How Loudspeaker Impulse Responses Become Sound Pressure Level (SPL) Plots:

The Good, the Bad, and the Ugly

Matthias S. Brennwald

Document version October 26, 2025

1   Introduction

Every loudspeaker designer or tinkerer wants to know how well their loudspeaker performs acoustically. Does it reproduce all frequencies evenly, or does it emphasise some and hide others? The answer lies in its SPL frequency response curve, which shows how loud the speaker is at each frequency.

Understanding and visualising the SPL response is central to almost every aspect of loudspeaker design and evaluation—from developing new drivers to characterising enclosures and interpreting room interactions. The goal seems simple: determine how a loudspeaker responds to different frequencies. The challenge lies in doing so accurately and efficiently, without (unintended) cheating.

An intuitive approach to determine the SPL curve of a loudspeaker would be to play a sine wave with a given frequency at a time, measure the resulting sound pressure level (SPL), repeat the process at the next frequency, and plot the SPL results across frequency. However, this method is very time-consuming and works reliably only in an anechoic environment, as room reflections and echoes can severely distort the measurement. Modern techniques achieve the same result in a single measurement step: they excite the loudspeaker with a short, broadband signal and use the Fourier transform to reveal the frequency response from the measured time-domain data.

This document introduces the concepts that link time-domain measurements (such as impulse responses) to frequency-domain data (such as SPL curves). It aims to provide a practical, intuitive explanation of what the Fourier transform does for loudspeaker measurements—and why understanding its limitations is just as important as applying it. Later sections explore both the strengths and the pitfalls of this process: the good, the bad, and occasionally the ugly.

2 Time domain vs. frequency domain: the Fourier transform

Consider a sound pressure signal evolving as a function of time, $x(t)$. The signal is recorded by a microphone and digitized by an A/D converter with sampling rate F_s. As a result, we obtain N discrete samples $x_i = x(t_i)$ at times

$$t_i = (i - 1) \times \tau, \qquad \tau = \frac{1}{F_s}, \quad i = 1, 2, \dots, N.$$

The idea of the Fourier transform is to describe each time-domain data point x_i as a combination of sine waves, each with amplitude A_j and phase φ_j, at discrete frequencies

$$f_j = \frac{F_s}{N} j, \quad j = 1, 2, \dots, M,$$

where $M = (N - 1)/2$ if N is odd, or $M = N/2 - 1$ if N is even.

The discrete Fourier expansion is then written as:

$$\text{For odd } N: \quad x_i = A_0 + \sum_{j=1}^{M} A_j \cos\left(2\pi f_j t_i + \varphi_j\right), \tag{1}$$

$$\text{For even } N: \quad x_i = A_0 + \sum_{j=1}^{M} A_j \cos\left(2\pi f_j t_i + \varphi_j\right) + A_{N/2} \cos\left(\pi t_i / \tau\right). \tag{2}$$

Each amplitude–phase pair (A_j, φ_j) represents one Fourier component at frequency $f_j = j F_s / N$. The A_0 coefficient is the DC component of the x_i data. Note that the total number of coefficients A_j (including A_0) and φ_j together equals N. In other words, the number of data points in the frequency domain is the same as in the time domain. The above equations therefore represent a fully determined system. The coefficients (A_j, φ_j) can be uniquely obtained from the measured samples x_i by solving this system, typically using efficient numerical algorithms such as the *Fast Fourier Transform (FFT)*.

Note that the Fourier transform is fully reversible. Given the frequency-domain data (A_j, φ_j), the corresponding time-domain signal x_i can be reconstructed exactly using the same equations. In other words, the time-domain data describe how the signal evolves over time, while the frequency-domain data represent the same signal in terms of its amplitude and phase distribution across frequencies. Both domains contain precisely the same information; they are simply two equivalent representations of the very same signal.

2.1 Periodic signals!

An important (but often overlooked) aspect of the discrete Fourier transform is that it implicitly assumes the signal to be periodic with period $T = N$. This follows directly from the use of sinusoidal functions, which are themselves periodic. In other words, when representing the discrete time-domain data x_i as a sum of sinusoidal waves, the discrete Fourier transform assumes that this pattern repeats indefinitely for $t < 0$ and

$t > T$. However, in many practical cases the measured signal is not actually periodic—for example, the impulse response discussed in Sec. 3. This assumption is not inherently a problem, but it does have consequences: if the start and end points of the time-domain data are not at the same amplitude, the implied periodic continuation introduces an artificial discontinuity at the boundary between repetitions. In the frequency domain, that discontinuity appears as additional spectral content (spectral leakage).

2.2   Where did the other frequencies go?

In practice, most real-world signals are not composed of a finite set of perfectly aligned frequencies f_j. Instead, their spectral energy almost always lies between the discrete frequencies sampled by the Fourier transform.

