An *M*-ary Coherent Optical Receiver for the Free-Space Channel

V. A. Vilnrotter
Telecommunications Systems Section

The free-space channel is an ideal medium for communicating by means of spatially and temporally coherent optical fields. Here we derive the structure of a coherent optical receiver for M-ary signals assuming Poisson detection statistics. Receiver performance is evaluated for an M-ary signal set consisting of orthogonal Walsh-functions under the assumption of high-intensity symbol counts. The asymptotic performance bound is examined in the limit as both the dimension of the signal-set and the receiver bandwidth become arbitrarily large and it is shown that on the average a maximum of nearly three bits of information can be encoded onto each received photon using the above modulation scheme.

I. Introduction

In recent years, considerable effort has been devoted to the problem of deep-space communication by means of modulated optical fields (Refs. 1,2). Due to the inherently high gain of diffraction-limited optical antennas, optical communication systems offer the potential for high data-rate communication over interplanetary distances. However, since optical fields are susceptible to attenuation and distortion by the terrestrial atmosphere, current efforts have concentrated on communication systems operating over the free-space channel, which provides a distortionless medium for the propagation of optical fields. This type of free-space system could be implemented by placing the optical receiver outside of the terrestrial atmosphere (i.e., in Earth orbit) and relaying the received data to ground by means of weather-independent RF data-links. In space, the relay receiver observes undistorted optical fields, and can therefore be designed to take advantage of the spatial and temporal coherence of the received field. In the following sections, we derive the structure and evaluate the performance of an optical receiver that coherently combines the received field with a locally generated optical field prior to photodetection, and attempts to retrieve the transmitted data by processing the detector's response to the combined field. We develop a mathematical model for the detection process, derive the structure of a maximum a posteriori (MAP) decoder and evaluate receiver performance for the case of orthogonal M-ary signals based on the well-known Walsh functions, a signaling scheme that allows for considerable simplifications in the subsequent analysis.

II. Coherent Optical Receiver Model

The coherent optical receiver considered here adds a strong, locally generated field coherently to the received optical field prior to photodetection, as shown in Fig. 1. The received field can be modelled as a plane-wave:

$$U_R(t) = U_R m(t) \exp\left[j\left(2\pi\nu t + \phi_R(t)\right)\right] \qquad (1)$$

where $2\pi\nu$ is the radian frequency of the received field, U_R is the (real) field amplitude, $m(t)$ is the modulation ($|m(t)| \leqslant 1$), and $\phi_R(t)$ is a random phase process associated with the received field. Similarly, we model the local field as

$$U_L(t) = U_L \exp\left[j\left(2\pi\nu t + \phi_L(t)\right)\right] \qquad (2)$$

where now U_L is an equivalent (real) local field amplitude referred to the receiver aperture, and $\phi_L(t)$ is again a random phase process. The detector responds to the instantaneous power of the sum field $s(t)$ by releasing free electrons from its active surface. If the detector bandwidth is sufficiently great, then for a given modulation function the detector output process $x(t)$ can be modelled as an inhomogeneous Poisson process, represented by a sequence of randomly occurring impulses. If the phase of the local field can be made to match that of the received field ($\phi_R(t) \cong \phi_L(t)$), then the modulation function $m(t)$ gives rise to an intensity function $\lambda(t)$, where

$$\begin{aligned}\lambda(t) &= \frac{\eta}{h\nu} A_R\, |s(t)|^2 \\ &\cong \frac{\eta}{h\nu} A_R \left(U_L^2 + 2\, U_L U_R m(t)\right) \triangleq \lambda_L + \lambda_s(t) \qquad (3)\end{aligned}$$

where we have assumed that $U_L >> U_R$. (In Eq. (3) η is the detector quantum efficiency, h is Planck's constant, ν is the frequency of the optical field and A_R is the area of the collecting aperture.) The count intensity therefore contains a constant term due to the local field and a modulation-dependent term due to the crossproduct of the local and received fields.

