

ZZT-domain Immiscibility of the Opening and Closing Phases of the LF GFM under Frame Length Variations

C.F. Pedersen, O. Andersen, P. Dalsgaard

Department of Electronic Systems, Aalborg University, Aalborg, Denmark

{cfp, oa, pd}@es.aau.dk

Abstract

Current research has proposed a non-parametric speech waveform representation (rep) based on zeros of the z-transform (ZZT) [1] [2]. Empirically, the ZZT rep has successfully been applied in discriminating the glottal and vocal tract components in pitch-synchronously windowed speech by using the unit circle (UC) as discriminant [1] [2]. Further, similarity between ZZT reps of windowed speech, glottal flow waveforms, and waveforms of glottal flow opening and closing phases has been demonstrated [1] [3]. Therefore, the underlying cause of the separation on either side of the UC can be analyzed via the individual ZZT reps of the opening and closing phase waveforms; the waveforms are generated by the LF glottal flow model (GFM) [1]. The present paper demonstrates this cause and effect analytically and thereby supplement the previous empirical works. Moreover, this paper demonstrates that immiscibility is variant under changes in frame lengths; lengths that maximize or minimize immiscibility are presented.

Index Terms: Zeros of the z-transform, LF glottal flow model, opening/closing phase separation

1. Introduction

A non-parametric speech signal representation (rep) dubbed ZZT (zeros of the z-transform) has been proposed in recent studies [1] [2] [4]. In the ZZT rep, the signal sequence represent a coefficient sequence in a real univariate polynomial, and the ZZT rep is obtained by estimating the zeros of this polynomial function.

Empirical studies on source/filter separation by ZZT, demonstrate that the ZZT rep related to the glottal flow opening phase lie outside the unit circle (UC), and the ZZT rep related to the vocal tract filter and glottal flow closing phase lie inside the UC when a suitable pitch synchronous windowing technique is employed [1] [4]. With the windowing technique, ZZT reps of speech waveforms are similar to ZZT reps of glottal flow waveforms alone and the separation around the UC in the complex plane remains [1]. Moreover, similarity between ZZT reps of entire glottal flow waveforms and ZZT reps of their individual glottal flow opening and closing phases have been demonstrated [3]. Therefore, analysis of the underlying cause of the source/filter separation on either side of the UC can be conducted on the individual ZZT reps of the opening and closing phases. In the studies [1] and [2], the non-interactive parametric LF glottal flow model (GFM) [5] was used to generate the glottal flow waveforms. In order to extend the previous studies, the LF GFM is also used in the present paper.

In the present paper, it is analytically demonstrated that the ZZT rep of the LF GFM opening phase lie outside the UC and the ZZT rep of the closing phase lie inside the UC. This prop-

erty is dubbed *immiscibility* as the ZZT reps are disjoint sets in the complex plane. Further, this paper demonstrates that the ZZT domain immiscibility is variant under changes in the frame length, i.e. the z-transform coefficient sequence length, N . Values of N that maximize or minimize immiscibility are found.

The remainder of this paper is organized by describing the fundamentals of the study in section 2; section 3 presents an analysis of immiscibility as a function of the z-transform coefficient sequence length, and in section 4 the analysis is exemplified. The results in section 5 present coefficient sequence lengths that maximize or minimize immiscibility, and in the closing section, section 6, the results are discussed along with future perspectives.

2. Fundamentals

2.1. LF glottal flow model

The LF GFM is defined by the derivative of the glottal flow as,

Definition 1 LF glottal flow model [5]

$$e(t) = \begin{cases} E_0 e^{\alpha t} \sin(\omega_g t), & t_0 \leq t \leq t_e \\ -\frac{E_e}{\epsilon t_a} \left(e^{-\epsilon(t-t_e)} - e^{-\epsilon(t_c-t_e)} \right), & t_e < t \leq t_c \\ 0, & t_c < t \leq T \end{cases}$$

Figure 1 illustrates a LF GFM waveform period, T , along with the commonly associated temporal landmarks.

