Efficient Tracking of the Cross-Correlation Coefficient

Ronald M. Aarts, *Senior Member, IEEE*, Roy Irwan, and Augustus J. E. M. Janssen

Abstract—In many (audio) processing algorithms, involving manipulation of discrete-time signals, the performance can vary strongly over the repertoire that is used. This may be the case when the signals from the various channels are allowed to be strongly positively or negatively correlated. We propose and analyze a general formula for tracking the (time-dependent) correlation between two signals. Some special cases of this formula lead to classical results known from the literature, others are new. This formula is recursive in nature, and uses only the instantaneous values of the two signals, in a low-cost and low-complexity manner; in particular, there is no need to take square roots or to carry out divisions. Furthermore, this formula can be modified with respect to the occurrence of the two signals so as to further decrease the complexity, and increase ease of implementation. The latter modification comes at the expense that not the actual correlation is tracked, but, rather, a somewhat deformed version of it. To overcome this problem, we propose, for a number of instances of the tracking formula, a simple warping operation on the deformed correlation. Now we obtain, at least for sinusoidal signals, the correct value of the correlation coefficient. Special attention is paid to the convergence behavior of the algorithm for stationary signals and the dynamic behavior if there is a transition to another stationary state; the latter is considered to be important to study the tracking abilities to nonstationary signals. We illustrate tracking algorithm by using it for stereo music fragments, obtained from a number of digital audio recordings.

Index Terms—Audio, cross-correlation coefficient, real-time tracking algorithm, stereophonic signals.

I. INTRODUCTION

WE PROPOSE to use the cross-correlation coefficient in digital audio as a means to acquire statistical information about the input signals with the aim to support the development of audio processing algorithms, for which we envisage numerous applications. We believe that for these algorithms, knowledge of the cross-correlation coefficient is essential to counteract the dependency of their performance on the particular input audio signals. For example, in sound reproduction stereo-base widening systems [1], negative correlation between the audio channels is introduced, while in multichannel audio systems the tracked correlation is used to mix the amount of ambient sounds to the surround channels [2].

Furthermore, correlation techniques are used in room acoustics as a measure for the diffuseness of a sound field [3], as a judgment of the quality of a sound field [4], as a tool in the

field of architectural acoustics [5]–[7], LPC analysis for speech coding [8], time delay of arrival (TDOA) [9], feature detector [10], and system identification [11]. An overview of tracking applications in audio–video object localization is given by [12]. Similar recursions as some of the ones we derive in this paper can also be found in [13]–[19], while [20] gives an overview of these methods.

However, we develop recursions for the cross-correlation coefficient (instead of only the cross-correlation) without utilizing any models while striving for maximum efficiency, by avoiding division, trigonometric operations such as FFT—which also necessitates the use of buffers—and the like. Furthermore we pay special attention to the convergence behavior of the algorithm for stationary signals and the dynamic behavior if there is a transition to another stationary state, the latter is considered to be important to study the tracking abilities to nonstationary signals.

The standard formula for the cross-correlation coefficient between two signals $x = x_l, y = y_l$, with integer time index l

$$\rho = \frac{\frac{1}{K+1} \sum_l (x_l - \overline{x})(y_l - \overline{y})}{\sqrt{\frac{1}{K+1} \sum_l (x_l - \overline{x})^2 \frac{1}{K+1} \sum_l (y_l - \overline{y})^2}} \qquad (1)$$

with $\overline{x}$ and $\overline{y}$ denoting mean values and summations being taken over a segment of length $K + 1$, suffers from the fact that it requires the operations of division and taking a square root. These two operations are unattractive from the point of view of real-time computation, and low-cost implementation. Furthermore, (1) is not optimal for tracking purposes, where the rectangular window of length $K + 1$ is shifted one sample at a time, because of the required administration at the beginning and at the end of the segment.

In Section II, we define the correlation of x and y at time instant k using an exponential window as

$$\rho(k) = \frac{S_{xy}(k)}{(S_{xx}(k)S_{yy}(k))^{1/2}}, \qquad \text{integer } k \qquad (2)$$

where

$$S_{xy}(k) = \sum_{l=0}^{\infty} c \, \mathrm{e}^{-\eta l} x_{k-l} y_{k-l}$$
$$c = 1 - \mathrm{e}^{-\eta} \qquad (3)$$

and S_{xx}, S_{yy} are defined similarly. In (3), we have taken for η a small but positive number that should be adjusted to the particular circumstances for which tracking of the cross-correlation coefficient is required.

S_{xy} can be considered as the output of a first-order low-pass filter with η as the time constant and $x_k y_k$ as input signal. We have omitted in (3), as opposed to (1), the mean values $\overline{x}$ and $\overline{y}$ since these vanish in most audio applications, or simple measures can be taken ensuring that they do vanish. We point out, however, that many of the developments given in this paper apply to the more general case of nonzero, and even time-varying, mean values; see the end of Section II for more details. Furthermore, for stationary signals, the limiting value for $\eta \downarrow 0$ of ρ as given in (2) equals the ρ from the standard formula in (1) when the segment length $K+1$ tends to infinity; see Appendix II for the proof.

We shall show in Section II that the ρ defined in (2) satisfies to a good approximation the recursion

$$
\begin{aligned}
\hat{\rho}(k) &= \hat{\rho}(k-1) + \gamma[\delta_k - \beta_k \hat{\rho}(k-1)], \quad \text{integer } k, \\
\delta_k &= 2x_k y_k \\
\beta_k &= \alpha x_k^2 + \alpha^{-1} y_k^2
\end{aligned}
\tag{4}
$$

where α and γ are determined in a simple fashion by η and the average signal powers of x and y. The actual influence of these average signal powers on the limiting behavior of $\hat{\rho}(k)$ turns out to be rather modest. For small values of γ (or, which is the same, small values of η) these signal powers manifest themselves in the convergence speed, and hence determine the tracking behavior of $\hat{\rho}(k)$, but not in the actual value of $\lim_{k \to \infty} \hat{\rho}(k)$ in the stationary case. Furthermore, in many audio engineering applications, where x and y are, respectively, the sound signals from the left and right channels, one can assume that x and y have equal average signal power, and in that case we have $\alpha = 1$ as we shall see. Hence, in many cases actual knowledge of the signal powers is not necessary since a rough estimate of it is sufficient for getting a good estimate of the limiting behavior of $\hat{\rho}(k)$.

Equation (4) is the basis for our approach of recursively tracking the cross-correlation coefficient. We shall also propose variants of (4), in which the δ_k and β_k are replaced by certain simple transforms, such as their sign and/or modulus. We thus obtain tracking formulas that are even more attractive from a computational point of view. However, by these transformations of δ_k and β_k, the tracking characteristics are changed as well. It may even become an issue "what is being tracked" by the modified (4). In Section V, we propose certain warping operations to correct this situation.

In Section III, we shall analyze the solution of (4), starting from an initial value $\hat{\rho}_0$ at $k = 0$, when $\gamma \downarrow 0$ (that is, when $\eta \downarrow 0$), and we shall indicate conditions under which

$$
\lim_{\gamma \downarrow 0} \left[\lim_{k \to \infty} \hat{\rho}(k) \right] = \rho
\tag{5}
$$

where ρ is the true cross-correlation coefficient defined as in (1) with segment lengths $K + 1$ tending to infinity, and x, y being stationary, zero-mean signals. The details of this analysis are presented in Appendix I where we switch, solely for notational convenience, to the differential equation corresponding to the difference equation in (4). The results of Section III apply equally well to the case that δ_k and β_k are replaced by certain transforms [yielding the variants of (4) mentioned above].

In order to study the difference equation of (4) we could have used the z-transform [21], however to avoid the cumbersome back transformation and to gain insight in the convergence behavior of the recursion we use a different approach. In Section IV, we consider the case of sinusoidal input signals x and y, and we compute explicitly the left-hand side of (5) for the solution of (4) and its variants. It turns out that the unmodified recursion [(4)] yields the correct value ρ for the left-hand side of (5), while some of the aforementioned variants produce certain deformed versions of ρ. The latter effect can be compensated for by applying a simple warping operation on the quantity at the left-hand side of (5). This warped quantity then gives the correct value of the cross-correlation coefficient for the important case of sinusoidal input signals, and it should be expected to yield a considerable enhancement of the performance of our algorithms for many other, nonsinusoidal, input signals. Section V discusses this warping operation in details.

Section VI deals with measuring the step response which is important to study the in-transition phenomenon of the algorithms. This phenomenon occurs simply because the algorithms need a certain time to adapt to sudden (statistical) changes in the input signals.

The proposed algorithms are also tested for audio signals coming from digital audio recordings, and the test results are presented in Section VII.

Finally, conclusions and future work are given in Section VIII.

