Mathematics and Music

Dave Benson

Department of Mathematics, University of Georgia, Athens GA 30602, USA

Home page: http://www.math.uga.edu/~djb/index.html

E-mail address: djb@byrd.math.uga.edu

Date: October 15, 2003

 Please email comments and corrections to the above email address. The latest version can be found at

http://www.math.uga.edu/~djb/html/math-music.html

in postscript and acrobat pdf formats.

To Christine Natasha

Ode to an Old Fiddle

From the Musical World of London (1834);[1]

THE POOR FIDDLER'S ODE TO HIS OLD FIDDLE

Torn
Worn
Oppressed I mourn
B a d
S a d
Three-quarters mad
Money gone
Credit none
Duns at door
Half a score
Wife in lain
Twins again
Others ailing
Nurse a railing
Billy hooping
Betsy crouping
Besides poor Joe
With fester'd toe.
Come, then, my Fiddle,
Come, my time-worn friend,
With gay and brilliant sounds
Some sweet tho' transient solace lend,
Thy polished neck in close embrace
I clasp, whilst joy illumines my face.
When o'er thy strings I draw my bow,
My drooping spirit pants to rise;
A lively strain I touch—and, lo!
I seem to mount above the skies.
There on Fancy's wing I soar
Heedless of the duns at door;
Oblivious all, I feel my woes no more;
But skip o'er the strings,
As my old Fiddle sings,
"Cheerily oh! merrily go!
"PRESTO! good master,
"You very well know
"I will find Music,
"If you will find bow,
"From E, up in alto, to G, down below."
Fatigued, I pause to change the time
For some *Adagio*, solemn and sublime.
With graceful action moves the sinuous arm;
My heart, responsive to the soothing charm,
Throbs equably; whilst every health-corroding care
Lies prostrate, vanquished by the soft mellifluous air.
More and more plaintive grown, my eyes with tears o'erflow,
And Resignation mild soon smooths my wrinkled brow.
Reedy Hautboy may squeak, wailing Flauto may squall,
The Serpent may grunt, and the Trombone may bawl;
But, by Poll,* my old Fiddle's the prince of them all.
Could e'en Dryden return, thy praise to rehearse,
His Ode to Cecilia would seem rugged verse.
Now to thy case, in flannel warm to lie,
Till call'd again to pipe thy master's eye.
*Apollo.

[1] Quoted in Nicolas Slonimsky's *Book of Musical Anecdotes*, reprinted by Schirmer, 1998.

Contents

Preface

This volume is a much expanded and augmented version of the course notes for the undergraduate course "Mathematics and Music," given at the University of Georgia. The prerequisites for the first time I gave the course were differential and integral calculus, as well as an elementary knowledge of music notation. The second time, I also required either calculus of several variables or ordinary differential equations. Many parts of the notes require a little more mathematical background. Exactly what parts of these notes are taught in the course depends on who turns up, what they are interested in, and the extent of their mathematical background.

These notes are not really designed for sequential reading from end to end, although it is possible to read them this way. In particular, the level of mathematical and musical sophistication required of the reader varies dramatically from section to section. I have tried to write in such a way that sections which use mathematics beyond the reach of the reader can be skipped without compromising possibly more elementary later sections.

It is quite possible, and the reader is encouraged to do so, to read sections of interest as though they were independent of each other, and then refer back, and to the index to fill in the gaps. The index has been made unusually comprehensive with this in mind.

I particularly encourage the reader to skip some of the later parts of the first chapter, because to be honest, the later chapters are more interesting.

Introduction

> *Why do we consider sine waves to represent "pure" notes, and all other tones to be made up out of sine waves?*

The answer to this question begins with a discussion of the human ear. This will lead to a discussion of harmonic motion, and explain the relevance of sine waves. This discussion gives rise to another more challenging question.

> *How is it that a string under tension can vibrate with a number of different frequencies at the same time?*

This question will focus our attention on the analysis of musical notes into sine waves, or Fourier analysis. Our discussion of Fourier series does not include proofs, but it does include rigorous statements of the relevant theorems. We discuss the distinction between convergence and uniform convergence, through the example of the Gibbs phenomenon. Fourier analysis, together with d'Alembert's remarkably elegant solution to the wave equation in one dimension, answers the question of the vibrating string.

> *Why does the vibration of a drum result in a spectrum which does not consist of multiples of a fundamental frequency in the way that the spectrum of a vibrating string does?*

We discuss Bessel functions as an example of Fourier analysis, for two purposes. First, the Bessel functions describe the motion of a vibrating circular membrane such as a drum, and are essential to understanding why the sound of a drum results in an inharmonic spectrum. The other purpose is as preparation for the discussion of frequency modulation synthesis in a later chapter.

> *Why does the modern western scale consist of twelve equally spaced notes to an octave?*

Spectrum and Fourier analysis will form the point of departure for our discussion of the development of scales. The emphasis is on the relation between the arithmetic properties of rational approximations to irrational numbers, and musical intervals. We concentrate on the development of the standard Western scales, from the Pythagorean scale and just intonation, through the meantone scale, to the irregular temperaments of Werckmeister and others, and finally the equal tempered scale. We also discuss a number of other scales. These include scales not based on the octave, such as the Bohlen–Pierce scale based on odd harmonics only, and the alpha, beta and gamma scales of Wendy Carlos.

> *How can a bunch of zeros and ones on a computer represent music? How does this affect the way we understand and manipulate music?*

In Chapter 7 and we discuss digital representations of signals, digital signal processing, and in Chapter 8 we discuss synthesis and computer music. The emphasis is on how the ideas involved in synthesis and signal processing reflect back into an understanding of structural elements of sound. Interesting sounds do not have a static frequency spectrum, and the goal is to understand the evolution of spectrum with time. We discuss the relevance of Bessel functions to FM synthesis, as well as Nyquist's theorem on aliasing above half the sample rate, the MIDI protocol, internet resources, and so on.

> *How can we formalize the pervasive use of pattern in music? Is there anything mathematics has to say about this?*

In Chapter 9, we discuss how the idea of symmetry relates to music. This leads us into the realms of group theory and combinatorics.

> *Why do rhythms and melodies, which are composed of sound, resemble the feelings, while this is not the case for tastes, colors or smells? Can it be because they are motions, as actions are also motions? Energy itself belongs to feeling and creates feeling. But tastes and colors do not act in the same way.*

Aristotle, *Prob.* xix. 29

Deryck Cooke, who reconstructed from Mahler's sketches a performing edition of his tenth symphony, has written a wonderful book [**22**] in which he describes the musical vocabulary and how it conspires to transmit mood. That is not the subject of this text, as mathematics has little to say about the correspondence between mood and musical form. We are interested instead in the mathematical theory behind music. Nonetheless, we can never lose sight of the evocative power of music if we are to reach any understanding of the context for the theory.

Books

I have included an extensive annotated bibliography, and have also indicated which books are still in print. This information may be slightly out of date by the time you read this.

There are a number of good books on the physics and engineering aspects of music. Dover has kept some of the older ones in print, so they are available at relatively low cost. Among them are Backus [**3**], Benade [**11**], Berg and Stork [**12**], Campbell and Greated [**16**], Fletcher and Rossing [**40**], Hall [**50**], Helmholtz [**54**], Jeans [**60**], Johnston [**64**], Morgan [**90**], Nederveen [**93**], Olson [**95**], Pierce [**103**], Rigden [**112**], Roederer [**118**], Rossing [**121**], Rayleigh [**109**], Taylor [**132**].

Books on psychoacoustics include Buser and Imbert [**15**], Cook (Ed.) [**21**], Deutsch (Ed.) [**32**], Helmholtz [**54**], Howard and Angus [**57**], Moore [**88**], Sethares [**129**], Von Békésy [**10**], Winckel [**140**], Yost [**143**], and Zwicker and Fastl [**144**]. A decent book on physiological aspects of the ear and hearing is Pickles [**102**].

Books including a discussion of the development of scales and temperaments include Asselin [**2**], Barbour [**6**], Blackwood [**13**], Daniélou [**30**], Deva [**33**], Devie [**34**], Helmholtz [**54**], Hewitt [**55**], Isacoff [**59**], Jedrzejwski [**61**], Jorgensen [**65**], Lattard [**71**], Lindley and Turner-Smith [**77**], Lloyd and Boyle [**78**], Mathieu [**83**], Moore [**89**], Neuwirth [**94**], Padgham [**97**], Partch [**98**], Pfrogner [**100**], Rameau [**108**], Ruland [**125**], Vogel [**135, 136, 137**], Wilkinson [**139**] and Yasser [**142**]. Among these, I particularly recommend the books of Barbour and Helmholtz. The Bohlen–Pierce scale is described in Chapter 13 of Mathews and Pierce [**82**].

There are a number of good books about computer synthesis of musical sounds. See for example Dodge and Jerse [**35**], Moore [**89**], and Roads [**113, 114**]. For FM synthesis, see also Chowning and Bristow [**18**]. For computers and music (which to a large extent still means synthesis), there are a number of volumes consisting of reprinted articles from the Computer Music Journal (M.I.T. Press). Among these are Roads [**117**], and Roads and Strawn [**116**]. Other books on electronic music and the role of computers in music include Cope [**23, 24, 25, 26**], Mathews and Pierce [**82**], Moore [**89**] and Roads [**113**]. Some books about MIDI (Musical Instrument Digital Interface) are Rothstein [**124**], and de Furia and Scacciaferro [**42**]. A standard work on digital audio is Pohlmann [**104**].

Books on random music and fractal music include Xenakis [**141**], Johnson [**63**] and Madden [**80**].

Popular magazines about electronic and computer music include "Keyboard" and "Electronic Musician" which are readily available at magazine stands.

Acknowledgements

I would like to thank Manuel Op de Coul for reading an early draft of these notes, making some very helpful comments on Chapters 5 and 6, and making me aware of some fascinating articles and recordings (see Appendix R). Thanks to Paul Erlich and Herman Jaramillo for emailing me various corrections and other helpful comments. Thanks to Robert Rich for responding to my request for information about the scales he uses in his recordings (see §6.1 and Appendix R). Thanks to Heinz Bohlen for taking an interest in these notes and for numerous email discussions regarding the Bohlen–Pierce scale §6.7. Thanks to my students, who patiently listened to my attempts at explanation of this material, and who helped me to clean up the text by understanding and pointing out improvements, where it was comprehensible, and by not understanding where it was incomprehensible.

This document was typeset with AMSLaTeX. The musical examples were typeset using MusicTeX, the graphs were made as encapsulated postscript (eps) files using MetaPost, and these and other pictures were included in the text using the graphicx package.

Essays

During the term, I shall expect each student to write one essay, on a topic to be chosen by the student and approved by me. This will be collected during or before the tenth week of the semester. For undergraduates, the essay will consist of between 5 and 20 typed pages. For graduate students, I shall expect between 10 and 40 pages.

These essays will be graded for grammar, style and use of English, as well as mathematical content. If you use a mathematical formula, make sure it is part of a complete sentence. In general, as a matter of style, a sentence should not begin with a mathematical symbol. Comprehensibility is a key issue too. To ascertain whether what you have written makes sense, I would recommend asking a friend to read through what you have written. If your friend asks you what something means, that's probably an indication that you should include more explanation.

Some examples of topics which the student may like to consider are as follows. This should be regarded as an indication of what sort of topics are likely to be regarded as acceptable. It should not be viewed as precluding a topic which the student may come up with.

Psychoacoustics

The Ear and Cochlear Mechanics

Concert hall acoustics

Sound compression

CSound

MIDI

Digital synthesis algorithms

History of scales from some particular culture
(e.g., Indian, Greek, Arabic, Chinese, Balinese, etc.)

Bessel functions

Combinatorics of twelve tone music

Relation of spectrum to scale

The Fourier transform

Wavelets (requires a strong mathematics background)

Formants and the human voice

Cross interleaved Reed–Solomon codes and the compact disc

The physics of some class of musical instruments
(stringed, woodwind, brass, percussive, etc.)

The phase vocoder

Symmetry in music

Get hold of a technical article from the *Computer Music Journal*, *Acustica*, or the *Journal of the Acoustical Society of America* and explain it assuming only a mathematical background at the level of this course.

CHAPTER 1

Waves and harmonics

1.1. What is sound?

The medium for the transmission of music is sound. A proper understanding of music entails at least an elementary understanding of the nature of sound and how we perceive it.

Sound consists of vibrations of the air. To understand sound properly, we must first have a good mental picture of what air looks like. Air is a gas, which means that the atoms and molecules of the air are not in such close proximity to each other as they are in a solid or a liquid. So why don't air molecules just fall down on the ground? After all, Galileo's principle states that objects should fall to the ground with equal acceleration independently of their size and mass.

The answer lies in the extremely rapid motion of these atoms and molecules. The mean velocity of air molecules at room temperature under normal conditions is around 450–500 meters per second (or somewhat over 1000 miles per hour), which is considerably faster than an express train at full speed. We don't feel the collisions with our skin, only because each air molecule is extremely light, but the combined effect on our skin is the air pressure which prevents us from exploding!

The mean free path of an air molecule is 6×10^{-8} meters. This means that on average, an air molecule travels this distance before colliding with another air molecule. The collisions between air molecules are perfectly elastic, so this does not slow them down.

We can now calculate how often a given air molecule is colliding. The collision frequency is given by

$$\text{collision frequency} = \frac{\text{mean velocity}}{\text{mean free path}} \sim 10^{10} \text{ collisions per second.}$$

So now we have a very good mental picture of why the air molecules don't fall down. They don't get very far down before being bounced back up again. The effect of gravity is then observable just as a gradation of air pressure, so that if we go up to a high elevation, the air pressure is noticeably lower.

So air consists of a large number of molecules in close proximity, continually bouncing off each other to produce what is perceived as air pressure. When an object vibrates, it causes waves of increased and decreased pressure. These waves are perceived by the ear as sound, in a manner to be investigated in the next section, but first we examine the nature of the waves themselves.

Sound travels through the air at about 340 meters per second (or 760 miles per hour). This does not mean that any particular molecule of air is moving in the direction of the wave at this speed (see above), but rather that the local disturbance to the pressure propagates at this speed. This is similar to what is happening on the surface of the sea when a wave moves through it; no particular piece of water moves along with the wave, it is just that the disturbance in the surface is propagating.

There is one big difference between sound waves and water waves, though. In the case of the water waves, the local movements involved in the wave are up and down, which is at right angles to the direction of propagation of the wave. Such waves are called *transverse waves*. Electromagnetic waves are also transverse. In the case of sound, on the other hand, the motions involved in the wave are in the same direction as the propagation. Waves with this property are called *longitudinal waves*.

Longitudinal waves

$\longrightarrow$ Direction of motion

Sound waves have four main attributes which affect the way they are perceived. The first is *amplitude*, which means the size of the vibration, and is perceived as loudness. The amplitude of a typical everyday sound is very minute in terms of physical displacement, usually only a small fraction of a millimeter. The second attribute is *pitch*, which should at first be thought of as corresponding to frequency of vibration. The third is *timbre*, which corresponds to the shape of the frequency spectrum of the sound. The fourth is *duration*, which means the length of time for which the note sounds.

These notions need to be modified for a number of reasons. The first is that most vibrations do not consist of a single frequency, and naming a "defining" frequency can be difficult. The second related issue is that these attributes should really be defined in terms of the perception of the sound, and not in terms of the sound itself. So for example the perceived pitch of a sound can represent a frequency not actually present in the waveform. This phenomenon is called the "missing fundamental," and is part of a subject called psychoacoustics.

Attributes of sound

Physical	Perceptual
Amplitude	Loudness
Frequency	Pitch
Spectrum	Timbre
Duration	Length

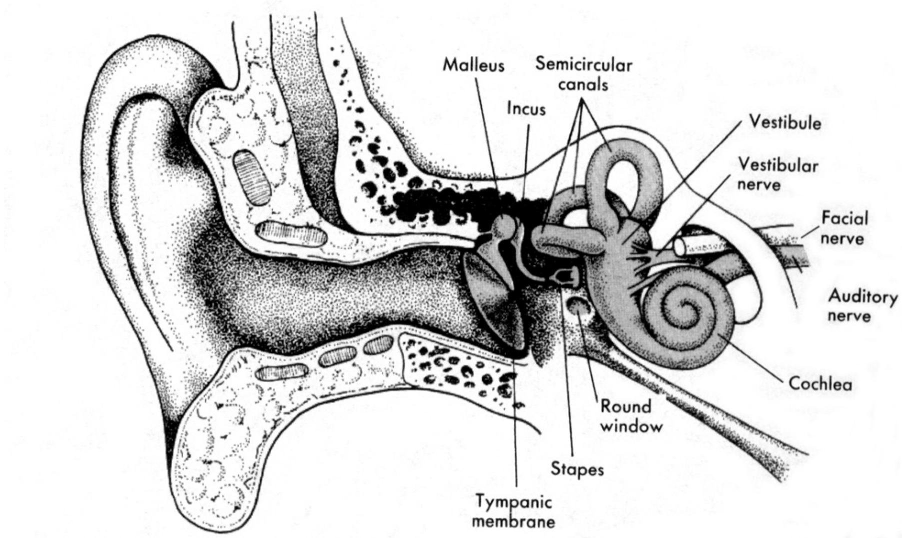
In order to get much further with understanding sound, we need to study its perception by the human ear. This is the topic of the next section.

1.2. The human ear

In order to understand the origins of the mathematical construction of scales, we must begin by understanding the physiological structure of the human ear. I have borrowed extensively from Gray's *Anatomy* for this description.

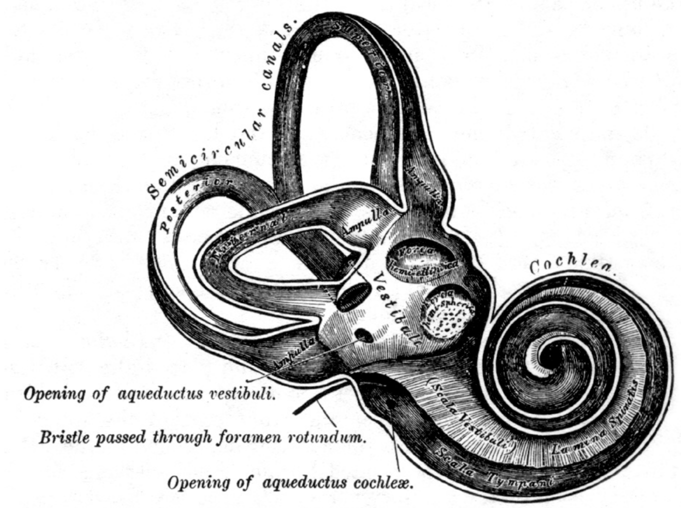
The ear is divided into three parts, called the outer ear, the middle ear or *tympanum* and the inner ear or *labyrinth*. The outer ear is the visible part on the outside of the head, called the *pinna* (plural *pinnæ*) or *auricle*, and is ovoid in form. The hollow middle part, or *concha* is associated with focusing and thereby magnifying the sound, while the outer rim, or *helix* appears to be associated with vertical spatial separation, so that we can judge the height of a source of sound.

The concha channels the sound into the auditory canal, called the *meatus auditorius externus* (or just *meatus*). This is an air filled tube, about 2.7 cm long and 0.7 cm in diameter. At the inner end of the meatus is the ear drum, or *tympanic membrane*.



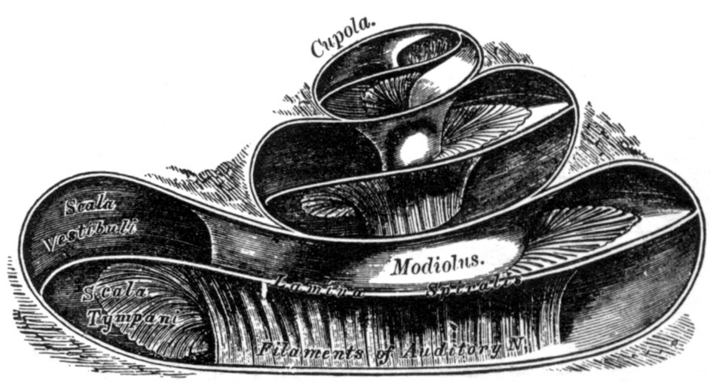
Picture from Berne and Levy, *Physiology*, 1993.

The ear drum divides the outer ear from the middle ear, or *tympanum*, which is also filled with air. The tympanum is connected to three very small bones (the *ossicular chain*) which transmit the movement of the ear drum to the inner ear. The three bones are the hammer, or *malleus*, the anvil, or *incus*, and the stirrup, or *stapes*. These three bones form a system of levers connecting the ear drum to a membrane covering a small opening in the inner ear. The membrane is called the *oval window*.



The osseous labyrinth laid open. (Enlarged.)

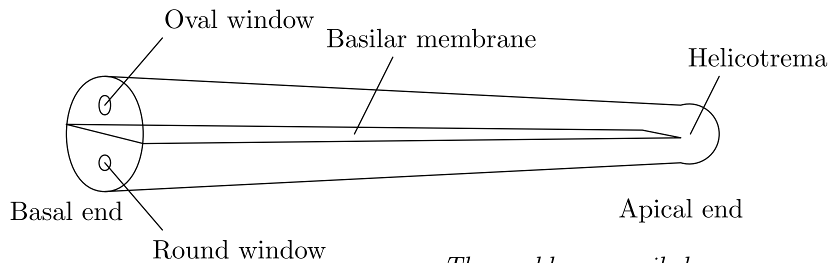
The inner ear, or *labyrinth*, consists of two parts, the *osseous labyrinth*,[1] consisting of cavities hollowed out from the substance of the bone, and the *membranous labyrinth*, contained in it. The osseous labyrinth is filled with various fluids, and has three parts, the *vestibule*, the *semicircular canals* and the *cochlea*. The vestibule is the central cavity which connects the other two parts and which is situated on the inner side of the tympanum. The semicircular canals lie above and behind the vestibule, and play a role in our sense of balance. The cochlea is at the front end of the vestibule, and resembles a common snail shell in shape. The purpose of the cochlea is to separate out sound into various components before passing it onto the nerve pathways. It is the functioning of the cochlea which is of most interest in terms of the harmonic content of a single musical note, so let us look at the cochlea in more detail.



The cochlea laid open. (Enlarged.)

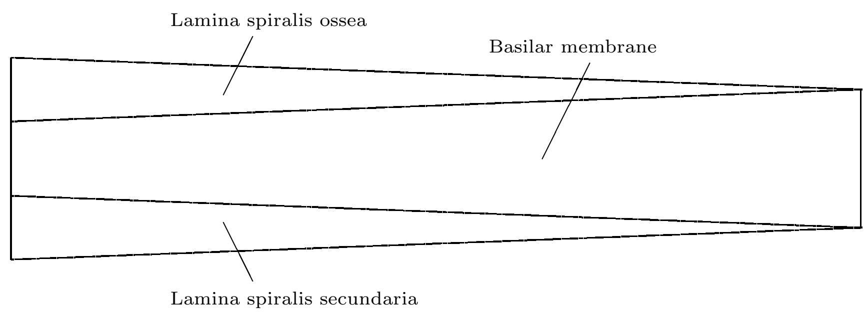
[1](Illustrations taken from the 1901 edition of *Anatomy, Descriptive and Surgical*, Henry Gray, F.R.S.)

The cochlea twists roughly two and three quarter times from the outside to the inside, around a central axis called the *modiolus* or *columnella*. If it could be unrolled, it would form a tapering conical tube roughly 30 mm (a little over an inch) in length.



The cochlea, uncoiled

At the wide (*basal*) end where it meets the rest of the inner ear it is about 9 mm (somewhat under half an inch) in diameter, and at the narrow (*apical*) end it is about 3 mm (about a fifth of an inch) in diameter. There is a bony shelf or ledge called the *lamina spiralis ossea* projecting from the modiolus, which follows the windings to encompass the length of the cochlea. A second bony shelf called the *lamina spiralis secundaria* projects inwards from the outer wall. Attached to these shelves is a membrane called the *membrana basilaris* or *basilar membrane*. This tapers in the opposite direction than the cochlea, and the bony shelves take up the remaining space.



The basilar membrane divides the interior of the cochlea into two parts with approximately semicircular cross-section. The upper part is called the *scala vestibuli* and the lower is called the *scala tympani*. There is a small opening called the *helicotrema* at the apical end of the basilar membrane, which enables the two parts to communicate with each other. At the basal end there are two windows allowing communication of the two parts with the vestibule. Each window is covered with a thin flexible membrane. The stapes is connected to the membrane called the *membrana tympani secundaria* covering the upper window; this window is called the *fenestra rotunda* or *oval window*, and has an area of 2.0–3.7 mm^2. The lower window is called the *round window*, with an area of around 2 mm^2, and the membrane covering it is not connected to anything apart from the window. There are small hair cells along the basilar membrane which are connected with numerous nerve endings for the auditory nerves. These transmit information to the brain via a complex system of neural pathways.

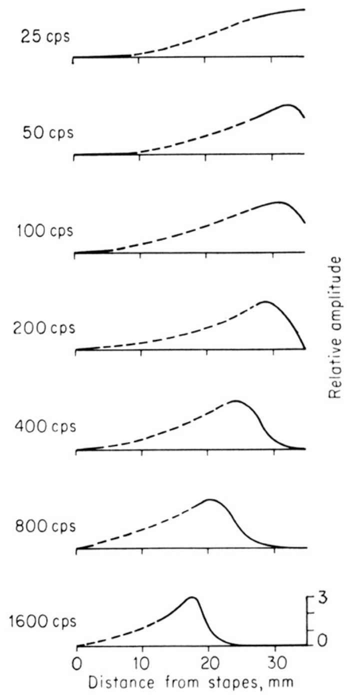
Now consider what happens when a sound wave reaches the ear. The sound wave is focused into the meatus, where it vibrates the ear drum. This causes the hammer, anvil and stapes to move as a system of levers, and so the stapes alternately pushes and pulls the membrana tympani secundaria in rapid succession. This causes fluid waves to flow back and forth round the length of the cochlea, in opposite directions in the scala vestibuli and the scala tympani, and causes the basilar membrane to move up and down.

Let us examine what happens when a pure sine wave is transmitted by the stapes to the fluid inside the cochlea. The speed of the wave of fluid in the cochlea at any particular point depends not only on the frequency of the vibration but also on the area of cross-section of the cochlea at that point, as well as the stiffness and density of the basilar membrane. For a given frequency, the speed of travel decreases towards the apical end, and falls to almost zero at the point where the narrowness causes a wave of that frequency to be too hard to maintain. Just to the wide side of that point, the basilar membrane will have to have a peak of amplitude of vibration in order to absorb the motion. Exactly where that peak occurs depends on the frequency. So by examining which hairs are sending the neural signals to the brain, we can ascertain the frequency of the incoming sine wave. This description of how the brain "knows" the frequency of an incoming sine wave is due to Hermann Helmholtz, and is known as the place theory of pitch perception.

Measurements made by von Békésy in the 1950s support this theory. The drawings at the top of page 7 are taken from his 1960 book [**10**] (Fig. 11-43). They show the patterns of vibration of the basilar membrane of a cadaver for various frequencies.

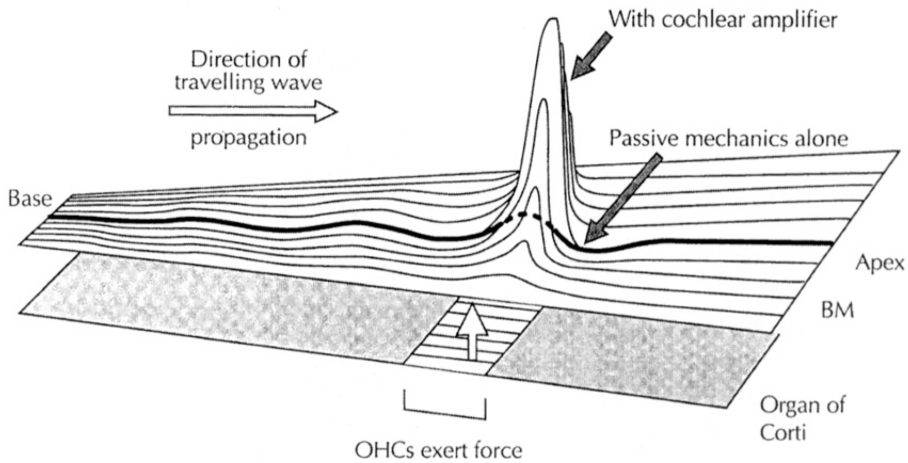
The spectacular extent to which the ear can discriminate between frequencies very close to each other is not completely explained by the *passive* mechanics of the cochlea alone, as reflected by von Békésy's measurements. More recent research shows that a sort of psychophysical feedback mechanism sharpens the tuning and increases the sensitivity. In other words, there is information carried both ways by the neural paths between the cochlea and the brain, and this provides active amplification of the incoming acoustic stimulus. The outer hair cells are not just recording information, they are actively stimulating the basilar membrane. See the figure at the bottom of page 7.

One result of this feedback is that if the incoming signal is loud, the gain will be turned down to compensate. If there is very little stimulus, the gain is turned up until the stimulus is detected. An annoying side effect of this is that if mechanical damage to the ear causes deafness, then the neural feedback mechanism turns up the gain until random noise is amplified, so that singing in the ear, or *tinnitus* results. The deaf person does not even have the consolation of silence.



Von Békésy's drawings of patterns of vibration of the basilar membrane. The solid lines are from measurements, while the dotted lines are extrapolated.

The phenomenon of *masking* is easily explained in terms of Helmholtz's theory. Alfred Meyer (1876) discovered that an intense sound of a lower pitch prevents us from perceiving a weaker sound of a higher pitch, but an intense sound of a higher pitch never prevents us from perceiving a weaker sound of a lower pitch. The explanation of this is that the excitation of the basilar membrane caused by a sound of higher pitch is closer to the basal end of the



Feedback in the cochlea, picture from Jonathan Ashmore's article in [**69**]. In this figure, OHC stands for "outer hair cells" and BM stands for "basilar membrane."

cochlea than that caused by a sound of lower pitch. So to reach the place of resonance, the lower pitched sound must pass the places of resonance for all higher frequency sounds. The movement of the basilar membrane caused by this interferes with the perception of the higher frequencies.

Further reading:

James Keener and James Sneyd, *Mathematical physiology*, Springer-Verlag, Berlin/New York, 1998. Chapter 23 of this book describes some fairly sophisticated mathematical models of the cochlea.

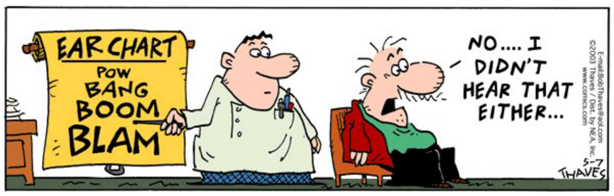
F. Richard Moore, *Psychology of hearing* [**88**].

James O. Pickles, *An introduction to the physiology of hearing* [**102**].

William A. Yost, *Fundamentals of hearing. An introduction* [**143**].

Eberhard Zwicker and H. Fastl, *Psychoacoustics: facts and models* [**144**].

1.3. Limitations of the ear



In music, frequencies are measured in Hertz (Hz), or cycles per second. The approximate range of frequencies to which the human ear responds is usually taken to be from 20 Hz to 20,000 Hz. For frequencies outside this range, there is no resonance in the basilar membrane, although sound waves of frequency lower than 20 Hz may often be *felt* rather than heard.[2] For comparison, here is a table of hearing ranges for various animals.[3]

Species	Range (Hz)
Turtle	20–1,000
Goldfish	100–2,000
Frog	100–3,000
Pigeon	200–10,000
Sparrow	250–12,000
Human	20–20,000
Chimpanzee	100-20,000
Rabbit	300–45,000
Dog	50–46,000
Cat	30–50,000
Guinea pig	150–50,000
Rat	1,000–60,000
Mouse	1,000–100,000
Bat	3,000–120,000
Dolphin (*Tursiops*)	1,000–130,000

Sound intensity is measured in *decibels* or dB. Zero decibels represents a power intensity of 10^{-12} watts per square meter, which is somewhere in the region of the weakest sound we can hear. Adding ten decibels (one *bel*) multiplies the power intensity by a factor of ten. So multiplying the power by a factor of b adds $10\log_{10}(b)$ decibels to the level of the signal. This means

[2]But see also: Tsutomi Oohashi, Emi Nishina, Norie Kawai, Yoshitaka Fuwamoto and Hiroshi Imai, *High-frequency sound above the audible range affects brain electric activity and sound perception*, Audio Engineering Society preprint No. 3207 (91st convention, New York City). In this fascinating paper, the authors describe how they recorded gamelan music with a bandwidth going up to 60 KHz. They played back the recording through a speaker system with an extra tweeter for the frequencies above 26 KHz, driven by a separate amplifier so that it could be switched on and off. They found that the EEG (Electroencephalogram) of the listeners' response, as well as the subjective rating of the recording, was affected by whether the extra tweeter was on or off, even though the listeners denied that the sound was altered by the presence of this tweeter, or that they could hear anything from the tweeter played alone. They also found that the EEG changes persisted afterwards, in the absence of the high frequency stimulation, so that long intervals were needed between sessions.

Another relevant paper is: Martin L. Lenhardt, Ruth Skellett, Peter Wang and Alex M. Clarke, *Human ultrasonic speech perception*, Science, Vol. 253, 5 July 1991, 82-85. In this paper, they report that bone-conducted ultrasonic hearing has been found capable of supporting frequency discrimination and speech detection in normal, older hearing-impaired, and profoundly deaf human subjects. They conjecture that the mechanism may have to do with the *saccule*, which is a small spherical cavity adjoining the scala vestibuli of the cochlea.

Research of James Boyk has shown that unlike other musical instruments, for the cymbal, roughly 40% of the observable energy of vibration is at frequencies between 20 kHz and 100 kHz, and showed no signs of dropping off in intensity even at the high end of this range. This research appears in *There's life above 20 kilohertz: a survey of musical-instrument spectra up to 102.4 kHz*, published on the Caltech Music Lab web site in 2000.

[3]Taken from R. Fay, *Hearing in Vertibrates. A Psychophysics Databook.* Hill-Fay Associates, Winnetka, Illinois, 1988.

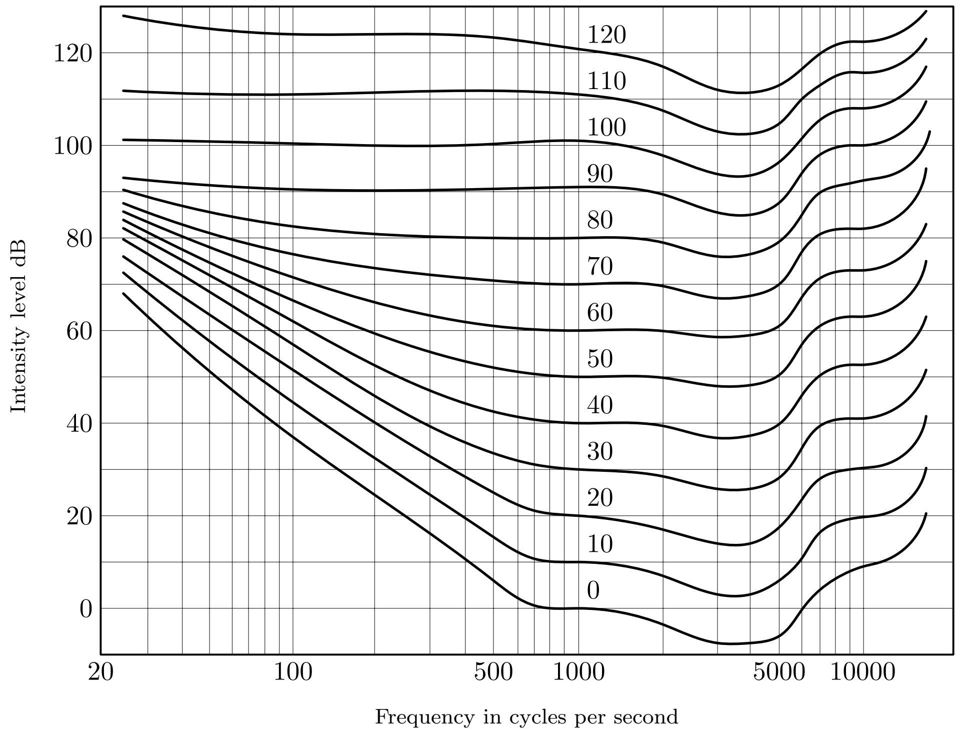
that the scale is logarithmic, and n decibels represents a power density of $10^{(n/10)-12}$ watts per square meter.

Often, decibels are used as a relative measure, so that an intensity *ratio* of ten to one represents an increase of ten decibels. As a relative measure, decibels refer to ratios of powers whether or not they directly represent sound. So for example, the power gain and the signal to noise ratio of an amplifier are measured in decibels. It is worth knowing that $\log_{10}(2)$ is roughly 0.3 (to five decimal places it is 0.30103), so that a power ratio of 2:1 represents a difference of about 3 dB. To distinguish from the relative measurement, the notation dB SPL (Sound Pressure Level) is sometimes used to refer to the absolute measurement of sound described above. It should also be mentioned that rather than using dB SPL, use is often made of a weighting curve, so that not all frequencies are given equal importance. There are three standard curves, called A, B and C. It is most common to use curve A, which has a peak at about 2000 Hz and drops off substantially to either side. Curves B and C are flatter, and only drop off at the extremes. Measurements made using curve A are quoted as dBA, or dBA SPL to be pedantic.

The *threshold of hearing* is the level of the weakest sound we can hear. Its value in decibels varies from one part of the frequency spectrum to another. Our ears are most sensitive to frequencies a little above 2000 Hz, where the threshold of hearing of the average person is a little above 0 dB. At 100 Hz the threshold is about 50 dB, and at 10,000 Hz it is about 30 dB. The average whisper is about 15–20 dB, conversation usually happens at around 60–70 dB, and the threshold of pain is around 130 dB.

The relationship between sound pressure level and perception of loudness is frequency dependent. The following graph, due to Fletcher and Munson[4] shows equal loudness curves for pure tones at various frequencies.

[4] H. Fletcher and W. J. Munson, *Loudness, its definition, measurement and calculation*, J. Acoust. Soc. Am. 5 (1933), 82–108.



The unit of loudness is the *phon*, which is defined as follows. The listener adjusts the level of the signal until it is judged to be of equal intensity to a standard 1000 Hz signal. The phon level is defined to be the signal pressure level of the 1000 Hz signal of the same loudness. The curves in this graph are called *Fletcher–Munson curves*, or *isophons*.

The amount of power in watts involved in the production of sound is very small. The clarinet at its loudest produces about one twentieth of a watt of sound, while the trombone is capable of producing up to five or six watts of sound. The average human speaking voice produces about 0.00002 watts, while a bass singer at his loudest produces about a thirtieth of a watt.

The *just noticeable difference* or *limen* is used both for sound intensity and frequency. This is usually taken to be the smallest difference between two successive tones for which a person can name correctly 75% of the time which is higher (or louder). It depends in both cases on both frequency and intensity. The just noticeable difference in frequency will be of more concern to us than the one for intensity, and the following table is taken from Pierce [**103**]. The measurements are in cents, where 1200 cents make one octave (for further details of the system of cents, see §5.4).

Frequency (Hz)	Intensity (dB) 5	10	15	20	30	40	50	60	70	80	90
31	220	150	120	97	76	70					
62	120	120	94	85	80	74	61	60			
125	100	73	57	52	46	43	48	47			
250	61	37	27	22	19	18	17	17	17	17	
550	28	19	14	12	10	9	7	6	7		
1,000	16	11	8	7	6	6	6	6	5	5	4
2,000	14	6	5	4	3	3	3	3	3	3	
4,000	10	8	7	5	5	4	4	4	4		
8,000	11	9	8	7	6	5	4	4			
11,700	12	10	7	6	6	6	5				

It is easy to see from this table that our ears are much more sensitive to small changes in frequency for higher notes than for lower ones. When referring to the above table, bear in mind that it refers to *consecutive* notes, not simultaneous ones. For simultaneous notes, the corresponding term is the *limit of discrimination*. This is the smallest difference in frequency between simultaneous notes, for which two separate pitches are heard. We shall see in §1.7 that simultaneous notes cause beats, which enable us to notice far smaller differences in frequency. This is very important to the theory of scales, because notes in a scale are designed for harmony, which is concerned with clusters of simultaneous notes. So scales are much more sensitive to very small changes in tuning than might be supposed.

Vos[5] studied the sensitivity of the ear to the exact tuning of the notes of the usual twelve tone scale, using two-voice settings from Michael Praetorius' *Musæ Sioniæ*, Part VI (1609). His conclusions were that scales in which the intervals were not more than 5 cents away from the "just" versions of the intervals (see §5.5) were all close to equally acceptable, but then with increasing difference the acceptability decreases dramatically. In view of the fact that in the modern equal tempered twelve tone system, the major third is about 14 cents away from just, these conclusions are very interesting. We shall have much more to say about this subject in Chapter 5.

Exercises

1. Power intensity is proportional to the square of amplitude. How many decibels represent a doubling of the amplitude of a signal?

2. (Multiple choice) Two independent 70 dB sound sources are heard together. How loud is the resultant sound, to the nearest dB?

(a) 140 dB, (b) 76 dB, (c) 73 dB, (d) 70 dB, (e) None of the above.

[5]J. Vos, *Subjective acceptability of various regular twelve-tone tuning systems in two-part musical fragments*, J. Acoust. Soc. Am. **83** (1988), 2383–2392.

1.4. Why sine waves?

What is the relevance of sine waves to the discussion of perception of pitch? Could we make the same discussion using some other family of periodic waves, that go up and down in a similar way?

The answer lies in the differential equation for simple harmonic motion, which we discuss in the next section. To put it briefly, the solutions to the differential equation

$$\frac{d^2y}{dt^2} = -\kappa y$$

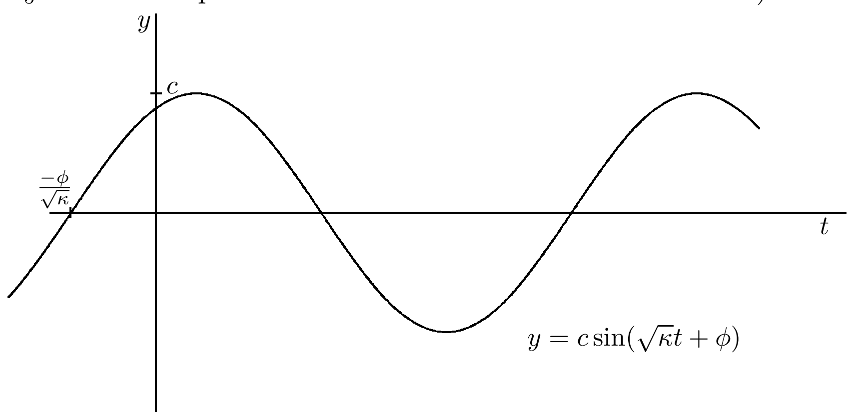
are the functions

$$y = A\cos\sqrt{\kappa}t + B\sin\sqrt{\kappa}t,$$

or equivalently

$$y = c\sin(\sqrt{\kappa}t + \phi)$$

(see §1.7 for the equivalence of these two forms of the solution).



The above differential equation represents what happens when an object is subject to a force towards an equilibrium position, the magnitude of the force being proportional to the distance from equilibrium.

In the case of the human ear, the above differential equation may be taken as a close approximation to the equation of motion of a particular point on the basilar membrane, or anywhere else along the chain of transmission between the outside air and the cochlea. Actually, this is inaccurate in several regards. The first is that we should really set up a second order partial differential equation describing the motion of the surface of the basilar membrane. This does not really affect the results of the analysis much except to explain the origins of the constant κ. The second inaccuracy is that we should really think of the motion as *forced damped harmonic motion* in which there is a damping term proportional to velocity, coming from the viscosity of the fluid and the fact that the basilar membrane is not perfectly elastic. In §§1.9–1.10, we shall see that forced damped harmonic motion is also sinusoidal, but contains a rapidly decaying transient component. There is a *resonant frequency* corresponding to the maximal response of the damped system to the incoming sine wave. The third inaccuracy is that for loud enough sounds the restoring force may be nonlinear. This will be seen to be the possible origin of some interesting acoustical phenomena. Finally, most musical notes do not consist of a single sine wave. For example, if a string is plucked, a periodic wave will result, but it will usually consist of a sum of sine waves with various amplitudes. So there will be various different peaks of amplitude of vibration of the basilar membrane, and a more complex signal is sent to the brain. The decomposition of a periodic wave as a sum of sine waves is called Fourier analysis, which is the subject of Chapter 2.

1.5. Harmonic motion

Consider a particle of mass m subject to a force F towards the equilibrium position, $y = 0$, and whose magnitude is proportional to the distance y from the equilibrium position,

$$F = -ky.$$

Here, k is just the constant of proportionality. Newton's laws of motion give us the equation

$$F = ma$$

where

$$a = \frac{d^2y}{dt^2}$$

is the acceleration of the particle and t represents time. Combining these equations, we obtain the second order differential equation

$$\frac{d^2y}{dt^2} + \frac{ky}{m} = 0. \tag{1.5.1}$$

We write $\dot{y}$ for $\dfrac{dy}{dt}$ and $\ddot{y}$ for $\dfrac{d^2y}{dt^2}$ as usual, so that this equation takes the form

$$\ddot{y} + ky/m = 0.$$

The solutions to this equation are the functions

$$y = A\cos(\sqrt{k/m}\,t) + B\sin(\sqrt{k/m}\,t). \tag{1.5.2}$$

The fact that these are the solutions of this differential equation is the explanation of why the sine wave, and not some other periodically oscillating wave, is the basis for harmonic analysis of periodic waves. For this is the differential equation governing the movement of any particular point on the basilar membrane in the cochlea, and hence governing the human perception of sound.

Exercises

1. Show that the functions (1.5.2) satisfy the differential equation (1.5.1).

2. Show that the general solution (1.5.2) to equation (1.5.1) can also be written in the form

$$y = c\sin(\sqrt{k/m}\,t + \phi).$$

Describe c and ϕ in terms of A and B. (If you get stuck, take a look at §1.7).

1.6. Vibrating strings

Consider a vibrating string, anchored at both ends. Suppose at first that the string has a heavy bead attached to the middle of it, so that the mass m of the bead is much greater than the mass of the string. Then the string exerts a force F on the bead towards the equilibrium position, and whose magnitude, at least for small displacements, is proportional to the distance y from the equilibrium position,

$$F = -ky.$$

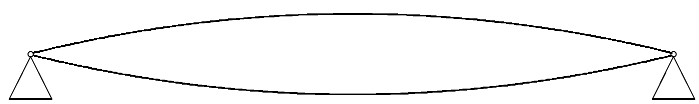
According to the last section, we obtain the differential equation

$$\frac{d^2y}{dt^2} + \frac{ky}{m} = 0.$$

whose solutions are the functions

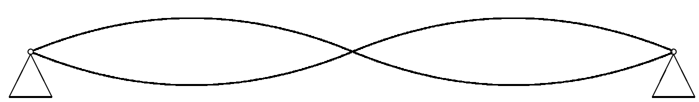
$$y = A\cos(\sqrt{k/m}\,t) + B\sin(\sqrt{k/m}\,t),$$

where the constants A and B are determined by the initial position and velocity of the string.



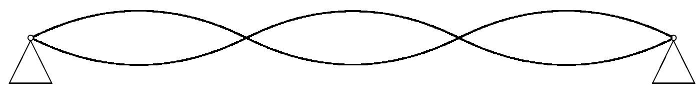
If the mass of the string is uniformly distributed, then more vibrational "modes" are possible. For example, the midpoint of the string can remain stationary while the two halves vibrate with opposite phases. On a guitar, this can be achieved by touching the midpoint of the string while plucking and then immediately releasing. The effect will be a sound exactly an octave above the natural pitch of the string, or exactly twice the frequency. The use of harmonics in this way is a common device among guitar players. If each half is vibrating with a pure sine wave then the motion of a point other than the midpoint will be described by the function

$$y = A\cos(2\sqrt{k/m}\,t) + B\sin(2\sqrt{k/m}\,t).$$



If a point exactly one third of the length of the string from one end is touched while plucking, the effect will be a sound an octave and a perfect fifth above the natural pitch of the string, or exactly three times the frequency. Again, if the three parts of the string are vibrating with a pure sine wave, with the middle third in the opposite phase to the outside two thirds, then the motion of a non-stationary point on the string will be described by the function

$$y = A\cos(3\sqrt{k/m}\,t) + B\sin(3\sqrt{k/m}\,t).$$



In general, a plucked string will vibrate with a mixture of all the modes described by multiples of the natural frequency, with various amplitudes. The amplitudes involved depend on the exact manner in which the string is plucked or struck. For example, a string struck by a hammer, as happens in a piano, will have a different set of amplitudes than that of a plucked string. The general equation of motion of a typical point on the string will be

$$y = \sum_{n=1}^{\infty} \left(A_n \cos(n\sqrt{k/m}\,t) + B_n \sin(n\sqrt{k/m}\,t) \right).$$

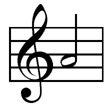
This leaves us with a problem, to which we shall return in the next two chapters. How can a string vibrate with a number of different frequencies at the same time? This forms the subject of the theory of Fourier series and the wave equation. Before we are in a position to study Fourier series, we need to understand sine waves and how they interact. This is the subject of the next section. We shall return to the subject of vibrating strings in §3.2, where we shall develop the wave equation and its solutions.

