A METHOD OF BLIND EQUALIZATION

Y. SATO, S. HASHIMOTO AND H. ODA

TOHÔ UNIVERSITY FACULTY OF SCIENCE

RESUME

Aveugle égalisation, ou ce qu'on appelle réexercice à Data, est la fonction significative à transmettre des données en très vite. Ce problème est fondamentalement à identifier le lien systèmes inconnus sans mesurer le système de puissance. Même si le système prise de puissance est IID, il devient très difficile à égaliser aveuglement en cas généraux où le système n'est pas causal. Quelques travaux précédents prétendent que le plan à réduire en minimum

\[ E(y_k - \text{sign}(y_k))^2, \]

où \( y_k \) est à égaliser débit de puissance. Le renouvellement des paramètres de l'égaliseur s'établissait en on-line operation, où ils sont corrige en petite valeur à chaque \( k \) chance à recevoir le signal.

En ce papier, l'aveugle égaliseur d'off-line operation aprés enregistrer la longue suite de système débit de puissance est présenté. Cette off-line operation va être banales dans autres champs sauf transmettre des données où l'information faut être opérée en l'heure réelle. Ce off-line algorithme accompli le application à résoudre

\[ E(y_k - \text{sign}(y_k))^2 \rightarrow 0 \ (k \neq 0). \]

Résultantement, cette idée conduit à la méthode d'enlever la dépendance entre égaliseur débits des puissances \( y_k \).

Section 1, on décrit l'histoire d'aveugle égaliseur. Section 2, on discute de deux idées; la première est à égaliser la fonction de probabilité densité d'égaliseur débit de puissance à celle de système prise de puissance, et la deuxième est à enlever la dépendance entre égaliseur débits des puissances. En section 3, on propose l'off-line algorithme. En section 4, notre off-line algorithme est confirmé en cas limité. Section 5, on montre quelque résultats de la simulation de computeur.

SUMMARY

Blind equalization, or called Retrain on Data is significant function in high speed data transmission. This problem is basically to identify the unknown linear system without measuring system input. Even if system input is IID, it becomes very difficult to blind-equalize in general cases where the system is not causal. Several previous works presented the scheme to minimize

\[ E(y_k - \text{sign}(y_k))^2, \]

where \( y_k \) is equalizer output. The updating of equalizer tap-weights was established in on-line operation, where they are corrected by small increment at every \( k \)th signal receiving chance.

In this paper, an off-line blind equalization processed after recording a long time sequence of system output is reported. Such an off-line operation will be conventional in other fields except data transmission where the information must be processed in real time. The present off-line algorithm performs a contraction mapping to solve

\[ E(y_k - \text{sign}(y_k))^2 \rightarrow 0 \ (k \neq 0). \]

Resultantly, this idea leads to a method of removing time-dependency in equalizer output \( y_k \).

Section 1 describes a past history of the blind equalization. Section 2 discusses the concept of two ideas; the first is to equalize probability density function of equalizer output with that of system input, and the second is to remove time-dependency in equalizer output. Section 3 proposes an off-line algorithm. In section 4, the present off-line algorithm is confirmed for the limited case. Section 5 shows the several results of computer simulation.
1. Introduction

Blind equalization, compelled to converge to an optimum equalizer state without referring the sending data sequence, is of interest to research engineers in several fields: image processing; signal detection; biophysics, in particular, neuro-physiology and statistical mathematics. Starting from the original work in this subject of the high speed data transmission by Y. Sato; A. Benveniste framed out the theoretical structure and his work is appreciated highly. The blind equalization is called "Retrain on data" in CCITT V.27 and V.29 recommendations, then, with high speed data modems must be equipped. For the data transmission cases, the original work achieved a blind equalization with theoretical treatment and computer simulation under limited situations, where the sending sequence is independent and has identically uniform distribution (IID) through the data scrambling and multi-level coding. The extended and general work by A. Benveniste derived theories for blind equalizable functionals having the steepest descent lines which converge to the optimum, and they mean to lighten the restrictions for sending sequence distribution and distortion of linear system. A primal difficulty comes from nonminimum phase property of linear system, since the system not causal. In order to overcome it, new techniques except application of the classical adaptive linear prediction to the autoregressive systems had been required.