II. DERIVATION OF TRACKING FORMULAS

In this section, we consider ρ as defined in (2) and (3), and we show that ρ satisfies to a good approximation (when η is small) the recursion in (4) with α and γ given by

$$
\begin{aligned}
\alpha &= \frac{y_{\text{RMS}}}{x_{\text{RMS}}} \\
\gamma &= \frac{c\,e^{\eta}}{2x_{\text{RMS}} y_{\text{RMS}}}.
\end{aligned}
\tag{6}
$$

Here $c = 1 - e^{-\eta}$ as in (3), and the subscripts RMS refer to the root mean-square values of x and y. Furthermore, we modify the recursion in (4) by replacing δ_k and β_k by computationally more attractive quantities. More specifically, we consider the modifications of (4) in which

$$
\beta_k = \begin{cases} 1, & \text{for } \rho_S \\ |y_k|, & \text{for } \rho_R \\ |x_k y_k|, & \text{for } \rho_M \end{cases}
\tag{7}
$$

and

$$
\delta_k = \begin{cases} \text{sign}\{x_k y_k\}, & \text{for } \rho_S \\ y_k\,\text{sign}\{x_k\}, & \text{for } \rho_R \\ x_k y_k, & \text{for } \rho_M \end{cases}
\tag{8}
$$

with the subscripts S, R, M representing the "sign," "relay correlation," and "modulus," respectively.

Equation (4) will lead to the classic sign algorithm in the case for ρ_S, see, e.g., [14]–[16]. We conclude this section by presenting some observations for the case that we have signals x and y that have nonzero mean values which need to be tracked

as well, and for the case that we use rectangular windows instead of exponentially decaying ones.

We now start to show that ρ of (2) and (3) satisfies to a good approximation the recursion in (4) with α and γ given by (6). To this end, we note that

$$S_{xy}(k) = \mathrm{e}^{-\eta} S_{xy}(k-1) + c x_k y_k, \quad \text{integer } k \qquad (9)$$

while similar recursions hold for S_{xx} and S_{yy}. Hence, from the definition in (2)

$$\rho(k) = \frac{S_{xy}(k-1) + c\,\mathrm{e}^{\eta} x_k y_k}{[\{S_{xx}(k-1) + c\,\mathrm{e}^{\eta} x_k^2\}\{S_{yy}(k-1) + c\,\mathrm{e}^{\eta} y_k^2\}]^{1/2}}. \qquad (10)$$

Since we consider small values of η we have that $c = 1 - \mathrm{e}^{-\eta}$ is small as well. Expanding the right-hand side of (10) in powers of c and retaining only the constant and the linear term, we get after some calculations

$$\rho(k) = \rho(k-1) + \frac{c\mathrm{e}^{\eta}}{2(S_{xx}(k-1)S_{yy}(k-1))^{1/2}}$$
$$\cdot \left\{ 2x_k y_k - \left[\left(\frac{S_{yy}(k-1)}{S_{xx}(k-1)} \right)^{1/2} x_k^2 \right. \right.$$
$$\left. \left. + \left(\frac{S_{xx}(k-1)}{S_{yy}(k-1)} \right)^{1/2} y_k^2 \right] \rho(k-1) \right\} + O(c^2). \qquad (11)$$

Then, deleting the $O(c^2)$ term, we obtain the recursion in (4) with α and γ given by (6) when we identify

$$\begin{aligned} x_{\mathrm{RMS}}^2 &= S_{xx}(k) \\ y_{\mathrm{RMS}}^2 &= S_{yy}(k) \end{aligned} \qquad (12)$$

for a sufficiently large k.

We observe at this point that we have obtained the recursion in (4) by applying certain approximations [as in (12)] and neglecting higher order terms. Therefore, it may very well be, that the actual ρ of (2) and (3) and the solution $\hat{\rho}$ of the recursion in (4) do not have very much to do with one another anymore, certainly when k is getting large. In Section III, however, we shall show that $\hat{\rho}$ shares some important properties with the true ρ. In particular, for strictly stationary and ergodic signals x and y the limiting values of $\hat{\rho}$ and ρ for $k \to \infty$ coincide when $\gamma \downarrow 0$ (i.e., $\eta \downarrow 0$). As already said, it is shown in Appendix II that under these circumstances

$$\lim_{\eta \downarrow 0} \rho(k) = \rho \qquad (13)$$

with ρ given by (1) (with $\overline{x} = \overline{y} = 0$) where the segment length $K + 1$ tends to infinity.

Instead of the δ_k and β_k used in (4), we may use the modifications as given by (7) and (8). In all cases, we assume that $\alpha = 1$, which is a reasonable assumption for audio signals, as will be shown in Section VII. This yields $\hat{\rho}_R$, $\hat{\rho}_M$, and $\hat{\rho}_S$, respectively. The advantage of using (7) and (8) is that they are computationally very efficient, while the limiting values for sinusoidal x and y are independent of their amplitude. The tracking behavior, specifically their convergence speed, differs in all cases

as shown in Section IV, but this can be used to detect special effects in the music recording.

We conclude this section with some notes and extensions of our methodology. The first comment deals with the matter of how to handle signals x and y that have nonzero, and actually time-varying, mean values. In those cases, we still define $\rho(k)$ as in (3), however, with the S_{xy} replaced by

$$S_{xy}(k) = \sum_{l=0}^{\infty} c\mathrm{e}^{-\eta l} (x_{k-l} - \overline{x}(k))(y_{k-l} - \overline{y}(k)) \qquad (14)$$

where

$$\overline{x}(k) = \sum_{l=0}^{\infty} c\mathrm{e}^{-\eta l} x_{k-l}$$
$$\overline{y}(k) = \sum_{l=0}^{\infty} c\mathrm{e}^{-\eta l} y_{k-l} \qquad (15)$$

and the S_{xx}, S_{yy} changed accordingly. It can then be shown that

$$\begin{aligned} \overline{x}(k) &= \mathrm{e}^{-\eta} \overline{x}(k-1) + c x_k \\ \overline{y}(k) &= \mathrm{e}^{-\eta} \overline{y}(k-1) + c y_k \end{aligned} \qquad (16)$$

and

$$S_{xy}(k) = \mathrm{e}^{-\eta} S_{xy}(k-1) + c p_k q_k \qquad (17)$$

where $p_k = x_k - \overline{x}(k-1)$, $q_k = y_k - \overline{y}(k-1)$, while similar recursions as in (17) hold for S_{xx}, S_{yy}. This then yields

$$\rho(k) = \frac{S_{xy}(k-1) + c p_k q_k}{[\{S_{xx}(k-1) + c p_k^2\}\{S_{yy}(k-1) + c q_k^2\}]^{1/2}}$$
$$= \rho(k-1) + \frac{c}{2(S_{xx}(k-1)S_{yy}(k-1))^{1/2}}$$
$$\cdot \left\{ 2p_k q_k - \left[\left(\frac{S_{yy}(k-1)}{S_{xx}(k-1)} \right)^{1/2} p_k^2 \right. \right.$$
$$\left. \left. + \left(\frac{S_{xx}(k-1)}{S_{yy}(k-1)} \right)^{1/2} q_k^2 \right] \rho(k-1) \right\} + O(c^2). \qquad (18)$$

From this point onwards, comparing with (11), one can proceed to give many, if not all, of the developments given in this paper for this more general situation.

We have considered thus far exponential windows for the definition of ρ in (2) and (3). We shall now give some observations for the case that we use rectangular windows. With the starting point of the signal segments fixed at $l = 0$, we first consider the ρ in (2) and (3) with

$$S_{xy}(k) = \frac{1}{k+1} \sum_{l=0}^{k} (x_l - \overline{x}(k))(y_l - \overline{y}(k)) \qquad (19)$$

where

$$\overline{x}(k) = \frac{1}{k+1} \sum_{l=0}^{k} x_l$$
$$\overline{y}(k) = \frac{1}{k+1} \sum_{l=0}^{k} y_l \qquad (20)$$

and the S_{xx}, S_{yy} defined similarly. Now the recursions are

$$\overline{x}(k) = \frac{k}{k+1}\,\overline{x}(k-1) + \frac{1}{k+1}\,x_k$$

$$\overline{y}(k) = \frac{k}{k+1}\,\overline{y}(k-1) + \frac{1}{k+1}\,y_k \qquad (21)$$

and

$$S_{xy}(k) = \frac{k}{k+1}\,S_{xy}(k-1) + \frac{k}{(k+1)^2}\,p_k q_k \qquad (22)$$

where $p_k = x_k - \overline{x}(k-1)$, $q_k = y_k - \overline{y}(k-1)$, with similar recursions as in (22) for S_{xx}, S_{yy}. This then leads, as before, to the conclusion that $\rho(k)$ satisfies to a good approximation the recursion

$$\hat{\rho}(k) = \hat{\rho}(k-1) + \gamma\left[2p_k q_k - \left(\alpha p_k^2 + \alpha^{-1} q_k^2\right)\hat{\rho}(k-1)\right] \qquad (23)$$

where

$$\alpha = \left(\frac{S_{yy}(k-1)}{S_{xx}(k-1)}\right)^{1/2}$$

$$\gamma = \frac{1}{2k(S_{xx}(k-1)S_{yy}(k-1))^{1/2}}. \qquad (24)$$

Note that the γ in (24) has a decay like $1/k$ while the γ of (6) is approximately constant (at least when x and y are stationary).