1.7. Trigonometric identities and beats

Since angles in mathematics are measured in radians, and there are 2π radians in a cycle, a sine wave with frequency ν in Hertz, peak amplitude c and *phase* ϕ will correspond to a sine wave of the form

$$c\sin(2\pi\nu t + \phi). \tag{1.7.1}$$

The quantity $\omega = 2\pi\nu$ is called the *angular velocity.* The role of the angle ϕ is to tell us where the sine wave crosses the time axis (look back at the graph in §1.4). For example, a cosine wave is related to a sine wave by the equation $\cos x = \sin(x + \frac{\pi}{2})$, so a cosine wave is really just a sine wave with a different phase.



440 Hz

For example, modern concert pitch[6] places the note A above middle C at 440 Hz so this would be represented by a wave of the form

$$c \sin(880\pi t + \phi).$$

This can be converted to a linear combination of sines and cosines using the standard formulas for the sine and cosine of a sum:

$$\sin(A + B) = \sin A \cos B + \cos A \sin B \tag{1.7.2}$$

$$\cos(A + B) = \cos A \cos B - \sin A \sin B. \tag{1.7.3}$$

So we have

$$c \sin(\omega t + \phi) = a \cos \omega t + b \sin \omega t$$

where

$$a = c \sin \phi \qquad b = c \cos \phi.$$

Conversely, given a and b, c and ϕ can be obtained via

$$c = \sqrt{a^2 + b^2} \qquad \tan \phi = a/b.$$

What happens when two pure sine or cosine waves are played at the same time? For example, why is it that when two very close notes are played simultaneously, we hear "beats"? Since this is the method by which strings on a piano are tuned, it is important to understand the origins of these beats.

The answer to this question also lies in the trigonometric identities (1.7.2) and (1.7.3). Since $\sin(-B) = -\sin B$ and $\cos(-B) = \cos B$, replacing B by $-B$ in equations (1.7.2) and (1.7.3) gives

$$\sin(A - B) = \sin A \cos B - \cos A \sin B \tag{1.7.4}$$

$$\cos(A - B) = \cos A \cos B + \sin A \sin B. \tag{1.7.5}$$

Adding equations (1.7.2) and (1.7.4)

$$\sin(A + B) + \sin(A - B) = 2 \sin A \cos B \tag{1.7.6}$$

which may be rewritten as

$$\sin A \cos B = \tfrac{1}{2}(\sin(A + B) + \sin(A - B)). \tag{1.7.7}$$

[6] Historically, this was adopted as the U.S.A. Standard Pitch in 1925, and in May 1939 an international conference in London agreed that this should be adopted as the modern concert pitch. Before that time, a variety of standard frequencies were used. For example, in the time of Mozart, the note A had a value closer to 422 Hz, a little under a semitone flat to modern ears. Before this time, in the Baroque and earlier, there was even more variation. For example, in Tudor Britain, secular vocal pitch was much the same as modern concert pitch, while domestic keyboard pitch was about three semitones lower and church music pitch was more than two semitones higher.

Similarly, adding and subtracting equations (1.7.3) and (1.7.5) gives

$$\cos(A+B)+\cos(A-B)=2\cos A\cos B \tag{1.7.8}$$

$$\cos(A-B)-\cos(A+B)=2\sin A\sin B, \tag{1.7.9}$$

or

$$\cos A\cos B=\tfrac{1}{2}(\cos(A+B)+\cos(A-B)) \tag{1.7.10}$$

$$\sin A\sin B=\tfrac{1}{2}(\cos(A-B)-\cos(A+B)). \tag{1.7.11}$$

This enables us to write any product of sines and cosines as a sum or difference of sines and cosines. So for example, if we wanted to integrate a product of sines and cosines, this would enable us to do so.

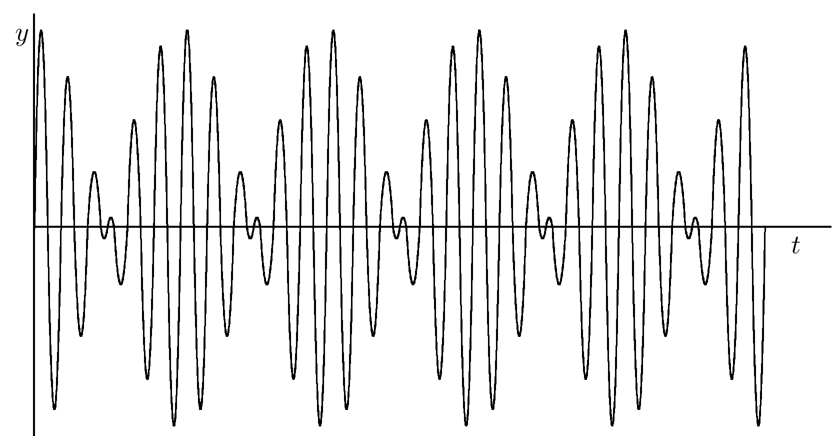
We are actually interested in the opposite process. So we set $u=A+B$ and $v=A-B$. Solving for A and B, this gives $A=\frac{1}{2}(u+v)$ and $B=\frac{1}{2}(u-v)$. Substituting in equations (1.7.6), (1.7.8) and (1.7.9), we obtain

$$\sin u+\sin v=2\sin\tfrac{1}{2}(u+v)\cos\tfrac{1}{2}(u-v) \tag{1.7.12}$$

$$\cos u+\cos v=2\cos\tfrac{1}{2}(u+v)\cos\tfrac{1}{2}(u-v) \tag{1.7.13}$$

$$\cos u-\cos v=2\sin\tfrac{1}{2}(u+v)\sin\tfrac{1}{2}(u-v) \tag{1.7.14}$$

This enables us to write any sum or difference of sine waves and cosine waves as a product of sines and cosines. Exercise 1 at the end of this section explains what to do if there are mixed sines and cosines.



$y=\sin(12t)+\sin(10t)=2\sin(11t)\cos(t)$

So for example, suppose that a piano tuner has tuned one of the three strings corresponding to the note A above middle C to 440 Hz. The second string is still out of tune, so that it resonates at 436 Hz. The third is being damped so as not to interfere with the tuning of the second string. Ignoring phase and amplitude for a moment, the two strings together will sound as

$$\sin(880\pi t)+\sin(872\pi t).$$

Using equation (1.7.12), we may rewrite this sum as

$$2\sin(876\pi t)\cos(4\pi t).$$

This means that we perceive the combined effect as a sine wave with frequency 438 Hz, the average of the frequencies of the two strings, but with the amplitude modulated by a slow cosine wave with frequency 2 Hz, or half the difference between the frequencies of the two strings. This modulation is what we perceive as beats. The amplitude of the modulating cosine wave has two peaks per cycle, so the number of beats per second will be four, not two. So the number of beats per second is exactly the difference between the two frequencies. The piano tuner tunes the second string to the first by tuning out the beats, namely by adjusting the string so that the beats slow down to a standstill.

If we wish to include terms for phase and amplitude, we write

$$c\sin(880\pi t + \phi) + c\sin(872\pi t + \phi').$$

where the angles ϕ and ϕ' represent the phases of the two strings. This gets rewritten as

$$2c\sin(876\pi t + \tfrac{1}{2}(\phi + \phi'))\cos(4\pi t + \tfrac{1}{2}(\phi - \phi')),$$

so this equation can be used to understand the relationship between the phase of the beats and the phases of the original sine waves.

If the amplitudes are different, then the beats will not be so pronounced because part of the louder note is "left over." This prevents the amplitude going to zero when the modulating cosine takes the value zero.

Exercises

1. Use the equation $\cos\theta = \sin(\pi/2 + \theta)$ and equations (1.7.12)–(1.7.13) to express $\sin u + \cos v$ as a product of trigonometric functions.

2. A piano tuner comparing two of the three strings on the same note of a piano hears five beats a second. If one of the two notes is concert pitch A (440 Hz), what are the possibilities for the frequency of vibration of the other string?

3. Evaluate $\displaystyle\int_0^{\pi/2} \sin(3x)\sin(4x)\,dx$.

4. (a) Setting $A = B = \theta$ in formula (1.7.10) gives the double angle formula

$$\cos^2\theta = \tfrac{1}{2}(1 + \cos(2\theta)). \qquad (1.7.15)$$

Draw graphs of the functions $\cos^2\theta$ and $\cos(2\theta)$. Try to understand formula (1.7.15) in terms of these graphs.

(b) Setting $A = B = \theta$ in formula (1.7.11) gives the double angle formula

$$\sin^2\theta = \tfrac{1}{2}(1 - \cos(2\theta)). \qquad (1.7.16)$$

Draw graphs of the functions $\sin^2\theta$ and $\cos(2\theta)$. Try to understand formula (1.7.16) in terms of these graphs.

5. In the formula (1.7.1), the factor c is called the *peak amplitude*, because it determines the highest point on the waveform. In sound engineering, it is often more useful to know the *root mean square*, or RMS amplitude, because this is what determines things like power consumption. The RMS amplitude is calculated by integrating the square of the value over one cycle, dividing by the length of the cycle to obtain the mean square, and then taking the square root. For a pure sine wave given by formula (1.7.1), show that the RMS amplitude is given by

$$\sqrt{\nu \int_0^{\frac{1}{\nu}} [c \sin(2\pi\nu t + \phi)]^2 \, dt} = \frac{c}{\sqrt{2}}.$$

6. Use equation (1.7.11) to write $\sin kt \, \sin \frac{1}{2}t$ as $\frac{1}{2}(\cos(k - \frac{1}{2})t - \cos(k + \frac{1}{2})t)$. Show that

$$\sum_{k=1}^{n} \sin kt = \frac{\cos \frac{1}{2}t - \cos(n + \frac{1}{2})t}{2 \sin \frac{1}{2}t} = \frac{\sin \frac{1}{2}(n+1)t \, \sin \frac{1}{2}nt}{\sin \frac{1}{2}t}. \tag{1.7.17}$$

Similarly, show that

$$\sum_{k=1}^{n} \cos kt = \frac{\sin(n + \frac{1}{2})t - \sin \frac{1}{2}t}{2 \sin \frac{1}{2}t} = \frac{\cos \frac{1}{2}(n+1)t \, \sin \frac{1}{2}nt}{\sin \frac{1}{2}t}. \tag{1.7.18}$$

7. Two pure sine waves are sounded. One has frequency slightly greater or slightly less than twice that of the other. Would you expect to hear beats?

1.8. Superposition

Superposing two sounds corresponds to adding the corresponding wave functions. This is part of the concept of *linearity*. In general, a system is linear if two conditions are satisfied. The first, *superposition*, is that two simultaneous independent input signals should give rise to the sum of the two outputs. The second condition, *homogeneity*, says that magnifying the input level by a constant factor should multiply the output level by the same constant factor.

Superposing harmonic motions of the same frequency works as follows. Two simple harmonic motions with the same frequency, but possibly different amplitudes and phases, always add up to give another simple harmonic motion with the same frequency. We saw some examples of this in the last section. In this section, we see that there is an easy graphical method for carrying this out in practice.

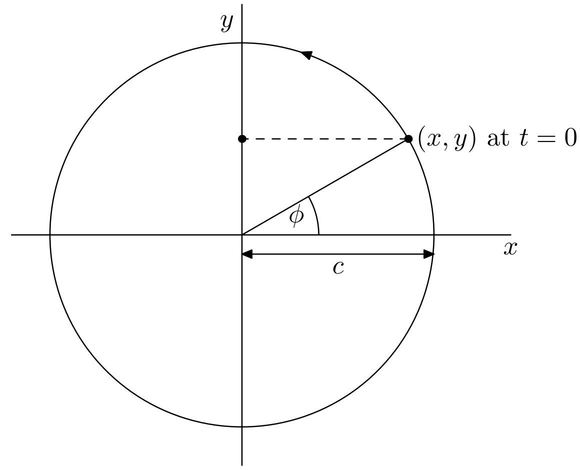
Consider a sine wave of the form $c \sin(\omega t + \phi)$ where $\omega = 2\pi\nu$. This may be regarded as the y-component of circular motion of the form

$$x = c \cos(\omega t + \phi)$$
$$y = c \sin(\omega t + \phi).$$

Since $\sin^2 \theta + \cos^2 \theta = 1$, squaring and adding these equations shows that the point (x, y) lies on the circle

$$x^2 + y^2 = c^2$$

with radius c, centered at the origin. As t varies, the point (x, y) travels counterclockwise round this circle ν times in each second, so ν is really measuring the number of cycles per second around the origin, and ω is measuring the angular velocity in radians per second. The phase ϕ is the angle, measured counterclockwise from the positive x-axis, subtended by the line from $(0,0)$ to (x,y) when $t=0$.

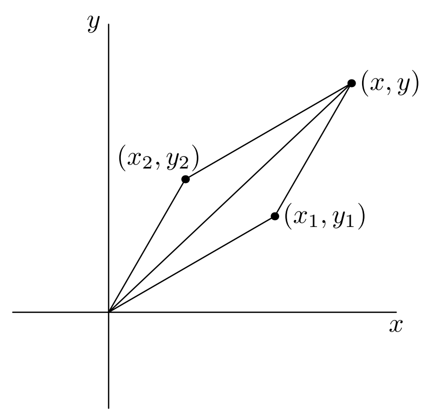


Now suppose that we are given two sine waves of the same frequency, say $c_1\sin(\omega t+\phi_1)$ and $c_2\sin(\omega t+\phi_2)$. The corresponding vectors at $t=0$ are

$$\begin{aligned}(x_1,y_1)&=(c_1\cos\phi_1,c_1\sin\phi_1)\\(x_2,y_2)&=(c_2\cos\phi_2,c_2\sin\phi_2).\end{aligned}$$

To superpose (i.e., add) these sine waves, we simply add these vectors to give

$$\begin{aligned}(x,y)&=(c_1\cos\phi_1+c_2\cos\phi_2,c_1\sin\phi_1+c_2\sin\phi_2)\\&=(c\cos\phi,c\sin\phi).\end{aligned}$$



We draw a copy of the line segment $(0,0)$ to (x_1,y_1) starting at (x_2,y_2), and a copy of the line segment $(0,0)$ to (x_2,y_2) starting at (x_1,y_1), to form a parallelogram. The amplitude c is the length of the diagonal line drawn from the origin to the far corner (x,y) of the parallelogram formed this way. The angle ϕ is the angle subtended by this line, measured as usual counterclockwise from the x-axis.

Exercises

1. Write the following expressions in the form $c\sin(2\pi\nu t+\phi)$:

(i) $\cos(2\pi t)$

(ii) $\sin(2\pi t)+\cos(2\pi t)$

(iii) $2\sin(4\pi t+\pi/6)-\sin(4\pi t+\pi/2)$.

2. Read Appendix C. Use equation (C.1) to interpret the graphical method described in this section as motion in the complex plane of the form

$$z = ce^{i(\omega t+\phi)}.$$

1.9. Damped harmonic motion

Damped harmonic motion arises when in addition to the restoring force $F=-ky$, there is a frictional force proportional to velocity,

$$F=-ky-\mu\dot{y}.$$

For positive values of μ, the extra term damps the motion, while for negative values of μ it promotes or forces the harmonic motion. In this case, the differential equation we obtain is

$$m\ddot{y}+\mu\dot{y}+ky=0. \tag{1.9.1}$$

This is what is called a linear second order differential equation with constant coefficients. To solve such an equation, we look for solutions of the form

$$y=e^{\alpha t}.$$

Then $\dot{y}=\alpha e^{\alpha t}$ and $\ddot{y}=\alpha^2 e^{\alpha t}$. So for y to satisfy the original differential equation, α has to satisfy the *auxiliary equation*

$$mY^2+\mu Y+k=0. \tag{1.9.2}$$

If the quadratic equation (1.9.2) has two different solutions, $Y=\alpha$ and $Y=\beta$, then $y=e^{\alpha t}$ and $y=e^{\beta t}$ are solutions of (1.9.1). Since equation (1.9.1) is linear, this implies that any combination of the form

$$y=Ae^{\alpha t}+Be^{\beta t}$$

is also a solution. The *discriminant* of the auxiliary equation (1.9.2) is

$$\Delta=\mu^2-4mk.$$

If $\Delta>0$, corresponding to large damping or forcing term, then the solutions to the auxiliary equation are

$$\alpha=(-\mu+\sqrt{\Delta})/2m$$
$$\beta=(-\mu-\sqrt{\Delta})/2m,$$

and so the solutions to the differential equation (1.9.1) are

$$y=Ae^{(-\mu+\sqrt{\Delta})t/2m}+Be^{(-\mu-\sqrt{\Delta})t/2m}. \tag{1.9.3}$$

In this case, the motion is so damped that no sine waves can be discerned. The system is then said to be *overdamped*, and the resulting motion is called *dead beat*.

If $\Delta < 0$, as happens when the damping or forcing term is small, then the system is said to be *underdamped*. In this case, the auxiliary equation (1.9.2) has no real solutions because Δ has no real square roots. But $-\Delta$ is positive, and so it has a square root. In this case, the solutions to the auxiliary equation are

$$\begin{aligned}\alpha &= (-\mu + i\sqrt{-\Delta})/2m \\ \beta &= (-\mu - i\sqrt{-\Delta})/2m,\end{aligned}$$

where $i = \sqrt{-1}$. See Appendix C for a brief introduction to complex numbers. So the solutions to the original differential equation are

$$y = e^{-\mu t/2m}(Ae^{it\sqrt{-\Delta}/2m} + Be^{-it\sqrt{-\Delta}/2m}).$$

We are really interested in real solutions. To this end, we use relation (C.1) to write this as

$$y = e^{-\mu t/2m}((A+B)\cos(t\sqrt{-\Delta}/2m) + i(A-B)\sin(t\sqrt{-\Delta}/2m)).$$

So we obtain real solutions by taking $A' = A + B$ and $B' = i(A - B)$ to be real numbers, giving

$$y = e^{-\mu t/2m}(A' \sin(t\sqrt{-\Delta}/2m) + B' \cos(t\sqrt{-\Delta}/2m)). \tag{1.9.4}$$

The interpretation of this is harmonic motion with a damping factor of $e^{-\mu t/2m}$.

The special case $\Delta = 0$ has solutions

$$y = (At + B)e^{-\mu t/2m}. \tag{1.9.5}$$

This borderline case resembles the case $\Delta > 0$, inasmuch as harmonic motion is not apparent. Such a system is said to be *critically damped*.

Examples

1. The equation

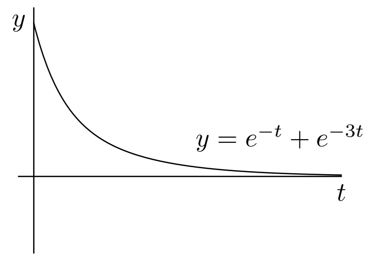
$$\ddot{y} + 4\dot{y} + 3y = 0 \tag{1.9.6}$$

is overdamped. The auxiliary equation

$$Y^2 + 4Y + 3 = 0$$

factors as $(Y + 1)(Y + 3) = 0$, so it has roots $Y = -1$ and $Y = -3$. It follows that the solutions of (1.9.6) are given by

$$y = Ae^{-t} + Be^{-3t}.$$



2. The equation

$$\ddot{y} + 2\dot{y} + 26y = 0 \tag{1.9.7}$$

is underdamped. The auxiliary equation is

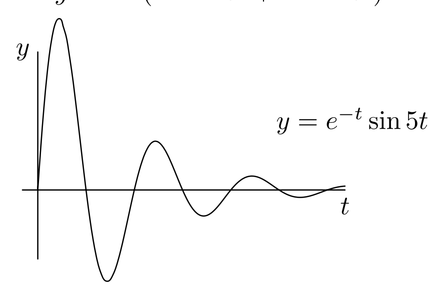
$$Y^2 + 2Y + 26 = 0.$$

Completing the square gives $(Y+1)^2 + 25 = 0$, so the solutions are $Y = -1 \pm 5i$. It follows that the solutions of (1.9.7) are given by

$$y = e^{-t}(Ae^{5it} + Be^{-5it}),$$

or

$$y = e^{-t}(A' \cos 5t + B' \sin 5t). \tag{1.9.8}$$



3. The equation

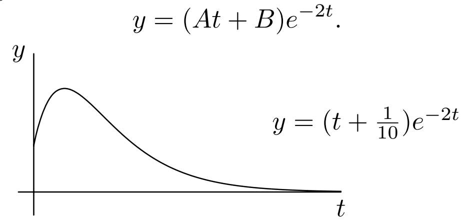
$$\ddot{y} + 4\dot{y} + 4y = 0 \tag{1.9.9}$$

is critically damped. The auxiliary equation

$$Y^2 + 4Y + 4 = 0$$

factors as $(Y+2)^2 = 0$, so the only solution is $Y = -2$. It follows that the solutions of (1.9.9) are given by

$$y = (At + B)e^{-2t}.$$



Exercises

1. Show that if $\Delta = \mu^2 - 4mk > 0$ then the functions (1.9.3) are real solutions of the differential equation (1.9.1).

2. Show that if $\Delta = \mu^2 - 4mk < 0$ then the functions (1.9.4) are real solutions of the differential equation (1.9.1).

3. Show that if $\Delta = \mu^2 - 4mk = 0$ then the auxiliary equation (1.9.2) is a perfect square, and the functions (1.9.5) satisfy the differential equation (1.9.1).

1.10. Resonance

Forced harmonic motion is where there is a forcing term $f(t)$ (often taken to be periodic) added into equation (1.9.1) to give an equation of the form

$$m\ddot{y} + \mu\dot{y} + ky = f(t). \tag{1.10.1}$$

This represents a damped system with an external stimulus $f(t)$ applied to it. We are particularly interested in the case where $f(t)$ is a sine wave, because this represents forced harmonic motion. Forced harmonic motion is responsible for the production of sound in most musical instruments, as well as the perception of sound in the cochlea. We shall see that forced harmonic motion is what gives rise to the phenomenon of *resonance*.

There are two steps to the solution of the equation. The first is to find the general solution to equation (1.9.1) without the forcing term, as described in §1.9, to give the *complementary function*. The second step is to find by any method, such as guessing, a single solution to equation (1.10.1). This is called a *particular integral*. Then the general solution to the equation (1.10.1) is the sum of the particular integral and the complementary function.

Examples

1. Consider the equation

$$\ddot{y} + 4\dot{y} + 5y = 10t^2 - 1. \tag{1.10.2}$$

We look for a particular integral of the form $y = at^2 + bt + c$. Differentiating, we get $\dot{y} = 2at + b$ and $\ddot{y} = 2a$. Plugging these into (1.10.2) gives

$$2a + 4(2at + b) + 5(at^2 + bt + c) = 10t^2 + t - 3.$$

Comparing coefficients of t^2 gives $5a = 10$ or $a = 2$. Then comparing coefficients of t gives $8a + 5b = 1$, so $b = -3$. Finally, comparing constant terms gives $2a + 4b + 5c = -3$, so $c = 1$. So we get a particular integral of $y = 2t^2 - 3t + 1$. Adding the complementary function (1.9.8), we find that the general solution to (1.10.2) is given by

$$y = 2t^2 - 3t + 1 + e^{-2t}(A' \cos t + B' \sin t).$$

2. As a more interesting example, to solve

$$\ddot{y} + 4\dot{y} + 5y = \sin 2t, \tag{1.10.3}$$

we look for a particular integral of the form

$$y = a \cos 2t + b \sin 2t.$$

Equating coefficients of $\cos 2t$ and $\sin 2t$ we get two equations:

$$\begin{aligned} -8a + b &= 1 \\ a + 8b &= 0. \end{aligned}$$

Solving these equations, we get $a = -\frac{8}{65}$, $b = \frac{1}{65}$. So the general solution to (1.10.3) is

$$y = \frac{\sin 2t - 8 \cos 2t}{65} + e^{-2t}(A' \cos t + B' \sin t).$$

The case of forced harmonic motion of interest to us is the equation

$$m\ddot{y} + \mu\dot{y} + ky = R\cos(\omega t + \phi). \tag{1.10.4}$$

This represents a damped harmonic motion (see §1.9) with forcing term of amplitude R and angular velocity ω.

We could proceed as above to look for a particular integral of the form

$$y = a\cos\omega t + b\sin\omega t$$

and proceed as in the second example above. However, we can simplify the calculation by using complex numbers (see Appendix C). Since this differential equation is linear, and since

$$Re^{i(\omega t+\phi)} = R(\cos(\omega t + \phi) + i\sin(\omega t + \phi))$$

it will be enough to find a particular integral for the equation

$$m\ddot{y} + \mu\dot{y} + ky = Re^{i(\omega t+\phi)}, \tag{1.10.5}$$

which represents a complex forcing term with unit amplitude and angular velocity ω. Then we take the real part to get a solution to equation (1.10.4).

We look for solutions of equation (1.10.5) of the form $y = Ae^{i(\omega t+\phi)}$, with A to be determined. We have $\dot{y} = Ai\omega e^{i(\omega t+\phi)}$ and $\ddot{y} = -A\omega^2 e^{i(\omega t+\phi)}$. So plugging into equation (1.10.5) and dividing by $e^{i(\omega t+\phi)}$, we get

$$A(-m\omega^2 + i\mu\omega + k) = R$$

or

$$A = \frac{R}{-m\omega^2 + i\mu\omega + k}.$$

So the particular integral, which actually represents the eventual "steady state" solution to the equation since the complementary function is decaying, is given by

$$y = \frac{Re^{i(\omega t+\phi)}}{-m\omega^2 + i\mu\omega + k}.$$

The bottom of this expression is a complex constant, and so this solution moves around a circle in the complex plane. The real part is then a sine wave with the radius of the circle as amplitude and with a phase determined by the argument of the bottom.

The amplitude of the resulting vibration, and therefore the degree of resonance (since we started with a forcing term of unit amplitude) is given by taking the absolute value of this solution,

$$|y| = \frac{R}{\sqrt{(k - m\omega^2)^2 + \mu^2\omega^2}}.$$

This amplitude magnification reaches its maximum when the derivative of $(k - m\omega^2)^2 + \mu^2\omega^2$ vanishes, namely when

$$\omega = \sqrt{\frac{k}{m} + \frac{\mu^2}{2m^2}},$$

when we have amplitude $mR/(\mu\sqrt{km+3\mu^2/4})$. The above value of ω is called the *resonant frequency* of the system. Note that this value of ω is slightly greater than the value which one may expect from Equation (1.9.4) for the complementary function:

$$\omega = \frac{\sqrt{-\Delta}}{2m} = \sqrt{\frac{k}{m} - \frac{\mu^2}{4m^2}},$$

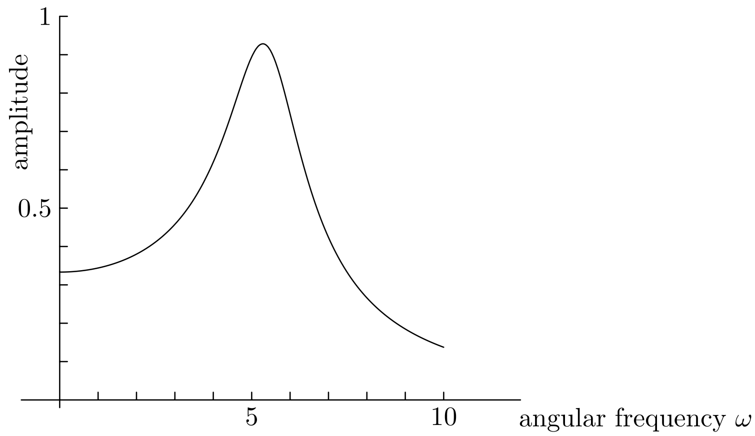
or even than the value of ω for the corresponding undamped system:

$$\omega = \sqrt{\frac{k}{m}}.$$

Example. Consider the forced, underdamped equation

$$\ddot{y} + 2\dot{y} + 30y = 10\sin\omega t.$$

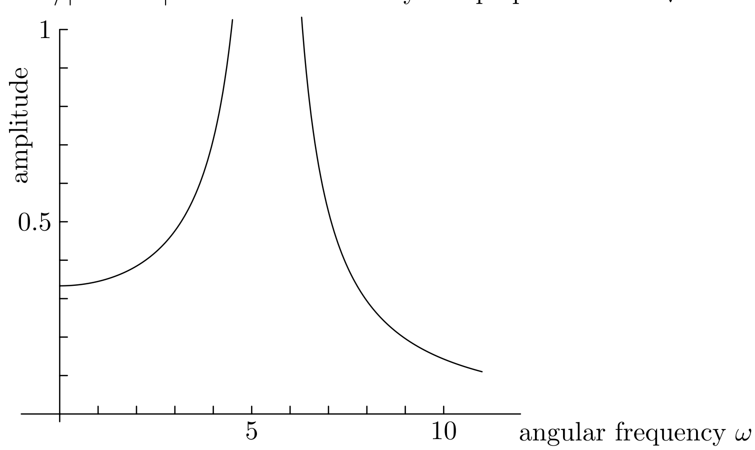
The above formula tells us that the amplitude of the resulting steady state sine wave solution is $10/\sqrt{900 - 56\omega^2 + \omega^4}$, which has its maximum value at $\omega = \sqrt{31}$.



Without the damping term, the amplitude of the steady state solution to the equation

$$\ddot{y} + 30y = 10\sin\omega t,$$

is equal to $10/|30-\omega^2|$. It has an "infinitely sharp" peak at $\omega = \sqrt{30}$.



At this stage, it seems appropriate to introduce the terms *resonant frequency* and *bandwidth* for a resonant system. The resonant frequency is the frequency for which the amplitude of the steady state solution is maximal. Bandwidth is a vague term, used to describe the width of the peak in the above graphs. So in the damped example above, we might want to describe the bandwidth as being from roughly $4\frac{1}{2}$ to $6\frac{1}{2}$, while for the undamped example it would be somewhat wider. Sometimes, the term is made precise by taking the interval between the two points either side of the peak where the amplitude is $1/\sqrt{2}$ times that of the peak. Since power is proportional to square of amplitude, this corresponds to a factor of two in the power, or a difference of $10\log_{10}(2)$ dB, or roughly 3 dB.

CHAPTER 2

Fourier theory

To be sung to the tune of Gilbert and Sullivan's *Modern Major General*:

I am the very model of a genius mathematical,
  For I can do mechanics, both dynamical and statical,
Or integrate a function round a contour in the complex plane,
  Yes, even if it goes off to infinity and back again;
Oh, I know when a detailed proof's required and when a guess'll do
  I know about the functions of Laguerre and those of Bessel too,
I've finished every tripos question back to 1948;
  There ain't a function you can name that I can't differentiate!
  There ain't a function you can name that he can't differentiate [Tris]

I've read the text books and I can extremely quickly tell you where
  To look to find Green's Theorem or the Principal of d'Alembert
Or I can work out Bayes' rule when the loss is not Quadratical
  In short I am the model of a genius mathematical!
For he can work out Bayes' rule when the loss is not Quadratical
  In short he is the model of a genius mathematical!

Oh, I can tell in seconds if a graph is Hamiltonian,
  And I can tell you if a proof of $4CC$'s a phoney 'un
I read up all the journals and I'm ready with the latest news,
  And very good advice about the Part II lectures you should choose.
Oh, I can do numerical analysis without a pause,
  Or comment on the far-reaching significance of Newton's laws
I know when polynomials are soluble by radicals,
  And I can reel off simple groups, especially sporadicals.
For he can reel off simple groups, especially sporadicals [Tris]

Oh, I like relativity and know about fast moving clocks
  And I know what you have to do to get round Russel's paradox
In short, I think you'll find concerning all things problematical
  I am the very model of a genius mathematical!
In short we think you'll find concerning all things problematical
  He is the very model of a genius mathematical!

Oh, I know when a matrix will be diagonalisable
  And I can draw Greek letters so that they are recognizable
And I can find the inverse of a non-zero quaternion
  I've made a model of a rhombicosidodecahedron;
Oh, I can quote the theorem of the separating hyperplane
  I've read MacLane and Birkoff not to mention Birkoff and MacLane
My understanding of vorticity is not a hazy 'un
  And I know why you should (and why you shouldn't) be a Bayesian!
  For he knows why you should (and why you shouldn't) be a Bayesian! [Tris]

I'm not deterred by residues and really I am quite at ease
  When dealing with essential isolated singularities,
In fact as everyone agrees (and most are quite emphatical)
  I am the very model of a genius mathematical!
In fact as everyone agrees (and most are quite emphatical)
  He is the very model of a genius mathematical!

—from CUYHA songbook, Cambridge (privately distributed) 1976.

2.1. Introduction

How can a string vibrate with a number of different frequencies at the same time? This problem occupied the minds of many of the great mathematicians and musicians of the seventeenth and eighteenth century. Among the people whose work contributed to the solution of this problem are Marin Mersenne, Daniel Bernoulli, the Bach family, Jean-le-Rond d'Alembert, Leonhard Euler, and Jean Baptiste Joseph Fourier.

In this chapter, we discuss Fourier's theory of harmonic analysis. This is the decomposition of a periodic wave into a (usually infinite) sum of sines and cosines. The frequencies involved are the integer multiples of the fundamental frequency of the periodic wave, and each has an amplitude which can be determined as an integral. A superb book on Fourier series and their continuous frequency spectrum counterpart, Fourier integrals, is Tom Körner [**68**]. The reader should be warned, however, that the level of sophistication of Körner's book is much greater than the level of these notes.

2.2. Fourier coefficients

Engraving of Jean Baptiste Joseph Fourier
(1768–1850) by Boilly (1823)
Académie des Sciences, Paris

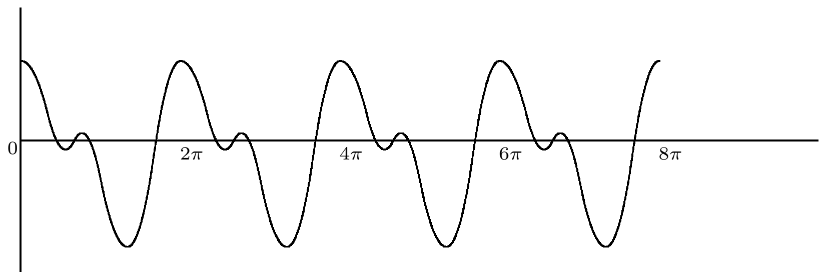
Fourier introduced the idea that periodic functions can be analyzed by using trigonometric series as follows.[1] The functions $\cos\theta$ and $\sin\theta$ are *periodic* with period 2π, in the sense that they satisfy

$$\cos(\theta + 2\pi) = \cos\theta$$
$$\sin(\theta + 2\pi) = \sin\theta.$$

In other words, translating by 2π along the θ axis leaves these functions unaffected. There are many other functions $f(\theta)$ which are periodic of period 2π, meaning that they satisfy the equation

$$f(\theta + 2\pi) = f(\theta).$$

We need only specify the function f on the half-open interval $[0, 2\pi)$ in any way we please, and then the above equation determines the value at all other values of θ.



A periodic function with period 2π

Other examples of such functions are the constant functions, and the functions $\cos(n\theta)$ and $\sin(n\theta)$ for any positive integer n. Negative values of n give us no more, since

$$\cos(-n\theta) = \cos(n\theta),$$
$$\sin(-n\theta) = -\sin(n\theta).$$

More generally, we can write down any series of the form

$$f(\theta) = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n\cos(n\theta) + b_n\sin(n\theta)). \tag{2.2.1}$$

Here, $\frac{1}{2}a_0$ is just a constant—the reason for the factor of $\frac{1}{2}$ will be explained in due course. Such a series is called a *trigonometric series*. Provided that there are no convergence problems, such a series will always define a function satisfying $f(\theta + 2\pi) = f(\theta)$.

[1]The basic ideas behind Fourier series were introduced in Jean Baptiste Joseph Fourier, *La théorie analytique de la chaleur*, F. Didot, Paris, 1822. Fourier was born in Auxerre, France in 1768 as the son of a taylor. He was orphaned in childhood and was educated by a school run by the Benedictine order. He was politically active during the French Revolution, and was almost executed. After the revolution, he studied in the then new Ecole Normale in Paris with teachers such as Lagrange, Monge and Laplace. In 1822, with the publication of the work mentioned above, he was elected *secretaire perpetuel* of the Académie des Sciences in Paris. Following this, his role seems principally to have been to encourage younger mathematicians such as Dirichlet, Liouville and Sturm, until his death in 1830.

The question which naturally arises at this stage is, to what extent can we find a trigonometric series whose sum is equal to a given periodic function? To begin to answer this question, we first ask: given a function defined by a trigonometric series, how can the coefficients a_n and b_n be recovered?

The answer lies in the formulae (for $m \geq 0$ and $n \geq 0$)

$$\int_0^{2\pi} \cos(m\theta)\sin(n\theta)\,dt = 0 \tag{2.2.2}$$

$$\int_0^{2\pi} \cos(m\theta)\cos(n\theta)\,dt = \begin{cases} 2\pi & \text{if } m = n = 0 \\ \pi & \text{if } m = n > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.2.3}$$

$$\int_0^{2\pi} \sin(m\theta)\sin(n\theta)\,dt = \begin{cases} \pi & \text{if } m = n \\ 0 & \text{otherwise} \end{cases} \tag{2.2.4}$$

These equations can be proved by using equations (1.7.7)–(1.7.11) to rewrite the product of trigonometric functions inside the integral as a sum before integrating.[2] The extra factor of two in (2.2.3) for $m = n = 0$ will explain the factor of $\frac{1}{2}$ in front of a_0 in (2.2.1).

This suggests that in order to find the coefficent a_m, we multiply $f(\theta)$ by $\cos(m\theta)$ and integrate. Let us see what happens when we apply this process to equation (2.2.1). Provided we can pass the integral through the infinite sum, only one term gives a nonzero contribution. So for $m > 0$ we have

$$\int_0^{2\pi} \cos(m\theta)f(\theta)\,d\theta = \int_0^{2\pi} \cos(m\theta)\Big(\tfrac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n\cos(n\theta) + b_n\sin(n\theta))\Big)\,d\theta$$

$$= \tfrac{1}{2}a_0\int_0^{2\pi}\cos(m\theta)\,d\theta + \sum_{n=1}^{\infty}\Big(a_n\int_0^{2\pi}\cos(m\theta)\cos(n\theta)\,d\theta + b_n\int_0^{2\pi}\cos(m\theta)\sin(n\theta)\,d\theta\Big)$$

$$= \pi a_m.$$

Thus we obtain, for $m > 0$,

$$\boxed{a_m = \frac{1}{\pi}\int_0^{2\pi}\cos(m\theta)f(\theta)\,d\theta.} \tag{2.2.5}$$

A standard theorem of analysis says that provided the sum converges absolutely (in other words, if the sum of the absolute values converges) then the integral can be passed through the infinite sum in this way. Under the same conditions, we obtain for $m > 0$

$$\boxed{b_m = \frac{1}{\pi}\int_0^{2\pi}\sin(m\theta)f(\theta)\,d\theta.} \tag{2.2.6}$$

[2]The relations (2.2.2)–(2.2.4) are sometimes called *orthogonality relations*. The idea is that the integrable periodic functions form an infinite dimensional vector space with an inner product given by $\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(\theta)g(\theta)\,d\theta$. With respect to this inner product, the functions $\sin(m\theta)$ $(m > 0)$ and $\cos(m\theta)$ $(m \geq 0)$ are *orthogonal*, or perpendicular.

The functions a_m and b_m defined by equations (2.2.5) and (2.2.6) are called the *Fourier coefficients* of the function $f(\theta)$.

We can now explain the appearance of the coefficient of one half in front of the a_0 in equation (2.2.1). Namely, since π is one half of 2π and $\cos(0\theta) = 1$ we have

$$a_0 = \frac{1}{\pi}\int_0^{2\pi} \cos(0\theta) f(\theta)\, d\theta \tag{2.2.7}$$

which means that the formula (2.2.5) for the coefficient a_m holds for all $m \ge 0$.

It would be nice to think that when we use equations (2.2.5), (2.2.6) and (2.2.7) to define a_m and b_m, the right hand side of equation (2.2.1) always converges to $f(\theta)$. This is true for nice enough functions f, but unfortunately, not for all functions f. In Section 2.4, we investigate conditions on f which ensure that this happens.

Of course, any interval of length 2π, representing one complete period, may be used instead of integrating from 0 to 2π. It is sometimes more convenient, for example, to integrate from $-\pi$ to π:

$$a_m = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos(m\theta) f(\theta)\, d\theta$$

$$b_m = \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(m\theta) f(\theta)\, d\theta.$$

In practice, the variable θ will not quite correspond to time, because the period is not necessarily 2π seconds. If the fundamental frequency (the reciprocal of the period) is ν then the correct substitution is $\theta = 2\pi\nu t$. Setting $F(t) = f(2\pi\nu t) = f(\theta)$ and substituting gives a Fourier series of the form

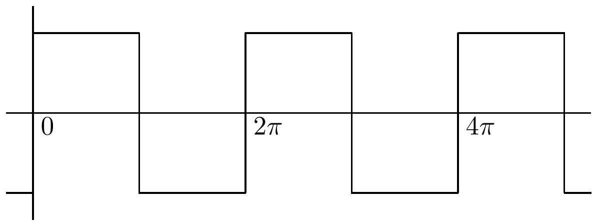
$$F(t) = \tfrac{1}{2}a_0 + \sum_{n=0}^{\infty}(a_n \cos(2n\pi\nu t) + b_n \sin(2n\pi\nu t)),$$

and the following formula for Fourier coefficients.

$$a_m = 2\nu \int_0^{1/\nu} \cos(2m\pi\nu t) F(t)\, dt, \tag{2.2.8}$$

$$b_m = 2\nu \int_0^{1/\nu} \sin(2m\pi\nu t) F(t)\, dt. \tag{2.2.9}$$

Example. The square wave sounds vaguely like the waveform produced by a clarinet, where odd harmonics dominate. It is the function $f(\theta)$ defined by $f(\theta) = 1$ for $0 \le \theta < \pi$ and $f(\theta) = -1$ for $\pi \le \theta < 2\pi$ (and then extend to all values of θ by making it periodic, $f(\theta + 2\pi) = f(\theta)$).



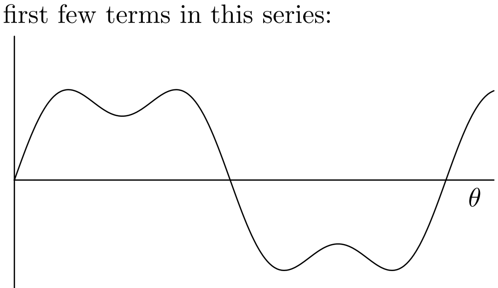
This function has Fourier coefficients

$$
\begin{aligned}
a_m &= \frac{1}{\pi}\left(\int_0^{\pi} \cos(m\theta)\, d\theta - \int_{\pi}^{2\pi} \cos(m\theta)\, d\theta\right) \\
&= \frac{1}{\pi}\left(\left[\frac{\sin(m\theta)}{m}\right]_0^{\pi} - \left[\frac{\sin(m\theta)}{m}\right]_{\pi}^{2\pi}\right) = 0 \\
b_m &= \frac{1}{\pi}\left(\int_0^{\pi} \sin(m\theta)\, d\theta - \int_{\pi}^{2\pi} \sin(m\theta)\, d\theta\right) \\
&= \frac{1}{\pi}\left(\left[-\frac{\cos(m\theta)}{m}\right]_0^{\pi} - \left[-\frac{\cos(m\theta)}{m}\right]_{\pi}^{2\pi}\right) \\
&= \frac{1}{\pi}\left(-\frac{(-1)^m}{m} + \frac{1}{m} + \frac{1}{m} - \frac{(-1)^m}{m}\right) \\
&= \begin{cases} 4/m\pi & (m \text{ odd}) \\ 0 & (m \text{ even}) \end{cases}
\end{aligned}
$$

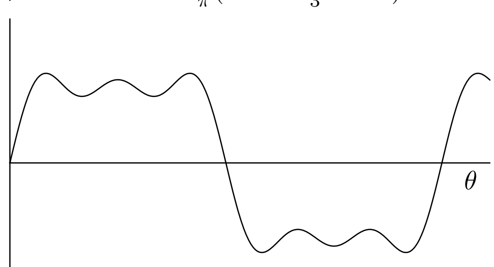
Thus the Fourier series for this square wave is

$$\tfrac{4}{\pi}(\sin\theta + \tfrac{1}{3}\sin 3\theta + \tfrac{1}{5}\sin 5\theta + \dots). \qquad (2.2.10)$$

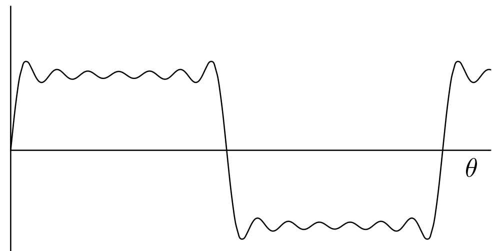
Let us examine the first few terms in this series:



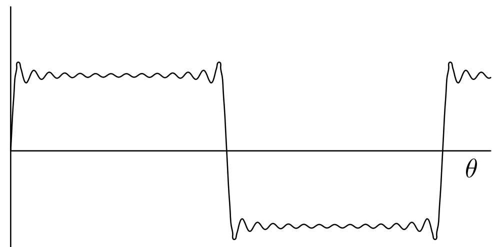
$\frac{4}{\pi}(\sin\theta + \frac{1}{3}\sin 3\theta)$



$\frac{4}{\pi}(\sin\theta + \frac{1}{3}\sin 3\theta + \frac{1}{5}\sin 5\theta)$



$\frac{4}{\pi}(\sin\theta + \frac{1}{3}\sin 3\theta + \cdots + \frac{1}{13}\sin 13\theta)$



$\frac{4}{\pi}(\sin\theta + \frac{1}{3}\sin 3\theta + \cdots + \frac{1}{27}\sin 27\theta)$

Some features of this example are worth noticing. The first observation is that these graphs seem to be converging to a square wave. But they seem to be converging quite slowly, and getting more and more bumpy in the process. Next, observe what happens at the point of discontinuity of the original function. The Fourier coefficients did not depend on what value we assigned to the function at the discontinuity, so we do not expect to recover that information. Instead, the series is converging to a value which is equal to the average of the higher and the lower values of the function. This is a general phenomenon, which we shall discuss later.

Finally, there is a very interesting phenomenon which is happening right near the discontinuity. There is an overshoot, which never seems to get any smaller.

Does this mean that the series is not converging properly? Well, not quite. At each given value of θ, the series is converging just fine. It's just when we look at values of θ closer and closer to the discontinuity that we find problems. This is because of a lack of *uniform* convergence. This overshoot is called the Gibbs phenomenon, and we shall discuss it in more detail in §2.5.

Exercises

1. Prove equations (2.2.2)–(2.2.4) by rewriting the products of trigonometric functions inside the integral as sums before integrating.

2. Are the following functions of θ periodic? If so, determine the smallest period, and which multiples of the fundamental frequency are present. If not, explain why not.

(i) $\sin\theta + \sin\frac{5}{4}\theta$.

(ii) $\sin\theta + \sin\sqrt{2}\,\theta$.

(iii) $\sin^2\theta$.

(iv) $\sin(\theta^2)$.
(v) $\sin\theta + \sin(\theta + \frac{\pi}{3})$.

3. Draw graphs of the functions $\sin(220\pi t)+\sin(440\pi t)$ and $\sin(220\pi t)+\cos(440\pi t)$. Explain why these sound the same, even though the graphs look quite different.

2.3. Even and odd functions

A function $f(\theta)$ is said to be *even* if $f(-\theta) = f(\theta)$, and it is said to be *odd* if $f(-\theta) = -f(\theta)$. For example, $\cos\theta$ is even, while $\sin\theta$ is odd. Of course, many functions are neither even nor odd. If a function happens to be both even and odd, then it is zero, because we have $f(\theta) = f(-\theta) = -f(\theta)$.

Given any function $f(\theta)$, we can obtain an even function by taking the average of $f(\theta)$ and $f(-\theta)$, i.e., $\frac{1}{2}(f(\theta) + f(-\theta))$. Similarly, $\frac{1}{2}(f(\theta) - f(-\theta))$ is an odd function. These add up to give the original function $f(\theta)$, so we have written $f(\theta)$ as a sum of its *even part* and its *odd part*,

$$f(\theta) = \frac{f(\theta) + f(-\theta)}{2} + \frac{f(\theta) - f(-\theta)}{2}.$$

To see that this is the unique way to write the function as a sum of an even function and an odd function, let us suppose that we are given two expressions $f(\theta) = g_1(\theta) + h_1(\theta)$ and $f(\theta) = g_2(\theta) + h_2(\theta)$ with g_1 and g_2 even, and h_1 and h_2 odd. Rearranging $g_1 + h_1 = g_2 + h_2$, we get $g_1 - g_2 = h_2 - h_1$. The left side is even and the right side is odd, so their common value is both even and odd, and hence zero. This means that $g_1 = g_2$ and $h_1 = h_2$.

