Fourier Series and Musical Theory

A. Visintin – Course of Fourier Analysis, a.a. 2016-17

Music is the sound of mathematics [1]

1 Fourier Series and Musical Scales

Music has a remarkable impact on our society, in particular on the cultural, social and economical life. It may then be useful to know at least the most elementary aspects of the musical construction. [2] In this section we address some elements of musical theory, which involve acoustic and Fourier theory.

1. Pitch. The musical signal is characterized by frequency, intensity, rhythm, and timbre. With good approximation, the acoustic response of the ear depends linearly on the logarithm of the frequency of the impacting sound wave, rather than the frequency itself. [3]

The note is a basic musical element. At variance with noise, a note is characterized by being close to periodic, of some period T; it may thus be represented as a Fourier series. [4] Functions of period T may have *fundamental period* T, $T/2$, ..., T/n, and so on. Any sufficiently regular periodic function of period T can be represented as a Fourier series of sinusoids (or complex exponentials) of fundamental period T, $T/2$, ..., T/n.

Let us review some basic notions of wave kinematic. For a sinusoidal wave, let us denote the sound speed by v, the wavelength (the minimal distance over which the wave repeats in space) by ℓ, the fundamental period by T, and the fundamental frequency (expressed in cycles per unit of time) by f_0. [5] We have

$$v = \ell/T, \qquad f_0 = v/\ell = 1/T. \tag{1.1}$$

A wave of period T/n thus has fundamental frequency $n/T = nf_0$. The components of frequency f_0, $2f_0$, ..., nf_0 are called the *harmonics* (or *partials,* or *overtones)* of that frequency. The amplitude of these components, that is the modulus of the Fourier coefficients (or spectrum) of the given function, is proportional to their acoustical intensity. The acoustical energy associated to that component is proportional to the square of the amplitude, and the total acoustical energy is the sum of these contributions.

Let us now transform the frequency nf_0 to the *pitch* [6] $\log nf_0 = \log n + \log f_0$, with respect to a fixed logarithmic basis larger than 1; the specific choice of the basis determines just a factor, and

[1] Leibniz wrote: “Musica est exercitium arithmeticae occultum nescientis se numerare animi.” (Music is a hidden arithmetic exercise of the soul, which does not know that it is counting.)

[2] In Antiquity and Middle Age, music was regarded as one of the seven *liberal arts.* Arithmetic, geometry, music and astronomy formed the *quadrivium;* jointly with grammar, logic and rethoric (the *trivium),* it constituted the basis of the education that was needed to access to the higher university studies of philosophy and theology.

[3] This is an instance of the *Fechner law:* a sensation is proportional to the logarithm of the stimulus. The same holds for the intensity of sound: in first approximation, we respond to the logarithm of the intensity. A similar law of logarithmic dependence of the response applies to several perception phenomena, e.g. sensibility to light. This is at the basis of our capability of recognizing a large range of stimuli; e.g., we can hear frequencies between 20 and 20.000 Hz (about): a span of three orders of magnitude.

[4] This requires assuming regularity conditions; in practice however in most of cases this is not restrictive.

[5] The frequency f is measured in hertz: 1 Hz = 1 cycle/second. The *angular frequency* (or radian frequency) is also used in literature; this is the number of radians per unit of time, and thus equals $2\pi f$.

[6] According to the Oxford Dictionary, *pitch* is the degree of highness or lowness of a note. Although pitch depends on frequency, it is related to our acoustic perception, rather than physical measurements. This mirrors the difference between *psychoacoustics* and *acoustics.* In this section however by pitch we shall refer to the logarithm of the fundamental frequency. (This use of the term however is not standard.)

A similar distinction applies to intensity: *loudness* is the degree of intensity of the sound. It is related to our acoustic perception, rather than physical measurements. Here also the span of our perceptions is rather large: from 1 decibel to (say) 130, up to the pain threshold. (1 decibel is 10 times the logarithm of the ratio between the actual is immaterial. By shifting the origin of the axis, we may eliminate the additive constant $\log f_0$. We may thus identify the pitch with $\log n$.

