RANDOMNESS AND THE REVERBERATION TIME, RT₆₀, OF ACOUSTIC RESPONSES

Ian J. Kelly and Francis M. Boland

Department of Electronic and Electrical Engineering, School of Engineering, Trinity College Dublin, Ireland.

ABSTRACT

This paper investigates an implication of the clustering, on the complex plane, of the roots of transfer function polynomials obtained from acoustic responses. These polynomials can be the high order transfer functions obtained from room impulse responses or the relatively lower order ones obtained from head related impulse responses. This clustering behavior is explained using results from the theory of the roots of random polynomials. It is demonstrated that, although they do not appear that way, the coefficients of acoustic polynomials can be modeled as random polynomials with certain constraints applied. In the case of room impulse responses, the median of their clustered roots is shown to be directly related to the reverberation time of the room. This is shown to be an accurate measure of the reverberation time by comparison with other estimation techniques.

Index Terms- Reverberation, Random Polynomials

1. INTRODUCTION

The work of Erdős and Turan [1] and later the work of Shepp and Vanderbei [2] among others, showed that polynomials of infinite order, whose coefficients were modeled as random variables distributed according to a small subset of stable probability distributions, would have roots which clustered uniformly about the unit circle. Continued work on this topic has extended the range of applicable distribution and has shown the results can apply to finite order random polynomials. In their 2008 paper Hughes and Nikeghbali [3] formulated a set of results, based upon Erdős and Turan's original proof, that required no independence restriction on the random variables. They simply require that the first and last coefficients of the finite order random polynomial be non-zero in order to calculate both a lower bound for the number of roots to be found within an annular region embracing the unit circle and a lower bound on the number of roots in each angular sub division of this annulus.

The authors have examined [4] the statistical distribution of the coefficients of common acoustic responses such as head related impulse responses (HRIRs) and room impulse responses (RIRs). We have shown that, despite exhibiting features such as exponential decay and onset delays, these responses' root distributions essentially behave in the same way as those of random polynomials. That investigation explained results, from our earlier work [5, 6], on extracting common sub-systems from sets of HRIRs through a process that amounts to seeking an approximate Greatest Common Divisor of the equivalent set of z-domain transfer functions. Our investigation of the role of root clustering in this approximate factorization of acoustic responses also introduced an application of this analysis to headphone equalization. In this paper we explore the relationship between the positioning and range of these annular clusters of roots for RIRs and we show that we can obtain an accurate measure of the room (RT_{60}) reverberation time from this information.

2. ACOUSTIC IMPULSE RESPONSES AS RANDOM POLYNOMIALS

To appraise the statistical distribution of the coefficients of common acoustic responses we examined the filter coefficients from 19074 HRIRs taken from the IRCAM *Listen* [7] HRIR data set. This amounts to 9765888 coefficients allowing for a good estimation of their overall distribution. Figure 1 shows a histogram of the coefficients from the HRIRs and superimposed on the histogram is an α -stable pdf which has been scaled to match the number of entries in the histogram. This approximate distribution was fitted using the McCulloch's method [8]. The approximate distribution is parametrized via its characteristic function [9]

$$\psi(t) = \mathbf{E}[e^{-\gamma^{\alpha}|t|^{\alpha}(1-j\beta sign(t)\tan\left(\frac{\pi\alpha}{2}\right))+j\delta t}]$$
(1)

Here the four parameters α , β , γ and δ represent the characteristic exponent describing the tail, the skewness, the scale and the location respectively. The parameters used to generate the distribution in Figure 1 were $\alpha = 1.2379$, $\beta = -0.2666$, $\gamma = 0.0039$, $\delta = -0.0013$. It is clear that the coefficients of HRIRs ,such as these, closely match a skewed α -stable distribution. Room impulse responses measured by the authors in

The authors would like to acknowledge support for this research from Trinity College Dublin, Google Inc. and Science Foundation Ireland Project(08/IN.1/I2112) - Content Aware Media Processing.