In the case that the starting point of the signal segments is allowed to vary as well, see (19) and (20), the recursion in (21)–(23) also involves sample values at these starting points, and are thus more complicated in nature.

III. ANALYSIS OF THE SOLUTION OF THE BASIC RECURSION

In this section, we consider the basic recursion in (4), and we analyze its solution $\hat{\rho}(k)$, given an initial value $\hat{\rho}_0$ at $k = 0$, when $\gamma \downarrow 0$. Here, we allow δ_k and β_k to be replaced by certain simple transforms such as those required for the definition of $\hat{\rho}_S$, $\hat{\rho}_R$, and $\hat{\rho}_M$ in Section II. Thus, we shall consider the recursion in (4) which we rewrite as

$$\hat{\rho}(k) = (1 - \gamma\beta_k)\hat{\rho}(k-1) + \gamma\delta_k \qquad (25)$$

for $k = 1, 2, \ldots$, with γ a small positive parameter and δ_k, β_k bounded sequences with $0 \le \beta_k \le 1$. By employing the recursion in (25) with $k, k-1, \ldots, 1$, one easily obtains [using $\hat{\rho}(0) = \hat{\rho}_0$]

$$\hat{\rho}(k) = \hat{\rho}_0 \prod_{l=1}^{k}(1 - \gamma\beta_l) + \gamma\sum_{l=1}^{k}\delta_l\prod_{j=l+1}^{k}(1 - \gamma\beta_j). \qquad (26)$$

Now set

$$A(k; \gamma) = \frac{-1}{\gamma}\sum_{l=1}^{k}\log(1 - \gamma\beta_l), \qquad k = 1, 2, \ldots \qquad (27)$$

and $A(0; \gamma) = 0$. Then, we have

$$\hat{\rho}(k) = \hat{\rho}_0 e^{-\gamma A(k;\gamma)} + \gamma\sum_{l=1}^{k}\delta_l e^{-\gamma(A(k;\gamma)-A(l;\gamma))} \qquad (28)$$

for $k = 0, 1, \ldots$.

We compare the formula in (28) for $\hat{\rho}(k)$ with the formula one gets for the solution of the differential equation

$$\hat{\rho}_c'(t) + \gamma\beta(t)\hat{\rho}_c(t) = \gamma\delta(t) \qquad (29)$$

with continuous time variable t. The solution of (29) with initial value $\hat{\rho}_c(0) = \hat{\rho}_0$ follows easily from basic calculus, and is given by

$$\hat{\rho}_c(t) = \hat{\rho}_0 e^{-\gamma A(t)} + \gamma\int_0^t \delta(s)e^{-\gamma(A(t)-A(s))}\,ds, \qquad t \ge 0 \qquad (30)$$

where

$$A(t) = \int_0^t \beta(s)\,ds, \qquad t \ge 0. \qquad (31)$$

While the analysis of the behavior of $\hat{\rho}(k)$ in (26)–(28) and that of $\hat{\rho}(t)$ in (29)–(31) proceed along the same lines, see Appendix I and the results (32)–(38), the analysis of $\hat{\rho}(t)$ is much less cumbersome since we have the elegant framework of integral calculus available here. Moreover, the quantities $M[\delta(t)]$, $M[\beta(t)]$ that appear in (38) and further on throughout the paper, are given in integral form and thus more convenient for computational purposes than the quantities $M[\delta_k]$, $M[\beta_k]$.

In comparing the solutions in (28) and (30) and the corresponding $A(k; \gamma)$ and $A(t)$ in (27) and (31), we consider the β_k and δ_k in the recursion (25) as sampled versions

$$\beta_k = \beta(kT_s)$$

$$\delta_k = \delta(kT_s), \qquad T_s = 1 \qquad (32)$$

of the continuous-time signals $\beta(t)$, $\delta(t)$ in (29) with sample epoch $T_s = 1$. We observe that

$$\lim_{\gamma\downarrow 0} A(k; \gamma) = \sum_{l=1}^{k}\beta_l. \qquad (33)$$

In Appendix I, we shall elaborate on the formulas given in (28) and (30) so as to obtain the limiting behavior of $\hat{\rho}(k)$ as $k \to \infty$, and of $\hat{\rho}_c(t)$ as $t \to \infty$ when $\gamma > 0$ is small. This we do under an assumption (slightly stronger than required) that the mean values

$$b_0(\gamma) = M\left[\frac{-1}{\gamma}\log(1 - \gamma\beta_k)\right]$$

$$= \lim_{K\to\infty}\frac{1}{K}\sum_{l=1}^{K}\frac{-1}{\gamma}\log(1 - \gamma\beta_l)$$

$$d_0 = M[\delta_k] = \lim_{K\to\infty}\frac{1}{K}\sum_{l=1}^{K}\delta_l \qquad (34)$$

for the discrete-time case and

$$b_0 = M[\beta(t)] = \lim_{T\to\infty}\frac{1}{T}\int_0^T \beta(s)\,ds$$

$$d_0 = M[\delta(t)] = \lim_{T\to\infty}\frac{1}{T}\int_0^T \delta(s)\,ds \qquad (35)$$

for the continuous-time case, exist. [Because of the relations in (32) that exist between β_k, δ_k and $\beta(t)$, $\delta(t)$, we have that the

two d_0s in (34) and (35) are equal, while the $b_0(\gamma)$ in (34) tends to b_0 in (35) when $\gamma \downarrow 0$.] We show in Appendix I that under these assumptions for any number $b < b_0(\gamma)$ we have

$$\hat{\rho}(k) = \frac{d_0}{b_0(\gamma)} + O\left(e^{-\gamma bk}\right) + \epsilon_1, \qquad k \geq 0 \qquad (36)$$

for the discrete-time case, while for any number $b < b_0$ we have

$$\hat{\rho}_c(t) = \frac{d_0}{b_0} + O\left(e^{-\gamma bt}\right) + \epsilon_2, \qquad t \geq 0 \qquad (37)$$

for the continuous-time case. In (36) and (37), ϵ_1 and ϵ_2 are quantities that tend to 0 uniformly in $k \geq 0$, $t \geq 0$ when $\gamma \downarrow 0$. Thus formulas (36) and (37) show how the convergence speed can be traded off against accuracy by varying $\gamma > 0$, and this translates naturally into an assessment of the capabilities of our methods.

Since $b_0(\gamma) \to b_0$ as $\gamma \downarrow 0$ we see from (36) and (37) that

$$\lim_{\gamma \downarrow 0} \left[\lim_{k \to \infty} \hat{\rho}(k) \right] = \lim_{\gamma \downarrow 0} \left[\lim_{t \to \infty} \hat{\rho}_c(t) \right]$$

$$= \frac{M[\delta_k]}{M[\beta_k]} = \frac{d_0}{b_0} = \frac{M[\delta(t)]}{M[\beta(t)]}. \qquad (38)$$

This result is basic for the further developments in this paper.

As a consequence of the basic result in (38), we show below that for the particular case

$$\beta_k = \alpha x_k^2 + \alpha^{-1} y_k^2$$
$$\delta_k = 2x_k y_k \qquad (39)$$

see (4), with strictly stationary and ergodic signals x_k, y_k, the left-hand side double-limit in (38) has the correct value

$$\rho = \frac{M[x_k y_k]}{(M[x_k^2] M[y_k^2])^{1/2}}. \qquad (40)$$

Indeed, we have in this case

$$M[\beta_k] = \alpha M[x_k^2] + \alpha^{-1} M[y_k^2]$$
$$M[\delta_k] = 2M[x_k y_k] \qquad (41)$$

and since

$$\alpha = \frac{y_{\text{RMS}}}{x_{\text{RMS}}} = \left(\frac{M[y_k^2]}{M[x_k^2]} \right)^{1/2} \qquad (42)$$

we easily obtain

$$M[\beta_k] = 2 \left(M[x_k^2] M[y_k^2] \right)^{1/2}. \qquad (43)$$

Therefore

$$\lim_{\gamma \downarrow 0} \left[\lim_{k \to \infty} \hat{\rho}(k) \right] = \frac{d_0}{b_0} = \frac{M[\delta_k]}{M[\beta_k]} = \frac{M[x_k y_k]}{(M[x_k^2] M[y_k^2])^{1/2}} \qquad (44)$$

as required.

Similar double-limit relations as in (44) can be obtained for the case that formula in (4) is modified, and in Section IV we shall work this out for sinusoidal signals x and y, and with modified recursion in (4) yielding $\hat{\rho}_S$, $\hat{\rho}_R$, $\hat{\rho}_M$ introduced in Section II.