Multiplication of even and odd functions works as you might expect: even times even or odd times odd gives even, and even times odd or odd times even gives odd.

Now for any odd function $f(\theta)$, and for any $a > 0$, we have

$$\int_{-a}^{0} f(\theta)\, d\theta = -\int_{0}^{a} f(\theta)\, d\theta$$

so that

$$\int_{-a}^{a} f(\theta) = 0.$$

So for example, if $f(\theta)$ is even and periodic with period 2π, then $\sin(m\theta)f(\theta)$ is odd, and so the Fourier coefficients b_m are zero, since

$$b_m = \frac{1}{\pi}\int_0^{2\pi} \sin(m\theta)f(\theta)\, d\theta = \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(m\theta)f(\theta)\, d\theta = 0.$$

Similarly, if $f(\theta)$ is odd and periodic with period 2π, then $\cos(m\theta)f(\theta)$ is odd, and so the Fourier coefficients a_m are zero, since

$$a_m = \frac{1}{\pi}\int_0^{2\pi} \cos(m\theta)f(\theta)\, d\theta = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos(m\theta)f(\theta)\, d\theta = 0.$$

This explains, for example, why $a_m = 0$ in the example on page 33. The square wave is not quite an even function, because $f(\pi) \neq f(-\pi)$, but changing the value of a function at a finite set of points in the interval of integration never affects the value of an integral, so we just replace $f(\pi)$ and $f(-\pi)$ by zero.

There is a similar explanation for why $b_{2m} = 0$ in the same example, using a different symmetry. The discussion of even and odd functions depended on the symmetry $\theta \mapsto -\theta$ of order two. For periodic functions of period 2π, there is another symmetry of order two, namely $\theta \mapsto \theta + \pi$. The functions $f(\theta)$ satisfying $f(\theta+\pi) = f(\theta)$ are *half-period symmetric*, while functions satisfying $f(\theta + \pi) = -f(\theta)$ are *half-period antisymmetric*. Any function $f(\theta)$ can be decomposed into half-period symmetric and antisymmetric parts:

$$f(\theta) = \frac{f(\theta) + f(\theta + \pi)}{2} + \frac{f(\theta) - f(\theta + \pi)}{2}.$$

Multiplying half-period symmetric and antisymmetric functions works in the same way as for even and odd functions.

If $f(\theta)$ is half-period antisymmetric, then

$$\int_{\pi}^{2\pi} f(\theta)\, d\theta = -\int_{0}^{\pi} f(\theta)\, d\theta$$

and so

$$\int_{0}^{2\pi} f(\theta)\, d\theta = 0.$$

Now the functions $\sin(m\theta)$ and $\cos(m\theta)$ are both half-period symmetric if m is even, and half-period antisymmetric if m is odd. So we deduce that if $f(\theta)$ is half-period symmetric, $f(\theta + \pi) = f(\theta)$, then the Fourier coefficients with odd indices (a_{2m+1} and b_{2m+1}) are zero, while if $f(\theta)$ is antisymmetric, $f(\theta + \pi) = -f(\theta)$, then the Fourier coefficients with even indices a_{2m} and b_{2m} are zero (check that this holds for a_0 too!). This corresponds to the fact that half-period symmetry is really the same thing as being periodic with half the period, so that the frequency components have to be even multiples of the defining frequency; while half-period antisymmetric functions only have frequency components at odd multiples of the defining frequency.

In the example on page 33, the function is half-period antisymmetric, and so the coefficients a_{2m} and b_{2m} are zero.

Exercises

1. Evaluate $\displaystyle\int_{0}^{2\pi} \sin(\sin\theta)\sin(2\theta)\, d\theta$.

2. Using the theory of even and odd functions, and the theory of half-period symmetric and antisymmetric functions, which Fourier coefficients of $\tan\theta$ have to be zero? Find the first nonzero coefficient.

3. Which Fourier coefficients vanish for a periodic function $f(\theta)$ of period 2π satisfying $f(\theta) = f(\pi - \theta)$? What about $f(\theta) = -f(\pi - \theta)$?

[Hint: Consider the symmetry $\theta \mapsto \pi - \theta$, and compare $\int_{-\pi/2}^{\pi/2} f(\theta)\,d\theta$ with $\int_{\pi/2}^{3\pi/2} f(\theta)\,d\theta$ for antisymmetric functions with respect to this symmetry.]

2.4. Conditions for convergence

Unfortunately, it is not true that if we start with a periodic function $f(\theta)$, form the Fourier coefficients a_m and b_m according to equations (2.2.5) and (2.2.6) and then form the sum (2.2.1), then we recover the original function $f(\theta)$. The most obvious problem is that if two functions differ just at a single value of θ then the Fourier coefficients will be identical. So we cannot possibly recover the function from its Fourier coefficients without some further conditions. However, if the function is nice enough, it can be recovered in the manner indicated. The following is a consequence of the work of Dirichlet.

THEOREM 2.4.1. *Suppose that $f(\theta)$ is periodic with period 2π, and that it is continuous and has a bounded continuous derivative except at a finite number of points in the interval $[0, 2\pi]$. If a_m and b_m are defined by equations* (2.2.5) *and* (2.2.6) *then the series defined by equation* (2.2.1) *converges to $f(\theta)$ at all points where $f(\theta)$ is continuous.*

PROOF. See Körner [**68**], Theorem 1 and Chapters 15 and 16. □

An important special case of the above theorem is the following. A C^1 function is defined to be a function which is differentiable with continuous derivative. If $f(\theta)$ is a periodic C^1 function with period 2π, then $f'(\theta)$ is continuous on the closed interval $[0, 2\pi]$, and hence bounded there. So $f(\theta)$ satisfies the conditions of the above theorem.

It is important to note that continuity, or even differentiability of $f(\theta)$ is not sufficient for the Fourier series for $f(\theta)$ to converge to $f(\theta)$. Paul DuBois-Reymond constructed an example of a continuous function for which the coefficients a_m and b_m are not bounded. The construction is by no means easy and we shall not give it here. The reader may form the impression at this stage that the only purpose for the existence of such functions is to beset theorems such as the above with conditions, and that in real life, all functions are just as differentiable as we would like them to be. This point of view is refuted by the observation that many phenomena in real life are governed by some form of Brownian motion. Functions describing these phenomena will tend to be everywhere continuous but nowhere differentiable.[3]

[3]The first examples of functions which are everywhere continuous but nowhere differentiable were constructed by Weierstrass, *Abhandlungen aus der Functionenlehre*, Springer (1886), p. 97. He showed that if $0 < b < 1$, a is an odd integer, and $ab > 1 + \frac{3\pi}{2}$ then $f(t) = \sum_{n=1}^{\infty} b^n \cos a^n (2\pi\nu)t$ is a uniformly convergent sum, and that $f(t)$ is everywhere continuous but nowhere differentiable. G. H. Hardy, *Weierstrass's non-differentiable function*, Trans. Amer. Math. Soc. 17 (1916), 301–325, showed that the same holds if the bound on ab is replaced by $ab > 1$. Manfred Schroeder, *Fractals, chaos and power laws*, W. H. Freeman and Co., 1991, p. 96, points out that functions of this form can be thought of as *fractal*

In music, noise is an example of the same phenomenon. Many of the functions employed in musical synthesis are not even continuous. Sawtooth functions and square waves are typical examples.

However, the question of convergence of the Fourier series is not the same as the question of whether the function $f(\theta)$ can be reconstructed from its Fourier coefficients a_n and b_n. At the age of 19, Fejér proved the remarkable theorem that any continuous function $f(\theta)$ can be reconstructed from its Fourier coefficients. His idea was that if the partial sums s_m defined by

$$s_m = \tfrac{1}{2}a_0 + \sum_{n=1}^{m}(a_n \cos(n\theta) + b_n \sin(n\theta)) \tag{2.4.1}$$

converge, then their averages

$$\sigma_m = \frac{s_0 + \cdots + s_m}{m+1}$$

converge to the same limit. But it is conceivable that the σ_m could converge without the s_m converging. This idea for smoothing out the convergence had already been around for some time when Fejér approached the problem. It had been used by Euler and extensively studied by Cesàro, and goes by the name of Cesàro summability.

Theorem 2.4.2 (Fejér). *If $f(\theta)$ is a Riemann integrable periodic function then the Cesàro sums σ_m converge to $f(\theta)$ as m tends to infinity at every value of θ where $f(\theta)$ is continuous.*[4]

Proof. We shall prove this theorem in Section 2.7. See also Körner [**68**], Chapter 2. □

We shall interpret this theorem as saying that every continuous function may be reconstructed from its Fourier coefficients. But the reader should bear in mind that if the function does not satisfy the hypotheses of Theorem 2.4.1 then the reconstruction of the function is done via Cesàro sums, and not simply as the sum of the Fourier series.

There are other senses in which we could ask for a Fourier series to converge. One of the most important ones is *mean square convergence.*

Theorem 2.4.3. *Let $f(\theta)$ be a continuous periodic function with period 2π. Then among all the functions $g(\theta)$ which are linear combinations of $\cos(n\theta)$ and $\sin(n\theta)$ with $0 \le n \le m$, the partial sum s_m defined in equation (2.4.1) minimizes the mean square error of $g(\theta)$ as an approximation to $f(\theta)$,*

$$\frac{1}{2\pi}\int_0^{2\pi} |f(\theta) - g(\theta)|^2 \, d\theta.$$

waveforms. For example, if we set $a = 2^{13/12}$, then doubling the speed of this function will result in a tone which sounds similar to the original, but lowered by a semitone and softer by a factor of b. This sort of self-similarity is characteristic of fractals. It is ironic that Weierstrass, in contrast with the vast majority of mathematicians, held a dislike for music.

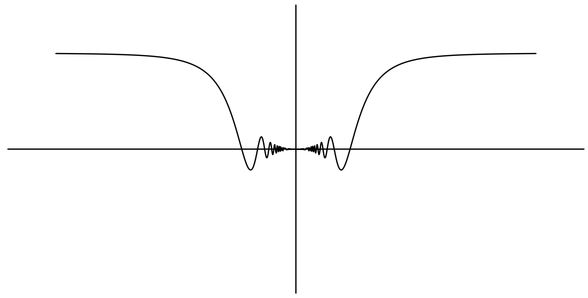
[4]Continuous functions are Riemann integrable, so Fejér's theorem applies to all continuous periodic functions.

Furthermore, in the limit as m tends to infinity, the mean square error of s_m as an approximation to $f(\theta)$ tends to zero.

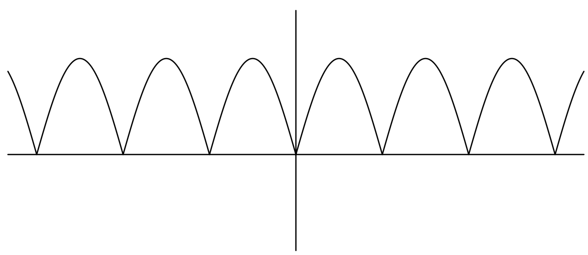
PROOF. See Körner [**68**], Chapters 32–34. $\square$

Exercises

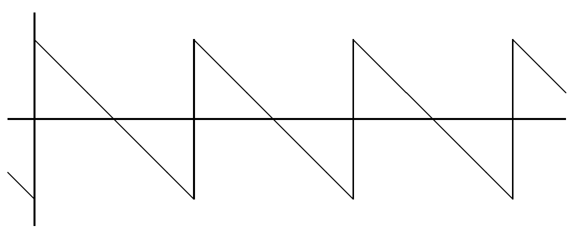
1. Show that the function $f(x) = x^2 \sin(1/x^2)$ is differentiable for all values of x, but its derivative is unbounded around $x = 0$.



2. Find the Fourier series for the periodic function $f(\theta) = |\sin\theta|$ (the absolute value of $\sin\theta$). In other words, find the coefficients a_m and b_m using equations (2.2.5) and (2.2.6). You will need to divide the interval from 0 to 2π into two subintervals in order to evaluate the integral.



3. Let $\phi(\theta)$ be the periodic sawtooth function with period 2π defined by $\phi(\theta) = (\pi - \theta)/2$ for $0 < \theta < 2\pi$ and $\phi(0) = \phi(2\pi) = 0$. Find the Fourier series for $\phi(\theta)$.[5]

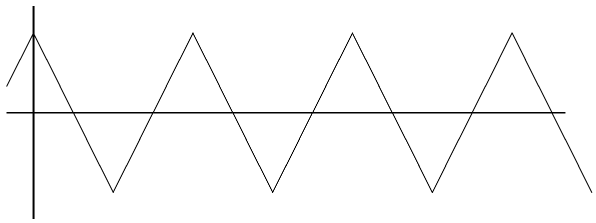


[5]The sawtooth waveform is approximately what is produced by a violin or other bowed string instrument. This is because the bow pulls the string, and then suddenly releases it when the coefficient of static friction is exceeded. The coefficient of dynamic friction is smaller, so once the string is released by the bow, it will tend to continue moving rapidly until the other extreme of its trajectory is reached. See Section 3.4.

4. Find the Fourier series of the continuous periodic triangular wave function defined by

$$f(\theta) = \begin{cases} \frac{\pi}{2} - \theta & 0 \le \theta \le \pi \\ \theta - \frac{3\pi}{2} & \pi \le \theta \le 2\pi \end{cases}$$

and $f(\theta + 2\pi) = f(\theta)$.



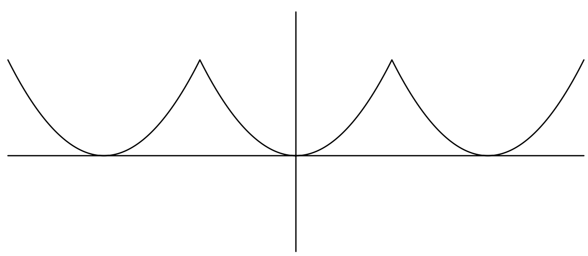
5. (a) Show that if $f(\theta)$ is a bounded (and Riemann integrable) periodic function with period 2π then the Fourier coefficients a_m and b_m defined by (2.2.5)–(2.2.7) are bounded.

(b) If $f(\theta)$ is a differentiable periodic function with period 2π, find the relationship between the Fourier coefficients $a_m(f)$, $b_m(f)$ for $f(\theta)$ and the Fourier coefficients $a_m(f')$, $b_m(f')$ for the derivative $f'(\theta)$. [Hint: use integration by parts]

(c) Show that if $f(\theta)$ is a k times differentiable periodic function with period 2π, and the kth derivative $f^{(k)}(\theta)$ is bounded, then the Fourier coefficients a_m and b_m of $f(\theta)$ are bounded by a constant multiple of $1/m^k$.

Thus, smoothness of $f(\theta)$ is reflected in rapidity of decay of the Fourier coefficients.

6. Find the Fourier series for the function $f(\theta)$ defined by $f(\theta) = \theta^2$ for $-\pi \le \theta \le \pi$ and and then extended to all values of θ by periodicity, $f(\theta + 2\pi) = f(\theta)$. Evaluate your answer at $\theta = 0$ and at $\theta = \pi$, and use your answer to find $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}$ and $\sum_{n=1}^{\infty} \frac{1}{n^2}$.



2.5. The Gibbs phenomenon

A function defined on a closed interval is said to be *piecewise continuous* if it is continuous except at a finite set of points, and at those points the left limit and right limit exist although they may not be equal. When we talk of the *size* of a discontinuity of a piecewise continuous function $f(\theta)$ at $\theta = a$, we mean the difference $f(a^+) - f(a^-)$, where

$$f(a^+) = \lim_{\theta \to a^+} f(\theta), \qquad f(a^-) = \lim_{\theta \to a^-} f(\theta)$$

denote the left limit and the right limit at that point. A periodic function is said to be piecewise continuous if it is so on a closed interval forming a period of the function.

Many of the functions encountered in the theory of synthesized sound are piecewise continuous but not continuous. These include waveforms such as the square wave and the sawtooth function.

Denote by $\phi(\theta)$ the piecewise continuous periodic function defined by $\phi(\theta) = (\pi - \theta)/2$ for $0 < \theta < 2\pi$, $\phi(0) = 0$, and $\phi(\theta + 2\pi) = \phi(\theta)$. Then given any piecewise continuous periodic function $f(\theta)$, we may add some finite set of functions of the form $C\phi(\theta+\alpha)$ (with C and α constants) to make the left limits and right limits at the discontinuities agree. We can then just change the values of the function at the discontinuities, which will not affect the Fourier series, to make the function continuous. It follows that in order to understand the Fourier series for piecewise continuous functions in general, it suffices to understand the Fourier series of continuous functions together with the Fourier series of the single function $\phi(\theta)$. The Fourier series of this function (see Exercise 3 of §2.4) is

$$\phi(\theta) = \sum_{n=1}^{\infty} \frac{\sin n\theta}{n}.$$

At the discontinuity ($\theta = 0$), this series converges to zero because all the terms are zero. This is the average of the left limit and the right limit at this point. It follows that for any piecewise continuous periodic function, the Cesàro sums σ_m described in §2.4 converge everywhere, and at the points of discontinuity σ_m converges to the average of the left and right limit at the point:

$$\lim_{m\to\infty} \sigma_m(a) = \tfrac{1}{2}(f(a^+) + f(a^-)).$$

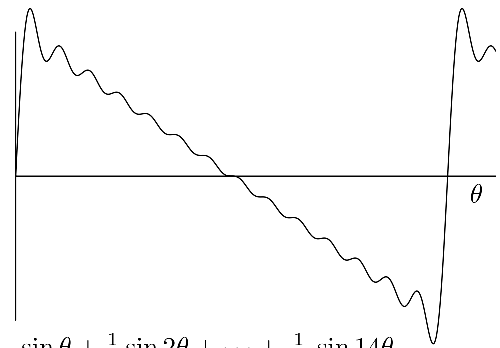
A further examination of the function $\phi(\theta)$ shows that the convergence around the point of discontinuity is not as straightforward as one might suppose. Namely, setting

$$\phi_m(\theta) = \sum_{n=1}^{m} \frac{\sin n\theta}{n}, \qquad (2.5.1)$$

although it is true that we have *pointwise convergence*, in the sense that for each point a we have $\lim_{m\to\infty} \phi_m(a) = \phi(a)$, this convergence is not uniform.

The definition in analysis of pointwise convergence is that given a value a of θ and given $\varepsilon > 0$, there exists N such that $m \geq N$ implies $|\phi_m(a) - \phi(a)| < \varepsilon$. Uniform convergence means that given $\varepsilon > 0$, there exists N (independent of a) such that for all values a of θ, $m \geq N$ implies $|\phi_m(a) - \phi(a)| < \varepsilon$. What happens with the Fourier series for the above function ϕ is that there is an overshoot, the size of which does not tend to zero as m gets larger. The peak of the overshoot gets closer and closer to the discontinuity though, so that for any particular value a of θ, convergence holds.

But choosing ε smaller than the size of the overshoot shows that uniform convergence fails. This overshoot is called the Gibbs phenomenon.[6]



$\sin\theta + \frac{1}{2}\sin 2\theta + \cdots + \frac{1}{14}\sin 14\theta$

To demonstrate the reality of the overshoot, we shall compute its size in the limit. The first step is to differentiate $\phi_m(\theta)$ to find its local maxima and minima. We concentrate on the interval $0 \le \theta \le \pi$, since $\phi_m(2\pi - \theta) = -\phi_m(\theta)$. We have

$$\phi'_m(\theta) = \sum_{n=1}^{m} \cos n\theta = \frac{\sin \frac{1}{2}m\theta \, \cos \frac{1}{2}(m+1)\theta}{\sin \frac{1}{2}\theta}$$

(see Exercise 6 of §1.7). So the zeros of $\phi'_m(\theta)$ occur at $\theta = \frac{(2k+1)\pi}{m+1}$ and $\theta = \frac{2k\pi}{m}$, $0 \le k \le \lfloor \frac{m-1}{2} \rfloor$.[7]

Now $\sin \frac{1}{2}\theta$ is positive throughout the interval $0 \le \theta \le 2\pi$. At $\theta = \frac{(2k+1)\pi}{m+1}$, $\sin \frac{1}{2}m\theta$ has sign $(-1)^k$ while $\cos \frac{1}{2}(m+1)\theta$ changes sign from $(-1)^k$ to $(-1)^{k+1}$, so that $\phi'_m(\theta)$ changes from positive to negative. It follows that $\theta = \frac{(2k+1)\pi}{m+1}$ is a local maximum of ϕ_m. Similarly, at $\theta = \frac{2k\pi}{m}$, $\cos \frac{1}{2}(m+1)\theta$ has sign $(-1)^k$ while $\sin \frac{1}{2}m\theta$ changes sign from $(-1)^{k-1}$ to $(-1)^k$, so that $\phi'_m(\theta)$ changes from negative to positive. It follows that $\theta = \frac{2k\pi}{m}$ is a local minimum of $\phi_m(\theta)$. These local maxima and minima alternate.

The first local maximum value of $\phi_m(\theta)$ for $0 \le \theta \le 2\pi$ happens at $\frac{\pi}{m+1}$. The value of $\phi_m(\theta)$ at this maximum is

$$\phi_m\left(\tfrac{\pi}{m+1}\right) = \sum_{n=1}^{m} \tfrac{1}{n} \sin\left(\tfrac{n\pi}{m+1}\right) = \frac{\pi}{m+1} \sum_{n=1}^{m} \frac{\sin\left(\frac{n\pi}{m+1}\right)}{\left(\frac{n\pi}{m+1}\right)}.$$

This is the Riemann sum for

$$\int_0^{\pi} \frac{\sin\theta}{\theta}\, d\theta$$

---

[6] Josiah Willard Gibbs described this phenomenon in a series of letters to *Nature* in 1898 in correspondence with A. E. H. Love. He seems to have been unaware of the previous treatment of the subject by Henry Wilbraham in his article *On a certain periodic function*, Cambridge & Dublin Math. J. **3** (1848), 198–201.

[7] The notation $\lfloor x \rfloor$ denotes the largest integer less than or equal to x.

with $m+1$ equal intervals of size $\frac{\pi}{m+1}$ (note that $\lim\limits_{\theta\to 0}\dfrac{\sin\theta}{\theta}=1$ so that we should define the integrand to be 1 when $\theta=0$ to make a continuous function on the closed interval $0\le\theta\le\pi$). Therefore the limit as m tends to infinity of the height of the first maximum point of the sum of the first m terms in the Fourier series for $\phi(\theta)$ is

$$\int_0^\pi \frac{\sin\theta}{\theta}\,d\theta \approx 1.8519370.$$

This overshoots the maximum value $\frac{\pi}{2}\approx 1.5707963$ of the function $\phi(\theta)$ by a factor of 1.1789797. Of course, the size of the discontinuity is not $\frac{\pi}{2}$ but π, so that as a proportion of the size of the discontinuity, the overshoot is about 8.9490%.[8] It follows that for any piecewise continuous function, the overshoot of the Fourier series just after a discontinuity is this proportion of the size of the discontinuity.

After the function overshoots, it then returns to undershoot, then overshoot again, and so on, each time with a smaller value than before. An argument similar to the above shows that the value at the kth critical point of $\phi_m(\theta)$ tends to $\int_0^{k\pi}\frac{\sin\theta}{\theta}\,d\theta$ as m tends to infinity. Thus for example the first undershoot ($k=2$) has a value with a limit of about 1.4181516, which undershoots $\frac{\pi}{2}$ by a factor of 0.9028233. The undershoot is therefore about 4.8588% of the size of the discontinuity.

The Gibbs phenomenon can be interpreted in terms of the response of an amplifier as follows. No matter how good your amplifier is, if you feed it a square wave, the output will overshoot at the discontinuity by roughly 9%. This is because any amplifier has a frequency beyond which it does not respond. Improving the amplifier can only increase this frequency, but cannot get rid of the limitation altogether.

Manufacturers of cathode ray tubes also have to contend with this problem. The beam is being made to run across the tube from left to right linearly and then switch back suddenly to the left. Much effort goes into preventing the inevitable overshoot from causing problems.

As mentioned above, the Gibbs phenomenon is a good example to illustrate the distinction between pointwise convergence and uniform convergence. For pointwise convergence of a sequence of functions $f_n(\theta)$ to a function $f(\theta)$, it is required that for each value of θ, the values $f_n(\theta)$ must converge to $f(\theta)$. For uniform convergence, it is required that the distance between $f_n(\theta)$ and $f(\theta)$ is bounded by a quantity which depends on n and not on θ, and which tends to zero as n tends to infinity. In the above example, the distance between the nth partial sum of the Fourier series and the original function can at best be bounded by a quantity which depends on n and

[8]This value was first computed by Maxime Bôcher, *Introduction to the theory of Fourier's series.* Ann. of Math. (2) **7** (1905–6), 81–152. A number of otherwise reputable sources overstate the size of the overshoot by a factor of two for some reason probably associated with uncritical copying.

not on θ, but which tends to roughly 0.28114. So this Fourier series converges pointwise, but not uniformly.

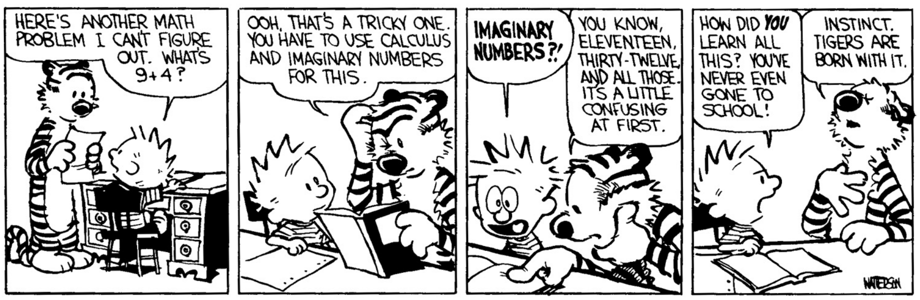
Exercises

1. Show that

$$\int_0^x \frac{\sin\theta}{\theta}\,d\theta = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)(2n+1)!}.$$

Use this formula to verify the approximate value of $\int_0^\pi \frac{\sin\theta}{\theta}\,d\theta$ given in the text.

2.6. Complex coefficients



The theory of Fourier series is considerably simplified by the introduction of complex exponentials. See Appendix C for a quick summary of complex numbers and complex exponentials. The relationships (C.1)–(C.3)

$$e^{i\theta} = \cos\theta + i\sin\theta \qquad \cos\theta = \frac{e^{i\theta}+e^{-i\theta}}{2}$$

$$e^{-i\theta} = \cos\theta - i\sin\theta \qquad \sin\theta = \frac{e^{i\theta}-e^{-i\theta}}{2i}$$

mean that equation (2.2.1) can be rewritten as[9]

$$f(\theta) = \sum_{n=-\infty}^{\infty} \alpha_n e^{in\theta} \tag{2.6.1}$$

where $\alpha_0 = \frac{1}{2}a_0$, and for $m > 0$, $\alpha_m = \frac{1}{2}a_m + \frac{1}{2i}b_m$ and $\alpha_{-m} = \frac{1}{2}a_m - \frac{1}{2i}b_m$. Conversely, given a series of the form (2.6.1) we can reconstruct the series (2.2.1) using $a_0 = 2\alpha_0$, $a_m = \alpha_m + \alpha_{-m}$ and $b_m = i(\alpha_m - \alpha_{-m})$ for $m > 0$.

[9]Note that we are dealing with complex valued functions of a real periodic variable, and not with functions of a complex variable here.

Equations (2.2.2)–(2.2.4) are replaced by the single equation[10]

$$\int_0^{2\pi} e^{im\theta} e^{in\theta}\, d\theta = \begin{cases} 2\pi & \text{if } m = -n \\ 0 & \text{if } m \neq -n \end{cases}$$

and equations (2.2.5)–(2.2.7) are replaced by

$$\alpha_m = \frac{1}{2\pi} \int_0^{2\pi} e^{-im\theta} f(\theta)\, d\theta. \tag{2.6.2}$$

Exercises

1. For the square wave example discussed in §2.2, show that

$$\alpha_m = \begin{cases} 2/im\pi & m \text{ odd} \\ 0 & m \text{ even.} \end{cases}$$

so that the Fourier series is

$$\sum_{n=-\infty}^{\infty} \frac{2}{i(2n+1)\pi} e^{i(2n+1)\theta}.$$

2.7. Proof of Fejér's Theorem

We are now in a position to prove Fejér's Theorem 2.4.2. This section may safely be skipped on first reading.

In terms of the complex form of the Fourier series, the partial sums (2.4.1) become

$$s_m = \sum_{n=-m}^{m} \alpha_n e^{in\theta}, \tag{2.7.1}$$

and so the Cesàro sums σ_m are given by

$$\begin{aligned}
\sigma_m(\theta) &= \frac{s_0 + \cdots + s_m}{m+1} \\
&= \frac{1}{m+1} \sum_{j=0}^{m} \sum_{n=-j}^{j} \alpha_n e^{in\theta} \\
&= \frac{1}{m+1} \big(\alpha_{-m} e^{-im\theta} + 2\alpha_{-(m-1)} e^{-i(m-1)\theta} + 3\alpha_{-(m-2)} e^{-i(m-2)\theta} + \dots \\
&\qquad + \cdots + m\alpha_{-1} e^{-i\theta} + (m+1)\alpha_0 e^0 + m\alpha_1 e^{i\theta} + \cdots + \alpha_m e^{im\theta} \big) \\
&= \sum_{n=-m}^{m} \frac{m+1-|n|}{m+1} \alpha_n e^{in\theta}. \\
&= \sum_{n=-m}^{m} \frac{m+1-|n|}{m+1} \left(\frac{1}{2\pi} \int_0^{2\pi} e^{-inx} f(x)\, dx \right) e^{in\theta}
\end{aligned}$$

[10]Over the complex numbers, to interpret this equation as an orthogonality relation (see the footnote on page 32), the inner product needs to be taken to be $\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(\theta)\overline{g(\theta)}\, d\theta$.

$$= \frac{1}{2\pi} \int_0^{2\pi} f(x) \left(\sum_{n=-m}^{m} \frac{m+1-|n|}{m+1} e^{in(\theta - x)} \right) dx$$

$$= \frac{1}{2\pi} \int_0^{2\pi} f(x) K_m(\theta - x)\, dx$$

where

$$K_m(y) = \sum_{n=-m}^{m} \frac{m+1-|n|}{m+1} e^{iny}.$$

The functions K_m are called the Fejér kernels.

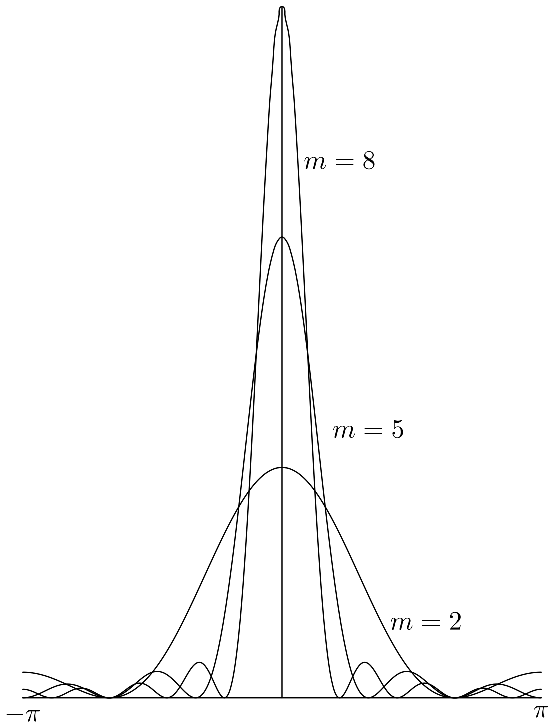
The substitution $y = \theta - x$ shows that

$$\frac{1}{2\pi} \int_0^{2\pi} f(x) K_m(\theta - x)\, dx = \frac{1}{2\pi} \int_0^{2\pi} f(\theta - y) K_m(y)\, dy$$

By examining what happens when a geometric series is squared, for $y \neq 0$ we have

$$\begin{aligned} K_m(y) &= \frac{1}{m+1}\big(e^{-imy} + 2e^{-i(m-1)y} + \cdots + (m+1)e^0 + \cdots + e^{imy}\big) \\ &= \frac{1}{m+1}(e^{-i\frac{m}{2}y} + e^{-i(\frac{m}{2}-1)y} + \cdots + e^{i\frac{m}{2}y})^2 \qquad (2.7.2) \\ &= \frac{1}{m+1}\left(\frac{e^{i\frac{m+1}{2}y} - e^{-i\frac{m+1}{2}y}}{e^{i\frac{1}{2}y} - e^{-i\frac{1}{2}y}}\right)^2 \\ &= \frac{1}{m+1}\left(\frac{\sin\frac{m+1}{2}y}{\sin\frac{1}{2}y}\right)^2, \end{aligned}$$

and $K_m(0) = m+1$ can also be read off from (2.7.2). Here are the graphs of $K_m(y)$ for some small values of m.



The function $K_m(y)$ satisfies $K_m(y) \geq 0$ for all y; for any $\delta > 0$, $K_m(y) \to 0$ uniformly as $m \to \infty$ on $[\delta, 2\pi - \delta]$; and $\int_0^{2\pi} K_m(y)\,dy = 2\pi$. So

$$\begin{aligned}\sigma_m(\theta) = \frac{1}{2\pi}\int_0^{2\pi} f(\theta - y)K_m(y)\,dy &\approx \frac{1}{2\pi}\int_{-\delta}^{\delta} f(\theta - y)K_m(y)\,dy \\ &\approx f(\theta)\left(\frac{1}{2\pi}\int_{-\delta}^{\delta} K_m(y)\,dy\right) \approx f(\theta).\end{aligned}$$

If $f(\theta)$ is continuous at θ, then by choosing δ small enough, the second approximation may be made as close as desired (independently of m). Then by choosing m large enough, the first and third approximations may be made as close as desired. This completes the proof of Fejér's theorem.

Exercises

1. (i) Substitute equation (2.6.2) in equation (2.7.1) to show that

$$s_m(\theta) = \frac{1}{2\pi}\int_0^{2\pi} f(x)D_m(\theta - x)\,dx$$

where

$$D_m(y) = \sum_{n=-m}^{m} e^{iny}.$$

The functions D_m are called the Dirichlet kernels.

(ii) Use a substitution to show that

$$s_m(\theta) = \frac{1}{2\pi}\int_0^{2\pi} f(\theta - y)D_m(y)\,dy.$$

(iii) By regarding the formula for $D_m(y)$ as a geometric series, show that

$$D_m(y) = \frac{\sin(m+\frac{1}{2})y}{\sin\frac{1}{2}y}.$$

(iv) Show that $|D_m(y)| \le |\operatorname{cosec}\frac{1}{2}y|$

(v) Sketch the graphs of the Dirichlet kernels for small values of m. What happens as m gets large?

2.8. Bessel functions

Bessel functions[11] are the result of applying the theory of Fourier series to the functions $\sin(z\sin\theta)$ and $\cos(z\sin\theta)$ as functions of θ, where z is to be thought of at first as a real (or complex) constant, and later it will be allowed to vary. We shall have two uses for the Bessel functions. One is understanding the vibrations of a drum in §3.6, and the other is understanding the amplitudes of side bands in FM synthesis in §8.8.

Now $\sin(z\sin\theta)$ is an odd periodic function of θ, so its Fourier coefficients a_n (2.2.1) are zero for all n (see §2.3). Since

$$\sin(z\sin(\pi+\theta)) = -\sin(z\sin\theta),$$

the Fourier coefficients b_{2n} are also zero (see §2.3 again). The coefficients b_{2n+1} depend on the parameter z, and so we write $2J_{2n+1}(z)$ for this coefficient. The factor of two simplifies some later calculations. So the Fourier expansion (2.2.1) is

$$\boxed{\sin(z\sin\theta) = 2\sum_{n=0}^{\infty} J_{2n+1}(z)\sin(2n+1)\theta.} \tag{2.8.1}$$

Similarly, $\cos(z\sin\theta)$ is an even periodic function of θ, so the coefficients b_n are zero. Since

$$\cos(z\sin(\pi+\theta)) = \cos(z\sin\theta)$$

we also have $a_{2n+1} = 0$, and we write $2J_{2n}(z)$ for the coefficient a_{2n} to obtain

$$\boxed{\cos(z\sin\theta) = J_0(z) + 2\sum_{n=1}^{\infty} J_{2n}(z)\cos 2n\theta.} \tag{2.8.2}$$

[11] Friedrich Wilhelm Bessel was a German astronomer and a friend of Gauss. He was born in Minden on July 22, 1784. His working life started as a ship's clerk. But in 1806, he became an assistant at an astronomical observatory in Lilienthal. In 1810 he became director of the then new Prussian Observatory in Königsberg, where he remained until he died on March 17, 1846. The original context (around 1824) of his investigations of the functions that bear his name was the study of planetary motion, see Section 2.11.

The functions $J_n(z)$ giving the Fourier coefficients in these expansions are called the *Bessel functions* of the first kind.

Equations (2.2.5) and (2.2.6) allow us to find the Fourier coefficients $J_n(z)$ in the above expansions as integrals. We obtain

$$2J_{2n+1}(z) = \frac{1}{\pi}\int_0^{2\pi} \sin(2n+1)\theta \sin(z\sin\theta)\, d\theta.$$

The integrand is an even function of θ, so the integral from 0 to 2π is twice the integral from 0 to π,

$$J_{2n+1}(z) = \frac{1}{\pi}\int_0^{\pi} \sin(2n+1)\theta \sin(z\sin\theta)\, d\theta.$$

Now the function $\cos(2n+1)\theta\cos(z\sin\theta)$ is negated when θ is replaced by $\pi-\theta$, so

$$\frac{1}{\pi}\int_0^{\pi} \cos(2n+1)\theta \cos(z\sin\theta)\, d\theta = 0.$$

Adding this into the above expression for $J_{2n+1}(z)$, we obtain

$$\begin{aligned} J_{2n+1}(z) &= \frac{1}{\pi}\int_0^{\pi} [\cos(2n+1)\theta\cos(z\sin\theta) + \sin(2n+1)\theta\sin(z\sin\theta)]\, d\theta \\ &= \frac{1}{\pi}\int_0^{\pi} \cos((2n+1)\theta - z\sin\theta)\, d\theta. \end{aligned}$$

In a similar way, we have

$$2J_{2n}(z) = \frac{1}{\pi}\int_0^{2\pi} \cos 2n\theta \cos(z\sin\theta)\, d\theta$$

which a similar manipulation puts in the form

$$J_{2n}(z) = \frac{1}{\pi}\int_0^{\pi} \cos(2n\theta - z\sin\theta)\, d\theta.$$

This means that we have the single equation for all values of n, even or odd,

$$\boxed{J_n(z) = \frac{1}{\pi}\int_0^{\pi} \cos(n\theta - z\sin\theta)\, d\theta} \tag{2.8.3}$$

which can be taken as a definition for the Bessel functions for integers $n \geq 0$. In fact, this definition also makes sense when n is a negative integer,[12] and gives

$$J_{-n}(z) = (-1)^n J_n(z). \tag{2.8.4}$$

This means that (2.8.1) and (2.8.2) can be rewritten as

$$\sin(z\sin\theta) = \sum_{n=-\infty}^{\infty} J_{2n+1}(z)\sin(2n+1)\theta \tag{2.8.5}$$

[12] For non-integer values of n, the above formula is not the correct definition of $J_n(z)$. Rather, one uses the differential equation (2.10.1). See for example Whittaker and Watson, *A course in modern analysis*, Cambridge University Press, 1927, p. 358.

$$\cos(z\sin\theta) = \sum_{n=-\infty}^{\infty} J_{2n}(z)\cos 2n\theta. \tag{2.8.6}$$

We also have

$$\sum_{n=-\infty}^{\infty} J_{2n}(z)\sin 2n\theta = 0$$

$$\sum_{n=-\infty}^{\infty} J_{2n+1}(z)\cos(2n+1)\theta = 0,$$

because the terms with positive subscript cancel with the corresponding terms with negative subscript. So we can rewrite equations (2.8.5) and (2.8.6) as

$$\sin(z\sin\theta) = \sum_{n=-\infty}^{\infty} J_n(z)\sin n\theta \tag{2.8.7}$$

$$\cos(z\sin\theta) = \sum_{n=-\infty}^{\infty} J_n(z)\cos n\theta. \tag{2.8.8}$$

So using equation (1.7.2) we have

$$\begin{aligned}\sin(\phi + z\sin\theta) &= \sin\phi\cos(z\sin\theta) + \cos\phi\sin(z\sin\theta)\\ &= \sin\phi\sum_{n=-\infty}^{\infty} J_n(z)\cos n\theta + \cos\phi\sum_{n=-\infty}^{\infty} J_n(z)\sin n\theta\\ &= \sum_{n=-\infty}^{\infty} J_n(z)(\sin\phi\cos n\theta + \cos\phi\sin n\theta).\end{aligned}$$

Finally, recombining the terms using equation (1.7.2), we obtain

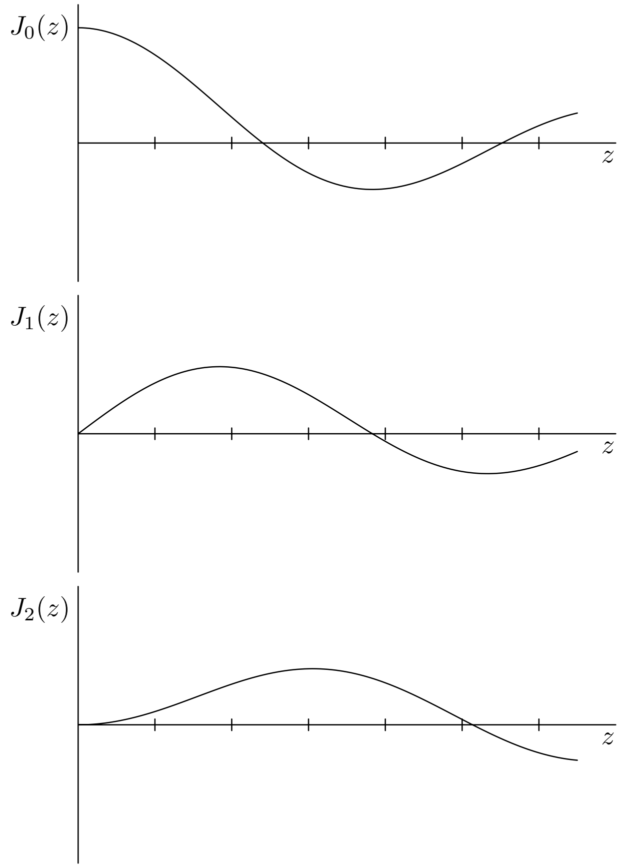
$$\boxed{\sin(\phi + z\sin\theta) = \sum_{n=-\infty}^{\infty} J_n(z)\sin(\phi + n\theta).} \tag{2.8.9}$$

This equation will be of fundamental importance for FM synthesis in §8.8. A similar argument gives

$$\cos(\phi + z\sin\theta) = \sum_{n=-\infty}^{\infty} J_n(z)\cos(\phi + n\theta), \tag{2.8.10}$$

which can also be obtained from equation (2.8.9) by replacing ϕ by $\phi + \frac{\pi}{2}$, or by differentiating with respect to ϕ, keeping z and θ constant.

Here are graphs of the first few Bessel functions:



Exercises

1. Replace θ by $\frac{\pi}{2} - \theta$ in equations (2.8.1) and (2.8.2) to obtain the Fourier series for $\sin(z\cos\theta)$ and $\cos(z\cos\theta)$.

2. Deduce equation (2.8.10) from equation (2.8.9).

2.9. Properties of Bessel functions

From equation (2.8.9), we can obtain relationships between the Bessel functions and their derivatives, as follows. Differentiating (2.8.9) with respect to z, keeping θ and ϕ constant, we obtain

$$\sin\theta\,\cos(\phi + z\sin\theta) = \sum_{n=-\infty}^{\infty} J_n'(z)\sin(\phi + n\theta) \qquad (2.9.1)$$

On the other hand, multiplying equation (2.8.10) by $\sin\theta$ and using (1.7.7), we have

$$\sin\theta\,\cos(\phi + z\sin\theta) = \sum_{n=-\infty}^{\infty} J_n(z).\tfrac{1}{2}\big(\sin(\phi + (n+1)\theta) - \sin(\phi + (n-1)\theta)\big)$$

$$= \sum_{n=-\infty}^{\infty} \tfrac{1}{2}\big(J_{n-1}(z) - J_{n+1}(z)\big)\sin(\phi + n\theta). \tag{2.9.2}$$

In the last step, we have split the sum into two parts, reindexed by replacing n by $n-1$ and $n+1$ respectively in the two parts, and then recombined the parts.

We would like to compare formulas (2.9.1) and (2.9.2) and deduce that

$$\boxed{J_n'(z) = \tfrac{1}{2}\big(J_{n-1}(z) - J_{n+1}(z)\big)} \tag{2.9.3}$$

In order to do this, we need to know that the functions $\sin(\phi + n\theta)$ are independent. This can be seen using Fourier series as follows.

LEMMA 2.9.1. *If*

$$\sum_{n=-\infty}^{\infty} a_n \sin(\phi + n\theta) = \sum_{n=-\infty}^{\infty} a_n' \sin(\phi + n\theta),$$

as an identity between functions of ϕ and θ, where a_n and a_n' do not depend on θ and ϕ, then each coefficient $a_n = a_n'$.

PROOF. Subtracting one side from the other, we see that we must prove that if $\sum_{n=-\infty}^{\infty} c_n \sin(\phi + n\theta) = 0$ (where $c_n = a_n - a_n'$) then each $c_n = 0$. To prove this, we expand using (1.7.2) to give

$$\sum_{n=-\infty}^{\infty} c_n \sin\phi\cos n\theta + \sum_{n=-\infty}^{\infty} c_n \cos\phi \sin n\theta = 0.$$

Putting $\phi = 0$ and $\phi = \frac{\pi}{2}$ in this equation, we obtain

$$\sum_{n=-\infty}^{\infty} c_n \cos n\theta = 0, \tag{2.9.4}$$

$$\sum_{n=-\infty}^{\infty} c_n \sin n\theta = 0. \tag{2.9.5}$$

Multiply equation (2.9.4) by $\cos m\theta$, integrate from 0 to 2π and divide by π. Using equation (2.2.3), we get $c_m + c_{-m} = 0$. Similarly, from equations (2.9.5) and (2.2.4), we get $c_m - c_{-m} = 0$. Adding and dividing by two, we get $c_m = 0$. □

This completes the proof of equation (2.9.3). As an example, setting $n = 0$ in (2.9.3) and using (2.8.4), we obtain

$$J_1(z) = -J_0'(z). \tag{2.9.6}$$

In a similar way, we can differentiate (2.8.9) with respect to θ, keeping z and ϕ constant to obtain

$$z\cos\theta\,\cos(\phi + z\sin\theta) = \sum_{n=-\infty}^{\infty} nJ_n(z)\cos(\phi + n\theta). \tag{2.9.7}$$

On the other hand, multiplying equation (2.8.10) by $z\cos\theta$ and using (1.7.10), we obtain

$$\begin{aligned} z\cos\theta\ \cos(\phi + z\sin\theta) \\ &= \sum_{n=-\infty}^{\infty} J_n(z).\tfrac{z}{2}\big(\cos(\phi+(n+1)\theta)+\cos(\phi+(n-1)\theta)\big) \\ &= \sum_{n=-\infty}^{\infty} \tfrac{z}{2}\big(J_{n-1}(z)+J_{n+1}(z)\big)\cos(\phi+n\theta). \end{aligned} \tag{2.9.8}$$

Comparing equations (2.9.7) and (2.9.8) and using Lemma 2.9.1, we obtain the recurrence relation

$$\boxed{J_n(z) = \frac{z}{2n}\big(J_{n-1}(z)+J_{n+1}(z)\big).} \tag{2.9.9}$$

Exercises

1. Show that $\displaystyle\int_0^\infty J_1(z)\,dz = 1$.
[You may use the fact that $\lim\limits_{z\to\infty} J_0(z) = 0$]

2.10. Bessel's equation and power series

Using equations (2.9.3) and (2.9.9), we can now derive the differential equation (2.10.1) for the Bessel functions $J_n(z)$. Using (2.9.3) twice, we obtain

$$\begin{aligned} J_n''(z) &= \tfrac{1}{2}(J_{n-1}'(z) - J_{n+1}'(z)) \\ &= \tfrac{1}{4}J_{n-2}(z) - \tfrac{1}{2}J_n(z) + \tfrac{1}{4}J_{n+2}(z). \end{aligned}$$

On the other hand, substituting (2.9.9) into (2.9.3), we obtain

$$\begin{aligned} J_n'(z) &= \tfrac{1}{2}\left(\tfrac{z}{2(n-1)}(J_{n-2}(z)+J_n(z)) - \tfrac{z}{2(n+1)}(J_n(z)+J_{n+2}(z))\right) \\ &= \tfrac{z}{4(n-1)}J_{n-2}(z) + \tfrac{z}{2(n^2-1)}J_n(z) - \tfrac{z}{4(n-1)}J_{n+2}(z). \end{aligned}$$

In a similar way, using (2.9.9) twice gives

$$\begin{aligned} J_n(z) &= \tfrac{z}{2n}\left(\tfrac{z}{2(n-1)}(J_{n-2}(z)+J_n(z)) + \tfrac{z^2}{2(n+1)}(J_n(z)+J_{n+2}(z))\right) \\ &= \tfrac{z}{4n(n-1)}J_{n-2}(z) + \tfrac{z^2}{n^2-1}J_n(z) + \tfrac{z^2}{4n(n+1)}J_{n+2}(z). \end{aligned}$$

Combining these three formulas, we obtain

$$J_n''(z) + \tfrac{1}{z}J_n'(z) - \tfrac{n^2}{z^2}J_n(z) = -J_n(z),$$

or

$$\boxed{J_n''(z) + \frac{1}{z}J_n'(z) + \left(1 - \frac{n^2}{z^2}\right)J_n(z) = 0.} \tag{2.10.1}$$

We now discuss the general solution to *Bessel's Equation*, namely the differential equation

$$f''(z) + \frac{1}{z}f'(z) + \left(1 - \frac{n^2}{z^2}\right) f(z) = 0. \tag{2.10.2}$$

This is an example of a second order linear differential equation, and once one solution is known, there is a general procedure for obtaining all solutions. In this case, this consists of substituting $f(z) = J_n(z)g(z)$, and finding the differential equation satisfied by the new function $g(z)$. We find that

$$\begin{aligned} f'(z) &= J_n'(z)g(z) + J_n(z)g'(z), \\ f''(z) &= J_n''(z)g(z) + 2J_n'(z)g'(z) + J_n(z)g''(z). \end{aligned}$$

So substituting into Bessel's equation (2.10.2), we obtain

$$\left(J_n''(z) + \frac{1}{z}J_n'(z) + \left(1 - \frac{n^2}{z^2}\right) J_n(z)\right) g(z) + \\ \left(2J_n'(z) + \frac{1}{z}J_n(z)\right) g'(z) + J_n(z)g''(z) = 0.$$

The coefficient of $g(z)$ vanishes by equation (2.10.1), and so we are left with

$$\left(2J_n'(z) + \frac{1}{z}J_n(z)\right) g'(z) + J_n(z)g''(z) = 0, \tag{2.10.3}$$

This is a separable first order equation for $g'(z)$, so we separate the variables

$$\frac{g''(z)}{g'(z)} = -2\frac{J_n'(z)}{J_n(z)} - \frac{1}{z}$$

and integrate to obtain

$$\ln|g'(z)| = -2\ln|J_n(z)| - \ln|z| + C$$

where C is the constant of integration. Exponentiating, we obtain

$$g'(z) = \frac{B}{zJ_n(z)^2}$$

where $B = \pm e^C$. Alternatively, we could have obtained this directly from equation (2.10.3) by multiplying by $zJ_n(z)$ to see that the derivative of $zJ_n(z)^2g'(z)$ is zero.

