UNIVERSITY OF CALIFORNIA SAN DIEGO

Impulse Response Measurement Using Complementary Sequences

A dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Music

by

Christopher Warren

Committee in charge:

Professor Miller Puckette, Chair
Professor F. Richard Moore, Co-Chair
Professor Anthony Burr
Professor Tamara Smyth
Professor Shahrokh Yadegari

2014

The Dissertation of Christopher Warren is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

Co-Chair

_____

Chair

University of California San Diego

2014

DEDICATION

Dedicated to my wife Ariana and my parents.

EPIGRAPH

“I call architecture frozen music.”

— Johann Wolfgang von Goethe

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to acknowledge Professor Miller Puckette for his overwhelming generosity and tremendous support.

I would also like to acknowledge Professor Emeritus F. Richard Moore, who continues to allow me to interrupt his retirement for discussions about acoustics and outer space.

Vita

1997 Bachelor of Arts in Music Composition, Brandeis University

2008 Master of Arts in Music, Science, and Technology, Stanford University

2014 Doctor of Philosophy in Music, University of California San Diego

Publications

The Flote: An Instrument For People With Limited Mobility
A. Aziz, C. Warren, H. Bursk, S. Folmer
Assets, Halifax, Nova Scotia, Canada
ACM SIGACCESS 978-1-59593-976-0/08/10

The Pond: Interactive Multimedia Installation
S. Follmer, C. Warren, A. Marquez-Borbon
NIME, Genoa, Italy ISBN:13-978-88-901344-6-3

Fields of Study

Major Field: Computer Music

Studies in Digital Audio Signal Processing

Professor Miller Puckette

Studies in Acoustics

Professor F. Richard Moore

ABSTRACT OF THE DISSERTATION

Impulse Response Measurement Using Complementary Sequences

by

Christopher Warren

Doctor of Philosophy in Music

University of California, San Diego, 2014

Professor Miller Puckette, Chair
Professor F. Richard Moore, Co-Chair

Although several methods have been proposed for impulse response measurement in acoustic spaces, there is no ideal one. Each responds differently to environmental noise present in the space being measured. Here a novel technique using complementary sequences is proposed, which offers improved accuracy over the commonly used methods under certain noise conditions. This method is compared with existing methods in the presence of several different noises, focusing on situations where the amplitudes of the noises are comparable to that of the measuring signal. Musical applications of the impulse responses generated by the new method are presented, focusing on their use in artificial reverberation. A library of impulse responses that have been created using this technique is discussed.

Notation

The following notation is used to succinctly illustrate the use of complementary pairs and various transformations.

Complementary pairs will be designated A_k and B_k. Each of these represents a sequence of numbers $A_k(1,\ldots,N)$ and $B_k(1,\ldots,N)$, where the subscript k is used to index the generation and will consequently also indicate the power of two which will produce the length N of each sequence, $N=2^k$ (e.g. for A_3, $N=2^3$).

Sequences are created iteratively through concatenation, with each generation doubling the length of the previous generation. Brackets denote the concatenation of two or more sequences. For two sequences A_k and B_k, each of length N, $[A_k\ B_k]$ denotes the $2N$ sequence:

$$[A_k B_k] = A_k(1), \ldots, A_k(N), B_k(1), \ldots, B_k(N)$$

Curly braces denote the interleaving of two or more sequences.

$$\{A_k B_k\} = A_k(1), B_k(1), \ldots, A_k(N), B_k(N)$$

Overbars denote a reversal in time.

$$\overline{A}_k = A_k(N), \ldots, A_k(1)$$

U_m denotes integer upsampling where m zeros are placed between each sample.

$$U_m(A_k) = \begin{cases} A_k(n/m), & \text{when } n/m \text{ is an integer} \\ 0, & \text{otherwise} \end{cases}$$
$$n = 1, \ldots, N$$

D_m denotes integer downsampling where every m^{th} sample is retained and the rest are discarded.

$$D_m(A_k) = A_k(nm)$$
$$n = 1, \ldots, N/m$$

S_m denotes a circular rotation m places to the right.

$$S_m(A_k) = A_k((n-m) mod N)$$

Convolution and cross-correlation are used in both their linear (acyclic) and circular (cyclic) forms.

linear (acyclic) convolution: $*$

circular (cyclic) convolution: $\circledast$

linear (acyclic) cross-correlation: $\star$

circular (cyclic) cross-correlation: $\circledast$

1.0 Introduction

The practice of recording music often bears little resemblance to the practice of listening to music. Microphones serve as proxies for ears, but are often placed far closer to their focus than any listener would care to be. Multitrack recording allows musicians to be recorded at different times, potentially in very different rooms, yet the expectation remains that their final product will appear to exist within a cohesive sonic space. Many sounds may be created electronically, and therefore without any inherent room ambience.

It is therefore not surprising that artificial reverberation is among the most widely used audio effects. The ability to alter the perceived space in which music exists is essential to the practice of audio engineering. To this end, a broad range of sonic tools have been developed which replicate the natural process of reverberation in controllable ways.

1.1 Artificial reverberators

Many solutions simply pass the signal though a physical system of notable resonance. This may be as straightforward as playing the signal back into a particularly reverberant room and recording the result. Echo chambers built specifically for this purpose use speakers to play back signals while microphones capture the response of the room. Highly reflective surfaces such as concrete and metal can be used to maximize the resonance of the space, prolonging the reverberation. Relocation of the microphone and speaker within the space can easily and controllably alter the results.

The echo chambers beneath Capitol Studios in Los Angeles are notable examples of this practice. Double concrete walls, a floating floor, and a suspended

ceiling minimize outside noise. Irregular trapezoidal construction encourages rich and thorough mixing by mitigating standing waves. A broadband decay time of five seconds can be reduced using absorbent baffles (Bayless 74-76).

While this technique may produce exceptional results, it is only possible at great expense. Great care must be taken to ensure that the rooms maintain a constantly low noise floor, since any noise present will be directly included in the final mix. Purpose-built architecture is not easily moved, geographically constraining the effect. The current relative scarcity of this once-common technique speaks to its poor cost-effectiveness as compared to other approaches.

Reverberation is, of course, not limited to vibrations in air. A suspended metal plate can be employed as a resonator, with electromagnetic transducers used to both excite the system and capture its resonance by translating electronic signals to physical movement and back again. These plates are typically large and heavy, their masses far too great to be substantially disturbed by vibrations in the air. While this makes them quite resistant to any extraneous environmental noise, it severely limits their portability and consequently restricts their use. The mass and malleability of the plate will determine the character of its resonance. Since neither of these properties are easily or quickly altered, external dampers are often used to quickly dissipate energy from the plate to reduce the decay time of the reverberation. This allows for limited timbral control over the output of the system, by attenuating higher frequencies more than lower frequencies (Eargle 233-235).

Stretched springs are particularly flexible solids, requiring significant time to return to a resting state once excited. As with plate reverberators, stretched springs use electromagnetic transducers to convert electronic signals to physical motion and back

again. This method is both portable and inexpensive, and the low ratio of surface area to mass makes it difficult for vibrations in air to interfere, all of which explain their omnipresence in guitar amplifiers. Like plate reverberators, the output is dependent on the physical properties of the device, making it difficult to effectively alter once built (Parker, Bilbao 216-218).

Digital signal processing can be used to create virtual reverberators whose behavior is not restricted by the laws of physics. While these algorithms can be used to emulate the dispersion of sound energy either directly or through modeling any of the artificial reverberators described above, they can also be used to create reverberators which transcend these limitations. Manfred Schroeder's early investigations demonstrated that efficient reverberators can be built from simple algorithms such as comb filters and allpass filters; multiple stages of these sections can be cascaded to produce rich results while requiring relatively few computational operations and little memory (Schroeder, "Colorless Artificial Reverberation" 192-197).

Delay lines can be used to directly displace signals in time, creating discrete echoes. Controlled feedback of the output to the input can reuse delay lines to produce subsequent echoes requiring only an additional multiply and add. Adding multiple read points, or taps, per delay line can also increase the number of virtual reflections without proportionally increasing the amount of memory used. Delay lines of differing lengths can be interconnected at their feedback stage so that the output from each is added into the input of each. By allowing signals to be passed through multiple delay lines in any combination, this can efficiently model the echo density of higher order reflections and the interaction between multiple surfaces that they represent. An energy preserving matrix such as a Hadamard can be used to distribute each output to each input. This

will limit the resonances caused by repeated cycles through a single delay line and ensure that the feedback will be easy to control (Stautner, Puckette 52-54; Smith, “Physical Audio…” 63-65).

Unlike their physical counterparts, digital processing methods may readily be manipulated in real time, allowing for substantial control over the output. The widespread integration of these algorithms into existing sonic tools effectively eliminates the issue of portability. Still, all of these digital methods rely on conceptual models which mimic the behavior of acoustic spaces, creating algorithms in their image in hopes of producing similar results.

1.2 Convolution Reverberators

Convolution reverberation algorithms have the advantage over other digital techniques of being based directly upon measurements of an actual space. An impulse response serves as a complete description of any band-limited linear time-invariant system, and can therefore be used to make any recorded sound seem to exist within that space. While no room has perfectly linear behavior, any nonlinearities are often weak enough that a linear measurement can serve as a useful approximation.

High quality impulse responses are crucial to producing a good result with convolution reverberators. When impulse responses are convolved with musical material, any noise or error within them will also be convolved, adding noise to the result and diminishing the realism of the effect. There are a variety of disparate techniques for impulse response measurement, but each struggles when presented with substantial environmental noise. The available libraries of impulse responses therefore tend to describe spaces with unusually low levels of environmental noise, such as concert halls,

recording studios, and churches. While these are undoubtedly useful in audio production, they represent a minuscule subset of all available spaces.

This dissertation presents a novel method for collecting impulse responses that can produce excellent results in even the noisiest of spaces. The EchoThief Impulse Response Library, which was created using this technique, is discussed in Chapter 8.

2.0 Cross-Correlation and Convolution

Cross-correlation and convolution are two very similar operations, each the sum of the products of every sample in one signal by every sample in the other. The sole difference between the two is a reversal in time of one of the signals. The operations can therefore be substituted for one another if one of the input signals is reversed. These operations are essential in acoustic measurement and of particular use when working with complementary sequences.

Cross-correlation is a measure of the similarity of two signals as a function of relative time offset, a property which makes it useful in detecting echoes.

$$h(n) \star s(n) = \sum_{m=-\infty}^{\infty} h(m)s(n+m)$$

For every possible offset in time between the two signals, their time-aligned pairs of samples are multiplied and these products are summed (Smith “Mathematics…” 129).

Convolution is mathematically similar to cross-correlation, differing only in the reversal in time of one of the signals.

$$h(n) * s(n) = \sum_{m=-\infty}^{\infty} h(m)s(n-m)$$

Any linear, time-invariant system can be modeled as a convolution operation, making this especially useful when studying acoustics (Smith, “Digital Filters” 184).

2.1 Acyclic and Cyclic Convolution

Convolution is a combination of two signals as a function of their relative time alignment. Representing every potential overlap of two signals of lengths *M* and *N* will require *M+N-1* samples, totaling their combined length less the single sample required as the minimum overlap. Assuming that each input signal is more than one sample in length, this will produce an output whose length exceeds either input. How this discrepancy between input signal length and output signal length will be handled depends on assumptions made about the input signals themselves.

Each signal can be assumed to be a finite, non-repeating sequence. In this case, each is treated as though all values outside their finite range are zeros, and the output signal can extend its full length of *M+N-1*. This approach is called acyclic or linear convolution.

$$h(n) * s(n) = \sum_{m=0}^{M+N-1} h(m)s(n-m)$$

Alternately, one of the input signals can be assumed to repeat indefinitely, with its finite length representing a single period. The output will be periodic at the same period, with excess signal wrapped around modulo the period and summed with the rest of the signal. This approach is called cyclic or circular convolution, indicating its periodic nature.

$$h(n) \otimes s(n) = \sum_{m=0}^{N-1} h(m)s((n-m) mod\, N)$$

2.2 Fast Convolution

The Convolution Theorem shows that convolution in one domain (time or frequency) is equivalent to multiplication in the other (Smith, “Discrete Fourier Transform” 152). ‘Fast convolution’ is performed by computing the Fourier transform of each signal, multiplying the spectra, and calculating the inverse Fourier transform of the product.

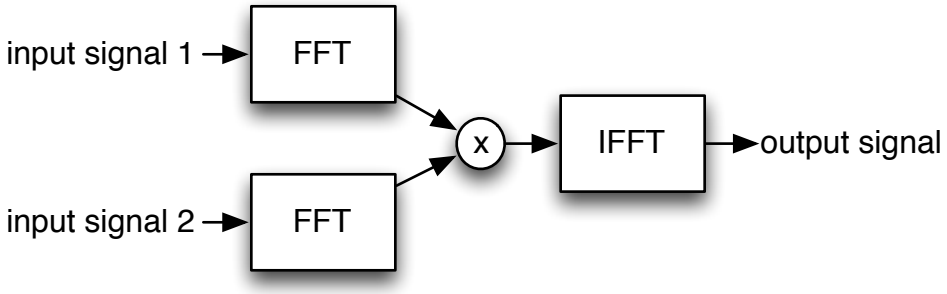

Figure 2.1: Fast convolution

As the output of a linear convolution operation will exceed the length of either input, care must be taken when using Fourier transforms to avoid cropping or wrapping part of the signal. The output of a linear convolution operation will be one sample less than the sum of the lengths of the inputs, so finding the smallest power of two larger than that sum will provide a size that is appropriate for all three transforms (two forward fast Fourier transforms and one inverse fast Fourier transform). Zero padding can be used to extend the signals to match the input size of the Fourier transform (Smith, “Discrete Fourier Transform”, 159).