In case the sample epoch T_s in (32) is not equal to 1, the formulas in (36) and (37) must be changed accordingly. Retaining γ as in (29), the γ in (25)–(28) should be replaced by γT_s, and (36) becomes [with b any number $< b_0(\gamma T_s)$]

$$\hat{\rho}(k) = \frac{d_0}{b_0(\gamma T_s)} + O\left(e^{-\gamma bk T_s}\right) + \epsilon_1, \qquad k \geq 0. \qquad (45)$$

This shows that the time constant for the tracking behavior is given by

$$\tau = \frac{1}{\gamma T_s \, b_0(\gamma T_s)} \qquad (46)$$

in the discrete-time case. With the above choice of retaining γ as in (29), the time-constant for the tracking behavior of $\hat{\rho}_c(t)$ is given by

$$\tau = \frac{1}{\gamma b_0} \qquad (47)$$

since the formula in (37) remains the same with this choice. We finally observe that, since $b_0(\gamma) \to b_0$ as $\gamma \downarrow 0$, the formulas in (46) and (47) for the time constants agree (apart from the factor T_s) asymptotically as $\gamma \downarrow 0$.

IV. SINUSOIDAL INPUT SIGNALS

In this section, we test the algorithms derived in Section II, and analyzed in Section III with respect to their steady state behavior, for sinusoidal input signals. Hence, we take

$$x_k = A_1 \sin(\omega k) \qquad (48)$$

and

$$y_k = A_2 \sin(\omega k + \phi) \qquad (49)$$

where $\omega = 2\pi f$, f is the frequency, and ϕ is an arbitrary phase-shift between the two sinusoids, and A_1 and A_2 are the amplitudes of the two sinusoids with $A_1 \neq A_2$ in general.

The recursion in (4), with β_k given by $\alpha x_k^2 + \alpha^{-1} y_k^2$, involves the ratio α of the RMS values of the input powers of x_k and y_k. For the audio applications we keep in mind that these input powers are not known, but can be assumed to be equal to one another. Therefore, we shall use in (4) with $\alpha = 1$. Evidently, when x_k and y_k are as in (48) and (49) with $A_1 \neq A_2$, we then cannot expect the double limit in (38) to yield the true correlation $\cos \phi$ between x_k and y_k anymore, and also the time constants for the convergence behavior of $\hat{\rho}$ are affected by changing α into 1. Note that the other quantities $\hat{\rho}_S$, $\hat{\rho}_R$, $\hat{\rho}_M$ do not involve α at all, so changing α into 1 is only an issue for $\hat{\rho}$ in (4), and not for its modifications $\hat{\rho}_S$, $\hat{\rho}_R$, $\hat{\rho}_M$.

To relate to the discrete-time signal sampled by f_s, we have now $\omega = 2\pi f_s/m$, where m (integer > 2) is the number of samples in one period of the sine. Using this discrete-time signal, substituting (48) and (49) into (1), and averaging it over an integer number M of periods of the sine where $K + 1 = mM$, we obtain

$$\rho = \cos \phi \qquad (50)$$

which obviously does not depend on A_1 and A_2, but on the phase difference ϕ only.

A. Behavior of $\hat{\rho}_c(t)$ at Sinusoidal Input

If the sinusoidal input signals given by (48) and (49) are used in (29), we have by using (38) with $\alpha = 1$

$$\lim_{\gamma \downarrow 0} \left[\lim_{t \to \infty} \hat{\rho}_c(t) \right] = \frac{2A_1 A_2 \cos \phi}{A_1^2 + A_2^2}. \tag{51}$$

Clearly, in the case $A_1 = A_2$, this simplifies to $\cos \phi$ which is the same result as in (50).

As (36) and (37) show, the deviation from the steady state value depends on γ. Using (47) and $b_0 = (A_1^2 + A_2^2)/2$, it appears that the time constant of the tracking behavior is equal to

$$\tau_1 = 2/(\gamma(A_1^2 + A_2^2)) \tag{52}$$

for the continuous-time case, and

$$\tau_{d1} = 2T_s/(\gamma(A_1^2 + A_2^2)) \tag{53}$$

for the discrete-time case.

B. Behavior of $\hat{\rho}_{S,c}(t)$ at Sinusoidal Input

The approach in Section II can be applied to the stationary solution of ρ_S yielding

$$\lim_{\gamma \downarrow 0} \left[\lim_{t \to \infty} \hat{\rho}_{S,c}(t) \right] = \frac{M[\delta_S(t)]}{M[\beta_S(t)]}. \tag{54}$$

If the sinusoidal input signals given by (48) and (49) are applied to (7), (8), we get using (54)

$$\lim_{\gamma \downarrow 0} \left[\lim_{t \to \infty} \hat{\rho}_{S,c}(t) \right] = \rho_S = \mathrm{H}(\phi) \tag{55}$$

where

$$\mathrm{H}(\phi) = 1 - 2|\phi|/\pi, \qquad |\phi| \le \pi \tag{56}$$

is a periodic function with period 2π and is shown as the dashed line in Fig. 1.

Using (7), (8), (35), and (47), it appears that the time constant of the tracking behavior is equal to

$$\tau_S = 1/\gamma \tag{57}$$

for the continuous-time case, and

$$\tau_{dS} = T_s/\gamma \tag{58}$$

for the discrete-time case, which is clearly independent of A_1 and A_2, as opposed to (53).

C. Behavior of $\hat{\rho}_{R,c}(t)$ at Sinusoidal Input

For the stationary solution of ρ_R, we have from Section II

$$\lim_{\gamma \downarrow 0} \left[\lim_{t \to \infty} \hat{\rho}_{R,c}(t) \right] = \frac{M[\delta_R(t)]}{M[\beta_R(t)]}. \tag{59}$$

If the sinusoidal input signals given by (48) and (49) are applied to (7), (8), we get using (59)

$$\lim_{\gamma \downarrow 0} \left[\lim_{t \to \infty} \hat{\rho}_{R,c}(t) \right] = \cos \phi \tag{60}$$
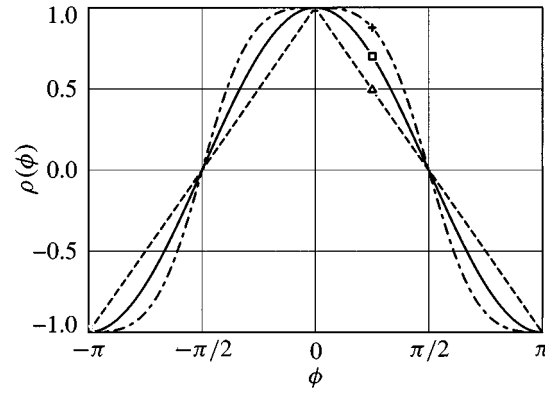
which is the same result as (50).

Fig. 1. Solid line (square marker): ρ, ρ_R, and $\cos \phi$, dashed line (triangular marker): ρ_S, and dashed-dotted line (plus marker): ρ_M.

Using (7) and (8), (35), and (47), we get $b_0 = 2A_2 t\pi$, and therefore

$$\tau_R = \pi/(2A_2\gamma) \tag{61}$$

for the continuous-time case, and

$$\tau_{dR} = \pi T_s/(2A_2\gamma) \tag{62}$$

for the discrete-time case.

D. Behavior of $\hat{\rho}_{M,c}(t)$ at Sinusoidal Input

For the stationary solution of $\hat{\rho}_{M,c}(t)$, we have

$$\lim_{\gamma \downarrow 0} \left[\lim_{t \to \infty} \hat{\rho}_{M,c}(t) \right] = \frac{M[\delta_M(t)]}{M[\beta_M(t)]}. \tag{63}$$

If the sinusoidal input signals given by (48) and (49) are applied to (7) and (8), we get using (63)

$$\lim_{\gamma \downarrow 0} \left[\lim_{t \to \infty} \hat{\rho}_{M,c}(t) \right] = \rho_M = \frac{\cos \phi}{f(\phi)} \tag{64}$$

where

$$f(\phi) = \frac{(\pi - 2|\phi|) \cos \phi}{\pi} + \frac{2 \sin |\phi|}{\pi}. \tag{65}$$

The function $f(\phi)$ is a smooth cosine-like function of ϕ, with period π, and is shown in Fig. 2. It appears that ρ_M behaves similarly as the various other ρs as can be seen in Fig. 1, where (64) is plotted (dashed-dotted line) together with the various other ρs.