Integrating again, we obtain

$$g(z) = A + B\int \frac{dz}{zJ_n(z)^2}$$

where the integral sign denotes a chosen antiderivative. Finally, it follows that the general solution to Bessel's equation is given by

$$f(z) = AJ_n(z) + BJ_n(z)\int \frac{dz}{zJ_n(z)^2}. \tag{2.10.4}$$

The function

$$Y_n(z) = \frac{2}{\pi} J_n(z) \int \frac{dz}{zJ_n(z)^2},$$

for a suitable choice of constant of integration, is called Neumann's Bessel function of the second kind, or Weber's function. The factor of $2/\pi$ is introduced (by most, but not all authors) so that formulas involving $J_n(z)$ and $Y_n(z)$ look similar; we shall not go into the details. From the above integral, it is not hard to see that $Y_n(z)$ tends to $-\infty$ as z tends to zero from above; we shall be more explicit about this towards the end of this section.

Next, we develop the power series for $J_n(z)$. We begin with $J_0(z)$. Putting $z = \theta = 0$ in equation (2.8.2), we see that $J_0(0) = 1$. By (2.8.4), $J_0(z)$ is an even function of z, so we look for a power series of the form

$$J_0(z) = 1 + a_2z^2 + a_4z^4 + \cdots = \sum_{k=0}^{\infty} a_{2k}z^{2k}$$

where $a_0 = 1$. Then

$$J_0'(z) = 2a_2z + 4a_4z^3 + \cdots = \sum_{k=1}^{\infty} 2ka_{2k}z^{2k-1},$$

$$J_0''(z) = 2 \cdot 1\, a_2 + 4 \cdot 3\, a_4z^2 + \cdots = \sum_{k=1}^{\infty} 2k(2k-1)a_{2k}z^{2k-2}.$$

Putting $n = 0$ in equation (2.10.1) and comparing coefficients of a_{2k-2}, we obtain

$$2k(2k-1)a_{2k} + 2ka_{2k} + a_{2k-2} = 0,$$

or

$$(2k)^2a_{2k} = -a_{2k-2}.$$

So starting with $a_0 = 1$, we obtain $a_2 = -1/2^2$, $a_4 = 1/(2^2 \cdot 4^2)$, ..., and by induction on k, we have

$$a_{2k} = \frac{(-1)^k}{2^2 \cdot 4^2 \ldots (2k)^2} = \frac{(-1)^k}{2^k(k!)^2}.$$

So we have

$$J_0(z) = 1 - \frac{z^2}{2^2} + \frac{z^4}{2^2 \cdot 4^2} - \frac{z^6}{2^2 \cdot 4^2 \cdot 6^2} + \cdots = \sum_{k=0}^{\infty} \frac{(-1)^k \left(\frac{z}{2}\right)^{2k}}{(k!)^2}. \qquad (2.10.5)$$

Since the coefficients in this power series are tending to zero very rapidly, it has an infinite radius of convergence.[13] So it is uniformly convergent, and can be differentiated term by term. It follows that the sum of the power series satisfies Bessel's equation, because that's how we chose the coefficients. We have already seen that there is only one solution of Bessel's equation with

[13]For any value of z, the ratio of successive terms tends to zero, so by the ratio test the series converges.

value 1 at $z = 0$, which completes the proof that the sum of the power series is indeed $J_0(z)$.

Differentiating equation (2.10.5) term by term and using (2.9.6), we see that

$$J_1(z) = \frac{z}{2} - \frac{z^3}{2^2 \cdot 4} + \frac{z^5}{2^2 \cdot 4^2 \cdot 6} - \cdots = \sum_{k=0}^{\infty} \frac{(-1)^k \left(\frac{z}{2}\right)^{1+2k}}{k!(1+k)!}.$$

Now using equation (2.9.9) and induction on n, we find that

$$\boxed{J_n(z) = \sum_{k=0}^{\infty} \frac{(-1)^k (\frac{z}{2})^{n+2k}}{k!(n+k)!},} \tag{2.10.6}$$

with infinite radius of convergence.

From the power series, we can get information about $Y_n(z)$ as $z \to 0^+$. For small positive values of z, $J_n(z)$ is equal to $z^n/2^n n!$ plus much smaller terms. So $\frac{1}{zJ_n(z)^2}$ is equal to $2^{2n}(n!)^2 z^{-2n-1}$ plus much smaller terms, and $\int \frac{1}{zJ_n(z)^2}\,dz$ is equal to $-2^{2n-1}n!(n-1)!z^{-2n}$ plus much smaller terms. Finally, $Y_n(z)$ is equal to $-2^n(n-1)!z^{-n}/\pi$ plus much smaller terms. In particular, this shows that $Y_n(z) \to -\infty$ as $z \to 0^+$.

Exercises

1. Show that $y = J_n(\alpha x)$ is a solution of the differential equation

$$\frac{d^2y}{dx^2} + \frac{1}{x}\frac{dy}{dx} + \left(\alpha^2 - \frac{n^2}{x^2}\right) y = 0.$$

Show that the general solution to this equation is given by $y = AJ_n(\alpha x) + BY_n(\alpha x)$.

2. Show that $y = \sqrt{x}J_n(x)$ is a solution of the differential equation

$$\frac{d^2y}{dx^2} + \left(1 + \frac{\frac{1}{4} - n^2}{x^2}\right) y = 0.$$

Find the general solution of this equation.

3. Show that $y = J_n(e^x)$ is a solution of the differential equation

$$\frac{d^2y}{dx^2} + (e^{2x} - n^2)y = 0.$$

Find the general solution of this equation.

4. The following exercise is another route to Bessel's differential equation (2.10.1).

(a) Differentiate equation (2.8.9) twice with respect to z, keeping ϕ and θ constant.

(b) Differentiate equation (2.8.9) twice with respect to θ, keeping z and ϕ constant.

(c) Divide the result of (b) by z^2 and add to the result of (a), and use the relation $\sin^2\theta + \cos^2\theta = 1$. Deduce that

$$\sum_{n=-\infty}^{\infty} \left(J_n''(z) + \frac{1}{z}J_n'(z) + \left(1 - \frac{n^2}{z^2}\right) J_n(z)\right) \sin(\phi + z\theta) = 0.$$

(d) Finally, use Lemma 2.9.1 to show that Bessel's equation (2.8.9) holds.

(The following exercises suppose some knowledge of complex analysis in order to give an alternative development of the power series and recurrence relations for the Bessel functions)

5. Show that

$$\begin{aligned} J_n(z) &= \frac{1}{2\pi}\int_0^{\pi} e^{i(n\theta - z\sin\theta)}\,d\theta + \frac{1}{2\pi}\int_0^{\pi} e^{-i(n\theta - z\sin\theta)}\,d\theta \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i(n\theta - z\sin\theta)}\,d\theta. \end{aligned}$$

Substitute $t = e^{i\theta}$ (so that $\frac{1}{2i}(t - \frac{1}{t}) = \sin\theta$) to obtain

$$J_n(z) = \frac{1}{2\pi i}\oint t^{-n-1} e^{\frac{1}{2}z(t-\frac{1}{t})}\,dt \tag{2.10.7}$$

where the contour of integration goes counterclockwise once around the unit circle. Use Cauchy's integral formula to deduce that $J_n(z)$ is the coefficient of t^n in the Laurent expansion of $e^{\frac{1}{2}z(t-\frac{1}{t})}$:

$$e^{\frac{1}{2}z(t-\frac{1}{t})} = \sum_{n=-\infty}^{\infty} J_n(z)t^n.$$

6. Substitute $t = 2s/z$ in (2.10.7) to obtain

$$J_n(z) = \frac{1}{2\pi i}\left(\frac{z}{2}\right)^n \oint s^{-n-1} e^{s - \frac{z^2}{4s}}\,ds.$$

Discuss the contour of integration. Expand the integrand in powers of z to give

$$J_n(z) = \frac{1}{2\pi i}\sum_{k=0}^{\infty} \frac{(-1)^k}{k!}\left(\frac{z}{2}\right)^{n+2k} \oint s^{-n-k-1} e^s\,ds$$

and justify the term by term integration. Show that the residue of the integrand at $s = 0$ is $1/(n+k)!$ when $n + k \geq 0$ and is zero when $n + k < 0$. Deduce the power series (2.10.6).

7. (a) Use the power series (2.10.6) to show that

$$J_n(z) = \tfrac{z}{2n}(J_{n-1}(z) + J_{n+1}(z)).$$

(b) Differentiate the power series (2.10.6) term by term to show that

$$J_n'(z) = \tfrac{1}{2}(J_{n-1}(z) - J_{n+1}(z)).$$

Further reading on Bessel functions:

Milton Abramowitz and Irene A. Stegun, *Handbook of mathematical functions*, National Bureau of Standards, 1964, reprinted by Dover in 1965 and still in print. This contains extensive tables of many mathematical functions including $J_n(z)$ and $Y_n(z)$.

Frank Bowman, *Introduction to Bessel functions*, reprinted by Dover in 1958 and still in print.

G. N. Watson, *A treatise on the theory of Bessel functions* [**138**] is an 800 page tome on the theory of Bessel functions. This work contains essentially everything that was known in 1922 about these functions, and is still pretty much the standard reference.

E. T. Whittaker and G. N. Watson, *Modern Analysis*, Cambridge University Press, 1927, chapter XVII.

See also Appendix B for some tables and a summary of some properties of Bessel functions, as well as a C++ program for calculation.

2.11. Fourier series for FM feedback and planetary motion

We shall see in §8.8 that in the theory of FM synthesis, feedback is represented by an equation of the form

$$\phi = \sin(\omega t + z\phi), \tag{2.11.1}$$

where ω and z are constants with $|z| \le 1$.

In the theory of planetary motion, Kepler's laws imply that the angle θ subtended at the center (not the focus) of the elliptic orbit by the planet, measured from the major axis of the ellipse, satisfies

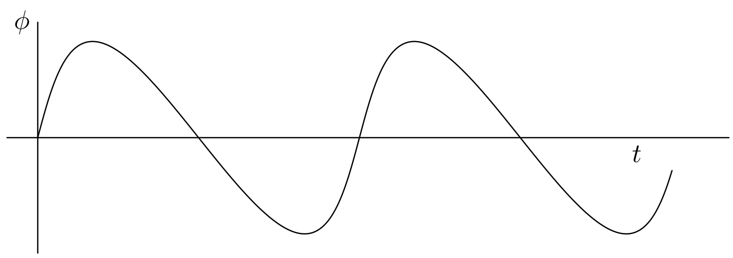
$$\omega t = \theta - z \sin\theta \tag{2.11.2}$$

where z is the eccentricity[14] of the ellipse, a number in the range $0 \le z \le 1$, and $\omega = 2\pi\nu$ is a constant which plays the role of average angular velocity.

Both of these equations define periodic functions of t, namely ϕ in the first case and $\sin\theta = (\theta - \omega t)/z$ in the second. In fact, they are really just different ways of writing the same equation. To get from equation (2.11.2) to (2.11.1), we use the substitution $\theta = \omega t + z\phi$. To go the other way, we use the inverse substitution $\phi = (\theta - \omega t)/z$.

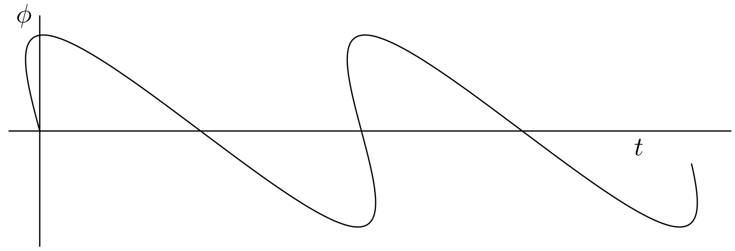
The same functions turn up in other places too. In an exercise at the end of this section, we describe the relevance to nonlinear acoustics.

To graph ϕ as a function of t, it is best to use θ as a parameter and set $t = (\theta - z\sin\theta)/\omega$, $\phi = \sin\theta$. Here is the result when $z = \frac{1}{2}$:



When $|z| > 1$, the parametrized form of the equation still makes sense, but it is easy to see that the resulting graph does not define ϕ uniquely as a function of t. Here is the result when $z = \frac{3}{2}$:

[14] The eccentricity of an ellipse is defined to be the distance from the center to the focus, as a proportion of the major radius.



In this section, we examine equation (2.11.2), and find the Fourier coefficients of $\phi = \sin\theta$ as a function of t, regarding z as a constant. The answer is given in terms of Bessel functions. In fact, the solution of this equation in the context of planetary motion was the original motivation for Bessel to introduce his functions $J_n(z)$.[15]

First, for convenience we write $T = \omega t$. Next, we observe that provided $|z| \le 1$, $\theta - z\sin\theta$ is a strictly increasing fuction of θ whose domain and range are the whole real line. It follows that solving equation (2.11.2) gives a unique value of θ for each T, so that θ may be regarded as a continuous function of T. Furthermore, adding 2π to both θ and T, or negating both θ and T does not affect equation (2.11.2), so $z\phi = z\sin\theta = \theta - T$ is an odd periodic function of T with period 2π. So it has a Fourier expansion

$$z\phi = \sum_{n=1}^{\infty} b_n \sin nT. \tag{2.11.3}$$

The coefficients b_n can be calculated directly using equation (2.2.6):

$$b_n = \frac{1}{\pi}\int_0^{2\pi} z\phi \sin nT\,dT = \frac{2}{\pi}\int_0^{\pi} z\phi\sin nT\,dT.$$

Integrating by parts gives

$$b_n = \frac{2}{\pi}\left[-z\phi\frac{\cos nT}{n}\right]_0^{\pi} + \frac{2}{\pi}\int_0^{\pi} z\,\frac{d\phi}{dT}\,\frac{\cos nT}{n}\,dT.$$

We have $\phi = 0$ when $T = 0$ or $T = \pi$, so the first term vanishes. Rewriting the second term, we obtain

$$b_n = \frac{2}{n\pi}\int_0^{\pi}\cos nT\,\frac{d(z\phi)}{dT}\,dT.$$

Now $\displaystyle\int_0^{\pi}\cos nT\,dT = 0$, so we can rewrite this as

$$\begin{aligned} b_n &= \frac{2}{n\pi}\int_0^{\pi}\cos nT\,\frac{d(z\phi+T)}{dT}\,dT = \frac{2}{n\pi}\int_0^{\pi}\cos nT\,\frac{d\theta}{dT}\,dT \\ &= \frac{2}{n\pi}\int_0^{\pi}\cos nT\,d\theta. \end{aligned}$$

[15] Bessel, *Untersuchung der Theils der planetarischen Störungen, welcher aus der Bewegung der Sonne entsteht*, Berliner Abh. (1826), 1–52.

In the last step, we have used the fact that as T increases from 0 to π, so does θ. Substituting $T = \theta - z\sin\theta$ now gives

$$b_n = \frac{2}{n\pi}\int_0^\pi \cos(n\theta - nz\sin\theta)\,d\theta.$$

Comparing with equation (2.8.3) finally gives

$$b_n = \frac{2}{n}J_n(nz).$$

Substituting back into equation (2.11.3) gives

$$\boxed{\phi = \sin\theta = \sum_{n=1}^{\infty}\frac{2J_n(nz)}{nz}\sin n\omega t.} \tag{2.11.4}$$

So this equation gives the Fourier series relevant to feedback in FM synthesis (2.11.1), planetary motion (2.11.2), and nonlinear acoustics (2.11.5).

Exercises

1. Show that if a function ϕ satisfying equation (2.11.1) is regarded as a function of z and t, and ω is regarded as a constant, then ϕ is a solution of the partial differential equation

$$\frac{\partial\phi}{\partial z} = \frac{\phi}{\omega}\frac{\partial\phi}{\partial t} \tag{2.11.5}$$

(See Appendix P for a brief review of partial derivatives). Show that if α is a nonzero constant, then the dilation which replaces ϕ by ϕ/α and replaces z by αz gives another solution to this equation.

[Warning: This equation is *nonlinear*: adding solutions does not give another solution, and multiplying a solution by a scalar does not give another solution]

This equation turns out to be relevant to nonlinear acoustics. In this context, the solutions given by applying the above dilation to equation (2.11.4) are called *Fubini solutions*,[16] in spite of the fact that they were described by Bessel more than a century earlier. The picture given on page 59 now represents the solution for $|\alpha z| > 1$, and describes an acoustic shock wave (in this context, αz is called the *distortion range variable*).

2.12. Pulse streams

In this section, we examine streams of square pulses. The purpose of this is twofold. First, in analog synthesizers[17] one method for obtaining a time varying frequency spectrum is to use pulse width modulation (PWM). A low frequency oscillator (LFO, §8.2) is used to control the pulse width of a square wave, while keeping the fundamental frequency constant. The second point

[16]Eugene Fubini, *Anomalies in the propagation of acoustic waves at great amplitude* (in Italian), Alta Frequenza 4 (1935), 530–581. Eugene Fubini (1913–1997) was son of the mathematician Guido Fubini (1879–1943), after whom Fubini's theorem is named.

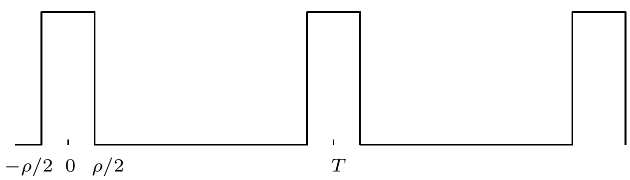
[17]This is also used in some of the more modern analog modeling synthesizers such as the Roland JP-8000/JP-8080.

is that by keeping the pulse width constant and decreasing the frequency, we motivate the definition of Fourier transform, to be introduced in §2.13.

Let us investigate the frequency spectrum of the square wave given by

$$f(t) = \begin{cases} 1 & 0 \le t \le \rho/2 \\ 0 & \rho/2 < t < T - \rho/2 \\ 1 & T - \rho/2 \le t < T \end{cases}$$

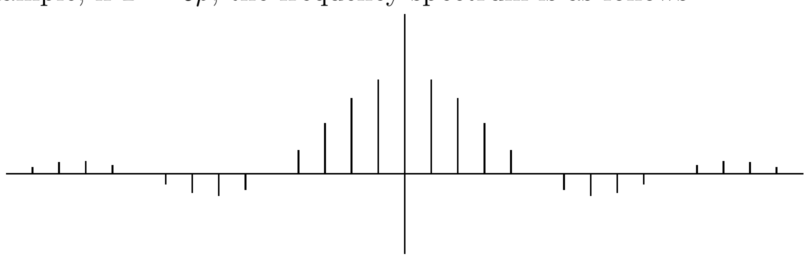
where ρ is some number between 0 and T, and $f(t+T) = f(t)$.



The Fourier coefficients are given by

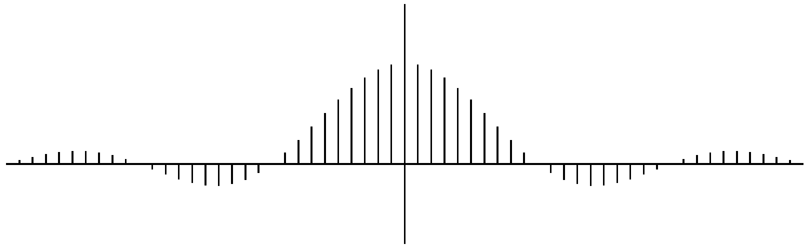
$$\alpha_m = \frac{1}{T}\int_{-\rho/2}^{\rho/2} e^{-2m\pi i t/T}\,dt = \frac{1}{m\pi}\sin(m\pi\rho/T).$$

For example, if $T = 5\rho$, the frequency spectrum is as follows



If we keep ρ constant and increase T, the shape of the spectrum stays the same, but vertically scaled down in proportion, so as to keep the energy density along the horizontal axis constant. It makes sense to rescale to take account of this, and to plot $T\alpha_m$ instead of α_m. If we do this, and increase T while keeping ρ constant, all that happens is that the graph fills in. So for example, removing every second peak from the original square wave

then the spectrum fills in as follows.



Letting T tend to infinity while keeping ρ constant, we obtain the *Fourier transform* of a single square pulse, which (after suitable scaling) is the function $\sin(\nu)/\nu$. Here, ν is a continuously variable quantity representing frequency.

2.13. The Fourier transform

The theory of Fourier series, as described in §§2.2–2.4, decomposes periodic waveforms into infinite sums of sines and cosines, or equivalently (§2.6) complex exponential functions of the form e^{int}. It is often desirable to analyse nonperiodic functions in a similar way. This leads to the theory of Fourier transforms. The theory is more beset with conditions than the theory of Fourier series. In particular, without the introduction of *generalized functions* or *distributions*, the theory only describes functions which tend to zero for large positive or negative values of the time variable t. To deal with this from a musical perspective, we introduce the theory of windowing. The point is that any actual sound is not really periodic, since periodic functions have no starting point and no end point. Moreover, we don't really want a frequency analysis of, for example, the whole of a symphony, because the answer would be dominated by extremely phase sensitive low frequency information. We'd really like to know at each instant what the frequency spectrum of the sound is, and to plot this frequency spectrum against time. Now, it turns out that it doesn't really make sense to ask for the instantaneous frequency spectrum of a sound, because there's not enough information. We really need to know the waveform for a time window around each point, and analyse that. Small window sizes give information which is more localized in time, but the frequency components are smeared out along the spectrum. Large window sizes give information in which the frequency components are more accurately described, but more smeared out along the time axis. This limitation is inherent to the process, and has nothing to do with how accurately the waveform is measured. In this respect, it resembles the Heisenberg uncertainty principle.[18]

[18] In fact, this is more than just an analogy. In quantum mechanics, the probability distributions for position and momentum of a particle are related by the Fourier transform, with an extra factor of Planck's constant $\hbar$. The Heisenberg uncertainty principle applies to the expected deviation from the average value of any two quantities related by the Fourier transform, and says that the product of these expected deviations is at least $\frac{1}{2}$. So in the quantum mechanical context the product is at least $\hbar/2$, because of the extra factor.

If $f(t)$ is a real or complex valued function of a real variable t, then its *Fourier transform* $\hat{f}(\nu)$ is the function of a real variable ν defined by[19]

$$\boxed{\hat{f}(\nu) = \int_{-\infty}^{\infty} f(t)e^{-2\pi i\nu t}\,dt.} \tag{2.13.1}$$

Existence of a Fourier transform for a function assumes convergence of the above integral, and this already puts restrictions on the function $f(t)$. A reasonable condition which ensures convergence is the following. A function $f(t)$ is said to be L^1, or *absolutely integrable* on $(-\infty,\infty)$ if the integral $\int_{-\infty}^{\infty}|f(t)|\,dt$ converges. In particular, this forces $f(t)$ to tend to zero as $|t|\to\infty$ (except possibly on a set of measure zero, which may be ignored), which makes integrating by parts easier. In §2.17, we shall see how to extend the definition to a much wider class of functions using the theory of distributions. For example, we would at least like to be able to take the Fourier tranform of a sine wave.

Calculating the Fourier transform of a function is usually a difficult process. As an example, we now calculate the Fourier transform of $e^{-\pi t^2}$. This function is unusual, in that it turns out to be its own Fourier transform.

THEOREM 2.13.1. *The Fourier transform of* $e^{-\pi t^2}$ *is* $e^{-\pi\nu^2}$.

PROOF. Let $f(t) = e^{-\pi t^2}$. Then

$$\begin{aligned}\hat{f}(\nu) &= \int_{-\infty}^{\infty} e^{-\pi t^2}e^{-2\pi i\nu t}\,dt\\ &= \int_{-\infty}^{\infty} e^{-\pi(t^2+2i\nu t)}\,dt\\ &= \int_{-\infty}^{\infty} e^{-\pi((t+i\nu)^2+\nu^2)}\,dt.\end{aligned}$$

[19]There are a number of variations on this definition to be found in the literature, depending mostly on the placement of the factor of 2π. The way we have set it up means that the variable ν directly represents frequency. Most authors delete the 2π from the exponential in this definition, which amounts to using the angular velocity ω instead. This means that they either have a factor of $1/2\pi$ appearing in formula (2.13.3), causing an annoying asymmetry, or a factor of $1/\sqrt{2\pi}$ in both (2.13.1) and (2.13.3).

Strictly speaking, the meaning of equation (2.13.1) should be

$$\lim_{a\to-\infty}\lim_{b\to\infty}\int_a^b f(t)e^{-2\pi i\nu t}\,dt.$$

However, under some conditions this double limit may not exist, while

$$\lim_{R\to\infty}\int_{-R}^{R} f(t)e^{-2\pi i\nu t}\,dt$$

may exist. This weaker symmetric limit is called the *Cauchy principal value* of the integral. Principal values are often used in the theory of Fourier transforms.

Substituting $x = t + i\nu$, $dx = dt$, we obtain

$$\hat{f}(\nu) = \int_{-\infty}^{\infty} e^{-\pi(x^2+\nu^2)}\,dx. \tag{2.13.2}$$

This form of the integral makes it obvious that $\hat{f}(\nu)$ is positive and real, but it is not obvious how to evaluate the integral. It turns out that it can be evaluated using a trick. The trick is to square both sides, and then regard the right hand side as a double integral.

$$\begin{aligned}\hat{f}(\nu)^2 &= \int_{-\infty}^{\infty} e^{-\pi(x^2+\nu^2)}\,dx \int_{-\infty}^{\infty} e^{-\pi(y^2+\nu^2)}\,dy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\pi(x^2+y^2+2\nu^2)}\,dx\,dy.\end{aligned}$$

We now convert this double integral over the (x, y) plane into polar coordinates (r, θ). Remembering that the element of area in polar coordinates is $r\,dr\,d\theta$, we get

$$\hat{f}(\nu)^2 = \int_0^{2\pi}\int_0^{\infty} e^{-\pi(r^2+2\nu^2)}\,r\,dr\,d\theta.$$

We can easily perform the integration with respect to θ, since the integrand is constant with respect to θ. And then the other integral can be carried out explicitly.

$$\begin{aligned}\hat{f}(\nu)^2 &= \int_0^{\infty} 2\pi r e^{-\pi(r^2+2\nu^2)}\,dr \\ &= \left[-e^{-\pi(r^2+2\nu^2)}\right]_0^{\infty} \\ &= e^{-2\pi\nu^2}.\end{aligned}$$

Finally, since equation (2.13.2) shows that $\hat{f}(\nu)$ is positive, taking square roots gives $\hat{f}(\nu) = e^{-\pi\nu^2}$ as desired. □

The following gives a formula for the Fourier transform of the derivative of a function.

THEOREM 2.13.2. *The Fourier transform of $f'(t)$ is $2\pi i\nu\hat{f}(\nu)$.*

PROOF. Integrating by parts, we have

$$\begin{aligned}\int_{-\infty}^{\infty} f'(t)e^{-2\pi i\nu t}\,dt &= \left[f(t)e^{-2\pi i\nu t}\right]_0^{\infty} - \int_{-\infty}^{\infty} f(t)(-2\pi i\nu)e^{-2\pi i\nu t}\,dt \\ &= 0 + 2\pi i\nu\hat{f}(\nu).\end{aligned}$$

□

The inversion formula is the following, which should be compared with Theorem 2.4.1.

THEOREM 2.13.3. *Let $f(t)$ be a piecewise C^1 function (i.e., on any finite interval, $f(t)$ is C^1 except at a finite set of points) which is also L^1.*

Then at points where $f(t)$ is continuous, its value is given by the **inverse Fourier transform**

$$f(t) = \int_{-\infty}^{\infty} \hat{f}(\nu)e^{2\pi i\nu t}\,d\nu. \tag{2.13.3}$$

(Note the change of sign in the exponent from equation (2.13.1)*)*
At discontinuities, the expression on the right of this formula gives the average of the left limit and the right limit, $\frac{1}{2}(f(t^+) + f(t^-))$, just as in §2.5.

Just as in the case of Fourier series, it is not true that a piecewise continuous L^1 function satisfies the conclusions of the above theorem. But a device analogous to Cesàro summation works equally well here. The analogue of averaging the first n sums is to introduce a factor of $1 - |\nu|/R$ into the integral defining the inverse Fourier transform, before taking principal values.

Theorem 2.13.4. *Let $f(t)$ be a piecewise continuous L^1 function. Then at points where $f(t)$ is continuous, its value is given by*

$$f(t) = \lim_{R\to\infty} \int_{-R}^{R} \Big(1 - \frac{|\nu|}{R}\Big)\hat{f}(\nu)e^{2\pi i\nu t}\,d\nu.$$

At discontinuities, this formula gives $\frac{1}{2}(f(t^+) + f(t^-))$.

2.14. Proof of the inversion formula

The purpose of this section is to prove the Fourier inversion formula, Theorem 2.13.3. This says that under suitable conditions, if a function $f(t)$ has Fourier transform

$$\hat{f}(\nu) = \int_{-\infty}^{\infty} f(t)e^{-2\pi i\nu t}\,dt \tag{2.14.1}$$

then the original function $f(t)$ can be reconstructed as the Cauchy principal value of the integral

$$f(t) = \int_{-\infty}^{\infty} \hat{f}(\nu)e^{2\pi i\nu t}\,d\nu. \tag{2.14.2}$$

First of all, we have the same difficulty here as we did with Fourier series. Namely, if we change the value of $f(t)$ at just one point, then $\hat{f}(\nu)$ will not change. So the best we can hope for is to reconstruct the average of the left and right limits, if this exists, $\frac{1}{2}(f(t^+) + f(t^-))$.

To avoid using t both as a variable of integration and the independent variable, let us use τ instead of t in (2.14.2). Then the Cauchy principal value of the right hand side of (2.14.2) becomes

$$\lim_{A\to\infty} \int_{-A}^{A} \left(\int_{-\infty}^{\infty} f(t)e^{-2\pi i\nu t}\,dt\right) e^{2\pi i\nu\tau}\,d\nu.$$

So this is the expression we must compare with $f(\tau)$, or rather with $\frac{1}{2}(f(\tau^+)+f(\tau^-))$. Since the outer integral just involves a finite interval, and the inner integral is absolutely convergent, we may reverse the order of integration to see that (2.14.2) is equal to

$$\begin{aligned}
&\lim_{A\to\infty}\int_{-\infty}^{\infty} f(t)\int_{-A}^{A} e^{2\pi i\nu(\tau-t)}\,d\nu\,dt \\
&= \lim_{A\to\infty}\int_{-\infty}^{\infty} f(t)\left[\frac{1}{2\pi i(\tau-t)}e^{2\pi i\nu(\tau-t)}\right]_{\nu=-A}^{A} dt \\
&= \lim_{A\to\infty}\int_{-\infty}^{\infty} f(t)\frac{\sin 2\pi A(\tau-t)}{\pi(\tau-t)}\,dt
\end{aligned}$$

where we've used (C.3) to rewrite the complex exponentials in terms of sines.

Substituting $x = t-\tau$, $t = \tau + x$ in the $\int_0^\infty$ part, and substituting $x = \tau - t$, $t = \tau - x$ in the $\int_{-\infty}^0$ part of the above integral, we find that (2.14.2) is equal to

$$\lim_{A\to\infty}\int_0^{\infty}(f(\tau+x)+f(\tau-x))\frac{\sin 2\pi Ax}{\pi x}\,dx. \tag{2.14.3}$$

So we really need to understand the behavior of $\frac{\sin 2\pi Ax}{\pi x}$ and its integral, as A gets large. We do this in the following theorem.

THEOREM 2.14.1. (i) *For $A > 0$, we have* $\displaystyle\int_0^\infty \frac{\sin 2\pi Ax}{\pi x}\,dx = \tfrac{1}{2}$,

(ii) *For any $\varepsilon > 0$, we have*

$$\lim_{A\to\infty}\int_0^{\varepsilon}\frac{\sin 2\pi Ax}{\pi x}\,dx = \tfrac{1}{2} \quad \textit{and} \quad \lim_{A\to\infty}\int_{\varepsilon}^{\infty}\frac{\sin 2\pi Ax}{\pi x}\,dx = 0.$$

PROOF. To see that the integral converges, write

$$I_n = \int_{n/2A}^{(n+1)/2A}\frac{\sin 2\pi Ax}{\pi x}\,dx.$$

Then the I_n alternate in sign and monotonically decrease to zero, so their sum converges. To find the value of the integral, we first find

$$\begin{aligned}
\int_0^{\frac{\pi}{2}}\frac{\sin(2n+1)u}{\sin u}\,du &= \int_0^{\frac{\pi}{2}}\frac{e^{(2n+1)iu}-e^{-(2n+1)iu}}{e^{iu}-e^{-iu}}\,du \\
&= \int_0^{\frac{\pi}{2}}(e^{2niu}+e^{2(n-1)iu}+\cdots+e^{-2niu})\,du \\
&= \frac{\pi}{2}.
\end{aligned} \tag{2.14.4}$$

For the last step, the terms in the integral cancel out in pairs, so that the only term giving a nonzero contribution is the middle one, which is $e^0 = 1$.

Now $\frac{1}{\sin u} - \frac{1}{u} \to 0$ as $u \to 0$ (combine and use l'Hôpital's rule, for example), so this expression defines a nonnegative, uniformly continuous function on $[0, \frac{\pi}{2}]$. An elementary estimate of the difference between consecutive positive and negative areas then shows that

$$\lim_{n\to\infty}\int_0^{\frac{\pi}{2}}\left(\frac{1}{\sin u}-\frac{1}{u}\right)\sin(2n+1)u\,du=0.$$

Combining with (2.14.4) gives

$$\lim_{n\to\infty}\int_0^{\frac{\pi}{2}}\frac{\sin(2n+1)u}{u}\,du=\frac{\pi}{2}.$$

Now substitute $(2n+1)u=2\pi Ax$ and divide by π to get

$$\frac{1}{\pi}\int_0^{\frac{\pi}{2}}\frac{\sin(2n+1)u}{u}\,du=\int_0^{\frac{2n+1}{4A}}\frac{\sin 2\pi Ax}{\pi x}\,dx\to\tfrac{1}{2}$$

as $n\to\infty$. For any given $A>0$, letting $n\to\infty$ gives (i). Given $\varepsilon>0$, set $A=\frac{2n+1}{4\varepsilon}$ and let $n\to\infty$ to get (ii). ☐

To prove Theorem 2.13.3, we first note that if $f(t)$ is L^1 then the Fourier integral makes sense, and our task is to understand (2.14.2), or equivalently (2.14.3). The idea is to use the above theorem to say that for any $\varepsilon>0$,

$$\lim_{A\to\infty}\int_\varepsilon^\infty(f(\tau+x)+f(\tau-x))\frac{\sin 2\pi Ax}{\pi x}\,dx=0,$$

so that (2.14.3) is equal to

$$\lim_{A\to\infty}\int_0^\varepsilon(f(\tau+x)+f(\tau-x))\frac{\sin 2\pi Ax}{\pi x}\,dx.$$

So at any point where $\lim\limits_{x\to0}(f(\tau+x)+f(\tau-x))$ exists, the theorem shows that the above integral is equal to $\frac{1}{2}\lim\limits_{x\to0}(f(\tau+x)+f(\tau-x))$. In particular, this holds for piecewise continuous functions.

2.15. Spectrum

How does the Fourier transform tell us about the frequency distribution in the original function? Well, just as in §2.6, the relations (C.1)–(C.3) tell us how to rewrite complex exponentials in terms of sines and cosines, and vice-versa. So the values of $\hat{f}$ at ν and at $-\nu$ tell us not only about the magnitude of the frequency component with frequency ν, but also the phase. If the original function $f(t)$ is real valued, then $\hat{f}(-\nu)$ is the complex conjugate $\hat{f}(\nu)$. The *energy density* at a particular value of ν is defined to be the square of the amplitude $|\hat{f}(\nu)|$,

$$\text{Energy Density}=|\hat{f}(\nu)|^2.$$

Integrating this quantity over an interval will measure the total energy corresponding to frequencies in this interval. But note that both ν and $-\nu$ contribute to energy, so if only positive values of ν are used, we must remember to double the answer.

The usual way to represent the *frequency spectrum* of a real valued signal is to represent the amplitude and the phase of $\hat{f}(\nu)$ separately for positive values of ν. Recall from Appendix C that in polar coordinates, we can write $\hat{f}(\nu)$ as $re^{i\theta}$, where $r = |\hat{f}(\nu)|$ is the amplitude of the corresponding frequency component and θ is the phase. So r is always nonnegative, and we take θ to lie between $-\pi$ and π. Then $\hat{f}(-\nu) = \overline{\hat{f}(\nu)} = re^{-i\theta}$, so we have already represented the information about negative values of ν if we have given both amplitude and phase for positive values of ν.

Parseval's identity states that the total energy of a signal is equal to the total energy in its spectrum:

$$\int_{-\infty}^{\infty} |f(t)|^2 \, dt = \int_{-\infty}^{\infty} |\hat{f}(\nu)|^2 \, d\nu.$$

More generally, if $f(t)$ and $g(t)$ are two functions, it states that

$$\int_{-\infty}^{\infty} f(t)\overline{g(t)} \, dt = \int_{-\infty}^{\infty} \hat{f}(\nu)\overline{\hat{g}(\nu)} \, d\nu. \tag{2.15.1}$$

The term *white noise* refers to a waveform whose spectrum is flat; for pink noise, the spectrum level decreases by 3dB per octave, while for brown noise (named after Brownian motion), the spectrum level decreases by 6dB per octave.

The windowed Fourier transform was introduced by Gabor,[20] and is described as follows. Given a windowing function $\psi(t)$ and a waveform $f(t)$, the windowed Fourier transform is the function of two variables

$$\mathcal{F}_\psi(f)(p, q) = \int_{-\infty}^{\infty} f(t)e^{-2\pi iqt}\psi(t - p) \, dt,$$

for p and q real numbers. This may be thought of as using all possible time translations of the windowing function, and pulling out the frequency components of the result.

Exercises

1. This exercise is for people who run the Windows operating system. Download a copy of Sound Frequency Analyzer from

www.relisoft.com/freeware/index.htm

This is a freeware realtime audio frequency analysing program for a PC running Windows 95 or higher. Plug a microphone into the audio card on your PC and use this program to watch a windowed frequency spectrum analysis of sounds such as your voice, any musical instruments you may have around, and so on. The program uses the fast Fourier transform, see §7.9.

The Windows Media Player contains an elementary oscilloscope. Use "Windows Update" to make sure you have at least version 7 of the Media Player, play your favorite CD, and under View → Visualizations, choose Bars and Waves → Scope.

[20] D. Gabor, *Theory of communication*, J. Inst. Electr. Eng. 93 (1946), 429–457.

Notice how it is almost impossible to get much meaningful information about how the waveform will sound, just by seeing the oscilloscope trace.

2. Find $\int_{-\infty}^{\infty} e^{-x^2}\,dx$.
[Hint: Square the integral and convert to polar coordinates, as in the proof of Theorem 2.13.1]

3. Show that if a is a constant then the Fourier transform of $f(at)$ is $\frac{1}{a}\hat{f}(\frac{\nu}{a})$.

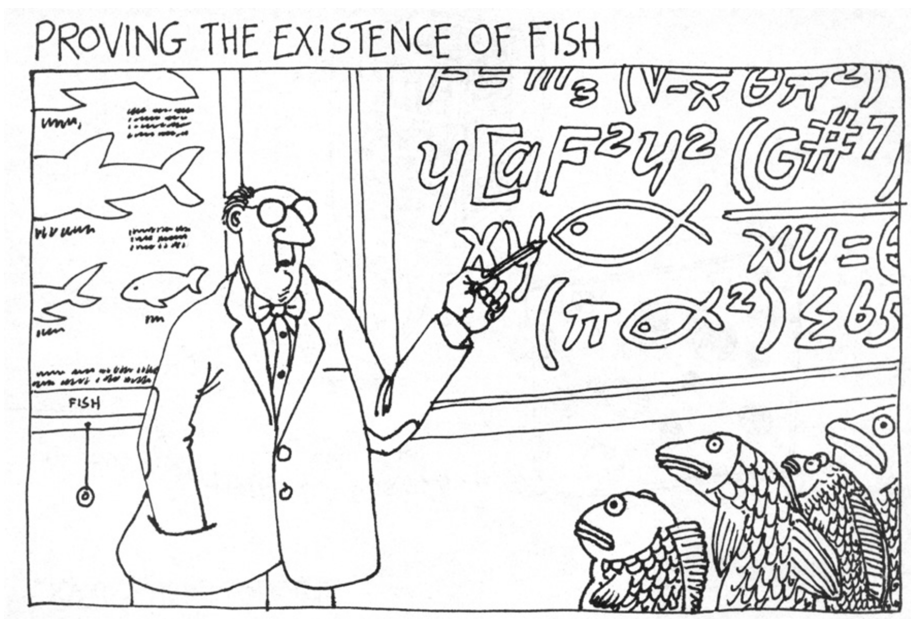
4. Show that if a is a constant then the Fourier transform of $f(t-a)$ is $e^{-2\pi i\nu a}\hat{f}(\nu)$.

5. Find the Fourier transform of the square wave pulse of §2.12

$$f(t) = \begin{cases} 1 & \text{if } -\rho/2 \le t \le \rho/2 \\ 0 & \text{otherwise.} \end{cases}$$

6. Using Theorem 2.13.1 and integration by parts, show that the Fourier transform of $2\pi t^2 e^{-\pi t^2}$ is $(1-2\pi\nu^2)e^{-\pi\nu^2}$.
[Hint: Substitute $x = t + i\nu$ in the integral.]

2.16. The Poisson summation formula



B. Kliban

When we come to study digital music in Chapter 7, we shall need to use the Poisson summation formula.

Theorem 2.16.1 (Poisson's summation formula).

$$\sum_{n=-\infty}^{\infty} f(n) = \sum_{n=-\infty}^{\infty} \hat{f}(n). \tag{2.16.1}$$

PROOF. Define

$$g(\theta) = \sum_{n=-\infty}^{\infty} f\left(\frac{\theta}{2\pi} + n\right).$$

Then the left hand side of the desired formula is $g(0)$. Furthermore, $g(\theta)$ is periodic with period 2π, $g(\theta + 2\pi) = g(\theta)$. So we may apply the theory of Fourier series to $g(\theta)$. By equation (2.6.1), we have

$$g(\theta) = \sum_{n=-\infty}^{\infty} \alpha_n e^{in\theta}$$

and by equation (2.6.2), we have

$$\begin{aligned}
\alpha_m &= \frac{1}{2\pi}\int_0^{2\pi} g(\theta)e^{-im\theta}\,d\theta \\
&= \frac{1}{2\pi}\int_0^{2\pi} \sum_{n=-\infty}^{\infty} f\left(\frac{\theta}{2\pi} + n\right) e^{-im\theta}\,d\theta \\
&= \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} \int_0^{2\pi} f\left(\frac{\theta}{2\pi} + n\right) e^{-im\theta}\,d\theta \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} f\left(\frac{\theta}{2\pi}\right) e^{-im\theta}\,d\theta \\
&= \int_{-\infty}^{\infty} f(t)e^{-2\pi imt}\,dt \\
&= \hat{f}(m).
\end{aligned}$$

The third step above consists of piecing together the real line from segments of length 2π. The fourth step is given by the substitution $\theta = 2\pi t$. Finally, we have

$$\sum_{n=-\infty}^{\infty} f(n) = g(0) = \sum_{n=-\infty}^{\infty} \alpha_n = \sum_{n=-\infty}^{\infty} \hat{f}(n). \qquad \square$$

2.17. The Dirac delta function

Dirac's delta function $\delta(t)$ is defined by the following properties:

(i) $\delta(t) = 0$ for $t \neq 0$, and

(ii) $\displaystyle\int_{-\infty}^{\infty} \delta(t)\,dt = 1$.

Think of $\delta(t)$ as being zero except for a spike at $t = 0$, so large that the area under it is equal to one. The awake reader will immediately notice that these properties are contradictory. This is because changing the value of a function at a single point does not change the value of an integral, and the function is zero except at one point, so the integral should be zero. Later in this section, we'll explain the resolution of this problem, but for the moment, let's continue as though there were no problem, and as though equations (2.13.1) and (2.13.3) work for functions involving $\delta(t)$.

It is often useful to shift the spike in the definition of the delta function to another value of t, say $t = t_0$, by using $\delta(t - t_0)$ instead of $\delta(t)$. The fundamental property of the delta function is that it can be used to pick out the value of another function at a desired point by integrating. Namely, if we want to find the value of $f(t)$ at $t = t_0$, we notice that $f(t)\delta(t-t_0) = f(t_0)\delta(t-t_0)$, because $\delta(t - t_0)$ is only nonzero at $t = t_0$. So

$$\int_{-\infty}^{\infty} f(t)\delta(t - t_0)\,dt = \int_{-\infty}^{\infty} f(t_0)\delta(t - t_0)\,dt = f(t_0)\int_{-\infty}^{\infty} \delta(t - t_0)\,dt = f(t_0).$$

Next, notice what happens if we take the Fourier transform of a delta function. If $f(t) = \delta(t - t_0)$ then by equation (2.13.1)

$$\hat{f}(\nu) = \int_{-\infty}^{\infty} \delta(t - t_0)e^{-2\pi i\nu t}\,dt = e^{-2\pi i\nu t_0}.$$

In other words, the Fourier transform of a delta function $\delta(t - t_0)$ is a complex exponential $e^{-2\pi i\nu t_0}$. In particular, in the case $t_0 = 0$, we find that the Fourier transform of $\delta(t)$ is the constant function 1. The Fourier transform of $\frac{1}{2}(\delta(t - t_0) + \delta(t + t_0))$ is

$$\tfrac{1}{2}(e^{-2\pi i\nu t_0} + e^{2\pi i\nu t_0}) = \cos(2\pi\nu t_0)$$

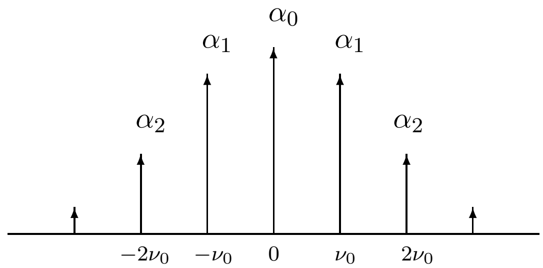
(see equation (C.2)).