Circular convolution produces an output the same length as its input, so zero padding is unnecessary. Accordingly, the size of the Fourier transform should be the same as the period of the cycle to ensure correct alignment of the wraparound.

2.3 Direct and Fast Convolution Speed

All convolution algorithms consist mostly of a large number of multiplies. Direct convolution employs the point-by-point multiplication of two arrays while fast convolution requires many multiplies to perform the forward and inverse Fourier transforms. The performance time of either algorithm is roughly proportional to the total number of multiplies required, so estimations of these can be used to compare their relative speeds.

Direct convolution of two real signals each of length N requires N^2 real multiplies, demonstrating a quadratic time complexity. Fast convolution has a linearithmic time complexity, a product of a linear and a logarithmic term ($Nlog_2N$). The two forward transforms and one inverse transform each consist of $Nlog_22N$ complex multiplies. Spectral multiplication requires an additional $2N$ complex multiplies. Each complex multiply consists of four real multiplies, for a final total of $12Nlog_22N+8N$ real multiplies (Burrus, Parks 90-95).

The quadratic time of direct convolution will be faster than the linearithmic time of fast convolution for only the briefest of signals (Strum, Kirk 519-521). All of the signals discussed here will have a length in the hundreds of thousands of samples; for signals of this length, fast convolution is significantly faster.

2.4 Mathematical Properties

Like their spectral cousin multiplication, both convolution and cross-correlation are commutative, associative, and distributive.

Commutativity allows operands to be reordered without changing the result.

$$a(n) * b(n) = b(n) * a(n)$$

Associativity shows that, when convolving three or more signals, the order of operations will not affect the result. Any pair of the input signals can be convolved to create an intermediate result which is then convolved with the third signal.

$$(a(n) * b(n)) * c(n) = a(n) * (b(n) * c(n))$$

A set of convolution operations cascaded in any order will reach the same result (less any considerations of numerical accuracy), an important distinction when considering the many stages of convolution represented by the typical acoustic measurement.

The distributive property demonstrates that the summation of two signals which have each been convolved with a common third signal is equivalent to the convolution of the sum of the original two signals with the third signal.

$$(a(n) * b(n)) + (a(n) * c(n)) = a(n) * (b(n) + c(n))$$

3.0 Impulses

The ideal impulse represented by the Dirac delta is both infinitely loud and infinitely brief, two properties which are, of course, unachievable in practice. Discretization of time at a uniform sampling rate and bit depth imposes finite limits on resolution in both amplitude and duration, making it trivial to define the discrete time approximation of this ideal impulse. The Kronecker delta, or unit impulse, is a piecewise function consisting of a single positive normalized sample at the origin and zeros everywhere else.

$$\delta(n) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}$$

This is the shortest impulse with the greatest amplitude that can be represented in the digital domain: one solitary normalized sample. It also serves as the identity element of the discrete convolution operation, returning the input unchanged as the output, making it a useful tool in measuring the acoustic properties of spaces. Three properties of the Kronecker delta make this possible: duration, amplitude, and spectrum.

As a single sample, the Kronecker delta represents the shortest duration and therefore the most precise temporal location that can be described within the limitations of the discrete medium: a single tick of the sample clock. Even acyclic convolution with the Kronecker delta will not smear the other input signal. Convolution with a signal of length *N* will produce a result of length *N*, regardless of whether acyclic or cyclic convolution is used.

Unit amplitude represents perfect normalization, the identity element of multiplication. A closer look at the indices, *m* and *n*, demonstrates how this causes the

Kronecker delta to be the identity element of the convolution operation. Replacing *s(n)* with $\delta(n)$ in the linear convolution formula produces

$$h(n) * \delta(n) = \sum_{m=0}^{M+N-1} h(m)\delta(n-m)$$

The Kronecker delta has a nonzero value only at time 0, so $\delta(n-m)$ is nonzero only when *m=n*; for all other values of *m*, the zero will nullify the product. This removes the summation and reduces the equation to

$$h(n) * \delta(n) = h(n)$$

An ideal impulse contains equal proportions of every frequency, giving it a flat spectrum. The same is true of the discrete time unit impulse, although it is band-limited to the Nyquist frequency. Each Fourier bin will therefore have the same value, so multiplication by this constant will not alter the spectral balance of the result.

3.1 Impulse Responses

An impulse response is the response of a system to an impulse. When the system is linear and time-invariant (LTI), this impulse response can be used to reproduce the effect of the space. The systems measured in this document are exclusively acoustical spaces, but the same process will work for any LTI system. The resulting impulse response describes the band-limited reverberation of the system completely (Smith “Digital Filters” 100-101). If the behavior of an LTI system in response to a single impulse is known, from it can be inferred its response to any sequence of

impulses, for example a digitally sampled sound. It is this property that makes this technique useful for artificial reverberation in the digital domain.

3.2 Anatomy of an Impulse Response

An impulse is an idealization of a compression wave traveling outward from its point source, a sphere expanding at the speed of sound. Each time it encounters a boundary, some portion of the energy is absorbed and the remainder is reflected back, folding the sphere back on itself. In any reasonably reverberant space, such reflected portions of this sphere will pass a nearby listener hundreds of times from many directions before its energy has dissipated below the threshold of audibility.

The listener will first hear the direct sound of the impulse, followed soon thereafter by reflections, then reflections of reflections, and so on. These reflections are categorized by the number of surfaces they have reflected against: first order reflections have reflected against a single surface before reaching the listener; second order reflections have reflected twice. Each reflection represents a loss of overall energy.

Plotting the amplitude of these reflections on a time axis (Figure 3.1) illustrates several characteristics common to all impulse responses. The direct sound necessarily has the shortest path to travel, making it the first to arrive at the listener; the timeline begins here with t_0. Reflected sounds necessarily take a longer path than the direct sound and therefore arrive later, beginning at t_1. The period of time between the arrival of the direct sound (t_0) and the arrival of the first reflection (t_1) is referred to as the initial time delay gap (Brook, Uzzle 173). Further early reflections quickly follow the first, their increasing number reducing the time between arrivals. Initially, these may be discernible as individual echoes. As more echoes arrive and the interval between subsequent

arrivals diminishes, this increased echo density makes it impossible to distinguish between individual reflections. Here t_2 marks transition from discrete early reflections to the wash of late reflections, typically on the order of 70-80 milliseconds from t_0 (Brook, Uzzle 174).

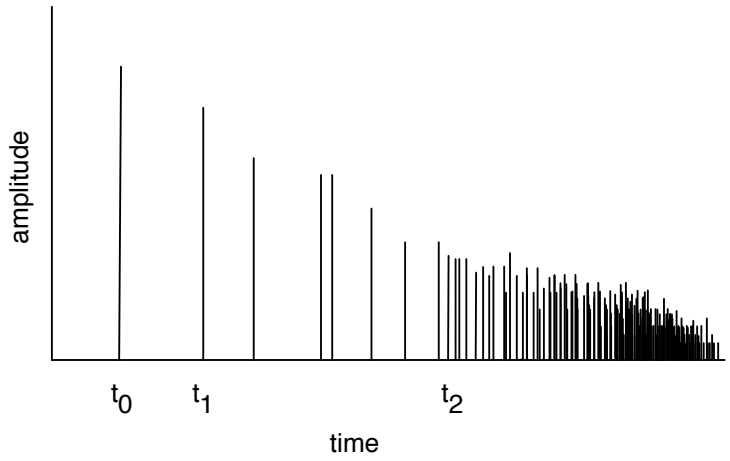

Figure 3.1: Relative amplitudes of individual reflections

Impulse responses of acoustic spaces all share the same overall amplitude envelope shape: excitation of the space creates a quick crescendo to an early peak, which is followed by an exponential decay as this energy dissipates.

The direct sound necessarily has the greatest amplitude of any single element of the impulse response. All of the reflections have lost more energy by traveling a greater distance and reflecting off a surface, losing energy in the transaction. However, if early reflections are aligned in such a way as to reinforce one another, they can collectively produce a louder peak than the direct sound.

In most enclosures, the late reflections are far greater in number than the early reflections. As a result, they sum more predictably, following an exponential curve as the

amplitude of the response wanes. The greater the echo density of the space, the more closely this will match an ideal exponential curve.

The arrival of direct sound provides an unambiguous marker identifying the beginning of the impulse response. However, the exponential decay of energy makes it difficult to define a precise end point of the impulse response. Sabine's early investigations into reverberation time illustrate this difficulty. Lacking any precision equipment with which to measure sound pressure level, he relied upon his ears to determine when the decaying reverberation of his test tone crossed the threshold of audibility, marking this time with a stopwatch (Sabine 10). Results obtained through this technique vary greatly depending on the spectral content of the measurement signal, the acuity of the listener's hearing, and the reverberance and noise floor of the room.

Rather than rely on the sensitivity of the measurement technique or the noise floor of the space, either of which may be highly variable, a simple ratio can be used to determine the effective end of the impulse response. The t_{60} of an impulse response (also sometimes written as RT_{60}) is the duration from the peak amplitude to the point at which the amplitude has decreased by 60 dB. This represents a diminishing to one thousandth of the peak amplitude and one millionth of the peak power. This duration may be as brief as a handful of samples or may be as long as several seconds.

Most LTI systems will respond differently to different input frequencies. For example, acoustical spaces tend to absorb higher frequencies more quickly than lower frequencies. A broadband t_{60} can establish the overall effective duration of an impulse response, but a more useful description will include the time it takes for energy to dissipate at multiple specific frequencies. A commonly used set of measurement frequencies are the six octaves of 125 Hz, 250 Hz, 500 Hz, 1 kHz, 2 kHz, and 4 kHz.

These frequencies span six of the ten octaves of human hearing, focusing on the more easily perceptible (logarithmic) center of the spectrum while disregarding the extremes on either end. A single value known as the midrange average may also be used, which is typically an average of the 500 Hz and 1 kHz times. This narrow focus is most appropriate when measuring spaces primarily used for speech (Davis 1175-1178)

3.3 Absorption

The Sabine Equation describes the relationship between a space's volume (*V*) in m^3, total surface area (*S*) in m^2, average absorption coefficient (*a*), and t_{60} in seconds.

$$t_{60} = (0.161 \cdot V)/(S \cdot a)$$

The product *Sa,* representing the total surface absorption, is given the unit of metric Sabins (Everest 99).

Spaces which are composed of more than one material can be modeled by dividing the area into sections and multiplying each by its own absorption coefficient.

$$t_{60} = (0.161 \cdot V)/((S_1 \cdot a_1) + (S_2 \cdot a_2) + ... + (S_n \cdot a_n))$$

4.0 Impulse Response Measurement

The impulse response of a room can be measured as easily as clapping in the space and recording the result. This loud, brief sound is a fair approximation of an idealized impulse. Several more elaborate techniques for impulse response measurement exist which differ only in the methods used to create the impulse. These methods can be divided into two types: physical and electronic.

4.1 Physical Methods

Physical methods employ a loud, brief sound as the impulse. All environmental sounds present at the recording location will be included in the impulse response, so the signal-to-noise ratio of the result will depend on the amplitude of the impulse relative to the ambient noise level. For this reason, louder impulse methods are preferable. Commonly used physical impulses include hand claps, balloon pops, gunshots, or spark gaps. However, none of these impulses has the flat spectrum ideal for measurement, so resulting impulse responses will retain the frequency character of the impulse. While our ears may perceive these impulses as instantaneous, each of these is substantially longer in duration than an idealized impulse. Moreover, the spectral envelope of each sound changes over its duration, further complicating the results. The further this impulse is from an ideal Kronecker delta, the more its character will be evident in the resulting impulse response.

Table 4.1: Sound pressure levels of impulses

impulse	dB SPL
loud hand clapping at 1m distance	130
balloon pop	157
handgun	166

Table 4.2: Sound pressure levels of environments

environment	dB SPL
auditory threshold	0
quiet library	30
average home	50
normal conversation	60

4.2 Electronic Methods

Electronic methods of impulse response measurement are necessarily more abstract. Speakers are poorly designed to produce individual forceful impulses without also producing continuing oscillations. Playing a single normalized sample through a sound system at a volume comparable to the physical methods above would require substantial amplification and introduce a further set of noise. A better solution would be to smear the energy of a single impulse over time so that it all remains, but doesn't all have to occur at once. This allows an enormous amount of energy to be contained within the dynamic limitations of a normalized digital audio system. Several methods

which decompose a powerful virtual impulse into a signal of normalized amplitude are described below.