Using (7), (8), (31), (47), and (65), we get $b_0 = A_1 A_2 f(\phi) t/2$, and therefore (for an average value of $f \approx 0.8$)
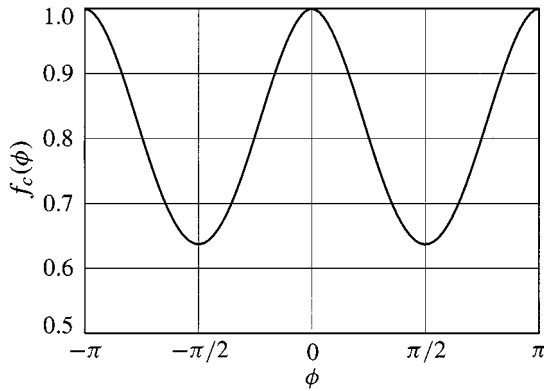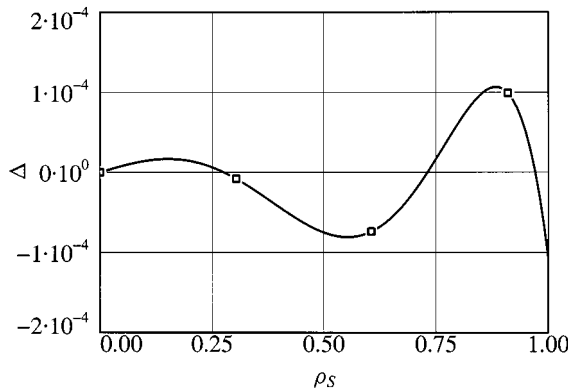
$$\tau_M = 2.5/(\gamma A_1 A_2) \tag{66}$$

for the continuous-time, and

$$\tau_{dM} = 2.5 T_s/(\gamma A_1 A_2) \tag{67}$$

for the discrete-time case.

V. WARPING OF ρ_S AND ρ_M

As Fig. 1 shows, ρ_M and ρ_S are similar but not identical to ρ and ρ_R. We propose to correct this by warping, where one ρ is mapped to another. As an example we warp ρ_S and ρ_M to ρ.

Fig. 2.   Smooth cosine-like function $f(\phi)$.

Fig. 3.   Error $(\Delta = \sin(\pi \rho_S / 2) - \rho_{cS})$ using (69).

A simple polynomial mapping is used for ρ_S and ρ_M. We want to determine a function g such that $g(\rho_S(\phi)) = \cos(\phi)$. Using (56) we get for the corrected ρ_S

$$\rho_{cS} = \sin\left(\frac{\pi}{2}\rho_S\right) \tag{68}$$

where the subscripts cS denotes the corrected version of ρ_S.

This relation was first reported by Van Vleck [22], later in [19], [14]; and Sullivan [20] where Sullivan calls the warping function f in his Table I. For efficiency reasons, the sine function can be approximated [23] yielding to a good approximation

$$\rho_{cS} = \sum_{i=1}^{n} a_i \rho_S^i \tag{69}$$

where for $n = 5$: $a_1 = 1.570\,626\,8$, $a_3 = -0.643\,229\,2$, $a_5 = 0.072\,710\,2$ and $a_2 = a_4 = 0$. Fig. 3 shows the error in this approximation $(\Delta = \sin(\pi \rho_S / 2) - \rho_{cS})$.

For the correction of ρ_M the following polynomial is used:

$$\rho_{cM} = \sum_{i=1}^{n} a_i \rho_M^i \tag{70}$$

where the subscripts cM denotes the corrected version of ρ_M. To maintain the even symmetry of ρ_{cM}, and to ensure that $\rho_{cM}(\pi) = 1$, we require $a_n = \sum_{i=1}^{n-1} a_i$. By means of a least square method we find for $n = 3$: $a_1 = 0.448$ and $a_2 = 1 - a_1$. For $n = 5$ we get $a_1 = 0.7574$, $a_3 = -0.5259$ and $a_5 = -0.7685$, while the as with even index are equal to
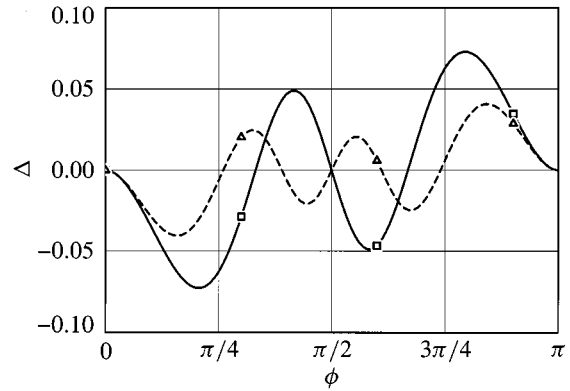
Fig. 4.   Error $(\Delta = \cos \phi - \rho_{cM}(\phi))$ for $n = 3$ (solid line) and $n = 5$ (dashed line) using (70).

zero. Fig. 4 shows the error $(\Delta = \cos \phi - \rho_c(\phi))$ for $n = 3$ (solid line) and $n = 5$ (dashed line) using (70).

We give now some comments on the influence of warping on a step response of ρ_S and ρ_M. If we assume a rising step response of ρ_S as

$$\rho_S = 1 - e^{-t/\tau_S} \tag{71}$$

and apply the warping of (68), we get another time constant given by

$$\tau_{cSr} = -\tau_S \ln\left(1 - \frac{2}{\pi} \arcsin\left(1 - e^{-1}\right)\right) \approx 0.572\,\tau_S. \tag{72}$$

If we assume a decaying step response of ρ_S as

$$\rho_S = e^{-t/\tau_S} \tag{73}$$

and apply the warping of (68) we get yet another time constant given by

$$\tau_{cSd} = -\tau_S \ln\left(\frac{2}{\pi} \arcsin\left(e^{-1}\right)\right) \approx 1.43\,\tau_S. \tag{74}$$

Note that the ratio $\tau_{cSd}/\tau_{cSr} \approx 2.5$, or in other words the rising curve becomes steeper while the decaying curve becomes less steep.

VI. MEASURING STEP RESPONSE

The algorithms for the various ρs were tested by the signals given by (48) and (49) with the following data:

$$\phi = \begin{cases} 0, & t \leq 100 \text{ ms} \\ \dfrac{\pi}{2}, & t > 100 \text{ ms} \end{cases} \tag{75}$$

and

$$A_1 = A_2 = 1; \quad \rho(0) = 0; \quad f_s = 44.1 \text{ kHz}. \tag{76}$$

The values for γ were chosen such that the rise times were approximately equal, as shown in Fig. 5.

It appeared that the time constants (the time to reach the value $1 - e^{-1} \approx 0.632$) are $\tau = 22.5$ ms, $\tau_{cM} = 26.8$ ms, $\tau_{cS} = 21.8$ ms, and $\tau_R = 18.2$ ms. These values correspond well with the values predicted by the formula in Section IV-A. Clearly, we see the slower decay of ρ_{cS} as discussed in Section V.
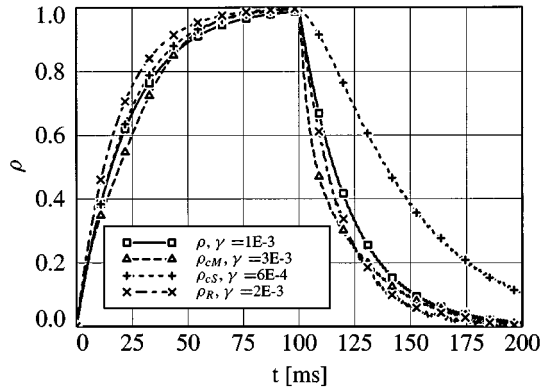
Fig. 5. Step response of the various algorithms for ρ ($f_s = 44.1$ kHz).
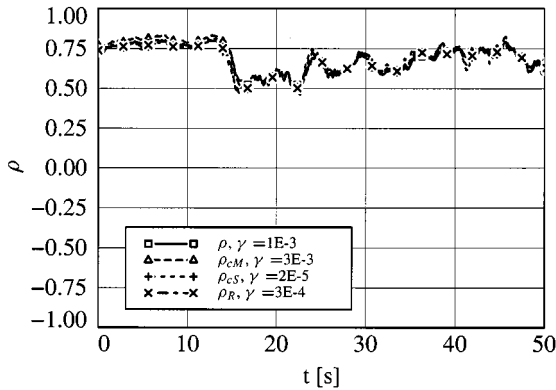

Fig. 6. Tracked cross-correlation coefficients obtained from "The Great Pretender," by Freddy Mercury ($f_s = 44.1$ kHz, α is set to one). Here we see typical behavior of the various ρs, they are all very similar. There are clear stereo effects audible, but there is made only modest use of anti-phase signals.

Fig. 7. Difference between the true ρ given in (10) and the approximated ρs shown in Fig. 6. The time constant η in (10) was 5.10^{-6}. The time constants and labeling for the other four ρs are the same as in Fig. 6. For the readability, the curves of ρ and ρ_R are shifted vertically by 0.075 and -0.075, respectively.

Fig. 8. Tracked cross-correlation coefficients obtained from "Live to Tell," by Madonna. The "Madonna Immaculate Collection" is recorded with special effects which basically widen a degree of stereo. Lower correlations are noticeable in the first 20 s, and even a few negative correlations occur.

VII. APPLICATION TO AUDIO SIGNALS

To demonstrate the behavior of the proposed techniques, consider the following stereo music fragments coming from digital audio signals.