Denote the distribution of the system input of IID by \( v \) and the equalizer output by \( y_k \), the theorem by A. Benveniste says "If the distribution of \( y_k \) is \( v \) too, the linear system must be transparent and \( y \) must be equal to the system input with time shift ambiguity, where \( v \) is not Gaussian." Every previous work intended to force implicitly the equalizer to depend on the distribution of \( v \), beside directly aiming intersymbol interference reduction or time independency in the sequence \( y_k \). The problem of blind equalization remains a large and serious part. Questions are, for example, what is solved for the whole permitted distribution \( v \) associated with blind algorithms and, furthermore, what kind of knowledge of \( v \) is substantial for this problem. It is evident that, by means of any schemes from standpoints of minimizing the correlation \( E[y_v y_{v-i}]/v_0 \) of intersymbol interference cannot be removed. It should be remarked that the independency as

\[
P(y_k = \alpha | y = \text{constant}) = P(y_k)
\]

(for all \( \alpha \) \( 0 \) and the constant is arbitrary) is equivalent to vanishing intersymbol interference. This paper presents a new type of off-line algorithm to realize the above equations. To simplify the algorithm, the estimated averages of each \( y_k \) when \( \alpha > 0 \) are used to adjust the equalizer parameter, and iterative contraction toward the optimum is accomplished. The computer simulation for several typical models are added.

2. A concept of blind equalization

Presuming that sending data sequence \( a_k \) in Figure 1 is time independent and has identical distribution, two strategies stated below have the same destination that the total response satisfies

\[
t_k \neq 0, \quad t_k = 0 \quad (k \neq 0) \quad (t_k = \sum_{k=n}^{n} a_k \cdot \mathbb{1}).
\]

(1) The first strategy is to equalize the probability density function (p.d.f.) of equilizer output \( y_k \) with that of \( a_k \).

(2) The second is remove the time dependency in \( y_k \) in sense that \( y_k \) is strictly uncorrelated : the p.d.f. of \( y_k \) is \( f_{y_k}(y_k) = \sum_{i=0}^{\infty} f_{x_k}(y_k) \), and any rules on the value of \( y_k \) is selected using any rules on the

The schemes of algorithm design based on above strategies are different in two cases of application when p.d.f. of \( a_k \) is prior unnown and unknown. In the first application, we can establish the more reliable blind equalization using knowledge of the p.d.f.. Along the first strategy of p.d.f. matching, for example, the maximum likelihood estimation \( \hat{\theta}_k \) of \( \theta_k \) after equalizing \( y_k \) will be derived by evaluating the \( y_k \) histogram, then the equalizer tap-weight updating scheme, employing the estimated steepest descent of \( (y_k \cdot \Delta a_k) \) as in decision directed manner is one of the possible blind equalization. In general application cases where the p.d.f. of \( a_k \) is unknown, we cannot take any efficient information of \( a_k \) from a 1th sampling value of \( y_k \). Major question in these situations must be for what class of p.d.f. a presented robust and simple algorithm possesses the desired destination.

The every previous works treated the simple algorithm which minimizes \( E[(y_k \cdot \text{sign}(y_k))^2] \) as

\[
w_k = w_k - \alpha_k \cdot (y_k - \text{sign}(y_k)) \quad \text{where}
\]

\[
\gamma = E[|y_k|] / E[|a_k|] \quad \text{or}
\]

\[
\gamma = E[|y_0|] / E[|a_0|]
\]

Y. Sato proved that this converges to Eq.(1) in case of data communication where \( a_k \) is uniformly distributed and initial peak distortion \( t_k / |t_k| < 1 \). The extended work by A. Benveniste derived that the algorithm of Eq.(2) is expected to smooth out the source noise and converge to Eq.(1). From the distribution matching principle suggested later by A. Benveniste, to minimize \( E[(y_k \cdot \text{sign}(y_k))^2] \), which means to minimize the dispersion around some constant \( y \) in the positive region of \( y_k \), should identify p.d.f. of \( y_k \) with that of \( a_k \). To give further discussion for the concept of blind equalization, decompose \( E[(y_k \cdot \text{sign}(y_k))^2] \) as follows,

\[
E[(y_k \cdot \text{sign}(y_k))^2] = E[y_k^2] - 2\gamma |y_k|^2 + \gamma^2
\]