4.3 Time Delay Spectrometry

Time delay spectrometry uses a sine wave with a fixed amplitude and a frequency which sweeps across the entire range of human hearing at a constant rate. The Fourier transform of the recorded result is divided by that of the original signal to produce an impulse response of the space.

$$y(n) = h(n) * x(n)$$
$$Y(k) = H(k)X(k)$$
$$H(k) = Y(k)/X(k)$$

This linear 'chirp' signal spends an equal amount of time at each frequency, assuring an overall flat spectrum across its range. However, human perception of pitch is logarithmic, with each successive octave spanning twice the frequency range of the one below it. The linear chirp spends half of its measurement time in the highest octave of human hearing, a quarter in the octave below, an eighth in the next, and so on. The results will consequently be more accurate for higher frequencies than lower ones.

4.4 Exponential Sine Sweeps

Exponential sine sweeps (ESS) rise in frequency exponentially, spending an equal amount of time within each octave. This is more congruent with our perception of pitch than the linear chirp. However, since this does not spend an equal amount of time

on each frequency, it does not have a flat spectrum like an ideal impulse or linear chirp. Post-processing equalization to correct this imbalance will introduce time smearing.

4.5 Maximum Length Sequences

Maximum length sequences (MLS) are bit sequences whose autocorrelation is a scaled periodic Kronecker delta function, missing only the energy at DC (0 Hz) (Schroeder “Integrated Impulse” 497). When the response of a system to this signal is cross-correlated with the sequence, the result is an impulse response of the system. These sequences can be generated using exclusive-or (XOR) linear feedback shift registers and efficiently cross-correlated using a fast Walsh-Hadamard transform (Borish and Angell 479). MLS are often referred to as pseudo-random white noise, indicating their flat spectrum, except for the missing DC term (Borish and Angell 484). Since these sequences contain only values of ± 1, they have the advantage of having the minimum possible crest factor (the ratio of peak value to quadratic mean), making full use of all available dynamic range (Vanderkooy 220). By comparison, normalized sinusoidal methods have a crest factor of $\sqrt{2}$.

4.6 Complementary Sequences

Complementary sequences (CS) are pairs of sequences of equal length which have the special property that their respective autocorrelations sum to zero for all values except for a single impulse. By convention, these are designated the *A* and *B* sequences. The sum of their autocorrelations is a Kronecker delta function scaled in height to the total length of the two sequences combined (Golay, "Complementary Series" 82).

$$(A \star A) + (B \star B) = 2N\delta$$

where *N* is the length of each sequence.

First described by mathematician and information theorist Marcel Golay, complementary sequences were initially used in multislit spectrometry, investigating the spectral patterns of light. Scott Foster, a researcher at Hewlett-Packard Laboratories, demonstrated their use in impulse response measurement of acoustic spaces (Foster 1986). Their ability to use normalized signals to create an enormous virtual impulse indicates a potential use in situations where a suitable signal to noise ratio would be difficult to achieve by other means.

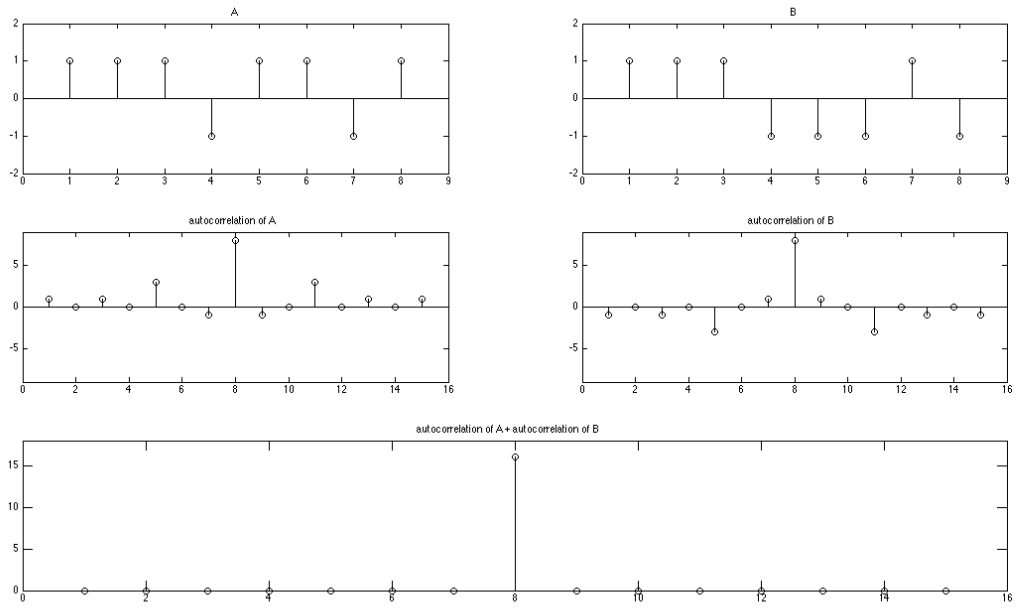

Figure 4.1: A_3 and B_3 pair, their linear autocorrelations, and the resulting sum

5.0 Measurement Using Complementary Sequences

To measure a space, the *A* and *B* sequences are each individually played into it while recording the respective responses. Care should be taken to ensure that the excitation from one sequence has dissipated before beginning the next sequence so that the measurements do not corrupt one another. If the system is both linear and time-invariant, these responses will be a convolution of the respective code and the space itself, *h*.

$$A * h = A_h$$
$$B * h = B_h$$

Each response is then cross-correlated with its corresponding original sequence and the results are summed.

$$\begin{aligned}(A \star A_h) + (B \star B_h) = \\ (A \star (A * h)) + (B \star (B * h)) = \\ ((A \star A) + (B \star B)) * h = \\ 2N\delta * h\end{aligned}$$

This result is *2Nδ*h*, the impulse response of the space convolved with a Kronecker delta function scaled by the total length of the two sequences.

Normalization can remove this scaling to ensure that each sample value *s* lies within the maximum limits of digital audio, $-1 \leq s \leq 1$. Since the Kronecker delta is the identity element of the convolution operation, it can be removed from the equation,

leaving a normalized impulse response of the space itself, *h* (Golay, "Complementary Series" 82-83).

5.1 Creating Complementary Pairs

Complementary pairs of increasing lengths can be created iteratively beginning with a common seed. If this seed is set to 1, the result will be a normalized bilevel sequence.

$$A_0 = [1]$$
$$B_0 = [1]$$

Further generations can be created iteratively by concatenating earlier generations and, in the case of the *B* sequence, inverting the latter half (Golay, "Complementary Series" 83).

$$A_{n+1} = [A_n \;\; B_n]$$
$$B_{n+1} = [A_n \; -B_n]$$

In this manner, sequences of any length 2^n can be created. Solutions also exist for any length $2^a 10^b 26^c$, where *a*, *b*, and *c* are integers, but are beyond the scope of this document (Golay, "Complementary Series" 83; Golay, "Note on Complementary Series" 84; Borwein 112).

Table 5.1: A and B sequences for generations 1-4

generation	length(samples)	A	B
1	2	[+ +]	[+ -]
2	4	[+ + + -]	[+ + - +]
3	8	[+ + + - + + - +]	[+ + + - - -+ -]
4	16	[+ + + - + + - + + + + + - - - + -]	[+ + + - + + - + - - - + + + - +]

In addition to concatenation, new sequences can be created by interleaving earlier generations (Golay, "Complementary Series" 83).

$$A_{n+1} = \{A_n \;\; B_n\}$$
$$B_{n+1} = \{A_n \; -B_n\}$$

Pairs can also be created by finding the sum and difference of sequences (Golay, "Complementary Series" 83).

$$A_{n+1} = A_n + B_n$$
$$B_{n+1} = A_n - B_n$$

Like the other methods, this will double the length relative to the previous generation. However, half of each sequence will consist of zeros (the second half of the *A* sequence and the first half of the *B* sequence, if the prior generations were created using the concatenation method), similar to zero-padding. Addition and subtraction can introduce values other than the seed; a single iteration of this process will double all of the nonzero values in both sequences.

5.2 Transforming Complementary Pairs

Several operations are described below which can be used to transform pairs while retaining their complementary properties (Golay 83-86; Tseng and Liu 644-648).

reversal:

$$A \mapsto \overline{A}$$
$$B \mapsto \overline{B}$$

inversion:

$$A \mapsto -A$$
$$B \mapsto -B$$

scaling:

$$A \mapsto Ax$$
$$B \mapsto Bx$$

alternating inversion:

$$A \mapsto (-1)^n A$$
$$B \mapsto (-1)^n B$$

integer upsampling:

$$A \mapsto U(A)$$
$$B \mapsto U(B)$$

integer downsampling:

$$A \mapsto D(A)$$
$$B \mapsto D(B)$$

rotation:

$$A \mapsto S_m(A)$$
$$B \mapsto S_m(B)$$

Pairs modified in these ways will remain complementary, but their output may be altered in the process. For example, scaling both sequences will produce an impulse response scaled by twice the amount. In many cases, the modification can be performed on one sequence before cross-correlation and after cross-correlation on its pair. Inversion of one sequence, for example, can be cancelled by inverting the cross-correlation of the other.

5.3 Multiple Measurements

Architecture remains stationary, while noise is by definition changing. Repeated measurements of the same space taken at a fixed location will differ principally in their noise. This suggests an opportunity for differentiation between the effects of a room on a signal and any noise accumulated during the measurement process (Schroeder, “Integrated Impulse Method" 499). Averaging multiple measurements will average their noise. Assuming the noise to be reasonably decorrelated, averaging will reduce the overall amplitude of this noise, producing a more accurate impulse response (Foster 931).

5.4 Zero Padding

If a measurement is to be repeated through concatenation, its period *N* must exceed the t_{60} of the space to be measured in order to prevent each measurement from sustaining into the following period and corrupting it. If the duration of each sequence is less than that of this period, zero padding can be used to separate the sequences.

Foster used complementary sequences each 64 samples in length at a sample rate of 50 kHz, repeating the measurement 1000 times at intervals of 20 ms (932). This relatively short period was sufficient to measure the impulse response of a microphone placed directly in front of a speaker, ignoring the effects of the room to realize an extremely brief t_{60}. This experiment can be reproduced as a single sequence to illustrate the process of concatenation and zero padding.

Each period of this sequence is 20 ms in duration, 1000 samples at the 50 kHz sample rate. The first 64 of these consist of one of the complementary sequences with the remaining 934 samples all set to zero.

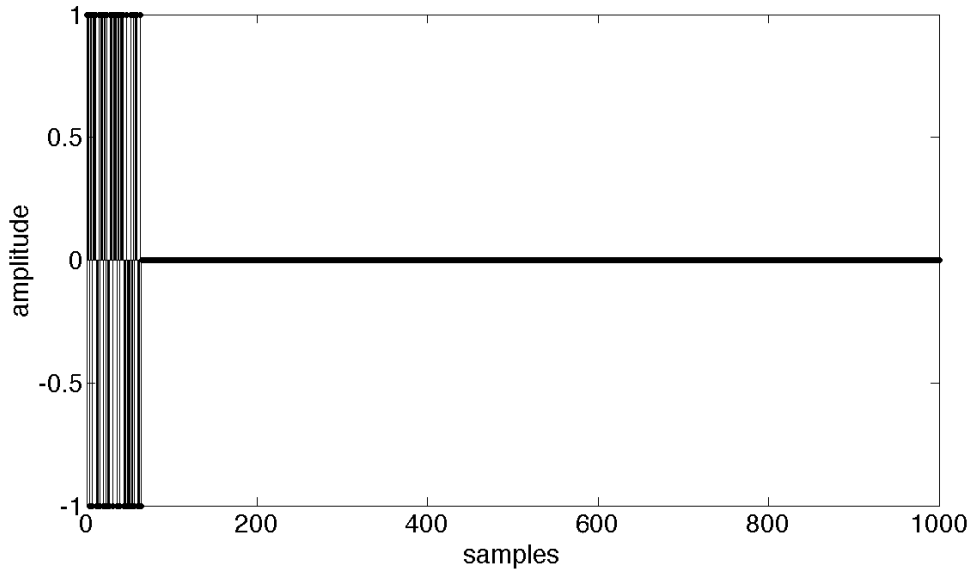

Figure 5.1: Foster measurement signal, single period

This period is repeated 1000 times with an equal number of A and B sequences. The total length of the resulting sequence is 1,000,000 samples (20 seconds at 50 kHz).

5.5 Concatenation

The impulse created by summation of autocorrelated normalized complementary sequences is equal in magnitude to the total length of both sequences, $2N\delta * h$. Longer sequences will produce an impulse of greater magnitude but require considerably more computation.

When Foster performed his experiment in 1986, the processor speed available at the time would make computing the results of signals substantially longer than 64 samples prohibitive. Modern processors are capable of processing much longer sequences with ease.

5.6 Periodic Sequences

Complementary pairs circularly rotated by equivalent amounts remain complementary. This periodicity allows sequences to be concatenated.

An impulse response produced by a complementary pair using cyclic convolution will be equal in length to either sequence. Resonance beyond this duration will wrap around circularly and be added to the beginning of the result. By overlapping the beginning of the impulse response, this makes the signal non-causal; from the perspective of the listener, which is necessarily delayed from the onset of the sample, the reverberation will occur prior to the sound which is causing it. To accurately measure

the resonance of a space without overlap, the period of concatenated sequences should substantially exceed the t_{60} of the space.