Conversely, if we apply the inverse Fourier transform (2.13.3) to the function $\hat{f}(\nu) = \delta(\nu - \nu_0)$, we obtain $f(t) = e^{2\pi i\nu_0 t}$. So we can think of the Dirac delta function concentrated at a frequency ν_0 as the Fourier transform of a complex exponential. Similarly, $\frac{1}{2}(\delta(\nu - \nu_0) + \delta(\nu + \nu_0))$ is the Fourier transform of a cosine wave $\cos(2\pi\nu_0 t)$ with frequency ν_0. We shall justify these manipulations towards the end of this section.

The relationship between Fourier series and the Fourier transform can be made more explicit in terms of the delta function. Suppose that $f(t)$ is a periodic function of t of the form $\sum_{n=-\infty}^{\infty} \alpha_n e^{in\theta}$ (see equation (2.6.1)) where $\theta = 2\pi\nu_0 t$. Then we have

$$\hat{f}(\nu) = \sum_{n=-\infty}^{\infty} \alpha_n \delta(\nu - n\nu_0).$$

So the Fourier transform of a real valued periodic function has a spike at plus and minus each frequency component, consisting of a delta function multiplied by the amplitude of that frequency component.



So what kind of a function is $\delta(t)$? The answer is that it really isn't a function at all, it's a *distribution*, sometimes also called a *generalized function*. A distribution is only defined in terms of what happens when we multiply by a function and integrate. Whenever a delta function appears, there is an implicit integration lurking in the background.

More formally, one starts with a suitable space of *test functions*,[21] and a distribution is defined as a continuous linear map from the space of test functions to the complex numbers (or the real numbers, according to context).

A function $f(t)$ can be regarded as a distribution, namely we identify it with the linear map taking $g(t)$ to $\int_{-\infty}^{\infty} f(t)g(t)\,dt$, as long as this makes sense. The delta function is the distribution which is defined as the linear map taking a test function $g(t)$ to $g(0)$. It is easy to see that this distribution does not come from an ordinary function in the above way. The argument is given at the beginning of this section. But we write distributions as though they were functions, and we write integration for the value of a distribution on a function. So for example the distribution $\delta(t)$ is defined by $\int_{-\infty}^{\infty} \delta(t)g(t)\,dt = g(0)$, and this just means that the value of the distribution $\delta(t)$ on the test function $g(t)$ is $g(0)$, nothing more nor less.

There is one warning that must be stressed at this stage. Namely, it does not make sense to multiply distributions. So for example, the square of the delta function does not make sense as a distribution. After all, what would $\int_{-\infty}^{\infty} \delta(t)^2 g(t)\,dt$ be? It would have to be $\delta(0)g(0)$, which isn't a number!

However, distributions can be multiplied by functions. The value of a distribution times $f(t)$ on $g(t)$ is equal to the value of the original distribution on $f(t)g(t)$. As long as $f(t)$ has the property that whenever $g(t)$ is a

[21]In the context of the theory of Fourier transforms, it is usual to start with the *Schwartz space* $\mathcal{S}$ consisting of infinitely differentiable functions $f(t)$ with the property that there is a bound not depending on m and n for the value of any derivative $f^{(m)}(t)$ times any power t^n of t ($m, n \geq 0$). So these functions are very smooth and all their derivatives tend to zero very rapidly as $|t| \to \infty$. An example of a function in $\mathcal{S}$ is the function e^{-t^2}. The sum, product and Fourier transform of functions in $\mathcal{S}$ are again in $\mathcal{S}$. For the purpose of saying what it means for a linear map on $\mathcal{S}$ to be continuous, the distance between two functions $f(t)$ and $g(t)$ in $\mathcal{S}$ is defined to be the largest distance between the values of $t^n f^{(m)}(t)$ and $t^n g^{(m)}(t)$ as m and n run through the nonnegative integers. The space of distributions defined on $\mathcal{S}$ is written $\mathcal{S}'$. Distributions in $\mathcal{S}'$ are called *tempered* distributions.

test function then so is $f(t)g(t)$, this makes sense. Test functions and polynomials satisfy this condition, for example.

Distributions can also be differentiated. The way this is done is to use integration by parts to give the *definition* of differentiation. So if $f(t)$ is a distribution and $g(t)$ is a test function then $f'(t)$ is defined via

$$\int_{-\infty}^{\infty} f'(t)g(t)\,dt = -\int_{-\infty}^{\infty} f(t)g'(t)\,dt.$$

So for example the value of the distribution $\delta'(t)$ on the test function $g(t)$ is $-g'(0)$.

To illustrate how to manipulate distributions, let us find $t\delta'(t)$. Integration by parts shows that if $g(t)$ is a test function, then

$$\int_{-\infty}^{\infty} t\delta'(t)g(t)\,dt = -\int_{-\infty}^{\infty} \delta(t)\frac{d}{dt}(tg(t))\,dt = -\int_{-\infty}^{\infty} \delta(t)(tg'(t)+g(t))\,dt.$$

Now $t\delta(t) = 0$, so this gives $-g(0)$. If two distributions take the same value on all test functions, they are by definition the same distribution. So we have

$$\boxed{t\delta'(t) = -\delta(t).}$$

The reader should be warned, however, that extreme caution is necessary when playing with equations of this kind. For example, dividing the above equation by t to get $\delta'(t) = -\delta(t)/t$ makes no sense at all. After all, what if we were to apply the same logic to the equation $t\delta(t) = 0$?

It is also useful at this stage to go back to the proof of Fejér's theorem give in §2.7. Basically, the reason why this proof works is that the functions $K_m(y)$ are finite approximations to the distribution $2\pi\delta(y)$. Approximations to delta functions, used in this way, are called *kernel functions*, and they play a very important role in the theory of partial differential equations, analogous to the role they play in the proof of Fejér's theorem.

The Fourier transform of a distribution is defined using Parseval's identity (2.15.1). Namely, if $f(t)$ is a distribution, then for any function $\overline{g(t)}$ the quantity $\int_{-\infty}^{\infty} f(t)\overline{g(t)}\,dt$ denotes the value of the distribution on $\overline{g(t)}$. We define $\hat{f}(\nu)$ to be the distribution whose value on $\overline{\hat{g}(\nu)}$ is the same quantity. In other words, the *definition* of $\hat{f}(\nu)$ is

$$\int_{-\infty}^{\infty} \hat{f}(\nu)\overline{\hat{g}(\nu)}\,d\nu = \int_{-\infty}^{\infty} f(t)\overline{g(t)}\,dt.$$

Notice that even if we are only interested in *functions*, this considerably extends the definition of Fourier transforms, and that the Fourier transform of a function can easily end up being a distribution which is not a function. For example, we saw earlier that the Fourier transform of the function $e^{2\pi i\nu_0 t}$ is the distribution $\delta(\nu - \nu_0)$.

Exercises

1. Find the Fourier transform of the sine wave $f(t) = \sin(2\pi\nu_0 t)$ in terms of the Dirac delta function.

2. Show that if C is a constant then

$$\delta(Ct) = \frac{1}{|C|}\delta(t).$$

3. The Heaviside function $H(t)$ is defined by

$$H(t) = \begin{cases} 1 & \text{if } t \geq 0 \\ 0 & \text{if } t < 0. \end{cases}$$

Prove that the derivative of $H(t)$ is equal to the Dirac delta function $\delta(t)$.
[Hint: Use integration by parts]

4. Show that $t\delta(t) = 0$.

5. Using Theorem 2.13.2, show that the Fourier transform of t^n is $\left(\frac{-1}{2\pi i}\right)^n \delta^{(n)}(\nu)$, where $\delta^{(n)}$ is the nth derivative of the Dirac delta function.

Further reading:

F. G. Friedlander and M. Joshi, *Introduction to the theory of distributions*, second edition, CUP, 1998.

A. H. Zemanian, *Distribution theory and transform analysis*, Dover, 1987.

2.18. Convolution

The Fourier transform does not preserve multiplication. Instead, it turns it into *convolution*. If $f(t)$ and $g(t)$ are two test functions, their convolution $f * g$ is defined by

$$(f * g)(t) = \int_{-\infty}^{\infty} f(s)g(t-s)\,ds.$$

The corresponding verb is to *convolve* the function f with the function g. The formula also makes sense if one of f and g is a distribution and the other is a test function. The result is a function, but not necessarily a test function. The convolution of two distributions sometimes but not always makes sense; for example, the convolution of two constant functions is not defined but the convolution of two Dirac delta functions is defined.

It is easy to check that the following properties of convolution hold whenever both sides make sense.

(i) (commutativity) $f * g = g * f$.

(ii) (associativity) $(f * g) * h = f * (g * h)$.

(iii) (distributivity) $f * (g + h) = f * g + f * h$.

(iv) (identity element) $\delta * f = f * \delta = f$.

Here, δ denotes the Dirac delta function.

THEOREM 2.18.1. (i) $\widehat{f * g}(\nu) = \hat{f}(\nu)\hat{g}(\nu)$,
(ii) $\widehat{fg}(\nu) = (\hat{f} * \hat{g})(\nu)$.

PROOF. To prove part (i), from the definition of convolution we have

$$\begin{aligned}\widehat{f * g}(\nu) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s)g(t-s)e^{-2\pi i\nu t}\,ds\,dt \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s)g(u)e^{-2\pi i\nu(s+u)}\,ds\,du \\ &= \left(\int_{-\infty}^{\infty} f(s)e^{-2\pi i\nu s}\,ds\right)\left(\int_{-\infty}^{\infty} g(u)e^{-2\pi i\nu u}\,du\right) \\ &= \hat{f}(\nu)\hat{g}(\nu).\end{aligned}$$

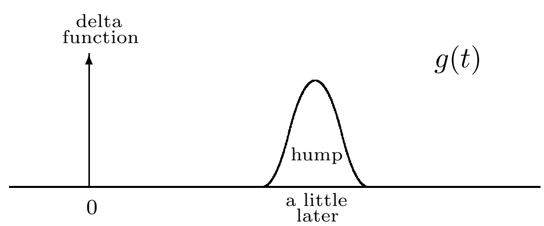
Here, we have made the substitution $u = t - s$. Part (ii) follows from part (i) by the Fourier inversion formula (2.13.3); in other words, by reversing the roles of t and ν. □

Part (i) of this theorem can be interpreted in terms of frequency filters. Applying a frequency filter to an audio signal is supposed to have the effect of multiplying the frequency distribution of the signal, $\hat{f}(\nu)$, by a filter function $\hat{g}(\nu)$. So in the time domain, this corresponds to convolving the signal $f(t)$ with $g(t)$, the inverse Fourier transform of the filter function.

The output of a filter is usually taken to depend only on the input at the current and previous times. Looking at the formula for convolution, this corresponds to the statement that $g(t)$, the inverse Fourier transform of the filter function, should be zero for negative values of its argument.

The function $g(t)$ for the filter is called the *impulse response*, because it represents the output when a delta function is present at the input. The statement that $g(t) = 0$ for $t < 0$ is a manifestation of *causality*.

For example, let $g(t)$ be a delta function at zero plus a hump a little later.



Then convolving a signal $f(t)$ with $g(t)$ will give $f(t)$ plus a smeared echo of $f(t)$ a short time later. The graph of $g(t)$ is interpreted as the impulse response, namely what comes out when a delta function is put in (in this case, crack — thump). These days, effects are often added to sound using a *digital filter*, which uses a discrete version of this process of convolution. See §7.8 for a brief description of the theory.

Exercises

1. Show that $\delta' * f = -f'$. Find a formula for $\delta^{(n)} * f$.

2. Prove the associativity formula $(f * g) * h = f * (g * h)$.

Further reading:

Curtis Roads, *Sound transformation by convolution*, appears as article 12 of Roads et al [**115**], pages 411–438.

2.19. Cepstrum

The idea of *cepstrum* is to look for periodicity in the Fourier transform of a signal, but measured on a logarithmic scale. So for example, this would pick out a series of frequency components separated by octaves. So the definition of the cepstrum of a signal is

$$\widehat{\ln \hat{f}}(\rho) = \int_{-\infty}^{\infty} e^{-2\pi i \rho \nu} \ln \hat{f}(\nu)\, d\nu.$$

This gives a sort of twisted up, backwards spectrum. The idea was first introduced by Bogert, Healy and Tukey, who introduced the terminology. The variable ρ is called *quefrency*, to indicate that it is a twisted version of frequency. Peaks of quefrency are called *rahmonics*.

If filtering a signal corresponds to multiplying its Fourier transform by a function, then *liftering* a signal is achieved by finding the cepstrum, multiplying by a function, and then undoing the cepstrum process. This process is often used in the analysis of vocal signals, in order to locate and extract formants.

Further reading:

B. P. Bogert, M. J. R. Healy and J. W. Tukey, *Quefrency analysis of time series for echoes: cepstrum, pseudo-autocovariance, cross-cepstrum and saphe cracking.* In *Proceedings of the Symposium on Time Series Analysis*, New York, Wiley 1963, pages 209–243.

Judith C. Brown, *Computer identification of wind instruments using cepstral coefficients*, Proceedings of the 16th International Congress on Acoustics and 135th Meeting of the Acoustical Society of America, Seattle, Washington (1998), 1889–1890.

Judith C. Brown, *Computer identification of musical instruments using pattern recognition with cepstral coefficients as features*, J. Acoust. Soc. Am. 105 (3) (1999), 1933–1941.

M. R. Schroeder, *Computer speech*, Springer Series in Information Sciences, Springer-Verlag, 1999, §10.14 and Appendix B.

Stan Tempelaars, *Signal processing, speech and music* [**133**], §7.2.

2.20. The Hilbert transform and instantaneous frequency

Although the notion of instantaneous *frequency spectrum* of a signal makes no sense (because of the Heisenberg uncertainty principle), there is a notion of *instantaneous frequency* of a signal at a point in time. The idea is to use the Hilbert transform. If $f(t)$ is the signal, its Hilbert transform $g(t)$ is defined to be the Cauchy principal value[22] of the integral

$$g(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{f(\tau)}{t-\tau}\,d\tau.$$

This makes an *analytic* signal $f(t) + ig(t)$.

For example, if $f(t) = c\cos(\omega t + \phi)$ then $g(t) = c\sin(\omega t + \phi)$ and $f(t) + ig(t) = ce^{i(\omega t+\phi)}$. In this case, $f(t) + ig(t)$ is rotating counterclockwise around the origin of the complex plane at a rate of ω radians per unit time. This suggests that the instantaneous angular frequency $\omega(t)$ is defined as the rate at which $f(t) + ig(t)$ is rotating around the origin. The angle $\theta(t)$ satisfies[23]

$$\tan\theta = g(t)/f(t),$$

so differentiating, we obtain

$$\sec^2\theta\,\frac{d\theta}{dt} = \frac{f(t)g'(t) - g(t)f'(t)}{f(t)^2}.$$

Using the relation

$$\sec^2\theta = 1 + \tan^2\theta = \frac{f(t)^2 + g(t)^2}{f(t)^2},$$

we obtain

$$\omega(t) = \frac{d\theta}{dt} = \frac{f(t)g'(t) - g(t)f'(t)}{f(t)^2 + g(t)^2}.$$

So the instantaneous frequency is given by

$$\nu(t) = \frac{\omega(t)}{2\pi} = \frac{1}{2\pi}\frac{f(t)g'(t) - g(t)f'(t)}{f(t)^2 + g(t)^2}.$$

The same reasoning also leads to the notion of *instantaneous amplitude* whose value is $\sqrt{f(t)^2 + g(t)^2}$. This is not the same as $|f(t)|$, which fails to capture the notion of instantaneous amplitude of a signal even for a sine wave.

From the formula for Hilbert transform, it can be seen that the definitions of instantaneous frequency and amplitude depend mostly on information about the signal close to the point being considered, but they do also have small contributions from the behavior far away.

---

[22] i.e., $g(t) = \lim_{A\to\infty}\lim_{\varepsilon\to 0}\left(\int_{-A}^{-\varepsilon}\frac{f(\tau)}{t-\tau}\,d\tau + \int_{\varepsilon}^{A}\frac{f(\tau)}{t-\tau}\,d\tau\right)$.

[23] The formula $\theta = \tan^{-1}(g(t)/f(t))$ is incorrect. Why?

Further reading:

B. Boashash, *Estimating and interpreting the instantaneous frequency of a signal—Part I: Fundamentals*, Proc. IEEE 80 (1992), 520–538.

L. Rossi and G. Girolami, *Instantaneous frequency and short term Fourier transforms: Applications to piano sounds*, J. Acoust. Soc. Am. 110 (5) (2001), 2412–2420.

Zachary M. Smith, Bertrand Delgutte and Andrew J. Oxenham, *Chimaeric sounds reveal dichotomies in auditory perception*, Nature 416, 7 March 2002, 87–90. This article discusses the Fourier transform and Hilbert transform as models for auditory perception of music and speech, and concludes that both play a role.

2.21. Wavelets

The wavelet transform is a relative of the windowed Fourier transform, in which all possible time translations and dilations are applied to a given window, to give a function of two variables as the transform. The exponential functions used in the windowed Fourier transforms are no longer present, but in some sense they are replaced by the use of dilations on the windowing function.

To be more precise, a *wavelet* is a function $\psi(t)$ of a real variable t which satisfies the *admissibility condition*

$$0 < c_\psi < \infty$$

where c_ψ is the constant defined by

$$c_\psi = \int_{-\infty}^{\infty} \frac{|\widehat{\psi}(\nu)|^2}{|\nu|}\, d\nu.$$

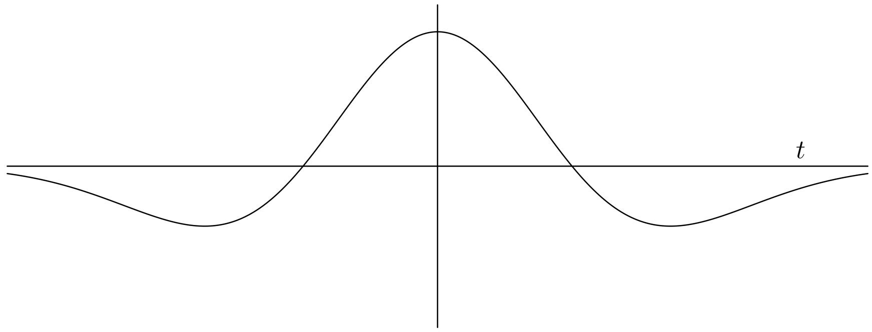
The wavelet ψ is chosen once and for all, and is interpreted as the shape of the window. The *wavelet transform* $L_\psi(f)$ of a waveform f is defined as the function of two variables

$$L_\psi(f)(a,b) = \frac{1}{\sqrt{|a|c_\psi}} \int_{-\infty}^{\infty} f(t)\psi\left(\frac{t-b}{a}\right) dt$$

for real $a \neq 0$ and b.

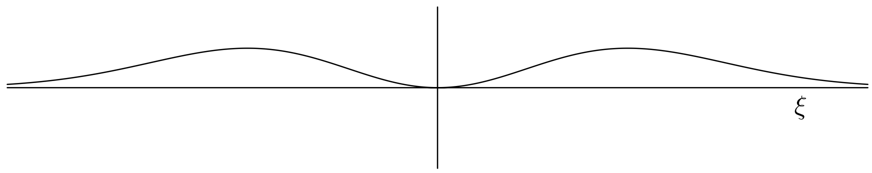
An example of a wavelet often used in practice is the *Mexican hat*, defined by

$$\psi(t) = (1 - 2\pi t^2)e^{-\pi t^2}.$$



The Fourier transform of the Mexican hat is

$$\widehat{\psi}(\nu) = 2\pi\nu^2 e^{-\pi\nu^2}$$



and we have $c_\psi = 1$.

The inverse wavelet transform L_ψ^* with respect to ψ is defined as follows. If $g(a,b)$ is a function of two real variables, then $L_\psi^*(g)$ is the function of the single real variable t defined by

$$L_\psi^*(g)(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{|a|c_\psi}} g(a,b)\psi\Big(\frac{t-b}{a}\Big)\,\frac{da\,db}{a^2}.$$

Note that at $a = 0$ the integrand is not defined, so the integral with respect to a simply misses out this value.

Theorem 2.21.1. *If $f(t)$ is a square integrable function of a real variable t then $L_\psi^* L_\psi f$ agrees with f at almost all values of t, and in particular, at all points where $f(t)$ is continuous.*

Further reading:

G. Evangelista, *Wavelet representations of musical signals*, appears as article 4 in Roads et al [**115**], pages 127–154.

R. Kronland–Martinet, *The wavelet transform for the analysis, synthesis, and processing of speech and music sounds*, Computer Music Journal 12 (4) (1988), 11–20.

A. K. Louis, P. Maaß and A. Rieder, *Wavelets, theory and applications*, Wiley, 1997. ISBN 0471967920.

Stéphane Mallat, *A wavelet tour of signal processing*, Academic Press, 1998. ISBN 0124666051.

P. Polotti and G. Evangelista, *Fractal additive synthesis via harmonic-band wavelets*, Computer Music Journal 25 (3) (2001), 22–37.

Curtis Roads, *The computer music tutorial* [**113**], pages 581–589.

CHAPTER 3

A mathematician's guide to the orchestra

3.1. Introduction

Ethnomusicologists classify musical instruments into five main categories, which correspond reasonably well to the mathematical description of the sound they produce.[1]

1. Idiophones, where sound is produced by the body of a vibrating instrument. This category includes percussion instruments other than drums. It is divided into four subcategories: struck idiophones such as xylophones and cymbals, plucked idiophones (lamellophones) such as the mbira and the balafon, friction idiophones such as the (bowed) saw, and blown idiophones such as the aeolsklavier (a nineteenth century German instrument in which wooden rods are blown by bellows).

2. Membranophones, where the sound is produced by the vibration of a stretched membrane; for example, drums are membranophones. This category also has four subdivisions: struck drums, plucked drums, friction drums, and singing membranes such as the kazoo.

3. Chordophones, where the sound is produced by one or more vibrating strings. This category includes not only stringed instruments such as the violin and harp, but also keyboard instruments such as the piano and harpsichord.

4. Aerophones, where the sound is produced by a vibrating column of air. This category includes woodwind instruments such as the flute, clarinet and oboe, brass instruments such as the trombone, trumpet and French horn, and also various more exotic instruments such as the bullroarer and the conch shell.

[1] This classification was due to Sachs and Hornbostel (Zeitschrift für Musik, 1914), who omitted the fifth category of electrophones. This last category was added in 1961 by Anthony Baines and Klaus P. Wachsmann in their translation of the article of Sachs and Hornbostel. This translation appears in *The Garland library of readings in ethnomusicology*, 6, ed. Kay K. Shelemay, 119–145, Garland, New York, 1961.

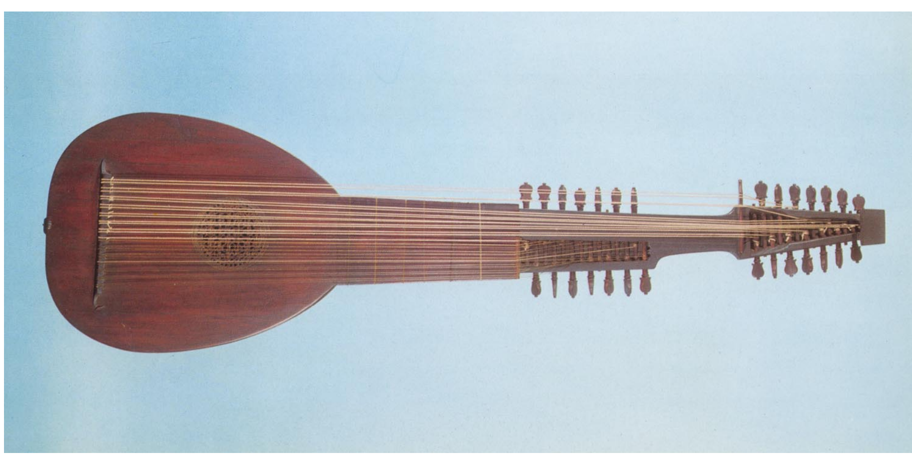
The Sachs–Hornbostel system had antecedents. A Hindu system dating back more than two thousand years divides instruments into four similar groups. Victor Mahillon, curator of the collection of musical instruments of the Brussels conservatoire, used a similar classification in his 1888 catalog of the collection.

5. Electrophones, where the sound is produced primarily by electrical or electronic means. This includes the modern electronic synthesizer (analog or digital) as well as sound generated by a computer program. One of the earliest electrophones was the theremin. An instrument such as an electric guitar, where the sound is produced mechanically and amplified and manipulated electronically, is not classified as an electrophone. An electric guitar is an example of a chordophone.

There are two main components that determine the nature of the sound coming from a musical instrument, namely the initial transient part of the sound, and the set of resonant frequencies making up the spectrum of the rest of the sound. Initial transients are notoriously difficult to describe mathematically, but have a profound effect on our perception of the sound. We shall return to this subject in Chapter 8. In this chapter, we shall concentrate on the description of resonant frequencies. It is this aspect of sound which is most relevant to the study of musical scales.

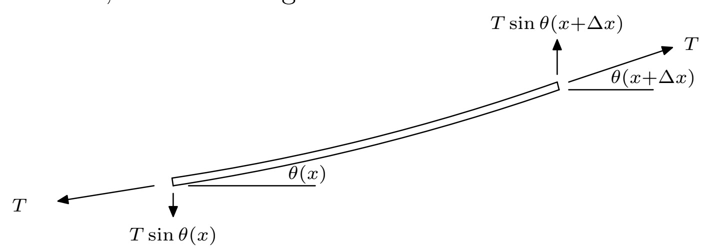
We begin with chordophones, where we need to understand the solutions of the one dimensional wave equation. This is followed by aerophones, which are mathematically very similar. Membranophones require us to solve the two dimensional wave equation, which gets us involved with Bessel functions. Finally, idiophones involve a more complicated equation of degree four. We leave electrophones until Chapter 8.

3.2. The wave equation for strings



Italian Theorboe, Musée Instrumental, Brussels, Belguim

In this section, we return to the subject of §1.6, and consider the relevance of Fourier series to the vibration of a string held at both ends. To make a more accurate analysis, we need to regard the displacement y as a function both of time t and position x along the string. Since y is being regarded as a function of two variables, the appropriate equations are written in terms of *partial derivatives*, and Appendix P gives a brief summary of partial derivatives. The equation describing the vibration of a string is called the *wave equation* in one dimension, which we now develop. This equation supposes that the displacement of the string is such that its slope at any point along its length at any time is small. For large displacements, the analysis is harder. Note that we are only concerned here with *transverse waves*, namely motion perpendicular to the string. Motion parallel to the string is called *longitudinal waves*, and will be ignored here.



Write T for the tension on the string (in newtons = $\mathrm{kg\,m/s^2}$), and ρ for the linear density of the string (in kg/m). Then at position x along the string, the angle $\theta(x)$ between the string and the horizontal will satisfy $\tan\theta(x) = \dfrac{\partial y}{\partial x}$. On a small segment of string from x to $x+\Delta x$, the vertical component of force at the left end will be $-T\sin\theta(x)$, and at the right end will be $T\sin\theta(x+\Delta x)$.

Provided that $\theta(x)$ is small, $\sin\theta(x)$ and $\tan\theta(x)$ are approximately equal. So the difference in vertical components of force between the two ends of the segment will be approximately

$$\begin{aligned} T\tan\theta(x+\Delta x) - T\tan\theta(x) &= T\left(\frac{\partial y(x+\Delta x)}{\partial x} - \frac{\partial y(x)}{\partial x}\right) \\ &= T\Delta x\,\frac{\dfrac{\partial y(x+\Delta x)}{\partial x} - \dfrac{\partial y(x)}{\partial x}}{\Delta x} \\ &\approx T\Delta x\frac{\partial^2 y}{\partial x^2}. \end{aligned} \tag{3.2.1}$$

The mass of the segment of string will be approximately $\rho\Delta x$. So Newton's law ($F = ma$) for the acceleration $a = \dfrac{\partial^2 y}{\partial t^2}$ gives

$$T\Delta x\frac{\partial^2 y}{\partial x^2} \approx (\rho\Delta x)\frac{\partial^2 y}{\partial t^2}.$$

Jean-le-Rond d'Alembert (1717–1783)

Cancelling a factor of Δx on both sides gives

$$T\frac{\partial^2 y}{\partial x^2} \approx \rho\frac{\partial^2 y}{\partial t^2}.$$

In other words, as long as $\theta(x)$ never gets large, the motion of the string is essentially determined by the wave equation

$$\boxed{\frac{\partial^2 y}{\partial t^2} = c^2\frac{\partial^2 y}{\partial x^2}} \tag{3.2.2}$$

where $c = \sqrt{T/\rho}$.

D'Alembert[2] discovered a strikingly simple method for finding the general solution to equation (3.2.2). Roughly speaking, his idea is to factorize the differential operator

$$\frac{\partial^2}{\partial t^2} - c^2\frac{\partial^2}{\partial x^2}$$

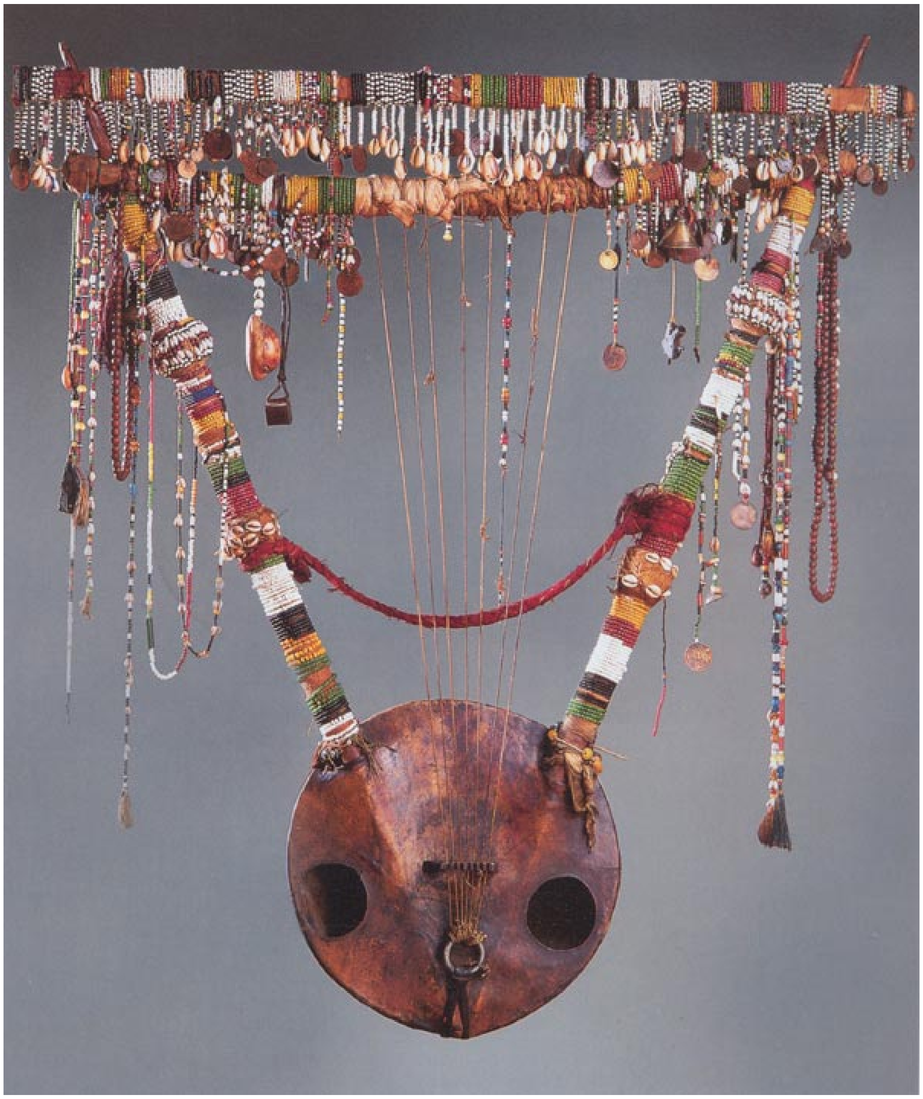
as

$$\Big(\frac{\partial}{\partial t} + c\frac{\partial}{\partial x}\Big)\Big(\frac{\partial}{\partial t} - c\frac{\partial}{\partial x}\Big).$$

More precisely, we make a change of variables

$$u = x + ct, \qquad v = x - ct.$$

[2]Jean-le-Rond d'Alembert was born in Paris on November 16, 1717, and died there on October 29, 1783. He was the illegitimate son of a chevalier by the name of Destouches, and was abandoned by his mother on the steps of a small church called St. Jean-le-Rond, from which his first name is taken. He grew up in the family of a glazier and his wife, and lived with his adoptive mother until she died in 1757. But his father paid for his education, which allowed him to be exposed to mathematics. Two essays written in 1738 and 1740 drew attention to his mathematical abilities, and he was elected to the French Academy in 1740. Most of his mathematical works were written there in the years 1743–1754, and his solution of the wave equation appeared in his paper: *Recherches sur la courbe que forme une corde tendue mise en vibration*, Hist. Acad. Sci. Berlin 3 (1747), 214–219.



19th century lyre found in Nuba Hills, Sudan. British Museum, London.

Then by the multivariable form of the chain rule, we have

$$\frac{\partial y}{\partial t} = \frac{\partial y}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial y}{\partial v}\frac{\partial v}{\partial t} = c\frac{\partial y}{\partial u} - c\frac{\partial y}{\partial v}.$$

Differentiating again, we have

$$\begin{aligned}\frac{\partial^2 y}{\partial t^2} &= \frac{\partial}{\partial u}\Big(\frac{\partial y}{\partial t}\Big)\frac{\partial u}{\partial t} + \frac{\partial}{\partial v}\Big(\frac{\partial y}{\partial t}\Big)\frac{\partial v}{\partial t} \\ &= c\Big(c\frac{\partial^2 y}{\partial u^2} - c\frac{\partial^2 y}{\partial u \partial v}\Big) - c\Big(c\frac{\partial^2 y}{\partial v \partial u} - c\frac{\partial^2 y}{\partial v^2}\Big) \\ &= c^2\left(\frac{\partial^2 y}{\partial u^2} - 2\frac{\partial^2 y}{\partial u \partial v} + \frac{\partial^2 y}{\partial v^2}\right).\end{aligned}$$

Similarly,

$$\begin{aligned}\frac{\partial y}{\partial x} &= \frac{\partial y}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial y}{\partial v}\frac{\partial v}{\partial x} = \frac{\partial y}{\partial u} + \frac{\partial u}{\partial x},\\ \frac{\partial^2 y}{\partial x^2} &= \frac{\partial^2 y}{\partial u^2} + 2\frac{\partial^2 y}{\partial u \partial v} + \frac{\partial^2 y}{\partial v^2}.\end{aligned}$$

Then equation (3.2.2) becomes

$$c^2\left(\frac{\partial^2 y}{\partial u^2} - 2\frac{\partial^2 y}{\partial u \partial v} + \frac{\partial^2 y}{\partial v^2}\right) = c^2\left(\frac{\partial^2 y}{\partial u^2} + 2\frac{\partial^2 y}{\partial u \partial v} + \frac{\partial^2 y}{\partial v^2}\right)$$

or

$$\boxed{\frac{\partial^2 y}{\partial u \partial v} = 0.}$$

This equation may be integrated directly to see that the general solution is given by $y = f(u) + g(v)$ for suitably chosen functions f and g. Substituting back, we obtain

$$\boxed{y = f(x + ct) + g(x - ct).}$$

This represents a superposition of two waves, one travelling to the right and one travelling to the left, each with velocity c.

Now the boundary conditions tell us that the left and right ends of the string are fixed, so that when $x = 0$ or $x = \ell$ (the length of the string), we have $y = 0$ (independent of t). The condition with $x = 0$ gives

$$0 = f(ct) + g(-ct)$$

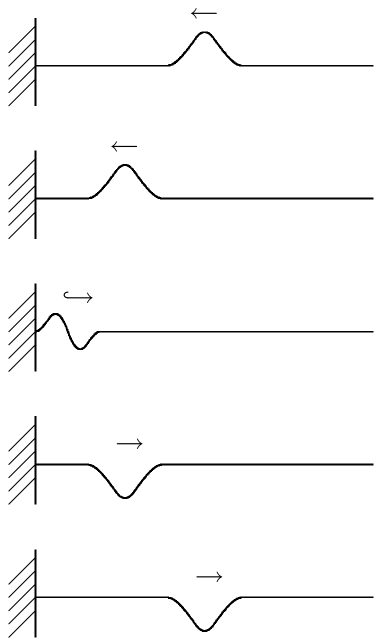
for all t, so that

$$g(\lambda) = -f(-\lambda) \tag{3.2.3}$$

for any value of λ. Thus

$$y = f(x + ct) - f(ct - x).$$

Physically, this means that the wave travelling to the left hits the end of the string and returns inverted as a wave travelling to the right. This is called the "principle of reflection."



Substituting the other boundary condition $x = \ell$, $y = 0$ gives $f(\ell+ct) = f(ct - \ell)$ for all t, so that

$$f(\lambda) = f(\lambda + 2\ell) \tag{3.2.4}$$

for all values of λ. We summarize all the above information in the following theorem.

THEOREM 3.2.1 (d'Alembert). *The general solution of the wave equation*

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2}$$

is given by

$$y = f(x + ct) + g(x - ct).$$

The solutions satisfying the boundary conditions $y = 0$ for $x = 0$ and for $x = \ell$, for all values of t, are of the form

$$y = f(x + ct) - f(-x + ct) \tag{3.2.5}$$

where f satisfies $f(\lambda) = f(\lambda + 2\ell)$ for all values of λ.

One interesting feature of d'Alembert's solution to the wave equation is worth emphasizing. Although the wave equation only makes sense for functions with second order partial derivatives, the solutions make sense for any *continuous* periodic function f. (Discontinuous functions cannot represent displacement of an unbroken string!) This allows us, for example, to make sense of the plucked string, where the initial displacement is continuous, but not even once differentiable. This is a common phenomenon when solving partial differential equations. A technique which is very often used is to

Marin Mersenne (1588–1648)

rewrite the equation as an integral equation, meaning an equation involving integrals rather than derivatives. Integrable functions are much more general than differentiable functions, so one should expect a more general class of solutions.

Equation (3.2.4) means that the function f appearing in d'Alembert's solution is periodic with period 2ℓ, so that f has a Fourier series expansion. So for example if only the fundamental frequency is present, then the function $f(x)$ takes the form $f(x) = C\cos((\pi x/\ell) + \phi)$. If only the nth harmonic is present, then we have $f(x) = C\cos((n\pi x/\ell) + \phi)$,

$$y = C\cos\left(\frac{n\pi(x+ct)}{\ell} + \phi\right) - C\cos\left(\frac{n\pi(-x+ct)}{\ell} + \phi\right). \tag{3.2.6}$$

The theory of Fourier series allows us to write the general solution as a combination of the above harmonics, as long as we take care of the details of what sort of functions are allowed and what sort of convergence is intended.

Using equation (1.7.12), we can rewrite the nth harmonic solution (3.2.6) as

$$y = 2C\sin\left(\frac{n\pi x}{\ell}\right)\sin\left(\frac{n\pi ct}{\ell} + \phi\right). \tag{3.2.7}$$

This is Bernoulli's solution to the wave equation.[3] Thus the frequency of the nth harmonic is given by $2\pi\nu = n\pi c/\ell$, or replacing c by its value $\sqrt{T/\rho}$,

$$\boxed{\nu = (n/2\ell)\sqrt{T/\rho}.}$$

This formula for frequency was essentially discovered by Marin Mersenne[4] as his "laws of stretched strings." These say that the frequency of a stretched string is inversely proportional to its length, directly proportional to the square root of its tension, and inversely proportional to the square root of the linear density.

Exercises

1. Piano wire is manufactured from steel of density approximately 5,900 kg/m^3. The manufacturers recommend a stress of approximately 1.1×10^9 Newtons/m^2. What is the speed of propagation of waves along the wire? Does it depend on cross-sectional area? How long does the string need to be to sound middle C (262 Hz)?

2. By what factor should the tension on a string be increased, to raise its pitch by a perfect fifth? Assume that the length and linear density remain constant.
[A perfect fifth represents a frequency ratio of 3:2]

3. Read the beginning of Appendix M on music theory, and then explain why the back of a grand piano is shaped in a good approximation to an exponential curve.

3.3. Initial conditions

In this section, we see that in the analysis of the wave equation (3.2.2) described in the last section, specifying the initial position and velocity of each point on the string uniquely determines the subsequent motion.

Let $s_0(x)$ and $v_0(x)$ be the initial vertical and velocity of the string as functions of the horizontal coordinate x, for $0 \le x \le \ell$. These must satisfy $s_0(0) = s_0(\ell) = 0$ and $v_0(0) = v_0(\ell) = 0$ to fit with the boundary conditions at the two ends of the string.

The first step is to extend the definitions of s_0 and v_0 to all values of x using the reflection principle. If we specify that $s_0(-x) = -s_0(x)$ and $v_0(-x) = -v_0(x)$, so that s_0 and v_0 are odd functions of x, this extends the domain of definition to the values $-\ell \le x \le \ell$. The values match up at $-\ell$ and ℓ, so we can extend to all values of x by specifying periodicity with period 2ℓ; namely that $s_0(x + 2\ell) = s_0(x)$ and $v_0(x + 2\ell) = v_0(x)$.

---

[3] Daniel Bernoulli, *Réflections et éclairissements sur les nouvelles vibrations des cordes, Exposées dans les Mémoires de l'Academie de 1747 et 1748*, Royal Academy, Berlin, (1755), 147ff.

[4] Marin Mersenne, *Harmonie Universelle*, Sebastien Cramoisy, Paris, 1636–37. Translated by R. E. Chapman as *Harmonie Universelle: The Books on Instruments*, Martinus Nijhoff, The Hague, 1957. Also republished in French by the CNRS in 1975 from a copy annotated by Mersenne.

Now we simply substitute into the solution given by d'Alembert's theorem. Namely, we know that

$$y = f(x+ct) - f(-x+ct) \tag{3.3.1}$$

where f is periodic with period 2ℓ. Differentiating with respect to t gives the formula for velocity

$$\frac{\partial y}{\partial t} = cf'(x+ct) - cf'(x-ct).$$

Substituting $t = 0$ in both the equation and its derivative gives the following equations

$$f(x) - f(-x) = s_0(x) \tag{3.3.2}$$

$$cf'(x) - cf'(-x) = v_0(x). \tag{3.3.3}$$

Integrating equation (3.3.3) and noting that $v_0(0) = 0$, we obtain

$$cf(x) + cf(-x) = \int_0^x v_0(u)\,du.$$

We divide this equation by c to obtain a formula for $f(x) + f(-x)$. So we can then add equation (3.3.2) and divide by two to obtain $f(x)$. This gives

$$f(x) = \tfrac{1}{2}s_0(x) + \frac{1}{2c}\int_0^x v_0(u)\,du.$$

Putting this back into equation (3.3.1) gives

$$y = \tfrac{1}{2}(s_0(x+ct) - s_0(-x+ct)) + \frac{1}{2c}\left(\int_0^{x+ct} v_0(u)\,du - \int_0^{-x+ct} v_0(u)\,du\right).$$

Using the fact that v_0 is an odd function, we have

$$\int_{x-ct}^{-x+ct} v_0(u)\,du = 0.$$

So we can rewrite the solution as

$$\boxed{y = \tfrac{1}{2}(s_0(x+ct) - s_0(-x+ct)) + \frac{1}{2c}\int_{x-ct}^{x+ct} v_0(u)\,du.}$$

It is now easy to check that this is the unique solution satisfying both the initial conditions and the boundary conditions.

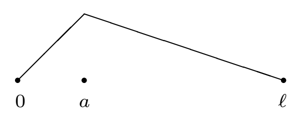
So for example, if the initial velocity is zero, as is the case for a plucked string, then the solution is given by

$$y = \tfrac{1}{2}(s_0(x+ct) - s_0(-x+ct)).$$

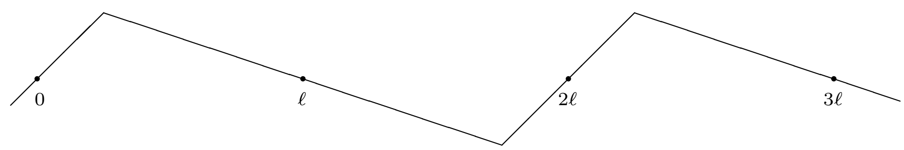
In other words, the initial displacement moves both ways along the string, with velocity c, and the displacement at time t is the average of the two travelling waves.

Let's see how this works in practice. Choose a satisfying $0 < a < 1$, and set

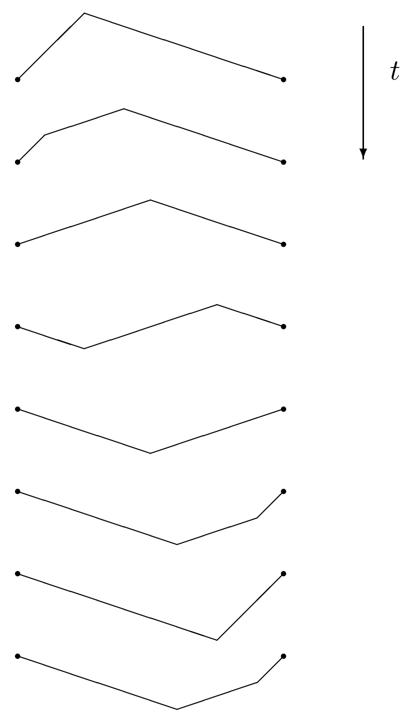
$$s_0(x) = \begin{cases} x/a & 0 \le x \le a \\ (\ell - x)/(\ell - a) & a \le x \le 1. \end{cases}$$



We use the reflection principle to extend this to a periodic function of period 2ℓ as described above.



Now we let this wave travel both left and right, and average the two resulting functions. Here is the resulting motion of the plucked string.



etc.

Exercises

1. (Effect of errors in initial conditions) Consider two sets of initial conditions for the wave equation (3.2.2), $s_0(x)$ and $v_0(x)$, $s_0'(x)$ and $v_0'(x)$, and let y and y' be the corresponding solutions. If we have bounds (not depending on x) on the distance between these initial conditions,

$$|s_0(x) - s_0'(x)| < \varepsilon_s, \qquad |v_0(x) - v_0'(x)| < \varepsilon_v,$$

show that the distance between y and y' satisfies

$$|y - y'| < \varepsilon_s + \frac{\ell \varepsilon_v}{2c}$$

(independently of x and t). This means, in particular, that the solution to the wave equation (3.2.2) depends *continuously* on the initial conditions.

Further reading:

J. Beament, *The violin explained: components, mechanism, and sound* [**8**].

R. Courant and D. Hilbert, *Methods of mathematical physics, I*, Interscience, 1953, §V.3.

L. Cremer, *The physics of the violin* [**27**].

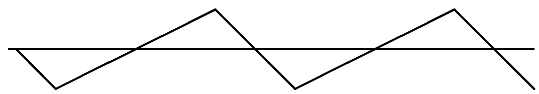
Neville H. Fletcher and Thomas D. Rossing, *The physics of musical instruments* [**40**]. Part III, String instruments.

T. D. Rossing, *The science of sound* [**121**], §10.

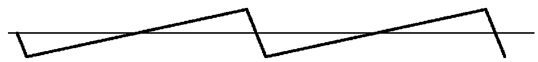
3.4. The bowed string

Ousainou Chaw on the *riti*, from Jacqueline Cogdell DjeDje, "Turn up the volume! A celebration of African music," UCLA 1999, p. 105.

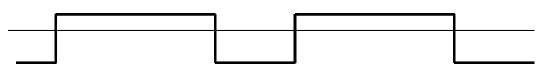
Helmholtz[5] carried out experiments on bowed violins, using a *vibration microscope* to produce *Lissajous* figures. He discovered that the motion of the string at every point describes a triangular pattern, but with slopes which depend on the point of observation. Near the bow, the displacement is as follows,



whereas nearer the bridge it looks as follows.



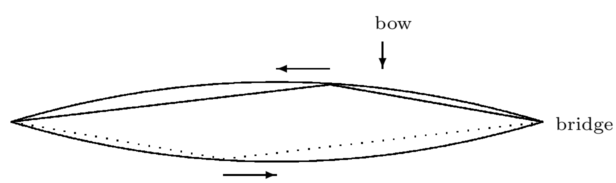
This means that the graph of velocity against time has the following form



where the area under the axis equals the area above, and the width of the trough decreases towards the bridge.

The interpretation of this motion is that the bowing action alternates between two distinct phases. In one phase, the bow sticks to the string and pulls it with it. In the other phase, the bow slides against the string. This form of motion reflects the fact that the coefficient of static friction is higher than the coefficient of dynamic friction.