In the discrete Fourier transform, the available frequencies form equally spaced bins centered at the f_j, each with width $\Delta f = F_s/N = F_s \times \tau$. A signal whose frequency does not coincide with one of these discrete f_j values will have its energy distributed over multiple bins, a phenomenon known as spectral leakage. While the total energy of the signal remains conserved, leakage can cause sharp spectral features to appear broadened or blurred across multiple frequency bins.

Sampling also imposes two important frequency limits. The lowest frequency present in the discrete Fourier transform ($f_1 = F_s/N$) is determined by the total length of the digitized signal, $T = N \times \tau$. Because the bin width Δf equals f_1, the signal duration also defines the frequency resolution of the time-domain data. The highest resolvable frequency ($f_M = F_s \times M/N$) is determined by the sampling rate: according to the Nyquist theorem, no frequency above $F_s/2$ can be represented in the digitized data. Any analog signal components above this limit will be folded back (or "aliased") to appear as spurious lower-frequency components in the sampled data.

3   The impulse response and its Fourier transform

Imagine a test signal that contains all frequencies at once, each with the same amplitude and zero phase; that is, $A_j = 1$ and $\varphi_j = 0$ for all $j = 1, \dots, M$. The DC component is $A_0 = 1$ and, for even N, the Nyquist component is $A_{N/2} = 1$. When such a signal is converted from the frequency domain to the time domain, the result is a very short pulse at time t_1 and zero at all other times $t_{i=2\dots N}$ — a Dirac delta. In other words, an ideal Dirac pulse excites all frequencies simultaneously, with equal amplitude and phase.

If we now drive a loudspeaker with this impulse, the microphone records its acoustic output over time. The resulting waveform, $h(t)$, is the impulse response, expressed in sound pressure (Pa) as a function of time. Because the input contains all frequencies equally, any variation in the measured response reflects how the loudspeaker modifies the amplitude and phase of those frequencies, revealing its deviation from a perfectly flat response.

The impulse response thus captures the complete acoustic behaviour of the loudspeaker (and the surrounding room): every resonance, reflection, and delay leaves its fingerprint in $h(t)$. Early peaks correspond to the direct sound from the driver, later

ripples often indicate cabinet or cone resonances, and delayed arrivals show reflections from walls or nearby objects.

4   The impulse response and its Fourier transform

Imagine a test signal that contains all frequencies at once, each with the same amplitude and zero phase; that is, $A_j = 1$ and $\varphi_j = 0$ for all $j = 1, ..., M$, together with a DC component $A_0 = 1$ (and for even N, a Nyquist component $A_{N/2} = 1$). When such a spectrum is converted from the frequency domain to the time domain using the above Fourier expansion, the result is $x_1 = N$ and $x_{i=2...N} = 0$, the discrete analogue of a Dirac delta. In other words, an ideal delta pulse contains all frequencies simultaneously and with equal amplitude and zero phase.

If we now drive a loudspeaker with this impulse, the microphone records its acoustic output over time. The resulting waveform, $h(t)$, is the impulse response, expressed in sound pressure (Pa) as a function of time. Because the input contains all frequencies equally, any variation in the measured response reflects how the loudspeaker modifies the amplitude and phase of those frequencies—revealing its deviation from a perfectly flat, ideal response.

The impulse response thus captures the complete acoustic behaviour of the loudspeaker (and the surrounding room): every resonance, reflection, and delay leaves its fingerprint in $h(t)$. Early peaks correspond to the direct sound from the driver, later ripples often indicate cabinet or cone resonances, and delayed arrivals show reflections from walls or nearby surfaces. Taken together, these features form the bridge between the time and frequency domains.

5   From impulse response to the SPL curve: gating and window functions (reality bites!)

SOMEWHERE MENTION RMS AND DECIBEL CONVERSION: $SPL_j = 20 \times \log_{10}\left(\frac{A_j/\sqrt{2}}{P_{\text{ref}}}\right)$ WITH $P_{\text{ref}} = 20\,\mu\text{Pa}$,