Periodic sequences created through direct concatenation have the obvious advantage over zero padded sequences of containing more energy, therefore resulting in impulse of greater magnitude. Since each instance (except the first) is preceded by an identical copy, the spillover from the previous period is identical to the wraparound from the period itself and can be safely truncated. Since the initial instance of each will be discarded (lacking a prior sequence and therefore circularity), a concatenation of *n* copies of *A* will produce *n-1* segments.

The *A* and *B* sequences can themselves be concatenated to produce a single sequence, simplifying the measurement process and enforcing consistent timing between complementary pairs. As above, the initial instance of each will not have equivalent wraparound and should therefore be discarded.

It is desirable to play the *A* and *B* sequences as close in time to one another as possible, in hopes of minimizing measurement error caused by gradual changes in the space. For example, a variance in temperature can significantly alter the acoustic properties of a space (Sabine 124-125). If each extracted *A* sequence is paired with its nearest *B* sequence, their relative temporal displacement will be determined by the lengths of each as well as the total number of repetitions.

5.7 Synchronization

The autocorrelations of complementary sequences have sidelobes which are perfectly out of phase. When summed, they will cancel each other out, producing a Kronecker delta. Any offset in time between the sequences will disrupt this balance and

degrade the result. The example below, second generation sequences are offset by a single sample before summation.

$$A_2 = [1\ 1\ 1\ -1]$$
$$B_2 = [1\ 1\ -1\ 1]$$
$$A_2 \star A_2 = [-1\ 0\ 1\ 4\ 1\ 0\ -1]$$
$$B_2 \star B_2 = [1\ 0\ -1\ 4\ -1\ 0\ 1]$$
$$(A_2 \star A_2) + (B_2 \star B_2) = [0\ 0\ 0\ 8\ 0\ 0\ 0]$$

Figure 5.2: Cross-correlations summed with no offset

$$S_1(A_2 \star A_2) = [0\ 1\ 4\ 1\ 0\ -1\ -1]$$
$$B_2 \star B_2 = [1\ 0\ -1\ 4\ -1\ 0\ 1]$$
$$S_1(A_2 \star A_2) + (B_2 \star B_2) = [1\ 1\ 3\ 5\ -1\ -1\ 0]$$

Figure 5.3: Cross-correlations summed with 1 sample offset

5.8 Averaging

Assuming consistent placement and volume, repeated measurements of a system will differ primarily in their noise. Averaging can be used to lower the overall noise level.

The noise found in impulse responses can be caused by several factors. Environmental noise present in the space being measured is likely to be the largest source. Amplification of electronic signals and their transmission along cables will introduce noise due to electromagnetic interference. The process of digitization will introduce quantization noise as signals are rounded to the nearest digital value. Quantization noise can largely be mitigated by increasing the word length (bit depth) used in the measurement process to reduce the rounding error. Foster warns that if the

quantization noise is not significantly quieter than the environmental noise it will not be independent from measurement to measurement and therefore not average away (932).

6.0 Example Implementation

The sequence B_{20} can be bisected to produce A_{19} and $-B_{19}$.

$$B_{20} = [\ A_{19}\ -B_{19}\ ]$$

Once the latter is negated, they form a complementary pair, A_{19} and B_{19}. Each of the component sequences is 524,288 samples (11.88 seconds at a 44100 Hz sampling rate), giving the overall concatenation a duration of 1,048,576 samples (23.78 seconds).

The inversion of the B_{19} sequence ensures that both component sequences contain B_{18} as their second half. Decomposing B_{20} into earlier generations illustrates this.

$$B_{20} = [\ A_{19}\ -B_{19}\ ] = [\ A_{18}\ B_{18}\ -A_{18}\ B_{18}\ ]$$

As a sequence is passed through a room, the resonance smears each sequence into the next. So long as the resonance does not exceed the period of the B_{18} sequence of 5.94 seconds, the spillover between 19th generation sequences will be equivalent. The common latter half allows the sequences to be directly concatenated without the need for an intermediary buffer, shortening the overall measurement time and reducing the time between pairs to 11.88 seconds. Most acoustic spaces have a t_{60} significantly shorter than half this, 5.94 seconds, making this sequence a good generalized solution.

Multiple B_{20} sequences can be concatenated, creating as many 19th-generation complementary pairs to be averaged. Schroeder demonstrated a similar process using repeated maximum length sequences, playing a periodic maximum length sequence at an unobtrusive level during his lectures on acoustics while displaying the cumulative

result (“Integrated Impulse…” 499). He warns of the difficulty of synchronization, proposing a simultaneous synch tone to detect tape stretch in analog recordings. Similar issues remain a concern when using complementary sequences, where clock drift can cause misalignment of pairs. Pairs extracted from within the same period are offset by half the period. Regardless of the overall length of the resulting sequence, the extracted pairs will retain a consistent temporal displacement from one another of 11.88 seconds, a brief enough duration to expect little drift.

The circularity of the compound sequences suggests a more concise method for extracting a large set of complementary pairs without proportionally extending the sequence. Considering the B_{20} sequence as a loop where N is the length of each 19th generation sequence (2^{19} samples), A_{19} and B_{19} pairs can be extracted at any integer rotation up to the length $N/2$ (the length of B_{18}). Once passed through a resonant system, the maximum rotation is reduced by the length of the resonance. Presuming a generous resonance where the t_{60} is $N/4$ (2.97 seconds), this provides a total of 131,072 possible pairs.

Circular rotation in this manner will introduce a discontinuity where the beginning of the sequence is concatenated with the end. To avoid this, a concatenation of two B_{20} sequences can be used, providing circularity within the signal itself so that a time offset is equivalent to rotation.

When one of these rotated sequences is cross-correlated with its unrotated counterpart, the result will share the rotation of the former. Each result should therefore be rotated back by an equal number of samples to align it, synchronizing the whole set of impulse responses before averaging them. Alternately, a set of correspondingly

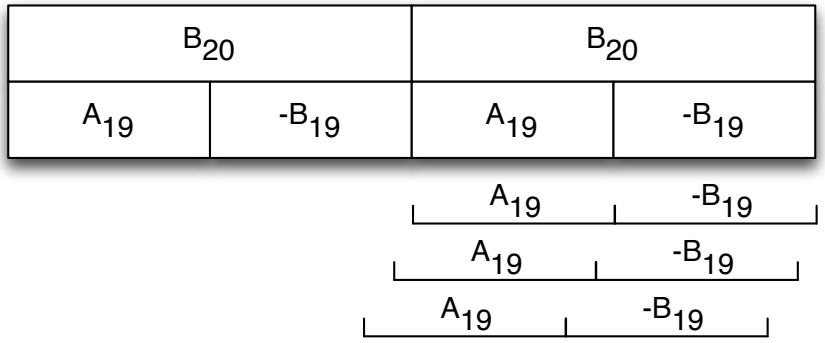

Figure 6.1: Extracting complementary pairs from concatenated B_{20} sequences

rotated sequences can be precompiled and Fourier transformed to reduce performance time.

It is neither practical nor necessary to compute an impulse response for every possible rotation. A more manageable subset of these can be selected based on the specific needs of the situation. For the experiment described in Chapter 7, 40 complementary pairs are extracted at intervals of 2000 samples.

6.1 Steady State

When energy is added to a system in the form of a measurement signal, the space will require some brief duration to reach a steady state. Lacking any preceding signal, measurement signal collected during this time cannot be circular, and will therefore need to be discarded. This duration will be at least as long as the t_{60} of the space. A full period of the B_{20} sequence will provide more than adequate time for most systems to reach steady state.

7.0 An Experimental Result

A comparison is made between three methods of impulse response measurement: exponential sine sweeps, maximum length sequences, and complementary sequences. A fourth method, popping a balloon, is included to provide a comparison between electronic and physical excitation methods. Inconsistencies inherent to this method (variable air pressure between balloons, the presence of a person in the room) should caution against too close a comparison with the other three methods.

7.1 Measurement Sequences

To accurately compare the three electronic methods, measurement sequences of equivalent length are used. Sequences have a period of 2^{20} samples (2^{20}-1 in the case of MLS). Two periods of each is played, the first to excite the room to steady state while the second is analyzed to produce the measured result (although some portion of the first period is used in the case of CS). This period is substantially longer than the t_{60} of the space, so shorter excerpts would suffice in exciting the space. However, full periods were employed to ensure equitable comparison between all methods. Each measurement signal has an overall duration of approximately 48 seconds followed by 10 seconds of silence.

The exponential sinusoid sweeps from 20 Hz to 20 kHz over 2^{20} samples. The maximum length sequence is a 20^{th} order bit sequence, producing a period of 2^{20}-1 samples. The complementary sequence is the periodic B_{20} sequence specified in Chapter 6.

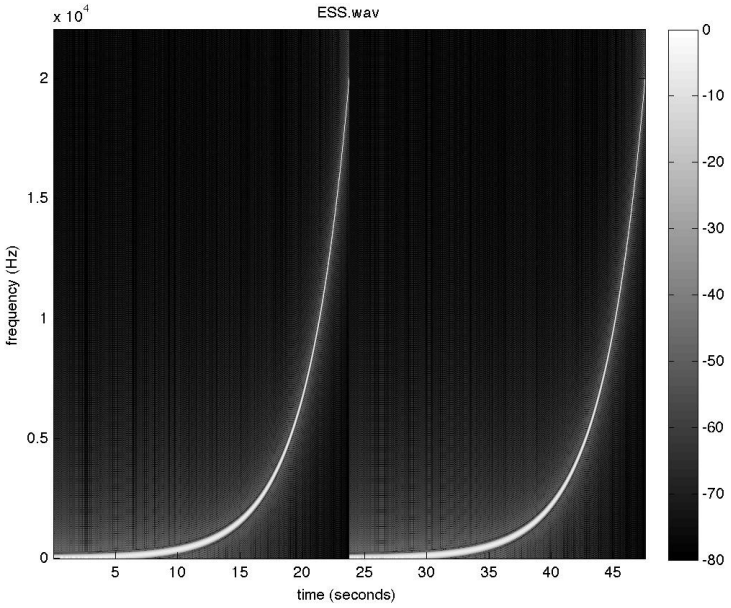

Figure 7.1: Spectrogram of exponential sine sweep (ESS)

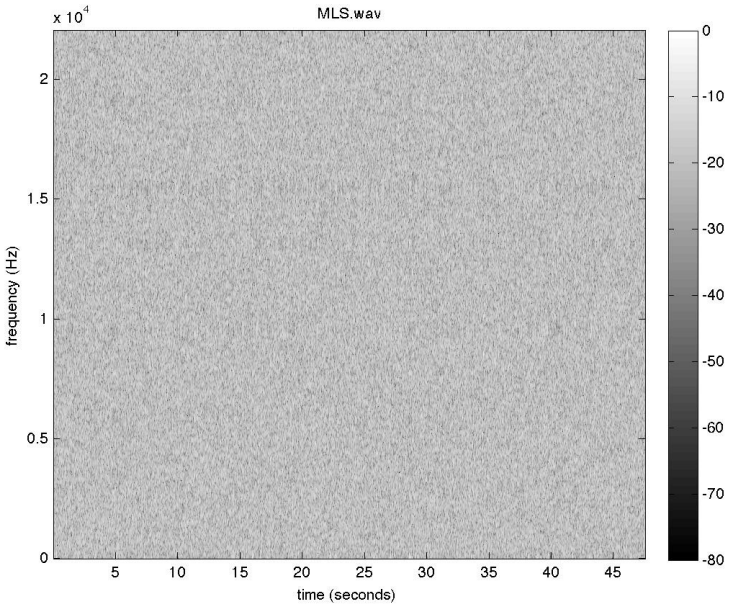

Figure 7.2: Spectrogram of maximum length sequence (MLS)

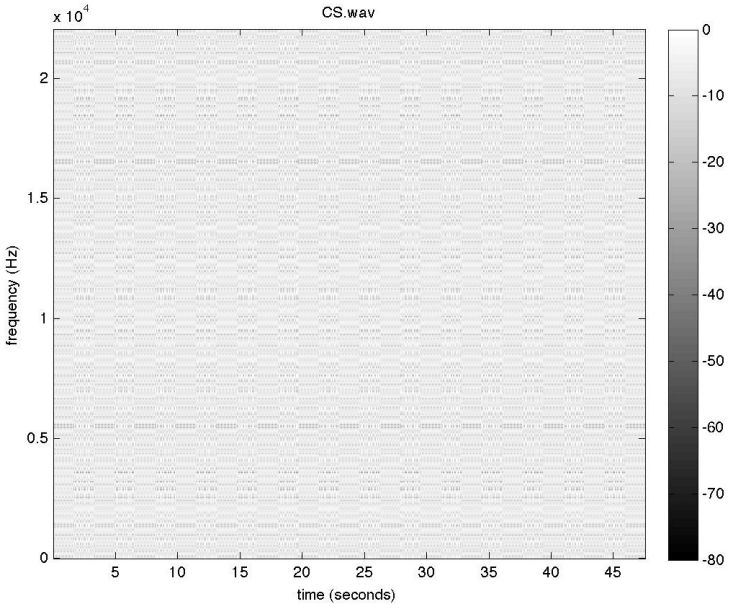

Figure 7.3: Spectrogram of complementary sequence (CS)

7.2 Noise

Each measurement method is repeated under five different noise conditions. The first measurement adds no environmental noise. The next three measurements add a single set of pseudo-randomly generated white noise presented at three dynamic levels (respective to the normalized measurement sequences): full normalization, -6 dB, and -12 dB. The final measurement adds a musical excerpt from the bridge of Megadeth's "Rust in Peace... Polaris" from the album *Rust in Peace*, chosen for its abrasive layered guitars and hyperactive kick drum.