From this formula, it is seen that the minimization is equivalent to maximizing under power constraint

\[
E[y_k^2] = \text{constant}
\]

To see details of \( E[y_k^2]/|v_0| \), denote \( v_0 \), describe it by \( E[a_k^2] \) when \( \alpha > 0 \) where we employ ensemble average when \( \alpha = 0 \) for the simplicity of notation. Then we have

\[
y_k = \sum_{i=1}^{\infty} a_k \cdot t_i \quad \text{(i)} \quad y_0 = \text{constant}
\]

Providing that the p.d.f. of \( a_k \) (denote \( v \) ) has zero-mean and is symmetric, \( a_k^2 \) for \( \alpha > 0 \) is written as

\[
a_k^2 = \frac{1}{2} \int_{-\infty}^{\infty} f_p(x) \text{dx} \quad a_0^2 = \frac{1}{2} \int_{-\infty}^{\infty} f_p(x) \text{dx}
\]

where \( p(x) \) is a normalized p.d.f. of all \( t_k \) except \( t_0 \) and its variance is determined by \( 1/t_0^{\alpha} \) from unit power constraint. Let us introduce the following expression, which is not strictly correct, but, will be accepted in limited case discussed later.

![Fig.1 Model of equalization system](https://example.com/model.png)
A METHOD OF BLIND EQUALIZATION

\[ \bar{u}_k = u(t_k) \]  \hspace{1cm} (9)

then

\[ \bar{v}_k = \frac{1}{t_k} \int \bar{u}_k(t) \, dt. \]  \hspace{1cm} (10)

It is noted that \( u(t) \) is monotonic and increasing function of \( t \) and \( u(0) = 0 \). Substituting power constraint \( \bar{v}_0 = 1 \) to Eq. (10), \( \bar{v}_0 \) is derived as

\[ \bar{v}_0 = \frac{u(t_0)}{t_0} = \frac{\int_0^{t_0} u(t) \, dt}{t_0} = \frac{\int_0^{t_0} u(t) \, dt}{t_0} \cdot \frac{1}{t_0} \]  \hspace{1cm} (11)

Out of the second term of Eq. (10), its component when \( t_0 = 0 \) is to be omitted. From Eq. (11), it is concluded that the condition

\[ \frac{u(t_0)}{t_0} > \frac{u(t_1)}{t_1} > 0 \quad (\text{for all } i \neq 0) \]  \hspace{1cm} (12)

is necessary and sufficient in order that \( \max \{ y(t) \} \) gives the desired destination \( y(t_0) = (i \neq 0) \). Furthermore, the condition (12) leads to another important result, i.e., the mean square distortion of \( a_k \) is smaller than that of \( t_k \).

\[ D_a = \left( \int_0^1 \frac{u(t)}{t} \, dt \right)^2 < D_t = \left( \int_0^1 \frac{1}{t} \, dt \right)^2. \]  \hspace{1cm} (13)

To find the p.d.f. of \( a_k \) satisfying the global condition (12) is the problem to give breakthrough toward blind equalizability and is remained as future problem. It is easy to examine for the limited case of local behavior around the solution \( t_k = 0 \) \((i \neq 0) \) and \( t_k = 1 \). At first, when \( t_0 \) tends to 1, denoting delta function by \( \delta(x) \),

\[ \frac{1}{t_0} \int_0^\infty \frac{a_0}{t_0} \frac{1}{p(x_0) \, dx_0} / t_0 = \int_0^\infty \frac{a_0}{t_0} \frac{1}{p(x_0) \, dx_0} / t_0 = \frac{1}{t_0} \int_0^\infty \frac{a_0}{t_0} \frac{1}{p(x_0) \, dx_0} / t_0 = E(a_0) \]  \hspace{1cm} (14)

The second, when all \( t_k \) tend to 0 and \( t_0 \) tends to 1,

\[ \lim_0 \int \frac{a_0 \, dx_0}{t_k} = \lim_0 \int \frac{a_0 \, dx_0}{t_k} = \lim_0 \int \frac{a_0 \, dx_0}{t_k} = \lim_0 \int \frac{a_0 \, dx_0}{t_k} = \lim_0 \int \frac{a_0 \, dx_0}{t_k} \]  \hspace{1cm} (15)