Let us set $L := \log 2$. For any $n \in \mathbf{N}$, let $m = m(n) \in \mathbf{N}$ and $r = r(n) \in [0, 1[$ be respectively the integer and the fractional part of $L^{-1} \log n$. Thus

$$\log n = (m + r)L \quad \text{with} \quad m \in \mathbf{N},\ 0 \leq r < 1. \tag{1.2}$$

The mapping $nf_0 \mapsto \log n = rL \pmod L$ operates as follows:

$$\begin{aligned}
f_0 &\mapsto \quad \log 1 = 0, \\
2f_0 &\mapsto \quad \log 2 = L = 0 \pmod L, \\
3f_0 &\mapsto \quad \log 3 = L + \log(3/2) = \log(3/2) \pmod L, \\
4f_0 &\mapsto \quad \log 4 = 2L = 0 \pmod L, \\
5f_0 &\mapsto \quad \log 5 = 2L + \log(5/4) = \log(5/4) \pmod L, \\
6f_0 &\mapsto \quad \log 6 = 2L + \log(3/2) = \log(3/2) \pmod L, \\
7f_0 &\mapsto \quad \log 7 = 2L + \log(7/4) = \log(7/4) \pmod L, \\
8f_0 &\mapsto \quad \log 8 = 3L = 0 \pmod L, \quad \text{and so on.}
\end{aligned} \tag{1.3}$$

In music one calls *octave* any pitch interval of length L, in particular the intervals $[mL, (m+1)L[$ for $m \in \mathbf{N}$. Each integer m thus determines an octave on the pitch scale. [7] One often neglects m, and identifies pitches modulus L, by what is called *octave reduction.* [8] Some authors call *chromas* the corresponding pitch classes. This reduction has a clear musical justification: if one *transposes* (i.e., shifts) the notes of a piece of music by an octave, the piece sounds somehow different, but is perfectly recognizable. Our brain indeed establishes a similarity among notes that are an integer number of octaves apart; this is often referred to as *chroma perception.* Something similar occurs when a piece of music is played e.g. by a violin and a cello, or when a man and a woman sing the same song, or pronounce the same words. [9] (For instance, when a choir sings *in unison,* usually men and women are singing an octave apart.)

Why do different musical instruments produce different sounds? As we saw, a note is close to periodic, of some period T, say. This sound corresponds to an evolution of air pressure $u = u(t)$, [10] which can thus be represented by a Fourier series of harmonics: [11]

$$u(t) = \sum_{k=0}^{\infty} a_k \cos(k2\pi t/T) + b_k \sin(k2\pi t/T) = \sum_{k=0}^{\infty} c_k e^{ik2\pi t/T}. \tag{1.4}$$

The k-th harmonic has intensity $I_k = \sqrt{a_k^2 + b_k^2} = |c_k|$, and $\sum_{k=0}^{\infty} I_k^2$ is the acoustic energy of this series. A real sound however may also include *inharmonic* (i.e., non-T-periodic) components

intensity and the minimal perceptible intensity. So any perceptible sound has an intensity of a nonnegative number of decibels. The span between perceptibility and the pain threshold is about 13 orders of magnitude.)

Anyway this picture is somehow simplified. For instance, loudness depends not only on intensity but also on frequency. Loosely speaking, the lower is the intensity of a sound, the less is our sensibility to its components of low and high frequencies. Dually, pitch depends not only on frequency but also on intensity.

[7] As we pointed out, our hearing sensibility ranges between 20 and 20.000 Hz. As $2^{10} = 1024$ is not far from 1000, this makes a span of almost 10 octaves. (The span of a piano keyboard is pretty large, but is just 7 octaves and a half; the span of an organ is 9 octaves. – Actually organs include several keyboards: three or four *manuals* and a *pedal,* respectively operated by hands and feet.)

[8] This reduction is shared by almost all cultures.

[9] In these cases however the sound are also different in the timbre, see ahead.

[10] Here we neglect space dependence.

[11] An *harmonic* is a sinuisodal function (namely, an exponential with imaginary exponent, thus a cosine or a sine is we consider the real or imaginary part), with frequency proportional to an integer multiple of the fundamental frequency. So is $Ae^{ik2\pi t/T}$ for any $A > 0$ and any $k \in \mathbf{N}$.

(see the next section). The smaller is the contribution of these inharmonic terms, the closer is $\sum_{k=0}^{\infty} I_k^2$ to the total energy of the sound.

As we saw, the pitch of a T-periodic sound is related to the fundamental frequency $f_0 = 1/T$. [12] The timbre of a musical instrument (or of a human voice) depends on the ratios $\{I_k/I_0 : k \in \mathbf{N}\}$ between the intensity of the respective harmonics and that of the fundamental. These ratios determine how the acoustic energy is distributed among the harmonics. A tuning fork just produces a *monochromatic* signal, namely a sound of a single frequency: a *pure note* without overtones; the sound of musical instruments is richer in harmonics. Remarkably, differences in phase of different harmonics play no role in our acoustic perception. All this depends on the human physiology of hearing.