Let the duration of each symbol be T seconds, and consider the time interval $[0,T)$. Assuming an infinite bandwidth detector, the output process can be represented as a sequence of randomly occurring impulses. With k such impulses occurring in $[0, T)$, a particular sample function of the output process can be represented as

$$x(t) = \begin{cases} 0 & ; \ k = 0 \\ \displaystyle\sum_{n=1}^{k} \delta(t - t_n) & ; \ k \geqslant 1 \end{cases} \qquad (4)$$

Following Snyder (Ref. 3) we define the "sample function density" for this inhomogeneous Poisson process as

$$p(x(t); 0 \leqslant t < T) = \begin{cases} \exp\left[-\displaystyle\int_0^T \lambda(t)\, dt\right] & ; \ k = 0 \\ \displaystyle\prod_{n=1}^{k} \lambda(t_n) \exp\left[-\int_0^T \lambda(t)\, dt\right] & ; \ k \geqslant 1 \end{cases} \qquad (5a)$$

or equivalently

$$p(x(t); 0 \leqslant t < T) = \exp\left\{-\int_0^T \left[\lambda(t) - x(t) \ln \lambda(t)\right] dt\right\} \qquad (5b)$$

where Eq. (5b) follows from the sifting property of delta functions. Note that for $k = 0, x(t) = 0$ in $[0,T)$, and therefore (5b) reduces to

$$\exp\left\{-\int_0^T \lambda(t)\, dt\right\}$$

as in (5a).

III. Decoding of *M*-ary Coherent Signals

Consider the case where the transmitter generates one of M symbols in the time interval $[0,T)$. If the ith symbol is transmitted, the modulation function $m_i(t)$ generates the ith Poisson intensity $\lambda_i(t)$. Under the maximum a posteriori (MAP) decoding criterion the decoder selects that symbol whose probability, conditioned on the received sample function $x(t)$, is the greatest. Denoting the ith hypothesis by H_i, this requires computing Prob $(H_i | x(t); 0 \leqslant t < T)$ for all i, and selecting the hypothesis corresponding to the greatest posterior probability. Letting $P(H_i)$ denote the a priori probability of H_i, and $p(x(t); 0 \leqslant t < T | H_i)$ denote the conditional sample function density, the decoder can equivalently compute

$$P(H_i)\, p(x(t); 0 \leqslant t < T | H_i) / p(x(t); 0 \leqslant t < T)$$

Keeping only hypothesis-dependent terms and taking the natural log, the decoder evaluates

$$\Lambda_i = Y_i + \int_0^T x(t) \ln \left(1 + \frac{\lambda_{si}(t)}{\lambda_L}\right) dt \tag{6}$$

where

$$Y_i = \ln P(H_i) - \int_0^T \lambda_{si}(t)\, dt$$

Based on Eq. (6), the MAP decoder for M-ary signals can be implemented as shown in Fig. 2. Upon observing a particular sample function of the detector output process $x(t)$, the decoder multiplies this output by a set of hypothesis-dependent logarithmic weighting functions, integrates for T-seconds, adds the appropriate hypothesis-dependent bias terms and selects the hypothesis corresponding to the largest resulting value.

IV. Receiver Performance

The receiver structure derived in the previous section applies to a general class of M-ary symbols, since no restrictions were imposed on the modulating functions (other than the normalization condition $|m(t)| \leqslant 1$). However, it is difficult to evaluate receiver performance for arbitrary modulation formats and a priori probabilities. The analysis may be greatly simplified for the case of equilikely orthogonal symbols generated by modulating functions that assume only a few discrete values. For purposes of analysis, a particularly simple signal set consists of the set of orthogonal Walsh functions $\{wal\ (i;t)\}$ which are defined and discussed in Appendix A. Each function assumes the values ± 1 in a given T-second interval. (We exclude the zeroth Walsh function $wal\ (0; t)$ from the signal set because it is dissimilar from all other Walsh elements in that its average value is not zero, and would therefore require special treatment in the analysis.) For equilikely Walsh symbols we let $P(H_i) = 1/M$ and $m_i(t) = wal\,(i;t)$, in which case the ith Poisson intensity function becomes $\lambda_i(t) = \lambda_L + \lambda_s\, wal\ (i;t)$, $i = 1, 2, \ldots, M$. Now the decoder computes only

$$\Lambda_i = \int_0^T x(t) \ln \left(1 + \frac{\lambda_s}{\lambda_L} wal\ (i;t)\right) dt \tag{7}$$

because the additive bias no longer depends on which hypothesis occurred.