If \( v(x) \) is uniform p.d.f. defined in \([-V, V]\), then

Eq. (14) = \( \frac{1}{V} \) and Eq. (15) = \( \frac{1}{3} \),

are derived and it is said that the condition (12) is satisfied in the neighborhood of the solution. For the Gaussian p.d.f.,

\[ u(t) = \frac{1}{\sqrt{2\pi} \sigma^2} \exp \left( -\frac{t^2}{2\sigma^2} \right) \]  \hspace{1cm} (16)

is derived. Therefore, from Eq. (11) it is said that the ambiguity of \( t_0 \) is remained and \( \max \{ y(t) \} \) cannot have the desired destination.

Now, apart from the first strategy of p.d.f. matching, let us recommend a following simple algorithm along the second strategy

\[ y_k = \frac{y(t)}{t_k} \]  \hspace{1cm} (17)

Equation (10) remarks that the k-th expectation \( \bar{u}_k \) has the same polarity of \( t_k \). From this fact, according to Eq. (18) is satisfied and namely the sequence of expectation \( \bar{a}_k \) is the inverse of sequence \( \bar{t}_k \), it is mostly expected that we have the trivial solution \( t_k = 0 \) \((i \neq 0) \).

3. OFF-LINE ALGORITHM

After receiving long sequence of signal \( x_k \), the present algorithm adjusts equalizer tap-weight \( w_k \) so that time averages \( \bar{y}_k \) \((i \neq 0) \) vanish, where

\[ \bar{y}_k = \frac{1}{L} \sum_{k=1}^{L} \text{sign}(y_k) y_{k-1} \]  \hspace{1cm} (18)

Let \( \bar{x}_k \) be time average of \( x_{k-1} \) when \( y_k > 0 \), i.e.,

\[ \bar{x}_k = \frac{1}{L} \sum_{k=1}^{L} \text{sign}(y_k) x_{k-1} \]  \hspace{1cm} (19)

then

\[ \bar{y}_k = \frac{1}{L} \sum_{k=1}^{L} \text{sign}(y_k) x_{k-1} \]  \hspace{1cm} (20)

\[ \bar{y}_k = \frac{1}{L} \sum_{k=1}^{L} \text{sign}(y_k) x_{k-1} \]  \hspace{1cm} (21)

While \( \bar{a}_k \) and \( \bar{t}_k \) are evaluated by time averaging, the equalizer is frozen. Hence, both of them are functional of equalizer tap-weight \( w_k \). From this reason, we must perform \( w_k \) optimization, for example, repeating to solve \( w_k \) satisfying \( \bar{y}_k = 0 \) \((i \neq 0) \) for the averaged \( x_{k-1} \) and to average \( x_{k-1} \) under the solved \( w_k \). Thus, an off-line procedure is described as

Step 1. calculate time average

\[ \bar{x}_k = \frac{1}{L} \sum_{k=1}^{L} \text{sign}(y_k) x_{k-1} \]  \hspace{1cm} (22)

Step 2. solve \( w_k \) to satisfy

\[ \frac{1}{L} \sum_{k=1}^{L} \text{sign}(y_k) x_{k-1} = 0 \]  \hspace{1cm} (23)

and return to step 1.

Through this procedure, the following equivalent one is implicitly carried.

Step 1. calculate time average

\[ \bar{x}_k = \frac{1}{L} \sum_{k=1}^{L} \text{sign}(y_k) x_{k-1} \]  \hspace{1cm} (24)

Step 2. solve \( t_k \) to satisfy

\[ \bar{t}_k = \frac{1}{L} \sum_{k=1}^{L} \text{sign}(y_k) x_{k-1} \]  \hspace{1cm} (25)

and return to step 1.
A METHOD OF BLIND EQUALIZATION

Figure 2 illustrates this implicit procedure, where $T_r$ and $A_r$ show responses $t_r$ and $a_r$, respectively obtained after $r$th iteration times.