The resulting motion of the entire string has the following form. The envelope of the motion is described by two parabolas, a lower one and an inverted upper one. Inside this envelope, at any point of time the string has two straight segments from the two ends to a point on the envelope. This point circulates around the envelope as follows.



To understand this behavior mathematically, we must solve the following problem. What are the solutions to the wave equation (3.2.2) satisfying not only the boundary conditions $y = 0$ for $x = 0$ and for $x = \ell$, for all values of t, but also the condition that the value of y as a function of t is prescribed for a particular value x_0 of x and for all t. Of course, the prescribed

[5]See section V.4 of [**54**].

motion at $x = x_0$ must have the right periodicity, because all solutions of the wave equation do:

$$y(x_0, t + 2\ell/c) = y(x_0, t).$$

It is tempting to try to solve this problem using d'Alembert's solution of the wave equation (Theorem 3.2.1). The problems we run into when we try to do this are interesting. For example, let's suppose that $x_0 = \ell/2$. Then we have

$$f(\ell/2 + ct) - f(-\ell/2 + ct) = y(\ell/2, t).$$

Replacing t by $t + \ell/c$ in this equation, we get

$$f(3\ell/2 + ct) - f(\ell/2 + ct) = y(\ell/2, t + \ell/c).$$

Adding, we get

$$f(3\ell/2 + ct) - f(-\ell/2 + ct) = y(\ell/2, t) + y(\ell/2, t + \ell/c).$$

But f is supposed to be periodic with period 2ℓ, so

$$f(3\ell/2 + ct) = f(-\ell/2 + ct).$$

This means that we have

$$y(\ell/2, t + \ell/c) = -y(\ell/2, t).$$

So not every periodic function with period $2\ell/c$ will work as the function $y(\ell/2, t)$. The function is forced to be half-period antisymmetric, so that only odd harmonics are present (see Section 2.3). This is only to be expected. After all, the even harmonics have a node at $x = \ell/2$, so how could we expect to involve even harmonics in the value of $y(x, t)$ at $x = \ell/2$?

Similar problems occur at $x = \ell/3$. The harmonics divisible by three are not allowed to occur in $y(\ell/3, t)$, because they have a node at $x = \ell/3$. This is a problem at every *rational* proportion of the string length.

It is becoming clear that Bernoulli's form (3.2.7) of the solution of the wave equation is going to be easier to use for this problem than d'Alembert's.

Since we are interested in functions $y(x_0, t)$ of the form shown in the diagrams at the beginning of this section, we may choose to measure time in such a way that $y(x_0, t)$ is an odd function of t, so that only sine waves and not cosine waves come into the Fourier series. So we set

$$y(x_0, t) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi ct}{\ell}\right).$$

Since the wave equation is linear, we can work with one frequency component at a time. So we set $y(x_0, t) = b_n \sin(n\pi ct/\ell)$. We look for solutions of the form

$$f(x) = c_n \cos\left(\frac{n\pi x}{\ell} + \phi_n\right)$$

and we want to determine c_n and ϕ_n in terms of b_n. We plug into d'Alembert's equation (3.2.5)

$$y(x_0, t) = f(x_0 + ct) - f(-x_0 + ct)$$

to get

$$b_n \sin\left(\frac{n\pi ct}{\ell}\right) = c_n \cos\left(\frac{n\pi(x_0+ct)}{\ell} + \phi_n\right) + c_n \cos\left(\frac{n\pi(-x_0+ct)}{\ell} + \phi_n\right).$$

Using equation (1.7.14), this becomes

$$b_n \sin\left(\frac{n\pi ct}{\ell}\right) = 2c_n \sin\left(\frac{n\pi x_0}{\ell}\right) \sin\left(\frac{n\pi ct}{\ell} + \phi_n\right).$$

Since this is supposed to be an identity between functions of t, we get $\phi_n = 0$ and

$$b_n = 2c_n \sin\left(\frac{n\pi x_0}{\ell}\right).$$

We now have a problem, very similar to the problem we ran into when we tried to use d'Alembert's solution. Namely, if $\sin(n\pi x_0/\ell)$ happens to be zero and $b_n \neq 0$, then there is no solution. So if x_0 is a rational multiple of ℓ then some frequency components are forced to be missing from $y(x_0, t)$. Apart from that, we have almost solved the problem. The value of c_n is

$$c_n = \frac{b_n}{2\sin(n\pi x_0/\ell)}$$

and so

$$f(x) = \sum_{n=1}^{\infty} \frac{b_n \sin(n\pi x/\ell)}{2\sin(n\pi x_0/\ell)}. \tag{3.4.1}$$

The solution of the wave equation is then given by plugging this into the formula (3.2.5). Using equation (1.7.12) we get

$$y = f(x+ct) - f(-x+ct) = \sum_{n=1}^{\infty} b_n \frac{\sin(n\pi x/\ell)\cos(n\pi ct/\ell)}{\sin(n\pi x_0/\ell)}.$$

The only thing that isn't clear so far is when the sum (3.4.1) converges. This is a point that we shall finesse by using Helmholtz's observation that in the case of a bowed string, for any chosen value of x_0 we have a triangular waveform

$$y(x_0, t) = \begin{cases} A\frac{t}{\alpha} & -\alpha \le t \le \alpha \\ A\dfrac{\ell - ct}{\ell - c\alpha} & \alpha \le t \le \frac{2\ell}{c} - \alpha \end{cases}$$

where α is some number depending on x_0, determining how long the leading edge of the triangular waveform lasts at the position x_0 along the string. The quantity A also depends on x_0, and represents the maximum amplitude of the vibration at that point. Using equation (2.2.9), we then calculate

$$\begin{aligned} b_n &= \frac{c}{\ell}\int_{-\alpha}^{\alpha} A\frac{t}{\alpha}\sin\left(\frac{n\pi ct}{\ell}\right)dt + \frac{c}{\ell}\int_{\alpha}^{\frac{2\ell}{c}-\alpha} A\frac{\ell - ct}{\ell - c\alpha}\sin\left(\frac{n\pi ct}{\ell}\right)dt \\ &= \frac{2A\ell^2}{n^2\pi^2 c\alpha(\ell - c\alpha)}\sin\left(\frac{n\pi c\alpha}{\ell}\right), \end{aligned}$$

so that

$$c_n = \frac{A\ell^2}{n^2\pi^2 c\alpha(\ell - c\alpha)} \frac{\sin(n\pi c\alpha/\ell)}{\sin(n\pi x_0/\ell)}.$$

Since the ratios of the c_n can't depend on the value of x_0 which we chose for our initial measurements, the only way this can work is if the two sine terms in this equation are equal, namely if

$$\frac{\pi c\alpha}{\ell} = \frac{\pi x_0}{\ell},$$

or

$$\alpha = x_0/c.$$

So if we measure the vibration at x_0 then the proportion $\alpha/(\ell/c)$ of the cycle spent in the trailing part of the triangular wave is equal to x_0/ℓ. In particular, if we measure at the bowing point, we obtain the following principle.

> *The proportion of the cycle for which the bow slips on the string is the same as the proportion of the string between the bow and the bridge.*

Now A is just some constant depending on x_0. Since c_n doesn't depend on x_0, the constant $A/c\alpha(\ell - c\alpha) = A/x_0(\ell - x_0)$ must be independent of x_0. If we write K for this quantity, we obtain a formula for amplitude in terms of position along the string,

$$A = Kx_0(\ell - x_0).$$

This formula explains the parabolic amplitude envelope for the vibration of the bowed string.

Further reading:

L. Cremer, *The physics of the violin* [**27**].

Joseph B. Keller, *Bowing of violin strings*, Comm. Pure and Appl. Math. 6 (1953), 483–495.

B. Lawergren, *On the motion of bowed violin strings*, Acustica 44 (1980), 194–206.

C. V. Raman, *On the mechanical theory of the vibrations of bowed strings and of musical instruments of the violin family, with experimental verification of the results: Part I*, Indian Assoc. Cultivation Sci. Bull. 15 (1918), 1–158.

J. C. Schelleng, *The bowed string and the player*, J. Acoust. Soc. Amer. 53 (1973), 26–41.

J. C. Schelleng, *The physics of the bowed string*, Scientific American 235 (1) (1974), 87–95. Reproduced in Hutchins, *The Physics of Music*, W. H. Freeman and Co, 1978.

Lily M. Wang and Courtney B. Burroughs, *Acoustic radiation from bowed violins*, J. Acoust. Soc. Amer. 110 (1) (2001), 543–555.

3.5. Wind instruments

To understand the vibration of air in a tube or pipe, we introduce two variables, displacement and acoustic pressure. Both of these will end up satisfying the wave equation, but with different phases.

We consider the air in the tube to have a rest position, and the wave motion is expressed in terms of displacement from that position. So let x denote position along the tube, and let $\xi(x,t)$ denote the displacement of the air at position x at time t. The pressure also has a rest value, namely the ambient air pressure ρ. We measure the *acoustic pressure* $p(x,t)$ by subtracting ρ from the absolute pressure $P(x,t)$, so that

$$p(x,t) = P(x,t) - \rho.$$

Hooke's law in this situation states that

$$p = -B\frac{\partial \xi}{\partial x}$$

where B is the *bulk modulus* of air. Newton's second law of motion implies that

$$\frac{\partial p}{\partial x} = -\rho\frac{\partial^2 \xi}{\partial t^2}.$$

Combining these equations, we obtain the equations

$$\frac{\partial^2 \xi}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 \xi}{\partial t^2} \tag{3.5.1}$$

and

$$\frac{\partial^2 p}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}. \tag{3.5.2}$$

where $c = \sqrt{B/\rho}$. These equations are the wave equation for displacement and acoustic pressure respectively.

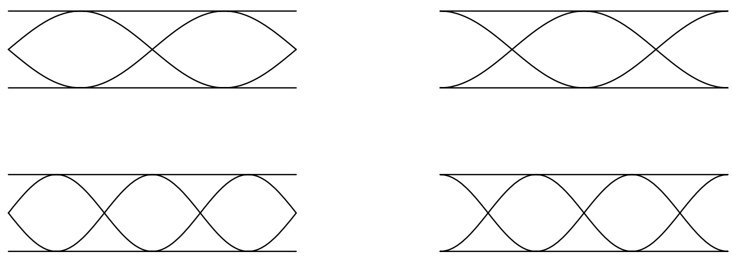
The boundary conditions depend upon whether the end of the tube is open or closed. For a closed end of a tube, the displacement ξ is forced to be zero for all values of t. For an open end of a tube, the acoustic pressure p is zero for all values of t.

Bone flute from Henan province, China, 6000 B.C. Picture from *Music in the age of Confucius*, p. 90. The oldest known (probable) flute is 45,000 years old, from the Neanderthal era, and was discovered in Slovenia in 1995. See the article of Kunej and Turk listed at the end of the section.

So for a tube open at both ends, such as the flute, the behavior of the acoustic pressure p is determined by exactly the same boundary conditions as in the case of a vibrating string. It follows that d'Alembert's solution given in Section 3.2 works in this case, and we again get integer multiples of a fundamental frequency. The basic mode of vibration is a sine wave, represented by the following diagram. The displacement is also a sine wave, but with a different phase.

pressure displacement

Bear in mind that the vertical axis in this diagram actually represents *horizontal* displacement or pressure, and not vertical, because of the longitudinal nature of air waves. Furthermore, the two parts of the graphs only represent the two extremes of the motion. In these diagrams, the nodes of the pressure diagram correspond to the antinodes of the displacement diagram and vice versa. The second and third vibrational modes will be represented by the following diagrams.

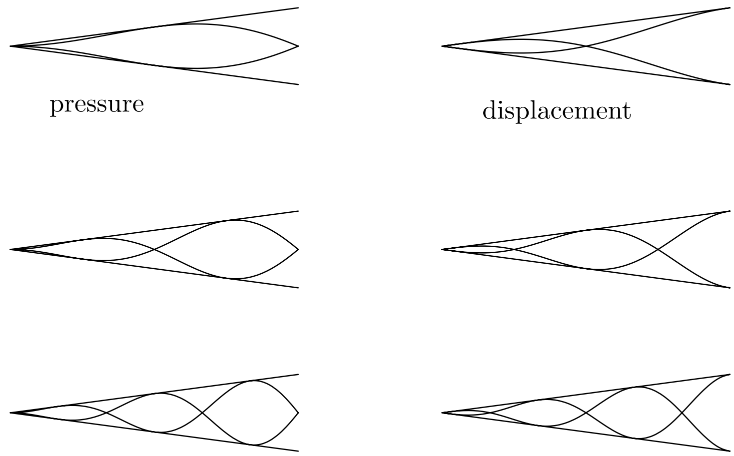


Tubes or pipes which are closed at one end behave differently, because the displacement is forced to be zero at the closed end. So the first two modes are as follows. In these diagrams, the left end of the tube is closed.

pressure displacement

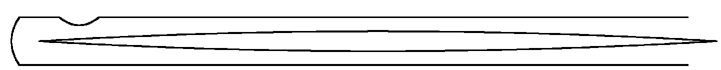
It follows that for closed tubes, odd multiples of the fundamental frequency dominate. For example, as mentioned above, the flute is an open tube, so all multiples of the fundamental are present. The clarinet is a closed tube, so odd multiples predominate.

Conical tubes are equivalent to open tubes of the same length, as illustrated by the following diagrams. These diagrams are obtained from the ones for the open tube, by squashing down one end.



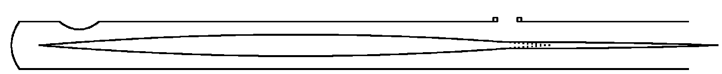
The oboe has a conical bore so again all multiples are present. This explains why the flute and oboe overblow at the octave, while the clarinet overblows at an octave plus a perfect fifth, which represents tripling the frequency. The odd multiples of the fundamental frequency dominate for a clarinet, although in practice there are small amplitudes present for the even ones from four times the fundamental upwards as well.

At this point, it should be mentioned that for an open end, $p = 0$ is really only an approximation, because the volume of air just outside of the tube is not infinite. A good way to adjust to make a more accurate representation of an actual tube is to work in terms of an *effective* length, and consider the tube to end a little beyond where it really does. The following diagram shows the effective length for the fundamental vibrational mode of a flute, with all holes closed.



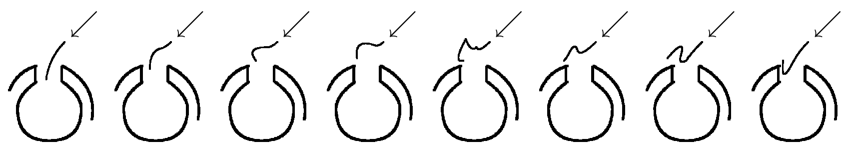
The *end correction* is the amount by which the effective length exceeds the actual length, and under normal conditions it is usually somewhere around three fifths of the width of the tube.

The effect of an open hole is to decrease the effective length of the tube. Here is a diagram of the first vibrational mode with one open hole.



The effective length of the tube can be seen by continuing the left part of the wave as though the hole doesn't exist, and seeing where the wave ends. This is represented by the dotted lines in the diagram. The larger the hole, the greater the effect on the effective length.

So what happens when the flutist blows into the mouthpiece of the flute? How does this cause a note to sound? The following pictures, adapted from stroboscopic experiments of Coltman using smoke particles, show how the airstream varies with time. The arrow represents the incoming air stream.



This air pattern results in a series of vortices being sent down the tube. When the vortices get to the end of the tube, they are reflected back up. They reach the beginning of the tube and are reflected again. Some of these will be out of phase with the new vortices being generated, and some will be in phase. The ones that are in phase reinforce, and feed back to build up a coherent tone. This in turn makes it more favorable for vortices to be formed in synchronization with the tone.

Further reading:

John W. Coltman, *Acoustics of the flute*, Physics Today 21 (11) (1968), 25–32. Reprinted in Rossing [**120**].

Neville H. Fletcher and Thomas D. Rossing, *The physics of musical instruments* [**40**]. Part IV, Wind instruments.

Ian Johnston, *Measured tones* [**64**], pages 207–233.

Drago Kunej and Ivan Turk, *New perspectives on the beginnings of music: Archeological and musicological analysis of a middle Paleolithic bone "flute,"* The origins of music, ed. Wallin, Merker and Brown, MIT Press (2000), 235–268.

C. J. Nederveen, *Acoustical aspects of woodwind instruments* [**93**].

T. D. Rossing, *The science of sound* [**121**], §12.

3.6. The drum

The Timpani (Gerard Hoffnung)

Consider a circular drum whose skin has area density (mass per unit area) ρ. If the boundary is under uniform tension T, this ensures that the entire surface is under the same uniform tension. The tension is measured in force per unit distance (newtons per meter).

To understand the wave equation in two dimensions, for a membrane such as the surface of a drum, the argument is analogous to the one dimensional case. We parametrize the surface with two variables x and y, and we use z to denote the displacement perpendicular to the surface. Consider a rectangular element of surface of width Δx and length Δy. Then the tension on the left and right sides is $T\Delta y$, and the argument which gave equation (3.2.1) in the one dimensional case shows in this case that the difference in vertical components is approximately

$$(T\Delta y)\left(\Delta x \frac{\partial^2 z}{\partial x^2}\right).$$

Similarly, the difference in vertical components between the front and back of the rectangular element is approximately

$$(T\Delta x)\left(\Delta y \frac{\partial^2 z}{\partial y^2}\right).$$

So the total upward force on the element of surface is approximately

$$T\Delta x\Delta y\left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}\right).$$

The mass of the element of surface is approximately $\rho\Delta x\Delta y$, so Newton's second law of motion gives

$$T\Delta x\Delta y\left(\frac{\partial^2 z}{\partial x^2}+\frac{\partial^2 z}{\partial y^2}\right)\approx(\rho\Delta x\Delta y)\frac{\partial^2 z}{\partial t^2}.$$

Dividing by $\Delta x\Delta y$, we obtain the wave equation in two dimensions, namely the partial differential equation

$$\rho\frac{\partial^2 z}{\partial t^2}=T\left(\frac{\partial^2 z}{\partial x^2}+\frac{\partial^2 z}{\partial y^2}\right).$$

As in the one dimensional case, we set $c=\sqrt{T/\rho}$, which will play the role of the speed of the waves on the membrane. So the wave equation becomes

$$\frac{\partial^2 z}{\partial t^2}=c^2\left(\frac{\partial^2 z}{\partial x^2}+\frac{\partial^2 z}{\partial y^2}\right).$$

Converting to polar coordinates (r,θ) using equation (P.4), we obtain

$$\frac{\partial^2 z}{\partial t^2}=c^2\left(\frac{\partial^2 z}{\partial r^2}+\frac{1}{r}\frac{\partial z}{\partial r}+\frac{1}{r^2}\frac{\partial^2 z}{\partial \theta^2}\right). \tag{3.6.1}$$

We look for *separable* solutions of this equation, namely solutions of the form

$$z=f(r)g(\theta)h(t).$$

The reason for looking for separable solutions will be explained further in the next section. Substituting this into the wave equation, we obtain

$$f(r)g(\theta)h''(t)=c^2\left(f''(r)g(\theta)h(t)+\frac{1}{r}f'(r)g(\theta)h(t)+\frac{1}{r^2}f(r)g''(\theta)h(t)\right).$$

Dividing by $f(r)g(\theta)h''(t)$ gives

$$\frac{h''(t)}{h(t)}=c^2\left(\frac{f''(r)}{f(r)}+\frac{1}{r}\frac{f'(r)}{f(r)}+\frac{1}{r^2}\frac{g''(\theta)}{g(\theta)}\right).$$

In this equation, the left hand side only depends on t, and is independent of r and θ, while the right hand side only depends on r and θ, and is independent of t. Since t, r and θ are three independent variables, this implies that the common value of the two sides is independent of t, r and θ, so that it has to be a constant. We shall see in the next section that this constant has to be a negative real number, so we shall write it as $-\omega^2$. So we obtain two equations,

$$h''(t)=-\omega^2 h(t), \tag{3.6.2}$$

$$\frac{f''(r)}{f(r)}+\frac{1}{r}\frac{f'(r)}{f(r)}+\frac{1}{r^2}\frac{g''(\theta)}{g(\theta)}=-\frac{\omega^2}{c^2}. \tag{3.6.3}$$

The general solution to equation (3.6.2) is a multiple of the solution

$$h(t)=\sin(\omega t+\phi),$$

where ϕ is a constant determined by the initial temporal phase. Multiplying equation (3.6.3) by r^2 and rearranging, we obtain

$$r^2\frac{f''(r)}{f(r)} + r\frac{f'(r)}{f(r)} + \frac{\omega^2}{c^2}r^2 = -\frac{g''(\theta)}{g(\theta)}.$$

The left hand side depends only on r, while the right hand side depends only on θ, so their common value is again a constant. This makes $g(\theta)$ either a sine function or an exponential function, depending on the sign of the constant. But the function $g(\theta)$ has to be periodic of period 2π since it is a function of angle. So the common value of the constant must be the square of an integer n, so that

$$g''(\theta) = -n^2 g(\theta)$$

and $g(\theta)$ is a multiple of $\sin(n\theta + \psi)$. Here, ψ is another constant representing spatial phase. So we obtain

$$r^2\frac{f''(r)}{f(r)} + r\frac{f'(r)}{f(r)} + \frac{\omega^2}{c^2}r^2 = n^2.$$

Multiplying by $f(r)$, dividing by r^2 and rearranging, this becomes

$$f''(r) + \frac{1}{r}f'(r) + \left(\frac{\omega^2}{c^2} - \frac{n^2}{r^2}\right) = 0.$$

Now Exercise 2 in §2.10 shows that the general solution to this equation is a linear combination of $J_n(\omega r/c)$ and $Y_n(\omega r/c)$. But the function $Y_n(\omega r/c)$ tends to $-\infty$ as r tends to zero, so this would introduce a singularity at the center of the membrane. So the only physically relevant solutions to the above equation are multiples of $J_n(\omega r/c)$. So we have shown that the functions

$$\boxed{z = AJ_n(\omega r/c)\sin(\omega t + \phi)\sin(n\theta + \psi)}$$

are solutions to the wave equation.

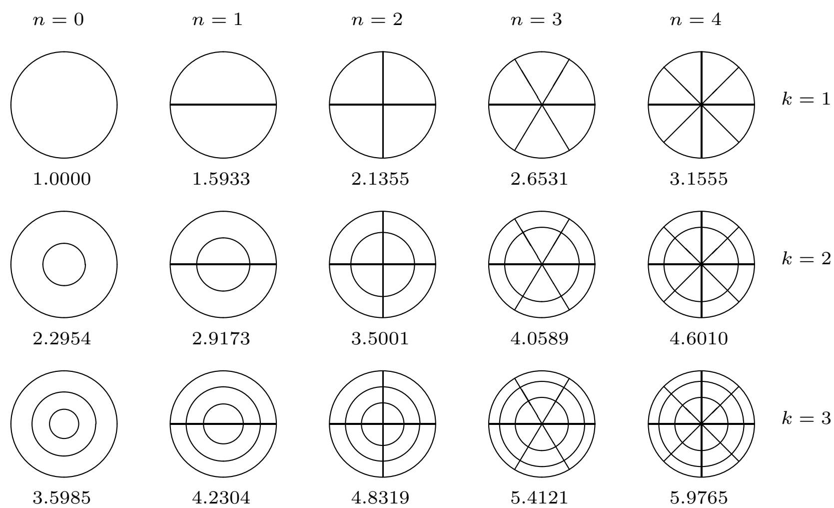
If the radius of the drum is a, then the boundary condition which we must satisfy is that $z = 0$ when $r = a$, for all values of t and θ. So it follows that $J_n(\omega a/c) = 0$. This is a constraint on the value of ω. The function J_n takes the value zero for a discrete infinite set of values of its argument. So ω is also constrained to an infinite discrete set of values.

It turns out that linear combinations of functions of the above form uniformly approximate the general, twice continuously differentiable solution of (3.6.1) as closely as desired, so that these form the drum equivalent of the sine and cosine functions of Fourier series.

Here is a table of the first few zeros of the Bessel functions. For more, see Appendix B.

k	J_0	J_1	J_2	J_3	J_4
1	2.40483	3.83171	5.13562	6.38016	7.58834
2	5.52008	7.01559	8.41724	9.76102	11.06471
3	8.65373	10.17347	11.61984	13.01520	14.37254

We have seen that to choose a vibrational mode, we must choose a nonnegative integer n and we must choose a zero of $J_n(z)$. Denoting the kth zero of J_n by $j_{n,k}$, the corresponding vibrational mode has frequency $(cj_{n,k}/a)$, which is $j_{n,k}/j_{0,1}$ times the fundamental frequency. The stationary points have the following pictures. Underneath each picture, we have recorded the value of $j_{n,k}/j_{0,1}$ for the relative frequency.

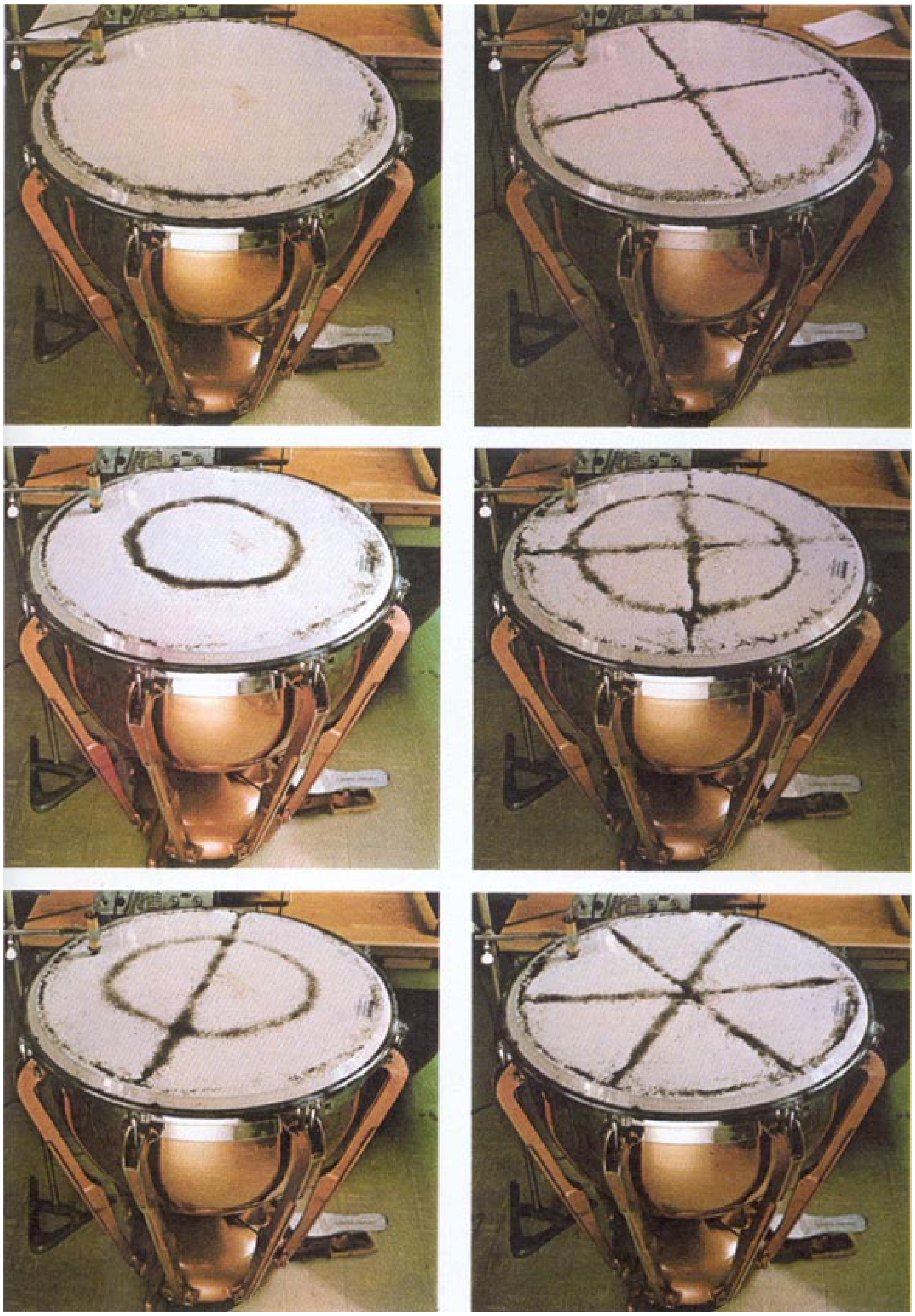


In the late eighteenth century, Chladni[6] discovered a way to see normal modes of vibration. He was interested in the vibration of plates, but the same technique can be used for drums and other instruments. He placed sand on the plate and then set it vibrating in one of its normal modes, using a violin bow. The sand collects on the stationary lines and gives a picture similar to the ones described above for the drum. A picture of Chladni patterns on a kettledrum can be found on page 105.

In practice, for a drum in which the air is confined (such as a kettledrum) the fundamental mode of the drum is heavily damped, because it involves compression and expansion of the air enclosed in the drum. So what is heard as the fundamental is really the mode with $n = 1$, $k = 1$, namely the second entry in the top row in the above diagram. The higher modes mostly involve moving the air from side to side. The inertia of the air has the effect of raising the frequency of the modes with $n = 0$, especially the fundamental, while the modes with $n > 0$ are lowered in frequency in such a way as to widen the frequency gaps. For an open drum, on the other hand, all the vibrational frequencies are lowered by the inertia of the air, but the ones of lower frequency are lowered the most.

The design of the orchestral kettledrum carefully utilises the inertia of the air to arrange for the modes with $n = 1$, $k = 1$ and $n = 2$, $k = 1$ to have

[6]E. F. F. Chladni, *Entdeckungen über die Theorie des Klanges*, 1787.



Chladni patterns on a kettledrum
from Risset, *Les instruments de l'orchestre*

frequency ratio approximating 3:2, so that what is perceived is a missing fundamental at half the actual fundamental frequency. Furthermore, the modes with $n = 3$, 4 and 5 (still with $k = 1$) are arranged to approximate frequency ratios of 4:2, 5:2 and 6:2 with the $n = 1$, $k = 1$ mode, thus accentuating the perception of the missing fundamental. The frequency of the $n = 1$, $k = 1$ mode is called the *nominal frequency* of the drum.

It is not true that the air in the kettle of a kettledrum acts as a resonator. A kettledrum can be retuned by a little more than a perfect fourth, whereas if the air were acting as a resonator, it could only do so for a small part of the frequency range. In fact, the resonances of the body of air are usually much higher in pitch, and do not have much effect on the overall sound. A more important effect is that the underside of the drum skin is prevented from radiating sound, and this makes the radiation of sound from the upper side more efficient.

Exercises

1. The women of Portugal (never the men) play a double sided square drum called an *adufe*. Find the separable solutions (i.e., the ones of the form $z = f(x)g(y)h(t)$) to the wave equation for a square drum. Write the answer in the form of an essay, with title: "What does a square drum sound like?". Try to integrate the words with the mathematics. Explain what you're doing at each step, and don't forget to answer the title question (i.e., describe the frequency spectrum).

Further reading:

Murray Campbell and Clive Greated, *The musician's guide to acoustics* [**16**], chapter 10.

R. Courant and D. Hilbert, *Methods of mathematical physics, I*, Interscience, 1953, §V.5.

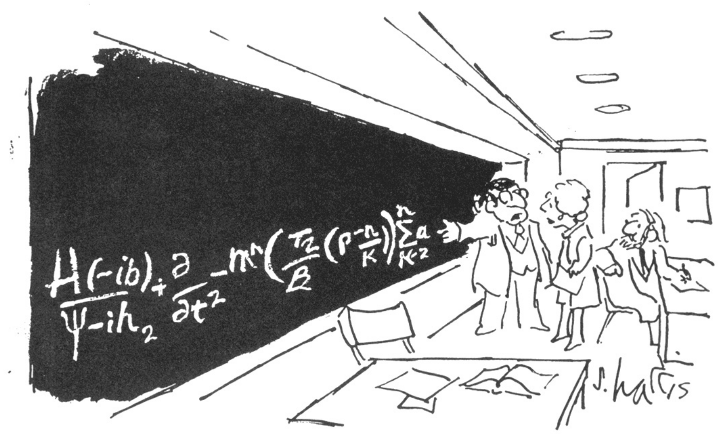
William C. Elmore and Mark A. Heald, *Physics of waves* [**37**], chapter 2.

Neville H. Fletcher and Thomas D. Rossing, *The physics of musical instruments* [**40**], §18.

C. V. Raman, *The Indian musical drums*, Proc. Indian Acad. Sci. A1 (1934), 179–188. Reprinted in Rossing [**120**].

Thomas D. Rossing, *Science of percussion instruments* [**122**].

3.7. Eigenvalues of the Laplace operator



"But this *is* the simplified version for the general public."

In this section, we put the discussion of the vibrational modes of the drum into a broader context. Namely, we explain the relationship between the shape of a drum and its frequency spectrum, in terms of the eigenvalues of the Laplace operator. This discussion explains the connection between the uses of the word "spectrum" in linear algebra, where it refers to the eigenvalues of an operator, and in music, where it refers to the distribution of frequency components. Parts of this discussion assume that the reader is familiar with elementary vector calculus and the divergence theorem.

We write ∇^2 for the operator $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$. This is known as the *Laplace operator* (in three dimensions the Laplace operator ∇^2 denotes $\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}$; the analogous operator makes sense for any number of variables). In this notation, the wave equation becomes

$$\frac{\partial^2 z}{\partial t^2} = c^2 \nabla^2 z.$$

We consider the solutions to this equation on a closed and bounded region Ω. So for the drum of the last section, Ω was a disc in two dimensions.

A *separable solution* to the wave equation is one of the form

$$z = f(x, y)h(t).$$

Substituting into the wave equation, we obtain

$$f(x, y)h''(t) = c^2 \nabla^2 f(x, y)\, h(t)$$

or

$$\frac{h''(t)}{h(t)} = c^2 \frac{\nabla^2 f(x, y)}{f(x, y)}.$$

The left hand side is independent of x and y, while the right hand side is independent of t, so their common value is a constant. We write this constant as $-\omega^2$, because it will transpire that it has to be negative. Then we have

$$g''(t) = -\omega^2 g(t), \tag{3.7.1}$$

$$\nabla^2 f(x,y) = -\frac{\omega^2}{c^2} f(x,y). \tag{3.7.2}$$

The first of these equations is just the equation for simple harmonic motion with angular frequency ω, so the general solution is

$$g(t) = A\sin(\omega t + \phi).$$

A nonzero, twice differentiable function $f(x,y)$ satisfying the second equation is called an *eigenfunction* of the Laplace operator ∇^2 (or more accurately, of $-\nabla^2$), with *eigenvalue*

$$\lambda = \omega^2/c^2. \tag{3.7.3}$$

There are two important kinds of eigenfunctions and eigenvalues. The *Dirichlet spectrum* is the set of eigenvalues for eigenfunctions which vanish on the boundary of the region Ω. The *Neumann spectrum* is the set of eigenvalues for eigenfunctions with vanishing derivative normal (i.e., perpendicular) to the boundary. The latter functions are important when studying the wave equation for sound waves, where the dependent variable is acoustic pressure (i.e., pressure minus the average ambient pressure).

For the benefit of the reader who knows vector calculus, in Appendix W we give a treatment of the solution of the wave equation, and justify the method of separation of variables. There, you can find the proof that the eigenvalues of $-\nabla^2$ (i.e., the values of λ for which $\nabla^2 z = -\lambda z$ has a nonzero solution) are positive and real, along with many other standard facts about the wave equation, which we now summarize.

We can choose Dirichlet eigenfunctions $f_1, f_2, \dots$ of $-\nabla^2$ on Ω with eigenvalues $0 < \lambda_1 \le \lambda_2 \le \dots$ with the following properties.

(i) Every eigenfunction is a finite linear combination of eigenfunctions f_i for which the λ_i are equal.

(ii) Each eigenvalue is repeated only a finite number of times.

(iii) $\lim_{n\to\infty} \lambda_n = \infty$.

(iv) (Completeness) Every continuous function can be written as a sum of an absolutely and uniformly convergent series of the form $f(x,y) = \sum_i a_i f_i(x,y)$.

The eigenvalue λ_i determines the frequency of the corresponding vibration via (3.7.3):

$$\omega_i = c\sqrt{\lambda_i}, \qquad \nu_i = c\sqrt{\lambda_i}/2\pi. \tag{3.7.4}$$

(recall that angular velocity ω is related to frequency ν by $\omega = 2\pi\nu$).

Initial conditions for the wave equation on Ω are specified by stipulating the values of z and $\frac{\partial z}{\partial t}$ for (x,y) in Ω, at $t = 0$. To solve the wave equation subject to these initial conditions, we use completeness to write $z = \sum_i a_i f_i(x,y)$

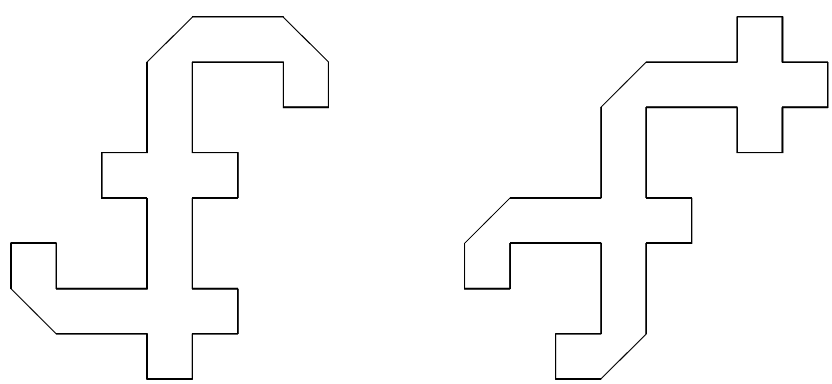
and $\frac{\partial z}{\partial t} = \sum_i b_i f_i(x,y)$ at $t = 0$. Then the unique solution is given by

$$z = \sum_\lambda f_i(x,y) \left(a_i \cos(c\sqrt{\lambda}\, t) + \frac{b_i}{c\sqrt{\lambda}} \sin(c\sqrt{\lambda}\, t) \right).$$

For further details, see Appendix W, and in particular, equation (W.37).

We have phrased the above discussion in terms of the two dimensional wave equation, but the same arguments work in any number of dimensions. For example, in one dimension it corresponds to the vibrational modes of a string, and we recover the theory of Fourier series.

An interesting problem, which was posed by Mark Kac in 1965 and solved by Gordon, Webb and Wolpert in 1991, is *whether one can hear the shape of a drum.* In other words, can one tell the shape of a simply connected closed region in two dimensions from its Dirichlet spectrum? Simply connected just means there are no holes in the region. Based on a method developed by Sunada a few years previously, Gordon, Webb and Wolpert found examples of pairs of regions with the same Dirichlet spectrum. The example which appears in their paper is the following.



Admittedly, it had probably not occurred to anyone to make drums using vibrating surfaces of these shapes, prior to this investigation. Many other pairs of regions with the same Dirichlet spectrum have been found. An example is worked out in detail at the end of Appendix W; this and many more can be found in the paper of Buser, Conway, Doyle and Semmler listed below. But it is still not known whether there are any *convex* examples.

Further reading:

P. Buser, J. H. Conway, P. Doyle and K.-D. Semmler, *Some planar isospectral domains*, International Mathematics Research Notices (1994), 391–400.

S. J. Chapman, *Drums that sound the same*, Amer. Math. Monthly 102 (2) (1995), 124–138.

Tobin Driscoll, *Eigenmodes of isospectral drums.* SIAM Rev. 39 (1997), 1-17.

Carolyn Gordon, David L. Webb, and Scott Wolpert, *One cannot hear the shape of a drum*, Bulletin of the Amer. Math. Soc. 27 (1992), 134–138.

Carolyn Gordon, David L. Webb, and Scott Wolpert, *Isospectral plane domains and surfaces via Riemannian orbifolds*, Invent. Math. 110 (1992), 1–22.

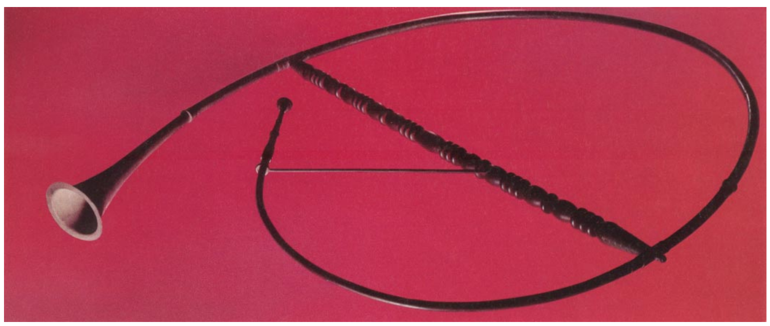
Mark Kac, *Can one hear the shape of a drum?*, Amer. Math. Monthly 73, (1966), 1–23.

M. H. Protter, *Can one hear the shape of a drum? Revisited.* SIAM Rev. 29 (1987), 185–197.

K. Stewartson and R. T. Waechter, *On hearing the shape of a drum: further results*, Proc. Camb. Phil. Soc. 69 (1971), 353–363.

T. Sunada, *Riemannian coverings and isospectral manifolds*, Ann. of Math. 121 (1985), 169–186.

3.8. The horn



Tuba curva, Pompeii, first century A.D.
Musée Instrumental, Brussels, Belguim

The horn, and other instruments of the brass family, can be regarded as a hard walled tube of varying cross-section. Fortunately, the cross-section matters more than the exact shape and curvature of the tube.

If $A(x)$ represents the cross-section as a function of position x along the tube, then assuming that the wavefronts are approximately planar and propagate along the direction of the horn, equation (3.5.2) can be modified to *Webster's horn equation*

$$\frac{1}{A(x)}\frac{\partial}{\partial x}\left(A(x)\frac{\partial p}{\partial x}\right) = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}, \tag{3.8.1}$$

or equivalently

$$\frac{\partial^2 p}{\partial x^2} + \frac{1}{A}\frac{dA}{dx}\frac{\partial p}{\partial x} = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}.$$

Solutions of this equation can be described using the theory of *Sturm–Liouville equations*. The theory of Sturm–Liouville equations is described in many standard texts on partial differential equations, and is a direct generalization of our discussion of the wave equation in Section 3.6 and Appendix W.

There is one particular form of $A(x)$ which is of physical importance because it gives a good approximation to the shape of actual brass instruments while at the same time giving an equation with relatively simple solutions. Namely, the *Bessel horn*, with cross section of radius and area

$$R(x) = bx^{-\alpha}, \qquad A(x) = \pi R(x)^2 = Bx^{-2\alpha}.$$

Here, the origin of the x coordinate and the constant b are chosen to give the correct radius at the two ends of the horn, and $B = \pi b^2$. Notice that the constant B disappears when $A(x)$ is put into equation (3.8.1). The parameter α is the "flare parameter" that determines the shape of the flare of the horn. The case $\alpha = 0$ gives a conical tube, and we shall usually assume that $\alpha \geq 0$. The solutions are sums of ones of the form

$$p(x,t) = x^{\alpha+\frac{1}{2}} J_{\alpha+\frac{1}{2}}(\omega x/c)(a\cos\omega t + b\sin\omega t). \tag{3.8.2}$$

Here, as usual, the angular frequency ω must be chosen so that the boundary conditions are satisfied at the ends of the horn.

Exercises

1. Verify that (3.8.2) is a solution of equation (3.8.1) with the given value of $A(x)$. You will need to use Bessel's differential equation (2.10.1) with n replaced by $\alpha + \frac{1}{2}$ and z replaced by $\omega x/c$.

Further reading:

E. Eisner, *Complete solutions of the "Webster" horn equation*, J. Acoust. Soc. Am. **41** (1967), 1126–1146.

Neville H. Fletcher and Thomas D. Rossing, *The physics of musical instruments* [**40**], §8.6.

Thomas D. Rossing, *The science of sound* [**121**], §11.

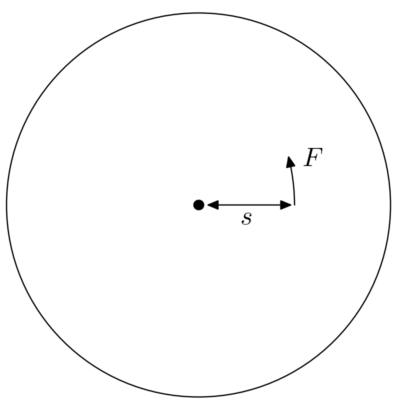
A. G. Webster, *Acoustical impedance, and the theory of horns and of the phonograph*, Proc. Nat. Acad. Sci. (US) 5 (1919), 275–282.

3.9. Xylophones and tubular bells

Xylophone, made by Yayi Coulibaly (1947), from Jacqueline Cogdell DjeDje, "Turn up the volume! A celebration of African music," UCLA 1999, p. 253.

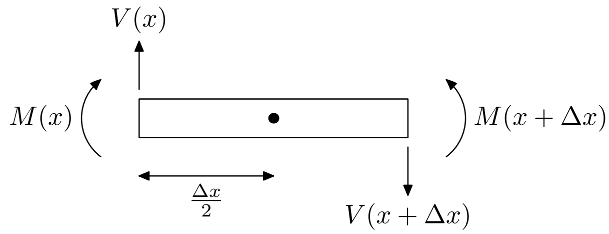
In this section we examine the theory of transverse waves in a slender stiff rod. This theory applies to instruments such as the xylophone and the tubular bells. We shall see that in this case, just as in the case of the drum, the vibrational modes do not consist of integer multiples of a fundamental frequency. Our goal will be to derive and solve the differential equation (3.9.2).

As well as the assumptions made in §3.2 about small angles, the basic assumption we shall make in order to obtain the appropriate differential equation is that terms coming from the resistance to motion caused by the rotational inertia of a segment of the rod are very small compared with terms coming from (vertical) linear inertia. This is only realistic for a slender rod. The upshot of this assumption is that the total *torque* on a segment of rod can be taken to be zero. Recall that if we try to twist an object about an axis, by applying a force F at distance s from the axis, then the torque applied is defined to be Fs. This is reasonable because the effect of such a turning force is proportional to the distance from the axis, as well as to the magnitude of the force.



Torque $= Fs$

Consider a segment of rod of length Δx, and let $V(x)$ be the vertical force (or *shearing force*) applied by the left end of the segment on the right end of the adjacent segment.



The torque on the segment due to this shearing force is

$$-V(x)\left(\frac{\Delta x}{2}\right) - V(x+\Delta x)\left(\frac{\Delta x}{2}\right) \approx -V(x)\Delta x$$

(the minus sign is because we regard counterclockwise as the positive direction for torque). Since we are regarding rotational inertia as negligible, this means that the torque, or *bending moment*, $M(x)$ applied by the segment on the adjacent segment satisfies

$$M(x+\Delta x) - M(x) - V(x)\Delta x \approx 0,$$

or

$$V(x) \approx \frac{M(x+\Delta x) - M(x)}{\Delta x}.$$

Taking limits as $\Delta x \to 0$, we obtain

$$V(x) = \frac{dM(x)}{dx}.$$

The upward force on the segment can now be calculated as

$$V(x) - V(x+\Delta x) \approx -\Delta x \frac{dV(x)}{dx} \approx -\Delta x \frac{d^2M(x)}{dx^2}.$$

Now the functions $V(x)$, $M(x)$, etc. are really functions of both x and t; we have suppressed the dependence on t in the above discussion. So we really need to write the total upwards force on the segment as

$$-\Delta x \frac{\partial^2 M(x,t)}{\partial x^2}.$$

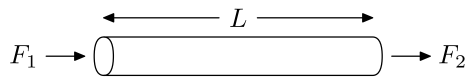
If the linear density of the rod is ρ (measured in kg/m) then the mass of the segment is $\rho\Delta x$. Writing y for the vertical displacement, Newton's second law of motion gives

$$-\Delta x \frac{\partial^2 M}{\partial x^2} = \rho\Delta x \frac{\partial^2 y}{\partial t^2},$$

or

$$\frac{\partial^2 y}{\partial t^2} + \frac{1}{\rho}\frac{\partial^2 M}{\partial x^2} = 0. \tag{3.9.1}$$

Now the bending moment M causes the rod to bend, and so there is a close relationship between M and $\partial^2 y/\partial x^2$. To understand this relationship, we must begin by introducing the concepts of stress, strain and Young's modulus. If a force $F = F_2 - F_1$ stretches or compresses a stiff slender rod of length L and cross-sectional area A,



then the length will increase by an amount ΔL. The *tension stress* (or just the *tension*) is defined to be

$$f = F/A.$$

The *tension strain* (or *extension*) is defined to be the proportional increase in length,

$$\epsilon = \Delta L/L.$$

Hooke's law for a stiff rod states that the extension is proportional to the tension,

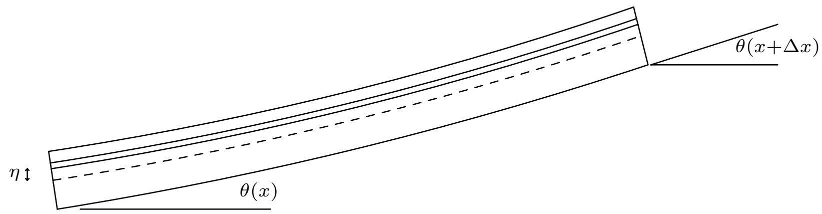
$$f = E\epsilon.$$

The constant of proportionality E is called the *Young's modulus*[7] (or *longitudinal elasticity*). Values for the Young's modulus for various materials at room temperature (18°C) are given in the following table.

