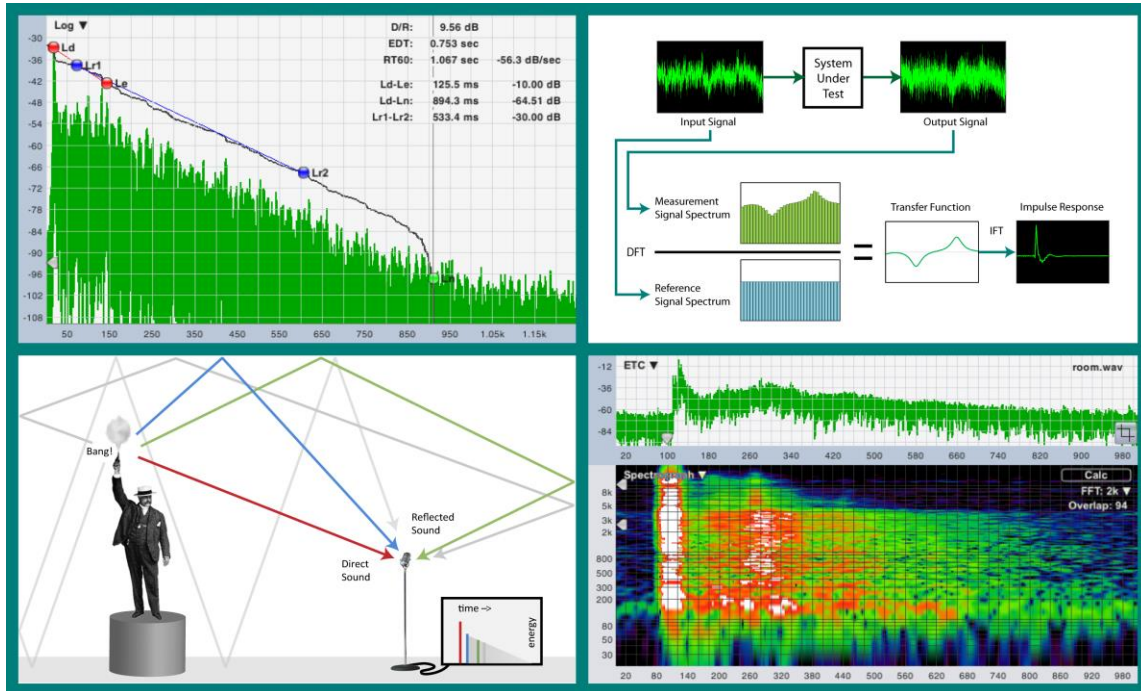


Smart 7 Impulse Response Measurement and Analysis Guide



Smart


rational acoustics

Introduction	1
Scope and purpose of document	1
1: What is an impulse response?	2
Anatomy of an acoustical impulse response	3
Propagation Delay	3
Arrival of Direct Sound	4
Discrete Reflections	4
Early Decay, Reverberant Build-up, and Reverberant Decay	4
Noise Floor	4
Uses for impulse response measurement data	5
Delay Time Measurement	5
Reflection Analysis	5
Reverberation Time (T60, RT60...)	5
Early Decay Time (EDT)	5
Early-to-late energy ratios	6
Speech Intelligibility Modeling	6
2. A Quick Tour of the Smart 7 Impulse Response Mode User Interface	7
1. Navigation pane	7
2. Main Graph Area	8
3. Cursor Readout	9
4. Sound Level Meter	9
5. Data Display Controls	9
6. Signal Generator	10
7. FFT size and Averaging Controls	10
8. Input Source Selector	10
9. Live Measurement Controls	10
10. Bandpass Filters	11
Additional Options for IR Measurement	11
General Options Pertaining to IR Measurement	11
Impulse Response Options	12
3: Analyzing Impulse Response Data	14

Time Domain Analysis	14
Logarithmic Time Domain Display	14
Linear Time Domain Display	15
Energy Time Curve (ETC)	17
Bandpass Filtering	19
Discrete Reflections	19
Reverberation Time	20
Reverse Time Integration	20
Evaluation Ranges (EDT, T20, T30)	21
Reporting Results for Reverberation Time	23
Early-to-Late Energy Ratios	24
Clarity Ratios (C35, C50, C80...)	24
The Histogram Display	25
The All Bands Table	26
Frequency Domain Analysis	26
The Spectrograph	27
Spectrograph Time and Frequency Resolution	29
Spectrograph Dynamic Range	31
Spectrograph Analysis of an Acoustical Impulse Response	31
Speech Intelligibility Metrics	32
STI and STIPA	32
ALCons	35
4: Measuring an Acoustical Impulse Response	36
What are we measuring, and why?	36
Direct vs Indirect IR measurement	36
Direct IR Measurement Using an Impulsive Stimulus	36
Indirect (Dual Channel) IR Measurement	37
Dual Channel IR measurement Using Period-Matched Signals	38
Dual Channel IR measurement Using Random Stimulus Signals	41
Selecting Excitation Sources and Positions	42
Minimum Distance from Sound Sources	43

Directional Loudspeakers and Reverberation Time.....	44
Selection of Measurement Positions	45
Selecting Measurement Parameters	46
Input source	47
Excitation Level	47
Input Levels	47
Measurement Duration	48
Averaging and Overlap.....	48
Pushing the Button and Making the Measurement	49
Saving Your Work.....	49
Recap: Recommended Default Settings for Dual-channel IR measurements	50
Signal Type	50
FFT size and Averaging.....	50
Excitation Level	50
Input Levels	50
Delay Time	50
Appendix A: Applicable Standards and Further Reading	52
Further Reading	52
Classic Reference Works for Acoustics	52
Appendix B: Room Volume, Absorption and Reverberation time (RT60)	53
Material absorption coefficients.....	53
Air absorption	53
Schroeder cut-off frequency	54
Practical measurement considerations	54
Appendix C: Sound source characteristics	56
Directivity factor (Q)	56
Sound pressure level at the listening position.....	56
Directivity index (DI)	57
Q and DI figures in practice.....	57
Conventional loudspeaker arrays	58
“Line” arrays.....	58

A note about echoes	59
Appendix D: Boundary effects	61
Acoustically small sources.....	61
Microphones near boundaries.....	61
Appendix E: Typical Measurement Rig Set-Ups.....	63
Stereo 2x2 Audio I-O	63
Multi-channel I-O	63
Traditional USB/FireWire Stand-Alone Interfaces	63
Network I-O.....	64

Introduction

Scope and purpose of document

Smaart 7 can operate in two distinct operating modes, real-time and IR (short for impulse response) mode. Both real-time and IR modes feature extensive capabilities for graphical analysis of audio data in the time and frequency domains, however real-time mode is primarily geared for analysis of streaming audio signals. IR mode is purpose built for acoustical analysis. In IR mode Smaart can work with time records several seconds in length – loaded from a file or actively measured – and can calculate metrics used in room acoustic and estimate speech intelligibility from IR data.

This document is intended as a practical guide to the basics of measuring and analyzing an acoustical impulse response (IR) using Rational Acoustics' Smaart. We intend to focus on the practical as much as we can without spending a lot of time on theory or mathematics. Our main objective is to help Smaart users become better acquainted with the process of selecting measurement parameters and microphone positions to get a good IR measurement, and hopefully extract some useful information from the measurement results.

1: What is an impulse response?

In the most basic terms, an impulse response (IR) can be defined as the time domain (time vs amplitude) response of a system under test (SUT) to an impulsive stimulus. The word “system” in this case could mean something as small as a microphone or a single transducer, something as simple as a single filter on an equalizer. Or, it might mean something as big as a concert hall or sports arena, as complicated an entire sound system or a combination of the two. Smart users of course are most often concerned with sound systems and their acoustical environments.

In the context of acoustical analysis, you might also think of an impulse response as the acoustical “signature” of a system. The IR contains a wealth of information about an acoustical system including arrival times and frequency content of direct sound and discrete reflections, reverberant decay characteristics, signal-to-noise ratio and clues to its ability to reproduce intelligible human speech, even its overall frequency response. The impulse response of a system and its frequency-domain transfer function turn out to be each other’s forward and inverse Fourier transforms.

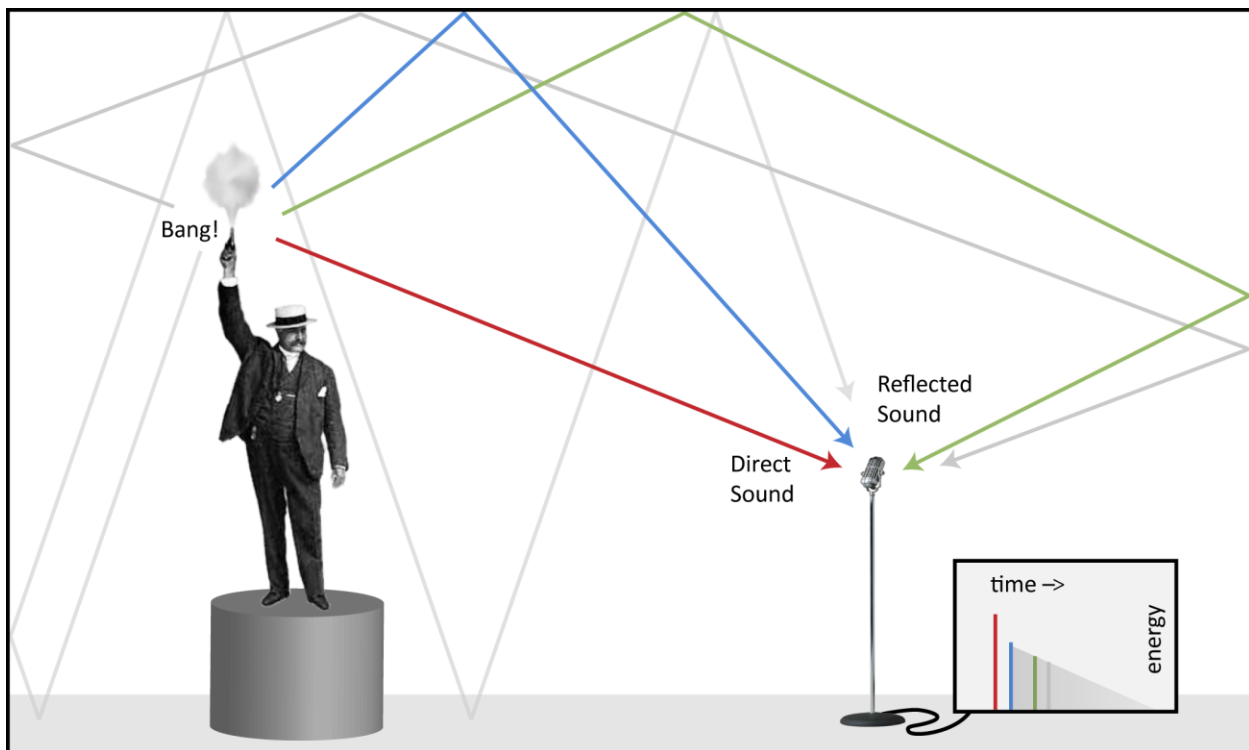


Figure 1: An acoustical impulse response consists of sound from an excitation source arriving at a measurement position by multiple pathways, both direct and reflected. Here we see the path of direct sound from the source to the microphone in red, followed by a first order reflection in blue, a second order reflection in green and higher order reflections in gray. Later arrivals tend to pile on top of each other forming a decay slope.

An acoustical impulse response is created by sound radiating outward from an excitation source and bouncing around the room. Sound traveling by the most direct path (a straight line from the source to a measurement position) arrives first and is expected to be the loudest. Reflected sound arrives later by a multitude of paths, losing energy to air and surface absorption along the way so that later arrivals tend

to come in at lower and lower levels. In theory this process goes on forever. In practice, the part we care about happens within a few seconds – perhaps less than a second in smaller rooms and/or spaces that have been acoustically treated to reduce their reverberation times.

The arrival of direct sound and probably some of the earliest arriving reflections will be clearly distinguishable on a time-domain graph of the impulse response. As reflected copies of the original sound keep arriving later and later, at lower and lower amplitude levels, they start to run together and form an exponential decay slope that typically looks like something close to a straight line when displayed on a graph with a logarithmic amplitude scale.

Anatomy of an acoustical impulse response

Although no two non-identical rooms ever have identical impulse responses, there are a few component features that we can identify in some combination in almost any acoustical impulse response. These include the arrival of direct sound, early reflections, reverberant build-up and decay, and the noise floor. Figure 2 shows an acoustical impulse with its component parts labeled. Descriptions for each follow.

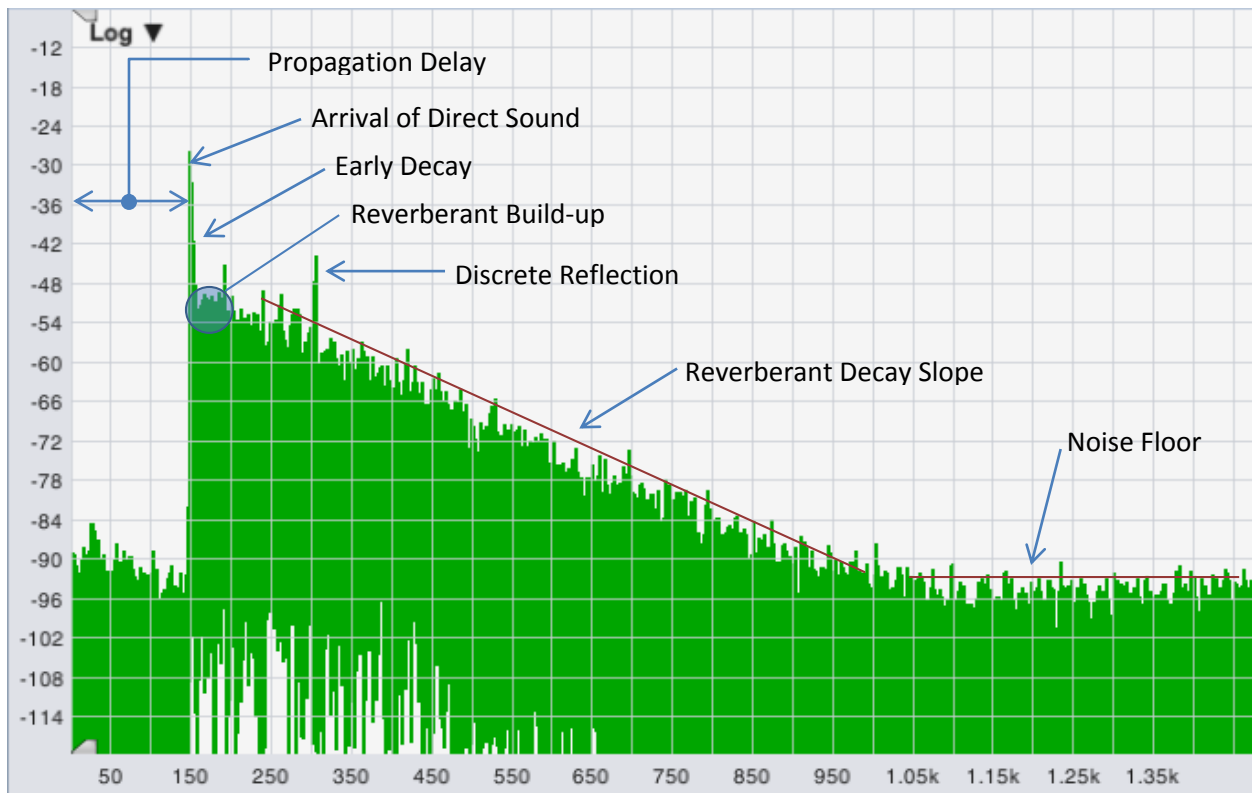


Figure 2: An acoustical impulse response with its common component parts labeled. This is a semi-log time domain chart with time in milliseconds on the x axis and magnitude in decibels on the y axis.

Propagation Delay

The time that it takes for direct sound from the sound source to reach the measurement position is the propagation delay time. This may include throughput delay for any DSP processors in the signal chain in addition to the time that it takes for sound to travel through the air.

Arrival of Direct Sound

Since the shortest distance between two points is always the straightest line, the first thing we expect to see when looking at an impulse response (IR) is the arrival of direct sound from whatever sound source we're using to stimulate the system under test. Depending on what we're trying to learn, the source could be an installed sound system, an omnidirectional loudspeaker brought in specifically for measurement purposes, a balloon pop or a shot from a blank pistol, or in a pinch, maybe hand claps or someone slamming a case lid shut.

In most cases we'd also expect the first arrival to be the loudest, and correspond to the highest peak we can see in the IR, and in most cases we'd be right. There can be occasional circumstances where that might not turn out to be strictly true, but in the vast majority of the cases it should.

Discrete Reflections

After the arrival direct sound, the next most prominent features we tend to see are sound arriving the next most direct paths; the lowest order reflections. Sound that bounces off one surface to get from the excitation source to a measurement position is called a first-order reflection, two bounces gives you a second order reflection and so on. Reflected sound can be useful or detrimental, depending on factors such as its relative magnitude and timing in relation to the direct sound and the extent to which it is clearly distinguishable from the diffuse reverberant sound.

Early Decay, Reverberant Build-up, and Reverberant Decay

Following the arrival of direct sound and the lowest order reflections, sound in a reverberant space will continue bouncing around a room for a while, creating higher and higher order reflections. At any given listening position, some of this reflected energy will combine constructively over a relatively short period of time, resulting in a build-up of reverberant sound, before air loss and absorption by the materials that make up reflecting surfaces begins to take their toll. At that point, the reverberant decay phase begins.

In practice, you may or may not be able to see the reverberant build-up in an impulse response as distinct from the direct sound and early reflections. Sometimes it can be quite clearly visible, other times not so much. By convention, the first 10 dB of decay after the arrival of direct sound in the reverse-time integrated IR (we will get to that in the next section) is considered to be early decay. Reverberant decay is conventionally measured over a range from 5 dB below the level of direct sound down to a point 30 dB below that on the integrated IR, or 20 dB in a pinch.

Noise Floor

In theory, the reverberant decay phase of the IR continues forever, as an ideally exponential curve that never quite reaches zero. In practice it reaches a point relatively quickly where we can no longer distinguish it from the noise floor of the measurement. Noise in an IR measurement can come from several sources, including ambient acoustical noise, and electrical noise in the SUT and the measurement system, quantization noise from digitizing the signal(s) for analysis, and process noise from DSP processes used for analysis.

Uses for impulse response measurement data

Delay Time Measurement

The most common use for impulse response measurements in Smaart is in finding delay times for signal alignment in transfer function measurements and for aligning loudspeaker systems. Each time you click the delay locator in Smaart an IR measurement runs in the background. In this case all we really care about is the initial arrival of direct sound, which is typically so prominent that you can pick it out with high confidence even when signal-to-noise ratio of the IR is poor, so we don't even bother displaying the results. Smaart simply scans for the highest peak and assumes that to be the first arrival, and most of the time that works very well.

Occasions where automatic delay measurements might not work well include measurements of low-frequency devices or any case where you're trying to measure a directional full-range system well off axis, in a location where a prominent reflection can dominate the high frequencies. In the latter case, it's possible for reflected HF energy to form a higher peak later than the arrival of direct sound, requiring you to visually inspect the IR data to find the first arrival.

Reflection Analysis

Another common use for IR measurements is in evaluating the impact of problematic discrete reflections. Reflected sounds can be beneficial or detrimental to a listener's perception of sound quality and/or speech intelligibility, depending on a number of factors. These factors include the type of program material being presented (generally, speech or music), the arrival time and overall level of the reflected sound relative to the level of direct sound, and the frequency content and the direction from which they arrive. As a general rule, the later they arrive and the louder they are (relative to direct sound) the more problematic they tend to be.

Reverberation Time (T60, RT60...)

Reverberation time is kind of the grandfather of quantitative acoustical parameters. First proposed by Walter Sabine more than 100 years ago, T60 or RT60 reverberation time is the time that it takes for reverberant in a room sound to decay by 60 decibels from an excited state (after the excitation signals stops). It is one of the most widely used (and in some cases perhaps misused) quantities in room acoustics. Although it is quite possible for two rooms with identical reverberation times to sound very different, when evaluated band-by-band it can still give you some idea as to the overall character of the reverberant field in a given room. In concert halls it can give you an idea of perceived warmth and spaciousness for music. In auditoriums, it is often used as a rough predictor of speech intelligibility.

Early Decay Time (EDT)

Early decay time ends up being the decay time for direct sound and earliest, lowest-order reflections. Since the earliest reflections tend to be the most beneficial in terms of separating sounds we want to hear from reverberation and background noise, EDT can give you some clues about overall clarity and intelligibility in a room and/or system. EDT, like RT60, is conventionally normalized to the time it would take for the system to decay 60 dB at the measured rate of decay.

Early-to-late energy ratios

Early to late energy ratios are a direct measure of the sound energy arriving within some specified interval following the arrival of direct sound, vs the energy in the remaining part of the IR. These provide a more direct method of evaluating the relationship between beneficial direct sound and early reflections that a listener hears versus the amount of (potentially detrimental) reverberation and noise, than inferences made from the early and reverberant decay rates.

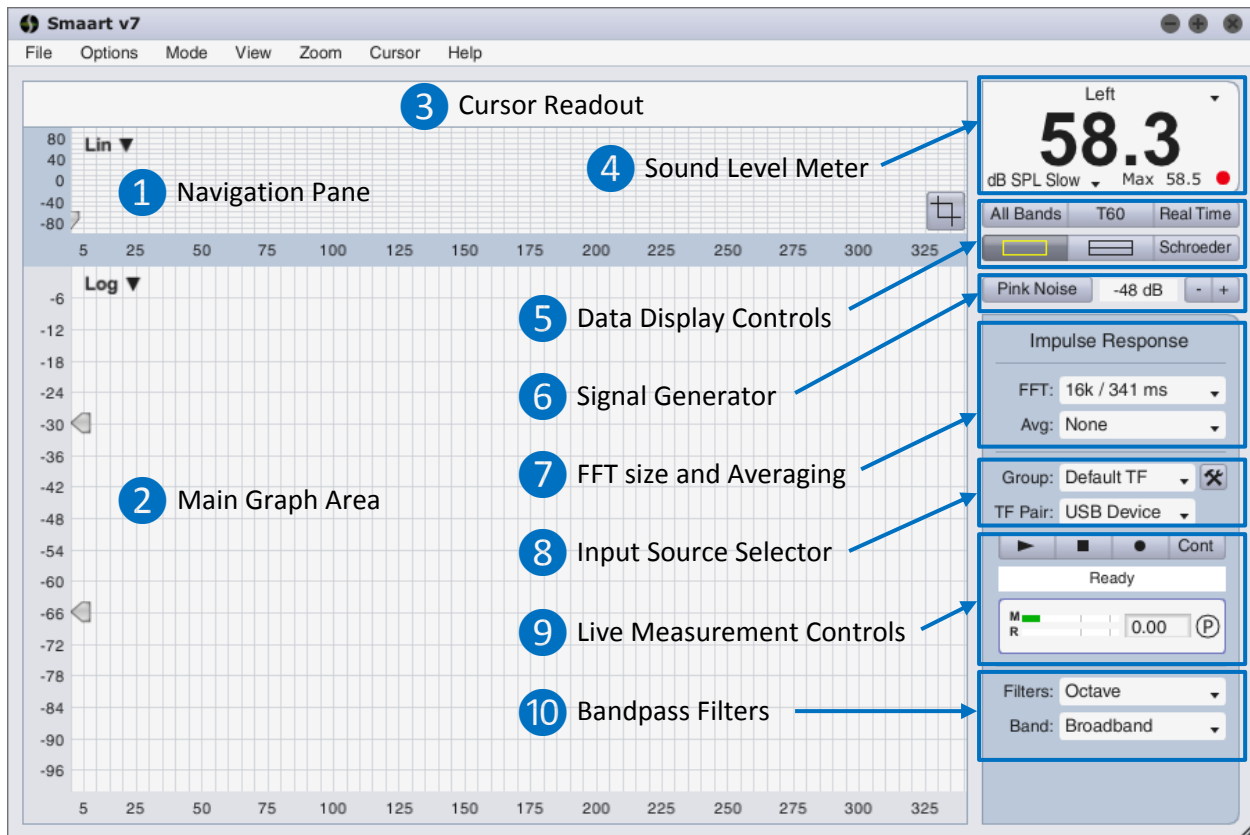
Speech Intelligibility Modeling

Early to late energy ratios such as C35 and C50 have long been used as objectively measurable predictors of subjective speech intelligibility. In the 1970s Victor Peutz came up with Articulation Loss of Consonants (ALCons), a predictive metric for intelligibility based in the volume of a room and its reverberation time, the directivity of loudspeakers and distance from source to the listener. Later on, Peutz revised the equation to use a direct-to-reverberant energy ratio in place of volume, distance and loudspeaker Q, making ALCons a directly measurable quantity. More recently, the speech transmission indexes (STI and STIPA) have emerged as metrics that are generally more robust. All of these can be calculated from the impulse response of a system.