Figs. 6 and 8 show some measurements of the cross-correlation coefficient using the four different algorithms presented in the previous sections. The measurements are shown in squares (ρ), triangle (ρ_{cM}), plus-sign (ρ_{cS}), and cross-sign (ρ_R), respectively. Furthermore, the measurements are performed within a time frame of 50 s. The length of 50 s is chosen such that the algorithms can demonstrate various audio mixes which have been done in a studio sufficiently. The variations in the audio mix can then be seen in the computed correlation coefficients. The parameter γ is set to 10^{-3}, 3×10^{-3}, 2×10^{-5}, and 3×10^{-4}, for the respective ρ, which are determined experimentally to achieve similar tracking behavior as discussed in Section VI.

Fig. 7 shows the error in the approximations for the same fragment and parameters as used for Fig. 6. Comparing each curve in Fig. 7, it can be seen that the difference between the ρ via (10) and the approximated ρ obtained using (4) is the smallest. However, for some particular a-typical cases the results differ slightly more although these are rather sparse; see Fig. 8. It is worth mentioning here that the differences between the algorithms can be used to detect some special effects in the recording. For example in Fig. 8 we see that ρ_{cS} behaves quite differently from other three algorithms in the first 10 s. Listening test confirms
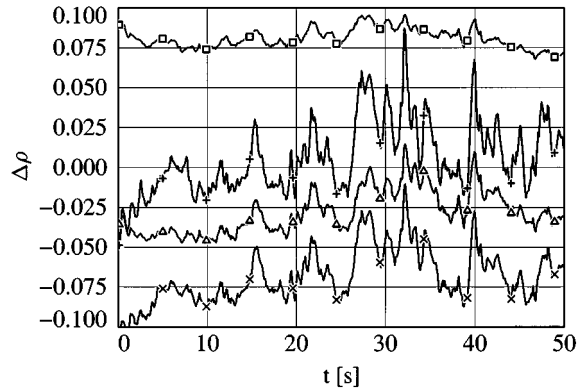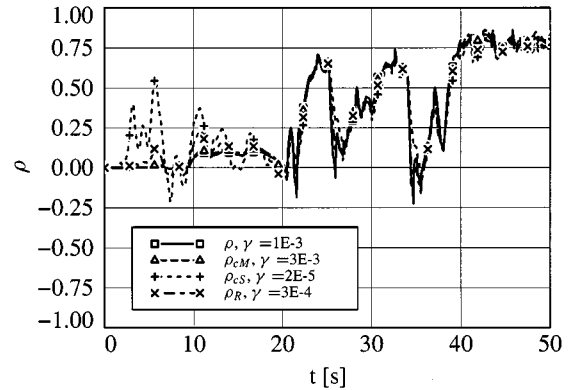
that at this time slot there is music playing at a very low signal level with a strongly varying interchannel phase and balance such that the other three algorithms cannot track these changes. In these cases, it might be beneficial for certain applications to track the difference between ρ_S and one of the three other ρs, in order to detect such special effects in the recording.

VIII. CONCLUSIONS

This paper has presented a formula for tracking the cross-correlation coefficient in real-time, and its modifications to increase ease of implementation. The proposed methods aim at lowering the computational complexity of the formal expression of the cross-correlation coefficient. It has been shown that the proposed methods contain only a few arithmetic operations, and are insensitive to the initial values.

We have formulated necessary and sufficient conditions to examine the behavior of the algorithms using differential equations where the validity of the algorithms have been shown for any nonstationary stochastic input signal. This behavior evaluation has been shown to provide satisfactory accuracy for sinusoidal inputs. Furthermore, the algorithms have also been tested in some music fragments. The derived tracking algorithms for the cross-correlation coefficient give results that strongly agree with what the standard formula ρ would give.

Future research will be directed toward extension to higher-order-statistics where more than two input signals are used.

APPENDIX I
STEADY-STATE BEHAVIOR OF THE SOLUTION OF (29) AS $\gamma \downarrow 0$

In this Appendix, we consider the difference equation

$$\hat{\rho}(k) = (1 - \gamma\beta_k)\hat{\rho}(k-1) + \gamma\delta_k, \qquad k \in \mathbf{Z} \quad (77)$$

and the associated differential equation

$$\hat{\rho}'_c(t) + \gamma\beta(t)\hat{\rho}_c(t) = \gamma\delta(t), \qquad t \in \mathbf{R} \quad (78)$$

where we consider $\beta_k = \beta(k)$, $\delta_k = \delta(k)$ as sampled versions of the continuous-time signals $\beta(t)$ and $\delta(t)$. We are particularly interested in the behavior of $\hat{\rho}(k)$, $\hat{\rho}_c(t)$ as $k \to \infty$, $t \to \infty$ for small $\gamma > 0$. In the analysis that follows, we shall restrict ourselves to the differential equation (78). The derivations for the difference equation (77) follow the same plan, but due to the discretization the developments for (77) have a more cumbersome presentation than those for (78). In the latter case, we can use the elegant framework of integral calculus; this can be mimicked throughout for the discrete-time case without big problems but with awkward presentation. A less trivial difference between the treatment of (77) and (78) is that the basic representations in (28) and (30) involve the quantities $A(k; \gamma)$ and $A(t)$ in (27) and (31), respectively. In particular, $A(k; \gamma)$ depends on γ while $A(t)$ does not. Accordingly, one should replace the bound in (83) on the deviations of $\beta(t)$ from its mean value b_0 by a bound

$$\left| \sum_{j=k-l}^{k} \left(\frac{-1}{\gamma} \log(1 - \gamma\beta_j) - b_0(\gamma) \right) \right|$$
$$\leq B(l+1)^a, \qquad k, l \geq 0 \quad (79)$$

on the deviation of $-\gamma^{-1} \log(1 - \gamma\beta_j)$ from its mean value $b_0(\gamma)$ with $aB > 0$ and $a \in [0, 1)$ that do not depend on $\gamma > 0$. As a consequence, the results for (77) take a somewhat different form as those for (78), and this is carried through in the presentation of the results for (77) in Section II.

For notational simplicity, we omit the symbol $\hat{\ }$ and the subscript c from $\hat{\rho}_c(t)$, and we thus consider the differential equation

$$\rho'(t) + \gamma\beta(t)\rho(t) = \gamma\delta(t), \qquad t \in \mathbf{R}. \quad (80)$$

In (80), the functions $\beta(t)$ and $\delta(t)$ are of the type as those considered in Section II. In particular, we assume that $\beta(t)$ and $\delta(t)$ are well-behaved (smooth) bounded functions for which the mean values

$$b_0 := \mathrm{M}[\beta(t)] = \lim_{T \to \infty} \frac{1}{T} \int_0^T \beta(s)\,\mathrm{d}s > 0 \quad (81)$$

$$d_0 := \mathrm{M}[\delta(t)] = \lim_{T \to \infty} \frac{1}{T} \int_0^T \delta(s)\,\mathrm{d}s \quad (82)$$

exist. In fact, we shall require somewhat more, i.e., the existence of $a \in [0, 1)$ and $B > 0$ such that

$$\left| \int_{t-s}^t (\beta(u) - b_0)\,\mathrm{d}u \right| \leq Bs^a$$
$$\left| \int_{t-s}^t (\delta(u) - d_0)\,\mathrm{d}u \right| \leq Bs^a \quad s, \geq 0. \quad (83)$$

The assumptions embodied by (81)–(83) are satisfied in case that $\beta(t)$ and $\delta(t)$ are bounded periodic functions. In that case, b_0 and d_0 are the DC-values of $\beta(t)$ and $\delta(t)$, and (83) is satisfied with $a = 0$ and

$$B = \int |\beta(u) - b_0|\,\mathrm{d}u$$
$$B = \int |\delta(u) - d_0|\,\mathrm{d}u \quad (84)$$

where the integrals are taken over one period of β, δ. Furthermore, the assumptions are satisfied by realizations $\beta(t)$, $\delta(t)$ of a large class of ergodic, strictly stationary processes, in which case one should expect the parameter a in (83) to be positive, typically 0.5 or slightly larger.

At the end of this Appendix, we present a connection between our results proven below, and Wiener's Tauberian theorem. For application of the latter theorem it is sufficient to only assume the existence of the two limits in (81) and (82). However, for validity of the double limit relation in (38), a certain control of the deviation of $\rho(t)$ from d_0/b_0 is necessary, and this is guaranteed when (83) is satisfied.

We now show that

$$\rho(t) = \gamma \int_0^\infty \delta(t-s)\mathrm{e}^{-\gamma b_0 s}\,\mathrm{d}s + O\left(\mathrm{e}^{-\gamma bt}\right) + O\left(\gamma^{1-a}\right) \quad (85)$$

where b is any number between 0 and b_0, and the last O-term is uniform in $t \geq 0$. Furthermore, we show that, uniformly in $t \geq 0$

$$\gamma \int_0^\infty \delta(t-s)\mathrm{e}^{-\gamma b_0 s}\,\mathrm{d}s = \frac{d_0}{b_0} + O\left(\gamma^{1-a}\right) \quad (86)$$

which establishes the results required in Section III.