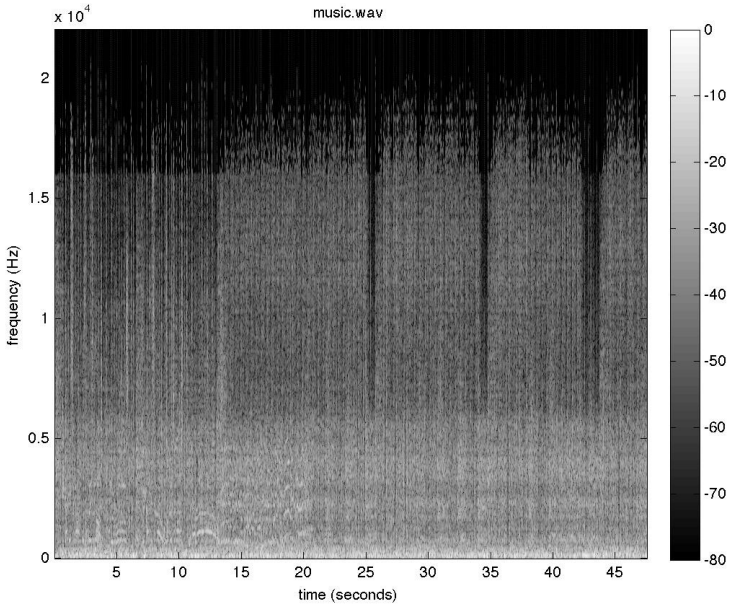

Figure 7.4: Spectrogram of "Rust in Peace... Polaris"

7.3 Location

The space used for this experiment is the recording studio in the Conrad Prebys Music Center at UC San Diego, with an impressively low noise floor of 28.7 dBA. Its floor plan is very nearly square, measuring approximately 9.4 by 9.3 meters with a ceiling at approximately 6.6 meters. Angled panels along the walls and ceiling diffuse reflections to mitigate standing waves. Absorbent panels are installed opposing reflective wooden walls, further inhibiting strong resonances. The space has a t_{60} of approximately half a second.

7.4 Equipment

The hardware described in the first configuration in Appendix I was used, providing stereo results at a 44100 Hz sampling rate and 24 bit word length. The microphone pair and measurement speaker are placed 5 meters apart, 1.5 and 1 meters from the ground respectively. A second speaker placed near the first provides noise signals. The balloons are popped immediately above the measurement speaker. A sound pressure level meter is placed between the microphones to compare the levels of each measurement technique.

7.5 Measurements

Each signal is played through the measurement speaker once to check the sound level present at the microphones. The maximum length sequence and complementary sequence predictably produce nearly the same level, 79.9 and 79.8 dBA respectively, approximately 50 dB above the noise floor of the space. The sine sweep is more difficult to accurately measure with frequency-dependent dBA weighting, but multiple measurements throughout the full range of the sweep average to 76.1 dBA, consistent with the lower crest factor of sinusoids relative to normalized bit sequences (expected to be 3 dB below the normalized bit sequences). The balloon pop produces a peak SPL of 86.5 dBA.

7.6 Comparison of Environmental Noises

The amplitude of an impulse response decays approximately exponentially until reaching the noise floor. In all of the test conditions, the noise floor was reached within the first second. RMS measurements of the responses beginning at one second and

lasting for one second are used to compare each measurement method against itself across four of the noise situations.

The balloon pop can serve as a rough baseline for comparison, as environmental noise is directly summed with the exciting impulse without any signal processing. Without added noise, it produces a noise floor of -76.4 dB. Adding normalized white noise raises this to -44.3 dB. Reducing the level of the white noise by 6 dB lowers the noise floor by 5.6 dB. A second reduction by the same amount lowers the noise floor by 13 dB, but this unexpected difference could likely be due to inconsistent levels produced by each balloon.

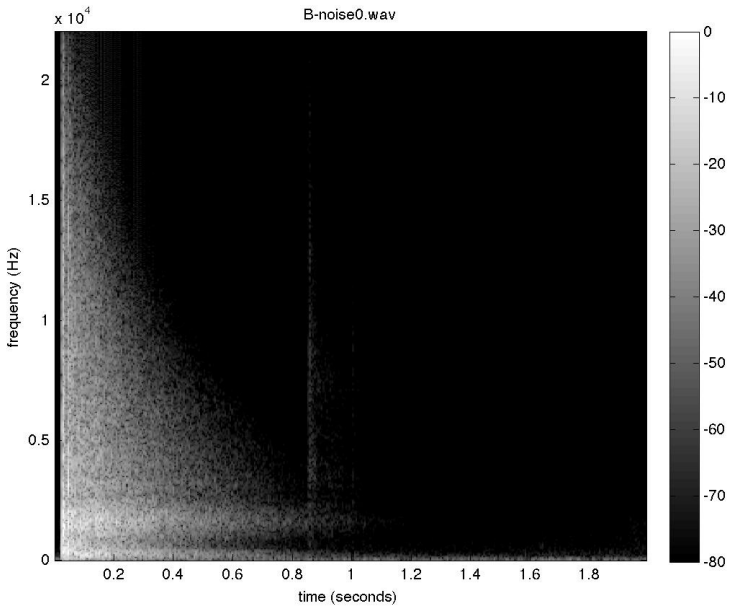

Figure 7.5: Spectrogram of balloon pop with no added noise

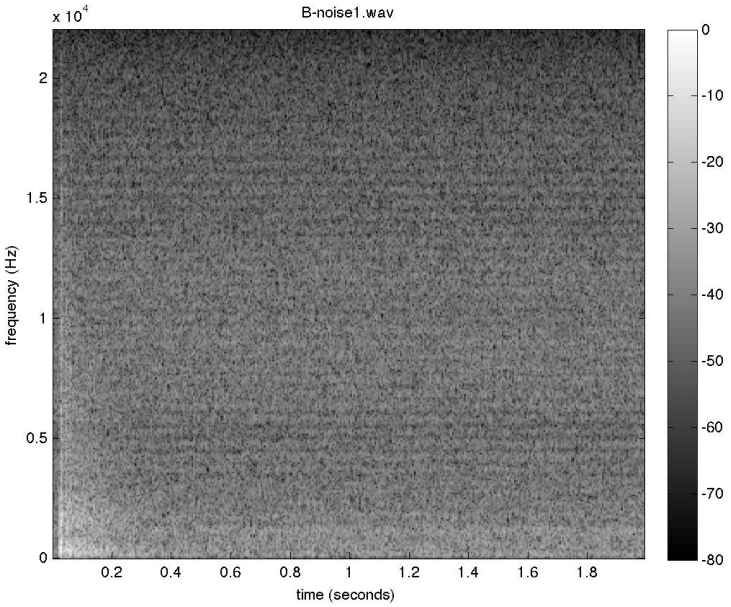

Figure 7.6: Spectrogram of balloon pop plus 0 dB white noise

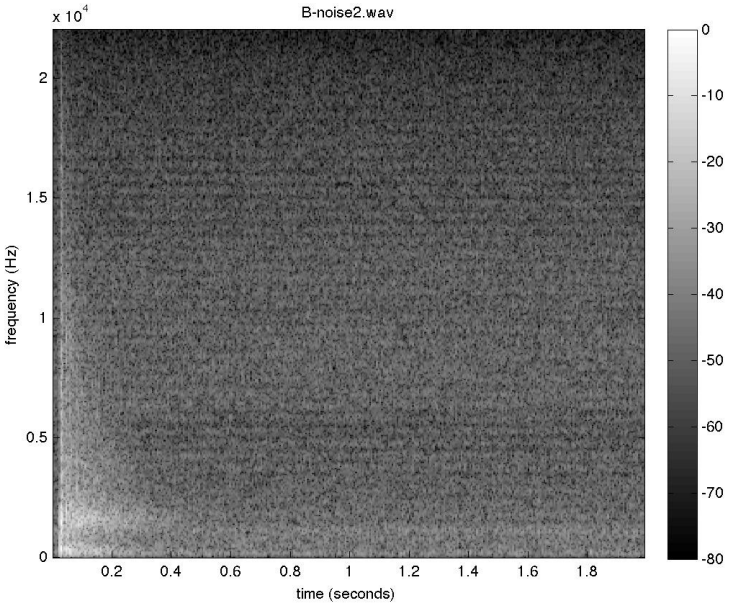

Figure 7.7: Spectrogram of balloon pop plus -6 dB white noise

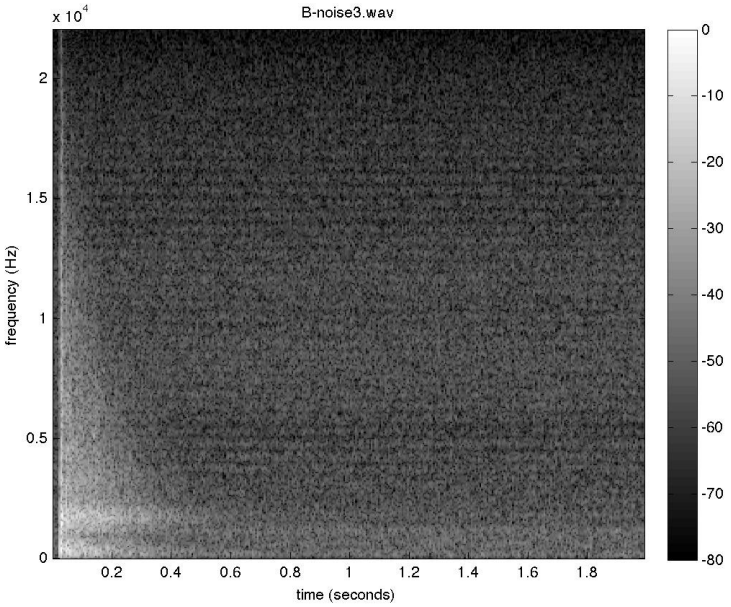

Figure 7.8: Spectrogram of balloon pop plus -12 dB white noise

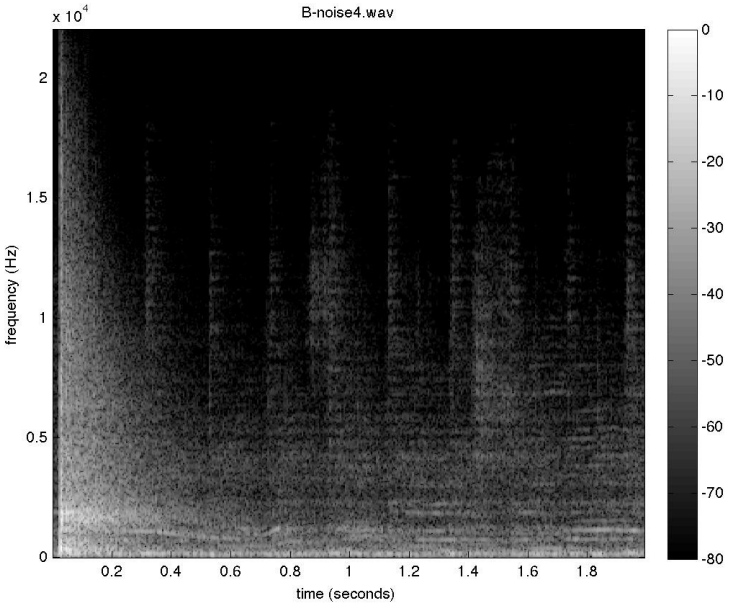

Figure 7.9: Spectrogram of balloon pop plus music

The results produced by electronic signals follow this pattern, with noise floors decreasing directly with lowered environmental noise. Complementary sequences provide the most consistent reduction, with each 6 dB lowering of the noise level matched by a 5.8 dB decrease in the noise floor.

Examining the impulse responses generated with MLS, several false impulses are evident after the noise floor has been reached (at approximately 63,000, 68,000 and 84,000 samples). These are obscured by the higher noise floor when normalized white noise is added, but apparent in the rest of the impulse responses. Since these occur after the noise floor has been reached, they can safely be truncated. However, in spaces with a longer t_{60}, this can be a concern.