2. A Quick Tour of the Smart 7 Impulse Response Mode User Interface

If you already know your way around IR mode in Smart 7 you can probably skip this section, but it might not hurt to at least skim over it. If you are new to IR mode then introductions are in order.

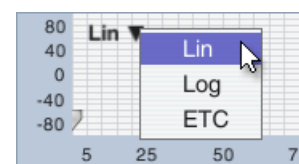
To get to IR mode in Smart, select *IR Mode* from the Mode menu, press the “I” key on your keyboard or click the *Impulse* button that appears in the upper right corner of the main window in real-time mode (just below the signal level /sound level meter). You will find yourself confronted with a screen like the one below. (The colors may be darker but the layout is the same.) On the right side of the window is a vertical strip of controls. The rest of the window is devoted to the graph areas and cursor readout.



1. Navigation pane

The small time-domain display in the upper part of the graph area is used for navigation and is always visible. Right-clicking and dragging across this (Ctrl + click and drag on Mac) panel selects a specific time range for display on the larger time-domain charts. The full IR time record remains visible in the navigation pane when you are zoomed in (unless you use the crop function). Clicking anywhere in the left margin of the plot clears the zoom range and returns any time-domain graphs in the main display pane(s) to the full IR time record.

The selector control in the upper right corner of the navigation pane selects the graph type to be displayed in this area. The navigation pane is limited to

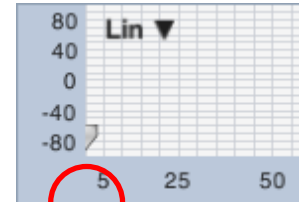


time-domain graph types only (Lin, Log or ETC).

The Crop button in the lower right corner of the navigation pane can crop a file for display purposes to show only the selected time range – a very useful feature when working with IR measurements with long noise tails. Cropping is non-destructive and can be undone – clicking the Crop button again on a cropped measurement restores the full extent of the original time record – however if you save the IR to a file while cropped, the cropped version is written to file.



The little arrow shaped widget positioned in the lower left corner of the navigation pane (circled in red in the screen clip show the right) is the *Time 0 marker*. When you record a dual-channel impulse response measurement in Smart, this marker is automatically set to match the reference signal delay time – if you are familiar with real-time transfer function measurement it’s analogous to the center point of the Live IR. For single-channel measurements or file-based data it is set to the beginning of the time record. Dragging it to the left or right moves the time-zero point for all time-domain graphs (Lin, Log or ETC) displayed in the main graph area.



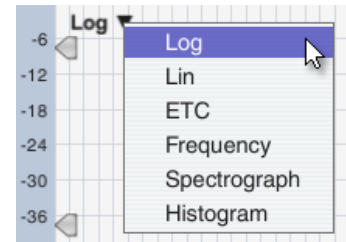
2. Main Graph Area

The larger (lower) portion of the graph area can be divided into one or two panes by clicking the buttons labeled with rectangles in area #5. Each pane can host any one of the six available main graph types:

- Lin, Log, ETC (time domain views)
- Spectrogram (frequency and level vs time),
- Frequency (spectrum)
- Histogram (bar chart of quantitative acoustics values by octave or 1/3-octave band.)

We will discuss each of these display types in the next section.

The graph type for each pane is selected by means of the drop-list control in the upper left corner of the pane. Some main graph types also have additional selector controls in their upper right corner that control display options specific to that graph. The two arrowhead shaped widgets to the left of the list of graph types appear on all time-domain and spectrograph plots. They control the dynamic range for the spectrograph.



As you probably noticed, the main graph area and the navigation pane are blank when you first enter IR mode. You won’t see any data in the graph areas until you record a measurement or load an impulse response from file (File menu > Load Impulse Response). Smart can open and analyze .wav and .aiff files containing any type of audio data, but IR mode is purpose-built for analyzing impulse responses. There’s no multi-channel file support or optimization for working with files more than a few seconds in length – and of course a lot the IR analysis capabilities are irrelevant for other type of audio data.

3. Cursor Readout

When measurement data is present, the cursor readout displays numeric coordinates for the cursor location(s) as you move your mouse over the graphs areas. Numeric coordinates are provided here for the cursor location in units of time, amplitude/magnitude and frequency, as applicable to graph type.



For time domain graphs (Lin, Log or ETC) in the main graph area(s), there are three sets of coordinates as show above. From left to right, they are the location of the locked cursor which typically marks the highest peak in the impulse response, the movable (mouse) cursor coordinates, and in brackets on the right, the difference between the locked and movable cursors. Note that time coordinates can optionally be displayed as both time (in milliseconds) and equivalent distance traveled, based on the currently specified speed of sound.

4. Sound Level Meter

The large numeric readout in the upper right corner of the Smart v.7 real-time mode window can be configured to function as a standard Sound Pressure Level (SPL) meter, an Equivalent Sound Level (LEQ) meter, or a peak signal level meter calibrated to digital full scale. The labels above and below the large numeric value in the center of the display function as menus that enable you to select the input channel driving the display and the type of measurement you wish to perform. When operating as an SPL or peak meter, the Max value can be reset by clicking the red circle (●) in the lower right corner of the display. In LEQ mode it changes from red to green (●) once the buffer for the rolling LEQ average is fully populated and clicking it will clear the buffer and restart the LEQ. Using the [K] hotkey displays a clock with the system time in the SPL meter window. Clicking anywhere in center portion of the meter opens the Sound Level Options dialog window where you will find additional options for configuring and calibrating the meter.



5. Data Display Controls

The first group of controls in the control strip on the right side of the IR mode window is devoted to data display functions. Starting from the top left of the screen clip on the right:

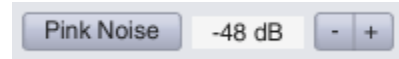


- The *All Bands* button opens the All Bands table, where all of the quantitative acoustical metrics that Smart can calculate are found.
- Clicking the *T60* displays the level marker widgets used for calculating reverberation time and early decay time on Log IR or ETC plot(s).
- The *Real Time* button exits IR mode and takes you back to real-time frequency domain measurement mode. (In real-time mode this changes to the *Impulse* button that will bring you back to IR mode.)
- The two buttons labeled with rectangles divide the main plot area (#2) into one or two graph panes – 1 rectangle, 1 pane; 2 rectangles, 2 panes.

- The *Schroeder* button displays a reverse time integration curve on Log IR or ETC plots.

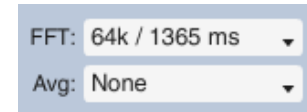
6. Signal Generator

The Signal generator controls work the same way in IR mode as in real-time mode. The leftmost button displays the currently selected signal type and turns the generator on and off. The +/- buttons on the right increase/decrease the output level. The center area displays the current output level and doubles as a button to open the signal generator control panel (also accessible by selecting Signal Generator from the Options menu) if you click it with your mouse. We will be talking more about signal generator options later on, as there are a few that are of particular interest in the context of IR measurements.



7. FFT size and Averaging Controls

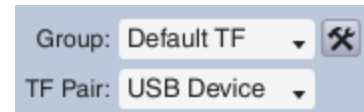
The FFT size and averaging controls together determine the measurement duration for dual-channel IR measurements. Notice that for each FFT size, the time constant is given along with the FFT size in samples. The FFT time constant, also called the time window, is the amount time it takes to record the required number of samples at the currently selected sampling rate.



Averages (Avg) sets the number of successive IR measurements to average together, to improve the signal-to-noise ratio of dual-channel measurements using random test signals. For deterministic IR measurements made using period-matched signals, the number of Averages is normally set “None”.

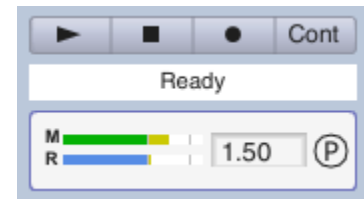
8. Input Source Selector

Dual-channel IR measurements in Smart are essentially transfer function measurements with the addition of an inverse Fourier transform at the end. Any of the reference and measurement signal pairs that you have set up for real-time transfer function measurements may be used for IR measurements. For single channel IR measurements (see #9 below) only the measurement channel is recorded. If you have more than one Transfer Function measurement group, the *Group* control is used to select the group containing the measurement that you want to use.



9. Live Measurement Controls

The live measurement controls in IR mode are analogous to a transfer function “measurement engine” in real-time mode, with a couple of extra twists. Starting from the top left, the buttons marked with a triangle (▶) and a square (■) start and stop a measurement. The button labeled with a circle (●) works like the record button on a tape deck or digital recorder but in this case it’s also a measurement mode control.



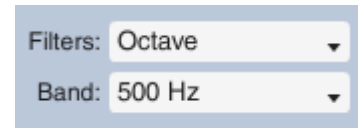
Clicking the start (▶) button *without* the record button punched in kicks-off a dual-channel IR measurement. *With* the record button (●) activated, Smart becomes a single-channel digital recorder and records just the measurement signal channel of the selected signal pair. The idea is that you would start recording and pop a balloon or fire your starter pistol (or whatever), then click the stop button to end recording and display your single-channel impulse response measurement.

The Continuous (Cont) button causes the dual-channel measurement routine to run continuously, starting over again automatically each time it finishes a measurement until you tell it to stop (by clicking the stop button). The results of the last measurement are displayed while it's recording and processing the next measurement. Click the stop button to end the measurement.

Underneath the recording controls there's a status bar to let you know what Smaart is doing when a measurement is in progress, since IR measurements can take a while to complete. Below that are input level meters and reference signal delay controls for the selected signal pair. The "@" button sets the delay time for the measurement to the highest peak in the impulse response when clicked. If you are measuring using a random stimulus signal and the delay is not already set you would run the measurement once to get the delay time, then click the @ button, then run the measurement again. You generally don't have to worry much about the delay time with deterministic IR measurements. They work fine with the delay set to zero. To zero out the delay time, type a zero in the delay time field and hit the enter key.

10. Bandpass Filters

The broadband impulse response is useful for finding delay times but for most acoustical analysis purposes, the IR needs to be filtered into octave bands or occasionally, 1/3-octave bands. Smaart includes complete sets of octave and 1/3-octave bandpass filters for impulse response analysis.



Bandpass filtering is non-destructive and is done on the fly whenever you need it. All you have to do is select the filter set (that you want to use (Octave or 1/3-Octave) using the *Filters* selector in the lower right corner of the IR mode window and then choose the center frequency for the band that you want to analyze from the *Band* list.

Additional Options for IR Measurement

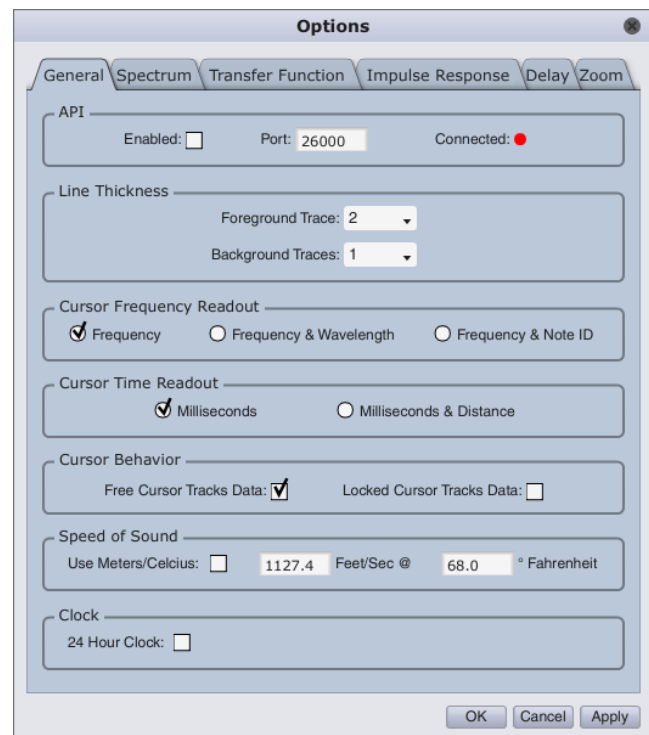
Some additional options pertaining to IR measurement that don't appear on the main screen can be found in the *General* and *Impulse Response* options pages, accessible from the *Options* menu.

General Options Pertaining to IR Measurement

Cursor Time Readout

The cursor time readout setting applies to the both IR Mode and the Live IR graph on Transfer Function displays.

- *Milliseconds* – displays time coordinates and relative time differences in



milliseconds only.

- *Milliseconds & Distance* – displays time coordinates as milliseconds and equivalent distance, based on the Speed of Sound settings (see below).

Speed of Sound

The settings in this section determine the speed of sound that Smart uses for calculating equivalent distances for time coordinates and also whether distances are displayed in feet or meters. It can also serve as a handy speed of sound calculator any time you need to know the speed of sound for a given air temperature.

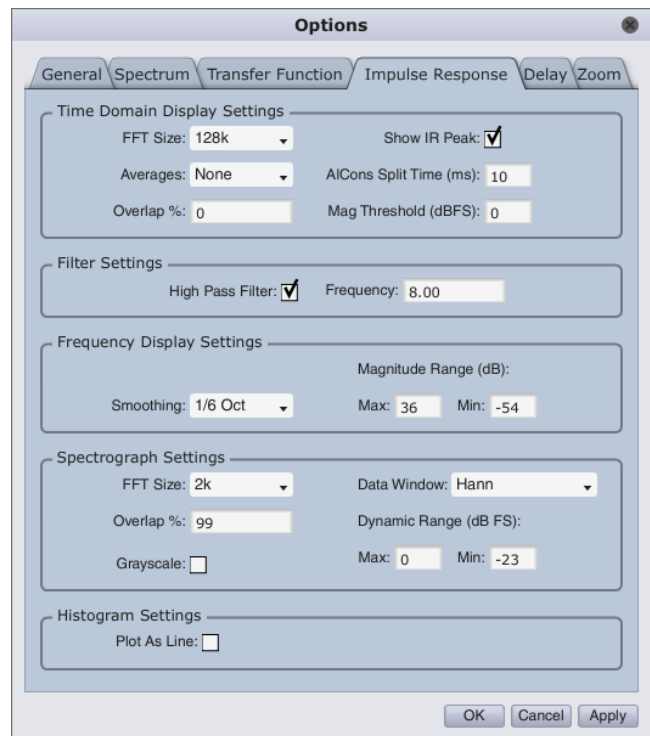
- *Use Meters/Celsius* – when this option is selected, Smart displays distances in meters and the temperature used for calculating speed of sound in degrees Celsius. Otherwise Smart displays distances in feet and uses degrees Fahrenheit for temperature.
- *Speed of sound ([unit]/sec) and temperature* – At elevations where humans would be comfortable breathing, the speed of sound is mainly a function of temperature, and so the two input are linked. Changing the temperature setting automatically recalculates the corresponding speed of sound and vice versa.

Impulse Response Options

Most of the controls in the Impulse Response options dialog tab echo the settings of the controls in the main window. We have already discussed those in some detail so we will concentrate here on the few that don't.

Time Domain Display Settings

- *FFT Size* and *Averages* echo the settings of the on-screen controls for dual-channel IR measurements that we talked about previously in topic #7.
- *Overlap %* (for Averaging, not to be confused with overlap for the Spectrograph display) – When Overlap is set to any value other than zero, each successive measurement going into an averaged dual-channel IR measurement shares the specified percentage of data with the previous frame(s).
- *Show IR Peak* sets the locked cursor in IR mode to the highest peak in the impulse response each time you run a new measurement.
- *ALCons Split Time* sets the split time for the early-to-late energy ratio used in calculating ALCons (a type of speech intelligibility estimation that can be calculated from an impulse response). There is no real standard for this parameter but common settings are 10 or 20 milliseconds.



- *Mag Threshold (dBFS)* is similar to magnitude thresholding in transfer function measurements. It is turned off when set to zero. When set to any other value (in dB FS), Smaart will zero out the transfer function at any frequency where the reference signal does not cross threshold before calculating the (dual-channel) impulse response.

Filter Settings

Smaart includes a sweepable highpass filter for IR measurements that can be handy when analyzing IR measurements that include a lot of very low-frequency noise or in cases where the reference signal being used in a dual-channel measurement is band-limited somehow. The filter is applied post process to IR data for display purposes – meaning it can be used for file-based or newly measured data – and only affects what you see on the screen. It does not change the underlying measurement. The cutoff frequency for the filter can be set to any value between 0 Hz and one half of the Nyquist frequency (equal to one half of the sampling frequency) for the currently selected audio sample rate.

Frequency Display Settings

- *Smoothing* echoes the setting of the Smoothing control in the upper right corner of the Frequency graph in IR mode.
- The *Magnitude* range controls are useful for setting a specific decibel range for the Frequency graph in IR mode, however you can also resize the range using the +/- keys or by right-click-and-drag mouse zooming, as you can with other graph types in Smaart.

Spectrograph Settings

- The *FFT Size* and *Overlap* controls echo the settings of the controls found in the upper right corner of the Spectrograph display in IR mode. Together they determine the time resolution of the spectrograph.
- *Grayscale* plots the spectrograph using varying shades of gray instead of color to represent magnitude.
- *Data Window* – sets the data window function used in calculating the individual FFTs used to create the spectrograph display. You can leave this set to Hann unless you have some good reason to change it.
- *Dynamic Range* echoes the settings of the slider control widgets found on the left edge of time domain and spectrograph displays on IR mode. The Spectrograph scales its color (or grayscale) spectrum to the range between the Min and Max values and plots decibel values above the Max thresholds in white and below the Min in black.

Histogram Settings

The Histogram chart in IR mode plots, the values found for any column in the All Bands table band-by-band for all octave or 1/3-octave. Selector controls are found in the upper right corner of the chart in the main window. By default, Smaart plots the histogram chart as a bar graph. Selecting *Plot as Line* in Histogram settings causes this chart to be plotted as a line graph instead.

3: Analyzing Impulse Response Data

Smaart provides a powerful set of tools for analyzing impulse response data in both the time and frequency domains. Time domain analysis tools include logarithmic and linear time-domain views, Energy Time Curves, octave and 1/3-octave bandpass filters, reverse time integration and automatic calculation of common acoustical parameters such as EDT, RT_{60} and clarity factors. Frequency domain analysis tools include spectrum analysis of arbitrary time ranges and the Spectrograph.

If we were discussing Smaart's real-time measurement and analysis mode, we would almost have to pause at this point to set-up and start actively measuring some kind of sound source in order to have something to analyze. But in IR mode, measuring and analyzing are generally two separate things that we can talk about separately. Data analysis in IR mode is an off-line, post-process affair that works the same whether we're onsite actively measuring a system or working with an impulse response recorded in a .wav or .aiff file. Since we just talked about the IR mode user interface in the previous section, let's dive right into actually using it while that discussion is still fresh in our minds.

Most of the examples in this section were created using a handful of .wav files that have been included with this document. Wherever applicable, we will tell you which file was used and how to duplicate our settings, so that you can gain a little hands-on experience as we go if you like. Our first example uses theater.wav, an IR measurement of a 400-seat historical vaudeville theater. The measurement was taken from about the main floor seating area using a small horn-loaded, PA speaker positioned on the stage lip as the excitation source. If you would like to load the file yourself, open up Smaart, switch to IR mode, then select Load Impulse Response from the File menu, and navigate to wherever the file resides on your hard disk to open it.

Time Domain Analysis

Logarithmic Time Domain Display

The time domain IR display with logarithmic (Log) amplitude scaling is probably the most familiar to anyone much accustomed to looking at acoustical impulse responses. In this view you can find the arrival times of direct sound and early reflections and overlay the reverse time integration of the IR, along with interactive widgets to calculate EDT and reverberation time (on Log and ETC displays only). Smaart provides octave- and 1/3-octave bandpass filters that you can use to filter the IR on the fly, to see how reverberant decay and other characteristics change with frequency.

The combination of locked and free cursors on time domain displays enables you to find the relative arrival time and amplitude differences between any two points on the plot. The difference between the two is shown in the cursor readout. If the *Milliseconds and Distance* option is selected in the *Cursor Time Readout* section of the *General* options page (Options menu > General), Smaart will also give you equivalent distances for time coordinates, based on the current *Speed of Sound* settings. To move the locked cursor to an arbitrary point on the plot, hold down the Ctrl key (Cmd key on Mac) on your keyboard while clicking with your mouse on a point that you want to mark. Pressing Ctrl/Cmd + "P" resets the locked cursor to the highest peak IR.

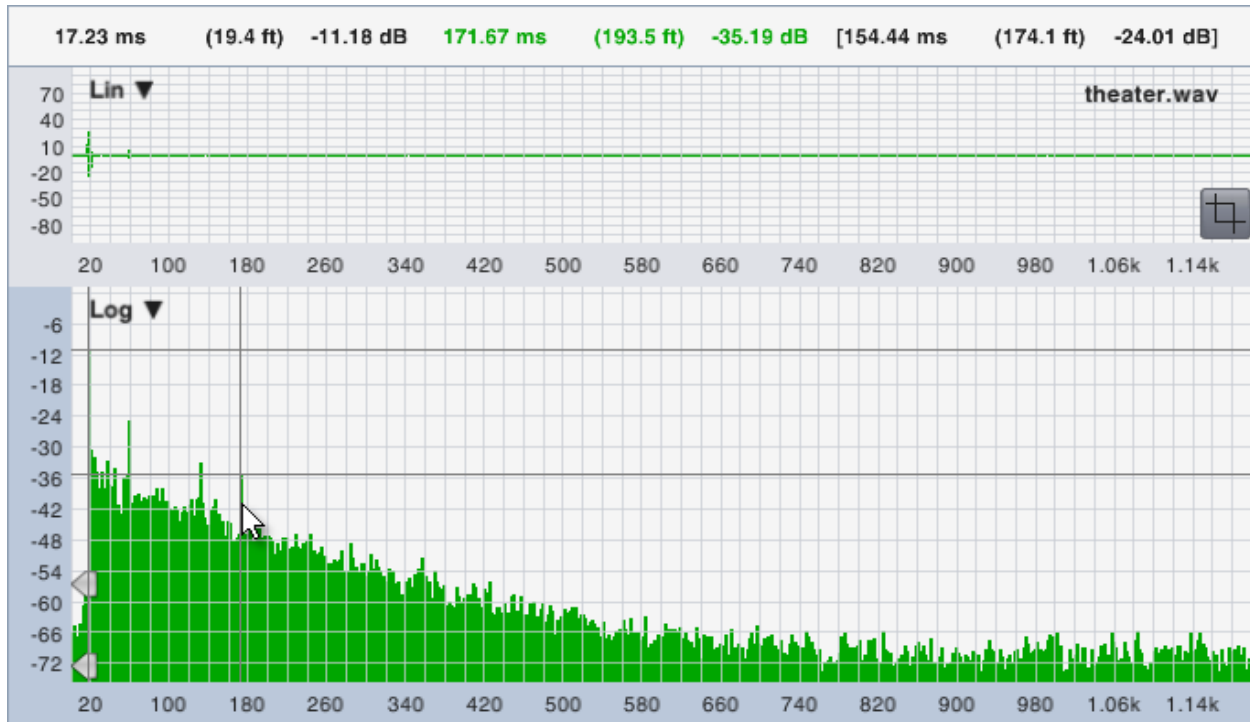


Figure 3: The logarithmic (Log) time domain graph plots time on the x axis and magnitude in decibels on the y axis. The combination of locked and movable cursors enables you to find time and level differences between any two points on the plot. Time coordinates can optionally be plotted with equivalent distances as shown above. The pair of coordinates on the far left in the cursor readout at the top of the frame is the locked cursor position, which is set to the highest peak in the IR. The middle pair of coordinates in green is the absolute location for the movable cursor and the rightmost pair in brackets is the difference between the first two.

Linear Time Domain Display

A linear (Lin) time domain chart plots the same data as the Log IR but on a normalized linear amplitude scale, where amplitude values are given as a percentage of digital full scale. This view tends to be of limited usefulness for acoustical analysis in general, however it can be a very good tool for finding discrete reflections, particularly when measuring in an empty hall before an audience arrives. In this case, using the linear IR view can help you to identify hard reflections that might be masked by the diffuse reverberant field on a logarithmic display, only to become much more obvious and audible (often on stage, to the consternation of opera singers) once there is an audience in place and the reverberant levels decrease.