Since $x(t)$ is a sample function of a random process, its weighted integrals Λ_i are random variables. The second characteristic function of these random variables may take on either of two forms, depending on the relation between the transmitted symbol and the weighting applied to the detector output: if H_i is true, then let $\psi_{ii}(\omega)$ denote the second characteristic function of Λ_i, and let $\psi_{i\ell}(\omega)$ denote the second characteristic function of Λ_ℓ, $\ell \neq i$. In order to determine $\psi_{ii}(\omega)$, we note from the properties of Walsh functions that $x(t)$ is a Poisson process with rate $(\lambda_L + \lambda_s)$ for a total of $T/2$ seconds, during which time a weighting factor of

$$\ln\left(1 + \lambda_s/\lambda_L\right)$$

is applied, while for the remaining $T/2$ seconds the rate parameter and weighting factor become $(\lambda_L - \lambda_s)$ and

$$\ln\left(1 - \lambda_s/\lambda_L\right)$$

respectively. Since intervals with different weighting factors are disjoint, integration over these regions yields independent weighted Poisson random variables. Hence, under H_i we can equivalently express Λ_i as

$$H_i;\ \Lambda_i = \int_0^{T/2} x^+(t) \ln\left(1 + \frac{\lambda_s}{\lambda_L}\right) dt + \int_{T/2}^{T} x^-(t) \ln\left(1 - \frac{\lambda_s}{\lambda_L}\right) dt \tag{8}$$

where $x^+(t)$ is a Poisson process with intensity $(\lambda_L + \lambda_s)$, and $x^-(t)$ is a Poisson process with intensity $(\lambda_L - \lambda_s)$. The characteristic function for Λ_i under hypothesis i therefore becomes the product of the characteristic functions corresponding to each integral in (8); hence the second characteristic function can be expressed as

$$\begin{aligned}\psi_{ii}(\omega) &= \ln E\ \{\exp\,(j\omega\Lambda_i)\} \\ &= \frac{T}{2}\lambda_L\left[\exp\left(j\omega \ln\left(1 + \frac{\lambda_s}{\lambda_L}\right)\right) + \exp\left(j\omega \ln\left(1 - \frac{\lambda_s}{\lambda_L}\right)\right) - 2\right] \\ &\quad + \frac{T}{2}\lambda_s\left[\exp\left(j\omega \ln\left(1 + \frac{\lambda_s}{\lambda_L}\right)\right) - \exp\left(j\omega \ln\left(1 - \frac{\lambda_s}{\lambda_L}\right)\right)\right]\end{aligned} \tag{9}$$

A similar argument can be used to derive $\psi_{i\ell}(\omega)$. For this case the Poisson intensity is still defined by the ith Walsh function, but now the weighting depends on the ℓth Walsh function, $\ell \neq i$. Now Λ_ℓ is obtained by integrating $x(t)$ over four disjoint intervals of equal duration during which all four possible combinations of count intensities and weights are assumed:

$$H_i: \Lambda_\ell = \int_0^{T/4} x^+(t) \ln \left(1 + \frac{\lambda_s}{\lambda_L}\right) dt + \int_{T/4}^{T/2} x^+(t) \ln \left(1 - \frac{\lambda_s}{\lambda_L}\right) dt + \int_{T/2}^{3T/4} x^-(t) \ln \left(1 + \frac{\lambda_s}{\lambda_L}\right) dt + \int_{3T/4}^{T} x^-(t) \ln \left(1 - \frac{\lambda_s}{\lambda_L}\right) dt \qquad (10)$$

In this case $\psi_{i\ell}$ reduces to

$$\psi_{i\ell}(\omega) = \frac{T}{2}\lambda_L \left[\exp\left(j\omega \ln \left(1 + \frac{\lambda_s}{\lambda_L}\right)\right) + \exp\left(j\omega \ln \left(1 - \frac{\lambda_s}{\lambda_L}\right)\right) - 2\right] \qquad (11)$$

In typical applications, the total power of the local field is great enough to generate a high intensity process at the detector output. Expanding the exponentials in (9) and (11) and letting the intensity due to the local field become arbitrarily large ($\lambda_L \to \infty$) we obtain the limiting forms

$$\psi_{ii}(\omega) \xrightarrow[\lambda_L \to \infty]{} j\omega (4K_s) - \frac{\omega^2}{2}(4K_s) \qquad (12a)$$

$$\psi_{i\ell}(\omega) \xrightarrow[\lambda_L \to \infty]{} - \frac{\omega^2}{2}(4K_s) \qquad (12b)$$

where we define K_s as the average symbol count that would be generated by the received field if it were direct-detected by the same receiver:

$$K_s = \frac{\eta}{h\nu} A_R \, T \, U_R^2$$

Equation (12a) is recognized as the second characteristic function of a Gaussian random variable with mean value $(4K_s)$ and variance $(4K_s)$, while (12b) corresponds to a Gaussian random variable with zero mean and variance $(4K_s)$. It follows that as the average intensity λ_L becomes suitably great, the random variables Λ_i may be approximated by Gaussian random variables with mean and variance as defined above.

