Some Mathematics in Music[1]

Twelve Keys Per Octave. The seven white keys are C, D, E, F, G, A, B and the five black keys are

$$C^\sharp = D^\flat,\ D^\sharp = E^\flat,\ F^\sharp = G^\flat,\ G^\sharp = A^\flat,\ A^\sharp = B^\flat.$$

Doubling the frequency means going an octave higher. A string vibrates so that the frequency is inversely proportional to the length: halving the length doubles the frequency. Back in ancient Greece, Pythagoras and his school noticed that the "nice sounds" are produced by "rational" changes in the length of the string:

- $G = 3/2$ (a fifth), in the sense that if the string that sounds a "middle C" has length $|OC|$, then, plucking the string at the point G so that $|OC|/|OG| = 3/2$ will produce the sound close to G;
- $F = 4/3$ (a forth);
- $D = 9/8$ (a tone);
- $EF = |OE|/|OF| = 16/15$ (semitone).

The problem is that, with this system, two semitones do not make a tone in the sense of the human ear: $(16/15)^2 \neq 9/8$, whereas the human ear hears sounds in a *multiplicative* way, reacting to *ratios* of frequencies. Alternatively, no number of fifths (the "best" interval after an octave) can produce octaves: because the number $\log_2 3$ is irrational (in fact, transcendental), no integer numbers m and n can satisfy the equation $(3/2)^n = 2^m$.

It was J. S. Bach (1685–1750) who, around 1720 realized (heard?) that the irrational number $2^{1/12} \approx 1.059$ is a very good approximation of the semitone $16/15 = 1.0(6)$ in the sense that all other intervals are well approximated in the multiplicative way:

- $2^{7/12} \approx 1.49 \approx 3/2$;
- $2^{5/12} \approx 1.33 \approx 4/3$;
- etc.

Because $2^{12/12} = 2$, this observation naturally leads to 12 semitones per octave, which translates to the 12 keys per octave on the piano.

The same conclusion (twelve semitones per octave) can be drawn using a *continued fraction* approximation of $\log_2 3$, which also shows that 12 is indeed an "optimal" number: looking at the best rational approximations of $\log_2 3$, we see that fewer than 12 keys would be 5: not enough; the next option after 12 keys is 41, which is too many.

Musical Instruments: The Wave Equation and Boundary Conditions. Emission and propagation of sound is described by the partial differential equation

$$u_{tt} = c^2 \mathbf{\Delta} u,$$

where $\mathbf{\Delta}$ is the Laplace operator and $u = u(t, x)$ describes the displacement, at time t and location x, of the medium that produces the sound; the medium also determines the number c, measured in the units of speed. The equation is solved in a bounded domain $G \subset \mathbb{R}^n$, $n = 1, 2, 3$; the general solution is

$$u(t,x) = \sum_{k=1}^{\infty} \big(A_k \cos(c\lambda_k t) + B_k \sin(c\lambda_k t)\big)\varphi(x), \tag{1}$$

where the numbers A_k, B_k are determined by the initial conditions $u(0, x)$ and $u_t(0, x)$; the numbers $0 < \lambda_1 < \lambda_2 \leq \lambda_3 \leq \cdots$ and the corresponding functions φ_k are determined from

$$\mathbf{\Delta}\varphi_k(x) = -\lambda_k^2 \varphi_k(x), \quad x \in G, \tag{2}$$

with suitable `boundary conditions`; the particular musical instrument determines the domain G and the corresponding boundary conditions. In particular, the `base/fundamental frequency` of the sound produced is

$$\nu_1 = \frac{c\lambda_1}{2\pi}. \tag{3}$$

For STRING INSTRUMENTS (violin, viola, cello, etc.),

- c^2 is the tension divided by the linear density;
- $G = (0, L)$, where L is the length of the string;
- the boundary conditions are $u(t, 0) = u(t, L) = 0$;
- $\lambda_k = \pi k/L$, $\varphi_k = \sin(\pi k x/L)$, so that $\nu_1 = \frac{1}{2L}\sqrt{\frac{\text{tension}}{\text{density}}}$.