Another thing the Linear IR can tell you that the Log and ETC graphs can't, is relative polarity. For example you could measure two midrange drivers or other like devices and determine if they are wired with the same or different polarity by noting which direction the prominent peaks in the impulse are pointed. Figure 4 shows a zoomed in view of the linear (Lin) scaled impulse response of a 2nd order Butterworth bandpass filter with normal and inverse polarity. Cutoff frequencies for the filter are 400 and 1600 Hz. It's easy to see that the peaks in the two IRs are pointed in different directions relative to each other. Unfortunately, this doesn't necessarily tell you which one is correct. But if you measured three like devices and one was different, you might reasonably say that the majority rules. Or if you

measured two like devices and found opposite polarity and one of them sounded better, it's possible you might have found the problem. Linear view can also come in handy for looking at other types of signals in the time domain other than impulse responses.

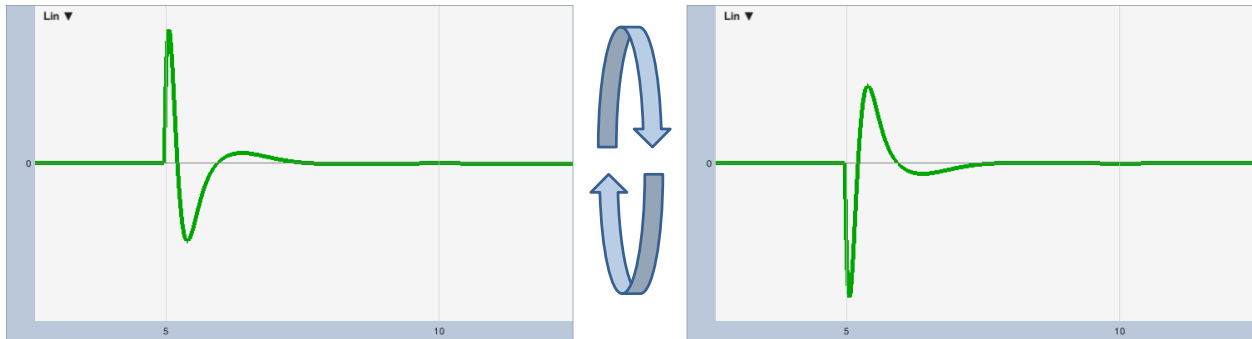


Figure 4: Zoomed in views of the linear impulse response of a bandpass filter, with normal and inverse polarity.

One thing the Linear IR *can* tell you that the Log and ETC graphs can't is relative polarity. For example you could measure two midrange drivers or other like devices and determine if they are wired with the same or different polarity by noting which direction the prominent peaks in the impulse are pointed. Figure 4 shows a zoomed in view of the linear (Lin) scaled impulse response of a 2nd order Butterworth bandpass filter with normal and inverse polarity. Cutoff frequencies for the filter are 400 and 1600 Hz. It's easy to see that the peaks in the two IRs are pointed in different directions relative to each other. Unfortunately this doesn't necessarily tell you which one is correct. But if you measured three like devices and one was different, you might reasonably say that the majority rules. Or if you measured two like devices and found opposite polarity and one of them sounded better, it's possible you might have found the problem.

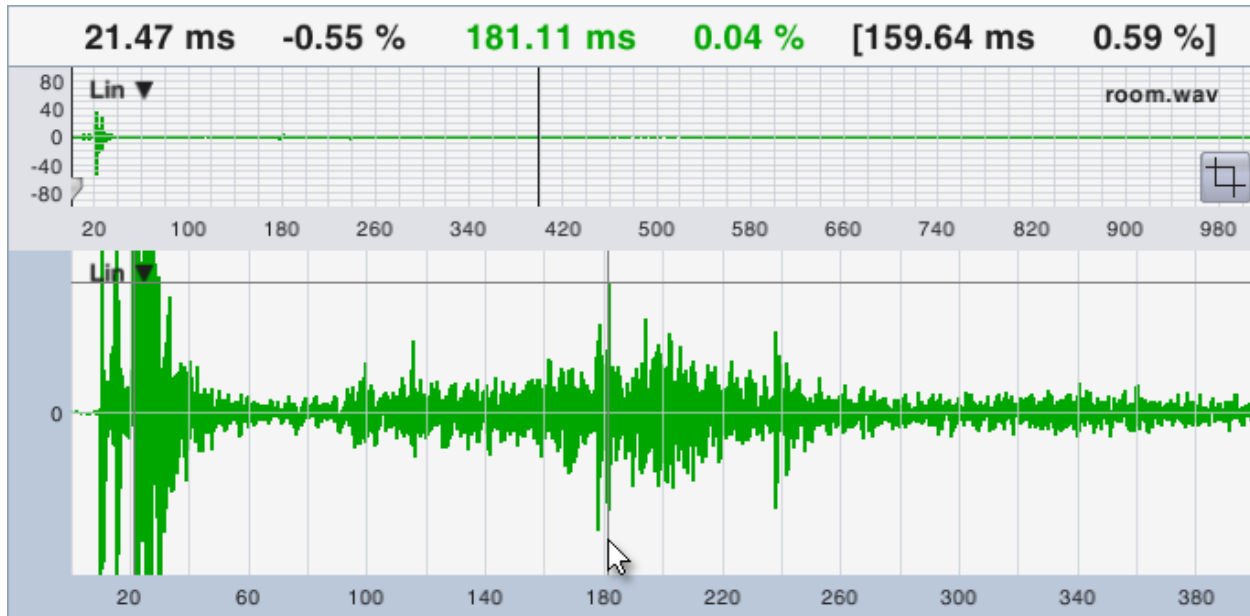


Figure 5: Zooming in on a Linear (Lin) time domain view of room.wav and using the cursor readout to find the relative arrival time of a prominent discrete reflection.

Energy Time Curve (ETC)

The impulse response represents a 2-D graph of a 3-D event: the magnitude and phase of the energy arrival over time. With magnitude on the vertical (y) axis and time on the horizontal (x) axis, phase ends up being represented on the z -axis, which is effectively lost in this view. Consequently, in the linear and log views of the IR, energy arrival that is 90° or 270° shifted shows as a zero crossings, thereby making a single arrival that is spread out over time and phase appear to be multiple arrivals.

The Energy Time Curve, also called envelope of the impulse response, represents the magnitude of the energy arrival over time by effectively ignoring phase. The textbook description is the real impulse response combined with its Hilbert transform – a copy of itself that has been rotated 90° in phase. In practical terms, the summation of the two tends to fill in zero crossings seen in the Log IR, producing a signal that can be a lot easier to look at than the Log IR by virtue of being less squiggly. At higher frequencies the Log IR and ETC may look very similar – both are plotted on a logarithmic magnitude scale – but the ETC is particularly useful sizing up the arrival of direct sound at low frequencies.

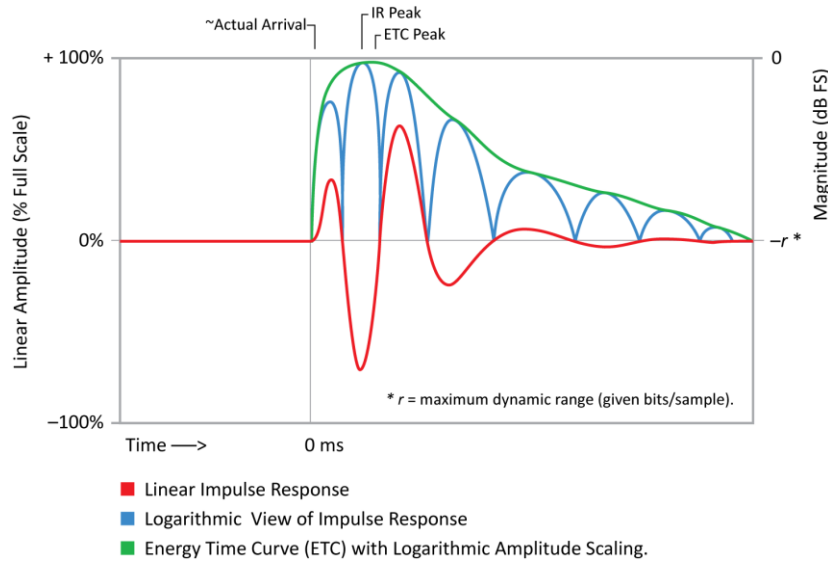


Figure 6: A comparison of the ETC and the impulse response with linear and logarithmic amplitude scaling.

If you zoom in on the first 250 ms of room.wav and switch to the 125 Hz octave band, the difference between the Log IR and the ETC is pretty striking (see Figure 7). To scale your display to look like Figure 7, press the plus [+] key on your keyboard a few times to zoom in on the magnitude range then use the up/down arrow keys to move the range up and down.

Note that when using peak locations to find delays, the ETC can sometimes give you a slightly different answer than the Log IR, because of the way it effectively interpolates between peaks in the IR. If you look at the smaller peak in the ETC, at about 124 ms in Figure 7, you can see that it falls in between two lobes in the Log IR. We have found that the ETC can be more effective than the Log IR tool for finding subwoofer delay times. But that is better done in real-time mode, using the ETC on the Live IR in conjunction with the frequency domain transfer function displays, where you can see phase as well as magnitude and watch changes happening in real time as you adjust processor settings. That's a little beyond the scope of this document.

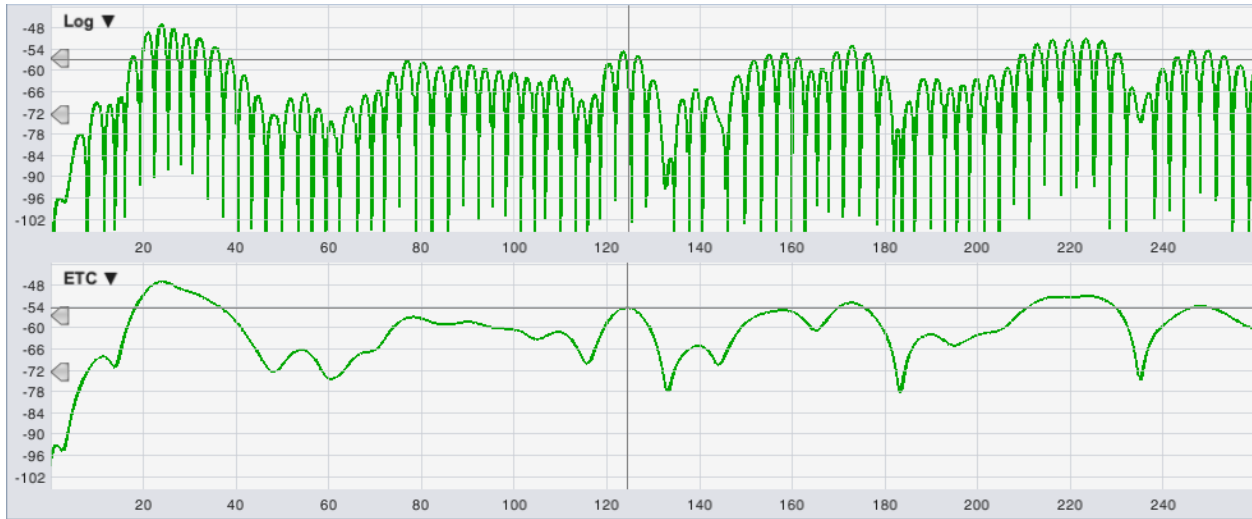
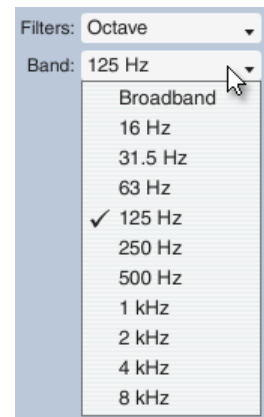


Figure 7: A comparison of the Log IR and ETC graphs in Smaart for the 125 Hz octave band in room.wav.

Bandpass Filtering

Up to now we have mainly been looking at the broadband IR, but quite a lot of acoustical analysis is conventionally done using octave, or sometimes 1/3-octave bands, especially as we get into reverberation times and early-to-late energy ratios. Smaart includes complete sets of octave and 1/3-octave bandpass filters for the octaves between 16 Hz and 16 kHz (assuming 48k or higher sampling rate, at lower sample rates you lose some of the upper bands). Bandpass filtering in Smaart is done non-destructively, on demand. To see a filtered version of the IR, select which set of filters to use (*Octave* or *1/3 Octave*) on the *Filters* selector, then select the band that you want to look at from the *Band* list.

Smaart's bandpass filters have linear phase response and their magnitude response satisfies the most stringent (Type 0) tolerances for octave and fractional-octave bandpass filters specified in IEC 61260 and ANSI S1.11. If you would like to see the magnitude response of the bandpass filters you can load the wave file *1samplePulse.wav* and bring up the *Frequency* graph, then step through the *Bands* list to see each filter. Bandpass filtering applies to all main display types except the Histogram chart (which is already filtered into bands). It does not affect the small graph in the navigation pane. Note that filtering the impulse response will clear the Spectrograph display if present and require a recalculation (by clicking the *Calc* button again).



Discrete Reflections

Reflections are a complicated subject because humans are very good at processing them. They may be useful or detrimental, depending on such factors as their arrival time and loudness relative to direct sound (the two biggies), their frequency content and even the angle they arrive from. Discrete reflections can cause audible problems ranging from coloration (timbre change) to image shift to audible echoes. But trying to figure out which reflections are friend or enemy by looking at squiggly lines on a computer screen can be a bit of a dicey prospect.

Short reflections arriving within the first 30 milliseconds or so after the direct sound at relatively high levels are notorious for producing comb filters that muck up our real-time frequency domain analysis; but humans actually find them beneficial, enhancing the intelligibility of speech and the clarity of music. Outside that early integration window, reflections can still contribute to subjective impressions of presence, warmth, spaciousness, etc. But the rules are a little different for speech vs music.

Individual broadband reflections arriving at 95 ms or more can destroy speech intelligibility and also make life difficult for presenters and performers if they reach the stage. At that point you're getting into the high end of the syllabic rate for very fast human speech, and people find it disorienting to be hearing the last syllable they spoke or sang as they are attempting to utter their next. That was actually the problem being investigated in the IR measurement shown in Figure 18, where a high-level reflection was arriving at about 160 ms, which is close to the average syllabic rate for normal conversational speech.

Low-order, early reflections may be visible on time domain plots as individual peaks following the arrival of direct sound. Later arrivals can show up as spikes protruding from the reverberant decay slope. On the Spectrograph plot, higher level broadband reflections can often be identified as distinct vertical streaks when you run the dynamic range controls up and down, particularly the Max setting. They tend to be most problematic when arriving at longer delay times and relatively high levels, compared to the level of the diffuse reverberant field.

A pretty good rule of thumb is that the later the arrival, the lower in level it needs to be in order to be perceived as beneficial or neutral. Another is that our tolerances for reflected sounds and reverberance tend to be wider for music than speech. In general, Smaart is very useful for identifying problematic reflections; however your ears are probably still the best tool for evaluating their relative significance or severity.

Reverberation Time

Reverberation time (commonly referred to as T60 or RT60, or somewhat less commonly as T30, T20 or simply T) is the time required for reverberant sound energy in a space to decay by 60 dB from an excited level. It is regarded as an important metric in the acoustics of musical performance spaces and also classrooms, auditoriums and cinemas, where it is used as a rough predictor of speech intelligibility.

Reverse Time Integration

Reverberation time is calculated from the reverse integration of an impulse response that has been filtered into octave bands. Conventionally, the 125 Hz to 4 kHz bands are evaluated. Reverse time integration is also called Schroeder integration, after Dr. Manfred Schroeder whose brain child it was. It is a simple thing in concept, but it can be a little tricky to do well.

In theory, you just start at the end of the time record and work your way back to the beginning, tallying up the squares of each sample in the IR as you go. A common problem however, is that the integration will flatten out when the reverberant decay slope runs into the noise floor of the IR. This can lead to overestimation of the reverberation time, particularly if the IR has limited dynamic range, and/or a lengthy noise tail.

The most straightforward solution for this problem is to find the point in the IR where the decay slope meets the noise floor, sometimes referred to as the “saddle point,” and begin the integration there, rather than at the end of the file. The location of the saddle point in an IR is notoriously difficult to estimate automatically though. Smaart (7.5 or higher) uses a proprietary algorithm for IR saddle point estimation that has proven quite robust, but it is not completely foolproof. So it’s still a good idea to check each band to make sure that you agree with the choices the software makes – particularly if there are any large anomalies in the tail of the IR, such as a prominent spike or distortion products piled up at the end of the record by a sweep signal.

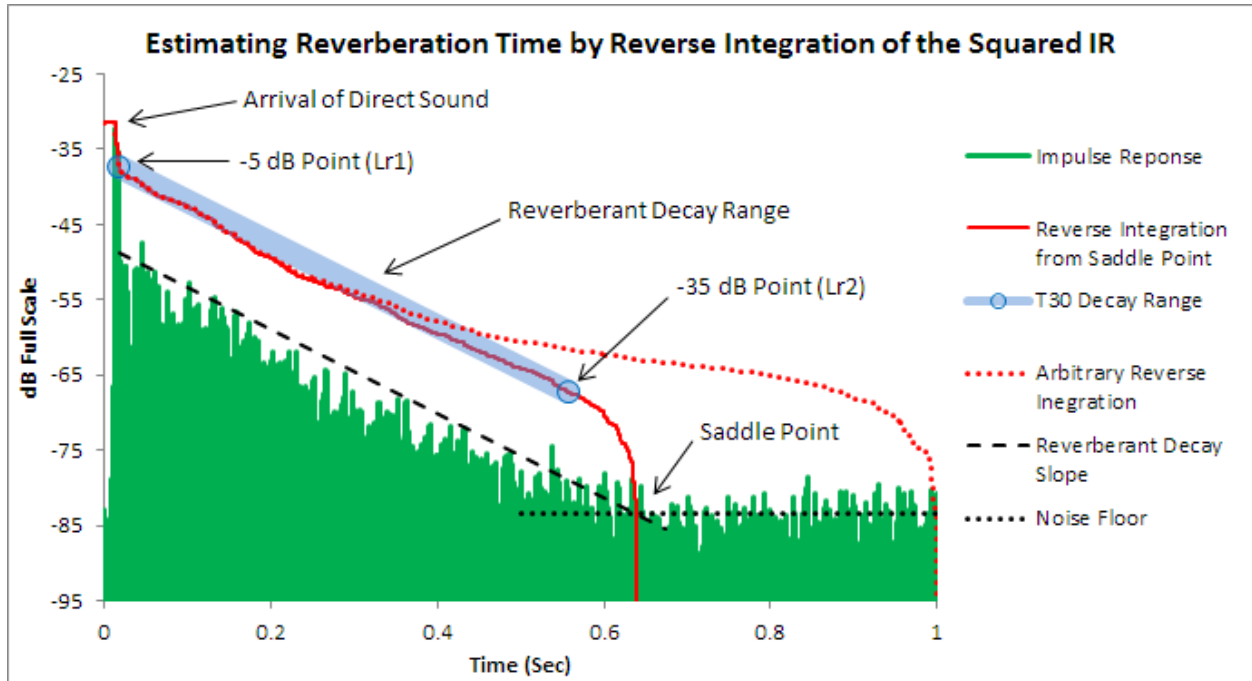


Figure 8: Estimating Reverberation time by reverse integration of the impulse response. Reverse integration of the IR from the “saddle point” – the approximate point where the reverberant decay slope meets the noise floor of the measurement – provides a very good estimation of reverberant decay time. Starting the reverse integration from an arbitrary point such as the end of the file, may result in overestimation of decay time.

Evaluation Ranges (EDT, T20, T30)

Because it is rarely possible to actually measure a full 60 dB of reverberant decay in acoustical systems, reverberation is typically evaluated over a smaller range. The starting point is always 5 dB down on the reverse integration curve from the point corresponding to the arrival of direct sound. The end point of the range is 30 dB down the curve from the starting point, provided that it is at least 10 dB above the noise floor – if not, a 20 dB range may be used. In either case, the measured decay time is extrapolated to the equivalent 60 dB decay time. In ISO 3382 parlance, these are referred to as T20 or T30. Early decay time (EDT) is conventionally measured from the arrival of direct sound down to 10 dB below it on the integration curve. Like reverberation time, EDT is also normalized to 60 dB decay time.

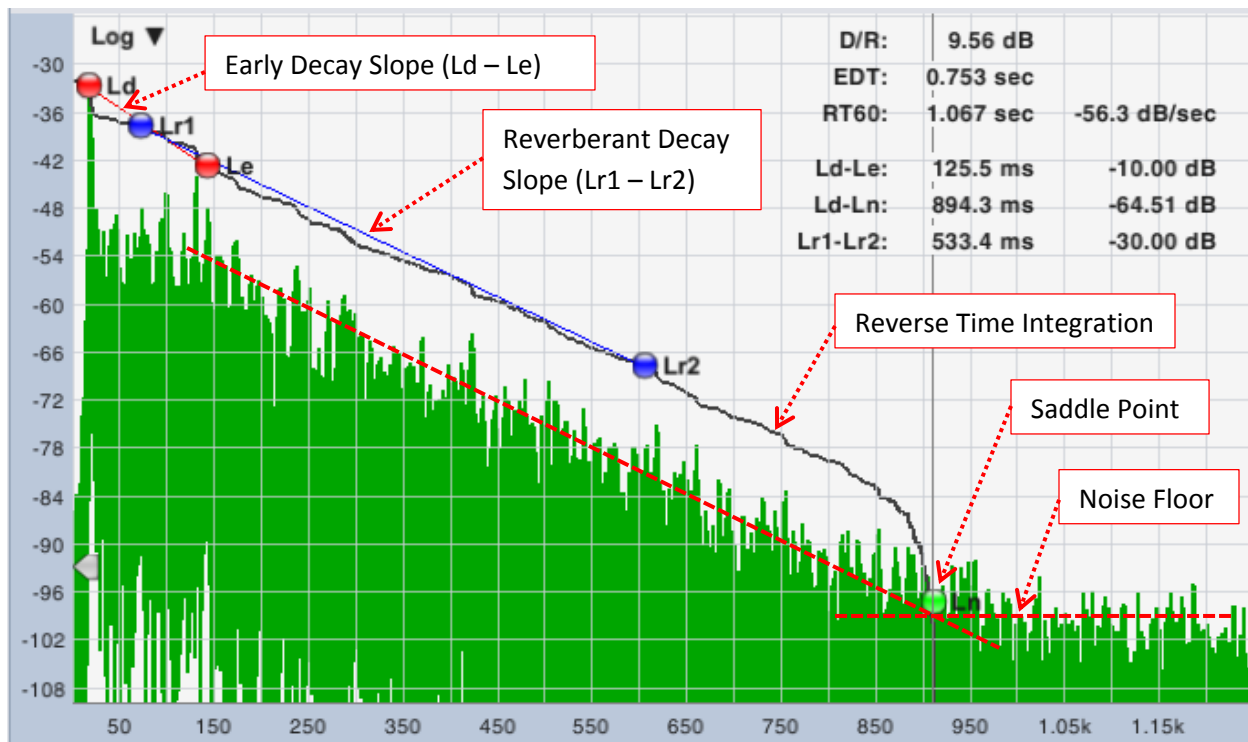


Figure 9: A log IR display with all the bells and whistles. The IR is the 500 Hz octave band of theater.wav. Clicking the Schroeder and RT60 buttons displays the reverse time integration curve and the start and end points for the EDT and RT60 evaluation ranges on Log or ETC charts. The positions of all the level markers (Ld, Le, Lr1, Lr2 and Ln) are user adjustable, however they shouldn't require very much adjusting under most circumstances.

Notice the five level marker widgets shown on the plot in Figure 9. If you were wondering about the cryptic labels, your secret decoder and the default positions for each of the markers is as follows.

- **Ld = Level Direct.** This marker is positioned on the reverse integration curve at the point corresponding to the arrival time of direct sound.
- **Le = Level Early (Decay).** This marker is automatically positioned 10 dB down from the Ld marker on the reverse integration curve. The slope between Ld and Le is used to calculate EDT.
- **Lr1 = Level Reverberant 1.** This marker designates the top of the reverberant decay range, 5 dB down on the reverse integration curve from the Ld marker. All of the level markers are user adjustable but positioning these three is pretty cut and dried. You should rarely find any need to touch them.
- **Lr2 = Level Reverberant 2.** This marker designates the end point for the reverberant decay slope. If there is sufficient dynamic range it should be placed 30 dB down the reverse integration curve from Lr1. If not, 20 dB will do. Lr2 is one of the two markers that you may sometimes want to adjust by hand; the other is Ln (below).
- **Ln = Level of Noise.** This is typically the most subjective of the five markers in terms of placement. The time location determines the start point for the reverse time integration curve, which is the basis for positioning all of the other markers. Ideally this will roughly correspond to the saddle point in the impulse response. The magnitude coordinate is used to estimate the

level of the noise floor of the measurement and the Lr2 marker needs to be at least 10 dB above that. As of version 7.5, Smart does a pretty good job of placing the Ln marker most of the time. However, it may still benefit from a human touch in some cases – particularly if the dynamic range of the measurement is marginal or there are significant distortion artifacts from a swept sine measurement or any other prominent anomalies in the noise tail of the IR being analyzed.