In def. 1, the LF GFM is defined by three constituents: (i) an opening phase, $e_o(t) = e(t)$ for $t_0 \leq t \leq t_e$, (ii) a closing phase, $e_c(t) = e(t)$ for $t_e < t \leq t_c$, and (iii) a shut phase, $e_s(t) = e(t)$ for $t_c < t \leq T$.

The opening phase models the opening of glottis from t_0 , the glottal opening instant (GOI), to t_e , the glottal closure instant (GCI) where the waveform yields its maximum negative amplitude, $-E_e$. The parameter $t_p = \pi/\omega_g$ marks the location of maximum glottal airflow.

The closing phase models the closing of glottis after the airflow termination at GCI, i.e. at t_e , to total closure at t_c . The parameter t_a determines the duration from t_e to the zero crossing of the tangent to $e_c(t_e)$, i.e. t_a is a measure of glottal closure abruptness. Finally, the shut phase simply models total closure of the glottal folds from the instant t_c to t_0' .

2.2. Zeros of the z-transform

In accordance with [1] [2] [4], the ZZT rep is defined as an all-zero rep of the z-transform of a signal sequence, i.e.

Definition 2 Zeros of the z-transform

The zeros of the z-transform of a sequence $(x_n)_{n=0}^{N-1} \subset \mathbb{R}$ are

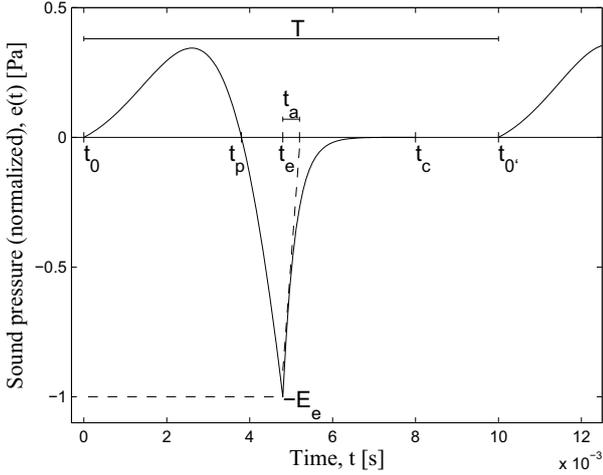


Figure 1: A LF GFM waveform period, T , with the commonly associated temporal landmarks.

defined as $z_1, z_2, \dots, z_m \in \mathbb{C} \setminus \{0\}$ such that $X(z_i) = \sum_{n=0}^{N-1} x_n z_i^{-n} = 0$ for $1 \leq i \leq m$.

Generally, a z-transform can be factored as,

$$X(z) = \sum_{n=0}^{N-1} x_n z^{-n} = \frac{x_0 \prod_{m=1}^{N-1} (z - z_m)}{z^{(N-1)}}$$

Provided $x_0 \neq 0$.

That is, the ZZT is an unordered sequence of the zeros of the assumed polynomial function in the numerator deducted by any poles, i.e. zeros at zero in this case, as these lead to an undefined z-transform.

The ZZT-transformation is denoted $\rho : \mathbb{R} \mapsto \mathbb{C}$, $\rho((x_n)_{n=1}^N) = (z_m)_{m=1}^{N-1-k}$, where x is a polynomial coefficient sequence ordered in descending powers, z is a sequence of non-zero zeros and k is the multiplicity of a zero at zero. The inverse ZZT transformation is either a true inverse (up to order and scaling), ρ^{-1} , if $k = 0$ as ρ becomes bijective, or an inverse relation, ρ_{ir} , if $k \neq 0$ as ρ becomes non-injective.