For instance, the clarinet produces mainly odd-numbered harmonics, whereas the guitar produces even as well as odd harmonics. This is due to the structure of these instruments: a tube open at one end for the clarinet, strings tied at both ends for the guitar. In many cases the intensity of the harmonic components decreases with the frequency, but this is not always the case. As we pointed out, the sound of a real musical instrument may also include inharmonic components; e.g., for the piano aside the harmonic components due to the strings and the (wooden) harmonic table, inharmonic components are due to the stiffness of the strings, as we briefly illustrate at the next section. [13]

2. The harmonic scale. Pythagoras, and the Chinese before him, proceeded to the following construction of the notes of the *harmonic scale.* [14] Let us assume that f_0 is the frequency of the note F (or *fa,* in Latin countries). So we associate the note F to the pitch $\log f_0$. By identifying pitches modulus L, by (1.3) both pitches $\log f_0$ and $2 \log f_0$ (actually, $2^n \log f_0$ for any $n \in \mathbf{N}$) correspond to the note F. A new note in the sequence of the harmonics of F comes with the third harmonic, i.e. $3f_0$; this is associated to the pitch $\log 3 = L + \log(3/2)$, that is, $\log(3/2)$ by octave reduction. One calls this note C, and says that it is a *fifth* higher than F (the reason of this denomination will be explained ahead).

At this point there are at least two ways to construct the diatonic scale. [15]

(i) One may consider the higher harmonics of f_0:

— the frequency $4f_0$ is associated to the pitch $\log 4 = 2L = 0 \pmod L$, so one retrieves F;

— the frequency $5f_0$ is associated to the pitch $\log 5 = 2L + \log(5/4) = \log(5/4) \pmod L$. This yields a new note, that is called A;

— the frequency $6f_0$ is associated to the pitch $\log 6 = 2L + \log(3/2) = \log(3/2) \pmod L$; so one retrieves C;

---

[12] This is octave dependent: here one cannot apply the octave reduction.

[13] There are also other differences. For instance, the clarinet can *sustain* (that is, maintain in time) all its harmonics, since it is a wind instrument that one blows into continuously; the harmonics of a plucked string instead start out loud, but then decay as the string loses energy. (To be precise, the clarinet harmonics don't stay completely constant either, but they don't just die away the way the string ones do.)

[14] A musical scale is a set of (contiguous) pitches. The most important scale in the Western tradition is the diatonic scale of seven notes (assuming octave reduction); these scales are usually labelled by the first note. There are 12 diatonic scales, one for each first note. For instance, the C Major diatonic scale consists of the pitch classes C, D, E, F, G, A, B; these correspond to the white keys of a keyboard. The diatonic scale in D Major consists of the notes $D, E, F^\sharp, G, A, B, C^\sharp$, and so on.

A large part of classical musical pieces is *essentially* diatonic; nondiatonic notes are often included to give flavor to a composition. This applies especially to the works of Haydn, Mozart, Beethoven (the main authors of the so-called *Classical Period),* and also to most of the written music of the so-called *Common Practice Period* (i.e., loosely speaking, from the beginning of '600 to the beginning of '900). Any musical composition of that period is thus associated to a diatonic scale, which is often indicated aside the name of the piece; e.g., Beethoven's Symphony N. 6 (also known as the pastoral symphony) is in F Major.

[15] On a string instrument different notes are produced by plucking strings of different length. E.g., by doubling the length, the frequency is halfed; by tripling the length, the frequency is multiplied by 2/3, and so on. This follows from (1.3).

— and so on, until the diatonic scale is complete. [16]

(ii) An alternative procedure, due to Pythagoras, consists in proceeding similarly to what we did in (1.2), but starting from $f_1 := 3f_0$ and proceeding by multiples of $3f_0$. Still neglecting the additive constant $\log f_0$, the mapping "frequency $\mapsto$ pitch class" reads

$$\begin{aligned} f_1 = 3f_0 \mapsto \quad & \log 3 = L + \log(3/2) = \log(3/2) \pmod L, \\ 2f_1 = 6f_0 \mapsto \quad & \log 6 = 2L + \log(3/2) = \log(3/2) \pmod L, \\ 3f_1 = 9f_0 \mapsto \quad & \log 9 = 3L + \log(9/8) = \log(9/8) \pmod L, \\ 4f_1 = 12f_0 \mapsto \quad & \log 12 = 3L + \log(3/2) = \log(3/2) \pmod L, \\ \text{and so on.} & \end{aligned} \tag{1.5}$$

The frequencies $f_1, 2f_1, ..., 2^n f_1, ...$ ($n \in \mathbf{N}$) correspond to the pitch $\log(3/2)$ by octave reduction, that we associated to the note C. Here the first new note after F comes with the frequency $3f_1$, which produces the pitch $\log(9/8)$ by octave reduction. One calls this note G, and says that is a *fifth* higher than C, since the difference of their pitches is

$$\log(9/8) - \log(3/2) = \log(3/4) = \log(3/2) - L = \log(3/2) \quad \mod L.$$

The reader will notice the analogy between the construction of C starting from F, and that of G starting from C: in both cases we moved up by $\log(3/2)$ (mod L), namely by a fifth, along the pitch scale.