When the level marker widgets for reverberation time are visible, a block of vital statistics also appears in the upper right corner of the plot. These include the 60 dB reverberation and early decay times (RT60 and EDT) and the time and level differences between three pairs of markers. The Ld-Le level difference should always be 10 dB. Lr1-Lr2 should be either 20 or 30 dB – this number is convenient for checking your work if you end up adjusting Lr2 is by hand. The Ld-Ln delta is interesting also, as it gives you the dynamic range of the measurement. D/R stands for direct/reverberant ratio. It is an early-to-late energy ratio that gets its split time from the time coordinate of the *Le* marker.

Saving your work

If you do end up adjusting any of the Level marker positions by hand, you can save their positions to a comma-separated values (.csv) text file, by selecting *Save Decay Markers* from the *File* menu. To reload them again later, first load the .wav or .aiff file containing the impulse response, then select *Load Decay Markers* from the *File* menu to open the marker position file.

Reporting Results for Reverberation Time

Reverberation time ideally should be measured from several locations throughout the room and the results from each measurement position averaged together, octave band by octave band to get an average decay time for each octave. Smart doesn't do that part for you, but the *All Bands* table does make it easy to get the data from each measurement into a spreadsheet.

Frequency Ranges

The standard evaluation range for reverberation time is the six one-octave bands from the 125 Hz to 4 kHz. Average times for each octave band can be presented in a table or on a graph. When presenting reverb times on a graph, the frequency axis of the graph should be labeled with the IEC standard nominal octave band center frequencies. The y-axis of the graph should have an origin of 0 and be labeled in seconds. It should be noted both in the table and on the graph whether T₂₀ or T₃₀ was used. ISO 3382-1 specifies that if a graph is presented it should be a line graph with a standardized aspect ratio of 2.5 cm per second and 1.5 cm per octave. ISO 3382-2 isn't so picky. It just says "a graph."

Reverberation times for the 125 and 250 Hz bands may be averaged together to get a T_{Low} figure. The average of the 500 Hz and 1 kHz band is called T_{Mid}. When a single number figure is given for reverberation time, it is assumed to be T_{Mid} unless otherwise stated. Smart calculates these values for you automatically and displays them in the *All Bands* Table.

Dividing T_{Low} by T_{Mid} gives you the Bass Ratio. Bass ratio quantifies the "warmth" of sound in a venue and is a particularly important parameter for concert halls. The word "Bass" in this case refers to vocal or instrument bass registers and should not be confused with PA-type sub-bass frequencies. Acceptable values are dependent on expectations. A Bass Ratio of 1.1 - 1.25 would be regarded as good for fairly

reverberant concert halls (RT 60 greater than 1.8 seconds) but the upper figure could be increased to 1.45 for less reverberant spaces.

As for what to look for in reverberation time results, preferred reverberation times vary according to room size and purpose and the type of program material being presented. In general you want to see shorter reverberation times for auditoriums, classrooms, theaters and cinemas – ideally from about 0.4 - 0.5 seconds for smaller rooms up to 0.8 - 1.2 seconds for larger rooms. Opera houses and mixed use performance spaces where both speech intelligibility and musical appreciation are equally important typically aim for the lower end of 1.2 – 1.8 second range. Spaces intended for symphonic performances and organ music can range from about 1.8 seconds up to three seconds or more in very large halls.

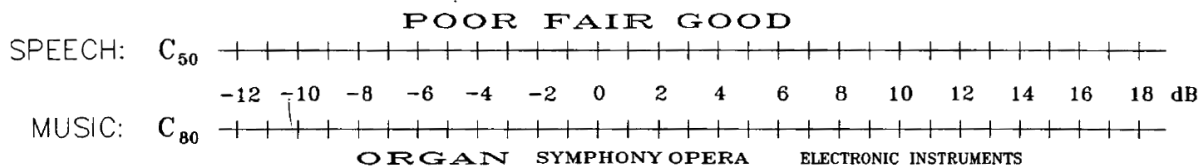
Reverberation times that are roughly equal across all frequencies are generally preferable for most purposes. The exceptions are things like choral, organ and romantic classical musical music, where a reverberation time curve weighted more toward the lower frequencies may be preferred. It’s pretty normal for higher frequencies to decay faster than lows but you don’t want to see times that are wildly different in neighboring octaves. In general though, acoustical treatments and/or physical changes to the sound system are typically required to effectively address problems any problems you may find.

Early-to-Late Energy Ratios

Early-to-late energy ratios are another way of objectively characterizing the reverberant characteristics of a room. They are arguably a better measure than reverberation time for any venue where a sound system is an organic part of the acoustical equation. They are simple to calculate automatically and are not subject to the kinds of complications that can make measurement of reverberation times somewhat subjective, but they are a more recent innovation and may be less widely understood than RT60.

Clarity Ratios (C35, C50, C80...)

Clarity indexes are early-to-late energy ratios that compare the integral of the energy arriving within the first n milliseconds of the arrival of direct sound (inclusive) to the energy in the remainder of the reverberant decay period. The two most commonly used are C50 and C80, which use at the 50 or 80 milliseconds respectively as their split times. The result of the comparison is expressed as a decibel ratio.



TEST FREQUENCIES:

SPEECH: 0.5k, 1k, 2k, & 4kHz octave-band C₅₀ values, intelligibility weighted & summed

MUSIC: 0.5, 1k, & 2kHz octave-band C₈₀ values, averaged

Figure 10: A scale for interpreting C50 and C80 measurement results for speech and music.

Shorter split times such as 35 or 50 ms are regarded as better predictors of speech intelligibility. C80 is more useful for music. In terms of what kinds of numbers to look for, Gerald Marshall provided the table shown in Figure 10 in the in a 1996 Journal of the AES article titled, *An Analysis Procedure for Room Acoustics and Sound Amplification Systems Based on the Early-to-Late Sound Energy Ratio*.

For the speech intelligibility side of the graph, Marshall used a weighted average of the 500 Hz to 4 kHz octave bands, with the following weights assigned to each band: 15% for 500 Hz, 25% for 1 kHz, 35% for 2 kHz and 25% for 4 kHz. Others have used the weighting tables for Articulation Index and STI and other scales of their own devising with similar results. For music, he used a simple average of the 500 Hz, 1 kHz and 2 kHz octave bands. We know of no applicable standards for this metric and it has been suggested extending the frequency ranges Marshall used an octave higher for speech and two octaves higher for music that it might be useful, but hopefully this example provides a useful starting point for evaluation.

The Histogram Display

Selecting *Histogram* as your display type for an IR mode graph plots a chart of all reverberation times or early-to-late energy ratios by octave or 1/3 octave bands. The type of data the Histogram displays and the resolution are selected by means of the list control in the upper right of the graph. You can change the Histogram to a line chart by opening the Impulse Response options page (*Options Menu > Impulse Response*) and selecting *Plot as Line* under *Histogram Settings*.

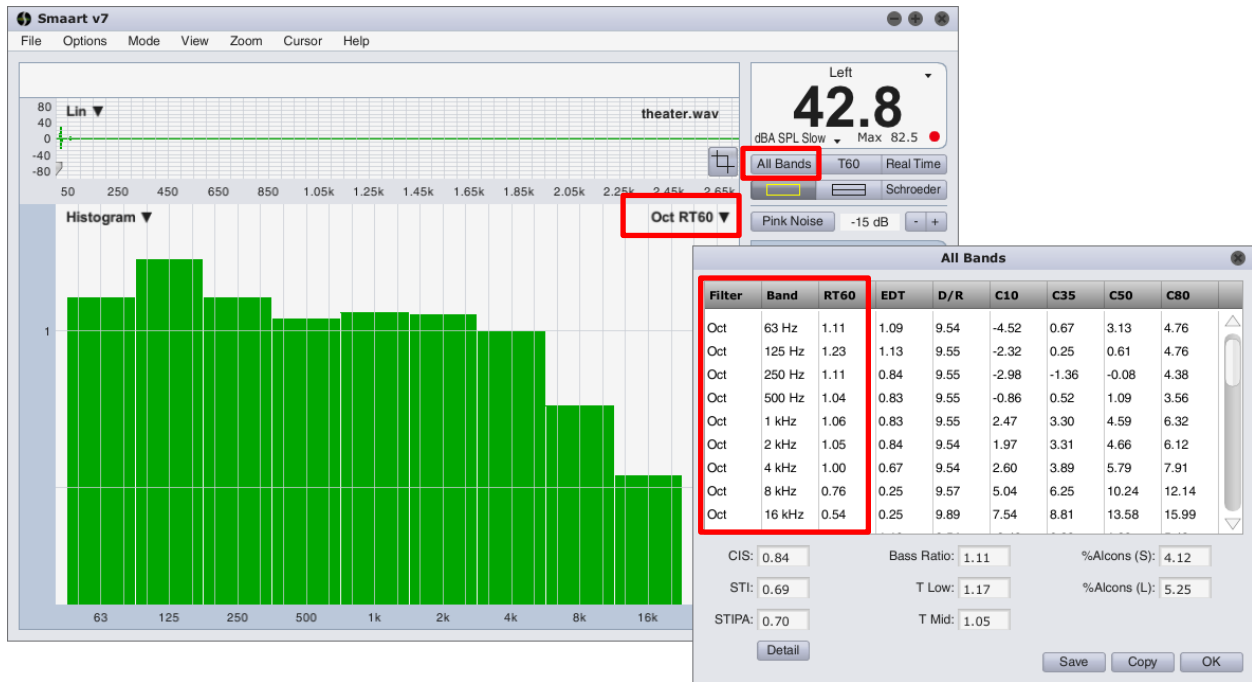


Figure 11: The Histogram graph and All Bands Table. The All Bands Table in Smaart collects reverberation times and early-to-late energy ratios for all octave and 1/3-octave bands in a single table. Speech intelligibility metrics (STI and ALCons) are displayed here as well. The Histogram chart can plot any column of the All Bands Table as a bar graph or line chart in octave or 1/3-octave resolution.

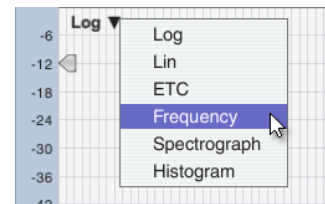
The All Bands Table

Clicking the *All Bands* button in Impulse mode in Smaart 7 brings up a report window containing just about every acoustical quantity that Smaart can calculate automatically from an IR, for each octave band and 1/3-octave band where applicable (see Figure 11). Speech intelligibility metrics (STI and ALCons) are displayed here along with Bass Ratio, T Mid and T Low, which are calculated from reverberation times for the 125 Hz to 1 kHz octave bands.

Clicking the *Copy* button in this window copies the entire table to the operating system's clipboard in tab-delimited ASCII format suitable for pasting into a spreadsheet or any other program that accepts ASCII text. You can also save it to a text file by clicking the *Save* button.

Frequency Domain Analysis

Selecting *Frequency* as your graph type in the main graph area automatically transforms the IR into the frequency domain to show you its spectrum. The Frequency graph has frequency in Hertz on the (horizontal) *x* axis and magnitude in decibels on (vertical) *y* axis. The *Smoothing* control in the upper left corner of the Frequency graph works exactly the same way as smoothing on the real-time transfer function display.

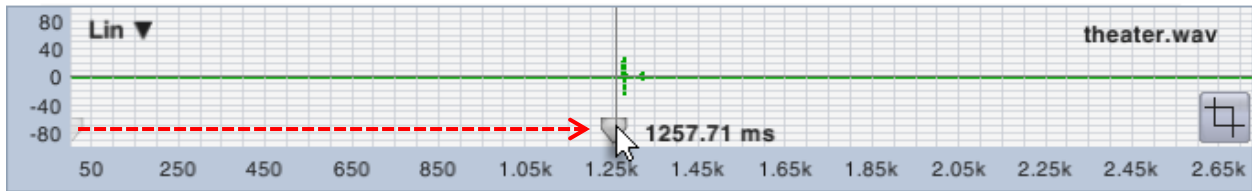


Smaart can calculate arbitrary-length DFTs in IR mode to give you the spectrum of virtually any subset of the IR time record that you care to zoom in on. Time and frequency domain displays are linked so that zooming in on a time-domain graph (Lin, Log or ETC) automatically changes the Frequency display to match. When the entire time record is selected, there's an assumption that you are analyzing a dual-FFT IR measured in Smaart and so no data window is used in calculating the spectrum in that case.

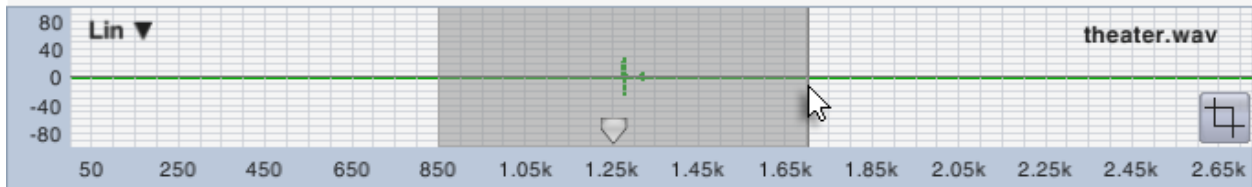
Smaart automatically uses a tapered data window when transforming any subset of the time record, so if you are analyzing an IR file from some other source or a file that has been cropped to less than its original length, you may see better results if you zoom in slightly in the time domain. Tapered data windows significantly attenuate data at the edges of a selected time range – many go all the way to 0 – so you generally want to position any peaks that you want to examine near the center of a selected range. The *Time 0* slider in the navigation pane can be used to move peak structures nearer to the center of the time window if they are too close to the edge to center up in the range. Clicking in the right margin of the navigation clears a time zoom.

If you zoom *way* in on the IR and select a very narrow time range centered on the arrival of direct sound it's possible to see the magnitude response of loudspeaker without comb filters caused by early reflections, at least at high frequencies. In practice, the usefulness of this strategy may be limited to how far away both the loudspeaker and microphone are from the nearest reflecting surfaces. The frequency response of a DFT is limited by its time constant, so you may find that by the time you squeeze the time window in enough to get rid of first order reflections, you can't really see much detail in the frequency domain. But it's something that people used to do quite a bit, back in the days before lab-measured anechoic response data for most professional loudspeakers became commonly available.

Step 1: Click and drag Time 0 slider to right.



Step 2: Right-click (Ctrl + click on Mac) and drag in navigation pane to select a desired time range.



Result: The spectrum of the selected time range is displayed in the Frequency graph.



Figure 12: Moving the time 0 point and selecting a time range for display. The time range selected in the navigation pane applies to the Frequency graph as well as time domain graphs (Lin, Log or ETC). Note that Smaart uses a tapered data window when transforming any subset of the full IR time range. We've drawn the outline of a Hann window in red on the navigation pane of the "result" portion of the illustration above to help visualize this.

The Spectrograph

Smaart's *Spectrograph* straddles the time and frequency domains giving you a birds-eye view of frequency domain characteristics of the IR over time. If you are familiar with Smaart's real-time spectrograph, the IR mode version is essentially the same display rotated 90°. If not, one way to think about how it works is to start with the idea of a spectrum analyzer. On a real-time spectrum analyzer (RTA) you typically have a bar graph or line chart with frequency on the x axis and magnitude in dB on the y axis, showing you the spectrum of some chunk of signal at a given moment in time – perhaps something like the one shown in Figure 13a.

An RTA is very useful tool, but if you want to gain a better understanding how the spectrum of a signal changes over *time*, you either need a really good memory or a different kind of graph. One solution might be to just keep sliding the old data to the back instead of erasing it as new data comes in, to form a 3-D graph with time on the z axis, as in Figure 13b. If you did this with a 3-D area chart instead of a bar graph it's commonly called a waterfall chart, but let's continue with the bar graph example, as both have the same limitation. The problem with this approach is that as new data comes in, higher-level values in front will cover up some of the data in back, so that you only get a partial picture, as in Figure 13.

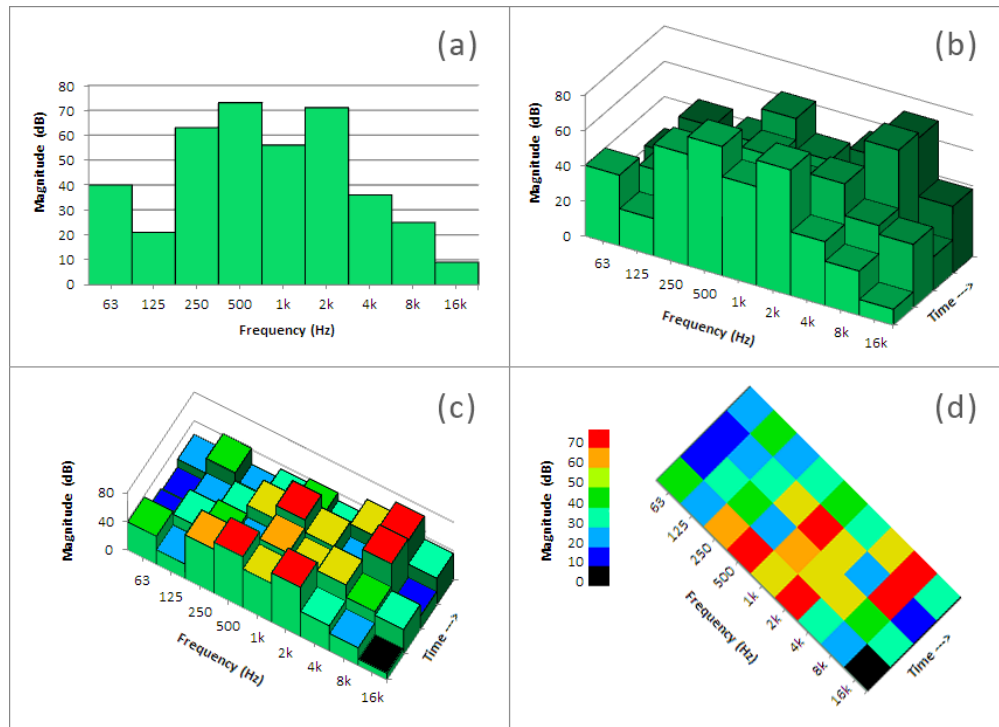


Figure 13: Turning a spectrum analyzer into a spectrograph.

You *could* rotate the graph in space until you can see all the bars (Figure 12c), but when you do that it becomes harder to discern how tall they all are. Assuming that you have a color display (waterfall charts were popular before anyone had color monitors) you might try painting the tops of the bars different colors to better show their magnitudes as we did. But at that point, it would much more readable if you just dispensed with the bar graph and plotted a 2-D chart instead, with frequency on one axis, time on the other, and magnitude indicated by color (Figure 13d). That's a spectrograph.

Generally, the “domain” of a graph is the independent variable, e.g., time or frequency, which is normally assigned to the horizontal (*x*) axis, but the spectrograph display has two independent variables. You can orient it whichever way is most convenient. In real-time mode in Smaart we want to relate the spectrograph to other frequency-domain graphs, so we plot it with frequency on the *x* axis and time on the *y* axis. In IR mode, we most often want to look at it in the context of other time-domain graphs, so we do it the other way around. Time goes on the *x* axis and frequency on the *y* axis.

To bring up the spectrograph in IR mode, click on the graph selector in the upper left corner of a main graph area pane and select Spectrograph. Initially, you are presented with a blank chart area until you click the *Calc* (calculate) button in the upper right corner of the graph pane. The Spectrograph can eat up a lot of graphics resources when it repaints and so we try to paint it only when necessary. Changing the time range selection in the navigation pane does not affect the spectrograph as it does the other time domain (Lin, Log, and ETC) and Frequency graphs in IR mode but moving Time 0, cropping the time record or filtering the IR will clear the spectrograph and require clicking the Calc button again. You can resize the spectrograph and move its range using the +/– arrow keys or right-click-and-drag on the plot to zoom in on a selected range as you can with any other graph in Smart. Also, as with other graph types in Smart, clicking in the left margin of the plot after zooming in will clear the zoom and return the plot to its previous x/y range.

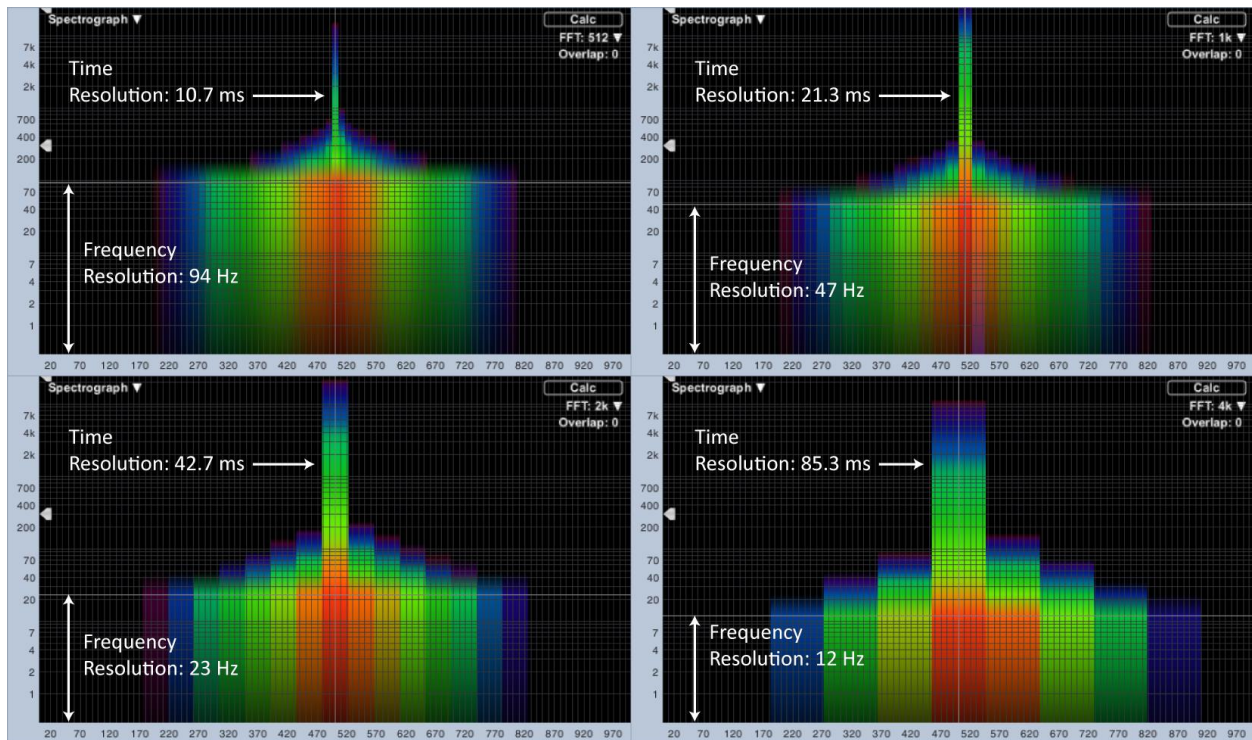
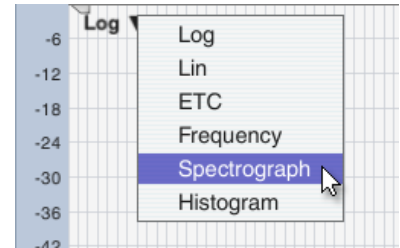
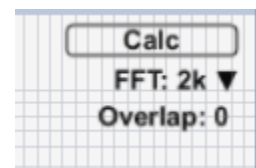


Figure 14: Spectrograph time and frequency resolution as a function of FFT size. FFT sizes ranging from 512 samples to 4K samples are compared, using 0% Overlap. As the FFT size is increased, frequency resolution improved but the peak of the IR is smeared out over a wider time range. The x axis of the graph is time, with frequency on the y axis.

Spectrograph Time and Frequency Resolution

In Figure 13 we used an octave band analyzer as an example for the sake of simplicity. Smart 7’s spectrograph is narrowband but the concept is the same. Each vertical stripe in the IR mode spectrograph represents one FFT, meaning the FFT size determines both the time and frequency resolution of the display. Larger



FFTs provide greater detail on the frequency axis but may mask transient events on the time axis, so it's a trade-off in that respect. The FFT and Overlap controls that appear below the *Calc* button determine the time and frequency resolution of the spectrograph display.