2.3. ZZT representation of opening and closing phase

Let $p(a, z)$ denote a univariate polynomial of $(N - 1)$ 'th degree with variable $z \in \mathbb{C}$ and the sequence of coefficients $(a_n)_{i=0}^{N-1} \subset \mathbb{R}$ ordered in descending powers. If $e_o(t)$ is discretized into the sequence $eo = (e_{o_n})_{n=0}^{N-1}$, the ZZT rep becomes

Expression 1 ZZT of LF GFM opening phase [3]

$$\begin{aligned} \text{ZZT} : z_m &= e^\alpha \rho(x_p), z_m \neq 0, e^{\alpha \pm ik}, m \in [1; N - 2] \\ \text{where} \\ p(x_p, z) &= \sin(k)z^N - \sin(kN)z + \sin(k(N - 1)), \\ k &= \omega_g = \pi/t_p \end{aligned}$$

Similarly, if $e_c(t)$ is discretized into the sequence $ec = (e_{c_n})_{n=0}^{N-1}$, the ZZT rep becomes

Expression 2 ZZT of LF GFM closing phase [3]

$$\begin{aligned} \text{ZZT} : z_m &= \rho(x_p), z_m \neq 0, 1, e^{-k}, m \in [1; N - 1] \\ \text{where} \\ p(x_p, z) &= (c_1 - c_2)z^{N+1} + (c_2e^{-k} - c_1)z^N + \\ &\quad (c_2 - c_1e^{-kN})z + (c_1e^{-kN} - c_2e^{-k}), \\ c_1 &= e^{\epsilon t_e}, c_2 = e^{-\epsilon(t_c - t_e)}, k = \epsilon \end{aligned}$$

In order to demonstrate the property of immiscibility of the ZZT reps in expressions 1 and 2, theorem 1, 2 and 3 are introduced as they state the modulus bounds of polynomial zeros.

Theorem 1 Cauchy bound [6]

All zeros of a complex polynomial,

$$p(a, z) = z^n + \sum_{k=0}^{n-1} a_k z^k$$

lie in the disk $|z| < \lambda$ where $\lambda = 1 + \max_{0 \leq k \leq n-1} \{|a_k|\}$

Theorem 2 Cauchy bounded annulus [7]

Let $p(a, z)$ be a polynomial with zeros z_1, \dots, z_m ordered as $0 < |z_1| \leq \dots \leq |z_m|$. Let λ^* denote the Cauchy bound of $p(a, z)$ and λ_* the Cauchy bound of $z^m p(a, 1/z)$. Then the following inequalities hold,

$$\frac{1}{\lambda_*} \leq |z_1| \leq \frac{1}{(2^{1/m} - 1)\lambda_*} \text{ and } (2^{1/m} - 1)\lambda_* \leq |z_m| \leq \lambda^*$$

Theorem 3 and 1 are equivalent, but theorem 3 does sometimes yield a tighter bound [8] which is the case in this paper.

Theorem 3 Alternative Cauchy bound [8]

All zeros of a n 'th degree complex polynomial,

$$p(a, z) = z^n + \sum_{k=0}^{n-1} a_k z^k$$

lie in the disk $|z| \leq \lambda_a$ where $\lambda_a = \max \{1, \sum_{i=0}^{n-1} |a_i|\}$

The subscript a on λ denotes the alternative Cauchy bound.

3. Immiscibility Invariance

In this section, the Cauchy bounds, λ , of the opening and closing phases are analysed. In order to investigate the immiscibility invariance under variations in polynomial coefficient sequence lengths, N , the bounds are expressed as functions of N . In the analysis $N \in \mathbb{R}_{++}$ although $N \in \mathbb{Z}_{++}$ in theorem 1, 2, and 3. Letting N be a non-integer real must only be considered as a reasonable theoretical interpolation.