The proof of these results are now presented. The solution $\rho(t)$ of (80) is given explicitly by

$$\rho(t) = \rho_0\mathrm{e}^{-\gamma A(t)} + \gamma \int_0^t \delta(s)\mathrm{e}^{-\gamma(A(t)-A(s))}\,\mathrm{d}s, \qquad t \in \mathbf{R}. \quad (87)$$

Here, we denote $\rho_0 = \rho(0)$ and

$$A(t) = \int_0^t \beta(s)\,\mathrm{d}s, \qquad t \in \mathbf{R}. \quad (88)$$

We analyze the right-hand side of (87) as $t \to \infty$. By the assumptions on β, we have that

$$A(t) \geq b_0 t - Bt^a, \qquad t \geq 0. \quad (89)$$

Hence, since $0 \leq a < 1$, we have that

$$\left| \rho_0\mathrm{e}^{-\gamma A(t)} \right| \leq |\rho_0|\mathrm{e}^{-\gamma b_0 t + B\gamma t^a} = O\left(\mathrm{e}^{-\gamma bt}\right) \quad (90)$$

for any $b \in (0, b_0)$.

As to the second term at the right-hand side of (87), we write

$$\gamma \int_0^t \delta(s) e^{-\gamma(A(t)-A(s))} \, ds$$

$$= \gamma \int_0^\infty \delta(t-s) e^{-\gamma(A(t)-A(t-s))} \, ds$$

$$- \gamma \int_t^\infty \delta(t-s) e^{-\gamma(A(t)-A(t-s))} \, ds. \quad (91)$$

For the second term at the right-hand side of (91) we observe that

$$A(t) - A(t-s) = \int_{t-s}^t \beta(u) \, du$$

$$= b_0 s + \int_{t-s}^t (\beta(u) - b_0) \, du$$

$$\geq b_0 s - B s^a, \qquad s \geq t. \quad (92)$$

Therefore, again since $0 \leq a < 1$

$$\left| \gamma \int_t^\infty \delta(t-s) e^{-\gamma(A(t)-A(t-s))} \, ds \right|$$

$$\leq \gamma \max|\delta| \int_t^\infty e^{-\gamma b_0 s + \gamma A s^a} \, ds = O\left(e^{-\gamma b t}\right) \quad (93)$$

for any $b \in (0, b_0)$, and $t \geq 0$.

Combining (90) and (93), we have now established that

$$\rho(t) = \gamma \int_0^\infty \delta(t-s) e^{-\gamma(A(t)-A(t-s))} \, ds + O\left(e^{-\gamma b s}\right),$$

$$t \geq 0 \quad (94)$$

for any $b \in (0, b_0)$. Next, we consider the first term at the right-hand side of (94). From (92), we have

$$\gamma \int_0^\infty \delta(t-s) e^{-\gamma(A(t)-A(t-s))} \, ds$$

$$= \gamma \int_0^\infty \delta(t-s) e^{-\gamma b_0 s} \, ds$$

$$+ \gamma \int_0^\infty \delta(t-s) e^{-\gamma b_0 s} \left\{ e^{-\gamma \int_{t-s}^t (\beta(u)-b_0) \, du} - 1 \right\} ds. \quad (95)$$

The second term at the right-hand side of (95) can be estimated as

$$\gamma \max|\delta| \int_0^\infty e^{-\gamma b_0 s} \left\{ e^{B\gamma s^a} - 1 \right\} ds \quad (96)$$

by the assumptions on β. By partial integration and the substitution $v = \gamma b_0 s$ we have

$$\gamma \int_0^\infty e^{-\gamma b_0 s} \left\{ e^{B\gamma s^a} - 1 \right\} ds$$

$$= \frac{-1}{b_0} \int_0^\infty \left\{ e^{B\gamma s^a} - 1 \right\} d\left(e^{-\gamma b_0 s}\right)$$

$$= D\gamma^{1-a} a \int_0^\infty v^{a-1} e^{-v} e^{\epsilon v^a} \, dv \quad (97)$$

where $D = B b_0^{-1-a}$ and $\epsilon = B\gamma^{1-a} b_0^{-a}$. Since $0 \leq a < 1$, the integral on the last line of (97) remains bounded when $\gamma \downarrow 0$,

and thus we see that the second term in the right-hand side of (95) is $O(\gamma^{1-a})$. Combining (94) and (95) we then obtain (85).

We finally show the validity of (86) by fixing $\gamma > 0$, and computing

$$\gamma \int_0^\infty \delta(t-s) e^{-\gamma b_0 s} \, ds = \frac{d_0}{b_0} + \gamma \int_0^\infty (\delta(t-s) - d_0) e^{-\gamma b_0 s} \, ds. \quad (98)$$

For the integral at the right-hand side of (98) we have by partial integration

$$\gamma \int_0^\infty (\delta(t-s) - d_0) e^{-\gamma b_0 s} \, ds$$

$$= \gamma \int_0^\infty e^{-\gamma b_0 s} \, d\left[\int_{t-s}^t (\delta(u) - d_0) \, du \right]$$

$$= \gamma e^{-\gamma b_0 s} \int_{t-s}^t (\delta(u) - d_0) \, du \Big|_0^\infty$$

$$+ \gamma^2 b_0 \int_0^\infty e^{-\gamma b_0 s} \left[\int_{t-s}^t (\delta(u) - d_0) \, du \right] ds$$

$$= \gamma^2 b_0 \int_0^\infty e^{-\gamma b_0 s} \left[\int_{t-s}^t (\delta(u) - d_0) \, du \right] ds. \quad (99)$$

To estimate the integral expression at the far right-hand side of (99) we use (83) to obtain uniformly in $t \geq 0$

$$\left| \gamma^2 b_0 \int_0^\infty e^{-\gamma b_0 s} \left[\int_{t-s}^t (\delta(u) - d_0) \, du \right] ds \right|$$

$$\leq B\gamma^2 b_0 \int_0^\infty e^{-\gamma b_0 s} s^a \, ds$$

$$= B b_0^{-a} \Gamma(1 + a) \gamma^{1-a}. \quad (100)$$

Hence, from (98)–(100) we obtain uniformly in $t \geq 0$, that

$$\gamma \int_0^\infty \delta(t-s) e^{-\gamma b_0 s} \, ds = \frac{d_0}{b_0} + O\left(\gamma^{1-a}\right) \quad (101)$$

which finally establishes (86).

In Appendix II we show, using Wiener's Tauberian theorem, the following. Let $h(s)$ be a bounded integrable function defined for $s \geq 0$. If one of the limits

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T h(s) \, ds$$

$$\lim_{\epsilon \downarrow 0} \epsilon \int_0^\infty h(s) e^{-\epsilon s} \, ds \quad (102)$$

exists, then so does the other with the same value. With $t \geq 0$ and $h(s) = \delta(t-s)$ it follows from

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T \delta(t-s) \, ds = d_0 \quad (103)$$

that

$$\lim_{\gamma \downarrow 0} \gamma \int_0^\infty \delta(t-s) e^{-\gamma b_0 s} \, ds = \frac{d_0}{b_0}. \quad (104)$$

The result in (101) gives somewhat more since we have uniformity in $t \geq 0$ and more precise information as to how fast the limit in (104) is approached. This comes at the expense of the additional assumption given in (83) that we had to make.

APPENDIX II
EXPONENTIAL WINDOWING, RECTANGULAR WINDOWING, AND WIENER'S TAUBERIAN THEOREM

In this Appendix, we show that for two bounded, strictly stationary, and ergodic discrete-time processes x_k and y_k the definition in (1) of ρ, based on rectangular windowing with segment length $K+1$ tending to infinity, and the definition in (2) and (3) of $\rho(k)$, based on exponential windowing with decay parameter $\eta \downarrow 0$, gives the same value for the cross-correlation coefficient. Evidently, by stationarity we may assume that $\overline{x} = \overline{y} = 0$. Also, in (3), we may assume that $k = 0$, and we may replace the $-l$ in $x_{-l}y_{-l}$ by l. Thus the equivalence of either definition will be proved when we can show the following. Let h_k, $k = 0, 1, \ldots$ be a bounded sequence. If one of the limits

$$\lim_{K \to \infty} \frac{1}{K+1} \sum_{l=0}^{K} h_l$$

$$\lim_{\eta \downarrow 0}(1 - e^{-\eta}) \sum_{l=0}^{\infty} e^{-\eta l} h_l \qquad (105)$$

exists, then so does the other, and it has the same value. Indeed, when we take $h_l = x_l y_l$, $|x_l|^2$, $|y_l|^2$, we see that any of the three quantities in the denominator and numerator of (1) has the same limit as $K \to \infty$ as the corresponding quantity in (3) as $\eta \downarrow 0$, and vice versa.