To investigate the convergence of the above procedure, normalize $a_r$ and $T_r$ by their peaks as $\tilde{a}_r/a_r$ and $\tilde{T}_r/T_r$ so that the convergence problem can be separately treated for peaks $\tilde{a}_r$ or $\tilde{T}_r$ and for distortion $D_0$ or $D_0'$. 

Rewrite the interaction procedure in Fig. 2 as

$$
\Delta T_{r-1} \rightarrow \Delta A_{r-1} \rightarrow \Delta T_r \rightarrow \Delta A_r \rightarrow \Delta T_{r+1},
$$

where above functions are defined after each peaks are subtracted from $T_r$, $A_r$, etc. and devide them by their peaks. Hence, accounting that time sequences are all real valued, we have, for example

$$
\frac{T_r}{T} \int_{-T_r/T}^{T_r/T} \Delta T_r dw = - \frac{T_r}{T} \int_{-T_r/T}^{T_r/T} \text{Re}\{\Delta T_r\} dw = 0, 
$$

(25)

$$
P_c^r = \frac{T_r}{2T} \int |T_r|^2 dw \ll \|\Delta T_r\|^2. 
$$

(26)

Now, we must confirm the inequalities

$$
\cdots > P_r^{r-1} > P_c^r > P_r^{r+1} > \cdots.
$$

Denote peak of $(1+\Delta T_r)^{-1}$ by $1+\delta t_0^r$, then

$$
\delta t_0^r = \frac{T_r}{2\pi} \int \frac{1}{1+\Delta T_r} \text{Re}\{-\Delta T_r\} dw. 
$$

(27)

Therefore,

$$
P_a^r = \frac{1}{(1+\delta t_0^r)^2} - (\delta t_0^r)^2 \frac{1}{(1+\delta t_0^r)^2}. 
$$

(28)

The deviation of peak value from $1,\delta t_0^r$ is approximately given by

$$
\delta t_0^r = \frac{1}{1 \neq t_{r-1}}. 
$$

(29)

Perform the similar derivation for $\Delta A_r \rightarrow \Delta T_{r+1}$, and taking account that the distortion reduction ratio through $\Delta A_r \rightarrow \Delta T_r \rightarrow \Delta A_r$ are given by $\tilde{u}_t$ (41) and $\tilde{u}_r$ (21) respectively, the reduction ratios from $\delta t_0^r$ to $\delta a_0^r$, whose polarities are the same, is approximated in order of $\tilde{u}_r$. From Eq. (29), and we obtain roughly

$$
P_a^r = \frac{1}{(1+\delta a_0^r)^2} - (\delta a_0^r)^2 \frac{1}{(1+\delta a_0^r)^2}. 
$$

(30)

$$
P_r^r = \left[ \frac{\tilde{u}_r^2}{\tilde{u}_r^2} - \left( \delta a_0^r \right)^2 \right] / \left( 1 + \delta a_0^r \right)^2. 
$$

(31)

Resultantly, in situations where $P_a^r$ is larger than $(\delta a_0^r)^2$, it can be said that $D_r^r < 0$. 

For the peaks $t_0$ and $a_0$, the recursive formula is

$$
T_0 = \frac{1}{\sigma_a} + \delta t_0 
$$

$$
a_0 = \mu_0 t_0 
$$

$$
T_0 = \frac{1}{\sigma_t} + \delta t_0 
$$

$$
a_0 = \mu_0 t_0 
$$

$$
T_0 = \frac{1}{\sigma_t} + \delta t_0 
$$

$$
a_0 = \mu_0 t_0 
$$

From these

$$
t_0^{r+2} = \frac{T_0^{r+1}}{\mu_0^{r+1} + \delta t_0^{r+1} + \frac{\delta t_0^{r+2}}{\sigma_t^{r+2}}} + \delta t_0^{r+2}. 
$$

(33)

In Eq. (33), if $\delta t_0$ tends to zero and $\mu_0^{r+2} / \sigma_t^{r+2}$ tends to 1, it is said that $t_0^{r+2}$ converges to some finite value.