3.1. Analysis of Cauchy bound of opening phase

The lowest Cauchy bound for an annulus containing all zeros is $\lambda_*^{-1}(N)$ (cf. th. 2). If $\lambda_*^{-1}(N) > 1$ all zeros lie outside the UC. As the e^α factor is just a real scaling of the zeros of $p(x_p, z)$ in exp. 1, $\lambda_*^{-1}(N)$ of $p(x_p, z)$ can be analysed in isolation heeding

$$e^{\alpha(h=1)} > (1/\lambda_*^{-1}(N) = \lambda_*(N)) \Leftrightarrow \alpha > \ln(\lambda_*(N)) \quad (3.1)$$

to ensure all zeros of the opening phase lie outside the UC. Note the sampling period $h = 1$ (cf. exp. 1). $\lambda_*^{-1}(N)$ of $p(x_p, z)$ yields;

$$\begin{aligned} \lambda_*^{-1}(N) &= \left(1 + \max \left\{ \left| \frac{\sin(k)}{\sin(k(N-1))} \right|, \left| \frac{\sin(kN)}{\sin(k(N-1))} \right| \right\} \right)^{-1} \\ \text{where } k &= \omega_g = \pi/t_p \end{aligned} \quad (3.2)$$

3.1.1. Global minima

The global minima points of $\lambda_*^{-1}(N)$ are

$$\left. \begin{aligned} \lim_{N \rightarrow a^\pm} \left| \frac{\sin(k)}{\sin(k(N-1))} \right| = \infty \\ \lim_{N \rightarrow a^\pm} \left| \frac{\sin(kN)}{\sin(k(N-1))} \right| = \infty \end{aligned} \right\} \Rightarrow \lim_{N \rightarrow a^\pm} \lambda_*^{-1}(N) = 0$$

where $a = (k + q\pi)/k = 1 + qt_p$, $q \in \mathbb{Z}$ (3.3)

3.1.2. Global maxima

The global maxima points of $\lambda_*^{-1}(N)$ are at the intersections

$$\left| \frac{\sin(k)}{\sin(k(N-1))} \right| = \left| \frac{\sin(kN)}{\sin(k(N-1))} \right| \Leftrightarrow |\sin(k)| = |\sin(kN)| \Leftrightarrow$$

$$N = \pm 1 + qt_p \vee N = t_p \pm 1 + qt_p \Rightarrow$$

$$N = -1 + qt_p \vee N = t_p - 1 + qt_p, q \in \mathbb{Z} \quad (3.4)$$

The solutions $N = 1 \vee N = t_p + 1$ are deducted as they coincide with the global minima points, i.e. singularities. The global maxima points lead to the following maximum value;

$$\left| \frac{\sin(k)}{\sin(k(N-1))} \right| \Big|_{\substack{N=-1 \vee \\ N=t_p-1}} = \left| \frac{1}{2\cos(\pi/t_p)} \right| \Rightarrow$$

$$\lambda_*^{-1}(-1) = \lambda_*^{-1}(t_p - 1) = \frac{|2\cos(\pi/t_p)|}{|2\cos(\pi/t_p)|+1} \quad (3.5)$$

Note that choosing N equal to a global max point does not guarantee the best possible separation, it merely puts the best possible lower bound on the separation distance; this is due to the inequalities in theorem 1, 2, and 3.

3.2. Analysis of Cauchy bound of closing phase

The upper modulus bound for the ZZT rep of the closing phase stated in expression 2 can be found by either of the theorems 1 or 3. Theorem 3 will, however, in some situations yield a tighter bound [8]; this is the case in the present closing phase analysis.

$$\lambda_a(N) = \max \left\{ 1, \left| \frac{c_2 e^{-k} - c_1}{c_1 - c_2} \right| + \left| \frac{c_2 - c_1 e^{-kN}}{c_1 - c_2} \right| + \left| \frac{c_1 e^{-kN} - c_2 e^{-k}}{c_1 - c_2} \right| \right\} =$$

$$\max \left\{ 1, e^{-\epsilon} + \frac{|e^{-\epsilon t_c} - e^{-\epsilon N}| + |e^{\epsilon(-t_c-1)} - e^{-\epsilon N}|}{1 - e^{-\epsilon t_c}} \right\}$$

where $c_1 = e^{t_e}$, $c_2 = e^{-\epsilon(t_c - t_e)}$, $k = \epsilon$

3.2.1. Global minimum

Clearly, the global minimum value of $\lambda_a(N)$ is 1; the minimum is achieved at the intersection

$$1 = e^{-\epsilon} + \frac{|e^{-\epsilon t_c} - e^{-\epsilon N}| + |e^{\epsilon(-t_c-1)} - e^{-\epsilon N}|}{1 - e^{-\epsilon t_c}} \Rightarrow$$

$$N = -\ln \left(\frac{1}{2} - \frac{1}{2} e^{-\epsilon} + e^{-\epsilon(t_c+1)} \right) / \epsilon \quad (3.7)$$