Fig. 1. Histogram of the coefficient values for 19074 IRCAM HRIRs, closely resembling a skewed α -stable distribution.

both office sized rooms and a small hall along with the room impulse responses from the AIR database [10], all had similarly distributed coefficients. All distributions were heavy tailed and leptokurtotic as is to be expected from exponentially decaying signals. The α -stable distribution generalizes the classical central limit theorem for sequences of random variables which may not be i.i.d. Figure 1 shows a histogram of the coefficients from 19074 HRIRs from the IRCAM Listen data set.

According to Hughes and Nikeghbali [3], for a polynomial defined as

$$P(z) = \sum_{k=0}^{N} a_k z^k \tag{2}$$

with randomly distributed coefficients a_k , the roots of the polynomial cluster uniformly about the unit circle if $L_N(P)$ is small compared to the polynomial order N where

$$L_N(P) = \log \sum_{k=0}^N |a_k| - \frac{1}{2} \log |a_0| - \frac{1}{2} \log |a_N|$$
(3)

provided a_0 and a_N are non zero.

Taking for example the (N = 512) IRCAM HRIRs, values of $L_N(P)$ for the 19074 responses had a mean value of 7.2346, with maximum and minimum 11.8950 and 5.3000 respectively.

Hughes and Nikeghbali derived an expression placing a lower bound on the number of roots v_N of a polynomial P(z) lying in an annulus bounded by $1 - \rho$ and $1/(1 - \rho)$ for $0 < \rho < 1$ as

$$v_N \ge N - 2\frac{L_N(P)}{\rho} \tag{4}$$

where

$$v_N \triangleq \#\{z_k : 1 - \rho \le |z_k| \le \frac{1}{1 - \rho}\}$$
 (5)

Applying this to the IRCAM responses using the mean value of $L_N(P)$, the number of roots falling within an annulus

around the unit circle parametrized by $\rho = 0.2$ would always be greater than or equal to 442 for HRIRs of length N = 512i.e. 86.5% of the roots.

On calculating the actual distribution of the roots of these responses however, the roots were found to cluster far more densely. The roots of 2325 randomly chosen IRCAM responses were calculated. The proportion of those within an annulus parametrized by $\rho = 0.2$ was calculated to be 1176955 out of 1188075 roots or 99.06% of the roots.

The reason behind this difference can be found by examining equation 3. It was stated that when $L_N(P)$ is small in relation to the order, N, of P(z), then the roots of P(z) will cluster uniformly about the unit circle. However, this depends heavily on the magnitudes of the first and last coefficients of P(z), a_0 and a_N . If either or both of these coefficients are close to zero $L_N(P)$ will grow much larger. However it can be shown that despite the fact that acoustic impulse responses often have low magnitude first and last coefficients, their roots will cluster no less closely about the unit circle.

To see why this is we must look closely at the reasons why acoustic impulse responses have low magnitude initial and final coefficients. Acoustic responses in general have some onset delay which encodes the source-receiver delay and, also in the case of HRIRs, the inter-aural time difference between left and right HRIR pairs. These onset delays are generally affected by measurement noise and transducer transfer functions, meaning the delay is seen as a set of low magnitude initial coefficients as opposed to a pure delay. This accounts for a low magnitude a_0 pushing up the value of $L_N(P)$. However the behavior of the roots is not greatly affected as the delay can be seen as a convolution with something approximating a time delayed Kronecker delta. This short polynomial denoted $\delta_R(z)$ of order $R \ll N$ will simply add an asymptotically negligible number of roots to those of an acoustic response despite increasing the value of $L_N(P)$.

This effect can easily be demonstrated with a HRIR with 22 samples of approximate onset delay stripped off as in Figure 2. The lower signal in Figure 2 was then deconvolved from the upper signal yielding a signal closely approximating a delayed Kronecker delta. The roots of this signal can be seen to form a ring outside the unit circle in Figure 3.

