Euler's playlist (why does music sound good?)

by Charlie Roberts

April 2026

1 Introduction

Music feels innately emotional, a deeply human expression. Yet beneath every melody lies a precise mathematical architecture. This essay aims to bridge that apparent contradiction: to show that musical beauty is not arbitrary, but is the auditory perception of mathematical order.

We begin by reducing sound to its fundamental variables: pitch, amplitude, timbre, and texture. Subsequently introducing Euler's 1739 **Tonnetz**: a geometric lattice as a form of discrete mathematics encoding relationships between notes through their frequency ratios. In this framework, chords become geometric polygons and progressions become transformations across a discrete space. The 3-4-5 Tonnetz is built on intervals whose ratios reduce to simple integers, and it is precisely this simplicity that Euler believed the human mind finds gratifying.

But the Tonnetz only maps *which* notes are played. To understand *why* they sound harmonious, we turn to the **Fourier transform** — a tool that decomposes any complex waveform into its constituent frequencies by identifying resonance spikes in the complex plane. This connects the geometric abstraction back to physical sound.

This leads us to Euler's *gradus suavitatis*: consonance arises when frequency ratios are simple. A perfect fifth (3:2) sounds smooth because its waveforms synchronize periodically; a dissonant interval does not. Mathematical order, perceived as beauty by the human ear.

2 The Basics

Music is conventionally written in Western staff notation, five staves populated with symbols encoding pitch, duration, and dynamics. For the purposes of this analysis, however, standard notation obscures more than it reveals. Instead, we reduce music to its fundamental physical variables:

- **1.Pitch** the frequency of a given sound measured in Hz (How high or low a note can sound)

- **2. Amplitude** this is the volume of a sound wave, the "intensity" (This determines the energy of a note)
- **3. Duration** the amount of time a note or song takes up
- **4. Timbre** This is the waveform complexity of a sound determined by its harmonics and describable via the Fourier transform. It is what allows the same note to sound distinct across two different instruments.
- **5. Texture** in essence this is the idea of density. (two separate instances of sound at the same time) this can lead to interference of the separate waves in certain contexts.

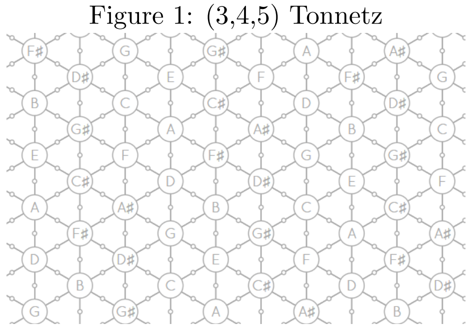
These five variables represent the raw data of sound, but their true complexity emerges through their interaction. By timbre and musical interaction of notes as a composite of periodic functions, we can move beyond music as mere artistic expression and analyses it as a rigorous geometric structure, a shift in perspective that underpins everything that follows.

3 Euler and the Tonnetz (a discrete, geometric abstraction of music)

In 1739, Leonhard Euler introduced the Tonnetz (tonal network) in his *Tentamen novae theoriae musicae*, seeking to visualise the relationships between musical intervals through their mathematical ratios rather than standard notation. The aim of this section is to recognize the differences and relations between musical notes through geometry, before modeling them as physical waves.

Central to Euler's project was defining **consonance** — the perceived sweetness of a sound — through prime numbers. He believed the human mind is gratified by simple mathematical ratios in sound, and sought a rigorous reason why certain combinations of notes sound agreeable. This question of consonance will motivate everything that follows, connecting the geometric structure of the Tonnetz to the frequency analysis of the Fourier transform.