Material	Young's modulus (N/m^2)
Aluminum	7.05×10^{10}
Brass	9.7–10.4×10^{10}
Copper	12.98×10^{10}
Gold	7.8×10^{10}
Iron	21.2×10^{10}
Lead	1.62×10^{10}
Silver	8.27×10^{10}
Steel	21.0×10^{10}
Zinc	9.0×10^{10}
Glass	5.1–7.1×10^{10}
Rosewood	1.2–1.6×10^{10}

Now we are ready to examine the segment of rod in more detail as it bends. There is a *neutral surface* in the middle of the rod, which is neither compressed nor stretched. It is represented by the dotted line in the diagram below. One side of this surface the horizontal filaments of rod are compressed, the other side they are stretched. Denote by η the distance from the neutral surface to the filament.

[7] Named after the British physicist and physician Thomas Young (1773–1829).



Write R for the radius of curvature of the neutral surface, so that the length of the segment at the neutral surface is $R\Delta\theta$. The length of the filament is $(R-\eta)\Delta\theta$, so the tension strain is $-(\eta\Delta\theta)/(R\Delta\theta) = -\eta/R$. So by Hooke's law, the tension stress on the filament is $-E\eta\Delta A/R$, where ΔA is the cross-sectional area of the filament.

Since the total horizontal force is supposed to be zero, we have

$$-\frac{E}{R}\int \eta\, dA = 0$$

so that $\int \eta\, dA = 0$. This says that the neutral surface passes through the *centroid* of the cross-sectional area. The total bending moment is obtained by multiplying by $-\eta$ and integrating:[8]

$$M = \frac{E}{R}\int \eta^2\, dA.$$

The quantity $I = \int \eta^2\, dA$ is called the *sectional moment* of the cross-section of the rod. So we obtain $M = -EI/R$. Now the formula for radius of curvature is $R = \left(1 + (\frac{dy}{dx})^2\right)^{\frac{3}{2}} / \frac{d^2y}{dx^2}$. Assuming that $\frac{dy}{dx}$ is small, this can be approximated by the formula $1/R = \frac{d^2y}{dx^2}$, so that

$$M(x,t) = EI\frac{\partial^2 y}{\partial x^2}.$$

Combining this with equation (3.9.1) gives

$$\boxed{\frac{\partial^2 y}{\partial t^2} + \frac{EI}{\rho}\frac{\partial^4 y}{\partial x^4} = 0.} \tag{3.9.2}$$

This is the differential equation which governs the transverse waves on the rod. It is known as the Euler–Bernoulli beam equation.

We look for separable solutions to equation (3.9.2). Setting

$$y = f(x)g(t)$$

we obtain

$$f(x)g''(t) + \frac{EI}{\rho}f^{(4)}(x)g(t) = 0$$

or

$$\frac{g''(t)}{g(t)} = -\frac{EI}{\rho}\frac{f^{(4)}(x)}{f(x)}.$$

[8] The minus sign comes from the fact that counterclockwise moment is positive.

Since the left hand side does not depend on x and the right hand side does not depend on t, both sides are constant. So

$$g''(t) = -\omega^2 g(t) \tag{3.9.3}$$

$$f^{(4)}(x) = \frac{\omega^2\rho}{EI} f(x). \tag{3.9.4}$$

Equation (3.9.3) says that $g(t)$ is a multiple of $\sin(\omega t + \phi)$, while equation (3.9.4) has solutions

$$f(x) = A \sin \kappa x + B \cos \kappa x + C \sinh \kappa x + D \cosh \kappa x$$

where

$$\kappa = \sqrt[4]{\frac{\omega^2\rho}{EI}} \tag{3.9.5}$$

(see Appendix C for the hyperbolic functions sinh and cosh). The general solution then decomposes as a sum of the normal modes

$$y = (A \sin \kappa x + B \cos \kappa x + C \sinh \kappa x + D \cosh \kappa x) \sin(\omega t + \phi). \tag{3.9.6}$$

The boundary conditions depend on what happens at the end of the rod. It is these boundary conditions which constrain ω to a discrete set of values. If an end of the rod is free, then the quantities $V(x,t)$ and $M(x,t)$ have to vanish for all t, at the value of x corresponding to the end of the rod. So $\partial^2 y/\partial x^2 = 0$ and $\partial^3 y/\partial x^3 = 0$. If an end of the rod is clamped, then the displacement and slope vanish, so $y = 0$ and $\partial y/\partial x = 0$ for all t at the value of x corresponding to the end of the rod.

We calculate

$$\begin{aligned}
\partial y/\partial x &= \kappa(A \cos \kappa x - B \sin \kappa x + C \cosh \kappa x + D \sinh \kappa x)\\
\partial^2 y/\partial x^2 &= \kappa^2(-A \sin \kappa x - B \cos \kappa x + C \sinh \kappa x + D \cosh \kappa x)\\
\partial^3 y/\partial x^3 &= \kappa^3(-A \cos \kappa x + B \sin \kappa x + C \cosh \kappa x + D \sinh \kappa x).
\end{aligned}$$

In the case of the xylophone or tubular bell, both ends are free. We take the two ends to be at $x = 0$ and $x = \ell$. The conditions $\partial^2 y/\partial x^2 = 0$ and $\partial^3 y/\partial x^3 = 0$ at $x = 0$ give $B = D$ and $A = C$. These conditions at $x = \ell$ give

$$\begin{aligned}
A(\sinh \kappa\ell - \sin \kappa\ell) + B(\cosh \kappa\ell - \cos \kappa\ell) &= 0\\
A(\cosh \kappa\ell - \cos \kappa\ell) + B(\sinh \kappa\ell + \sin \kappa\ell) &= 0.
\end{aligned}$$

These equations admit a nonzero solution in A and B exactly when the determinant

$$(\sinh \kappa\ell - \sin \kappa\ell)(\sinh \kappa\ell + \sin \kappa\ell) - (\cosh \kappa\ell - \cos \kappa\ell)^2$$

vanishes. Using the relations $\cosh^2 \kappa\ell - \sinh^2 \kappa\ell = 1$ and $\sin^2 \kappa\ell + \cos^2 \kappa\ell = 1$, this condition becomes

$$\cosh \kappa\ell \cos \kappa\ell = 1.$$

The values of $\kappa\ell$ for which this equation holds determine the allowed frequencies via the formula (3.9.5).

Set $\lambda = \kappa\ell$, so that λ has to be a solution of the equation

$$\cosh\lambda\cos\lambda = 1. \tag{3.9.7}$$

Then equation (3.9.5) shows that the angular frequency and the frequency are given by

$$\omega = \sqrt{\frac{EI}{\rho}}\,\frac{\lambda^2}{\ell^2}; \qquad \nu = \frac{\omega}{2\pi} = \sqrt{\frac{EI}{\rho}}\,\frac{\lambda^2}{2\pi\ell^2}. \tag{3.9.8}$$

Numerical computations for the positive solutions to equation (3.9.7) give the following values, with more accuracy than is strictly necessary.

$$\begin{aligned}
\lambda_1 &= 4.730040744862704026024048\,1\\
\lambda_2 &= 7.853204624095837556477066\,7\\
\lambda_3 &= 10.995607838001670906669032\,5\\
\lambda_4 &= 14.137165491257464177105917\,9\\
\lambda_5 &= 17.278759657399481438091074\,0\\
\lambda_6 &= 20.420352245626061090936411\,2
\end{aligned}$$

As n increases, $\cosh\lambda_n$ increases exponentially, and so $\cos\lambda_n$ has to be very small and positive. So λ_n is close to $(n+\frac{1}{2})\pi$, the nth zero of the cosine function. For $n \geq 5$, the approximation

$$\lambda_n \approx (n+\tfrac{1}{2})\pi - (-1)^n 2e^{-(n+\frac{1}{2})\pi} - 4e^{-2(n+\frac{1}{2})\pi} \tag{3.9.9}$$

holds to at least 20 decimal places.[9]

Using equation (3.9.8), we find that the frequency ratios as multiples of the fundamental are given by the quantities λ_n^2/λ_1^2:

n	λ_n^2/λ_1^2
1	1.00000000000000
2	2.75653850709996
3	5.40391763238332
4	8.93295035238193
5	13.34428669366689
6	18.63788788658119

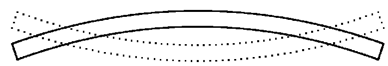
The resulting set of frequencies is certainly inharmonic, just as in the case of the drum. But as n increases, equation (3.9.9) shows that the higher partials have ratios approximating those of the squares of odd integers.

The vibrational modes described by the above values of λ correspond to the following pictures.

[9] This series continues as follows:

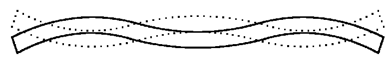
$$\lambda_n \approx (n+\tfrac{1}{2})\pi - (-1)^n 2e^{-(n+\frac{1}{2})\pi} - 4e^{-2(n+\frac{1}{2})\pi} - (-1)^n \tfrac{34}{3}e^{-3(n+\frac{1}{2})\pi} - \tfrac{112}{3}e^{-4(n+\frac{1}{2})\pi} - \cdots$$

The (difficult) challenge to the reader is to compute the next few terms! As a check, $m!$ times the fraction in front of the mth exponential term should be an integer. The answer to this challenge can be found in Appendix A.



λ_1

λ_2

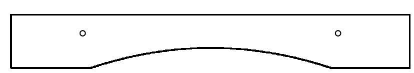


λ_3

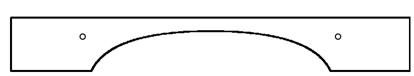
etc.

For actual instruments, rather than the idealized bar described above, the series of partials is somewhat different. Tubular bells are the closest to the ideal situation described above, with second and third partials at frequency ratios of 2.76:1 and 5.40:1 to the fundamental.

The bars of an orchestral xylophone are made of rosewood, or sometimes of more modern materials which are more durable and keep their pitch under more extreme conditions. There is a shallow arch cut out from the underside, with the intention of producing frequency ratios of 3:1 and 6:1 for the second and third partials with respect to the fundamental. These partials correspond to tones an octave and a perfect fifth, respectively two octaves and a perfect fifth above the fundamental.



The marimba is also made of rosewood, and the vibe is made from aluminum. For these instruments, a deeper arch is cut out from the underside, with the intention of producing frequency ratios of 4:1 and 10:1 with respect to the fundamental. These represent tones two octaves, respectively three octaves and a major third above the fundamental.



The tuning of the second partial can be made quite precise, because material removed from different parts of the bar affect different partials. Removing material from the end increases the fundamental and the partials. Taking material away from the sides of the arch lowers the second partial, while taking it from the center of the arch lowers the fundamental frequency. The third partial is harder to make accurate, and so in practice, while the fundamental and second partial are usually reasonably precise, the third is often further from the desired pitch. Tuning is carried out using stroboscopic equipment, which allows for tuning of the fundamental and second partial to within plus or minus one cent (a cent is a hundredth of a semitone).

Further reading:

Antoine Chaigne and Vincent Doutaut, *Numerical simulations of xylophones. I. Time-domain modeling of the vibrating bars*, J. Acoust. Soc. Am. 101 (1) (1997), 539–557.

R. Courant and D. Hilbert, *Methods of mathematical physics, I*, Interscience, 1953, §V.4.

William C. Elmore and Mark A. Heald, *Physics of waves* [**37**], Chapter 3.

Neville H. Fletcher and Thomas D. Rossing, *The physics of musical instruments* [**40**], §19.

D. Holz, *Investigations on acoustically important qualities of xylophone-bar materials: Can we substitute any tropical woods by European species?*, in *Proc. Int. Symp. Musical Acoustics*, Jouve, Paris (1995), 351–357.

A. M. Jones, *Africa and Indonesia: the evidence of the xylophone and other musical and cultural factors*, E. J. Brill, Leiden, 1964. This book contains a large number of measurements of the tuning of African and Indonesian xylophones. The author argues the hypothesis that there was Indonesian influence on African music, and therefore visitations to Africa by Indonesians, long before the Portuguese colonization of Indonesia.

James L. Moore, *Acoustics of bar percussion instruments*, Permus Publications, Columbus, Ohio, 1978.

Thomas D. Rossing, *Science of percussion instruments* [**122**], Chapters 5–7.

3.10. The mbira

At a lecture demonstration I once attended in Seattle, Washington, Dumisani Maraire, a visiting artist from Zimbabwe, walked onto the stage carrying a round-box resonator with a fifteen-key instrument inside. He turned toward the audience and raised the round-box over his head. "What is this?" he called out.

There was no response.

"All right," he said, "it is an mbira; M-B-I-R-A. Now what did I say it was?"

A few people replied, "Mbira; it is an mbira." Most of the audience sat still in puzzlement.

"What is it?" Maraire repeated, as if slightly annoyed.

More people called out, "Mbira."

"Again," Maraire insisted.

"Mbira!" returned the audience.

"Again!" he shouted. When the auditorium echoed with "Mbira," Maraire laughed out loud. "All right," he said with good-natured sarcasm, "that is the way the Christian missionaries taught me to say 'piano.' "

Paul F. Berliner, *The soul of mbira.*

The *mbira* is a popular melodic instrument of Africa, especially the Shona people of Zimbabwe. Other names for the instrument are *sanzhi*, *likembe* and *kalimba*; the general ethnomusicological category is the *lamellophone*. It consists of a set of keys on a soundboard, usually with some kind of resonator such as a gourd for amplifying and transmitting the sound. The keys are usually metal, clamped at one end and free at the other. They are depressed with the finger or thumb and suddenly released to produce the vibration.

The method of the §3.9 can be used to analyze the resonant modes of the keys of the mbira. There is no change up to the point where the boundary conditions are applied to equation (3.9.6). We take the clamped end to be at $x = 0$ and the free end at $x = \ell$. At $x = 0$, the condition $y = 0$ gives $D = -B$ and $\partial y/\partial x = 0$ gives $C = -A$. The conditions $\partial^2 y/\partial x^2 = 0$ and

Picture of mbira from Zimbabwe, from Jacqueline Cogdell DjeDje, "Turn up the volume! A celebration of African music," UCLA 1999, p. 240.

$\partial^3 y/\partial x^3 = 0$ at $x = \ell$ then give

$$-A(\sin\kappa\ell + \sinh\kappa\ell) - B(\cos\kappa\ell + \cosh\kappa\ell) = 0$$
$$-A(\cos\kappa\ell + \cosh\kappa\ell) + B(\sin\kappa\ell - \sinh\kappa\ell) = 0.$$

These equations admit a nonzero solution in A and B exactly when the determinant

$$-(\sin\kappa\ell + \sinh\kappa\ell)(\sin\kappa\ell - \sinh\kappa\ell) - (\cos\kappa\ell + \cosh\kappa\ell)^2$$

vanishes. This time, the equation reduces to

$$\cosh\kappa\ell\cos\kappa\ell = -1.$$

Setting $\lambda = \kappa\ell$ as before, we find that λ has to be a solution of the equation

$$\cosh\lambda\cos\lambda = -1. \tag{3.10.1}$$

Then the angular frequency and the frequency are again given by equation (3.9.8). The following are the first few solutions of equation (3.10.1).

$$\begin{aligned}
\lambda_1 &= 1.87510406871196116644530 82\\
\lambda_2 &= 4.69409113297417457643639 18\\
\lambda_3 &= 7.85475743823761256486100 86\\
\lambda_4 &= 10.9955407348754669906673491\\
\lambda_5 &= 14.1371683910464705809170468\\
\lambda_6 &= 17.2787595320882363335439284
\end{aligned}$$

Notice that these are approximately the same as the values found in the last section, except that there is one extra value playing the role of the fundamental. The analog of equation (3.9.9) is

$$\lambda_n \approx (n - \tfrac{1}{2})\pi - (-1)^n 2e^{-(n-\frac{1}{2})\pi} - 4e^{-2(n-\frac{1}{2})\pi}$$

which holds to at least 20 decimal places for $n \geq 6$. The frequency ratios as multiples of the fundamental are given by the quantities λ_n^2/λ_1^2:

n	λ_n^2/λ_1^2
1	1.00000000000000
2	6.26689302577067
3	17.54748193680844
4	34.38606115720300
5	56.84262292810201
6	84.91303597071318

Of course, the above figures are based on an idealized mbira with constant cross section for the keys. The keys of an actual mbira are very far from constant in cross section, and so the actual relative frequencies of the partials may be far from what is described by the above table. But the most prominent feature, namely that the frequencies of the partials increase quite rapidly, holds in actual instruments.

Further reading:

Paul F. Berliner, *The soul of mbira; music and traditions of the Shona people of Zimbabwe*, University of California Press, 1978. Reprinted by University of Chicago Press, 1993.

3.11. The gong

As a first approximation, the gong can be thought of as a circular flat stiff metal plate of uniform thickness. In practice, the gong is slightly curved, and the thickness is not uniform, but for the moment we shall ignore this. The stiff metal plate behaves like a mixture of the drum and the stiff rod. So the partial differential equation governing its motion is fourth order, as in the case of the stiff rod, but there are two directions in which to take partial derivatives, as in the case of the drum. If z represents displacement, and x and y represent Cartesian coordinates on the gong, then the equation is

$$\frac{\partial^2 z}{\partial t^2} + \frac{Eh^2}{12\rho(1-s^2)}\nabla^4 z = 0. \qquad (3.11.1)$$

This equation first appears (without the explicit value of the constant in front of the second term) in a paper of Sophie Germain.[10] In this equation, h is the thickness of the plate, and an easy calculation shows that $\frac{h^2}{12} = \frac{1}{h}\int_{-h/2}^{+h/2} z^2\,dz$ is the corresponding sectional moment in the one thickness direction (in the case of the stiff rod, there were two dimensions for the cross-section, so the case of the stiff plate is easier in this regard). The quantity E is the Young's modulus as before, ρ is area density, and s is *Poisson's ratio.* This is a measure of the ratio of sideways spreading to the compression. The extra factor of $(1-s^2)$ in the denominator on the right hand of the above equation does not correspond to any term in equation (3.9.2). It arises from the fact that when the plate is bent downwards in one direction, it causes it to curl up in the perpendicular direction along the plate.

The term $\nabla^4 z$ denotes

$$\nabla^2\nabla^2 z = \frac{\partial^4 z}{\partial x^4} + 2\frac{\partial^4 z}{\partial x^2 \partial y^2} + \frac{\partial^4 z}{\partial y^4}.$$

Observe the cross terms carefully. Without them, a rotational change of coordinates would not preserve this operation.

[10] Sophie Germain's paper, "Recherches sur la théorie des surfaces élastiques," written in 1815 and published in 1821, won her a prize of a kilogram of gold from the French Academy of Sciences in 1816. The paper contained some significant errors, but became the basis for work on the subject by Lagrange, Poisson, Kirchoff, Navier and others.

Sophie Germain is probably better known for having made one of the first significant breakthroughs in the study of Fermat's last theorem. She proved that if x, y and z are integers satisfying $x^5 + y^5 = z^5$, then at least one of x, y and z has to be divisible by 5. More generally, she showed that the same was true when 5 is replaced by any prime p such that $2p+1$ is also a prime.

Gong from the Music Research Institute in Beijing
From *The Musical Arts of Ancient China*, exhibit 20.

In the case of the stiff rod, we had to use the hyperbolic functions as well as the trigonometric functions. In this case, we are going to need to use the *hyperbolic Bessel functions*. These are defined by

$$I_n(z) = i^{-n} J_n(iz),$$

and bear the same relationship to the ordinary Bessel functions that the hyperbolic functions $\sinh x$ and $\cosh x$ do to the trigonometric functions $\sin x$ and $\cos x$.

Looking for separable solutions $z = Z(x, y)h(t) = f(r)g(\theta)h(t)$ to equation (3.11.1), we arrive at the equations

$$\nabla^4 Z = \kappa^4 Z \tag{3.11.2}$$

and

$$\frac{\partial^2 h}{\partial t^2} = -\omega^2 h \tag{3.11.3}$$

where ω and κ are related by

$$\kappa^4 = \frac{12\rho(1 - s^2)\omega^2}{Eh^2}.$$

We factor equation (3.11.2) as

$$(\nabla^2 - \kappa^2)(\nabla^2 + \kappa^2)z = 0. \tag{3.11.4}$$

So any solution to either the equation

$$\nabla^2 z = \kappa^2 z \tag{3.11.5}$$

or to the equation

$$\nabla^2 z = -\kappa^2 z \tag{3.11.6}$$

is also a solution to (3.11.2).

LEMMA 3.11.1. *Every solution z to equation* (3.11.2) *can be written uniquely as $z_1 + z_2$ where z_1 satisfies equation* (3.11.5) *and z_2 satisfies equation* (3.11.6).

PROOF. We use a variation of the even and odd function method. If $\nabla^4 z = \kappa^4 z$, we set

$$z_1 = \tfrac{1}{2}(z + \kappa^{-2}\nabla^2 z), \qquad z_2 = \tfrac{1}{2}(z - \kappa^{-2}\nabla^2 z).$$

Then

$$\begin{aligned}
\nabla^2 z_1 &= \tfrac{1}{2}(\nabla^2 z + \kappa^{-2}\nabla^4 z) = \tfrac{1}{2}(\nabla^2 z + \kappa^2 z) = \kappa^2 z_1, \\
\nabla^2 z_2 &= \tfrac{1}{2}(\nabla^2 z - \kappa^{-2}\nabla^4 z) = \tfrac{1}{2}(\nabla^2 z - \kappa^2 z) = -\kappa^2 z_2.
\end{aligned}$$

and $z_1 + z_2 = z$.

For the uniqueness, if z_1' and z_2' constitute another choice, then rearranging the equation $z_1 + z_2 = z_1' + z_2'$, we have $z_1 - z_1' = z_2' - z_2$. The common value z_3 of $z_1 - z_1'$ and $z_2' - z_2$ satisfies both equations (3.11.5) and (3.11.6). So $z_3 = \kappa^{-2}\nabla^2 z_3 = -z_3$, and hence $z_3 = 0$. It follows that $z_1 = z_1'$ and $z_2 = z_2'$. $\square$

Solving equation 3.11.5 is just the same as in the case of the drum, and the solutions are given as trigonometric functions of θ multiplied by Bessel functions of r. Equation 3.11.6 is similar, except that we must use the hyperbolic Bessel functions instead of the Bessel functions. We then have to combine the two classes of solutions in order to satisfy the boundary conditions, just as we did with the trigonometric and hyperbolic functions for the stiff rod. This leads us to solutions of the form

$$z = (AJ_n(\kappa r) + BI_n(\kappa r))\sin(\omega t + \phi)\sin(n\theta + \psi).$$

The boundary conditions for the gong require considerable care, and the first correct analysis was given by Kirchoff in 1850. His boundary conditions can be stated for any region with smooth boundary. Choosing coordinates in such a way that the element of boundary is a small segment of the y axis going through the origin, they are as follows.

$$\frac{\partial^2 z}{\partial x^2} + s\frac{\partial^2 z}{\partial y^2} = 0$$

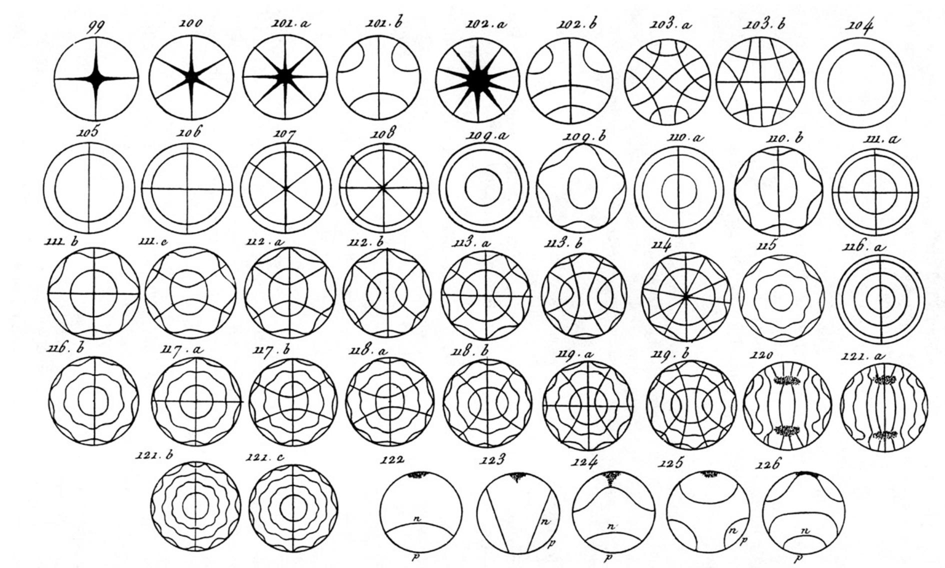
$$\frac{\partial^3 z}{\partial x^3} + (2 - s)\frac{\partial^3 w}{\partial x \partial y^2} = 0.$$

The derivation of these equations may be found in Chapter X of the first volume of Rayleigh's *The theory of sound* [**109**], §216, where these boundary conditions appear as equations (6). He goes on to find the normal modes and eigenvalues by Fourier series methods. The results are similar to those for the drum in §3.6, but the modes with $k = 0$ and $n = 0$ or $n = 1$ are missing; it is easy to see why if we try to imagine the corresponding vibration of a gong. So the fundamental mode is $k = 0$ and $n = 2$. The relative frequencies are tabulated in §3.6 of Fletcher and Rossing, *The physics of musical instruments*, and reproduced in the following table.

k	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
0	—	—	1.000	2.328	4.11	6.30
1	1.73	3.91	6.71	10.07	13.92	18.24
2	7.34	11.40	15.97	21.19	27.18	33.31

Vibrational frequencies for a free circular plate

Actual gongs in real life are not perfect circular plates. Many designs feature circularly symmetric raised portions in the middle of the gong. This modifies the frequencies of the normal modes and the character of the sound. Often, eigenvalues become close enough together to degenerate, and then normal modes can mix. This seems to be in evidence in Chladni's original drawings (see also page 104).



From E. F. F. Chladni, *Traité d'acoustique*, Courcier, Paris 1809.

Cymbals are similar in design, and the theory works in a similar way. Because of the deviation from flatness, the normal modes again tend to combine in interesting ways. For example, the mode $(n, k) = (7, 0)$ and the mode

$(2,1)$ or $(3,1)$ are often close enough in frequency to degenerate into a single compound mode (see Rossing and Peterson, 1982).

Further reading:

R. Courant and D. Hilbert, *Methods of mathematical physics, I*, Interscience, 1953, §V.6.

Neville H. Fletcher and Thomas D. Rossing, *The physics of musical instruments* [**40**], §§3.5–3.6 and §20.

Karl F. Graff, *Wave motion in elastic solids* [**48**].

Philip M. Morse and K. Uno Ingard, *Theoretical acoustics* [**91**], §5.3.

J. W. S. Rayleigh, *The theory of sound* [**109**], Chapter X.

Thomas D. Rossing, *Science of percussion instruments* [**122**], Chapters 8 and 9.

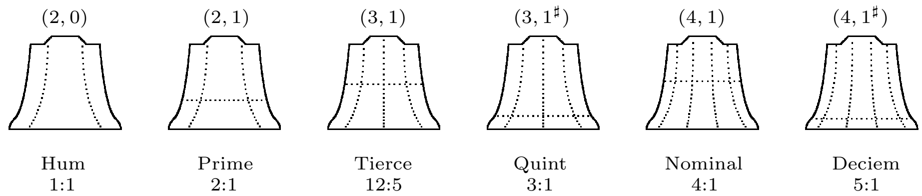
Thomas D. Rossing and Neville H. Fletcher, *Nonlinear vibrations in plates and gongs*, J. Acoust. Soc. Am. 73 (1983), 345–351.

Thomas D. Rossing and R. W. Peterson, *Vibrations of plates, gongs and cymbals*, Percussive Notes 19 (3) (1982), 31.

M. D. Waller, *Vibrations of free circular plates. Part I: Normal modes*, Proc. Phys. Soc. 50 (1938), 70–76.

3.12. The bell

Church bells are used for *change ringing*, a subject discussed further in §9.4. A bell can be thought of as a very deformed plate; its vibrational modes are similar in nature, but starting with $n = 2$. But the exact shape of the bell is made so as to tune the various vibrational modes relative to each other. There are five modes with special names, which are as follows. The mode $(n,k) = (2,0)$ is the fundamental, and is called the *hum*. The *prime* is $(2,1)$, and is tuned to twice the frequency, putting it an octave higher. There are two different modes $(3,1)$, one of which has the stationary circle around the waist, and the other nearer the rim. The one with the waist is called the *tierce*, and is tuned a minor third above the prime. The other mode, sometimes denoted $(3,1^\sharp)$, with the stationary circle nearer the rim is called the *quint*. It is pitched a perfect fifth above the prime. The *nominal* mode is $(4,1)$, tuned an octave above the prime, so that it is two octaves above the hum. The nominal mode is by far the one with the largest amplitude, so that this is the perceived pitch of the bell. Mode $(4,1^\sharp)$ is sometimes called the *deciem*, and is usually tuned a *major* third above the nominal. It can be imagined that a great deal of skill goes into the tuning of the vibrational modes of a bell. It is an art which has developed over many centuries. Particular attention is given to the construction of the thick ring near the rim. The information described above is summarized in the following diagram.



The singing bowl. Our discussion of the bell applies equally well to other objects such at the Tibetan *singing bowl*, which is used mainly for ritual purposes. The singing bowl is struck with a wooden mallet on the inside of the rim to set it vibrating. The tone can often last in excess of a minute before it is inaudible.

I have a Tibetan singing bowl in my living room, which is approximately 19 cm (or 7 inches) in diameter. It has two clearly audible partials, and some others too high to hear the pitch very precisely. The fundamental sounds at about 196 Hz, and the second partial sounds at about 549 Hz, giving a ratio of about 2.8:1.

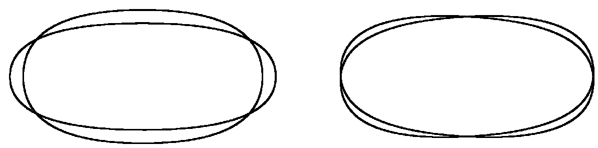
Singing bowl from Tibet

Chinese bells. In 1977, an extraordinary discovery was made in the Hubei province of China. A huge burial pit was found, containing over four thousand bronze items. This was the tomb of Marquis Yi of the state of Zeng, and inscriptions date it very precisely at 433 B.C. The tomb contains many musical instruments, but the most extraordinary is a set of sixty-five bronze bells. These are able to play all twelve notes of the chromatic scale over a range of three octaves, and further bells fill this out to a five octave range.

Each bell is roughly elliptical in cross-section. There are two separate strike points, and the bell is designed so that the normal modes excited at the strike points have essentially nothing in common. So the pereived pitches are quite different. The strike point for the lower pitch is called the *sui*, and the one for the higher pitch is the *gu*. The bells are tuned so that this difference is either a major third or a minor third. The separation of the modes is

Bell from the tomb of Marquis Yi
(middle tier, height 75cm, weight 32.2kg)
Picture from *Music in the age of Confucius*, p. 43.

achieved through the use of an elaborate set of nipples on the outer surface of the bell. See the picture on page 128. The values of n and k are the same for the *sui* and *gu* version of a vibrational mode, but the orientation is different. This may be illustrated as follows for the modes with $n = 2$, where the diagram represents the movement of the lower rim.



sui *gu*

It seems very hard to understand how these two tone bells were cast. The inscriptions naming the two tones were cast with the bell, so they must have been predetermined. What's more, the design does not just scale up proportionally, and there is no easy formula for how to produce a larger bell with the same musical interval. Modern physics does not lead to any understanding of the design procedures that were used to produce this set of bells.

Further reading:

Lothar von Falkenhausen, *Suspended Music: Chime-bells in the culture of bronze age China*, University of California Press, Berkeley, 1993.

Neville H. Fletcher and Thomas D. Rossing, *The physics of musical instruments* [**40**], §21.

Yuan-Yuan Lee and Sin-yan Shen, *Chinese musical instruments*, Chinese Music Society of North America, Chicago, USA, 1999.

Thomas D. Rossing, *The acoustics of bells*, Van Nostrand Reinhold, 1984.

Thomas D. Rossing, *The science of sound* [**121**], §13.4.

Thomas D. Rossing, *Science of percussion instruments* [**122**], Chapters 11–13.

Jenny So, *Eastern Zhou ritual bronzes from the Arthur M. Sackler collections*, Smithsonian Institution, 1995. This is a large format book with photographs and descriptions of Chinese bronzes from the Eastern Zhou. Pages 357–397 describe the two tone bells from the collection. There is also an extensive appendix (pages 431–484) titled "Acoustical and musical studies on the Sackler bells," by Lothar von Falkenhausen and Thomas D. Rossing. This appendix gives a great many technical details of the acoustics and tuning of two tone bells.

Jenny So (ed.), *Music in the age of Confucius*, Sackler Gallery, Washington, 2000. This beautifully produced book contains an extensive set of photographs of the set of bells from the tomb of Marquis Yi.

3.13. Acoustics

The basic equation of acoustics is the three dimensional wave equation, which describes the movement of air to form sound. The discussion is similar to the one dimensional discussion in §3.5. Recall that acoustic pressure p is measured by subtracting the (constant) ambient air pressure ρ from the absolute pressure P. In three dimensions, p is a function of x, y, z and t. This is related to the displacement vector field $\boldsymbol{\xi}(x, y, z, t)$ by two equations. The first is *Hooke's law*, which in this situation can be written as

$$p = -B\,\nabla.\,\boldsymbol{\xi}$$

where B is the bulk modulus of air. Newton's second law of motion implies that

$$\nabla p = -\rho \frac{\partial^2 \boldsymbol{\xi}}{\partial t^2}.$$

Putting these two equations together gives

$$\nabla^2 p = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} \tag{3.13.1}$$

where $c = \sqrt{B/\rho}$. So p satisfies the three dimensional wave equation.

In an enclosed space, the boundary conditions are given by $\nabla p = 0$ on the walls of the enclosure for all t. Looking for separable solutions leads to the theory of Dirichlet and Neumann eigenvalues, just as in the two dimensional case when we discussed the drum in §3.6. So there is a certain set of *resonant frequencies* for the enclosure, determined by the eigenvalues of ∇^2 in the region. The same reasoning as in §3.6 leads to the conclusion that the relationship between frequency and eigenvalue is $\nu = c\sqrt{\lambda}/2\pi$, see equation (3.7.4). For an enclosure of small total volume, the eigenvalues are widely spaced. But as the volume increases, the eigenvalues get closer together. So for example a concert hall has a large total volume, and the eigenvalues are typically at intervals of a few Hertz, and the spacing is somewhat erratic. Fortunately, the ear is performing a windowed Fourier analysis with a relatively short time window, so that in accordance with Heisenberg's uncertainty principle, fluctuations on a fine frequency scale are not noticed.[11]

There is one useful situation where we can explicitly solve the three dimensional wave equation, namely where there is complete spherical symmetry. This corresponds to a physical situation where sound waves are generated at the origin in an anisotropic fashion. In this case, we convert into spherical coordinates, and ignore derivatives with respect to the angles. Denoting radial distance from the origin by r, the equation becomes

$$\frac{\partial^2(rp)}{\partial r^2} = \frac{1}{c^2}\frac{\partial^2(rp)}{\partial t^2}.$$

---

[11]See pages 72–73 of Manfred Schroeder, *Fractals, chaos and power laws*, Springer-Verlag, 1991.

Regarding rp as the dependent variable, this is really just the one dimensional wave equation. So d'Alembert's Theorem 3.2.1 shows that the general solution is given by

$$p = (f(r + ct) + g(r - ct))/r.$$

The functions f and g represent waves travelling towards and away from the origin, respectively. Notice that the sound source needs to have finite size, so that we do not run into problems at $r = 0$.

Exercises

1. Show that if $\mathbf{u}$ is a unit vector in some direction in three dimensions, then the function

$$p(\mathbf{x}, t) = e^{i\omega(ct - \mathbf{u}.\mathbf{x})}$$

satisfies the three dimensional wave equation (3.13.1). This (or rather its real part) represents a sound wave travelling in the direction of $\mathbf{u}$ with speed c and angular velocity ω.

2. Find the solutions to the three dimensional wave equation for an enclosed region in the shape of a cuboid. Use separation of variables for all four variables, and place the origin at a corner of the region to make the calculations easier.

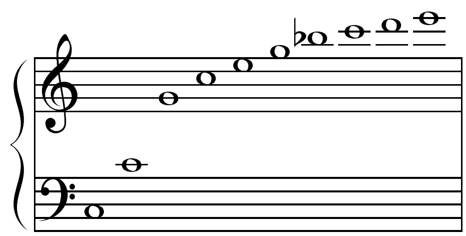
CHAPTER 4

Consonance and dissonance

In this chapter, we investigate the relationship between consonance and dissonance, and simple integer ratios of frequencies.

4.1. Harmonics

When a note on a stringed instrument or a wind instrument sounds at a certain pitch, say with frequency ν, all that really means is that the sound is (roughly) periodic with that frequency. The theory of Fourier series shows that such a sound can be decomposed as a sum of sine waves with various phases, at integer multiples of the frequency ν. The component of the sound with frequency ν is called the *fundamental*. The component with frequency $m\nu$ is called the mth *harmonic*, or the $(m-1)$st *overtone*. So for example if $m = 3$ we obtain the third harmonic, or the second overtone.[1]



1 2 3 4 5 6 7 8 9 10

This diagram represents the series of harmonics based on a fundamental at the C below middle C. The seventh harmonic is actually somewhat flatter than the B♭ above the treble clef. In the modern equally tempered scale, even the third and fifth harmonics are very slightly different from the notes G and E shown above—this is more extensively discussed in Chapter 5.

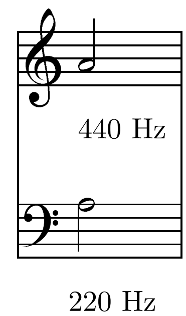
There is another word which we have been using in this context: the mth *partial* of a sound is the mth frequency component, counted from the bottom. So for example on a clarinet, where only the odd harmonics are present, the first partial is the fundamental, or first harmonic, and the second partial is the third harmonic. This term is very useful when discussing sounds where the partials are not simple multiples of the fundamental, such as for example the drum, the gong, or the various instruments of the gamelan.

[1] I find that the numbering of overtones is confusing, and I shall not use this numbering.

Exercises

1. Define the following terms, making the distinctions between them clear:
(a) the mth harmonic, (b) the mth overtone, (c) the mth partial.

4.2. Simple integer ratios



Why is it that two notes an octave apart sound consonant, while two notes a little more or a little less than an octave apart sound dissonant? An interval of one octave corresponds to doubling the frequency of the vibration. So for example, the A above middle C corresponds to a frequency of 440 Hz, while the A below middle C corresponds to a frequency of 220 Hz.

We have seen in Chapter 3 that if we play these notes on conventional stringed or wind (but not percussive) instruments, each note will contain not only a component at the given frequency, but also partials corresponding to multiples of that frequency. So for these two notes we have partials at:

440 Hz, 880 Hz, 1320 Hz, 1760 Hz, ...

220 Hz, 440 Hz, 660 Hz, 880 Hz, ...

On the other hand, if we play two notes with frequencies 440Hz and 225Hz, then the partials occur at:

440 Hz, 880 Hz, 1320 Hz, 1760 Hz, ...

225 Hz, 450 Hz, 675 Hz, 900 Hz, ...

The presence of components at 440 Hz and 450 Hz causes a sensation of roughness, which is interpreted by the ear as dissonance. We shall discuss at length, later in this chapter, the history of different explanations of consonance and dissonance, and why this should be taken to be the correct one.

Because of the extreme consonance of an interval of an octave, and its role in the series of partials of a note, the human brain often perceives two notes an octave apart as being "really" the same note but higher. This is so heavily reinforced by musical usage in every genre that we have difficulty imagining that it could be otherwise. When choirs sing "in unison," this usually means that the men and women are singing an octave apart.[2] The idea that notes differing by a whole number of octaves should be considered as equivalent is often referred to as *octave equivalence*.

[2]It is interesting to speculate what effect it would have on the theory of color if visible light had a span greater than an octave; in other words, if there were to exist two visible colors, one of which had exactly twice the frequency of the other. In fact, the span of human vision is just shy of an octave. This may explain why the colors of the rainbow seem to join up into a circle.

The Experiences of Pythagoras
(Gaffurius, 1492)

The musical interval of a perfect fifth[3] corresponds to a frequency ratio of 3:2. If two notes are played with a frequency ratio of 3:2, then the third partial of the lower note will coincide with the second partial of the upper note, and the notes will have a number of higher partials in common. If, on the other hand, the ratio is slightly different from 3:2, then there will be a sensation of roughness between the third partial of the lower note and the second partial of the upper note, and the notes will sound dissonant.

In this manner, small integer ratios of frequencies are picked out as more consonant than other intervals. We stress that this discussion only works for notes whose partials are at multiples of the fundamental frequency. Pythagoras essentially discovered this in the sixth century B.C.; he discovered that when two similar strings under the same tension are sounded together, they give a pleasant sound if the lengths of the strings are in the ratio of two small integers. This was the first known example of a law of nature ruled by the arithmetic of integers, and greatly influenced the intellectual

[3]We shall see in the next chapter that the fifth from C to G in the modern Western scale is not precisely a perfect fifth.

development of his followers, the Pythagoreans. They considered that a liberal education consisted of the "quadrivium," or four divisions: numbers in the abstract, numbers applied to music, geometry, and astronomy. They expected that the motions of the planets would be governed by the arithmetic of ratios of small integers in a similar way. This belief has become encoded in the phrase "the music of the spheres,"[4] literally denoting the inaudible sound produced by the motion of the planets, and has almost disappeared in modern astronomy (but see the remarks in Exercise 1 of Section 6.2).[5]

4.3. Historical explanations of consonance

In writing this section, I have drawn heavily on the work of Plomp and Levelt. The reference can be found at the end of the section.

The discovery of the relationship between musical pitch and frequency occurred around the sixteenth or seventeenth century, with the work of Galileo Galilei and (independently) Mersenne. Galileo's explanation of consonance was that if two notes have their frequencies in a simple integer ratio, then there is a regularity, or periodicity to the total waveform, not present with other frequency ratios, so that the ear drum is not "kept in perpetual torment."[6] The problem with this explanation is that it involves some circular reasoning—the notes are consonant because the ear finds them consonant! Furthermore, experimentation with tones produced using nonharmonic partials produce results which contradict this explanation, as we shall see in §4.6.

In the seventeenth century, it was discovered that a simple note from a conventional stringed or wind instrument had partials at integer multiples of the fundamental. The eighteenth century theoretician and musician Rameau ([**108**], chapter 3) regarded this as already being enough explanation for the consonance of these intervals, but Sorge[7] (1703–1778) was the first to consider roughness caused by close partials as the explanation of dissonance. It was not until the nineteenth century that Helmholtz (1821–1894) [**54**] sought to explain consonance and dissonance on a more scientific basis. Helmholtz based his studies on the structure of the human ear. His idea was that for small differences between the frequencies of partials, beats can

[4]Plato, *Republic*, 10.617, ca. 380 B.C.

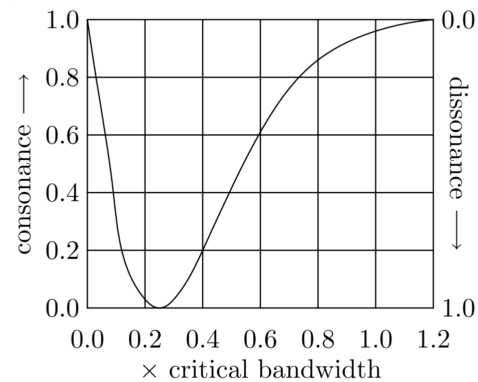
[5]The idea embodied in the phrase "the music of the spheres" is still present in the seventeenth century work of Kepler on the motion of the planets. He called his third law the "harmonic law," and it is described in a work entitled *Harmonices Mundi* (Augsburg, 1619). However, his law properly belongs to physics, and states that the square of the period of a planetary orbit is proportional to the cube of the maximum diameter. It is hard to find any recognizable connection with musical harmony or the arithmetic of ratios of small integers. Kepler's ideas are celebrated in Paul Hindemith's opera, *Die Harmonie der Welt*, 1956–7. The title is a translation of Kepler's.

[6]Galileo Galilei, *Discorsi e dimonstrazioni mathematiche interno à due nuove scienze attenenti alla mecanica ed i movimenti locali*, Elsevier, 1638. Translated by H. Crew and A. de Salvio as *Dialogues concerning two new sciences*, McGraw-Hill, 1963.

[7]G. A. Sorge, *Vorgemach der musicalischen Composition*, Verlag des Autoris, Lobenstein, 1745–1747

be heard, whereas for larger frequency differences, this turns into roughness. He claimed that for maximum roughness, the difference between the two frequencies should be 30–40 Hz, independently of the individual frequencies. For larger frequency differences, the sense of roughness disappears and consonance resumes. He then goes on to deduce that the octave is consonant because all the partials of the higher note are among the partials of the lower note, and no roughness occurs.

Plomp and Levelt, in the nineteen sixties, seem to have been the first to carry out a thorough experimental analysis of consonance and dissonance for a variety of subjects, with pure sine waves, and at a variety of pitches. The results of their experiments showed that on a subjective scale of consonance ranging from zero (dissonant) to one (consonant), the variation with frequency ratio has the shape shown in the graph below. The x axis of this graph is labeled in multiples of the critical bandwidth, defined below. This means that the actual scale in Hertz on the horizontal axis of the graph varies according to the pitch of the notes, but the shape of the graph remains constant; the scaling factor was shown by Plomp and Levelt to be proportional to critical bandwidth.



The salient features of the above graph are that the maximum dissonance occurs at roughly one quarter of a critical bandwidth, and consonance levels off at roughly one critical bandwidth.

It should be stressed that this curve is for pure sine waves, with no harmonics; also that consonance and dissonance is different from recognition of intervals. Anyone with any musical training can recognize an interval of an octave or a fifth, but for pure sine waves, these intervals sound no more nor less consonant than nearby frequency ratios.

Exercises

1. Show that the function $f(t) = A\sin(at) + B\sin(bt)$ is periodic when the ratio of a to b is a rational number, and nonperiodic if the ratio is irrational. [Hint: Differentiate twice and take linear combinations to get a single sine wave, to get information about possible periods]

Further reading:

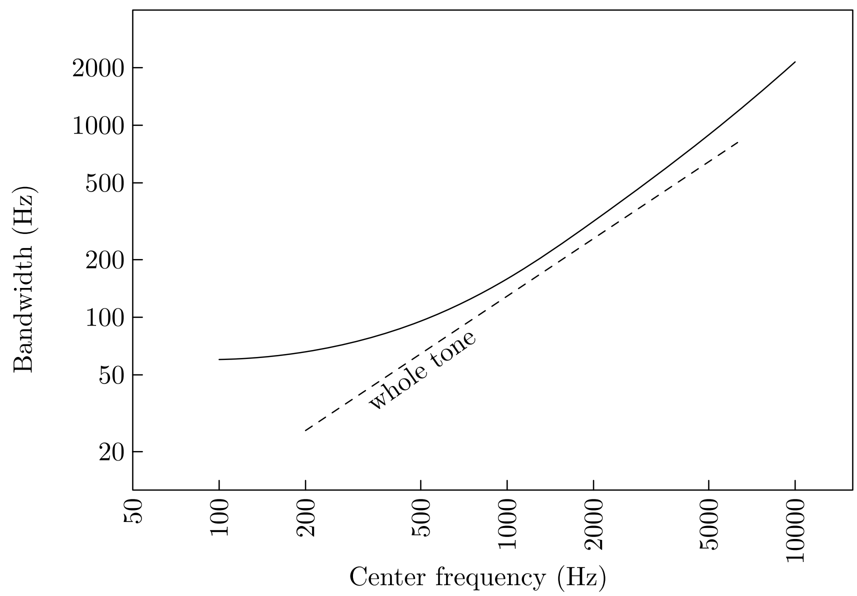
D. D. Greenwood, *Critical bandwidth and the frequency coordinates of the basilar membrane*, J. Acoust. Soc. Am. 33 (1961), 1344–1356.