A similar argument can be made for the low magnitude of the last coefficient of an acoustic response, a_N . Most acoustical impulse responses exhibit a statistical decaying magnitude similar to an exponential decay. It was proposed by Steiglitz and Dickinson [11] that the decay in acoustic response coefficients could be modeled as IIR filtering, adding only a fixed (asymptotically negligible) number of zeros and poles to the responses's z-transform. This means one can still use the white noise result on the distribution of roots. A windowed impulse response giving an approximate FIR filter would add just a fixed set of roots to the Argand plane. However such a model of the exponential decay is not suitable for the proposed application. The correlation that the IIR nature of such



Fig. 2. This figure shows a HRIR (above) and that same HRIR with approximate onset delay removed (below)



Fig. 3. Roots of the approximate delay polynomial

a filter would induce on the coefficients of the response would mean that their probability distribution would be drastically changed with a removal of independence This goes against the fact that the arrivals in a RIR should be independent of one another. This behavior renders Steiglitz and Dickinson's model unsuitable. Furthermore, such IIR based models do not have frequency responses which behave in a similar way to those of acoustic impulse responses. The authors instead propose to model the decay of acoustic signals like HRIRs and RIRs such that the expected value of the coefficients decay exponentially, i.e. each sample in a response h[n] can be modeled as

$$\mathbf{E}\left[h[n]\right] = e^{-\beta n} p[n] \tag{6}$$

and thus,

$$\mathbf{E}\left[H(z)\right] = P(e^{-\beta}z) \tag{7}$$

where $\beta << 1$ and as before p[n] is a random signal vector of length N. This means that the expected value of the polynomial coefficients in the ensemble also has an exponential decay. Such a model is reminiscent of Polack's statistical reverberation model [12]. This has the effect of reducing the expected value of the magnitude of the polynomials roots by a uniform factor, $e^{-\beta}$, without altering their angular location. Thus the roots remain just as closely clustered, but about a ring just within the unit circle and not about the unit circle itself. The predominantly minimum phase nature of HRIR roots closely matches this model. The exponential decay thus results in a low magnitude final coefficient a_N , increasing the value of $L_N(P)$ without disrupting the clustering behavior of the polynomial roots.

Concerning the angular distribution of the roots in this annulus about the unit circle; Let

$$v_{\theta\phi} \triangleq \#\{z_k : \theta \le \arg(z_k) < \phi\}$$
(8)

be the number of roots of polynomial P(z) whose argument lies between θ and ϕ where $0 \le \theta < \phi \le 2\pi$. According to Hughes and Nikeghbali, the quadratic inequality

$$v_{\theta\phi}^2 - \frac{N(\phi-\theta)}{\pi} v_{\theta\phi} + \frac{N^2(\phi^2 - 2\phi\theta + \theta^2)}{4\pi^2} - NCL_N(P) \le 0$$
(9)

holds where C is some constant. Placing a lower bound on the number of roots per angular subdivision

In fact for the 2325 IRCAM HRIRs studied, the average number of roots lying in each 1° segment was 1.4179. This fits very well with the ideal Erdős and Turan result which states that

$$\lim_{N \to \infty} \mathbf{E} \left[\frac{1}{N} v_{\theta \phi} \right] = \frac{\phi - \theta}{2\pi}$$
(10)

as

$$N\frac{\phi-\theta}{2\pi} = 1.4222$$
 (11)

for N = 512.

3. RELATING REVERBERATION TIME TO ROOT CLUSTER RADIUS

Consider the following model of a room impulse response. Let p[n] be a random signal vector of length N who's entries correspond to the coefficients of a random polynomial. We can multiply this signal with a decaying exponential window $w[n] = e^{-\beta n}$ also of length N. The room impulse response can thus be modeled as

$$h[n] = p[n] \otimes w[n] \tag{12}$$

where \otimes is the Hadamard product for vectors.