4. Analysis in case of uniform p.d.f. and $\left| t_0^{r+1} \right| \geq \left| t_1 \right|$

The uniqueness that $t_0 = 0 (\lambda \neq 0)$ satisfying $\bar{u}_0 = 0$ \((\bar{u}_0 = 0)\) is resulted from the following derivations (refer Appendix A for several derivations in this section)

$$
\bar{u}_0 = \mu(t_0) t_0, \quad \mu(t_0) = \frac{V}{T_0} (1 + \frac{1}{3\mu}) ,
$$

(34)

$$
\bar{u}_0 = \mu(t_0) t_0, \quad \mu(t_0) = \frac{V}{T_0} (1 + \frac{1}{3\mu}) ,
$$

(35)

$$
\bar{u}_0 = \mu(t_0) t_0, \quad \mu(t_0) = \frac{V}{T_0} (1 + \frac{1}{3\mu}) ,
$$

(36)

$$
\bar{u}_0 = \mu(t_0) t_0, \quad \mu(t_0) = \frac{V}{T_0} (1 + \frac{1}{3\mu}) ,
$$

(37)

$$
\bar{u}_0 = \mu(t_0) t_0, \quad \mu(t_0) = \frac{V}{T_0} (1 + \frac{1}{3\mu}) ,
$$

(38)

Where $t_0 = \frac{T_0^2}{2\lambda}$, and

$$
\lambda = \bar{u}_0^2 / \sigma_0. 
$$

(39)

$$
V_0 = \frac{T_0}{2}(1 + \frac{1}{3\mu}) ,
$$

(40)

$$
\lambda = \bar{u}_0^2 / \sigma_0. 
$$

(41)

$$
\lambda = \bar{u}_0^2 / \sigma_0. 
$$

(42)

In Fig. 2, the implicit procedure of the off-line algorithm is shown.
whose curve is shown in Fig. 3. In step 2, for the normalized \( T_\tau \) and \( A_\tau \), their inverses are roughly approximated by

\[
T_\tau^{-1}(\omega) = \frac{1}{1+2l T_\tau(\omega)} = 1-2A_\tau^{-1}(\omega) \tag{43}
\]

and similarly

\[
A_\tau^{-1}(\omega) = 1-l T_\tau^{-1}(\omega) \tag{44}
\]

where \( A_\tau^{-1}(\omega) \) and \( T_\tau^{-1}(\omega) \) are already normalized as have non-zero constant offset.

Permitting above approximation, it is shown that

\[
||A_\tau^{-1}||=||A_\tau||=||T_\tau^{-1}||=||T_\tau||
\]

and that \( ||A_\tau|| \) converges to zero (See Appendix B).

5. Computer simulations

Several computer simulations are performed, assuming simplified response of unknown system. The response was fixed as

\[
h_1=0.2, \ h_2=0.3, \ h_3=1.0, \ h_4=0.5, \ h_5=0.2
\]

and the number of tap-weights was 7. Time-averaging was calculated over 20000 samples.

Figure 4 shows the curves of \( D_{\tau} \) for the uniform p.d.f., and Figure 5 shows trajectory of \( T_\tau(\omega) \) at every even times of iteration. Similarly, Figures 6 and 7 are for triangle p.d.f. In triangle cases, the degradation of convergence is seen in Fig. 6. The reason of this degradation is imagined that \( T_\tau \) (\( \xi \neq 0 \)) converge too fast to zero and in neighborhood of the solution the solved tap-weights have ambiguity caused by inaccuracy of 20000 times time-averaging. To overcome this, the following polarity detection was employed,

\[
\varphi(y_\omega) = \begin{cases} 1 & \text{if } y_\omega > \Delta \\ 0 & \text{if } \Delta y_\omega > \Delta \\ -1 & \text{if } -\Delta y_\omega
\end{cases} \tag{45}
\]

Its simulation result is shown in Fig. 8. The convergence is more rapid than that of Fig. 6.