- Introduction: (1) A single measurement of the impulse response gets us the complete frequency-domain data via the Fourier transform. Sleek! (2) Gating the IR to remove measurement artifacts like echoes: nice, we get the anechoic response without needing an anechoic room. Sleek!
- However, there are limits to the method. Build on the theory section (periodicity, fLow=fResolution, etc.) and demonstrate them in a worked example: (1) Gating room echo by cropping echoic part of the IR data results in less data points, and hence the low-frequency limit and bin width is determined by the arrival/delay time of the first echo.
- Windowing with a taper so smooth out the end of the IR: a "massaging" therapy for the signal to be periodic ($x_N \approx x_1$); but it attenuates/messes the tail of the impulse response, so data is changed or lost.
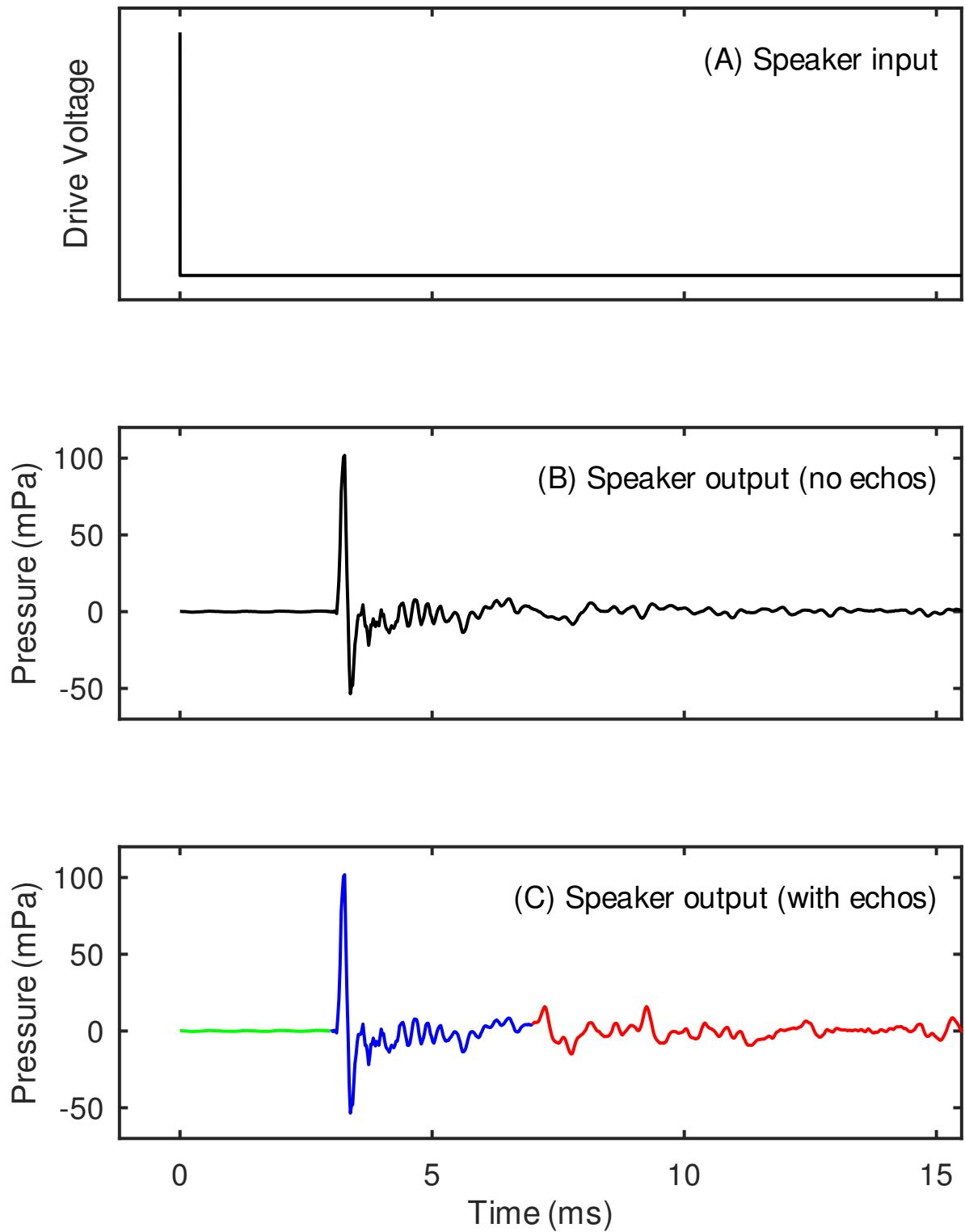
Figure 1: Loudspeaker impulse response. (A) The Dirac input to the loudspeaker, (B) loudspeaker impulse response in an anechoic environment, (C) impulse response of the same loudspeaker, but with echos from the room walls.

Fig. **??**

6 Dealing with reality: The Good, the Bad, and the Ugly

DISCUSS THE METHODS INTRODUCED IN PREVIOUS SECTIONS, DEMONSTRATE THEIR (AB)USE AND EFFECTS

The relationships described in Sec. 2 define the discrete time and frequency grid on which the Fourier transform operates. Real-world signals rarely align perfectly with this grid, which is why techniques to reduce leakage or to improve the visual appearance of and SPL curve are applied – sometimes for the better, sometimes not.

- The following sections discuss how real-world processing shapes this transformation:
  - **The Good:** gating and windowing to suppress reflections.
  - **The Bad:** excessive windowing that removes low-frequency content.
  - **The Ugly:** manipulation of data through zeroing, smoothing, or padding.
- Each step is an attempt to recover the true loudspeaker behaviour from a reflection-rich environment—but not all are equally honest.

This conceptual introduction explains how an impulse response represents a complete description of a loudspeaker's behaviour. Later sections explore how this data is processed, interpreted, and occasionally abused.