The reverberation time RT_{60} is the 60dB decay time for a RIR [13]. In the case of our model signal this can be easily derived from the envelope w[n] and can be obtained by solving

$$20 \log_{10} (e^{-\beta \mathbf{K} \mathbf{I}_{60}}) = -60 \text{ (dB)}$$

to get
$$\mathbf{RT}_{60} = \frac{1}{\beta} \ln (10^3).$$
 (13)

We know from Hughes and Nikeghbali [3], that the roots of p[n] cluster uniformly about the unit circle. Also by the

properties of the z-transform $H(z) = P(e^{-\beta}z)$ and so the magnitudes of the roots of P(z) are scaled by a factor of $e^{-\beta}$ or $e^{\frac{\ln(10^3)}{RT_{60}}}$ to become the roots of H(z).

By way of Sabine's formula [14] we can also roughly relate the radius of the ring of clustered roots from a room impulse response to the room's physical characteristics. Sabine's equation states

$$\mathbf{RT}_{60} = \frac{4V}{cS\bar{a}}\ln(10^6) \tag{14}$$

where V is the room's volume in m^3 , S is the room's total surface area in m^2 , c is the speed of sound in air, $c \approx 340.29 \ ms^{-1}$ and \bar{a} is the average absorption coefficient of the room's surfaces. Recalling that $\mathrm{RT}_{60} = \frac{1}{\beta} \ln{(10^3)}$, this allows us to relate these physical properties to β

$$\beta = \frac{cS\bar{a}}{8V}.$$
(15)

4. EXAMINATION OF RT₆₀ ESTIMATED FROM REAL AND MODELED RIRS

A set of Model RIRs were generated by means of a eight thousandth order, random, normally distributed coefficient vector, whose coefficients had been scaled by a known decaying exponential. In this implementation the user can specify a value for the simulated room's RT_{60} . Therefore it was possible to test whether one could correctly estimate the RT_{60} of a room from the radius at which the RIR's roots clustered. The results from this simulation are shown in Figure 4. Here the error in estimating RT_{60} from both the root location based method described in Section 3 and a method based upon ISO 3382 [15] are shown. The radius of the root annulus is taken to be the median value of the magnitudes of the roots of the response. A median value is chosen so as to disregard outlying roots.



Fig. 4. Error in reverberation time from the root location based method described in Section 3 and a method based upon ISO 3382 [15] for a set of 8000th order generated responses

This benchmark test was performed in order to ascertain the expected error in the RT_{60} estimations under fully con-

trolled conditions. Based upon the success of this as a benchmark it was then possible to compare the estimation of RT_{60} using both methods when applied to real RIRs from the AIR database [10]. This comparison is made in Table 1. A comparison is also made to the RT_{60} values provided by Jeub, Schäfer and Vary along with the the AIR database. They used the Schroeder method [16] to estimate the these values. As can be seen the RT_{60} values obtained via the root cluster based method are more consistent within each room.

Room	Schroeder	Root Method	ISO 3382
	0.0800	0.1423	0.0964
Studio	0.1100	0.1400	0.0955
	0.1800	0.1409	0.0976
	(2.6e-3)	(1.3e-6)	(1.1e-6)
	0.3700	0.3530	0.5479
Office	0.4400	0.3768	0.5366
	0.4800	0.3851	0.2610
	(3.1e-3)	(2.7e-4)	(2.6e-2)
	0.2100	0.3141	0.3359
	0.2200	0.3213	0.3565
Meeting	0.2100	0.3136	0.3686
	0.2400	0.3189	0.3754
	0.2500	0.3288	0.3526
	(3.3e-4)	(3.8e-5)	(2.3e-5)
Lecture	0.7000	0.7560	0.8569
	0.7200	0.7995	0.8647
	0.7900	0.8076	0.8910
	0.8000	0.7990	0.9007
	0.8100	0.8169	0.8902
	0.8300	0.8087	0.8405
	(2.7e-3)	(4.6e-4)	(5.5e-4)

Table 1. RT_{60} values for the first four rooms in the AIR database calculated via root locations, ISO 3382 and the Schroeder method. The values shown in brackets are the variances between RT_{60} values estimated in each room using each of the three methods.