The approximation $N \approx -\ln(\frac{1}{2})/\epsilon$ is sufficiently accurate for physically realistic LF GFM as $\epsilon > 300$ [3] (cf. eq. 4.4).

3.2.2. Global maximum

For completeness, the global maximum point of $\lambda_a(N)$ is 0; the maximum value is

$$\lim_{N \rightarrow 0^+} e^{-\epsilon} + \frac{|e^{-\epsilon t_c} - e^{-\epsilon N}| + |e^{\epsilon(-t_c-1)} - e^{-\epsilon N}|}{1 - e^{-\epsilon t_c}} =$$

$$e^{-\epsilon} + 1 + \frac{1 - e^{-\epsilon(t_c+1)}}{1 - e^{-\epsilon t_c}} \quad (3.8)$$

For practical LF GFM instances, the limit value is approximately 2 as $\epsilon > 300$ and $t_c > 0.002$ [3].

4. Numerical Experiment

In this experiment, the lower Cauchy bound of the opening phase and the upper Cauchy bound of the closing phase are illustrated as functions of the coefficient sequence length N . Again, $N \in \mathbb{R}_{++}$ although $N \in \mathbb{Z}_{++}$ in theorem 1, 2, and 3. Letting N be a non-integer real must only be considered as a reasonable theoretical interpolation.

The LF GFM parameters are set to the values applied in the waveform in figure 1; they are

$$\begin{aligned} t_0 &= 0.00000s & t_p &= 0.00380s & t_e &= 0.00480s \\ t_a &= 0.00031s & t_c &= 0.00800s & E_e &= 1.00000Pa \end{aligned}$$

4.1. Cauchy bound of opening phase

The lower Cauchy bound of the LF GFM opening phase (cf. eq. 3.2) is illustrated along with the lower bound of the alpha parameter (cf. ineq. 3.1) in figure 2.

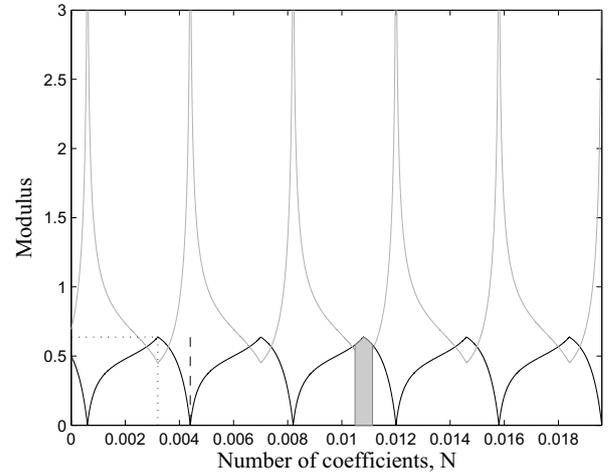


Figure 2: The black line illustrates the lower Cauchy bound of the LF GFM opening phase. A global min point (cf. eq. 3.3) is indicated by the vertical dashed line. The vertical and horizontal dotted lines indicate a global max point (cf. eq. 3.4) and the global max value (cf. eq. 3.5) respectively. The grey line shows the lower bound of α (cf. ineq. 3.1). The grey region exemplifies a feasible neighbourhood for N (cf. ineq. 4.3).