Figure 14 illustrates how this relationship works. It was created using the file *6dbOctImpulse.wav*; the impulse response of a linear phase lowpass filter with 6 dB per octave roll-off. The sharpest part of the peak in the impulse response, where most of the HF energy lives, is probably no more than a few milliseconds wide but notice how its energy is spread across the full FFT time constant in every case. At 512 points, time resolution (the FFT time constant) is a respectable 10.7 milliseconds but the FFT *frequency* bins are spaced almost 100 Hz apart. Increasing the FFT size to 4K points gets you 12 Hz frequency resolution but smears the peak in the IR over an 85 ms time range.

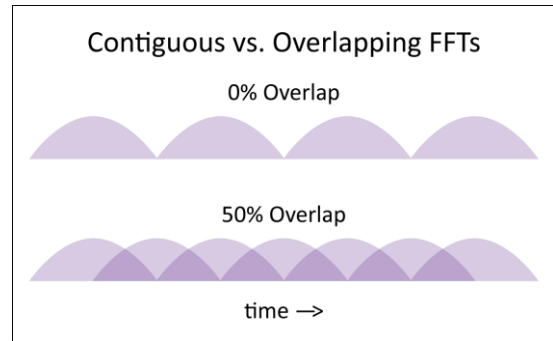


Figure 15: With FFT overlap set to 0%, each FFT is calculated from unique data. At 50% overlap, the darker shaded areas are shared by successive FFTs. Our "FFTs" are drawn as flattened half circles to suggest a tapered FFT data window.

The other factor affecting the time resolution of the spectrograph is the *Overlap* percentage that you specify in the upper right corner of the graph. When Overlap is set to zero as in Figure 15, each successive FFT "slice" of the Spectrograph is calculated from unique time domain data, each frame beginning where the last one ended. When Overlap is set to any non-zero value, each successive FFT frame shares some percentage of its data in common with the previous frame(s). The FFT time constant is still the FFT time constant but more overlap can sometimes allude to, if not exactly restore some missing detail on the time axis as FFT size is increased, in addition to producing smoother blending between "slices." (See Figure 16).

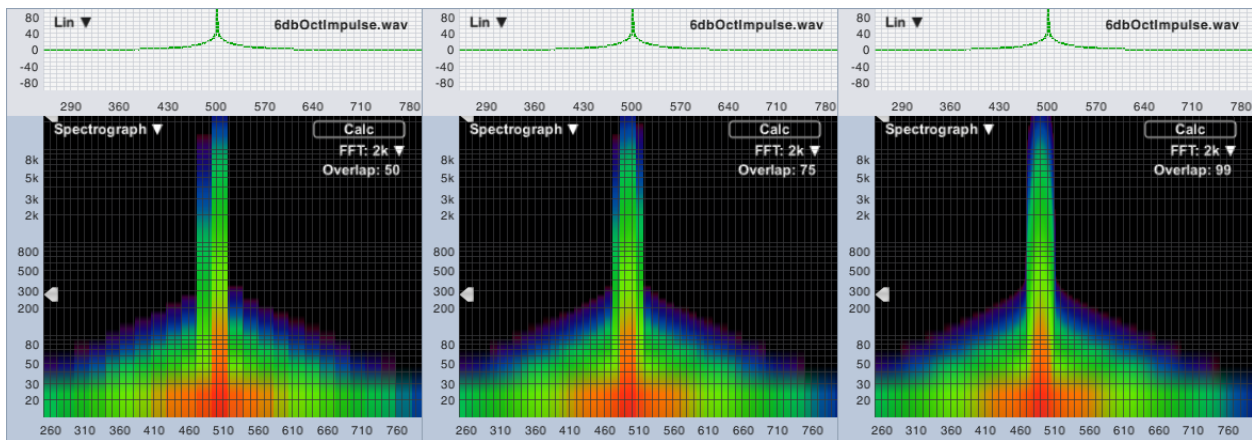
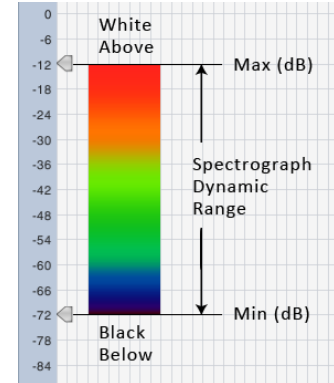


Figure 16: The 2K FFT example from Figure 14, shown with 50%, 75% and 99% overlap (left to right).

Spectrograph Dynamic Range

The dynamic range of the Spectrograph is controlled by two arrowhead shaped widgets that appear on the left edge of the Spectrograph chart. These controls are echoed on the Log IR and ETC displays where they relate directly to the decibel levels on the graph. The upper of the two widgets sets the maximum (Max) threshold; the lower one sets the minimum (Min). The Spectrograph scales its color spectrum between these two extremes. Any FFT bin whose magnitude exceeds the specified maximum is mapped to the color white. Values falling below the minimum are mapped to black.



Spectrograph Analysis of an Acoustical Impulse Response

Long-time Smaart users may actually recognize the impulse response shown in Figure 18. It was made using the file *room.wav* that was distributed with early versions of Smaart. A spectrograph of this file first appeared in the Smaart 1.0 user manual and it has popped up in innumerable places since then. The measurement was recorded on the stage of a 6000 seat performance space using an overhead PA cluster as the excitation source. It features a very prominent reverberant build-up phase and problematic late reflected energy arriving about 160 ms after the direct sound.

Figure 17: Spectrograph dynamic range and color mapping.

This is a good file to experiment with to see how changing the FFT size, Overlap and dynamic range can reveal different aspects of the IR. You can see that we've set the navigation pane graph type to ETC and moved *Time 0* to about 100 ms. The FFT size is 2K, overlap is 95% and dynamic range is -20 to -60 dB.

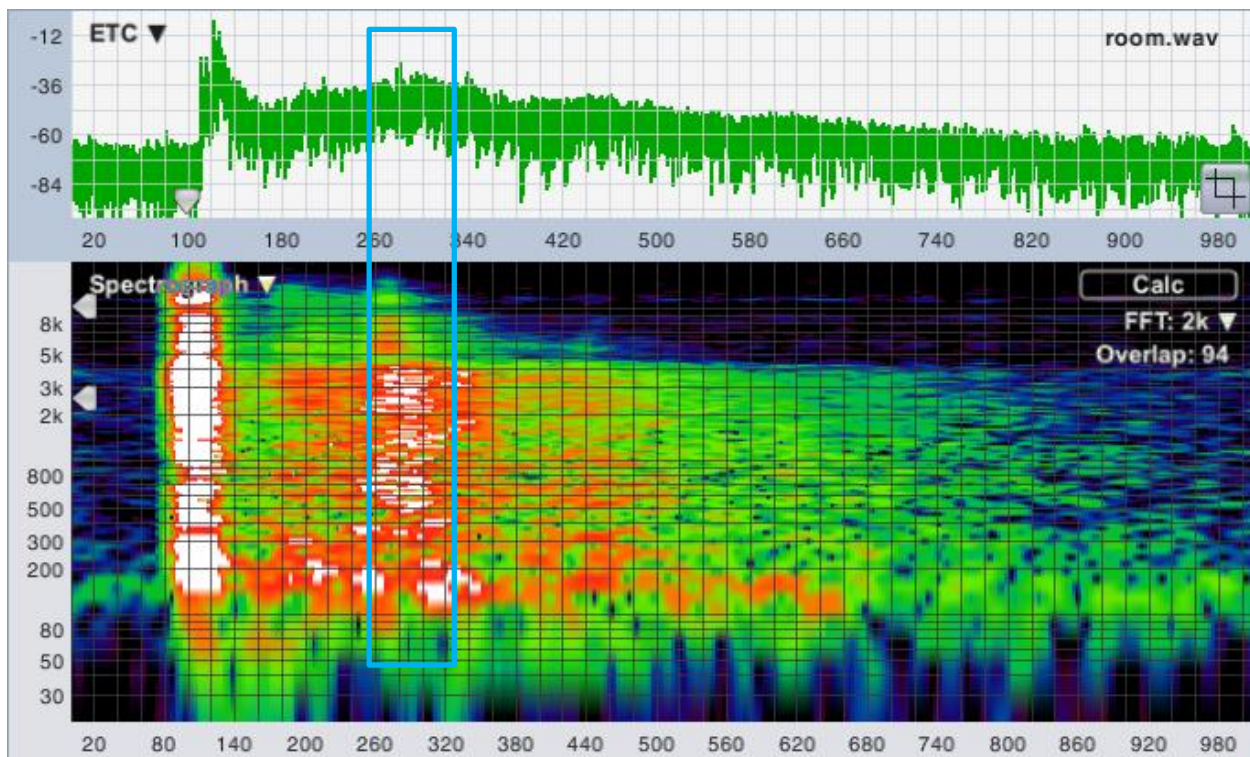


Figure 18: Broadband ETC and Spectrograph of a room impulse response showing a problematic back wall reflection. The spectrograph can be extremely useful for examining both the level and frequency content of features in the IR such as reverberant build-up and discrete reflections.

Speech Intelligibility Metrics

We have discussed the applicability of early-to-late energy ratios as good predictors of subjective speech intelligibility and RT60 as a somewhat rougher gauge. There also exist some purpose-built metrics created specifically for predicting speech intelligibility. Two of these, ALCons and STI, can be calculated from the impulse response of an acoustical system.

STI and STIPA

The Speech Transmission Index (STI) has emerged over the last decade or so as the go-to metric for objective estimation of speech intelligibility in acoustics. STI is a relative of the Articulation Index (AI), which is based on speech-to-noise ratios across a wide range of frequencies. But unlike AI and SII, STI works well for reverberant environments in addition to communication systems.

Rather than estimating intelligibility based on direct-to-reverberant or signal-to-noise ratios, STI starts with the concept of speech as a carrier frequency (sound from our vocal cords) that is modulated by very low frequency fluctuations as the speaker's mouth and tongue move and change shape to form words. Looking at Figure 19, a segment of actual human speech, it's not hard to see how someone might arrive at that conclusion.

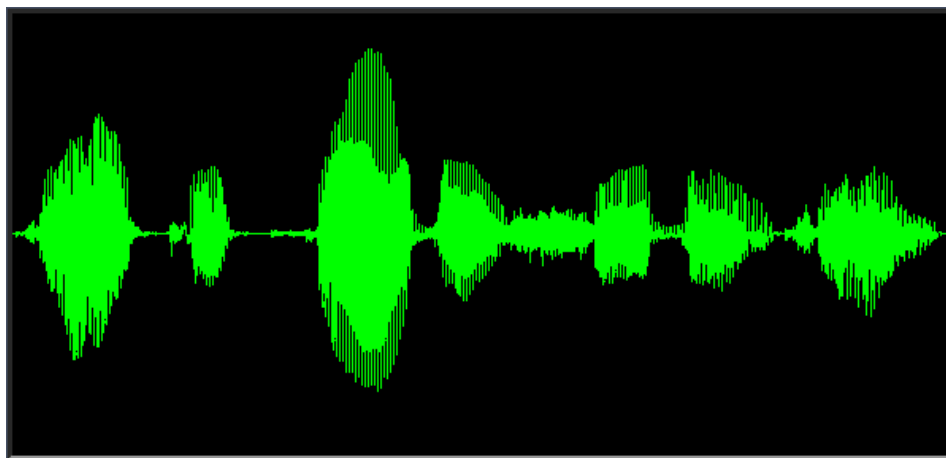


Figure 19: A recording of a male voice reciting, "Joe took father's shoe bench out."

The basic idea is that most of the information in speech is carried in these low-frequency modulations, and anything that reduces the depth of the modulations must negatively affect speech intelligibility. The real advantage of this approach is that STI ends up being sensitive to just about any factor that works to degrade speech intelligibility in a sound system and/or a room, including noise, excessive reverberation, distortion and audible echoes.

The basis for STI is the modulation transfer function (MTF) which quantifies the depth of modulation in the received signal relative to the transmitted signal at specified frequencies. The modulation transfer

function can be measured directly using specialized, “speech-like” test signals or calculated indirectly, from the impulse response or ETC of a system under test. In either case, it is measured over a range of seven octaves, from 125 Hz to 8 kHz, at fourteen modulation frequencies per band. The modulation frequencies range from 0.63 Hz to 12.5 Hz in 1/3-octave intervals.

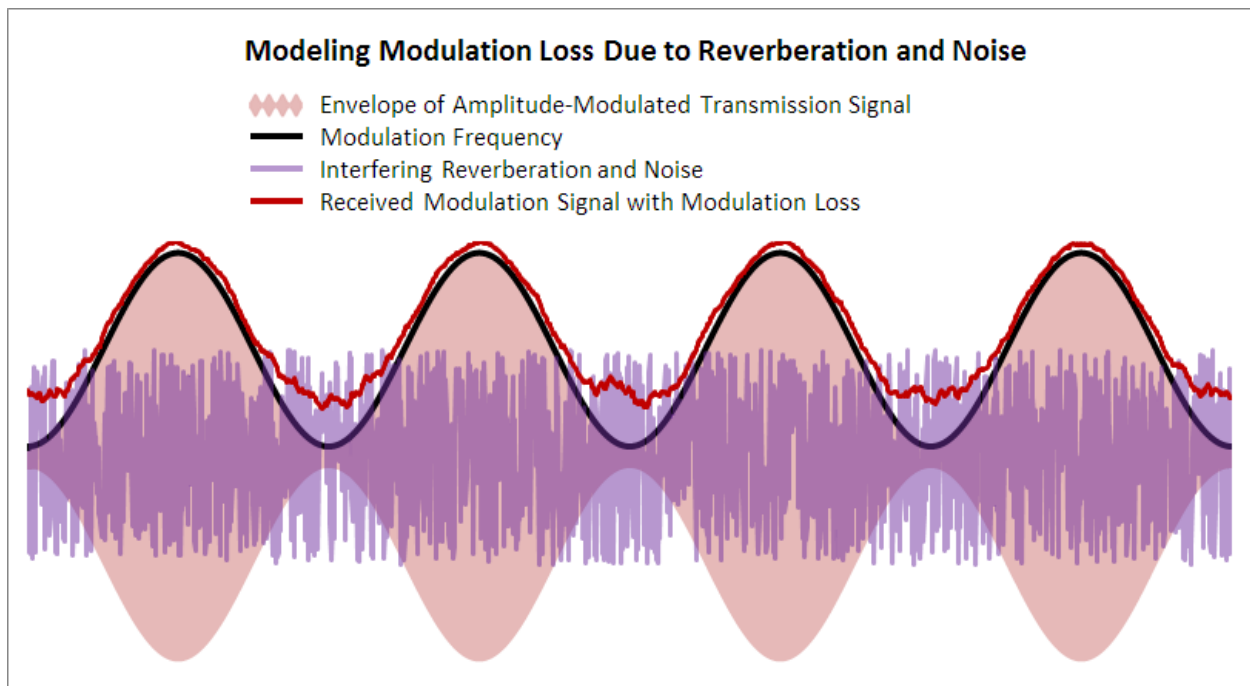


Figure 20: Estimating speech intelligibility based on modulation loss in an amplitude-modulated signal. The black line is the modulation signal for the transmitted signal. The red line is the modulation signal of the received transmission. The difference between the two tells you the loss of modulation through the transmission channel, due to interference from reverberation (in the case of an acoustical system) and noise.

The current IEC standard on STI (60268-16, Edition 4.0 2011-06) says that when measuring STI indirectly from an impulse response, “The duration of the impulse response shall not be less than half the reverberation time and at least 1,6 s to ensure a reliable calculation of the modulation indices for the lowest modulation frequency of 0,63 Hz.” However to us, this seems to overlook the fact that frequency bins in a DFT are linearly spaced, meaning two lowest bins are always an octave apart, whereas the modulation frequencies for STI are on 1/3-octave intervals. If the DFT time window is 1.6 seconds then the two lowest bins are at 0.63 and 1.25 Hz, whereas the first three STI modulation frequencies are at 0.63, 0.8 and 1.0 Hz.

This is to say that hitting all of the lowest STI modulation frequencies would require a measurement window significantly longer than 1.6 seconds. We experimented and found that 5 seconds was a sweet spot, producing data points very close to all of the STI modulation frequencies. That is the reason for the 5000 ms DFT size introduced in Smaart 7.5 for IR measurements. The DFT size in samples depends on the sampling rate you are using. For example at 48k sample rate, 5000 ms works out to 240,000 samples.

IEC 60268-16 qualifies IR-based STI measurement techniques for “noise-free” measurements, where a minimum signal-to-noise ratio of 20 dB in all seven octave bands is obtainable. The standard specifically qualifies MLS and swept sine test signals for used with indirect measurement techniques but also states that, “Theoretically, other mathematically deterministic pseudo-noise (random phase) signals could be employed.” Period matched pseudorandom noise in Smart fits that description and works well for STI.

There are a couple of potential advantages to using period-matched noise over sweeps. One is that it’s important to conduct STI measurements at sound levels that reflect actual use of the system under test, and the absolute level of sweep signals that start and stop can be ambiguous. Another is the issue of distortion products potentially piling up at the end of the time record FFT-based IR measurements made with logarithmic sweep signals. Those need to be dealt with in some fashion, either cropped off or windowed out, lest they corrupt the measurement. That isn’t an issue with period matched noise although you still need to take care not to overdrive the SUT.

General caveats to using STI are that it is sensitive to strongly fluctuating background noise levels, which can lead to overestimation of low intelligibility systems or underestimation of scores on the high end. When measuring in the presence of fluctuating background noise, at least three measurements should be taken and their results averaged to reduce measurement uncertainty. Also if the speech source and some prominent source of interfering background noise are widely separated, STI may underestimate intelligibility – human hearing can be smarter than machines about that kind of thing. STI is also sensitive to clipping or amplitude compression in the transmission channel, but in our case those would also violate the linear time-invariant system rule for transfer function measurements. So don’t do that.

Qualitative Thresholds

In terms of specific levels to look for, anything better than 0.80 STI is considered excellent. STI scores between 0.61 and 0.80 are “Good,” 0.45-0.60 is “Fair,” 0.35 to 0.44 is “Poor” and anything less than 0.35 is atrocious. Note that there are male and female versions of STI in the more recent versions of the standard. Any time an STI number is stated without specifying whether it’s for a male or female speaker, the assumption is that it’s the male version.

STIPA

One of the problems with direct measurement of STI is that it takes a lot of time to make a measurement. The modulation frequencies are so closely spaced that each one had to be measured separately. There are 98 modulation frequencies in all (14 x 7), and the full direct measurement takes about 15 minutes to perform as a result. STI for Public Address systems (STIPA) was developed to get around this problem.

STIPA is essentially the same measurement as STI, but uses a subset of its modulation frequencies; two per octave, for a total of 14. STIPA is typically measured directly, using a special test signal that excites all 14 frequencies at the same time, so that the measurement can be completed in a single pass. STIPA measurements can be completed in a few seconds and have been found to correlate very well with the more rigorous, full STI. STIPA is currently validated only for male speakers.

In Smaart of course, we measure STI *indirectly* from the impulse response and the full STI measurement takes no longer to perform than a typical direct measurement of STIPA. Smaart does provide figures for both STI and STIPA, however STIPA in our case is more properly termed STIPA(IR), since it is based on IR data rather than measured directly. It is provided for informational purposes, e.g., to facilitate comparison with readings from hand-held STIPA meters and is literally just a subset of the full STI measurement, calculated from exactly the same measurement data.

ALCons

ALCons, sometimes called %ALCons because it is stated as a percentage, stands for Articulation Loss of Consonants. Consonants are critical to speech intelligibility, as they are short in duration and tend to get lost more easily than vowel sounds that are voiced over a longer period of time.

ALCons was originally conceived as an estimate based on distance of the listener from a sound source, the directivity (Q) of the source, the volume and reverberation time of the room and the level of background noise. A longer form of the calculation suitable for direct measurement was later developed and found its way into various acoustical measurement platforms including Smaart. The long form uses direct vs reverberant energy – usually based on C20 or C10 – in place of Q, distance and volume. ALCons can be calculated for any frequency range but it is conventionally regarded as most meaningful in the octave band centered on 2 kHz, as this is where most of the energy in consonant sounds is found.

Advantages of ALCons include the fact that in its original form, it made estimation of intelligibility possible for sound system designs not yet installed, in rooms not yet been built, without the aid of acoustic modeling programs not yet available in the 1970s and '80s. In its long form, it could be directly measured in existing installations based on ETC or IR data produced by the TDS and MLS/FHT measurement systems that were prevalent before the computing power needed to calculate FFTs large enough for room acoustics work became widely accessible. Its main disadvantages are its reliance on assumptions that the sound field is statistically well behaved and without audible echoes, which may or may not be true.

Qualitative Thresholds

ALCons is upside down relative to other intelligibility metrics. Smaller numbers mean better scores. Anything less than 5% is considered excellent. Between 5-10% is “Good”, 10-15% is rated as “Fair” and anything less than 15% is problematic.

4: Measuring an Acoustical Impulse Response

You can make an impulse response measurement in Smaart by clicking a single button. Whether or not that measurement returns any useful information depends on decisions made beforehand. The process for measuring an acoustical impulse response can be summarized as follows:

- Selection of measurement technique and stimulus type.
- Selection of excitation source(s) and position(s).
- Selection of measurement position(s).
- Estimation of the reverberation time and background noise.
- Selection of measurement parameters (measurement length/duration, excitation level)
- Exciting the system and recording results

What are we measuring, and why?

Before you set out to make any acoustical measurement, it is always helpful to define your objectives clearly. The trip back to the site that you save may be your own. Obviously we want to measure the acoustical impulse response of a system under test (SUT) for some reason, but what exactly *is* the system? Is it a room? Is it a sound system? Is it a combination of a sound system its acoustical environment? What do you want to know about the system? What equipment and measurements will be needed to make sure you get the information you need?

If you want to measure the reverberation time of a room with an installed sound system, are you more interested in the room or the system? Consider that using a directional loudspeaker to excite the space may affect reverberation times in locations that are on axis with the speaker. Consider also that when using different speakers to measure from different points in the room, any significant differences between those speakers will show up in your measurement results. If the room is your target, the course of least resistance might be to bring in an omnidirectional loudspeaker specifically designed for acoustical measurement. On the other hand, if your objective is to measure the performance of a loudspeaker system installed in a room, you might be more concerned with early-to-late energy ratios and speech intelligibility metrics than reverberation time of the room, exclusive of the sound system.

Direct vs Indirect IR measurement

There are two basic ways to measure an impulse response in Smaart; direct or indirect. Or you could say there are three possible methods, because the indirect IR measurement method that Smaart employs can be used as a deterministic or non-deterministic measurement technique. Let's start with the simplest and easiest to understand – old fashioned direct IR measurement – and work our way up from there. This should not be construed as being in any sort of order of preference. In fact, we're starting with what is usually the least preferred method, but they all have their selling points.

Direct IR Measurement Using an Impulsive Stimulus

The most intuitive way to measure the impulse response of a system would be to use an impulsive stimulus of some kind and simply record what happens. And in fact, people have been doing just that for decades. The advantages are that you do not need a sound system or even a measurement system to

perform the measurement. All you really need is some way to make a loud bang and some way to record it. The main problem with this approach – other than the fact that it doesn't tell you anything about an installed sound system (if applicable) – is the scarcity of really good impulsive stimulus sources.

An ideal impulsive stimulus would be a perfectly instantaneous, perfectly omnidirectional burst of energy, containing equal proportions of energy at all audible frequencies. In the time domain, it would appear as a single vertical spike no more than one sample in width. In the frequency domain, it would produce a perfectly flat magnitude and phase trace. When evaluating the response of your system under test (SUT) to this ideal stimulus, you could then confidently assume that anything you saw in the IR that wasn't an instantaneous spike, or anything in the frequency domain that wasn't a flat line, must be the response of your system.

If you loaded the wave file *1samplePulse.wav* to look at the frequency response of Smaart's bandpass filters in the previous section, you were actually performing a direct IR measurement on the filters using an ideal impulse. Unfortunately, stimulus signals like that don't exist in the physical world. When we need to measure the impulse response of an acoustical system directly, we end up using stimulus sources that are less than ideal. Blank pistols and balloon pops are common sources. Signal cannon, spark gaps, fireworks and even spot welders have been used. The problem with all of these is that their spectral content is not uniform, their envelopes are not instantaneous, they may not really be as omnidirectional as one might guess, and all of these factors will vary to some extent from one measurement to the next. This introduces uncertainty from the start as to which part of the completed measurement is stimulus and which is response. It also limits the repeatability of test results. For this reason, systems such as Smaart that indirectly infer the response of a system to an ideal impulse have become more the tools of choice these days.

Indirect (Dual Channel) IR Measurement

Indirect impulse response measurements are made using dual-channel measurement techniques that mathematically estimate the response of an SUT using continuous or periodic test signals. Three of the four indirect IR measurement methods we can name require specialized test signals such as sweeps (Time Delay Spectrometry and Direct Convolution) or specialized noise (Hadamard Transform/MLS).