The result concerning the two limits in (105) is a well-known example of a Tauberian theorem. It is a consequence of the continuous-time Tauberian theorem that we already announced at the end of Appendix I. Indeed, when we choose in (102) for $h(s)$ the step function that assumes the value h_l on $[l, l+1)$, $l = 0, 1, \ldots$ and let $T \to \infty$ through integer values we easily get the result concerning the two limits in (105). Here it should also be noted that $1 - e^{-\eta} = \eta + O(\eta^2)$.

The proof concerning the two limits in (102) can be given by using Wiener's Tauberian theorem, see [24, Th. 4, pp. 73–74]. Given two absolutely integrable functions $K_1(x)$, $K_2(x)$ defined on $\mathbf{R}$ with unit integral and with Fourier transforms that do not vanish for real argument, [24, Th. 4] states the following. When $f(y)$, $y \in \mathbf{R}$, is a bounded function and if one of the limits

$$\lim_{x \to \infty} \int_{-\infty}^{\infty} K_1(x - y)f(y)\,\mathrm{d}y$$

$$\lim_{z \to \infty} \int_{-\infty}^{\infty} K_2(z - y)f(y)\,\mathrm{d}y \qquad (106)$$

exists, then so does the other with the same value. Taking

$$K_1(x) = \begin{cases} e^{-x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \qquad (107)$$

$$K_2(x) = e^{-x-e^{-x}}, \qquad x \in \mathbf{R} \qquad (108)$$

in [24, Th. 4], and

$$f(y) = h(e^y),$$
$$x = \log T,$$
$$z = -\log \epsilon \qquad (109)$$

while substituting $s = e^y$, we see that the two limits in (106) turn into the limits in (102). Hence we only have to show that the K_1 and K_2 in (107) and (108) have Fourier transforms that do not vanish for real argument. We compute

$$\int_{-\infty}^{\infty} e^{-2\pi ivx} K_1(x)\,\mathrm{d}x = \frac{1}{1 + 2\pi ivx}$$

$$\int_{-\infty}^{\infty} e^{-2\pi ivx} K_2(x)\,\mathrm{d}x = \Gamma(1 + 2\pi ivx) \qquad (110)$$

and none of these functions vanish for a real value of v.

REFERENCES

[1] R. M. Aarts, "Phantom sources applied to stereo-base widening," *J. Audio Eng. Soc.*, vol. 48, no. 3, pp. 181–189, Mar. 2000.

[2] R. Irwan and R. M. Aarts, "A method to convert stereo to multi-channel sound," in *Proc. AES 19th Int. Conf.*, Schloss Elmau (Kleis), Germany, June 21–24, 2001, pp. 139–143.

[3] R. K. Cook, R. V. Waterhouse, R. D. Berendt, S. Edelman, and M. C. Thompson, "Measurement of correlation coefficients in reverberant sound fields," *J. Acoust. Soc. Amer.*, vol. 27, no. 6, pp. 1072–1077, Nov. 1955.

[4] K. Kurozumi and K. Ohgushi, "The relationship between the cross-correlation coefficient two-channel acoustic signals and sound image quality," *J. Acoust. Soc. Amer.*, vol. 74, no. 6, pp. 1726–1733, Dec. 1983.

[5] M. Barron, "The effect of first reflections in concert halls—The need for lateral reflections," *J. Sound Vibr.*, vol. 15, no. 4, pp. 475–494, 1971.

[6] A. Czyzewski, "A method of artificial reverberation quality testing," *J. Audio Eng. Soc.*, vol. 38, no. 3, pp. 129–141, March 1990.

[7] O. Lundén and M. Bäckström, "Stirrer efficiency in FOA reverberation chambers: Evaluation of correlation coefficients and chi-squared tests," in *IEEE Int. Symp. Electromagnetic Compatibility*, vol. 1, 2000, pp. 11–16.

[8] T. P. Barnwell, "Recursive autocorrelation computation for LPC analysis," in *Proc. 1977 IEEE Int. Conf. Acoustics, Speech, Signal Processing*, May 9–11, 1977, pp. 1–4.

[9] G. C. Carter, Ed., *Coherence and Time Delay Estimation*. New York: IEEE Press, 1992.

[10] J. P. Lewis, "Fast template matching," *Vision Interface*, pp. 120–123, 1995. (An update of this paper "Fast normalized cross-correlation" is available [Online] http://www.idiom.com/zilla/Work/nvisionInterface/).

[11] L. Ljung, *System Identification, Theory For the User*. Englewood Cliffs, NJ: Prentice-Hall, 1999.

[12] N. Strobel, S. Spors, and R. Rabenstein, "Joint audio–video object localization and tracking," *IEEE Signal Processing Mag.*, vol. 18, pp. 22–31, Jan. 2001.

[13] D. Hertz, "A fast digital method of estimating the autocorrelation of a Gaussian stationary process," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, p. 329, Apr. 1982.

[14] K. J. Gabriel, "Comparison of three correlation coefficient estimators for Gaussian stationary processes," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 1023–1025, Aug. 1983.

[15] G. Jacovitti and R. Cusani, "An efficient technique for high correlation estimation," *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. ASSP-35, pp. 654–660, May 1987.

[16] R. Cusani and A. Neri, "A modified hybrid sign estimator for the normalized autocorrelation function of a Gaussian stationary process," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 1321–1324, Oct. 1985.

[17] T. Koh and E. J. Powers, "Efficient methods to estimate correlation functions of Gaussian stationary processes and their performance analysis," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 1032–1035, Aug. 1985.

[18] G. Jacovitti and R. Cusani, "Performance of the hybrid-sign correlation coefficient estimator for Gaussian stationary processes," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 731–733, June 1985.

[19] S. S. Wolff, J. B. Thomas, and T. R. Williams, "The polarity-coincidence correlator: A nonparametric detection device," *IRE Trans. Inform. Theory*, vol. IT-8, pp. 5–9, Jan. 1962.

[20] M. C. Sullivan, "Efficient autocorrelation estimation using relative magnitudes," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 445–447, Mar. 1989.

[21] R. P. Agarwal, *Difference Equations and Inequalities. Theory, Methods, and Applications*. New York: Dekker, 1991.

[22] J. H. van Vleck and D. Middleton, "The spectrum of clipped noise," *Proc. IEEE*, vol. 54, pp. 2–19, Jan. 1966.

[23] C. Hastings, *Approximations for Digital Computers*. Princeton, NJ: Princeton Univ. Press, 1955.

[24] N. Wiener, *The Fourier Integral and Certain of Its Applications*. New York: Dover, 1933.

Roy Irwan received the M.Sc. degree from Delft University of Technology, Delft, The Netherlands, in 1992, and the Ph.D. degree from the University of Canterbury, Christchurch, New Zealand in 1999, both in electrical engineering.

From 1993 to 1995, he was employed as a System Engineer at NKF b.v., where he was primarily involved in installation of fiber optics cables. In 1999, he joined the Digital Signal Processing Group at Philips Research Laboratories, Eindhoven, The Netherlands. He has published a number of refereed papers in international journals. His research interests include digital signal processing, image processing, optics, and inverse problems.

Ronald M. Aarts (SM'95) was born in 1956, in Amsterdam, The Netherlands. He received the B.Sc. degree in electrical engineering in 1977 and the Ph.D. degree from Delft University of Technology, Delft, The Netherlands, in 1995.

In 1977, he joined the Optics Group of Philips Research Laboratories, Eindhoven, The Netherlands, where he was engaged in research into servos and signal processing for use in both video long play players and compact disc players. In 1984, he joined the Acoustics Group of the Philips Research Laboratories and was engaged in the development of CAD tools and signal processing for loudspeaker systems. In 1994, he became a member of the DSP Group of the Philips Research Laboratories where he was engaged in the improvement of sound reproduction, by exploiting DSP and psycho-acoustical phenomena. He has published more than 100 papers and reports and is the holder of over a dozen U.S. patents in the aforementioned fields. He was a member of organizing committees and chairman for various conventions.

He is a fellow of the Audio Engineering Society, the Dutch Acoustical Society, and the Acoustical Society of America.

Augustus J. E. M. Janssen was born in 1953. He received the Eng. and Ph.D. degrees in mathematics from the Eindhoven University of Technology, Eindhoven, The Netherlands, in October 1976 and June 1979, respectively.

From 1979 to 1981, he was a Bateman Research Instructor in the Mathematics Department at the California Institute of Technology, Pasadena. In 1981, he joined Philips Research Laboratories, Eindhoven, where his principal responsibility is to provide high-level mathematical service and consultancy in mathematical analysis. His research interest is in Fourier analysis with emphasis on time-frequency analysis, in particular Gabor analysis. His current research interests include the Fourier analysis of nonlinear devices such as quantizers. He has published 95 papers in the fields of signal analysis, mathematical analysis, Wigner distribution and Gabor analysis, information theory, and electron microscopy. He has also published 35 internal reports and holds five U.S. patents.

Dr. Janssen received the prize for the best contribution to the Mathematical Entertainments column of the *Mathematical Intelligencer* in 1987 and the EURASIP's 1988 Award for the Best Paper of the Year in Signal Processing.