Figure 2 illustrates that $\lambda_*^{-1}(N) \rightarrow 0 \Rightarrow \alpha \rightarrow \infty$ why N must be chosen outside a small neighbourhood of the global min points and preferably inside a small neighbourhood of the global max points. By choosing N according to these guidelines, the constraint on α is minimized and the opening/closing phase separation guarantee around the UC maximized. Concretely, by def. 1 the α parameter is computed;

$$\alpha = \ln \left(\frac{e(t_e)}{E_0 \sin(\pi t_e / t_p)} \right) / t_e = 543.64280 \quad (4.1)$$

As $\frac{e(t_e)}{E_0}$ is constant under changes in absolute sound pressure level, α is a function of t_e and t_p . The α -value in eq. 4.1 leads to the following feasible neighbourhood for N ,

$$\lambda_*^{-1}(N) > (e^{-543.6428 \cdot (h=1)}) \approx 8 \cdot 10^{-237} \Rightarrow$$

$$N \in \mathbb{R}_{++} \setminus [1 + qt_p \pm \delta], q \in \mathbb{Z}, \delta \approx 0 \quad (4.2)$$

In this example, N can practically be chosen freely with exception of the global min points at $N = 1 + qt_p$. Of course, N must be an element of \mathbb{Z}_{++} and is chosen in combination with frame length and sampling frequency, e.g. 30ms of data sampled at 16kHz $\Rightarrow N = \frac{30 \cdot 16000}{1000} = 480$. Note the sampling period $h = 1$ in ineq. 4.2 (cf. exp. 1 and ineq. 3.1). Letting $h = 0.001$ yields;

$$\lambda_*^{-1}(N) > (e^{-543.6428 \cdot (h=0.001)} \approx 0.58062) \Rightarrow \quad (4.3)$$

$$N \in [-1 + qt_p - 82.68\%qt_p; -1 + qt_p + 86.80\%qt_p], q \in \mathbb{Z}$$

The feasible neighbourhood for N in ineq. 4.3 is shown in fig. 2.

4.2. Cauchy bound of closing phase

In order to compute the upper Cauchy bound of the LF GFM closing phase (cf. eq. 3.6), the parameter ϵ must be estimated. This is done iteratively by the following relationship [5],

$$\epsilon t_a = 1 - e^{-\epsilon(t_c - t_e)} \Leftrightarrow \epsilon \approx 3261.44143 \quad (4.4)$$

Figure 3 illustrates the upper Cauchy bound and that the global min point is reached at $N < t_c \ll 1$; thus, only the opening phase constraints on N must be considered when choosing a suitable sequence length.

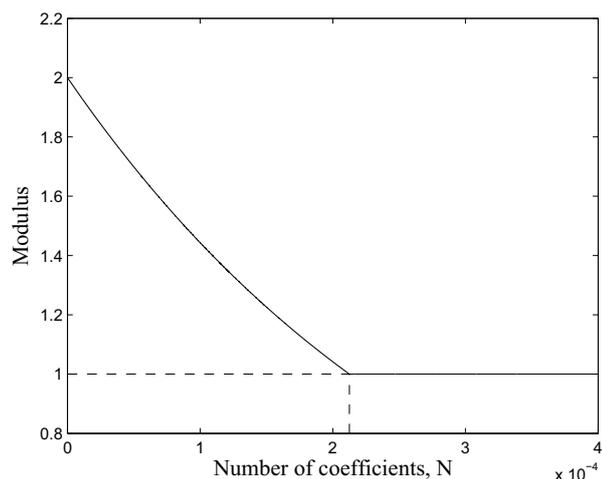


Figure 3: Upper Cauchy bound of LF GFM closing phase. The global min point in eq. 3.7 is indicated by the vertical dashed line; the horizontal dashed line indicates the global min value.

5. Results

In the present paper it has been demonstrated that the lower bound of ZTT-domain immiscibility of LF GFM opening and closing phases is periodically variant under changes in the z-transform coefficient sequence length, i.e. frame length, N .