The dual-channel transfer function method that Smaart uses for indirect IR measurement also works best using period matched test signals, but unlike the other three it can also produce very acceptable results using random test signals, provided that both the reference and measurement signals are captured. Transfer function-based IR measurement systems work by calculating the frequency-domain transfer function of a system under test (SUT) from the Fourier Transforms of two signals – the signal going into a system and the output of the system in response to this input – and then transforming the result back into the time domain using an inverse Fourier transform (IFT).

Remember that perfect impulse that we were lamenting didn't exist in the real world for direct IR measurements? Well, that happens to be what you get if you take the IFT of the transfer function of two identical signals. So it follows that when we take the transfer function of a stimulus signal and the SUT's response to it, we theoretically should get something very much like its response to an ideal impulse.

And in fact that's pretty much what happens in practice when you use a period-matched excitation signal. When you use this same technique with effectively random signals you also get a lot of extra noise, but repeating the measurement several times and averaging the results generally takes care of that, and Smart makes this easy to do.

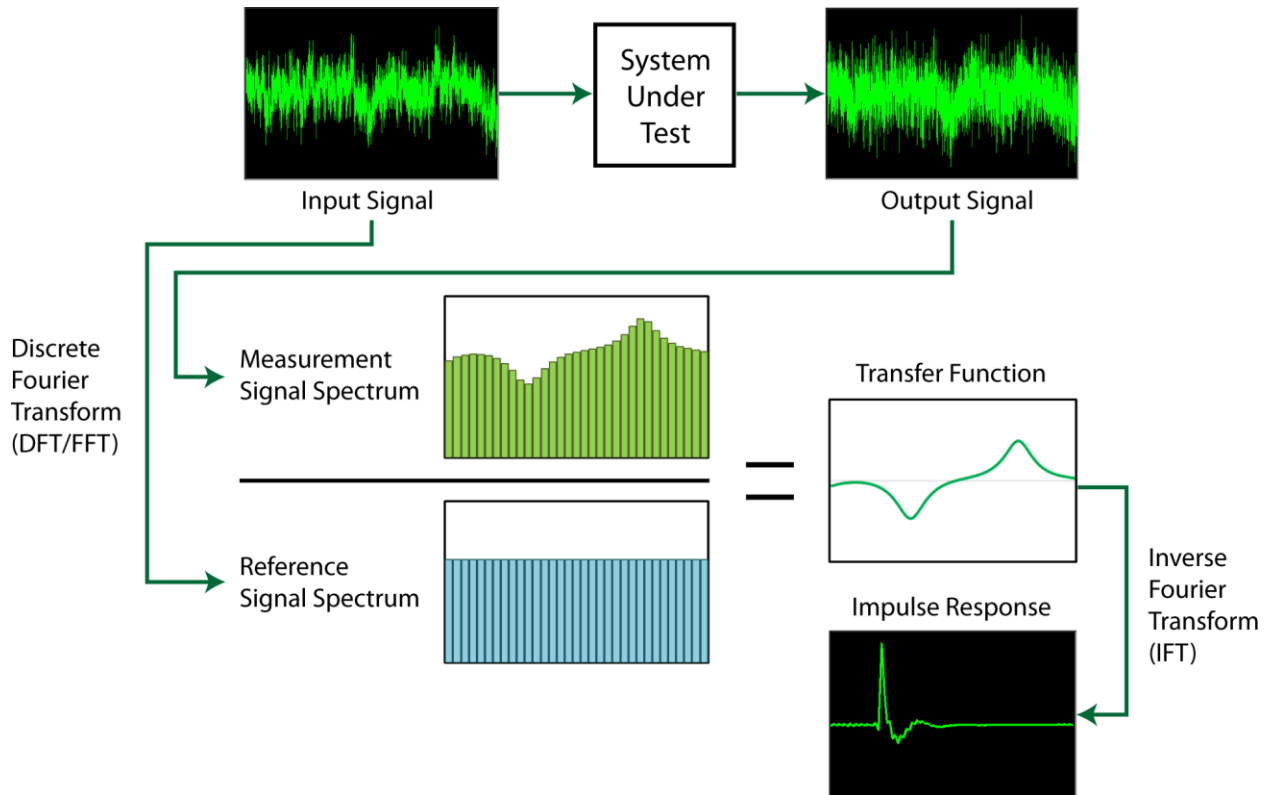


Figure 21: Block diagram of a Dual-FFT transfer function IR measurement.

Dual Channel IR measurement Using Period-Matched Signals

In a way, the discrete Fourier transform (DFT or FFT – all FFTs are DFTs, but not all DFTs are fast) is kind of a dirty mathematical trick. Fourier transforms of all types theoretically work only with signals of infinite length but the DFT gets around this by pretending that a finite chunk of signal being analyzed is really just one instance of an infinitely repeating series of chunks that look exactly like it.

The best way to get around this inherent assumption of this cyclicity in DFT/FFT analysis is to feed the DFT what it really wants to eat: a test signal that either fits completely within the measurement time window or cycles with periodicity equal to the length of the DFT time constant. Signals that meet these criteria can produce deterministic, highly repeatable measurements in a fraction of the time it takes to get comparable results using random signals.

When using matched periodic signals:

- No data window is required.

- Delay compensation is not a critical requirement.
- Little or no averaging is necessary – in fact, averaging may actually increase noise artifacts.
- Measurement time constants can be kept to reasonable lengths.
- Small time variances become less of a concern.
- Subjectivity in selecting measurement parameters is reduced.
- The measurement system doesn't necessarily need to be connected to the system under test. (When using a known test signal, the measurement system and the SUT can get their stimulus/reference signals from two different sources and the measurement will still work. You won't get an accurate propagation delay time without an audio feed from the signal source being used to excite the SUT, but if you don't really need delay times this can be a very handy option.)

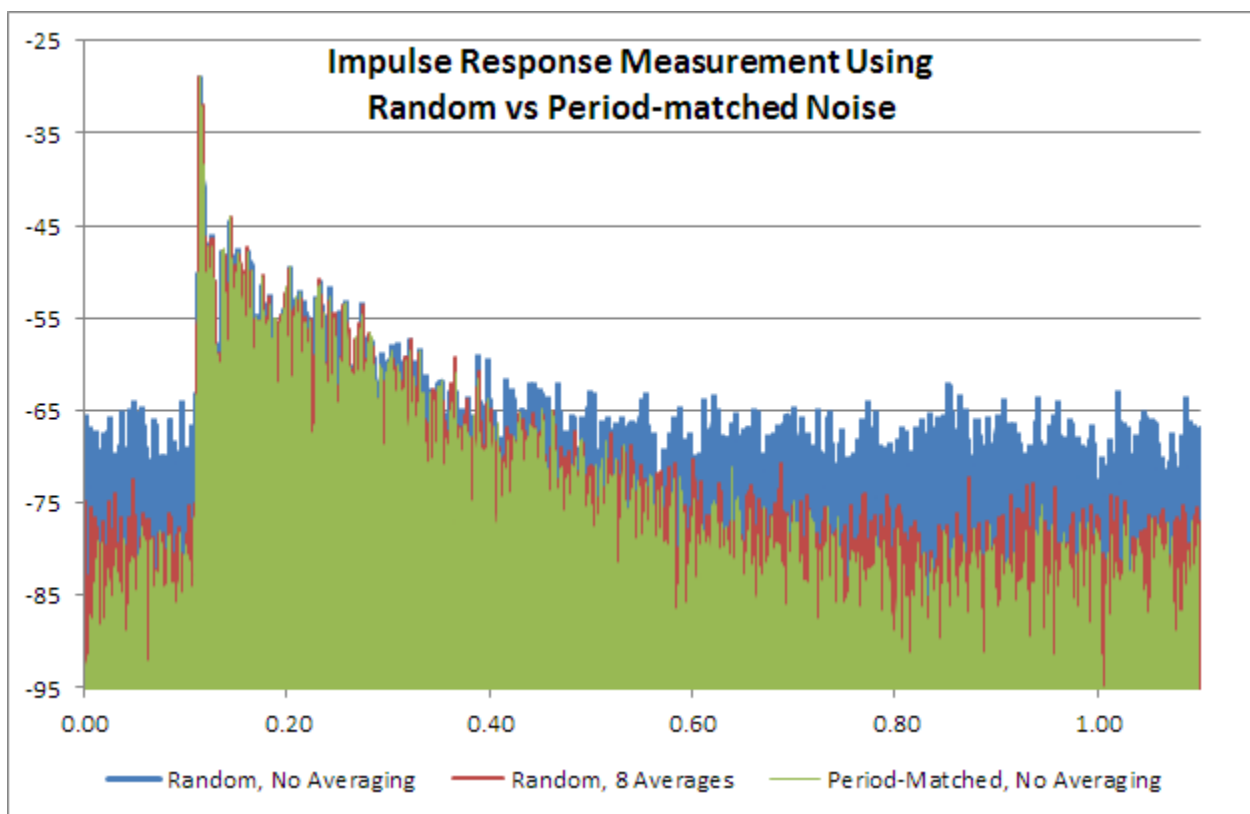


Figure 22: Three indirect IR measurements of the same room, taken from the same microphone position using effectively random noise vs period-matched pseudorandom noise. The period-matched noise measurement (in Green) takes the same amount of time as the unaveraged random measurement (in Blue) but has much better dynamic range. By repeating the random noise measurement eight times and averaging the results (the measurement in Red) we can greatly improve its signal-to-noise ratio, however the measurement takes eight times longer to perform.

Logarithmic Sweeps

Logarithmic sweeps are called *Pink Sweeps* in Smart. When you select this signal type in the signal generator, Smart drops the IR data window without being told. A data window in conjunction with a

sweep signal would act as a filter on its frequency content, since each frequency appears at only a single point in time during the measurement.

Sweeps can be used as a circular or aperiodic signal source. If the *Triggered by impulse response* option is enabled in Smaart's signal generator, the sweep signal is triggered by starting an IR measurement. When you kick off the measurement, Smaart will insert a short period of silence before the sweep in case there's any lag in starting the recording device, then run the sweep and insert another period of silence afterward to let the SUT ring out. If the *Triggered by impulse response* option is unchecked, the sweep runs continuously when the generator is turned on. In this case you would start the generator before starting the measurement as you would with other test signals.

A peculiarity of dual-FFT-based IR measurements made with logarithmic sweeps is that distortion products in the excitation loudspeaker/SUT are washed out of the IR and show up as "pre-arrivals". Because the DFT is a circular function, these typically end up wrapped around *past the beginning* of the measurement and pile up near the end of the time record. The practical implication is that you may need to make the measurement time window a little larger than you would for a matched noise measurement, to ensure these artifacts do not intrude on the reverberant decay slope.

As regards STI measurement, IEC 60628 states that, "When using a sine sweep technique, the distortion components that are inherent within the method shall be edited out or removed from the IR before calculation of the STI can be undertaken." It is our opinion however, that this requirement argues for the use of period-matched noise rather than sweeps for STI measurement, since the masking effect of high levels of distortion in lower-spec announcement systems can significantly affect speech intelligibility, and properly should be included in the measurement.

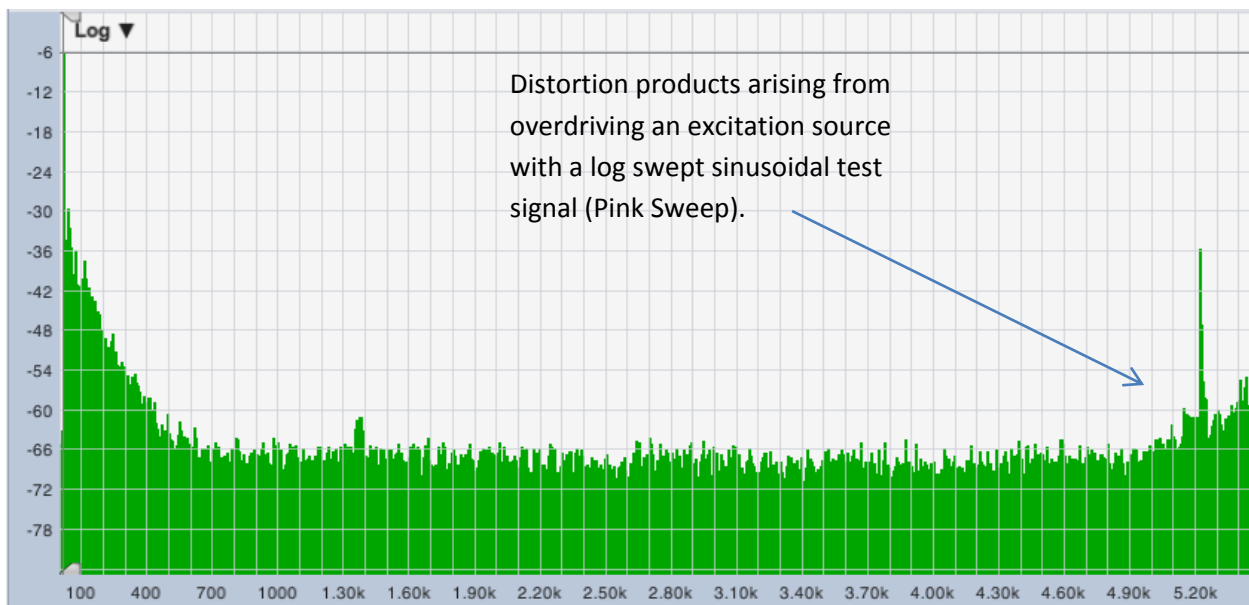


Figure 23: An impulse response measured using a log swept sine signal (Pink Sweep), showing harmonic distortion products from the excitation loudspeaker piled up at the end of the time record. In this case, the speaker being used to excite the room was overdriven and the distortion component was quite significant.

Dual Channel IR measurement Using Random Stimulus Signals

An excitation signal that is not completely contained within or, if continuous, has its periodicity precisely matched to the time constant of a discrete Fourier transform being used for analysis is effectively random as far as the DFT is concerned. In Smaart's signal generator, the *Random* pink noise option or any pseudorandom cycle length with periodicity longer than the FFT size used are effectively random. (Periods shorter than the FFT size should never be used because they won't contain energy at all FFT bin frequencies.) Other examples would include music or noise signals with arbitrary periodicity from sources other than Smaart.

Perhaps the best argument in favor of using random signals for IR measurement is, because you can. If you want to make a measurement with music instead of noise, you can. If it's easier to generate pink noise from a mixing board or in a processor than it would be to inject a test signal into the signal chain from Smaart, that will work. The only absolute requirements are that the measurement system needs to capture an exact copy of the signal going into the SUT and that signal must contain enough energy at all frequencies of interest to you to make a solid measurement.

The main caveats associated with random stimulus signals are a relatively high level of noise, meaning that you have to measure over a longer period than you would with a purpose-built signal to get comparable results. It is left up to the operator to decide how much averaging or how long a time window to use and the actual dynamic range of the SUT is ambiguous. This can be a critical factor in speech intelligibility measurements.

Reducing Noise in IR Measurements Made Using Random Stimulus Signals

There are three basic things you can do to improve the dynamic range of measurements made using random test signals. The first is to delay the reference signal to match the timing of the measurement signal, so that the data windows line up. You should always do this when measuring with random signals. The second is to evaluate the system over a longer period of time by increasing the DFT size or by averaging multiple measurements (or both). The third is to simply measure louder, which also applies to deterministic and direct IR measurements – in that case you're increasing the signal-to-noise ratio of the measurement by increasing the level of the actual signal, rather than statistically.

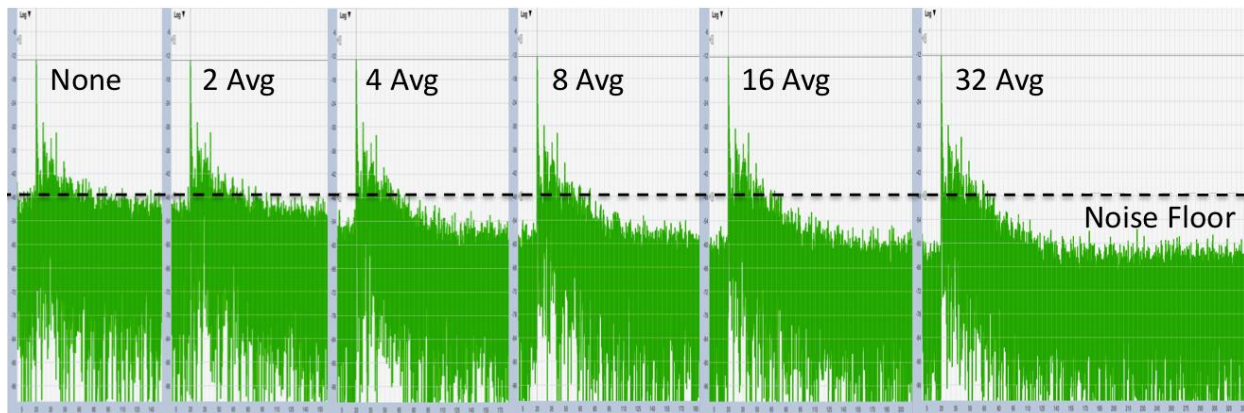


Figure 24: The effect of averaging on an IR measurement made using random stimulus signals. In theory, each doubling of the number of averages increases signal-to noise ratio by 3 dB.

Averaging works by inducing regression to the mean in random components of the IR (that is to say, the noisy part). Let's say you take a signal – any signal, maybe an impulse response – and mix it with random noise. Obviously, you get a noisy signal. There is no way to tell just by looking which part was signal and which was noise. But if you take several copies of the same signal and mix each one with *different* noise, then average all of them together; the noise component of each noisy signal (being random, and different in each case) should start to average toward zero – the theoretical arithmetic mean for random audio noise – while the signal parts (being the same in every case) should average out to themselves.

Of course, all of this depends on the assumption that the signal part of the signal is the same in every case. When working indoors that should generally be a safe assumption. After all, we are working with what we assume to be linear, time-invariant systems in a fairly controlled environment where the worst that could probably happen from one pass to the next is a blast of hot or cold air from an HVAC system causing a slight change to the speed of sound. It might be a larger concern if you needed to make an IR measurement outdoors under windy conditions for some reason. In any scenario where there might be a possibility of any significant time variance during the measurement period, you would probably be better off increasing the measurement time window and/or using a period-matched stimulus signal rather than upping the number of averages.

In theory, averaging two IR measurements or doubling the FFT size used for a single IR measurement should improve signal-to-noise ratio of the measurement by 3 dB. Note that both result in doubling the measurement time, which is really the key to the whole thing. Each additional doubling (2, 4, 8, 16...) of the measurement duration should theoretically get you another 3 dB, although in practice you might reach a point of diminishing returns at some point.

Selecting Excitation Sources and Positions

Excitation source positions should be places that sound would normally emanate from when the system under test is in service. If the loudspeakers you are using to excite the room are the places that sound normally comes from then you've got that part covered. Otherwise an omnidirectional sound source of

some kind should be placed on the stage, podium, lectern, pulpit or whatever location(s) that would best simulate normal use of the room/system, and at an appropriate height.

Minimum Distance from Sound Sources

Another general requirement for measurement positions used for room IR measurements is that they need to be located far enough away from loudspeakers or other sound sources being used to excite the room to ensure that the measurement is not unduly dominated by direct sound. ISO 3382-2 provides the formula shown in Eq. 1 for calculating the minimum distance (d_{\min}) for any measurement position from an excitation source.

$$d_{\min} = 2\sqrt{V/c\hat{T}} \quad (\text{Eq. 1})$$

where

V is the volume of the room in cubic meters

c is the speed of sound in meters/sec

\hat{T} is estimated reverberation time in seconds

But we did promise to keep the math to a minimum, so Figure 25 provides a handy graph of this relationship to go with it.

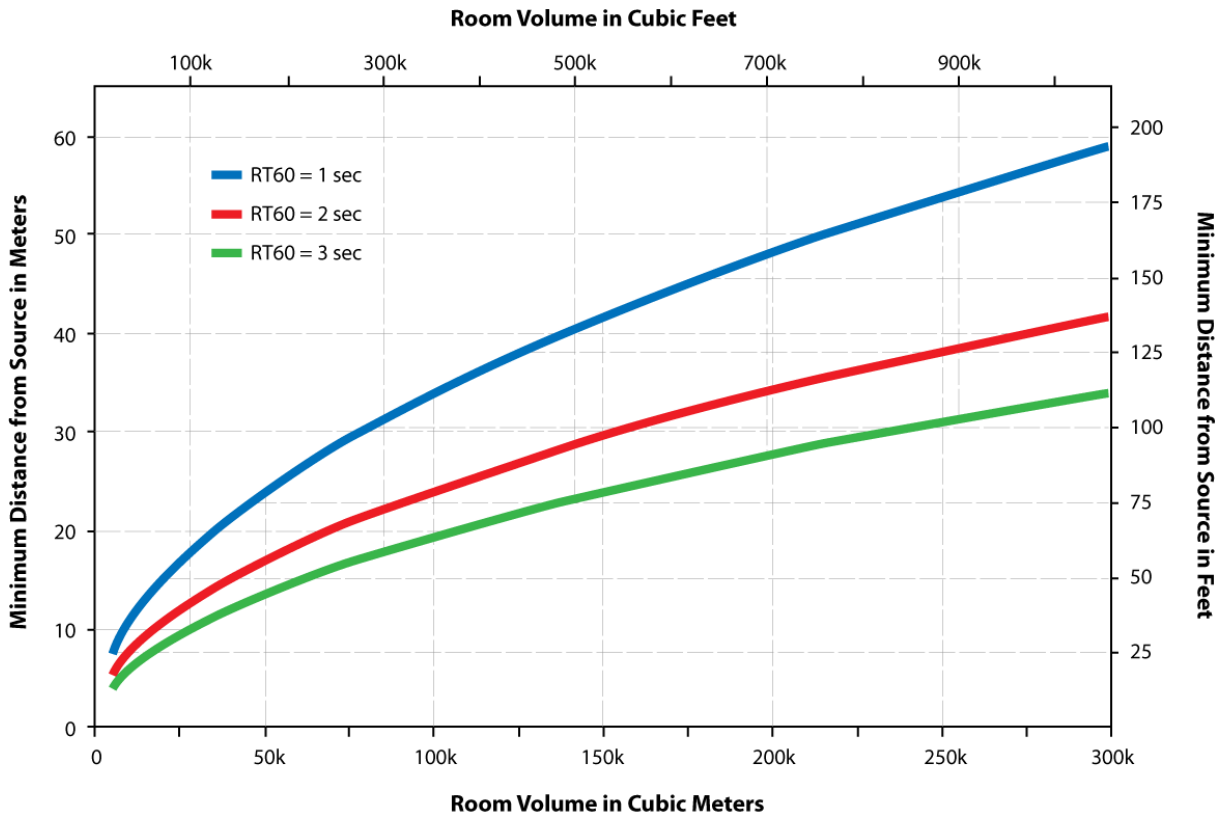


Figure 25: Minimum distance to any measurement position from the excitation source (e.g., a loudspeaker) used for reverberation time measurements. The minimum distance is a function of room volume, estimated reverberation time and the speed of sound, as described by Eq. 1 above. This example uses speed of sound at 20° C (68° F); i.e., 343.6 meters/sec or 1127.4 fps.

Directional Loudspeakers and Reverberation Time

For the specific purposes of reverberation time measurement, a potential complicating factor may arise if an installed sound system is to be used to excite the room. Impulse response measurements made with directional loudspeakers typically have higher direct to reverberant ratios than IR measurements made by other methods, which could affect reverberation times in the upper octaves. ISO 3382-1 unequivocally states that “the sound source shall be as close to omnidirectional as possible” and provides criteria for assessing the omnidirectionality of a prospective source. ISO 3382-2 specifies measurement procedures for three levels of accuracy in reverberation time measurements: Survey (quick and dirty), Engineering (pretty good) and Precision (very good). For the Precision method, the requirements for the excitation source are identical to those specified in 3382-1 but 3382-2 goes on to say that “For the survey and engineering measurements, there are no specific requirements for the directivity.”

Clearly, the spirit of the law is that omnidirectional sources are preferred for reverberation time measurement, but the standard does leave a little wiggle room when measuring in “ordinary rooms” (as opposed to formal musical performance spaces).

Differences between reverberation times measured using directional versus omnidirectional speakers will tend to be most significant in the frequency ranges covered by the most directional loudspeaker elements. As such, these are likely to be most problematic in the 1 kHz through 4 kHz octave bands (4 kHz being the highest octave band normally evaluated for reverberation time). An informal survey of IR measurements we happened to have on hand that were made using directional PA speakers revealed reverb times as much as half a second shorter in the upper octaves than you might expect to see with an omnidirectional source. In practice it may be possible to compensate for this difference by adjusting the upper end of the evaluation range to some point below the standard -5 dB start point, but of course that represents a departure from standard practice and introduces an element of subjectivity that the standardized procedure was designed to eliminate.