[1]Sergey Lototsky, USC

Here is the list of the open strings on the main string instruments and the corresponding frequencies (in Hz).
Violin: G3-D4-A4-E5 (196–293.66–440–659.25); Viola: C3-G3-D4-A4 (130.81–196–293.66–440);
Cello: C2-G2-D3-A3 (65.41–98–146.83–220); Double Bass: E1-A1-D2-G2 (41.2–55–73.4–98).

For wind instruments (both wood and brass), $c \approx 343$ meters per second, is the speed of sound in the air, but the region G and/or the boundary conditions are usually too complicated. Two exceptions are

(1) The flute:
- $G = (0, L)$, $L \approx 0.7$ meters;
- the boundary conditions are $u(t, 0) = u(t, L) = 0$ (both ends are open);
- $\lambda_k = \pi k/L$, $\varphi_k = \sin(\lambda_k x) = \sin(\pi k x/L)$, so that, by (3), $\nu_1 = c/(2L) \approx 250$Hz is rather close to reality;

(2) The clarinet:
- $G = (0, L)$, $L \approx 0.6$ meters;
- the boundary conditions are $u(t, 0) = u_x(t, L) = 0$ (blow into one end, where, for simplicity, $x = L$, the other end $x = 0$ is open);
- $\lambda_k = \frac{\pi}{2L} + \frac{\pi}{L}(k-1)$, $\varphi_k = \sin(\lambda_k x)$, so that, by (3), $\nu_1 = c/(4L) \approx 140$Hz, which is somewhat close to the actual 165Hz.

For DRUMS, c depends on the tension and density of the membrane, and the eigenvalue problem (2) is solved in the disk of radius R, with zero boundary conditions. The radial component of φ_k is now a suitable Bessel function J. In particular,

$$\nu_1 = \frac{c\,\alpha_{0,1}}{2\pi R},$$

where $\alpha_{0,1}$ is the first zero of the Bessel function J_0.

Music Recording.

MECHANICAL PROCESS: The main component is a needle vibrating by going through the grooves. The production of the original grooves and re-production of the sound can be facilitated using a microphone and electrical (vacuum tube based) or electronic (semi-conductor-based) amplifiers. Using lasers instead of a needle to reproduce the sound is also possible [first proposed in 1986; still used, but not very widely].

In the basic setting, sound is captured by the horn, transmitted via a diaphragm to the cutting stylus to produce the grooves.

- In the original phonograph of Edison (1877), the grooves were on a cylinder.
- Later (1888) a disc was introduced, which made it easier to copy the recording: via the process of electrotyping (galvanoplasty), the stamper is produced as a negative impression of the master disc; the stamper is then used to makes copies by pressing onto the heated material.
- The principle material for the discs was hard rubber, followed by shellac [starting around 1895], and then [from 1931] vinyl (polyvinyl chloride); other material were also used.
- The standard rotation speed for the discs (in revolutions per minute): 78, 45, $33\frac{1}{3}$ (usually referred to as LP, or long playing), $16\frac{2}{3}$, $8\frac{1}{3}$ (mostly for spoken word only).

DIGITAL PROCESS is based on the `sampling theorem` (Nyquist 1929, Kotelnikov 1933, Raabe 1939, Someya 1949, Shannon 1949): if the function $f = f(t)$, $t \in \mathbb{R}$, has a *compactly supported* Fourier transform, that is, there is an $\omega_0 > 0$ such that $|\hat{f}(\omega)| = 0$, $|\omega| > \omega_0$, where

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\mathrm{i}\omega t} f(t)\, dt, \ \mathrm{i} = \sqrt{-1},$$

then

$$f(t) = \sum_{k=-\infty}^{+\infty} \frac{\sin\big(\omega_0(t - k\tau)\big)}{\omega_0(t - k\tau)} f(k\tau), \quad \tau \le \frac{\pi}{\omega_0}.$$

[The proof is by writing $\hat{f}$ as a Fourier series and then changing summation and using the equality $f(t) = (1/\sqrt{2\pi}) \int_{-\infty}^{+\infty} e^{\mathrm{i}\omega t} \hat{f}(\omega)\, d\omega$.] The audio signal has the highest (linear) frequency of about 20kHz, which corresponds to taking at least 40,000 samples per second. For certain sounds, the highest frequency can be lower; compression can further reduce the storage requirements. The first CD (compact disc) appeared in 1982; by 1988, CDs outsold LPs. Digital recording can be also done on a magnetic tape (1978).