The probability of correct decoding can be found by computing the probability that given H_i, Λ_i exceeds all other Λ_ℓ, $\ell \neq i$. Making use of the Gaussian approximation, the probability of correct decoding, $P(C)$, can be expressed as

$$P(C) = \int_{-\infty}^{\infty} dx \frac{e^{-(x - 4K_s)^2/2(4K_s)}}{\sqrt{2\pi(4K_s)}} \times \left[\int_{-\infty}^{x} dy \frac{e^{-y^2/2(4K_s)}}{\sqrt{2\pi(4K_s)}}\right]^{(M-1)} \qquad (13)$$

while the symbol error probability becomes $P(E) = 1 - P(C)$. Equation (13) is well known in the context of M-ary orthogonal signal detection in additive Gaussian noise, a similarity that stems from the Gaussian approximation made in Eq. (12). Assuming a photodetector quantum efficiency of one ($\eta = 1$), it is convenient to define the "photon information rate" as $\rho \triangleq (\log_2 M)/K_s$, which can be interpreted as the average information in bits encoded onto each detected photon (Ref. 4). Since each symbol gives rise to $(M/2)/(M\text{-}1)$ bit errors, the average bit error probability $P_B(E)$ can be expressed as

$$P_B(E) = \left(\frac{M/2}{M-1}\right)\left\{1 - \int_{-\infty}^{\infty} dz \frac{e^{-z^2}}{\sqrt{\pi}} \times \left[1 - \frac{1}{2} Erfc\left(z + \sqrt{\frac{2\log_2 M}{\rho}}\right)\right]^{(M-1)}\right\} \qquad (14)$$

where we recognized the relation between the bracketed term in (13) and the complementary error function. Equation (14) has been tabulated extensively in the literature (Refs. 5, 6). Figure 3 shows graphs of the average bit-error probability $P_B(E)$ as a function of ρ for increasing M, obtained from the tabulated values in (Ref. 6). Note that photon information rates of roughly 1 bit per photon can be achieved in the range of bit error probabilities of interest ($P_B(E) \lesssim 10^{-3}$) with signal sets of moderate dimension ($2^5 < M < 2^{10}$). The limiting behavior of $P_B(E)$ as the number of transmitted symbols becomes arbitrarily large can be obtained from the corresponding results derived for additive Gaussian noise by a simple transformation of variables (Refs. 5, 6):

$$\lim_{M\to\infty} P_B(E) = \begin{cases} \frac{1}{2}; & \frac{2}{\rho} < \ln 2 \\ 0; & \frac{2}{\rho} > \ln 2 \end{cases} \tag{15}$$

It follows that arbitrarily low bit error probabilities can be achieved as M approaches infinity, as long as the inequality $\rho < 2/\ln 2 \cong 2.89$ bits/photon is satisfied. This implies that when operating with wideband, high quantum-efficiency photodetectors, the M-ary coherent optical system described above can theoretically transfer several bits of information per received photon on the average (depending on the required bit error probability), but the limiting value of 2.89 bits/photon cannot be exceeded, regardless of system complexity.

V. Conclusions

A coherent optical receiver model for M-ary orthogonal signals has been examined. The structure of the MAP decoder designed for arbitrary signals was developed, and receiver performance evaluated for the special case of M-ary orthogonal signals derived from Walsh functions. This specialization enabled the development of a mathematically rigorous solution for receiver performance in the limiting case of high intensity detection, where a Gaussian approximation could be invoked. It should be emphasized, however, that these results also apply to a more general class of orthogonal signals, as long as the Gaussian approximation for the test random variables $\{\Lambda_i\}$ can be justified. Receiver performance was evaluated in terms of the "photon information rate" ρ, which is a measure of the average information that can be encoded onto each detected photon. The equivalent bit error probability depends both on ρ and on the signal-set dimension M. It was found that reliable communication could not be achieved at photon information rates exceeding 2.89 bits/photon. However, reliable communication at photon information rates exceeding one bit per photon appears feasible with M-ary coherent optical systems employing high quantum-efficiency, wideband photodetectors, if negligibly small optical phase error can be maintained. The effects of phase error and of other external disturbances (such as background radiation) on the performance of M-ary coherent optical receivers remains to be examined in future studies.