6. Conclusion

An off-line algorithm based on idea of removing time-dependency in equalizer output was presented. Through derivations of the algorithms, several problems what is substantial for blind equalization was discussed and some of them was remained for future problem. They are

1. To solve the broad class of p.d.f. in explicit form satisfying the condition (12).
2. To confirm whether maximizing \( \mathbb{E}(y_\omega) \) under the power constraint implies equivalently \( y_\omega \neq 0(\xi \neq 0) \) or not.
3. To give a perfect proof of the convergence of the off-line algorithm for some general cases of p.d.f.
4. To find a general method to accelerate the convergence as tried in Fig. 8.
A METHOD OF BLIND EQUALIZATION

Fig. 8 Evolution of D_t for triangle p.d.f. employing polarity detection Eq. (45)

Appendix A
Let each of the sequence \{a_n\} be independent and uniformly distributed on the interval \([-V, V]\). Assuming \( \gamma_0 = E[\gamma_0] \), then
\[
\gamma_0 = a_0 \text{sign} (a_k t_k)
\]
\[
= \lim_{N \to \infty} \frac{1}{2N+1} \sum_{-N}^{N} a_0 \text{sign} (a_k t_k) \text{d}a_k \text{d}a_N
\]
\[
= \left( \frac{1}{2V} \right)^2 \int_{-V}^{V} \text{d}a_k \int_{0}^{1} a_0 \text{d}a_0
\]
\[
= \frac{V}{2} (1 - \frac{1}{3} D_t)
\]
Similarly,
\[
\gamma_k = \frac{\text{Vr}_k}{3t_0}
\]

Normalizing as \( \gamma_0 = 1 \), we have Eqs. (34) and (35). From these results, we obtain Eqs. (36) and (37) as follows
\[
\gamma_0 = \sum \gamma_k = \sum Vr_k \left( 1 - \frac{1}{3} D_t \right) t_k t_0 + \sum_V \frac{Vr_k}{3t_0} t_k t_0
\]
\[
= \frac{V}{2} (1 - \frac{1}{3} D_t)
\]
and
\[
\gamma_k = \sum a_{k-1} t_k - \gamma_k
\]
\[
= \frac{V}{2t_0} (1 - \frac{1}{3} D_t) t_k + \sum Vr_k \left( 1 - \frac{1}{3} D_t \right) t_k
\]
\[
= \frac{V}{2} (1 - \frac{1}{3} D_t)
\]

Appendix B
Owing to the approximation as Eqs. (43) and (44), in order to show \(|| \Delta \gamma_1 \| > || \Delta \theta_{o} ||\), it is sufficient to prove the inequality,
\[
\frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} |A_r^{-1}(\omega)|^2 d\omega < \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} |T^{-1}(\omega)|^2 d\omega
\]
or more simply,
\[
|A_r^{-1}(\omega)| < |T^{-1}(\omega)| (-\pi/T < \omega < \pi/T)
\]

On the other hand, from Eqs. (34) and (35), we have
\[
A_r(\omega) - 1 = \alpha (T_r(\omega) - 1),
\]
where
\[
\alpha = \frac{2}{3(1 - \frac{1}{3} D_t)}
\]

It is easy to check \( 2/3 < \alpha < 1 \) for \( 0 < D_t < 1 \), under which we discuss the proof of (B-2). Since Eq. (B-2) is equivalent to
\[
\frac{|A_r(\omega)|^2}{A_r(\omega)} < \frac{|T_r(\omega)|^2}{T_r(\omega)}
\]
then, from Eq. (B-3), the inequality to be proved becomes as follows
\[
\alpha^2 |T_r(\omega)|^2 < |A_r(\omega)|^2,
\]
i.e.,
\[
\alpha^2 |T_r(\omega)|^2 < |A_r(\omega)|^2.
\]
Taking \( \beta \neq 0 \)
\[
|T_r(\omega)|^2 = 1 + \beta^2 + 2\beta \cos \theta,
\]
\[
|A_r(\omega)|^2 = 1 + \alpha^2 + 2\alpha \cos \theta
\]
we have
\[
|A_r(\omega)|^2 - \alpha^2 |T_r(\omega)|^2 = 1 - \alpha^2 - 2\alpha (1 - \alpha) \cos \theta.
\]

Since \( \beta \neq 0 \) for \( \gamma_0 > \gamma_1 \), we obtain
\[
|A_r(\omega)|^2 - \alpha^2 |T_r(\omega)|^2 > 1 - \alpha^2 + 2\alpha (1 - \alpha) > 0,
\]
and the proof is completed.

References