Given a choice between a measurement made with a directional speaker or not making a measurement at all, a less than ideal measurement is usually better than none. However if any potential errors or subjectivity in evaluating reverberation time are a source of great concern for you, it might be necessary to bring in an omnidirectional measurement speaker and do it by the book – or at least record a few balloon pops, just to have a second opinion.

Reverberation times aside, it is worth mentioning that for most *other* purposes, IR measurements made using an installed sound system that is actually used for amplified performances in the space you are measuring will be more representative of actual use of the system than measurements made by any other means. In some cases, in order to get everything you need, you might need to make one set of measurements using an omnidirectional source positioned on the stage, another using the installed sound system and perhaps even a third using the house paging system, to estimate its intelligibility.

Selection of Measurement Positions

The first and most obvious rule for selecting measurement positions is that you generally want to measure from places where one would expect to find listeners when the system under test is in service. If a tree falls in the forest and no one is there to hear it, who really cares if it makes a sound? You might also want to give special attention to any areas where you think there could be problems. Other than that, it is kind of like taking an opinion poll. If we measure from a single position, we have one “opinion” of what the room sounds like. If we sample from several different locations, we might reasonably expect to see some consensus emerge as to the most common characteristics of the system response, and for position-dependent differences to begin to average out. The more measurement positions, the lower the theoretical margin of error, assuming the positions are chosen so as to be statistically valid.

For the Survey method in ISO 3382-2, a single stimulus source location is measured from at least two measurement locations, providing a theoretical margin of error of $\pm 10\%$ for octave bands. The Engineering method calls for at least two stimulus source positions and six independent source-microphone combinations for a nominal accuracy $\pm 5\%$ for octave bands or $\pm 10\%$ in 1/3-octave bands. The precision method calls for 12 independent source-microphone combinations using at least two different stimulus source locations and reduces measurement uncertainty to no more than $\pm 2.5\%$ for octave bands and $\pm 5\%$ for 1/3-octave bands.

ISO 3382-2 specifies that all measurement positions should be at least one-half wavelength apart and at least one-quarter wavelength from any reflecting surface including the floor. For example, if we wanted to measure as low as the 125 Hz octave band, the lower band edge is at ~90 Hz. At 68° F (20° C), the speed of sound in air is 1127.4 feet per second (343.6 mps) and so one wavelength at 90 Hz would be about 12.5 ft (3.8 m). From that, we could conclude that no two mic positions should be less than 6.25 ft (1.9 m) apart and all microphones should be at least 3.13 ft (0.95 m) above the floor and at least that far from any wall or other reflecting surface. For the 63 Hz band you would need to double those distances.

Of course ISO 3382-1 applies specifically to measurement of reverberation time in rooms, so what about acoustical measurements made for other purposes. Two additional standards we could look at as a guide to microphone placement are ANSI S1.2, Criteria for Evaluating Room Noise, and SMPTE 202M, the current standard for calibrating cinema sound systems. ANSI S1.2 has this to say about measurement positions:

“Sound measurements for rating room noise under this standard shall be made at locations that are near the average normal standing or seated height of human ears in the space: 5’-6” for standing and 4’-0” for seated adults—3’-6” standing and 2’-6” for seated children. The microphone shall be no closer than 2’-0” from any sound reflecting surface or 4’-0” from the intersection of two intersecting reflecting surfaces, or 8’-0” from the intersection of three intersecting reflecting surfaces.”

SMPTE 202M recommends that microphones be placed:

“In indoor theaters, at position S [...] and position R [...] should it exist, and at a sufficient number of other positions to reduce the standard deviation of measured position-to-position response to less than 3 dB, which will typically be achieved with four positions. [...] It is recommended that measurements be made at a normal seated ear height between 1.0 m and 1.2 m (3.3 ft and 4.0 ft), but not closer than 150 mm (6 in) from the top of a seat, and not closer than 1.5 m (4.9 ft) to any wall and 5.0 m (16.4 ft) from the loudspeaker(s).”

(Position “S” generally works out to be a little left or right of the approximate center of the room on the main floor. Position R is for balconies.)

We can see that these are all in general agreement (depending to some extent on frequency) even though one talking about reverberation time, another is for background noise surveys, and the third is for RTA measurements of cinema sound system. They are probably also not out of line with positions you would intuitively choose for frequency-domain transfer function measurements of a sound system.

Selecting Measurement Parameters

Once you have done all the groundwork of determining your source and measurement positions and figuring out which measurement technique to use, the part of the measurement procedure that directly involves Smaart is actually pretty easy. Basically, you just need to select your measurement parameters, turn on the signal generator (or other stimulus signal source) and kick off the measurement. The two main things you need to concern yourself with at that point are the stimulus level and the measurement duration, which will be some combination of the FFT size and the number of averages.

Input source

If you already have one or more Transfer Function measurements configured and will be using one of those to make your measurement, use the Group and TF Pair selectors shown in to select the one you want. To create a new TF pair, click the little hammer and wrench button next to the *Group* selector to open the Measurement Config window, then click the *New TF Measurement* button (see Figure 26). This pops up another dialog where you can select the input device and channels that you want to use and give the measurement pair a name. If you are unfamiliar with how to set up your measurement system for transfer function and dual-channel measurements, Appendix E has some example setup diagrams.

Excitation Level

The rule of thumb for setting the excitation level for IR measurements is that you would like to be able to get at least 40-50 dB above the background noise level. In reverberation time measurements, we evaluate reverberant decay over a range starting 5 dB down from the arrival of direct sound (normally the highest peak in the IR) and extending down another 20 or 30 dB from the start point.

A 30 dB range is preferred but 20 dB is OK if you can't get 30. Either way, the lower end of the range needs to be at least 10 dB above the noise floor of the IR measurement. When you add that all up, you're looking for a minimum of 45 dB of dynamic range for a 30 dB evaluation range, and at least 35 dB for a 20 dB range – and that's in a perfect world, with no noise artifacts from the measurement process itself. In the real world, adding another 5 to 10 dB on top of that would be a definite nice-to-have (unless of course that would drive the system into distortion or blow something up).

To figure out how loud you need to be, you can simply measure the background noise level. We are looking for a relative relationship so you don't even really need to be calibrated for SPL (unless perhaps you plan on doing an STI measurement). Just set the sound level meter in Smaart to Slow SPL and watch the meter for ten or twenty seconds with no output signal running to get a feel for the baseline noise level, then start the signal generator at a low level and gradually increase the gain until you reach the target excitation level (or as close to it as you can reasonably get).

Input Levels

Once you have your output levels nailed down, adjust your input levels (by whatever means available) until both the measurement and reference signal levels (labeled M and R on the control strip) are roughly even and running at a reasonable level. The yellow segment of the meters in Smaart runs from -12 dB to -6 dB full scale and that's the target zone. The meters are peak reading and we hard limit peaks for noise signals in Smaart, but you have to also allow for fluctuations due to background noise in acoustical measurements and if you use noise from another source you may see wider variations in the peak levels. With sinusoidal sweeps you can run the levels a little higher if you like, due to the lower crest factor of the signal but you always want to keep the levels out of the red. If you are doing a single channel measurement, you will probably have to waste a few balloons or fire off a few blank cartridges while adjusting the measurement channel gain to get a good

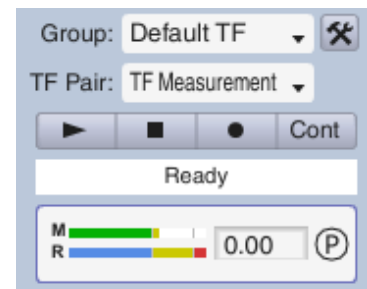


Figure 26: The measurement signal level (M) is running a comfortable level. The reference (R) channel is clipping.

solid signal level with no clipping on the input level meter.

Measurement Duration

If you are only concerned with measuring delay times then a good rule of thumb for how long the measurement needs to be is 3 times the longest delay time that you want to measure. If you want to measure reverberation time and early to late energy ratios, then the 60 dB decay time (RT60) of the system is a good target. This is kind of a functional requirement for period-matched dual-channel measurements but it's also a pretty good practical target regardless of how you are measuring. Ideally you would like to measure 30 dB of reverberant decay and the lower end of the evaluation range should be at least 10 dB above the noise floor, so that's 40 dB which of course is two thirds of 60. By the time you factor in propagation delay, early decay and maybe enough of a noise tail to see the dynamic range of the measurement, chances are you have eaten well into the remaining third.

Of course, both these rules require either knowing the delay time or RT60 before you've measured them. That typically means you have to guess, then measure, then possibly adjust your guess and measure again. For delay times you can use the distance to the source divided by the speed of sound as a starting point. For "guesstimating" purposes, you can use 1130 feet or 345 meters per second at typical room temperatures – the speed of sound increases with temperature so if it is very hot where you are working you might adjust your estimate upward a little, or downward if it's cold.

For reverberation times, one to two seconds should at least get you in the ballpark for most theaters and auditoriums. Stadiums and other large structures can have much longer reverb times. There's never any harm in measuring over too long a period, so you may want to err to the high side. If you make a preliminary measurement and you are happy with the results you might even be done. If not, you can adjust accordingly and measure again. Note that as a rule, lower frequencies tend to decay more slowly than highs, meaning that the limiting factor may be the reverberation times in the lowest octaves that your stimulus source can excite. So be sure to check the lower bands when estimating reverb times.

FFT Size (Time window)

For dual-channel measurements the duration of the measurement is determined by the FFT time constant – that is, the time required to record enough samples for a given FFT size at whatever sampling rate you are using. In Impulse mode, Smaart gives you the time constant in milliseconds, along with the frame size in samples for each available FFT/DFT size.

Averaging and Overlap

Averaging, as we discussed earlier in this section, is primarily something you concern yourself with when using effectively random stimulus signals. With random or effectively random stimulus signals, deciding

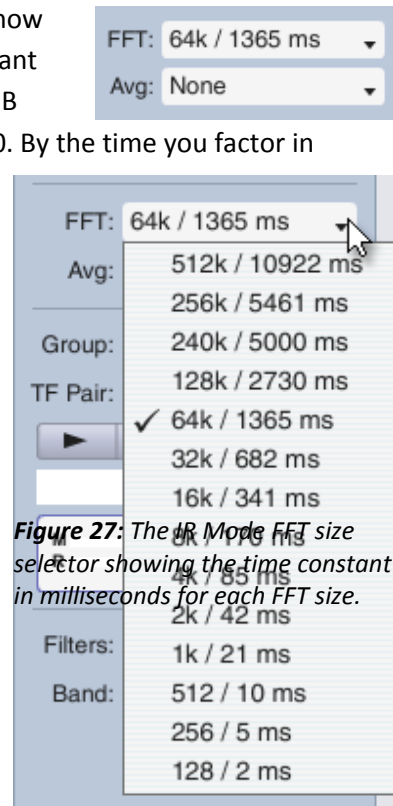


Figure 27: The IR Mode FFT size selector showing the time constant in milliseconds for each FFT size.

how much is averaging is enough is kind of judgment call but typical settings are in the 4-16 range. In very noisy environments, you may want to use a larger value and/or consider using a period matched signal. When measuring with period-matched noise or sweeps, averaging is normally set to “None” or 2, although it is still possible that a higher setting could prove helpful if measuring in an extremely noisy environment.

Another factor that affects how averaging works is the *Overlap* setting found in Impulse Response options (*Options menu > Impulse Response*). When overlap is set to 0% each FFT is calculated from unique data, giving you the maximum amount of noise reduction that you can get from a given number of averages. When you set the measurement Overlap to a non-zero value then successive FFTs share some data in common (see Figure 15 in the previous section) – remember that measurement overlap and Spectrograph overlap are two different things, but the principle is the same. If measurement overlap is set to 50%, it only takes a little longer to record 16 averages than it would for 8 at 0% overlap. You don’t get the full benefit of averaging 16 unique FFTs in that case and processing time increases but you should see at least a little better signal-to-noise than you would get 8 with some net time savings.

Delay Compensation

When making IR measurements with random signal sources you will get much better results if you compensate for the delay time through the system under test. So plan on making the measurement twice if you don’t already know the delay time; once to find the delay and a second time for a keeper. In Smaart 7.4 or higher, the button labeled “@” (for Peak) that appears next to the input level meters in IR mode (see Figure 26) sets the reference signal delay to the highest peak in the impulse response.

Pushing the Button and Making the Measurement

Having nailed down which measurement technique will get you the results you need, chosen your excitation sources and measurement positions, set your input and output levels, and selected your FFT length and number of averages (if applicable), all that’s left to do is push the button(s). For a dual-channel measurement, start your excitation signal (unless you’re using a triggered sweep in Smaart, in which case it will start by itself) and click the start (▶) button in the control strip on the right side of the main window (see Figure 26). Smaart will take it from there and display the measurement results when it has finished.

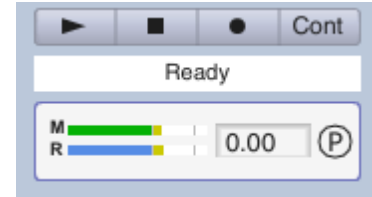
For a single-channel (direct IR) measurement, click the record (●) button, then click the start (▶) button, pop your balloon or fire your blank pistol (or whatever), give the system a few seconds to ring out, and then click the stop button (■) to end the recording and display the results.

Saving Your Work

When you measure an IR in Smaart does not automatically assume that you want to save it. Sometimes it takes a few tries to get it right and we don’t need a lot of old baggage piling up in the process. When you have a measurement that you are happy with and want to save, click in the *File* menu and select *Save Impulse Response* then select the directory where you want to store the file and give it a name. Note that if you have cropped the file for display purposes using the Crop function, only the displayed portion of the time record is written to file.

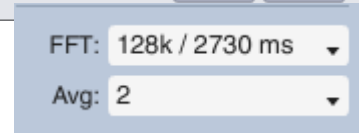
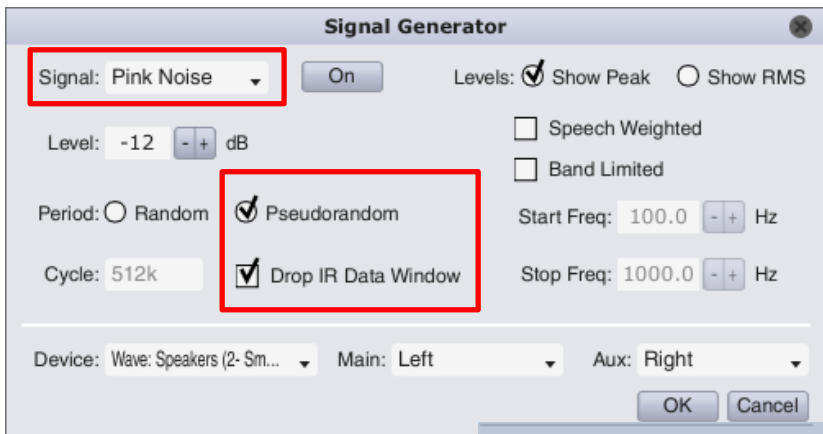
Recap: Recommended Default Settings for Dual-channel IR measurements

The following are some recommended “go-to” default settings for IR measurements in Smart that should typically work well for most rooms.



Signal Type

If you are able to use Smart’s signal generator as your stimulus signal source then period matched pseudorandom noise is a good all-around choice for signal type. To turn on this option, open the signal generator control panel, select *Pink Noise* as the signal type, then tick the boxes labeled *Pseudorandom* and *Drop IR Data Window*.



FFT size and Averaging

128K makes a good default setting for FFT size. At 48k sampling rate, that gives you almost 3 seconds of time window. You would generally have to be measuring in a pretty huge space to need more than that, but it’s not ridiculously long for smaller venues. For averages, a setting of 2 is a good when using period-matched noise. If you have to use a random signal source for some reason, up the number of averages to 8, or maybe 16 if you are working in a noisier environment. We are assuming 0% overlap for averaging.

Excitation Level

If you need to measure reverberation time, then your excitation level needs to be a minimum of 45 dB above the background noise level for T30 (preferred) or at least 35 dB above to get T20. For most other purposes, any excitation level that is comfortably above the background level should be fine.

Input Levels

When using random or pseudorandom noise signals, -10 to -12 dB or so is the preferred input level for any kind of measurement in Smart including IR measurements. -12 dB is the point where the input levels in Smart turn yellow.

Delay Time

When using period-matched noise as your excitation signal you can set the delay time to 0 if you don’t already have it set for your selected signal pair. If you do, then there’s no harm in leaving it alone. If you are using a random noise source and don’t already know the delay time for your signal pair, run the IR measurement once to find it, then click the “@” button to set it, then run the measurement again.

Appendix A: Applicable Standards and Further Reading

Several of the techniques, procedures and quantities discussed in this document are the subject of ISO and IEC standards. We highly recommend reading those standards, rather than relying on our interpretations and synopses.

The main one is ISO 3382, *Acoustics — Measurement of room acoustic parameters*, parts 1 and 2.

- *Part 1: Performance Spaces*, (ISO 3382-1) seems to be geared more toward concert halls, opera houses and theaters, and contains a quite a bit of information on acoustical quantities other than reverberation times not found in...
- *Part 2: Reverberation time in ordinary rooms* (ISO 3382-2)

We note that there is quite a bit of overlap between the two. In fact, the rationale for making it a two-part standard doesn't seem immediately obvious to us, but if you didn't want to buy both parts, part one is the more comprehensive of the two. Both parts contain *much* more information with regard to measurement procedures and statement of results than we have presented here.

For anyone interest in getting into the nuts and bolts of techniques used for indirect IR measurement, ISO 18233, *Acoustics — Application of new measurement methods in building and room acoustics* is a good place to start.

If you are interested in making speech intelligibility measurements using STI or STIPA, IEC 60268-16, *Sound system equipment – Part 16: Objective rating of speech intelligibility by speech transmission index* (Edition 4, 2011 or later) is really a must-read.

Further Reading

Two books that we found particularly useful in preparing this document were:

Ballou, Glenn. ed. *Handbook for Sound Engineers, Third Edition*. Focal Press.

Toole, Floyd. *Sound Reproduction, Loudspeakers and Rooms*, Focal Press.

Classic Reference Works for Acoustics

Beranek, Leo J, *Music, Acoustics and Architecture*, Wiley.

Olson, Harry F, *Modern Sound Reproduction*, Van Nostrand Reinhold.

Rigden, John, *Physics and the Sound of Music*, Wiley.

Jones, D.S, *Acoustics and Electromagnetic Waves*, Clarendon Press

Kleppe, J.A, *Engineering Applications of Acoustics*, Artech House

Rossi, Mario, *Acoustics and Electro-acoustics*, Artech House

Berg, Richard E. and Stork, David G., *The Physics of Sound*, 2nd Ed., Prentice Hall.

Borden, G. J. and Harris, K. S., *Speech Science Primer: Physiology, Acoustics and Perception of Speech*, 2nd Ed, Williams and Wilkins.

Appendix B: Room Volume, Absorption and Reverberation time (RT60)

Reverberation time is usually measured or calculated for a 60dB decay range (RT60). The following information will help you to appreciate the factors that affect reverberation time so that, after measurements have been made, appropriate changes can be suggested where necessary.

RT60 is heavily dependent on the room volume, the amount of absorption in the room and air absorption at higher frequencies. A popular approximation, based on Sabine's formula, is:

$$RT = 55.3V / c(A + 4mV)$$

where

- *RT* is the time taken for the reverberant sound energy to decay by 60dB (in seconds)
- *V* is the volume of the room in m³
- *c* is the speed of sound in m/s (varies with temperature)
- *A* is the total sound absorption of room materials in m² Sabines – i.e. the sum of surface areas, each multiplied by its respective sound absorption coefficient
- *m* is the intensity attenuation coefficient of air per meter (varies with temperature and humidity)

Material absorption coefficients

Typical absorption coefficients:

Material	Typical mid-band sound absorption coefficient
Marble	0.01
Plastered walls	0.02
Bare brick	0.03
10mm Plywood	0.09
10mm mineral wool	0.60
25mm polyurethane foam	0.70
Acoustic ceiling tile	0.72
Audience member on upholstered seat	0.88

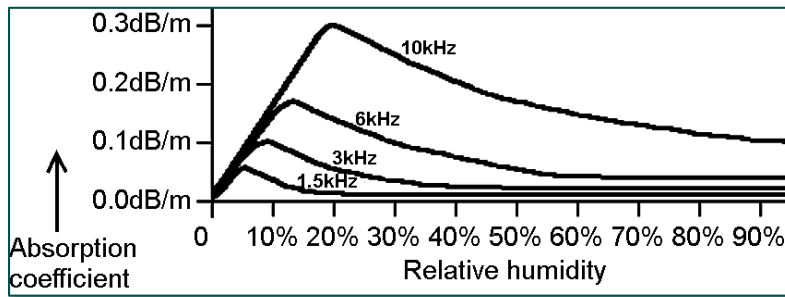
Note that absorption coefficients vary with frequency.

See: www.sengpielaudio.com/calculator-RT60Coeff.htm

Air absorption

Air absorbs sound quite significantly at higher frequencies due to molecular resonance. This high frequency air absorption varies with temperature and relative humidity in a way too complex to cover here. However, a useful air absorption calculator may be found at the link below.

www.doctorproaudio.com/doctor/calculadores_en.htm#calc_humid



Air absorption (in dB/m) vs RH at room temperature

Note

Sound absorption through air is *per meter* – not per doubling of distance as seen for radial attenuation. Sound in a room with a mid-band RT60 of, say, 2 seconds, will travel several hundred meters before decaying into the noise floor so air absorption is a significant additional factor in reducing high frequency reverberation times.

Schroeder cut-off frequency

RT60 assessments are usually restricted to the *Schroeder region* of the frequency range where mid and high frequency wavelengths are short compared to the room dimensions, and sound paths become dense, diffuse and fairly random. At lower frequencies individual modes occur. These modes are more related to room dimensions, less dense and don't follow the diffuse reverberation characteristics required for accurate RT analysis.

The crossover between the Schroeder region and the Modal region is known as the *Schroeder frequency*. This frequency only marks a cross-over region where modes start to overlap by a fairly arbitrary factor of three. It is not an exact science. A popular approximation based on Schroeder's work is:

$$f = 2000\sqrt{RT}/V$$

As before, RT is the time taken for the reverberant sound energy to decay by 60dB (in seconds) and V is the volume of the room in cubic meters. For room volume in cubic feet, the equation can be written as:

$$f = 11885\sqrt{RT}/V$$

Practical measurement considerations

Most of the above equations are approximations assuming an omnidirectional sound source, homogenous sound distribution - and the possession of accurate venue information! In reality, accurate venue information may be difficult to acquire and operating conditions may vary from event to event.

Reverberation characteristics are also likely to vary slightly across the audience area and early decay characteristics (see later) will vary significantly from seat to seat so accurate measurements, over a practical range of operating conditions, are recommended.

As most rooms have significant background noise levels (the noise floor), a full 60dB decay range isn't always available – especially when the direct sound is at normal speech levels or, as is possible with Smaart, we measure a system with an audience in place.

For convenience, instead of measuring the full 60 dB decay time, the decay slope (in dB per second) is measured over a 30 dB range (preferred) and then doubled to work out the RT60. The slope is conventionally measured between the -5dB and -35dB points, as long as the lower point is at least 10dB above the noise floor. If the noise floor is exceptionally high, it may sometimes be necessary to measure the decay slope over a smaller level range. In that case a 20 dB range may be measured instead and then tripled to normalize to equivalent 60 dB decay time.

According to ISO 3382, reverberation time measured over a 30 dB range is called T_{30} and T_{20} signifies a 20 dB measured range. In their notation scheme, the letter T , by itself, stands for 60 dB decay time and so hopefully it is understood that the 20 and 30 refer only to the measured range. The stated reverberation time in seconds for either figure is the equivalent 60 dB decay at the measured rate of decay.

Typical values are as follows: *(in practice, of course, acceptable ranges vary with expectations - based on the venue size)*

- | | |
|------------------------------|-----------------------|
| ▪ Talk studio | less than 0.5 seconds |
| ▪ Conference/classroom | up to 1 second |
| ▪ Musical theatres | 1 to 1.5 seconds |
| ▪ Chamber music/opera venues | 1.5 to 2 seconds |
| ▪ Symphony halls | 1.5 to 2.5 seconds |

Appendix C: Sound source characteristics

A sound source's positioning, orientation and directional characteristics will determine both its ability to cover an audience consistently and its propensity to excite non-audience areas of the room. The latter can enhance the listening experience if, for example, it adds a nice, clean "tail" of diffuse reverberation to an orchestral performance. But room excitation can also ruin the auditory experience if, for example; loudspeaker lobes are causing coloration or if loudspeaker arrays are poorly aimed causing late reflections, excessive reverberation and a loss of intelligibility.