5. CONCLUSION

In this paper it has been demonstrated that behavior of the root distribution of acoustic impulse responses is consistent with those of random polynomials. This has been shown to be the case despite issues such as exponential decay. It has also been shown that based upon the effects of scaling random polynomials with decaying exponential windows, it is possible to accurately deduce the RT_{60} of a room from the median of the root locations of a measured impulse response. The RT_{60} estimations from root locations were shown to be consistent with those made by standard methods and to be almost constant across different RIR measurements from within a room.

6. REFERENCES

- P. Erdos and P. Turan, "On the distribution of roots of polynomials," *The Annals of Mathematics*, vol. 51, no. 1, pp. 105–119, Jan 1950.
- [2] L.. Shepp and R.. Vanderbei, "The complex zeros of random polynomials," *Transactions of the American Mathematical Society*, vol. 347, no. 11, pp. 4365–4384, Jan 1995.
- [3] C. P. Hughes and A. Nikeghbali, "The zeros of random polynomials cluster uniformly near the unit circle.," *Compositio Mathematica*, vol. 144, pp. 734746, 2008.
- [4] Ian J. Kelly and Francis M. Boland, "Exploiting randomness in acoustic impulse responses to achieve headphone compensation through deconvolution," *J. Acoust. Soc. Am.*, vol. 133, no. 5, may 2013.
- [5] C. Masterson, G. Kearney, M. Gorzel, and F.M. Boland, "HRIR order reduction using approximate factorization," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 20, no. 6, pp. 1808 –1817, aug. 2012.
- [6] C. Masterson, S. Adams, G. Kearney, and F. Boland, "A method for head related impulse response simplification," in 17th European Signal Processing Conference (EUSIPCO 2009), Glasgow, Scotland, August 2009, EUSIPCO, pp. 2569–2573.
- [7] IRCAM Room Acoustics Team, "Listen HRTF database," http://recherche.ircam.fr/equipes/salles/listen/, Accessed: 08/06/12.
- [8] J. Huston McCulloch, "Simple consistent estimators of stable distribution parameters," *Communications in Statistics - Simulation and Computation*, vol. 15, no. 4, pp. 1109–1136, 1986.
- [9] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic models with infinite variance*, Chapman and Hall/CRC, 2000.
- [10] Marco Jeub, Magnus Schäfer, and Peter Vary, "A binaural room impulse response database for the evaluation of dereverberation algorithms," in *Proceedings of the 16th international conference on Digital Signal Processing*, Piscataway, NJ, USA, 2009, DSP'09, pp. 550– 554, IEEE Press.
- [11] K. Steiglitz and B. Dickinson, "Phase unwrapping by factorization," Acoustics, Speech and Signal Processing, IEEE Transactions on, vol. 30, no. 6, pp. 984–991, dec 1982.

- [12] Jean-Dominique Polack, La transmission de lenergie sonore dans les salles, Ph.D. thesis, Universite du Maine, Le Mans, 1988.
- [13] Wallace C Sabine, "Architectural acoustics," in *Proceedings of the American Academy of Arts and Sciences*. JSTOR, 1906, vol. 42, pp. 51–84.
- [14] Sabine W. C. and M. David Egan, "Collected papers on acoustics," J. Acoust. Soc. Am., vol. 95, no. 6, pp. 3679–3680, 1994.
- [15] ISO 3382:1997, "Acoustics measurement of the reverberation time of rooms with reference to other acoustical parameters," International Organisation for Standardisation, Geneva, Switzerland.
- [16] Manfred R Schroeder, "New method of measuring reverberation time," *The Journal of the Acoustical Society of America*, vol. 37, pp. 409, 1965.