To maximize the lower immiscibility bound, i.e. maximize the opening/closing phase separation guarantee, N must equal a global max point of the lower Cauchy bound of the LF GFM opening phase, (cf. eq. 3.4). At these t_p -periodically recurring global max points, the lower Cauchy bound of the opening phase ZTT rep, $\lambda_*^{-1}(N)$, reaches its max value (cf. eq. 3.5). To guarantee that the ZTT rep of the opening phase lies outside the UC, the constraint $\alpha > \ln(\lambda_*(N))$ must be satisfied (cf. def. 1 and eq. 3.1); by choosing N equal to a global max point, the constraint on α is minimized. In practical terms, a well-chosen

N is contained in a small neighbourhood around a max point. The size of the neighbourhood is determined by t_p and α .

To minimize the opening/closing phase separation guarantee, N must equal a global min point of the lower Cauchy bound of the opening phase (cf. eq. 3.3). At the t_p -periodically recurring global min points, the guarantee vanishes and the ZTT rep of the opening phase does not have a min modulus bound, i.e. the Cauchy bound is zero. In practical terms, N must be chosen outside a small neighbourhood around the min points as α is upper bounded in a physically realistic LF GFM instance; feasible neighbourhoods are exemplified in ineq. 4.2 and 4.3.

The global min value, 1, of the upper Cauchy bound of the closing phase is reached at $N < t_c \ll 1$ (cf. eq. 3.7); thus, the ZTT rep of the closing phase will always be within or on the UC. Therefore, only the opening phase constraints on N must be considered when choosing a suitable sequence length. The global maximum value, ≈ 2 for practical LF GFM instances, is only interesting from a theoretical point of view as it is reached at $N = 0$ (cf. eq. 3.8).

6. Discussion

Section 3 and 4 demonstrate how to choose a z-transform coefficient sequence length, i.e. frame length, N , that maximize or minimize the separation guarantee. It must be emphasized that the max separation guarantee does not guarantee the best possible separation, but the best possible lower bound on the separation distance (cf. th. 1, 2, and 3). In practical situations, the present authors have witnessed that opening/closing phase separation around the UC is the predominant situation even if guidelines on how to choose N is not considered.

The guidelines on how to choose N may be applied in studies as [1], [2] and [4] where separation maximization by ZTT rep is relevant. In these studies, speech waveforms are windowed synchronously with the GCI by windows of two-pitch-period length, i.e. frames from neighbouring GCIs are determined dynamically per pitch and analyzed; this results in varying frame and window lengths. To maximize separation, these frames could also be dynamically adjusted according to the guidelines.

7. References

- [1] B. Bozkurt, *Zeros of the z-transform (ZTT) representation and chirp group delay processing for the analysis of source and filter characteristics of speech signals*, Ph.D. dissertation, Faculte Polytechnique de Mons, Belgium, October 2005.
- [2] B. Bozkurt, B. Doval, C. D'Alessandro and Thierry Dutoit, *Zeros of Z-Transform Representation With Application to Source-Filter Separation in Speech*, IEEE Signal Processing Letters, vol. 12, No. 4, April 2005.
- [3] C.F. Pedersen, P. Dalsgaard and O. Andersen, *On Separability of the Opening and Closing Phases of the LF Glottal Flow Model by Zeros of the Z-transform*, IEEE Transactions on Audio, Speech, and Language Processing, submitted for publication, 2009.
- [4] B. Bozkurt, L. Couvreur and T. Dutoit, *Chirp group delay analysis of speech signals*, Elsevier Speech Communication, vol. 49, no. 3, pp. 159-176, March 2007.
- [5] G. Fant, J. Liljencrants and Q. Lin, *A four-parameter model of glottal flow*, STL-QPSR, vol. 26, no. 4, pages 1-13, 1985.
- [6] A.L. Cauchy, *Exercices de mathematique*, Oeuvres 2, vol. 9, 1829.
- [7] Q.I. Rahman and G. Schmeisser, *Analytic Theory of Polynomials*, Oxford University Press, 2002.
- [8] H.P. Hirst and W.T. Macey, *Bounding the Roots of Polynomials*, The College Mathematics Journal, vol. 28, No. 4, Mathematical Association of America, 1997.