Ignoring high frequency air absorption, a point source's direct sound pressure will be inversely proportional to the source-to-listener distance (i.e. 1/r or 6dB attenuation for every doubling of that distance). In echoic rooms, however, reverberation will add to the direct sound at the listener position depending on how directional the source is. An omnidirectional source, away from any boundaries, would radiate sound pressure spherically and cause strong room excitation.

Directivity factor (Q)

If a sound source is directional, so that its coverage is not fully spherical, its sound power gets concentrated into a smaller sector of a sphere. To help quantify that sound power concentration, sound sources are said to have a directivity factor (Q).

Directivity factor (Q) is:

- 1 for spherical source
- 2 for hemispherical source
- 4 for quarter spherical source...and so on.

Sound pressure level at the listening position

If we use directivity factor (Q) to quantify a sound source's directional properties, then, for on-axis listeners, the directivity-dependent relationship between source sound power level and the total (direct + reverberation) listener position sound pressure level approximates to:

$$L_p = L_w + \log_{10} \left(\frac{Q}{4\pi r^2} \right) + (4/R)$$

The directivity factor (Q) can also be used to estimate the critical distance (Dc).

$$Dc = 0.14 \times \sqrt{(QR)}$$

For the previous two equations:

- L_p is the total (direct + reverberation) sound pressure level at the listening position
- L_w is the source's sound power level in dB ref to 10^{-12} Watts
- Q is the directivity factor (*1 for spherical, 2 for hemispherical, 4 for quarter spherical etc.*)
- π is 3.142...
- r is the distance between the source and the listening position
- R is the room constant (*the room's ability to absorb sound - i.e. the product of surface area and absorption coefficient*)

We don't need to memorize these equations but understanding the relationship will help us troubleshoot problems.

For instance, we see that:

- Adding absorption will, obviously, reduce reverberation and increase the critical distance. Alternatively, we could reduce the effects of the reverberation by aiming highly directional loudspeakers into the audience but away from the walls and ceiling
- Conversely, using less directional sources and/or lower room absorption (preferably with plenty of diffusion and a good balance between early and late decay times) will add “air” to our sound and prevent it from becoming too “dry.”

Directivity index (DI)

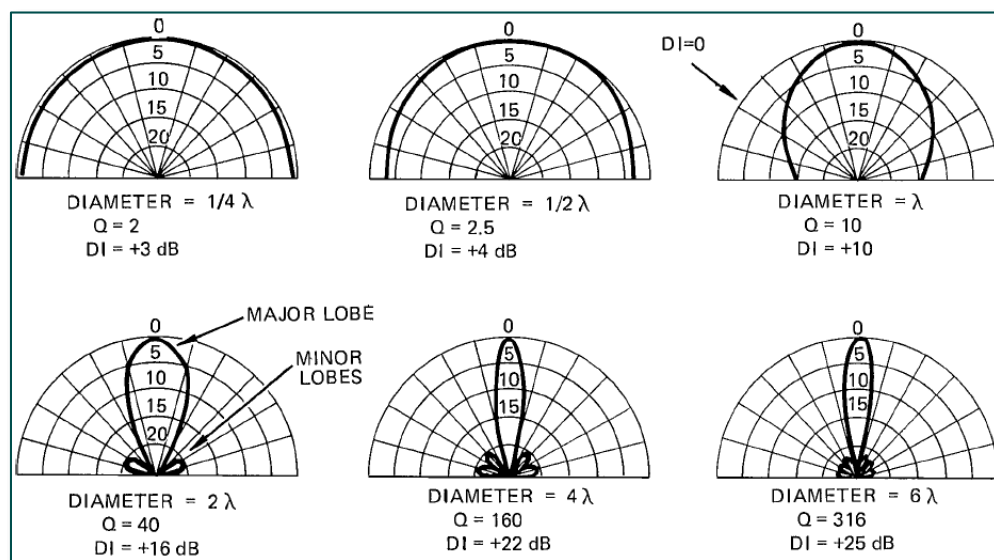
An alternative directivity figure that indicates the concentration of radiated power in dB is the Directivity Index (DI). Its value (in dB) is:

$$DI = 10 \times \log Q$$

So, assuming the listener is on axis, if $Q = 2$, $DI = +3\text{dB}$; if $Q = 4$, $DI = +6\text{dB}$ etc...

Q and DI figures in practice

Large radiating surfaces tend to be more directional than small ones. Whilst good far-field summation occurs perpendicular (on-axis) to large sources, less than perfect summation occurs off-axis due to phase variations between waves emanating from different areas of the surface. This can cause minor off-axis lobes at some frequencies and cancellations at others.



Courtesy: JBL Professional/Harman

The above illustration shows the polar patterns, plus the relevant directivity factor (Q) and directivity index (DI) of baffle-mounted piston-like sound sources (e.g. stiff loudspeaker cones) for various diameters with respect to the wavelength of sound being produced.

Reading from left to right, top to bottom, you can either think in terms of an increasing diameter at a fixed frequency for each successive example . . . or, you can think in terms of increasing frequency for a fixed diameter for each successive example.

The nice thing about the directivity index figure DI is that it can be used in conjunction with a source's polar plot (assuming a dB level scale) to indicate the true directivity index for an off-axis listener. This is shown in the above top right example. The on-axis DI is +10dB, but, at the off-axis point (arrowed) the polar response is 10dB down with respect to the on-axis position. The DI at the off-axis point arrowed is the on-axis DI minus the off-axis figure . . . i.e. the off-axis DI is $+10\text{dB} - 10\text{dB} = 0\text{dB}$.

Conventional loudspeaker arrays

A spherical array of identical high-Q loudspeaker horns will act like one large spherical radiator if the optimum inter-cabinet splay angles are used. Note, however, that the overall array directivity will tend to be lower than that of the individual horns at mid and high frequencies due to the array's wider overall coverage.

At lower frequencies, however, arraying can provide tighter pattern control by making the system acoustically larger. Things can get quite complicated and pretty difficult to predict – especially when you factor in manufacturing tolerances, marketing-optimized loudspeaker specifications, inaccurate venue drawings and a lively room's sensitivity to relatively small changes in array shape and tilt.

Prediction software can get us close to our design goal and will certainly enable us to work out requirements and budgets etc., but we need Smaart to check that installed systems are meeting spec. Smaart's ability to complete as-built measurements – even during shows with an audience in place, if required – makes it a very powerful verification tool.

“Line” arrays

It should be noted that the term “line” array has expanded in popular usage in recent years to become a term for a relatively long, vertical array of loudspeaker – typically just one cabinet wide. Manufacturers often imply that a line array system will give the user a text-book line source radiation characteristic and many users assume they'll benefit from a radial attenuation of 3dB per doubling of distance rather than the 6dB/doubling of a conventional array.

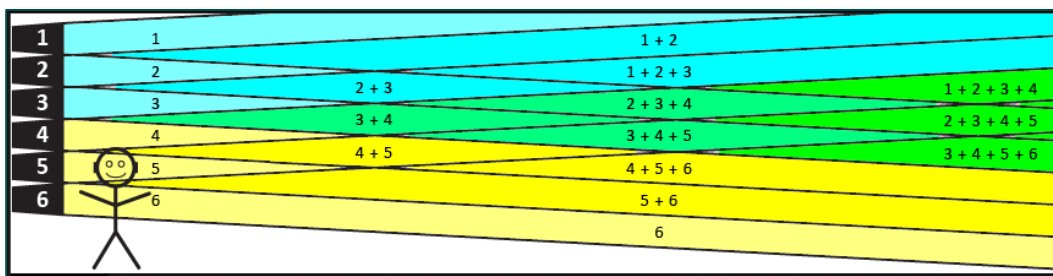
In reality, there is rarely the budget for a straight line source column as it would have to stretch from the floor to the highest seat. Most “line” arrays are much shorter for budget and safety reasons so have to be progressively curved from top to bottom in order to cover the seating areas from the furthest, highest seat down to the front floor areas beyond the front fill coverage.

However, the upper, straighter part vertical pattern control can be very tightly controlled at mid and high frequencies. And, if aimed hard into the audience, can reduce upper walls and ceiling excitation, reducing the reverberant level and increasing the critical distance.

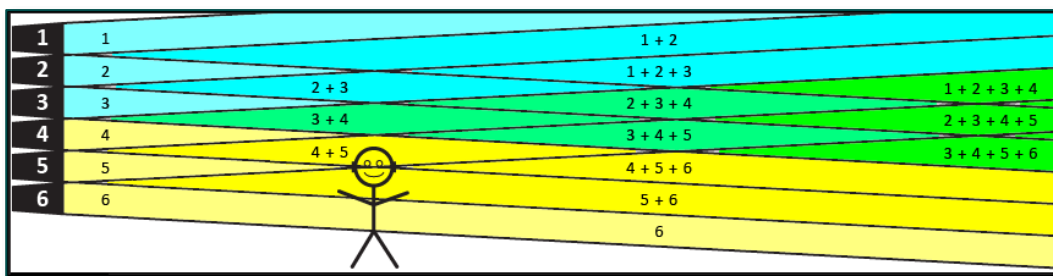
A note about echoes

The upper part of a line array provides tight mid-high pattern control above the array – minimising direct ceiling excitation. However the low radial attenuation rate can make them prone to create very audible echoes from distant boundaries if they aren't aimed accurately.

Always use Smart's linear (Lin) IR graph to check for echoes as they can often be hidden amongst the room's diffuse reverberation – only to become audible, especially on stage, when the audience is in place and reverberant energy levels have dropped.



A simple way of understanding the line array effect of reducing radial attenuation from 6dB/doubling of distance to, perhaps, 3 or 4dB/doubling of distance at mid-high frequencies is to imagine a listener very close to just one horn-loaded element of the array. In the figure above, the listener (let's call him Victor) mainly hears element 5, with some minor off-axis contributions from the other elements.

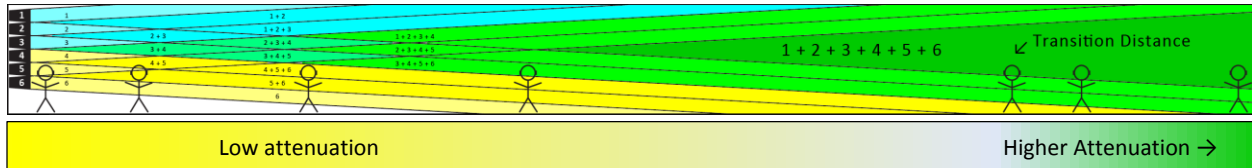


Each element (#1 - 6) is a point source and would, if measured in isolation, have the usual 6dB/doubling of distance radial attenuation characteristic. However, as Victor moves further away from element 5, he starts to benefit from the vector sum of elements 4 and 5 so the usual 6 dB/doubling characteristic is modified by the extra element coming into play. This phenomenon is similar to the near field effect, mentioned earlier, which occurs close to any large sound source.

As Victor continues his journey further away he then benefits from the vector sum of more and more elements and these partially compensate for the usual 6 dB/doubling of distance each element in

isolation would have. Victor will benefit from this low (3 - 4 dB/doubling of distance) attenuation region all the time there are extra elements available.

However, Victor will eventually run out of these extra elements and the radial attenuation characteristic will revert back to the usual 6dB/doubling of distance. This point in Victor’s journey is called the transition distance – see next illustration.



Note that the transition distance, for a true line source, is proportional to frequency and to the line length².

$$\text{Transition Distance} = \frac{\text{Line length}^2 \times \text{frequency}}{2 \times \text{speed of sound}}$$

Allowing for the fact that the speed of sound is temperature-dependent, transition distances for a typical 20 ft (~6m) straight long line source at 68° F (20° C) will be approximately:

Frequency	200Hz	400Hz	800Hz	1.6kHz	3.6kHz	7.2kHz
Transition distance (ft)	35.5	71.0	142	284	638.6	1277
Transition distance (m)	10.8	21.5	43.3	86.5	194.7	389.3

This suggests that high frequency propagation would be very efficient. But, in practice, HF energy will be significantly reduced by air absorption (by up to 0.1dB/meter, worse case) and, of course, imperfect coupling at high frequencies due to rigging tolerances etc.

It is particularly important to remember that transition distance is proportional to the square of the line length if spectral balance is to be maintained over a wide area. The length of the straight line section (at the top of most arrays) needs to be at least 6% of the maximum distance to be covered for spectrally balanced vocals.

- Q)** What has all this got to do with measuring room acoustics using Smaart’s Impulse mode?
- A)** Smaart’s ability to measure room phenomena “live” with an audience in place means that you will often be asked to assess acoustics under practical, amplified show conditions using an installed PA system. Many of these installed systems will be “line” arrays – whether appropriate or not – and a basic understanding of them will allow you to make intelligent recommendations.

Appendix D: Boundary effects

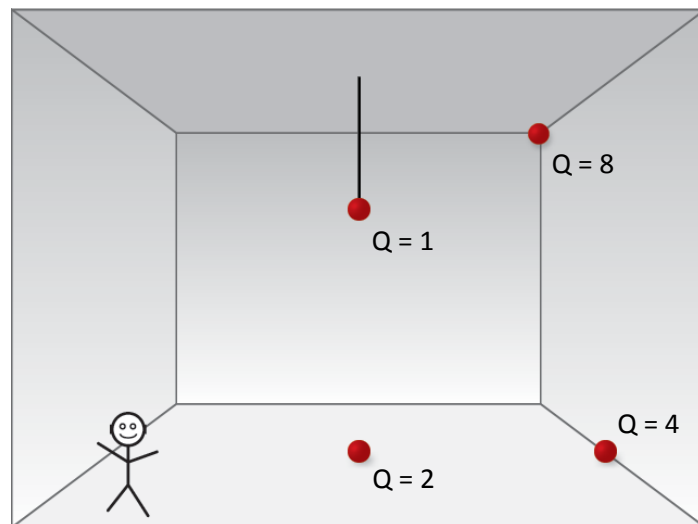
Acoustically small sources

Acoustically small sound sources are sources whose dimensions are small compared with the sound wavelengths they produce. Away from boundaries, acoustically small sources tend to be omnidirectional.

A single-driver 1m cube subwoofer would be a good example of an acoustically small sound source as its size would be less than a quarter of the wavelength at low bass frequencies. At 80Hz, for instance, the wavelength would be just over 14 ft (~4m).

The *effective* directivity of an acoustically small source tends to be governed by local boundaries. The illustration shows a large room with some identical, acoustically small (red spherical) sources in various positions with respect to the room boundaries. (*Ignore Victor. He's only there to avoid 3-D ambiguity*)

- The source that is dangling in free space has a Directivity Factor (Q) of 1 and radiates its acoustic power spherically – into full space.
- The mid-floor source (Q=2) has its acoustic power concentrated into a hemisphere. It radiates the same power but concentrated into half-space.
- The baseboard source (Q=4) has its acoustic power concentrated into a quarter-sphere. Again, it radiates the same power but, this time, concentrated, into quarter space.
- And the rear corner source (Q=8) has its acoustic power concentrated into an eighth-sphere - radiating the same power but, this time, concentrated, into eighth space.



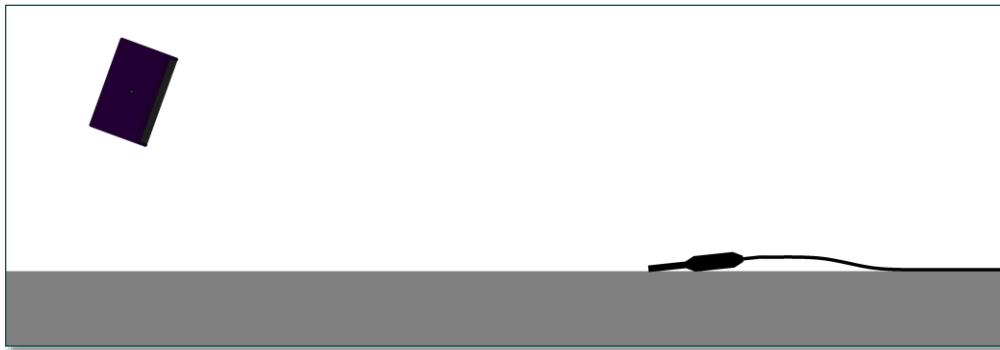
Useful on-axis free field sound pressure level/headroom increases, with respect to Q=1, will approach 6dB per boundary:

- +6dB for Q=2 (half space)
- +12dB for Q=4 (quarter space)
- +18dB for Q=8 (eighth space)

Microphones near boundaries

Similar spl increases apply to pressure microphones if they are placed acoustically close to a surface – i.e. much less than $1/6^{\text{th}}$ of the shortest wavelength of interest. At $1/6^{\text{th}}$ wavelength, the “round trip” will be $1/3^{\text{rd}}$ wavelength, causing a 120° phase shift and resulting in unity gain summation. Larger spacing would cause partial or full cancellation and combing. Placing a microphone capsule on the center of a

large surface can be useful when measuring, for instance, PA systems in an empty room devoid of seats, as the reflective qualities of the floor will only be visible as a 6dB increase in overall level.



This method has to be used with caution though, as the microphone will indicate mid-high frequency sound pressure levels some 6dB higher than would be perceived at ear height in a full audience.

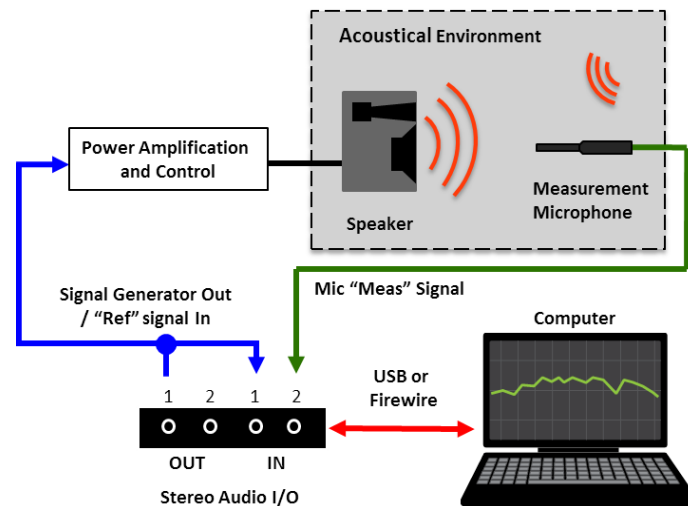
Audience members would, of course, still perceive a 6dB increase at low frequencies – because ear height is acoustically close to the floor at long wavelengths – but they'd perceive a 6dB reduction at mid and high frequencies when compared to the microphone's empty floor response.

Appendix E: Typical Measurement Rig Set-Ups

Below are some example Smaart measurement system setup diagrams for transfer function and dual-channel IR measurement. Basic connections are pretty much the same for any I-O you may encounter. We trust that you understand the audio cabling necessary connect your own equipment together.

Stereo 2x2 Audio I-O

Transfer Function or dual-channel impulse response measurements are the result of comparing a reference signal (pre system) and a measurement signal (post system). This measurement type is essential for aligning loudspeaker systems, as unlike RTA measurements, TF measurements are not “time-blind.” This makes it possible to display Phase (time), Coherence (stability), and Magnitude (response).



In this example, Smaart’s internal Signal Generator is used to excite the system. The output of the generator is routed to input 1 of the I-O. Using a Y-split, the signal is physically looped back (known as a hard-wired loop-back) into an input channel on the I-O. This input is then assigned to the Reference signal of a transfer function measurement.

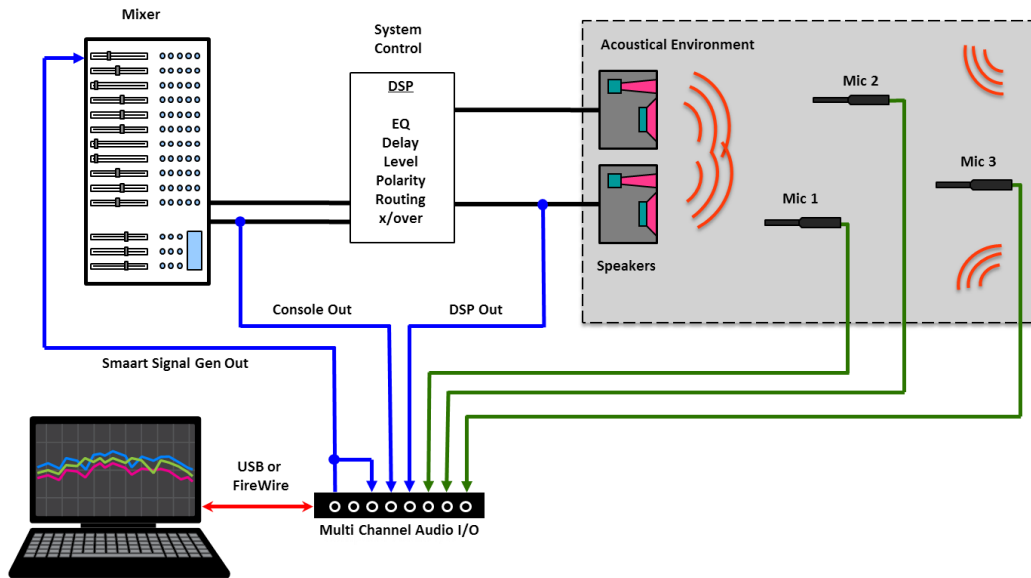
Multi-channel I-O

In version 7, Smaart introduced the capability to run and view as many simultaneous real-time Spectrum and Transfer Function measurement – as well as live averages of individual measurements – as your computer can handle. Having multi-channel I-O capability enables you to set up multiple microphones to compare different measurement positions without a lot of running around. This can be a huge advantage when working with larger, more complex systems.

Traditional USB/FireWire Stand-Alone Interfaces

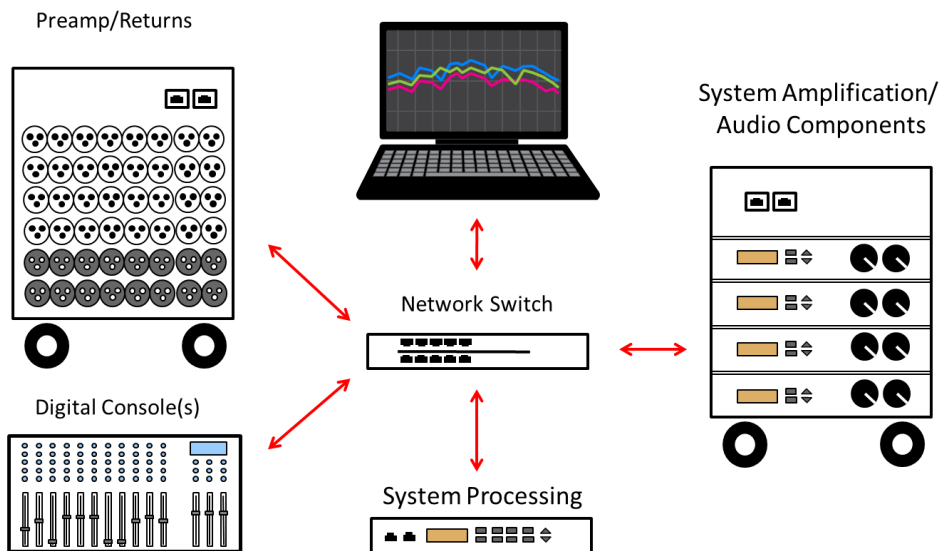
Eight-channel, single rack space packages are common for multi-channel audio interfaces. The single rack space format makes portability in a rack case, or in your backpack, equally viable. In addition to

fielding more microphones, having more inputs enables you to tap into electronic measurement points as well.



Above, we have three measurement microphones as well as electronic measurement points pre-system processing, pre-speaker system, and a hard-wired loopback pre-console. In this example, the Console Out and DSP Out can be used as reference signals that let you isolate the response of the system post-console or post-DSP, or as measurement signals, e.g., for graphing an EQ response in Smart.

Network I-O



The ability to fully digitize your audio stream from the stage pre-amp, all the way to the speaker system has become a reality as of late. In some professional systems nowadays, the only analog signals present

are those from the microphones to the stage conversion box; where they are digitized and distributed via the digital snake.

Using Dante Network Protocol as an example, in a setup such as this the computer operating Smaart sees Dante Virtual Soundcard as an audio I-O device. All routing for the soundcard is done within Dante's Controller application. The computer running Dante Controller (in this example is also running or hosting the Smaart application) has the ability to see any signal connected in the network. With the ability to analyze any signal on the network, this can provide for some extremely flexible measurement configuration possibilities.