Proceeding similarly along the so-called *progression of the fifths,* one constructs the note D and so on, until the diatonic scale is complete. By order of construction we get F, C, G, D, A, E, B. [17] Because of octave reduction, this list is not ordered by frequency: e.g. the pitch of C is smaller than that of F, after octave reduction. By ordering these pitches by increasing frequency and starting by C (as it is usually done), one gets the usual arrangement C, D, E, F, G, A, B, by alphabetic order but starting from C. This ordering explains the names given to the notes. [18]

As we saw, the progression of the fifths reads F, C, G, D, A, E, B. The distance between two consecutive elements is four notes, but musicians are used to regard such an interval as a fifth, since they include both extreme elements. For instance, by removing the octave reduction, one goes from F to C by F, G, A, B, C. Similarly, C and E are a third apart. [19]

On a keyboard the notes C, D, E, F, G, A, B correspond to the white keys (7 in each octave).[20] We shall see that the black keys (5 in each octave) represent intermediate pitches, that can be retrieved by taking further steps in the progression of the fifths.

The diatonic scale may also be constructed by an equivalent procedure, which we might call *progression of the seconds.* This consists in alternating a movement up of a fifth and one down of a fourth. The combined effect of these two steps is moving up of a second: so e.g. C, D, E, F, G, A, B.

[16] More in detail, starting by F (by C, by G, resp.) this construction yields the respective sequences of the overtones:

$$FFCFACE^\flat F.... \qquad CCGCEGB^\flat C.... \qquad GGDGBDFG... \qquad \text{and so on.}$$

(A musically knowledgeable reader will recognize that the first seven notes of each of these sequences form the *dominant* 7th *chord* of the previous key. E.g. GGDGBDFG yields $GBDF$, which is the dominant 7th chord of C.)

[17] Here we are confining ourselves to the *diatonic* scale. The progression goes on including alterations, that is, *sharps* and *flats* (see ahead).

[18] In Latin countries these notes are named "do, re, mi, fa, sol, la, si", respectively. The historical origin of these names is more interesting than that of the sequence C, D, E, F, G, A, B, which however goes back to ancient Greece.

[19] Musicians count this way; so for them a fifth plus a third is a seventh; for ordinary human beings instead $4 + 2 = 6$. The explanation stays in that they use ordinals instead of cardinals to measure intervals. As we saw, they also have their own alphabetic ordering: C, D, E, F, G, A, B; namely, they start from C. And, as we shall see, for them the diatonic scale that starts from C reads C, D, E, F, G, A, B, C (with the C repeated at the end); thus, in a way, eight items for a scale of seven notes. So, a world apart ...

[20] The name is due to the fact that, as we saw, musicians deal with scales of 8 notes, since they repeat the initial note.

Let us now extend the scale over the octaves by periodicity, namely removing the octave reduction:

$$...C, D, E, F, G, A, B, C, D, E, F, G, A, B... \tag{1.6}$$

This sequence is unbounded on both sides. But if we restrict ourselves to audible pitches, we get a finite interval of pitches; this extension may depend on various elements, anyway it should span no more than about ten octaves. For instance, standard keyboards have 88 keys, which means 7 octaves and 4 more keys. [21] Organs have a span of 9 octaves.

Melody and harmony. In western music one usually distinguishes four basic elements:

(i) the melody = the sequence of sounds that gives a musical meaning to a composition;

(i) the harmony = the notes that are superposed to the melody (i.e., the accompanying notes); (This provides a sort of background for melodic developments.)

(iii) the rhythm = the time structure of the composition.

(iv) the timbre = the character or quality of a musical sound, given by the musical instrument or voice that produces that sound. (As we saw, this corresponds on how the energy is distributed among the multiples of the fundamental frequencies.)

Loosely speaking, the melody is sequential, whereas the harmony is simultaneous. This has a clear interpretation on the staff, [22] which is interpreted as a Cartesian plane, with time in abscissa and logarithm of frequency in ordinate. The melody and the harmony are respectively represented by the development of the composition in the horizontal and vertical directions: melody and harmony may thus be respectively referred to as the horizontal and vertical dimension of music.

*** Chords.** Harmony rests upon *chords,* i.e., notes that are played or sang simultaneously; the basic chords are *triads,* i.e. chords of three note (more precisely, of three pitch classes). One distinguishes two basic classes: *major* and *minor triads.*

Triads, i.e., sets of three notes (e.g., CEG), are the basic elements of classical harmony. One distinguishes a certain number of types of triads, in particular major and minor triads.