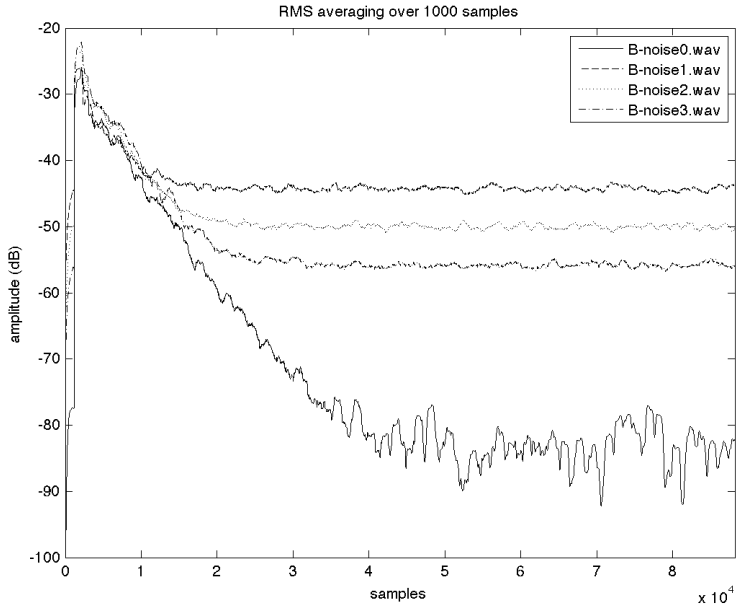

Figure 7.10: Balloon pop plus added noise

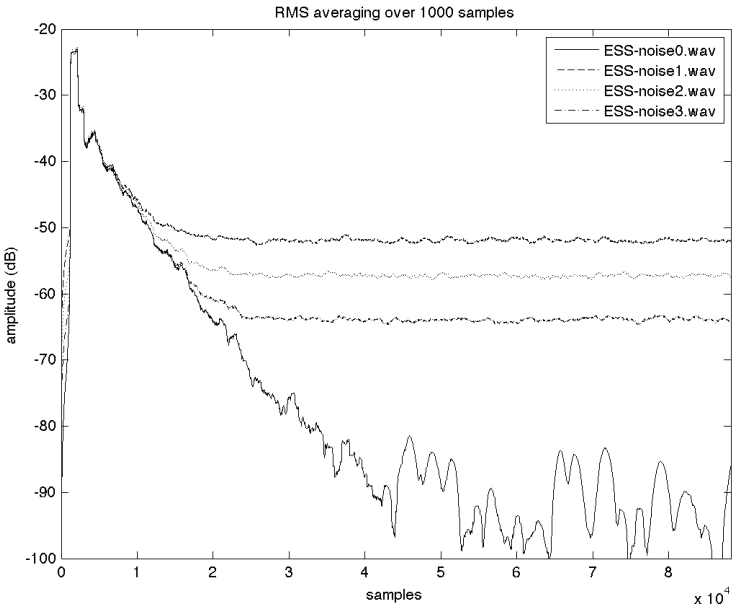

Figure 7.11: Exponential sine sweep plus added noise

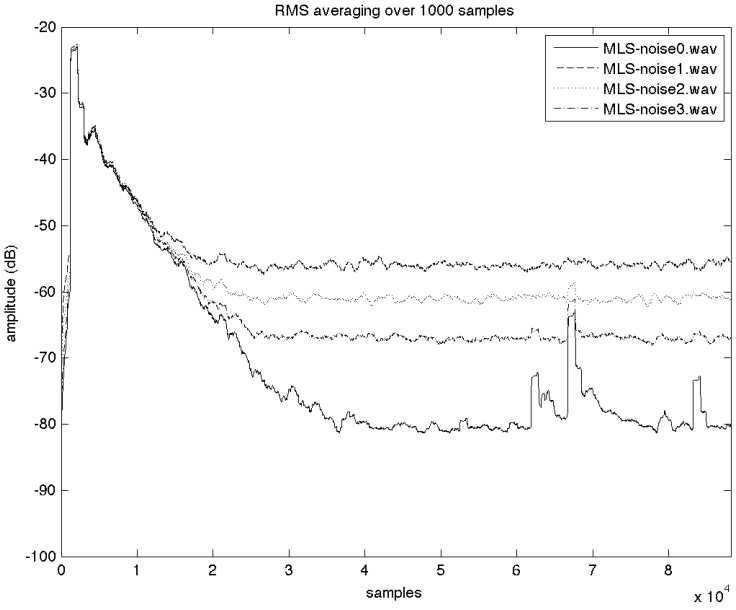

Figure 7.12: Maximum length sequence plus added noise

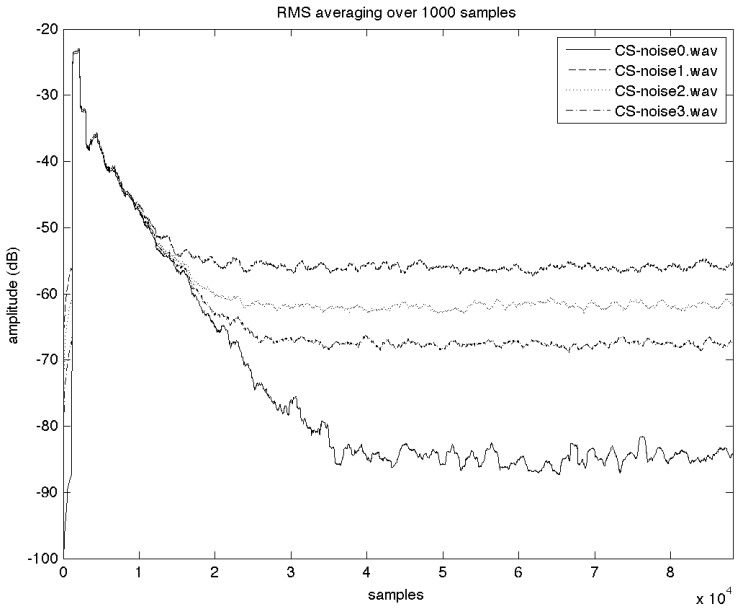

Figure 7.13: Complementary sequence plus added noise

7.7 Comparison Between Measurement Methods

All of the electronic methods perform significantly better than the balloon method under every noise condition. Consequently, the balloon method is not included in the following comparisons.

The performance of all of the electronic measurement methods are directly compared at each noise condition. All of the impulse responses are normalized to allow meaningful comparison of noise floors between methods. Note that the RMS averaging over 1000 samples will exhibit a substantially lower peak amplitude than the actual normalized peak; the true peak is at 0 dB.

Without any added noise, all three methods perform similarly until approximately 30,000 samples. In the latter half, the MLS produces the highest noise floor of -75.2 dB,

5.4 dB higher than CS and 14.2 dB higher than ESS. False impulses contribute to this relatively high noise floor.

With normalized white noise added, MLS and CS produce similar noise floors, measured at -55.8 dB and -56.2 dB respectively. ESS exhibits a significantly higher noise floor at -51.9 dB.

Lower amplitude white noise continues this trend, with MLS and CS performing similarly and ESS consistently producing a noise floor several decibels louder. When the white noise is attenuated by 6 dB, the false impulses produced by MLS begin to appear.

When the noise is musical, ESS provides the lowest noise floor, with MLS and CS producing noise floors 8.0 dB and 7.3 dB higher respectively.

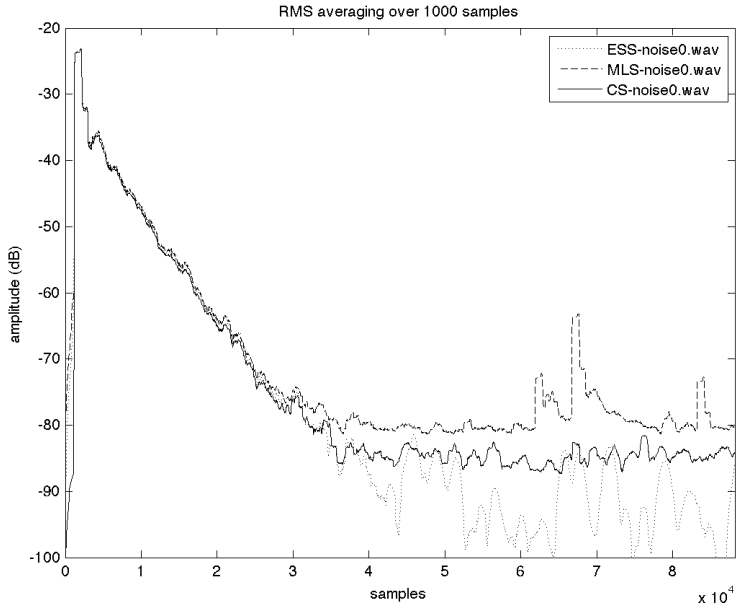

Figure 7.14: Comparison of measurement methods with no added noise

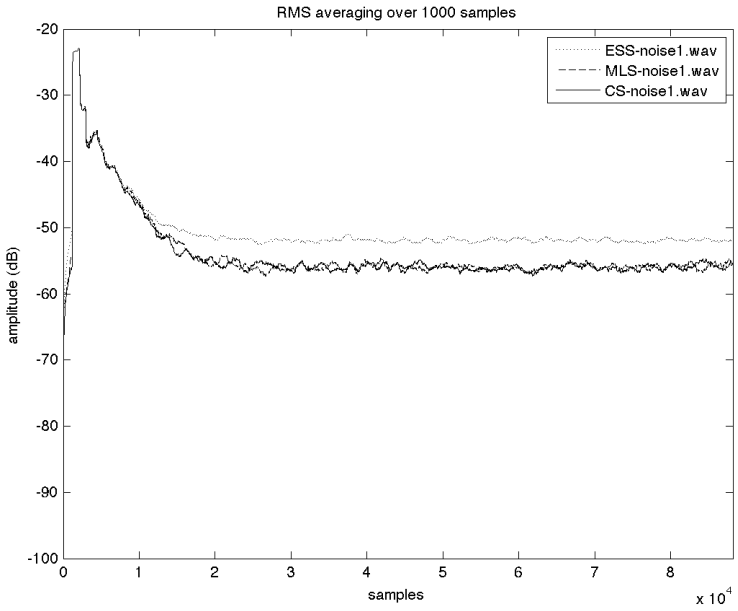

Figure 7.15: Comparison of measurement methods with white noise at 0 dB

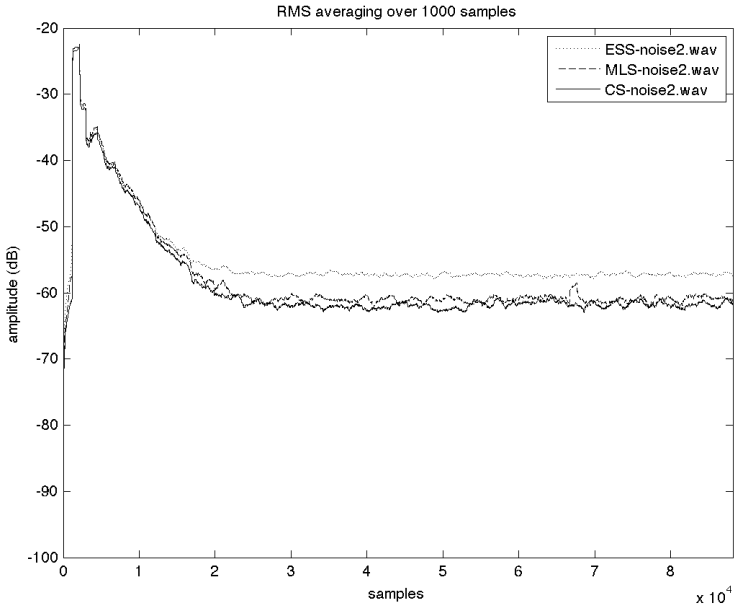

Figure 7.16: Comparison of measurement methods with white noise at -6 dB

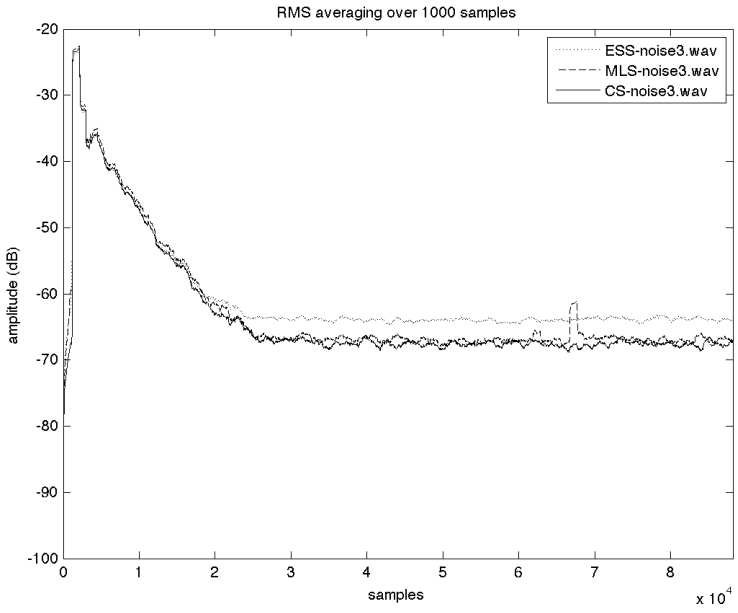

Figure 7.17: Comparison of measurement methods with white noise at -12 dB

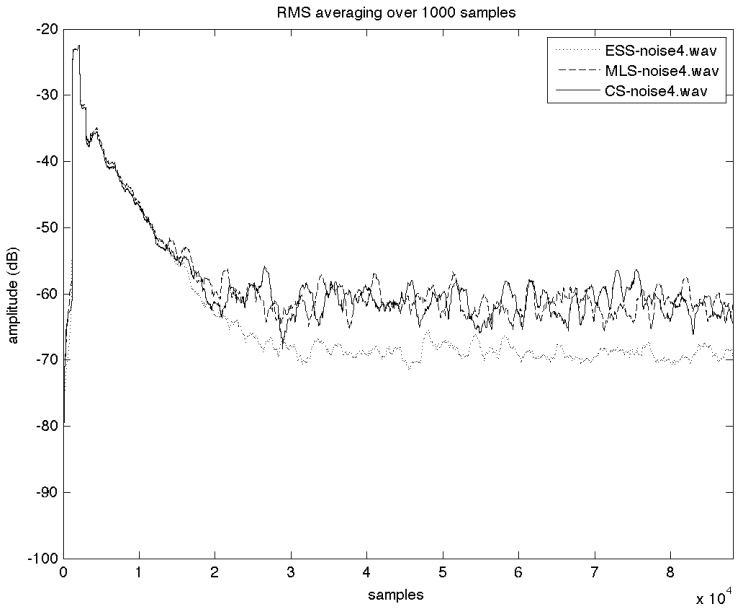

Figure 7.18: Comparison of measurement methods with music

Table 7.1: Noise floors of impulse response measurements (dB)

	balloon	ESS	MLS	CS
no added noise	-76.4	-89.4	-75.2	-80.6
0 dB white	-44.3	-51.9	-55.8	-56.2
-6 dB white	-49.9	-57.3	-60.8	-62.0
-12 dB white	-62.9	-63.9	-66.1	-67.8
music	-57.0	-69.1	-61.1	-61.8

7.8 Informal Listening Tests

While highly subjective, blind listening comparisons between the impulse responses can provide insight into the musical usefulness of each. Headphones are used to avoid room coloration and provide optimal stereo separation. Comparisons are made listening to the impulse responses directly and used as a convolution reverberator.