References

1. Vilnrotter, V. A., and Gagliardi, R. M., "Optical Communications Systems for Deep-Space Applications," Publication 80-7, Jet Propulsion Laboratory, Pasadena, Calif., Mar. 15, 1980.
2. Gagliardi, R. M., Vilnrotter, V. A., and Dolinar, S. J., "Optical Deep Space Communication via Relay Satellite," Publication 81-40, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 15, 1981.
3. Snyder, D. L., *Random Point Processes*, J. Wiley, New York, N.Y., 1975.
4. McEliece, R. J., and Welch, L. R., "Coding for Optical Channels with Photon Counting," in *The DSN Progress Report 42-52*, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 15, 1979.
5. Viterbi, A. J., *Principles of Coherent Communication*, McGraw-Hill, New York, N.Y., 1966.
6. Lindsey, W. C., and Simon, M. K., *Telecommunication Systems Engineering*, Prentice-Hall, Englewood Cliffs, N.J., 1973.
7. Harmuth, F. H., *Transmission of Information by Orthogonal Functions*, Springer-Verlag, Berlin, Germany, 1969.

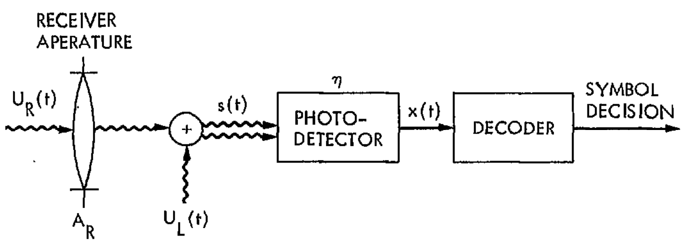

Fig. 1. Coherent optical receiver block diagram

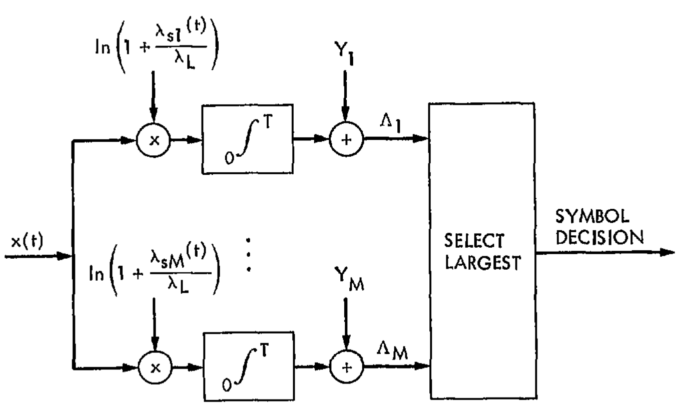

Fig. 2. MAP receiver structure

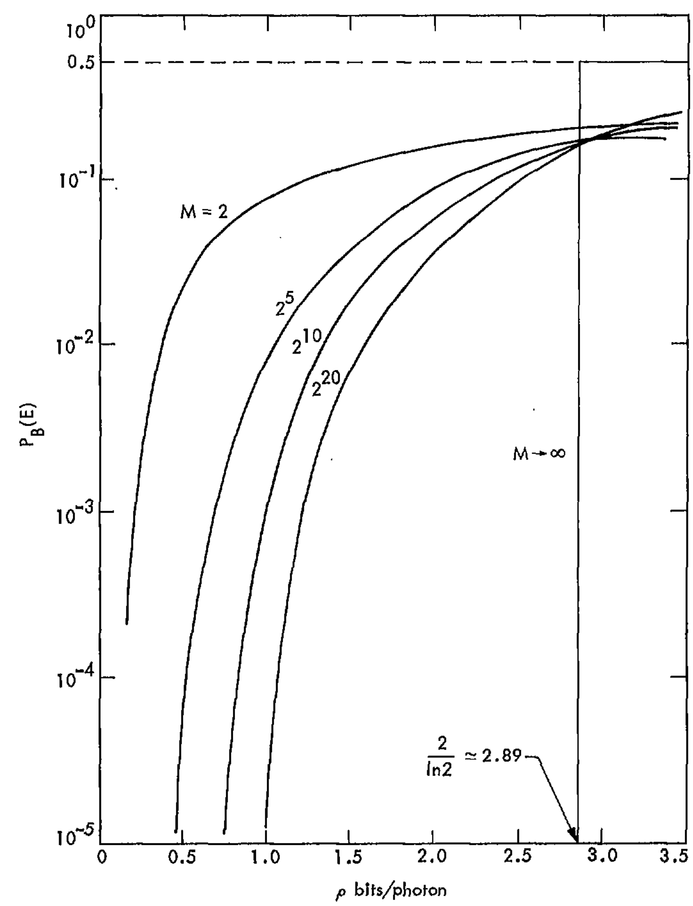

Fig. 3. *M*-ary coherent optical receiver performance: $P_B(E)$ vs ρ for various *M*

