φ_j takes alternatively the values $+1$ and -1, moreover the factor w_{2N-1} takes the values $+1, +1, -1, -1, +1, +1, -1, -1 \ldots$ It means that in both cases we have the same partitioning and factorizing possibilities, and a similar modularity, which is, however, simple in the Walsh-case, because all the factors take periodically the only values $+1$ or -1 at the basic points. Our proof is now complete.

Summary

For the classification of processes into a finite set of subclasses by the help of the time evolution of some measurable parameters, every orthonormal system is competitive for the projection into a finite dimensional space in lack of a priori knowledge of the probability degree in the space of the sample functions. In this case the Walsh-system has favourable properties, because it guarantees a good Chebyshev-approximation and the most effective computer use. These properties are presented and proven in this paper.

References

1. WALSH, J. L.: A closed set of normal orthogonal functions. American Journal of Math. 45 (1923) pp. 5–24.
2. ALEXITS, G.: Konvergenzprobleme der Orthogonalreihen. Budapest, Akadémiai Kiadó 1960.

† Prof. Dr. Tamás FREY (H-1521 Budapest)
Tibor NAGYPÁL, H-1145 Budapest XIV, Amerikai út 57. National Institute of Neurosurgery, Hungary