The balloon method consistently produces results with significantly more midrange than the other methods; by comparison, the remaining methods produce results with substantially flatter spectra.

The false impulses late in the MLS impulse responses are clearly audible, resulting in clear echoes when used in a convolution reverberator.

When combined with music, the ESS method produces sinusoidal artifacts which lower in pitch throughout the impulse response (visible in Figure 7.19). When used in reverberation, these artifacts produce audible pitch shifting.

7.9 Analysis of Results

Balloon pops do not favorably compare to any of the electronic methods under any of the noise conditions and perform exceptionally poorly when environmental noise is added. The nature of this method precludes precise time alignment necessary for averaging of multiple measurements. This technique should therefore not be considered for use in high-noise environments.

When the environmental noise is low, ESS produce the lowest noise floor, followed by CS and then MLS. When white noise is introduced, bit sequences (MLS and CS) produce the lowest noise floors, with ESS creating noise floors which are consistently substantially higher.

ESS produces the lowest noise floor when music is added. However, the strongly pitched artifacts this introduces can create modulation when used with a convolution reverb. For this reason, this method may be undesirable when creating impulse responses intended for reverberation.

8.0 EchoThief Impulse Response Library

Throughout the development of this method, I have put it to use capturing impulse responses in hundreds of spaces. I have compiled my favorites among these to create the EchoThief Impulse Response Library, in hopes that others will find them useful (http://www.echothief.com).

Impulse responses in this library are displayed as pins on a map to emphasize their origins as measurements of actual spaces. Seven additional maps present subsets of these based on broad categories: Underground, Brutalism, Nature, Stairwells, Venues, Underpasses, and Recreation. Clicking on a pin reveals a link to the impulse response as well as a panoramic photo of the space that has been measured and a description of the main component materials of that space.

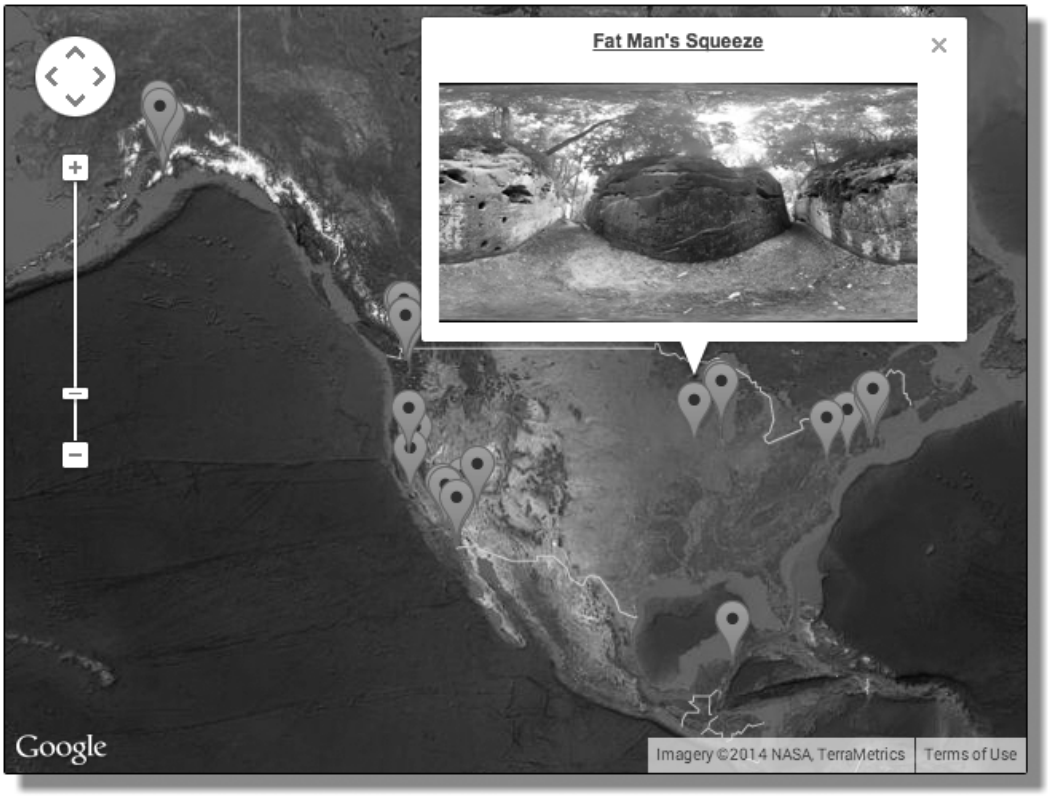

Figure 8.1: The EchoThief Impulse Response Library map

8.1 Battery Benson

My favorite place to capture impulse responses is Fort Worden, a decommissioned coastal artillery fortification on the Olympic Peninsula in Port Townsend, Washington. It has been converted into a state park, but the hulking remnants of concrete batteries and tunnels remain.

Battery Benson is a particularly labyrinthian warren, with irregular trapezoidal construction likely to disorient invading infantry. Thick concrete walls at uneven angles provide long reverberation and rich mixing with few standing waves, unintentionally similar to the Capitol Studios echo chambers discussed in Chapter 1. The broadband t_{60} of the space is 1.5 seconds with lower frequency components sustaining to nearly 2 seconds.

Figure 8.2: Panorama of Battery Benson

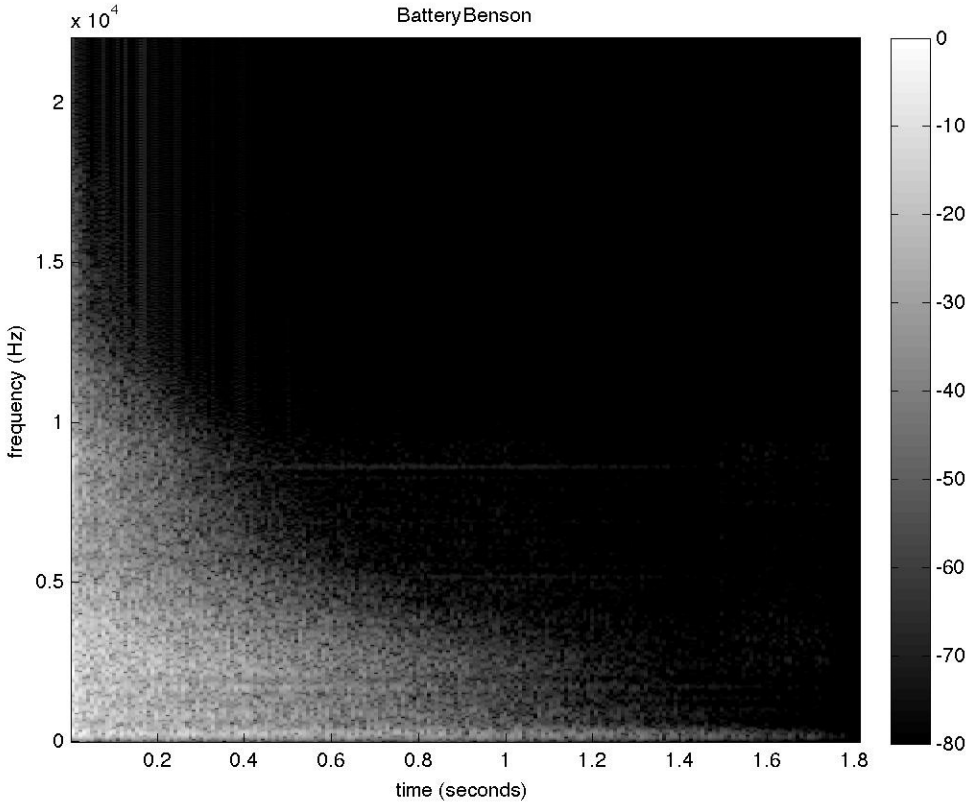

Figure 8.3: Spectrogram of Battery Benson impulse response

8.2 Wangenheim Rare Books Room

The Wangenheim Rare Books Room in the old San Diego Central Library provides an example of a much shorter reverberation time, with a t_{60} of merely 350 ms. Many absorbent and diffusive materials contribute to this brief duration: cork floors covered with heavy rugs, shelf upon shelf of books, thick silk wallpaper, plush furniture, and popcorn ceilings.

Figure 8.4: Panorama of Wangenheim Rare Books Room

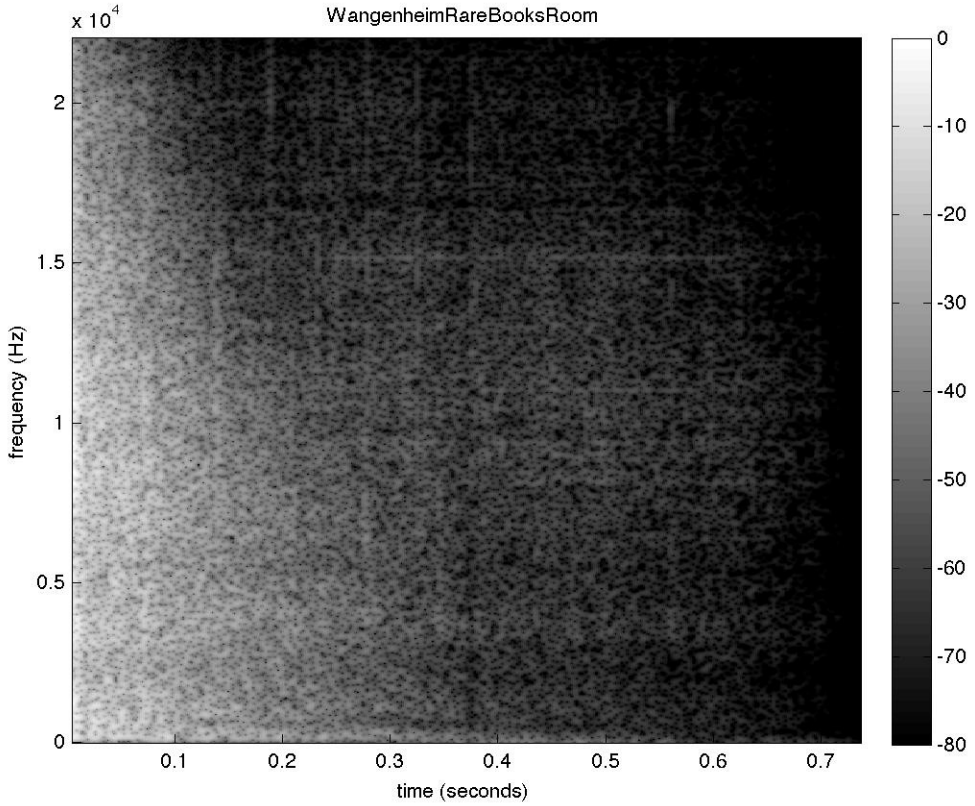

Figure 8.5: Spectrogram of Wangenheim Rare Books Room impulse response

8.3 Red Bridge

Only one covered bridge in Wisconsin remains standing: Red Bridge outside of Cedarburg, just north of Milwaukee. Built from rough sawn cedar with only a small amount of its original asphalt paving remaining, its lattice construction leaves little opportunity for standing waves. Openings at both ends and between the walls and roof provide many routes for sound to escape, resulting in a t_{60} of 500 ms.