Figure 1: (3,4,5) Tonnetz

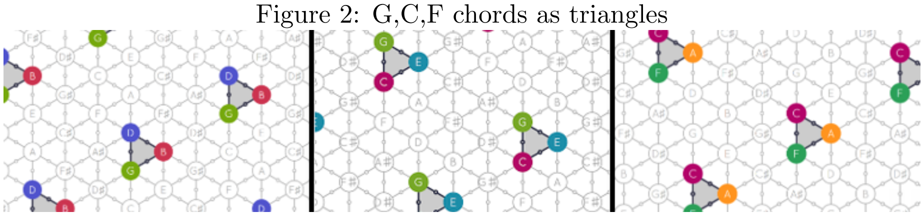


To understand music as a network, we move away from artistic notation toward Discrete Mathematics. We represent each musical note as a vertex in a finite set $V=\{0,1,\ldots,12$ where each number corresponds to one of the twelve notes in Western music. "C" is assigned 0 and "G" is assigned 7. There is no start or end to this set; it wraps around continuously like a clock face, forming a cycle called a Tonnetz (Tone Net).

Using a 3-4-5 Tonnetz}, we construct a lattice by connecting notes that are separated by specific intervals: minor thirds (3 steps), major thirds (4 steps), and perfect fourths (5 steps). An interval is simply the numerical distance between two notes in our set. These intervals are not chosen arbitrarily; when converted into physical frequency, the relationship between these steps reduces to a clean 3:2 ratio. As we will see, it is precisely this simple integer relationship that gives certain combinations of notes their pleasant, harmonious quality.

In this geometric framework, a chord (a set of notes played simultaneously) becomes a geometric polygon. A major chord, for example, occupies three vertices: 0, 4, and 7. When music moves from one chord to another, we are witnessing geometric transformations across this lattice:

Figure 2: G,C,F chords as triangles



Note: A "smooth" chord change is a **reflection** or **rotation** where (n-1) vertices remain fixed.

3.1 The Bridge: From Networks to Harmonics

The Tonnetz is the map. It tells us which notes and chords are being played and how they relate to one another geometrically. However, it does not tell us why certain combinations sound pleasant and others do not, nor does it explain what is physically happening when those notes are heard together. In reality we do not receive notes individually; we receive their sum. When a chord is played, each note produces a sine wave at its corresponding frequency, and these waves interfere with one another to produce a single complex waveform. This combined waveform is what we physically hear, and it corresponds to the timbre discussed earlier.

This presents a practical problem. Given a recording of a piece of music, with many different sounds layered on top of one another, how do we recover the individual frequencies that produced it? The resulting wave appears hopelessly entangled. Undoing it feels, at first, like trying to unmix colours from a tin of paint. The answer lies in the Fourier transform, which will allow us to decompose any complex waveform into its constituent frequencies. Once we can isolate those frequencies, we can examine their ratios and begin to answer Euler's original question: why does music sound good?

4 The Fourier transform

We begin with a visual approach to untangling a sum of sine waves. Consider first a single sine wave, where the yy y axis represents the intensity of the frequency and the xx x axis represents time. This simple model captures the essential behaviour of a single musical note, and will form the foundation for understanding what happens when multiple notes are combined.

$$y = \sin(6\pi x) \tag{1}$$

To convert from Cartesian to polar form, we substitute $x = r\cos(\theta)$ and $y = r\sin(\theta)$:

$$r\sin(\theta) = \sin(6\pi r\cos(\theta))$$

To convert this polar equation back into Cartesian form, we use the inverse substitutions $r = \sqrt{x^2 + y^2}$ and $\theta = \arctan\left(\frac{y}{x}\right)$:

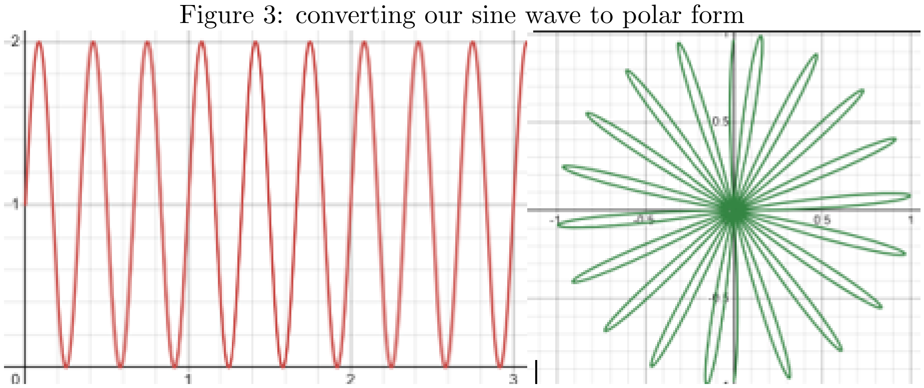
$$\sqrt{x^2 + y^2}\sin\left(\arctan\left(\frac{y}{x}\right)\right) = \sin(6\pi\sqrt{x^2 + y^2}\cos\left(\arctan\left(\frac{y}{x}\right)\right))$$

Which simplifies back to the original identity:

$$y = \sin(6\pi x)$$

This graph demonstrates how as we increase the value of x, the displacement

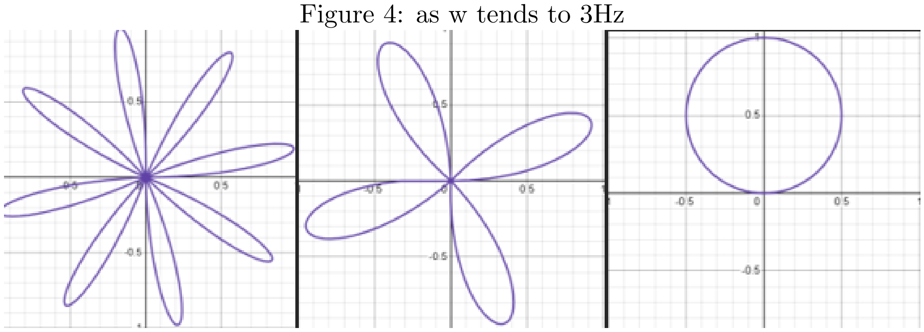
Figure 3: converting our sine wave to polar form



varies with a period of 1/3 You might be asking, "What's the reason for wrapping a wave around the origin?"

Well, by varying the cycles we wrap per second, we are searching for a mathematical resonance. When our winding speed matches the wave's frequency, the signal stops canceling itself out and "bunches up" on one side of the origin. In music, this is how we isolate overtones (another tone within a wave sum eg: sin2x +sin3x) ; it allows us to identify the specific frequency ratios that create an instrument's unique sound by tracking where the center of mass spikes. as seen below the average function value shifts dramatically when the cycles per second that we wrap our wave around the origin is equal to the frequency, here it is 3Hz. The reason for this sudden change of value is called a resonance spike, the wave simply is in sync with the rotational value (every loop the wave lands on the same spot). The brilliance of this is that for a sum of different waves, every frequency within it demonstrates this same pattern.

Figure 4: as w tends to 3Hz



Mathematical Derivation

To find the Fourier Transform of $g(x) = \sin(6\pi x)$ to identify the frequencies within this wave, we use the definition:

$$\hat{g}(f) = \int_{-\infty}^{\infty} g(x) e^{-i2\pi f x} dx$$

Step 1: Rewrite the Sine Wave

Using Euler's Formula, we express the sine function in terms of complex exponentials. Since $\sin(\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2i}$:

$$\sin(6\pi x) = \frac{e^{i6\pi x} - e^{-i6\pi x}}{2i}$$

Step 2: Substitute into the Transform

Substitute this expression into the Fourier integral:

$$\hat{g}(f) = \int_{-\infty}^{\infty} \left(\frac{e^{i6\pi x} - e^{-i6\pi x}}{2i} \right) e^{-i2\pi f x} dx$$

Factor out the constant $\frac{1}{2i}$ and distribute the exponential term:

$$\hat{g}(f) = \frac{1}{2i}\left[\int_{-\infty}^{\infty} e^{i6\pi x}e^{-i2\pi fx}dx - \int_{-\infty}^{\infty} e^{-i6\pi x}e^{-i2\pi fx}dx\right]$$