Appendix A
Properties of Walsh Functions

An extensive discussion of the properties of Walsh functions can be found in Ref. 7. For our interest here, it is convenient to define the Walsh functions $wal\,(i;t)$ over the interval $[0, T)$, indexed by i. The Walsh functions can be generated by the equation

$$wal\,(2m+p\,;t) = (-1)^{[m/2]+p}\left\{wal\,(m\,;2t) + (-1)^{m+p}\,wal\left(m\,;2\left(t-\frac{T}{2}\right)\right)\right\} \quad \text{(A-1)}$$

$p = 0$ or $1, m = 0,1,2,\ldots$

where

$$wal\,(0\,;t) = \begin{cases} 1;0 \leqslant t < T \\ 0\,;\text{elsewhere} \end{cases} \quad \text{(A-2)}$$

In Eq. (A-1) $[m/2]$ means "greatest integer less than or equal to $m/2$." As an example, let us generate $wal\,(1;t)$. This requires setting $m = 0, p = 1$. Hence

$$wal\,(1\,;t) = (-1)\left\{wal\,(0\,;2t) - wal\left(0\,;2\left(t-\frac{T}{2}\right)\right)\right\} \quad \text{(A-3)}$$

All subsequent elements are obtained from previous elements by applying scaling, shifting and sign reversal operations.

The product of two Walsh functions is another Walsh function:

$$wal\,(i\,;t)\,wal\,(\ell\,;t) = wal\,(i \oplus \ell\,;t) \quad \text{(A-4)}$$

where $\oplus$ stands for modulo 2 addition of the indices expressed in binary form. Note that the product of a Walsh function with itself always yields $wal\,(0;t)$.

Walsh functions are orthogonal and therefore obey the relation

$$\frac{1}{T}\int_0^T wal\,(i\,;t)\,wal\,(\ell\,;t)\,dt = \delta_{i\ell} \quad \text{(A-5)}$$

where $\delta_{i\ell}$ is the Kroenecker delta. Since each Walsh function $wal\,(i;t), i \geqslant 1$ is orthogonal to $wal\,(0;t)$, it follows that

$$\int_0^T wal\,(i\,;t)\,dt = 0\,;i \geqslant 1 \quad \text{(A-6)}$$

Therefore each of these functions assume the values +1 and −1 for exactly half the total duration, or $T/2$ seconds.

Consider any two Walsh functions $wal\,(i;t)$ and $wal\,(\ell;t)$ satisfying $i,\ \ell \geqslant 1,\ i < \ell$. We maintain that a point-by-point comparison yields all four possible combinations of +1 and −1 for exactly one quarter of the total duration, or $T/4$ seconds. This assertion can be proven as follows: the product of any two Walsh functions is another Walsh function, which must assume the values +1 and −1 for equal lengths of time in order to satisfy (A-6). The two Walsh components of the product must therefore agree in sign for $T/2$ seconds (to generate +1 in the product) and disagree in sign for $T/2$ seconds (to generate −1 in the product). Let d_{++} denote the total duration over which both component Walsh functions agree in sign and are positive, d_{--} denote the duration over which they agree in sign and are negative, and likewise d_{-+} and d_{+-} denote durations over which the component functions disagree in sign. (Let the first sign in the subscript refer to the Walsh function with the lower index, i.) Clearly, $d_{++} + d_{--} = d_{+-} + d_{-+} = T/2$ in order for the product to satisfy (A-6). It remains to be shown that

$$d_{++} = d_{--} = d_{+-} = d_{-+} = T/4 \quad \text{(A-7)}$$

Let $d_{++} = d_{--} + \Delta$ for some Δ. Then it must be true that $d_{-+} = d_{+-} + \Delta$ in order for $wal\,(i;t)$ to satisfy (A-6), but we must also have $d_{+-} = d_{-+} + \Delta$ since $wal\,(\ell;t)$ must also satisfy (A-6). Hence $\Delta = 0$, and $d_{++} = d_{--}$. A parallel argument shows that $d_{-+} = d_{+-}$, and Eq. (A-7) follows.