In major triads the notes are separated first by 4 semitones (as in the interval CE; this is called a *major third)* and then by 3 semitones (as in the interval EG; this is called a *minor third).* Major triads include the first elements of the harmonic sequence that is generated by the first note (called the root of the chord); e.g., starting by C_2 we get the sequence C_3, G_3, C_4, E_4, G_4, and so on; by octave identification and alphabetic rearrangement we get CEG.

In minor triads (e.g., ACE) the notes are separated first by 3 semitones (as in the interval AE) and then by 4 semitones (as in the interval CE). This simple modification (3+4 instead of 4+3...) gives the triad a rather different mood: loosely speaking, major triads tend to convey a sense of happiness, whereas minor triads give a mood of sadness. Remarkable, this distinction applies to triads built on any note.

Musical compositions are usually associated to a note; e.g., Beethoven's third symphony *is in* $E^\flat$ mayor; the nineth symphony *is in* D minor. Most musical pieces of the period 1600-1900 are characterized by a specific note, which is sort of fingerprint of the work; this called the key of the play. Often this note is the first as well as the last note of the composition; but this correspondence work → key is more subtle, and here we shall not enter those deep waters.

[21] These go from A_0 to C_8, if one does not apply the octave reduction and labels different notes by an index number. More precisely, the sequence of natural notes (namely white keys) is $A_0, B_0, C_1, D_1, ..., G_7, A_8, B_8, C_8$.

[22] Staff (US) or stave (UK) (plural for either: staves) is a set of five horizontal lines on which the tunes are annotated. This is called *pentagramma* in Italian (but in English *pentagram* has a quite different meaning: it is the David star). For instance for keyboards scores there are two staves, respectively for the right and the left hand. The music of the other instruments is written just on one staff.

Drawbacks of the Pythagorean tuning. C, D, E, F, G, A, B, C is known as the *major scale* in the key of C, which is the note it starts from. [23] If the scale starts by a different note, some keys must be changed. [24] Actually, one cannot simply shift the pitches, since by changing the scale one must preserve the ratios among the frequencies. On the logarithmic scale this means preserving the distances between the pitches, and as we saw these are not uniformly spaced.

It is often convenient to imagine that there are two types of intervals: whole notes (long intervals) and seminotes (short intervals), that here we denote by T and S. The major scale C, D, E, F, G, A, B, C corresponds to the intervals [25]

$$T, T, S, T, T, T, S. \tag{1.7}$$

All together, this makes $5T + 2S = 6T$: the equivalent of six whole tones (or 12 half tones) for 7 intervals. For instance, by starting the diatonic scale by E, one must look for the note that is a whole tone higher. As F is half a tone higher than C, and G is half a tone and a half higher than C, one must select an intermediate note between F and G. [26] This is $F^\sharp$, which corresponds to a black key on a keyboard. Proceeding similarly for the other scales, the notes become 12 (on the keyboard, 7 white keys plus 5 black keys in each octave). By starting by each of the 12 notes and reproducing the pattern (1.7), one gets the 12 major scales:

$$C, D, E, F, G, A, B, C \tag{1.8}$$
$$D, E, F^\sharp, G, A, B, C^\sharp, D \tag{1.9}$$
$$E, F^\sharp, G^\sharp, A, B, C^\sharp, D^\sharp, E \tag{1.10}$$
$$\ldots\ldots \tag{1.11}$$

There are also 12 *natural minor scales,* one for each note. They differ from the major scales by the sequence of the length of the intervals, that here reads T, S, T, T, S, T, T. The natural minor scales are [27]

$$A, B, C, D, E, F, G, A \tag{1.12}$$
$$G, A, B^\flat, C, D, E^\flat, F, G \tag{1.13}$$
$$\ldots\ldots \tag{1.14}$$

There are two further classes of minor scales: the (widely used) *harmonic minor scales,* and the *melodic minor scales;* these also are one for each note. So, major and minor scales all together are 48.

*** Modes.** We got the 12 major scales starting from each of the 12 notes, by reproducing the pattern T, T, S, T, T, T, S. The (natural) minor scales was similarly built by reproducing the

[23] Starting by A one gets A, B, C, D, E, F, G, A, which is called the (natural) *minor scale* of A (because of the pattern of its sequence of tones and semitones, see ahead). But here we mainly confine ourselves to the major scales.

The use of a specific scale (or, as musicians say, of a specific key) has a deep effect on the musical composition, at least according the classical rules of harmony.

[24] For instance, the diatonic scale in the key of D is $D, E, F^\sharp, G, A, B, C^\sharp, D$. On a keyboard this corresponds to using five white keys and two black keys.