Figure 8.6: Panorama of Red Bridge

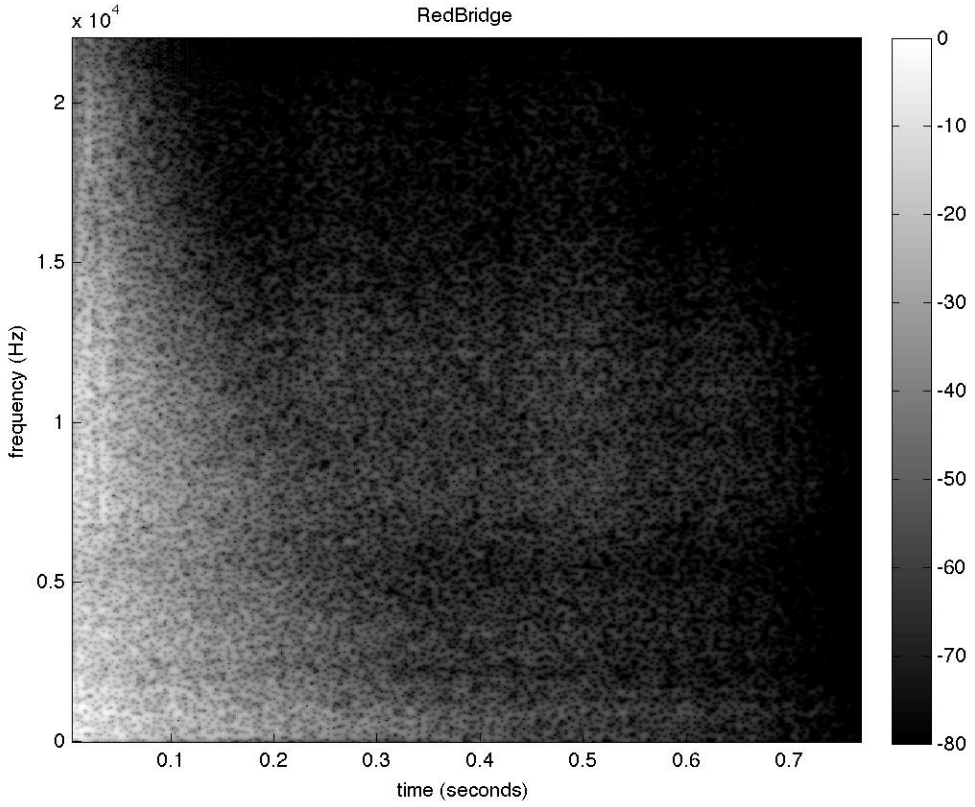

Figure 8.7: Spectrogram of Red Bridge impulse response

8.4 Tunnel to Heaven

Adjacent to Battery Benson is a long concrete foot tunnel through a hill, connecting it to the other batteries. Barely large enough to walk through, it has an arched ceiling and curves gently upward throughout its length. Above the entrance, the words “To Heaven” have been spray painted. Its unusual geometry results in the strangest of all of the impulse responses I have captured. The first 200 ms sounds similar to many of the other impulse responses. As the high frequencies dissipate, a noticeably pitched element is revealed which continues for the remainder of the

response. Unlike other impulse responses, where standing waves can produce pitched elements at fixed frequencies, here the pitch lowers throughout the duration.

Figure 8.8: Tunnel to Heaven

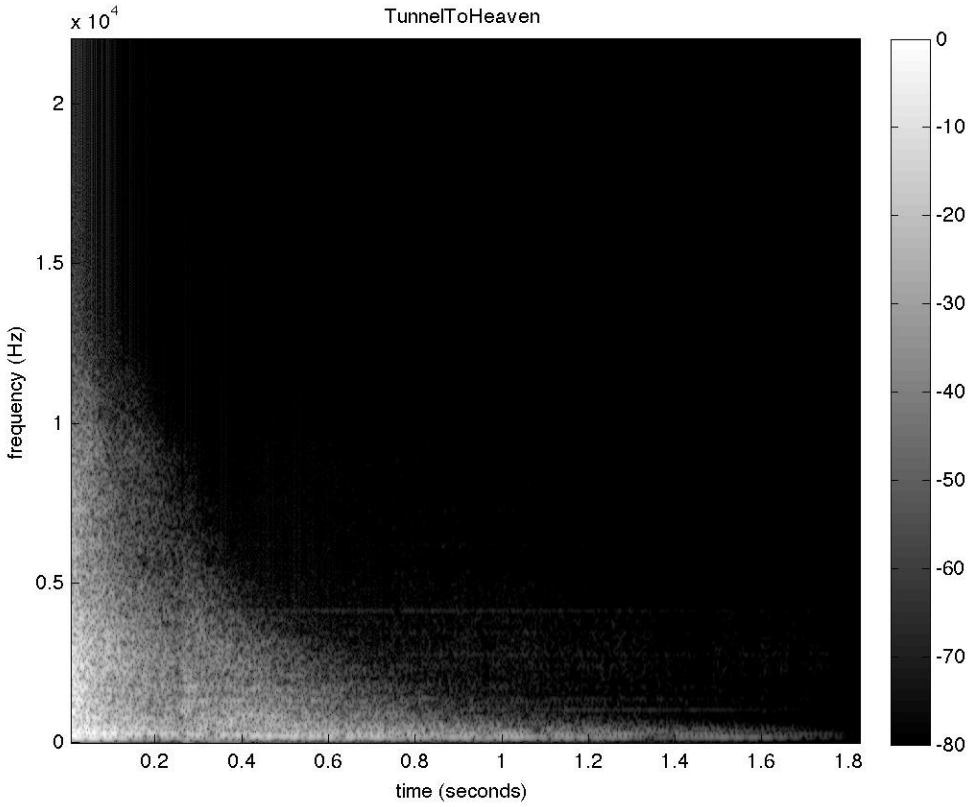

Figure 8.9: Spectrogram of Tunnel to Heaven impulse response

9.0 Summary

Complementary sequences can be extended beyond the typical t_{60} of acoustic spaces, allowing them to be repeated periodically to extract musically useful impulse responses from noisy spaces. This method performs well with added noise, producing impulse responses without the pitched artifacts created when using exponential swept sine methods.

The half-period similarity between *A* and *B* sequences, due to their iterative creation, provides an opportunity to create pairs which can be directly concatenated without the need for silence between measurements. This reduces the time between complementary pairs, minimizing the time in which the system under examination can change between measurements.

The circularity of complementary sequences can be exploited to extract a large set of complementary pairs from a brief composite sequences. These can be averaged to produce a lower noise floor.

All of these factors make complementary sequences preferable over existing techniques when capturing impulse responses in most real world situations where environmental noise is substantial.

Appendix I: Hardware

The hardware configuration used in the research and development of this technique evolved over numerous iterations. The two setups described below are the most recent versions of these, for indoor and outdoor use respectively. The former was used for the measurements described in Chapter 6.

Indoor configuration:

Recording/Playback:	2011 MacBook Pro
Processor:	2.7 GHz Intel Core i7
Memory:	8 GB 1333 MHz DDR3
Audio Interface:	MOTU UltraLite FireWire
Sample Rate:	44100 Hz
Word Length:	24 bits
Microphones:	Oktava MK012 condenser pair with foam windscreens
Amplifier	
Model:	Dayton Audio DTA-100a
Wattage:	50 watts/channel
Ohmage:	8 ohms
THD:	<0.01% at 30 watts
Speaker	
Model:	Tannoy Reveal 601p
Drivers:	6.5” LF/MF driver, 1" tweeter
Wattage:	50W RMS, 100W Program

Outdoor configuration:

Recording/Playback:	Zoom H4n
Sample Rate:	44100 Hz
Word Length:	16 bit
Microphones:	condenser pair with fur windscreen
Amplifier	
Model:	Dayton Audio DTA-1
Wattage:	10 watts/channel
Ohmage:	8 ohms
THD:	<0.1% THD at 10 watts
Speaker	
Model:	JBL Control 1
Drivers:	5.25” LF/MF driver, 0.75” tweeter
Wattage:	150W

Bibliography

Bartholomew, Wilmer T. *Acoustics of Music*. Prentice Hall, Inc.,1980.

Bayless, James W. "Innovations in Studio Design and Construction in the Capitol Tower Recording Studios." *Journal of the Audio Engineering Society* 5.2 (1957): 71-76.

Borish, Jeffrey, and James B. Angell. "An Efficient Algorithm for Measuring the Impulse Response Using Pseudorandom Noise." *Journal of the Audio Engineering Society* 31.7/8 (1983): 478-488.

Borwein, Peter. *Computational Excursions in Analysis and Number Theory*. Springer, 2002.

Bracewell, Ronald N. *The Fourier Transform and Its Applications*. McGraw-Hill, 1965.

Brook, Rollins and Ted Uzzle. "Rooms for Speech, Music, and Cinema." *Handbook for Sound Engineers*. Ed. Glen Ballou (1987): 155-197.

Burrus, CSS, and Thomas W Parks. *DFT/FFT and Convolution Algorithms: Theory and Implementation*. John Wiley & Sons, Inc., 1991.

Davis, Don. “Audio Measurements." *Handbook for Sound Engineers*. Ed. Glen Ballou (1987): 1175-1178.

Dodge, Charles, and Thomas A Jerse. *Computer Music: Synthesis, Composition and Performance*. Schirmer, 1985.

Doyle, Peter. *Echo and Reverb: Fabricating Space in Popular Music Recording, 1900-1960*. Wesleyan University Press, 2005.

Eargle, John. *Handbook of Recording Engineering*. Springer, 2005.

Everest, F. Alton. “Acoustical Design of Audio Rooms”. *Handbook for Sound Engineers*. Ed. Glen Ballou (1987): 155-197.

Farina, Angelo. "Simultaneous Measurement of Impulse Response and Distortion with a Swept- Sine Technique." *Audio Engineering Society Convention* 108 1 Feb. 2000.

Foster, Scott. "Impulse Response Measurement Using Golay Codes." *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'86*. Apr. 1986: 929-932.

Golay, Marcel. "Complementary Series." *Information Theory, IRE Transactions on* 7.2 (1961): 82-87.

Golay, Marcel. “Multislit Spectrometry,” *Journal of the Optical Society of America* 39 (1949): 437.

Golay, Marcel. "Note on ‘Complementary Series’". *Proceedings of the IRE* 50:1 (1962): 84.

Helmholtz, Hermann L. F. *On The Sensations of Tone.* English translation by Alexander J. Ellis, Dover Publications, Inc., 1954.

Heyser, Richard C. "Acoustical Measurements by Time Delay Spectrometry." *Journal of the Audio Engineering Society* 15.4 (1967): 370-382.

Knudsen, Vern O. and Cyril M Harris. *Acoustical Designing in Architecture.* John Wiley & Sons, Inc.,1950.

Moore, F. Richard. *Elements of Computer Music*. Prentice Hall, Inc., 1990.

Milner, Greg. *Perfecting Sound Forever: An Aural History of Recorded Music*. Faber and Faber, Inc., 2009.

Norcross, S, and John S Bradley. "Comparison of Room Impulse Response Measurement Methods." *Canadian Acoustics* 22.3 (1994): 47-48.

Oppenheim, Alan V, and Ronald W Schafer. *Digital Signal Processing*. Prentice Hall, Inc., 1975.

Parker, Julian, and Stefan Bilbao. "Spring Reverberation: A Physical Perspective." *Proceedings of the 12th International Conference on Digital Audio Effects (DAFx’09)* Sep. 2009: 416-421.

Puckette, Miller. *The Theory and Techniques of Electronic Music*. World Scientific, 2007.

Rossing, Thomas D., F. Richard Moore, and Paul A. Wheeler. *The Science of Sound, 3rd Edition*. Addison-Wesley, 2001.

Sabine, Wallace C. *Collected Papers on Acoustics*. Harvard University Press, 1922.

Schroeder, Manfred R. “Integrated Impulse Method Measuring Sound Decay Without Using Impulses." *The Journal of the Acoustical Society of America* 66 (1979): 497-500.

Schroeder, Manfred R. and Benjamin F Logan. “Colorless Artificial Reverberation." Journal of the Audio Engineering Society 9.3 (1961): 192-197.

Shaughnessy, Mary Alice. *Les Paul: An American Original*. W. Morrow, 1993.

Smith, Julius O. *Introduction to Digital Filters: With Audio Applications*. W3K Publishing, 2007.

Smith, Julius O. *Mathematics of the Discrete Fourier Transform (DFT): With Music and Audio Applications*. W3K Publishing, 2007.

Smith, Julius O. *Physical Audio Signal Processing: For Virtual Musical Instruments and Audio Effects*. W3K Publishing, 2010.

Smith, Julius O. *Spectral Audio Signal Processing*. W3K Publishing, 2007.

Smith, Steven W. *The Scientist and Engineer's Guide to Digital Signal Processing*. California Technical Pub., 1997.

Stan, Guy-Bart, Jean-Jacques Embrechts, and Dominique Archambeau. "Comparison of different impulse response measurement techniques." *Journal of the Audio Engineering Society* 50.4 (2002): 249-262.

Stautner, John, and Miller Puckette. "Designing Multi-channel Reverberators." *Computer Music Journal* 6.1 (1982): 52-65.

Strum, Robert D, and Donald E Kirk. *First Principles of Discrete Systems and Digital Signal Processing*. Addison-Wesley Longman Publishing Co., Inc., 1988.

Taylor, Charles A. *The Physics of Musical Sounds*. American Elsevier, 1965.

Tseng, Chin-Chong, and C Liu. "Complementary Sets of Sequences." *Information Theory, IEEE Transactions on* 18.5 (1972): 644-652.

Vanderkooy, John. "Aspects of MLS measuring systems." *Journal of the Audio Engineering Society* 42.4 (1994): 219-231.

Vigran, Tor Erik. *Building Acoustics*. Taylor & Francis, 2008.