Step 3: Simplify the Exponents

Combine the terms in the exponents by factoring out $-i2\pi x$:

$$\hat{g}(f) = \frac{1}{2i}\left[\int_{-\infty}^{\infty} e^{-i2\pi x(f-3)}dx - \int_{-\infty}^{\infty} e^{-i2\pi x(f+3)}dx\right]$$

Step 4: Identify the Delta Functions

The definition of the Dirac Delta function is $\delta(f - f_0) = \int_{-\infty}^{\infty} e^{-i2\pi x(f-f_0)}dx$. Applying this to our two integrals:

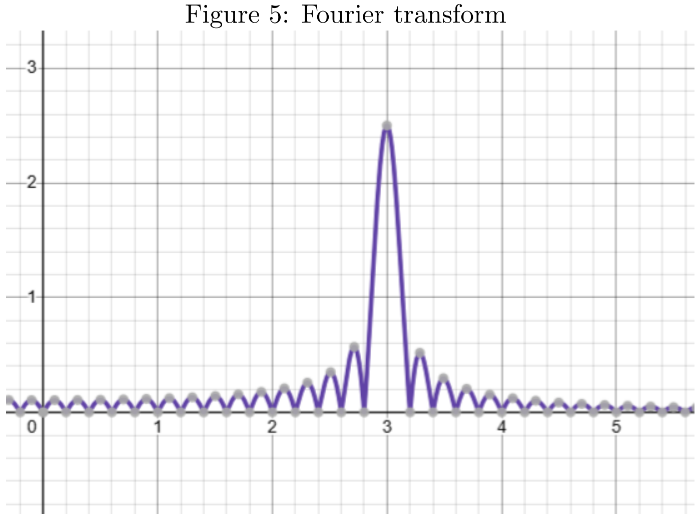
$$\hat{g}(f) = \frac{1}{2i}[\delta(f-3) - \delta(f+3)]$$

Final Result

The result shows two infinite spikes (impulses) at $f = 3$ and $f = -3$:

$$\hat{g}(f) = \frac{i}{2}\delta(f+3) - \frac{i}{2}\delta(f-3)$$

Figure 5: Fourier transform



This figure represents the transform with the x axis being frequencies within our wave defined above, here we can make reference back to our polar form graph where we wrapped the sine wave around the center, the function transformation above represents the coordinate of average mass of our sine wave when wrapped around the origin. In essence this means if we imagined the function like to have an equal thickness and hence a weight, the position of average mass is what is plotted above, the link here is that as our cycles per second vary, the average coordinate of the function hovers around zero. However, when the function reaches 3 cycles per second there is a sudden spike as the mass drastically shifts. This is exactly the frequency of our original wave.

What is so useful about this transform is if we have multiple sound waves interfering with each other to get a resulting sum for example in a chord, as represented by the 3,4,5 tonnetz. the spikes will also appear at the points where the frequency is one of the notes within a chord.