[25] This is the footprint of the so-called *major scale.* The (natural) *minor scale* is characterized by the sequence T, S, T, T, S, T, T.

[26] The half tone (or semitone) is the smallest musical interval commonly used in Western tonal music. It is usually divided in 100 cents; so the width of an octave is 12 half tones = 1200 cents.

[27] These are cyclical permutations (of shift 2) of the major scale with root a tone and a half higher. For instance the natural minor scale with root A (i.e., A, B, C, D, E, F, G, A) is obtained by shifting the notes of the major scale with root C (i.e., C, D, E, F, G, A, B, C).

So the family of major scales and that of (natural) minor scales consist of the same sequences of notes, and they differ just in the ordering. This raises an obvious question: why is then the musical flavor of major scales so different from that of minor scales? The answer essentially stays in the use that is made of these scales in harmony. (Explaining this is would take us too far...)

pattern T, S, T, T, S, T, T by starting by each of the 12 notes. Notice that this second pattern is obtained by applying a cyclic permutation to the first one.

It is natural to wonder what may we get if first we apply a different cyclic permutation to that pattern (e.g., T, S, T, T, T, S, T; or S, T, T, T, S, T, T), and then repeat the above procedure of selecting 7 notes starting by each of the 12 notes, but following one of these modified patterns.

This yields new scales, with different sharps and flats. [28]

These new scales are called *modes.* They characterize the medieval music, before introduction of the tonal system that we are used to. [29]

Anyway this is a rather simplified picture of tonality. The distinction between modal and tonal system may be less sharp: in several compositions of the 20th century they coexist. Moreover, besides these more or less precise objective aspects, other more qualitative features are involved; these include the presence in most of compositions of a note that plays a central role. Much is also related to the harmony (namely, loosely speaking, to notes played simultaneously). But delving with that would take us too far.

3. Equal temperament. Next we briefly illustrate why this procedure of constructing the scales is not quantitatively precise in terms of frequencies.

By further pursuing the progression of the fifths, one keeps on producing new notes. Still applying octave reduction, besides the so-called *natural notes* F, C, G, D, A, E, B, one gets five intermediate notes, that are denoted by appending a ♯ (sharp) or a ♭ (flat) to the notes that we already introduced. [30] The progression goes to infinity on both sides. [31]

Moreover, along this progression the frequency intervals are slightly changed at each change of key (of the scale): using a keyboard, one should then retune it. So in practice, by playing a keyboard tuned with the Pythagorean tuning, one is restricted to using a single scale. (Life is slightly easier for the human voice and for some other musical instruments, that may be retuned on the spot.)

We constructed the diatonic scale by using the progression of the fifths. One might use the progression of the thirds as an alternative, or use both. These procedures would yield qualitatively similar outcomes, with the same diatonic scale of notes [32] but each with slightly different frequencies, and still with unequally spaced pitches. None of these methods nor the many variants that have been suggested over the centuries has been able to overcome these drawbacks. Actually, there is a mathematical obstruction, which involves rational and irrational numbers. Next we try to illustrate this issue.

The progression of the fifths cannot be closed. If it were so, then after a certain number of steps one would retrieve a note X with a frequency ratio between X and F equal to a power of two. This is impossible, since $(3/2)^h \neq 2^k$ for any $h, k \in \mathbf{N}$. The same conclusion would be attained progressing by other intervals, since

$$\forall p, q, h, k \in \mathbf{N}, \quad \text{if } p/q \notin \mathbf{N} \quad \text{then } (p/q)^h \neq 2^k. \tag{1.15}$$

It would be much more convenient if the octave were divided into a family of equally spaced intervals (on the logarithmic scale, of course). But this division would not be attainable by applying

[28] This does not exhaust all the mathematically admissible selections of 7 different (unordered) notes out of 12. Computing this number is of course possible, but is not trivial. For instance, A may coexist with $B♭$, but not with $A^♯$, although in equal temperament $A^♯ = B♭$. So one should consider a set of 17 notes rather than 12. But e.g. C may not coexist with $B♯$, since they are the same note, although they are written differently.

[29] The tonal system is at the basis of the Western music of the *Common Practice Period.* Although the 20th century has seen many new trends, the tonal system still pervades popular music, including *canzonette...*

[30] On a keyboard, each black key is the sharp of the note on its left, and is the flat of the note on its right; e.g., $C^♯ = D^♭$. (This identity actually rests upon the equal temperament, that we are going to introduce.)

[31] To be precise, every step of the progression produces a new note. However these new notes are close to the twelve we already defined. (This error is at the basis of the present discussion.) One then neglects these differences, and repeats the twelve notes, identifying sharps and flats.

[32] Starting by C and progressing by thirds, one would get the ordering C, E, G, B, D, F, A.

rational increments of frequency. Let $m \in \mathbf{N}$, and divide the octave of length $L = \log 2$ into m equally spaced intervals of length $L/m = \log(2^{1/m})$. This corresponds to m multiplicative increments of ratio $2^{1/m}$ for the frequency, and this ratio is not rational; so it is not consistent with the Pythagorean tuning.

For instance, the Western culture is used to the 12-note scale that we introduced above. As we saw, the ratio between the frequencies of two contiguous Cs is 2, and there are 12 semitones in each octave; the uniform ratio between the frequencies of two notes that are a semitone apart should then be $2^{1/12}$ (thus $2^{1/6}$ is they are a whole-tone apart). On the logarithmic scale, the distance between two pitches (or two keys of the keyboard) would be a multiple of

$$\log(2^{1/12}) = (\log 2)/12 = 1,05946... \approx 1,05882... = 18/17. \tag{1.16}$$

The ratio between frequencies would then be a multiple of $2^{1/12}$, thus an irrational number, whereas in the Pythagorian theory it was a fraction:

18/17 for a semitone, 3/2 for a fifth, 4/3 for a major third, 5/4 for a minor third.

As a fifth, a major third and a minor third respectively correspond to 5, 4 and 3 semitones, with the system of uniform ratios instead the above ratios equal

$$2^{1/12},\ 5 \times 2^{1/12},\ 4 \times 2^{1/12},\ 3 \times 2^{1/12}, \text{ respectively.}$$

The pitch subdivision that we just outlined is called *equal temperament.* For each semitone, because of (1.16) it corresponds to a frequencies ratio equal to

$$2^{1/12} \approx e^{18/17}\ .$$

This has several advantages: by using it, no approximation is needed to close the progression of the fifths, which thus becomes what is named the *circle of the fifths.* The same twelve notes are thus used for all the scales. In this way one uses just twelve notes, and may transpose musical pieces without retuning keyboards. There is less variety, but more order, and of course this is much more practical.

Anyway equal temperament has a price: one reproduces the harmonic scale only approximately. We associated this scale to the Fourier expansion with frequencies that are in integer ratios: here all that is lost, or at least is just an approximation. From this point of view, every note is out of tune. But the error is minimized by spreading it uniformly over the twelve pitches. [33]

On music notation. It has been noticed that, by introducing the staff about one thousand years ago, the monk Guido d'Arezzo (partly a legendary figure) anticipated the Cartesian coordinates, and also provided a first *time-frequency analysis.* [34]

2 Further Results

Inharmonic components. As we pointed out, real sounds may also include *inharmonic* components. For instance, taking $T = 2\pi$, if we denote by $u = u(x,t)$ the transversal deformation of the string of a piano, the free oscillations fulfill a PDE of the form

$$D_t^2 u - c^2 D_x^2 u + \kappa^2 D_x^4 u = 0 \qquad \text{in } [-\pi,\pi] \times \mathbf{R}. \tag{2.1}$$

The coefficients c^2 and κ^2 respectively account for *elasticity* and *stiffness;* thus the material is purely elastic iff $\kappa^2 = 0$ (as for an ideal string), and purely rigid for $c^2 = 0$ (as for an ideal bar.) We shall assume that $c^2 \neq 0$, so that by setting $\tilde{t} = ct$ and $\gamma = \kappa^2/c^2$, we get

$$D_{\tilde{t}}^2 u - D_x^2 u + \gamma D_x^4 u = 0 \qquad \text{in } [-\pi,\pi] \times \mathbf{R}. \tag{2.2}$$

[33] Some musicians object that in this way we can play everything, but we play it poorly.

[34] Anyway, a large part of the many musical novelties that have been ascribed to Guido were already in use at his time (so he was not really the author).

Let us drop the tilde, and write t instead of $\tilde{t}$. Let us look for solutions of this equation of the form

$$u_k(x,t) = w_k(x)e^{ikt} \qquad \forall (x,t) \in [-\pi,\pi] \times \mathbf{R}, \text{ for some } k \in \mathbf{R}; \tag{2.3}$$

this is called an *Ansatz.* [35] By the method of separation of variables, the general solution u of the PDE (2.2) can be represented as a weighted average of these functions. Notice that each u_k is time-periodic of period 2π iff $k \in \mathbf{Z}$; therefore the sum of the u_ks is 2π-periodic in time iff $k \in \mathbf{Z}$ for all ks. If k_1, k_2 are two distinct rationals, $u_{k_1} + u_{k_2}$ exhibits *beats.* [36] If either k_1 or k_2 is irrational (and the other one is rational), $u_{k_1} + u_{k_2}$ is not even periodic.