Figure 6: Fourier transform

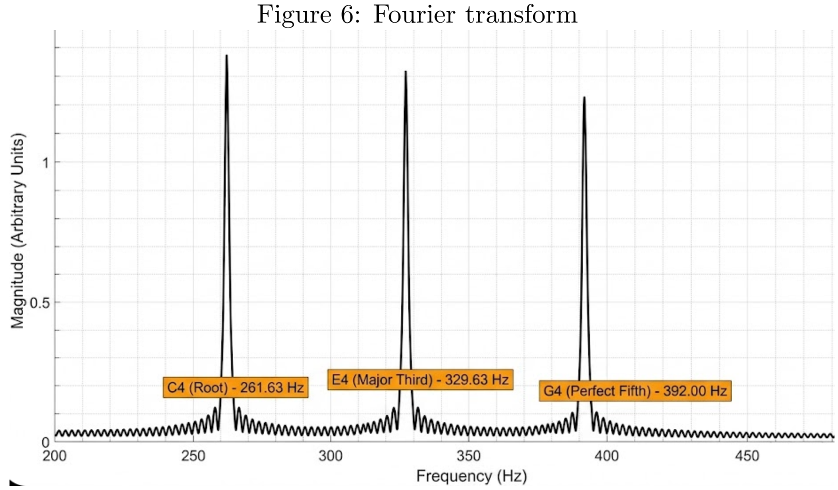
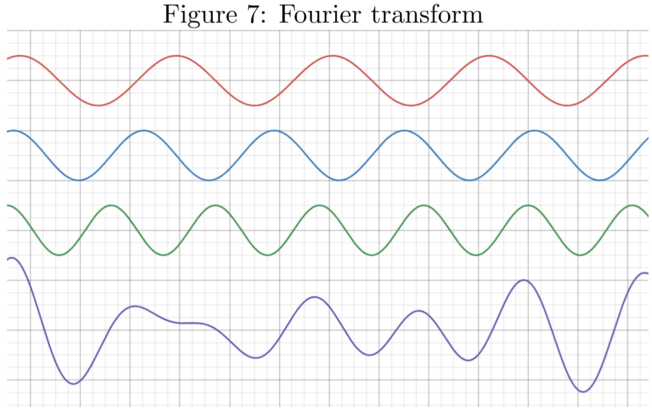


Figure 7: Fourier transform



5 Why an interaction of notes can sound consonant or dissonant

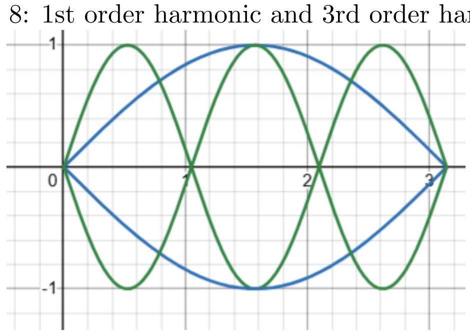
To bridge the gap between your geometric network and the physical wave, we must look at **frequency ratios**. In music, pitch is defined as frequency in Hz , and the intervals in your 3-4-5 Tonnetz—the minor third, major third, and perfect fourth—correspond to specific mathematical relationships. For example, a **perfect fifth** (the interval between C and G) is defined by a frequency ratio of **3:2**. This means that for every 3 cycles of a G note, the C note completes exactly 2 cycles. Music "sounds good" because the human mind is gratified by these simple, low-integer relationships; when notes share a simple ratio, their waveforms synchronize frequently in time. This synchronization is the physical reason the "center of mass" stabilizes in your polar plots—consonance is essentially the auditory perception of mathematical order.

Euler proposed that the simpler the ratio between frequencies, the more "perfect" the sound actually was, the simplest ratio being the **2:1 Octave**, where one frequency is double the next[cite: 30]. One place you may have seen this mathematical relationship is in the **harmonic series**:

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

In this series, you can see these integer relationships clearly. For example, a **1:3 ratio** (the fundamental and the third harmonic) creates a very strong, consonant resonance. Because these numbers are simple, the waves synchronize perfectly every few cycles. This physical alignment is precisely what causes the "resonance spikes" in your Fourier plots where the signal stops canceling itself out and the "center of mass" bunches up[. As Euler theorized with his *gradus suavitatis*, we perceive this mathematical synchronization as "sweetness" or consonance.

Figure 8: 1st order harmonic and 3rd order harmonic



This graph directly represents the 2nd octave of a given frequency (blue), meaning that the harmonic frequencies are all direct multiples of the original. This makes the additional frequencies act as overtones of diminishing amplitude.

These overtones are considered "perfect" because their peaks align precisely with the fundamental wave at regular intervals. In Euler's view, this synchronization creates a low *gradus suavitatis* (degree of sweetness), as the mind easily recognizes the mathematical pattern. This lack of "interference" or clashing between the waves is exactly what makes the octave the cleanest, most consonant sound in music.

6 Closing remarks