For any $k \in \mathbf{R}$, the Ansatz u_k (2.3) fulfills the equation (2.2) iff

$$-w_k'' + \gamma w_k'''' = k^2 w_k \qquad \text{in } [-\pi,\pi]; \tag{2.4}$$

this means that either $w_k \equiv 0$ (which of course is of no interest) or $w = w_k$ is an eigenfunction of the operator $-D_x^2 + \gamma D_x^4$, and k^2 is the corresponding eigenvalue. This problem has a real solution of the form $w_k(x) = \sin(hx + \theta_k)$ (with $h, \theta_k \in \mathbf{R}$) iff

$$h^2 + \gamma h^4 = k^2 \qquad (k \in \mathbf{R}). \tag{2.5}$$

This establishes a relation between k and h: if this holds we write $h = h_k$. Notice that this is independent of θ_k.

As the equation (2.4) is of fourth order, it is in order to prescribe four boundary conditions; for instance, the *homogeneous* conditions

$$w_k(-\pi) = 0, \quad w_k''(-\pi) = 0, \quad w_k(\pi) = 0, \quad w_k''(\pi) = 0. \tag{2.6}$$

The equation (2.4) is of the form $Aw_k = k^2 w_k$ with $A := -D^2 + \gamma D^4$, and this is an elliptic operator. By the theory of boundary problems for elliptic equations, the *eigenvalue problem* (2.4), (2.6) has a nontrivial solution for a countable family of ks.

Here data are such that $w_k(x) = \sin(hx)$ is a solution of this problem iff and (2.5) is fulfllled with $h \in \mathbf{Z}$. By (2.5), this entails that γ is rational; hence $h = h_k \notin \mathbf{Z}$ for any irrational γ. In this case the functions u_ks are still sinusoidal in time, but have no common period. These functions are called *inharmonic* components of the solution, since they do not coincide with the components of any Fourier decomposition.

For instance, we already know that for $\gamma = 0$ the solutions u_k of (2.3) are sinusoidal in time and all have a common period iff $k \in \mathbf{Z}$. For $\gamma \neq 0$, k need not be an integer; then the u_ks are still sinusoidal in time, but have no common period. These functions are called *inharmonic* components of the solution, since they do not coincide with the *harmonics,* namely the components of a Fourier decomposition.

Small English-Italian musical vocabulary:

beats = battimenti,

[35] This German term has many meanings; in mathematics it usually means *guess.*

[36] This acoustic phenomenon consists in a periodic oscillation of the amplitude of a wave combined with a high frequency oscillations. This is explained by the classical prosthaphaeresis formulas:

$$\sin x + \sin y = 2\sin\tfrac{x+y}{2}\cos\tfrac{x-y}{2} \qquad \forall x,y \in \mathbf{R}.$$

So the (additive) superposition of two oscillations of the same amplitude and of respective frequency $1/\alpha$ and $1/\beta$ produces the multiplicative superposition of oscillations of frequency proportional to $2/(\alpha+\beta)$ and $2/(\alpha-\beta)$]. If α and β are close, then $2/(\alpha-\beta)$ is large.

This is at the basis of *dissonance:* if two notes with sufficiently close frequencies are played simultaneously, then our hear perceives the beats, and this is rather disturbing. Dissonance occurs just if the closedness of the two frequencies is in a critical range: no beats are perceived if the two frequencies are either sufficiently close or sufficiently apart.

chord = accordo,
equal temperament = temperamento equabile,
flat = bemolle,
half tone = the interval between E ad F for instance,
key = tasto oppure chiave (ovvero tonalità di un brano),
note = nota,
octave = ottava,
pitch = altezza (di un suono),
pitch class (by octave reduction) = nota,
to play = suonare,
score (or sheet) = spartito (= partitura musicale),
sharp = diesis,
staff = pentagramma,
tone = the interval betweeen C ad D for instance,
tune = intonazione,
to tune = accordare,
tuning fork = diapason,
triad = triade (ovvero accordo di tre note).

References. A huge literature is devoted to musical theory.

In Italian language for instance there are the introductory text

Ziegenrücker: *ABC della musica.* Rugginenti, Torino 2000

and the even more elementary

Shanet: *Come imparare a leggere la musica.* Rizzoli, Milano 2014.

The following article is also of interest:

S. Isola: *Su alcuni rapporti tra matematica e scale musicali.* Matematica, Cultura e Società. Aprile 2016, 31–49

Much is available on the net, too. E.g., see the entertaining talk of P. Oddifreddi:

$http://scuola.repubblica.it/blog/video/odifreddi-la-musica-spiegata-con-la-matematica/?video_rep_id =$ 265205