R. Plomp and W. J. M. Levelt, *Tonal consonance and critical bandwidth*, J. Acoust. Soc. Am. 38 (1965), 548–560.

R. Plomp and H. J. M. Steeneken, *Interference between two simple tones*, J. Acoust. Soc. Am. 43 (1968), 883–884.

4.4. Critical bandwidth

To introduce the notion of *critical bandwidth*, each point of the basilar membrane in the cochlea is thought of as a band pass filter, which lets through frequencies in a certain band, and blocks out frequencies outside that band. The actual shape of the filter is certainly more complicated than this simplified model, in which the left, top and right edges of the envelope of the filter are straight vertical and horizontal lines. This is exactly analogous to the definition of bandwidth given in §1.10, and introducing a smoother shape for the filter does not significantly alter the discussion. The width of the filter in this model is called the critical bandwidth. Experimental data for the critical bandwidth as a function of center frequency is available from a number of sources, listed at the end of this section. Here is a sketch of the results.



Critical bandwidth as a function of center frequency

A rough calculation based on this graph shows that the size of the critical bandwidth is somewhere between a whole tone and a minor third throughout most of the audible range, and increasing to a major third for small frequencies.

Further reading:

B. R. Glasberg and B. C. J. Moore, *Derivation of auditory filter shapes from notched-noise data*, Hear. Res. 47 (1990), 103–138.

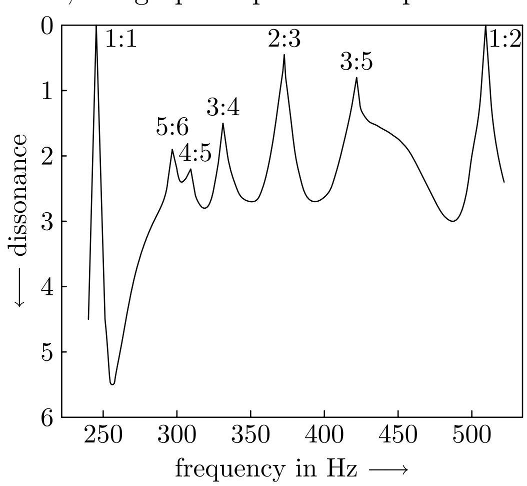
E. Zwicker, *Subdivision of the audible frequency range into critical bands (Frequenzgruppen)*, J. Acoust. Soc. Am. 33 (1961), 248.

E. Zwicker, G. Flottorp and S. S. Stevens, *Critical band width in loudness summation*, J. Acoust. Soc. Am. 29 (1957), 548–557.

E. Zwicker and E. Terhardt, *Analytical expressions for critical-band rate and critical bandwidth as a function of frequency*, J. Acoust. Soc. Am. 68 (1980), 1523–1525.

4.5. Complex tones

Plomp and Levelt took the analysis one stage further, and examined what would happen for tones with a more complicated harmonic content. They worked under the simplifying assumption that the total dissonance is the sum of the dissonances caused by each pair of adjacent partials, and used the above graph for the individual dissonances. They do a sample calculation in which a note has partials at the fundamental and its multiples up to the sixth harmonic. The graph they obtain is shown abelow. Notice the sharp peaks at the fundamental (1:1), the octave (1:2) and the perfect fifth (2:3), and the smaller peaks at ratios 5:6 (just minor third), 4:5 (just major third), 3:4 (perfect fourth) and 3:5 (just major sixth). If higher harmonics are taken into account, the graph acquires more peaks.



In order to be able to draw such Plomp–Levelt curves more systematically, we choose a formula which gives a reasonable approximation to the curve displayed on page 137. Writing x for the frequency difference in multiples of the critical bandwidth, we choose the dissonance function to be[8]

$$f(x) = 4|x|e^{1-4|x|}.$$

This takes its maximum value $f(x) = 1$ when $x = \frac{1}{4}$, as can easily be seen by differentiating. It satisfies $f(0) = 0$, and $f(1)$ is small (about $\frac{1}{5}$), but not zero. This last feature does not quite match the graph given by Plomp and Levelt, but a closer examination of their data shows that the value $f(1) = 0$ is not quite justified.

Further reading:

R. Plomp and W. J. M. Levelt, *Tonal consonance and critical bandwidth*, J. Acoust. Soc. Am. 38 (1965), 548–560.

4.6. Artificial spectra

So what would happen if we artificially manufacture a note having partials which are not exact multiples of the fundamental? It is easy to perform such experiments using a digital synthesizer. We make a note whose partials are at

440 Hz, 860 Hz, 1203 Hz, 1683 Hz, ...

and another with partials at

225 Hz, 440 Hz, 615 Hz, 860 Hz, ...

to represent slightly squeezed harmonics. These notes sound consonant, despite the fact that they are slightly less than an octave apart, whereas scaling the second down to

220 Hz, 430 Hz, 602 Hz, 841 Hz, ...

causes a distinctly dissonant sounding exact octave.

If we are allowed to change the harmonic content of a note in this way, we can make almost any set of intervals seem consonant. This idea was put forward by Pierce (1966, reference below), who designed a spectrum suitable for an equal temperament scale with eight notes to the octave. Namely, he used the following partials, given as multiples of the fundamental frequency:

$$1:1, \quad 2^{\frac{5}{4}}:1, \quad 4:1, \quad 2^{\frac{5}{2}}:1, \quad 2^{\frac{11}{4}}:1, \quad 8:1.$$

This may be thought of as a stretched version of the ordinary series of harmonics of the fundamental. When two notes of the eight tone equal tempered scale are played using synthesized tones with the above set of partials, what happens is that the partials either coincide or are separated by at least $\frac{1}{8}$ of an octave. Pierce's conclusion is that

[8]Sethares [**129**] takes for the dissonance function $f(x) = e^{-b_1 x} - e^{-b_2 x}$ where $b_1 = 3.5$ and $b_2 = 5.75$. This needs normalizing by multiplication by about 5.5, and then gives a graph very similar to the one I have chosen. The particular choice of function is somewhat arbitrary, because of a lack of precision in the data as well as in the subjective definition of dissonance. The main point is to mimic the visible features of the graph.

> ... by providing music with tones that have accurately specified but nonharmonic partial structures, the digital computer can release music from the tyrrany of 12 tones without throwing consonance overboard.

Further reading:

W. Hutchinson and L. Knopoff, *The acoustic component of western consonance*, Interface 7 (1978), 1–29.

A. Kameoka and M. Kuriyagawa, *Consonance theory I: consonance of dyads*, J. Acoust. Soc. Am. 45 (6) (1969), 1452–1459.

A. Kameoka and M. Kuriyagawa, *Consonance theory II: consonance of complex tones and its calculation method*, J. Acoust. Soc. Am. 45 (6) (1969), 1460–1469.

Jenö Keuler, *Problems of shape and background in sounds with inharmonic spectra*, Music, Gestalt, and Computing [**73**], 214–224, with examples from the accompanying CD.

Max V. Mathews and John R. Pierce, *Harmony and nonharmonic partials*, J. Acoust. Soc. Am. 68 (1980), 1252–1257.

John R. Pierce, *Attaining consonance in arbitrary scales*, J. Acoust. Soc. Am. 40 (1966), 249.

John R. Pierce, *Periodicity and pitch perception*, J. Acoust. Soc. Am. 90 (4) (1991), 1889–1893.

William A. Sethares, *Tuning, timbre, spectrum, scale* [**129**]. This book comes with a compact disc full of illustrative examples.

William A. Sethares, *Consonance-based spectral mappings.* Computer Music Journal 22 (1) (1998), 56–72.

Frank H. Slaymaker, *Chords from tones having stretched partials.* J. Acoust. Soc. Am. 47 (1970), 1569–1571.

E. Terhardt, *Pitch, consonance, and harmony.* J. Acoust. Soc. Am. 55 (1974), 1061–1069.

E. Terhardt and M. Zick, *Evaluation of the tempered tone scale in normal, stretched, and contracted intonation.* Acustica 32 (1975), 268–274.

4.7. Combination tones

When two loud notes of different frequencies f_1 and f_2 are played together, a note can be heard corresponding to the difference $f_1 - f_2$ between the two frequencies. This was discovered by the German organist Sorge (1744) and Romieu (1753). Later (1754) the Italian violinist Tartini claimed to have made the same discovery as early as 1714. Helmholtz (1856) discovered that there is a second, weaker note corresponding to the sum of the two frequencies $f_1 + f_2$, but that it is much harder to perceive. The general name for these sum and difference tones is *combination tones*, and the difference notes in particular are sometimes called *Tartini's tones*. The reason (overlooked by Helmholtz) why the sum tone is so hard to perceive is because of the phenomenon of masking discussed at the end of §1.2.

It is tempting to suppose that the combination tones are a result of a discussion similar to the discussion of beats in §1.7. However, this seems to be misleading, as this argument would seem more likely to give rise to notes of half the difference and half the sum of the notes, and this does not seem to be what occurs in practice. Moreover, when we hear beats, we are not hearing a *sound* at the beat frequency, because there is no corresponding place on the basilar membrane for the excitation to occur. Further evidence that these are different phenomena is that when the two tones are heard one with each ear, beats are still discernable, while combination tones are not.

Helmholtz [**54**] (Appendix XII) had a more convincing explanation of combination tones, based on the supposition that the sounds are loud enough for nonlinearities in the response of some part of the auditory system to come into effect.

In the presence of a quadratic nonlinearity, a damped harmonic oscillator with a sum of two sinusoidal forcing terms of different frequencies will vibrate with not only the two incoming frequencies but also with components at twice these frequencies and at the sum and difference of the frequencies. Intuitively, this is because

$$\begin{aligned}(\sin mt + \sin nt)^2 &= \sin^2 mt + 2\sin mt \sin nt + \sin^2 nt \\ &= \tfrac{1}{2}(1-\cos 2mt) + \tfrac{1}{2}(\cos(m-n)t - \cos(m+n)t) + \tfrac{1}{2}(1-\cos 2nt).\end{aligned}$$

So if some part of the auditory system is behaving in a nonlinear fashion, a quadratic nonlinearity would correspond to the perception of doubles of the incoming frequencies, which are probably not noticed because they look like overtones, as well as sum and difference tones corresponding to the terms $\cos(m+n)t$ and $\cos(m-n)t$.

Quadratic nonlinearities involve an asymmetry in the vibrating system, whereas cubic nonlinearities do not have this property. So it seems reasonable to suppose that the cubic nonlinearities are more pronounced in effect than the quadratic ones in parts of the auditory system. This would mean that combination tones corresponding to $2f_1 - f_2$ and $2f_2 - f_1$ would be more prominent than the sum and difference. This seems to correspond to what is experienced in practice. These cubic terms can be heard even at low volume, while a relatively high volume is necessary in order to experience the sum and difference tones.

Helmholtz's theory ([**54**], appendix XII) was that the nonlinearity giving rise to the distortion was occurring in the middle ear, and in particular the tympanic membrane. Measurements made by Guinan and Peake[9] have shown that the nonlinearities in the middle ear are insufficient to explain the

[9] J. J. Guinan and W. T. Peake, *Middle ear characteristics of anesthetized cats.* J. Acoust. Soc. Am. **41** (1967), 1237–1261.

phenomenon. Current theory favors an intracochlear origin for the nonlinearities responsible for the sum and difference tone. Furthermore, the distortions responsible for cubic effects are now thought to have their origins in psychophysical feedback, and are part of the normal auditory function rather than a result of overload.[10]

There is also a related concept of *virtual pitch* for a complex tone. If a tone has a complicated set of partials, we seem to assign a pitch to a composite tone by very complicated methods which are not well understood. Schouten[11] demonstrated that Helmholtz's discussion does not completely explain what happens for these more complex sounds. If the ear is simultaneously subjected to sounds of frequencies 1800 Hz, 2000 Hz and 2200 Hz then the subject hears a tone at 200 Hz, representing a "missing fundamental," and which might be interpreted as a combination tone. However, if the sounds have frequencies 1840 Hz, 2040 Hz and 2240 Hz then instead of hearing a 200 Hz tone as would be expected by Helmholtz's theory, the subject actually hears a tone at 204 Hz. Schouten's explanation for this has been disputed in more recent work, and it is probably fair to say that the subject is still not well understood.

Walliser[12] has given a recipe for determining the perceived missing fundamental, without supplying a mechanism which explains it. His recipe consists of determining the difference in frequency between two adjacent partials (or harmonic components of the sound), and then approximating this with as simple as possible a rational multiple of the lowest harmonic component. So in the above example, the difference is 200 Hz, so we take one nineth of 1840 Hz to give a missing fundamental of 204.4 Hz. This is an extremely good approximation to what is actually heard. Later authors have proposed minor modifications to Walliser's algorithm, for example by replacing the lowest partial with the most "dominant" in a suitable sense. A more detailed discussion can be found in chapter 5 of B. C. J. Moore's book [**88**].

Licklider[13] also cast doubt on Helmholtz's explanation for combination tones by showing that a difference tone cannot in practice be masked by a noise with nearby frequency, while it should be masked if Helmholtz's theory were correct.

Combination tones and virtual pitch remain among many interesting topics of modern psychoacoustics, and a current active area of research.

---

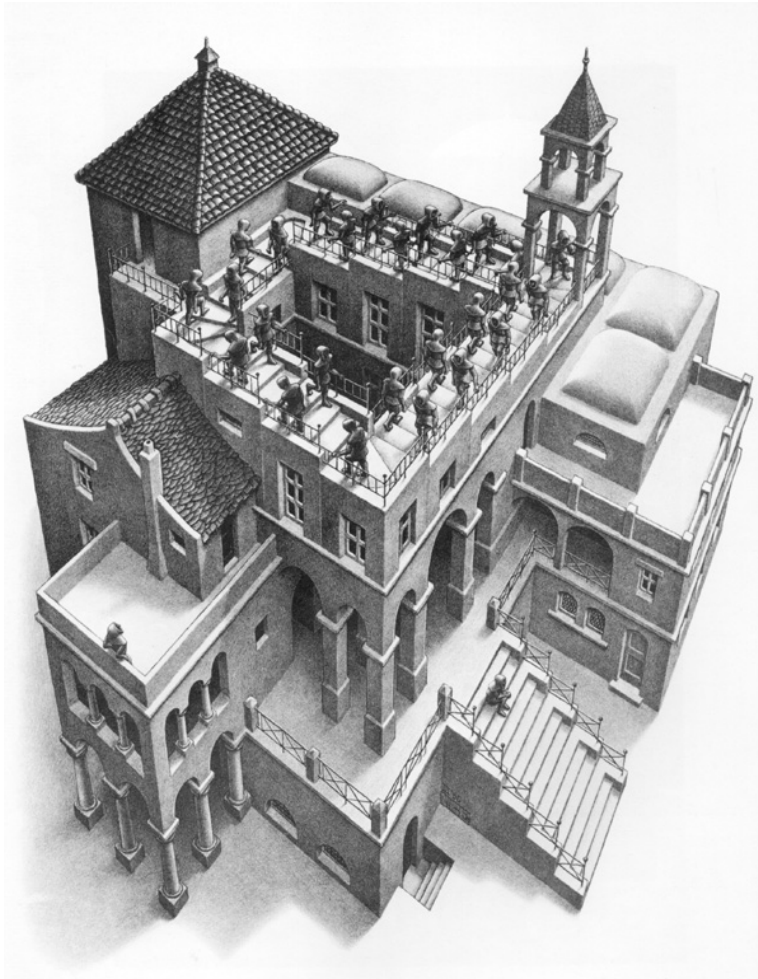
[10] See for example Pickles [**102**], pp. 107–109.

[11] J. F. Schouten, *The residue and the mechanism of hearing*, Proceedings of the Koningklijke Nederlandse Akademie van Wetenschappen **43** (1940), 991–999.

[12] K. Walliser, *Über ein Funktionsschema für die Bildung der Periodentonhöhe aus dem Schallreiz*, Kybernetik **6** (1969), 65–72.

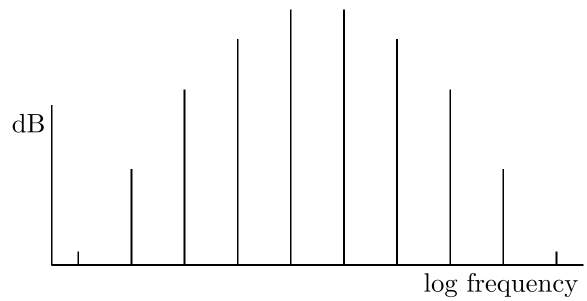
[13] J. C. R. Licklider, *Periodicity by "pitch" and place "pitch"*, J. Acoust. Soc. Am. **26**, (1954), 945.

4.8. Musical paradoxes



M. C. Escher, *Ascending and descending* (1960).

One of the most famous paradoxes of musical perception was discovered by R. N. Shepard, and goes under the name of the Shepard scale. Listening to the Shepard scale, one has the impression of an ever-ascending scale where the end joins up with the beginning, just like Escher's famous ever ascending staircase in his picture, *Ascending and descending.* This effect is achieved by building up each note out of a complex tone consisting of ten partials spaced at one octave intervals. These are passed through a filter so that the middle partials are the loudest, and they tail off at both the bottom and the top. The same filter is applied for all notes of the scale, so that after ascending through one octave, the dominant part of the sound has shifted downwards by one partial.



The partials present in this sound are of the form $2^n.f$, where f is the lowest audible frequency component.

A related paradox, discovered by Diana Deutsch (1975), is called the *tritone paradox*. If two Shepard tones are separated by exactly half an octave (a tritone in the equal tempered scale), or a factor of $\sqrt{2}$, then it might be expected that the listener would be confused as to whether the interval is ascending or descending. In fact, only some listeners experience confusion. Others are quite definite as to whether the interval is ascending or descending, and consistently judge half the possible cases as ascending and the complementary half as descending.

Diana Deutsch is also responsible for discovering a number of other paradoxes. For example, when tones of 400 Hz and 800 Hz are presented to the two ears with opposite phase, about 99% of subjects experience the lower tone in one ear and the higher tone in the other ear. When the headphones are reversed, the lower tone stays in the same ear as before. See her 1974 article in Nature for further details.

Further reading:

E. M. Burns, *Circularity in relative pitch judgments: the Shepard demonstration revisited, again*, Perception and Psychophys. 21 (1977), 563–568.

Diana Deutsch, *An auditory illusion*, Nature 251 (1974), 307–309.

Diana Deutsch, *Musical illusions*, Scientific American 233 (1975), 92–104.

Diana Deutsch, *A musical paradox*, Music Percept. 3 (1986), 275–280.

Diana Deutsch, *The tritone paradox: An influence of language on music perception*, Music Percept. 8 (1990), 335–347.

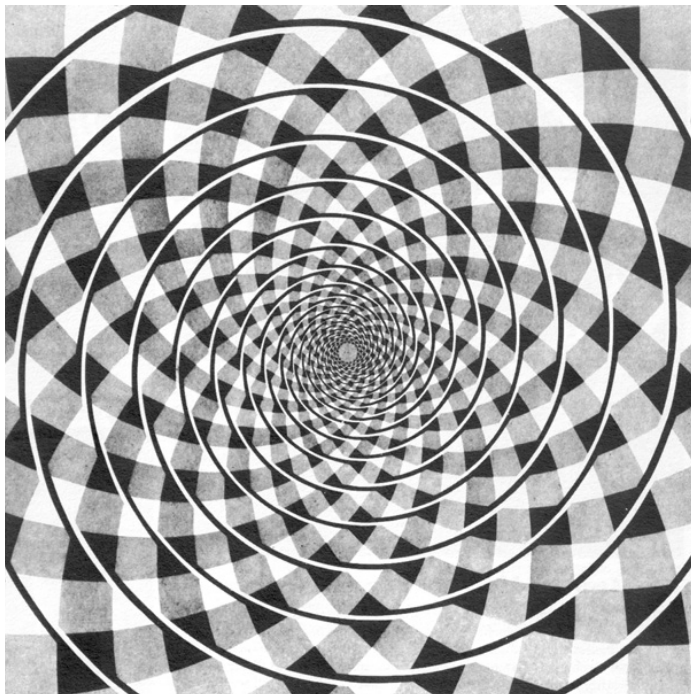
R. N. Shepard, *Circularity in judgments of relative pitch*, J. Acoust. Soc. Am. 46 (1960), 2346–2353.

Further listening: (See Appendix R)

Auditory demonstrations CD (Houtsma, Rossing and Wagenaars), track 52 is a demonstration of Shepard's scale, followed by an analogous continuously varying tone devised by Jean-Claude Risset.

CHAPTER 5

Scales and temperaments: the fivefold way



"A perfect fourth? cries Tom. Whoe'er gave birth
To such a riddle, should stick or fiddle
On his numbskull ring until he sing
A scale of perfect fourths from end to end.
Was ever such a noddy? Why, almost everybody
Knows that not e'en one thing perfect is on earth—
How then can we expect to find a perfect fourth?"

(Musical World, 1863)[1]

5.1. Introduction

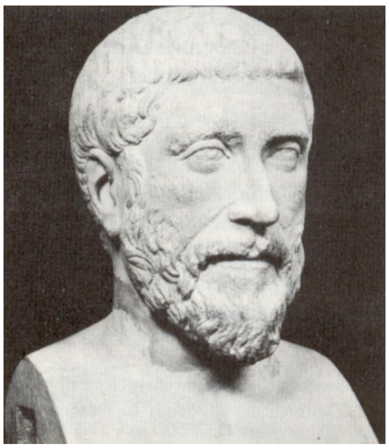
We saw in the last chapter that for notes played on conventional instruments, where partials occur at integer multiples of the fundamental frequency, intervals corresponding to frequency ratios expressable as a ratio of small integers are favored as consonant. In this chapter, we investigate how this gives rise to the scales and temperaments found in the history of western music.

[1]Quoted in Nicolas Slonimsky's *Book of Musical Anecdotes*, reprinted by Schirmer, 1998, p. 299. The picture comes from J. Frazer, *A new visual illusion of direction*, British Journal of Psychology, 1908. And yes, check it out, they *are* concentric circles, not a spiral.

Scales based around the octave are categorized by Barbour [**6**] into five broad groups: Pythagorean, just, meantone, equal, and irregular systems. The title of this chapter refers to this fivefold classification, and to the interval of a perfect fifth as the starting point for the development of scales. We shall try to indicate where these five types of scales come from.

In Chapter 6 we shall discuss further developments in the theory of scales and temperaments, and in particular, we shall study some scales which is not based around the interval of an octave. These are the Bohlen–Pierce scale, and the scales of Wendy Carlos.

5.2. Pythagorean scale



Pythagoras

As we saw in Section 4.2, Pythagoras discovered that the interval of a perfect fifth, corresponding to a frequency ratio of 3:2, is particularly consonant. He concluded from this that a convincing scale could be constructed just by using the ratios 2:1 and 3:2. Greek music scales of the Pythagorean school are built using only these intervals, although other ratios of small integers played a role in classical Greek scales.

So for example, if we use the ratio 3:2 twice, we obtain an interval with ratio 9:4, which is a little over an octave. Reducing by an octave means halving this ratio to give 9:8. Using the ratio 3:2 again will then bring us to 27:16, and so on.

What we now refer to as the Pythagorean scale is the one obtained by tuning a sequence of fifths

fa–do–so–re–la–mi–ti.

This gives the following table of frequency ratios for a major scale:[2]

note	do	re	mi	fa	so	la	ti	do
ratio	1:1	9:8	81:64	4:3	3:2	27:16	243:128	2:1

In this system, the two intervals between successive notes are a *major tone* of 9:8 and a *minor semitone* of 256:243 or 2^8:3^5. The semitone is not quite half of a tone in this system: two minor semitones give a frequency ratio of 2^{16}:3^{10} rather than 9:8. The Pythagoreans noticed that these were almost equal:

$$2^{16}/3^{10} = 1.10985715\ldots$$

[2]A Pythagorean minor scale can be constructed using ratios 32:27 for the minor third, 128:81 for the minor sixth and 16:9 for the minor seventh.

$$9/8 = 1.125$$

In other words, the Pythagorean system is based on the fact that

$$2^{19} \approx 3^{12}, \quad \text{or} \quad 524288 \approx 531441,$$

so that going up 12 fifths and then down 7 octaves brings you back to almost exactly where you started. The fact that this is not quite so gives rise to the *Pythagorean comma* or *ditonic comma*, namely the frequency ratio

$$3^{12}/2^{19} = 1.013643265\ldots$$

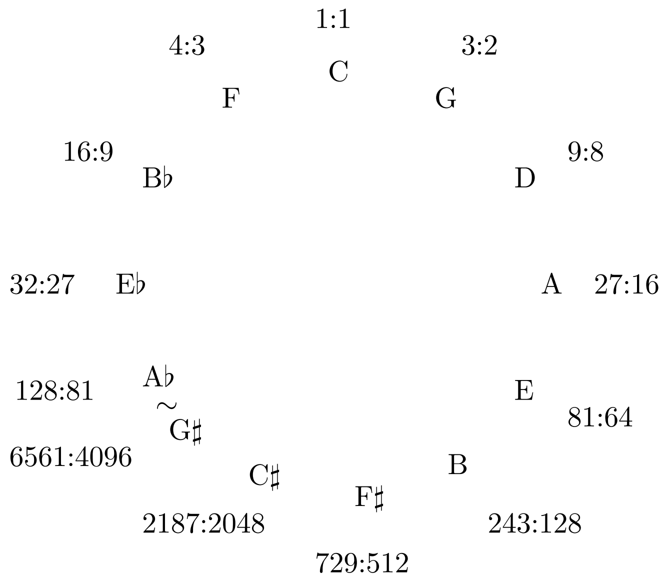
or just slightly more than one nineth of a whole tone.[3]

It seems likely that the Pythagoreans thought of musical intervals as involving the process of continued subtraction or *antanairesis*, which later formed the basis of Euclid's algorithm for finding the greatest common divisor of two integers (if you don't remember how Euclid's algorithm goes, it is described in Lemma 9.7.1). A 2:1 octave minus a 3:2 perfect fifth is a 4:3 perfect fourth. A perfect fifth minus a perfect fourth is a 9:8 Pythagorean whole tone. A perfect fourth minus two whole tones is a 256:243 Pythagorean minor semitone. It was called a *diesis* (difference), and was later referred to as a *limma* (remnant). A tone minus a diesis is a 2187:2048 Pythagorean major semitone, called an *apotomē*. An apotomē minus a diesis is a 531441:524288 Pythagorean comma.

5.3. The cycle of fifths

The Pythagorean tuning system can be extended to a twelve tone scale by tuning perfect fifths and octaves, at ratios of 3:2 and 2:1. This corresponds to tuning a "cycle of fifths" as in the following diagram:

[3]Musical intervals are measured logarithmically, so dividing a whole tone by nine really means taking the nineth root of the ratio, see §5.4.

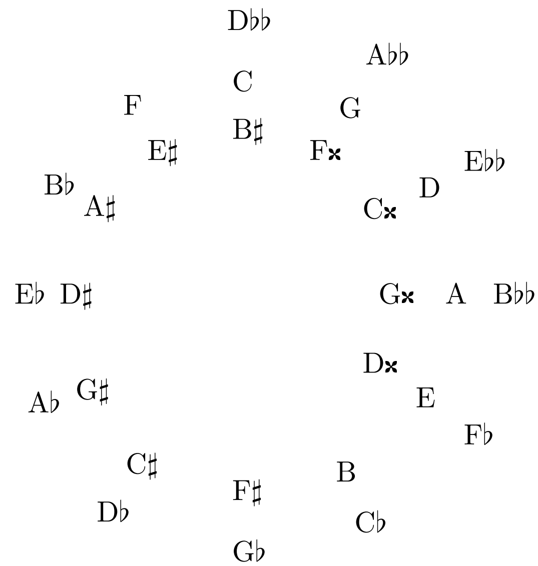


In this picture, the Pythagorean comma appears as the difference between the notes A♭ and G♯, or indeed any other *enharmonic* pair of notes:

$$\frac{6561/4096}{128/81} = \frac{3^{12}}{2^{19}} = \frac{531441}{524288}$$

In these days of equal temperament (see §5.14), we think of A♭ and G♯ as just two different names for the same note, so that there is really a *circle* of fifths. Other notes also have several names, for example the notes C and B♯, or the notes E♭♭, D and C𝄪.[4] In each case, the notes are said to be enharmonic, and in the Pythagorean system that means a difference of exactly one Pythagorean comma. So the Pythagorean system does not so much have a circle of fifths, more a sort of *spiral of fifths*.

[4]The symbol 𝄪 is used in music instead of ♯♯ to denote a double sharp.



So for example, going clockwise one complete revolution takes us from the note C to B♯, one Pythagorean comma higher. Going round the other way would take us to D♭♭, one Pythagorean comma lower. We shall see in Section §6.2 that the Pythagorean spiral never joins up. In other words, no two notes of this spiral are equal. The twelfth note is reasonably close, the 53rd is closer, and the 665th is very close indeed.

Exercises

1. What is the name of the note

(a) one Pythagorean comma lower than F,

(b) two Pythagorean commas higher than B,

(c) two Pythagorean commas lower than B?

Further listening: (See Appendix R)

Guillaume de Machaut, *Messe de Notre Dame*, Hilliard Ensemble, sung in Pythagorean intonation.

5.4. Cents

We should now explain the system of *cents*, first introduced by Alexander Ellis around 1875, for measuring frequency ratios. This is the system most often employed in the modern literature. This is a logarithmic scale[5] in which there are 1200 cents to the octave. Each whole tone on the modern

[5]See Appendix L for more on logarithms.

equal tempered scale (described below) is 200 cents, and each semitone is 100 cents. To convert from a frequency ratio of r:1 to cents, the value in cents is

$$1200 \log_2(r) = 1200 \ln(r)/\ln(2).$$

To convert an interval of n cents to a frequency ratio, the formula is

$$2^{\frac{n}{1200}} : 1.$$

In cents, the Pythagorean scale in the key of C major comes out as follows:

note	C	D	E	F	G	A	B	C
ratio	1:1	9:8	81:64	4:3	3:2	27:16	243:128	2:1
cents	0.000	203.910	407.820	498.045	701.955	905.865	1109.775	1200.000

In these notes, we shall usually give our scales in the key of C, and assign the note C a value of 0 cents. Everything else is measured in cents above the note C.

In France, rather than measuring intervals in cents, they use as their basic unit the *savart*, named after its proponent, the French physicist Félix Savart (1791–1841). In this system, a ratio of 10:1 is assigned a value of 1000 savarts. So a 2:1 octave is

$$1000 \log_{10}(2) \approx 301.030 \text{ savarts.}$$

One savart corresponds to a frequency ratio of $10^{\frac{1}{1000}}$:1, and is equal to

$$\frac{1200}{1000 \log_{10}(2)} = \frac{6}{5 \log_{10}(2)} \approx 3.98631 \text{ cents.}$$

Exercises

1. Show that to three decimal places, the Pythagorean comma is equal to 23.460 cents. What is it in savarts?

2. Convert the frequency ratios for the vibrational modes of a drum, given in §3.6, into cents above the fundamental.

3. Assigning C the value of 0 cents, what is the value of the note E♭♭ in the Pythagorean scale?

5.5. Just intonation

Just intonation refers to any tuning system that uses small, whole numbered ratios between the frequencies in a scale. This is the natural way for the ear to hear harmony, and it's the foundation of classical music theory. The dominant Western tuning system - equal temperament - is merely a 200 year old compromise that made it easier to build mechanical keyboards. Equal temperament is a lot easier to use than JI, but I find it lacks expressiveness. It sounds dead and lifeless to me. As soon as I began working microtonally, I felt like I moved from black & white into color. I found that certain combinations of intervals moved me in a

> *deep physical way. Everything became clearer for me, more visceral and expressive. The trade-off is that I had to be a lot more careful with my compositions, for while I had many more interesting consonant intervals to choose from, I also had new kinds of dissonances to avoid. Just intonation also opened me up to a greater appreciation of non-Western music, which has clearly had a large impact on my music.*

Robert Rich (synthesist)
from his FAQ at www.amoeba.com

After the octave and the fifth, the next most interesting ratio is 4:3. If we follow a perfect fifth (ratio 3:2) by the ratio 4:3, we obtain a ratio of 4:2 or 2:1, which is an octave. So 4:3 is an octave minus a perfect fifth, or a perfect fourth. So this gives us nothing new. The next new interval is given by the ratio 5:4, which is the fifth harmonic brought down two octaves.

If we continue this way, we find that the series of harmonics of a note can be used to construct scales consisting of notes that are for the most part related by small integer ratios. Given the fundamental role of the octave, it is natural to take the harmonics of a note and move them down a number of octaves to place them all in the same octave. In this case, the ratios we obtain are:

1:1 for the first, second, fourth, eighth, etc. harmonic,

3:2 for the third, sixth, twelfth, etc. harmonic,

5:4 for the fifth, tenth, etc. harmonic,

7:4 for the seventh, fourteenth, etc. harmonic,

and so on.

As we have already indicated, the ratio of 3:2 (or 6:4) is a perfect fifth. The ratio of 5:4 is a more consonant major third than the Pythagorean one, since it is a ratio of smaller integers. So we now have a *just major triad* (do–mi–so) with frequency ratios 4:5:6. Most scales in the world incorporate the major triad in some form. In western music it is regarded as the fundamental building block on which chords and scales are built. Scales in which the frequency ratio 5:4 are included were first developed by Didymus in the first century B.C. and Ptolemy in the second century A.D. The difference between the Pythagorean major third 81:64 and the Ptolemy–Didymus major third 5:4 is a ratio of 81:80. This interval is variously called the *syntonic comma*, *comma of Didymus*, *Ptolemaic comma*, or *ordinary comma*. When we use the word *comma* without further qualification, we shall always be referring to the syntonic comma.

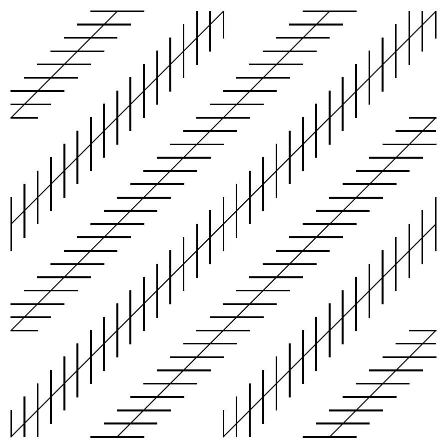
Just intonation in its most limited sense usually refers to the scales in which each of the major triads I, IV and V (i.e., C–E–G, F–A–C and G–B–D) is taken to have frequency ratios 4:5:6. Thus we obtain the following table of ratios for a just major scale:

note	do	re	mi	fa	so	la	ti	do
ratio	1:1	9:8	5:4	4:3	3:2	5:3	15:8	2:1
cents	0.000	203.910	386.314	498.045	701.955	884.359	1088.269	1200.000

The *just major third* is therefore the name for the interval (do–mi) with ratio 5:4, and the *just major sixth* is the name for the interval (do–la) with ratio 5:3. The complementary intervals (mi–do) of 8:5 and (la–do) of 6:5 are called the *just minor sixth* and the *just minor third*.

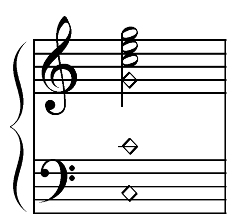
The differences between various versions of just intonation mostly involve how to fill in the remaining notes of a twelve tone scale. In order to qualify as just intonation, each of these notes must differ by a whole number of commas from the Pythagorean value. In this context, the comma may be thought of as the result of going up four perfect fifths and then down two octaves and a just major third. In some versions of just intonation, a few of the notes of the above basic scale have also been altered by a comma.

5.6. Major and minor



Parallel lines

In the last section, we saw that the basic building block of western music is the major triad, which in just intonation is built up out of the fourth, fifth and sixth notes in the harmonic series.



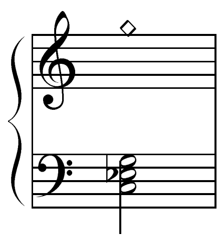
4:5:6

The minor triad is built by reversing the order of the two intervals, to obtain a chord of the form C–E♭–G. The ratios are 5:6 for the C–E♭ and 4:5

for the E♭–G. It seems futile to try to understand these as the harmonics of a common fundamental, because we would have to express the ratios as 10:12:15, making the fundamental 1/10 of the frequency of the C. It makes more sense to look at the harmonics of the notes in the triad, and to notice that all three notes have a common harmonic. Namely,

$$6 \times \mathrm{C} = 5 \times \mathrm{E}\flat = 4 \times \mathrm{G}.$$

So if we play a minor triad, if we listen carefully we can pick out this common harmonic, which is a G two octaves higher. For some subtle psychoacoustic reason, it sometimes sounds as though it's just one octave higher. It is probably the high common harmonic which causes us to associate minor chords with sadness.

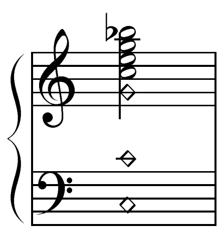


10:12:15

Another point of view regarding the minor triad is to view it as a *modification* of a major triad by slightly lowering the middle note to change the flavor. Music theory is full of modified chords, usually meaning that one of the notes in the chord has been raised or lowered by a semitone.

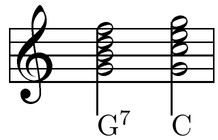
5.7. The dominant seventh

If we go as far as the seventh harmonic, we obtain a chord with ratios 4:5:6:7. This can be thought of as C–E–G–B♭, with a 7:4 B♭.



4:5:6:7

There is a closely related chord called the *dominant seventh* chord, in which the B♭ is the Pythagorean minor seventh, 16:9 higher than the C instead of 7:4. If we start this chord on G (3:2 above C) instead of C, we will obtain a chord G–B–D–F, and the F will be 4:3 above C. This chord has a strong tendency to resolve to C major, whereas the 4:5:6:7 version feels a lot more stable.



We shall have more to say about the seventh harmonic in §6.9.

Further Reading:

Martin Vogel, *Die Naturseptime* [**137**].

5.8. Commas and schismas

Recall from §5.2 that the Pythagorean comma is defined to be the difference between twelve perfect fifths and seven octaves, which gives a frequency ratio of 531441:524288, or a difference of about 23.460 cents. Recall also from §5.5 that the word *comma*, used without qualification, refers to the syntonic comma, which is a frequency ratio of 81:80. This is a difference of about 21.506 cents.

So the syntonic comma is very close in value to the Pythagorean comma, and the difference is called the *schisma*. This represents a frequency ratio of

$$\frac{531441/524288}{81/80} = \frac{32805}{32768},$$

or about 1.953 cents.

The *diaschisma*[6] is defined to be one schisma less than the comma, or a frequency ratio of 2048:2025. This may be viewed as the result of going up three octaves, and then down four perfect fifths and two just major thirds.

The *great diesis*[7] is one octave minus three just major thirds, or three syntonic commas minus a Pythagorean comma. This represents a frequency ratio of 128:125 or a difference of 41.059 cents.

The *septimal comma* is the amount by which the seventh harmonic 7:4 is flatter than the Pythagorean minor seventh 16:9. So it represents a ratio of $(16/9)(4/7) = 64/63$ or a difference of 27.264 cents.

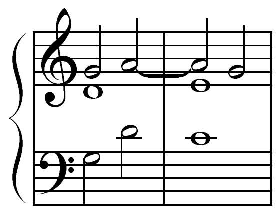
Exercises

1. Show that to three decimal places, the (syntonic) comma is equal to 21.506 cents and the schisma is equal to 1.953 cents.

[6] Historically, the Roman theorist Boethius (ca. 480–524 A.D.) attributes to Philolaus of Pythagoras' school a definition of schisma as one half of the Pythagorean comma and the diaschisma for one half of the diesis, but this does not correspond to the common modern usage of the terms.

[7] The word *diesis* in Greek means 'leak' or 'escape', and is based on the technique for playing the aulos, an ancient Greek wind instrument. To raise the pitch of a note on the aulos by a small amount, the finger on the lowest closed hole is raised slightly to allow a small amount of air to escape.

2. (G. B. Benedetti)[8] Show that if all the major thirds and sixths and the perfect fourths and fifths are taken to be just in the following harmonic progression, then the pitch will drift upwards by exactly one comma from the initial G to the final G.



$$\left(\tfrac{3}{4} \times \tfrac{3}{2} \times \tfrac{3}{5} \times \tfrac{3}{2} = \tfrac{81}{80}\right)$$

This example was given by Benedetti in 1585 as an argument against Zarlino's[9] assertion (1558) that unaccompanied singers will tend to sing in just intonation. For a further discussion of the syntonic comma in the context of classical harmony, see §5.11.

3. Here is a quote from Karlheinz Stockhausen[10] (*Lectures and interviews*, compiled by Robert Maconie, Marion Boyars publishers, London, 1989, pages 110–111):

> With the purest tones you can make the most subtle melodic gestures, much, much more refined than what the textbooks say is the smallest interval we can hear, namely the Pythagorean comma 80:81. That's not true at all. If I use sine waves, and make little glissandi instead of stepwise changes, then I can really feel that little change, going far beyond what people say about Chinese music, or in textbooks of physics or perception.
>
> But it all depends on the tone: you cannot just use any tone in an interval relationship. We have discovered a new law of relationship between the nature of the sound and the scale on which it may be composed. Harmony and melody are no longer abstract systems to be filled with any given sounds we may choose as material. There is a very subtle relationship nowadays between form and material.

(a) Find the error in this quote, and explain why it does not really matter.

(b) What is the new law of relationship to which Stockhausen is referring?

5.9. Eitz's notation

Eitz[11] devised a system of notation, used in Barbour [**6**], which is convenient for describing scales based around the octave. His method is to start with the Pythagorean definitions of the notes and then put a superscript describing how many commas to adjust by. Each comma multiplies the frequency by a factor of 81/80.

---

[8] G. B. Benedetti, *Diversarum speculationum*, Turin, 1585, page 282. The example is borrowed from Lindley and Turner-Smith [**77**], page 16.

[9] G. Zarlino, *Istitutione harmoniche*, Venice, 1558.

[10] Karlheinz Stockhausen has been much maligned in the German press in the months following September 2001. I urge anyone with a brain to go to his home page at www.stockhausen.org and find out what he really said, and what the context was. The full text of the original interview is there.

[11] Carl A. Eitz, *Das mathematisch-reine Tonsystem*, Leipzig, 1891. A similar notation was used earlier by Hauptmann and modified by Helmholtz [**54**].

As an example, the Pythagorean E, notated E^0 in this system, is 81:64 of C, while E^{-1} is decreased by a factor of 81/80 from this value, to give the just ratio of 80:64 or 5:4.

In this notation, the basic scale for just intonation is given by

$$\mathrm{C}^0 - \mathrm{D}^0 - \mathrm{E}^{-1} - \mathrm{F}^0 - \mathrm{G}^0 - \mathrm{A}^{-1} - \mathrm{B}^{-1} - \mathrm{C}^0$$

A common variant of this notation is to use subscripts rather than superscripts, so that the just major third in the key of C is E_{-1} instead of E^{-1}.

An often used graphical device for denoting just scales, which we use here in combination with Eitz's notation, is as follows. The idea is to place notes in a triangular array in such a way that moving to the right increases the note by a 3:2 perfect fifth, moving up and a little to the right increases a note by a 5:4 just major third, and moving down and a little to the right increases a note by a 6:5 just minor third. So a just major 4:5:6 triad is denoted

$$\begin{array}{ccc} & \mathrm{E}^{-1} & \\ \mathrm{C}^0 & & \mathrm{G}^0. \end{array}$$

A *just minor triad* has these intervals reversed:

$$\begin{array}{ccc} \mathrm{C}^0 & & \mathrm{G}^0 \\ & \mathrm{E}\flat^{+1} & \end{array}$$

and the notes of the just major scale form the following array:

$$\begin{array}{ccccccc} & \mathrm{A}^{-1} & & \mathrm{E}^{-1} & & \mathrm{B}^{-1} & \\ \mathrm{F}^0 & & \mathrm{C}^0 & & \mathrm{G}^0 & & \mathrm{D}^0 \end{array}$$

This method of forming an array is usually ascribed to Hugo Riemann,[12] although such arrays have been common in German music theory since the eighteenth century to denote key relationships and functional interpretation rather than frequency relationships.

It is sometimes useful to extend Eitz's notation to include other commas. Several different notations appear in the literature, and we choose to use p to denote the Pythagorean comma and z to denote the septimal comma. So for example $\mathrm{G}\sharp^{-p}$ is the same note as $\mathrm{A}\flat^0$, and the interval from C^0 to $\mathrm{B}\flat^{-z}$ is a ratio of $\frac{16}{9} \times \frac{63}{64} = \frac{7}{4}$, namely the seventh harmonic.

[12] Hugo Riemann, *Ideen zu einer 'Lehre von den Tonvorstellungen,'* Jahrbuch der Musikbibliothek Peters, 1914–1915, page 20; *Grosse Kompositionslehre*, Berlin, W. Spemann, 1902, volume 1, page 479.

Exercises

1. Show that in Eitz's notation, the example of §5.8, Exercise 2 looks like:

$$\begin{array}{cccccc}
G^{0} & D^{0} & A^{0} & E^{0} & & \\
 & & & C^{+1} & G^{+1} &
\end{array}$$

2. (a) Show that the schisma is equal to the interval between $D\flat\flat^{+1}$ and C^{0}, and the interval between C^{0} and $B\sharp^{-1}$.

(b) Show that the diaschisma is equal to the interval between C^{0} and $D\flat\flat^{+2}$.

(c) Give an example to show that a sequence of six overlapping chords in just intonation can result in a drift of one diaschisma.

(d) How many overlapping chords in just intonation are needed in order to achieve a drift of one schisma?

5.10. Examples of just scales

Using Eitz's notation, we list the examples of just intonation given in Barbour [**6**] for comparison. The dates and references have also been copied from that work.

Ramis' Monochord
(Bartolomeus Ramis de Pareja, *Musica Practica*, Bologna, 1482)

$$\begin{array}{cccccccccc}
 & & & D^{-1} & A^{-1} & E^{-1} & B^{-1} & F\sharp^{-1} & C\sharp^{-1} \\
A\flat^{0} & E\flat^{0} & B\flat^{0} & F^{0} & C^{0} & G^{0} & & &
\end{array}$$

Erlangen Monochord
(anonymous German manuscript, second half of fifteenth century)

$$\begin{array}{ccccccccc}
 & & & & & & & E^{-1} & B^{-1} \\
 & G\flat^{0} & D\flat^{0} & A\flat^{0} & E\flat^{0} & B\flat^{0} & F^{0} & C^{0} & G^{0} \\
E\flat\flat^{+1} & B\flat\flat^{+1} & & & & & & &
\end{array}$$

Erlangen Monochord, Revised
The deviations of $E\flat\flat^{+1}$ from D^{0}, and of $B\flat\flat^{+1}$ from A^{0} are equal to the schisma, as are the deviations of $G\flat^{0}$ from $F\sharp^{-1}$, of $D\flat^{0}$ from $C\sharp^{-1}$, and of $A\flat^{0}$ from $G\sharp^{-1}$. So Barbour conjectures that the Erlangen monochord was really intended as

$$\begin{array}{ccccccccc}
 & & & & E^{-1} & B^{-1} & F\sharp^{-1} & C\sharp^{-1} & G\sharp^{-1} \\
E\flat^{0} & B\flat^{0} & F^{0} & C^{0} & G^{0} & D^{0} & A^{0} & &
\end{array}$$

Fogliano's Monochord No. 1
(Lodovico Fogliano, *Musica theorica*, Venice, 1529)

F$\sharp^{-2}$ C$\sharp^{-2}$ G$\sharp^{-2}$

D^{-1} A^{-1} E^{-1} B^{-1}

B$\flat^{0}$ F^{0} C^{0} G^{0}

E$\flat^{+1}$

Fogliano's Monochord No. 2

F$\sharp^{-2}$ C$\sharp^{-2}$ G$\sharp^{-2}$

A^{-1} E^{-1} B^{-1}

F^{0} C^{0} G^{0} D^{0}

E$\flat^{+1}$ B$\flat^{+1}$

Agricola's Monochord
(Martin Agricola, *De monochordi dimensione*, in *Rudimenta musices*, Wittemberg, 1539)

F$\sharp^{-1}$ C$\sharp^{-1}$ G$\sharp^{-1}$ D$\sharp^{-1}$

B$\flat^{0}$ F^{0} C^{0} G^{0} D^{0} A^{0} E^{0} B^{0}

De Caus's Monochord
(Salomon de Caus, *Les raisons des forces mouvantes avec diverses machines*, Francfort, 1615, Book 3, Problem III)

F$\sharp^{-2}$ C$\sharp^{-2}$ G$\sharp^{-2}$ D$\sharp^{-2}$

D^{-1} A^{-1} E^{-1} B^{-1}

B$\flat^{0}$ F^{0} C^{0} G^{0}

Kepler's Monochord No. 1
(Johannes Kepler, *Harmonices mundi*, Augsburg, 1619)

E^{-1} B^{-1} F$\sharp^{-1}$ C$\sharp^{-1}$ G$\sharp^{-1}$

F^{0} C^{0} G^{0} D^{0} A^{0}

E$\flat^{+1}$ B$\flat^{+1}$

(Note: the G$\sharp^{-1}$ is incorrectly labeled G$\sharp^{+1}$ in Barbour, but his numerical value in cents is correct)

Kepler's Monochord No. 2

E^{-1} B^{-1} F$\sharp^{-1}$ C$\sharp^{-1}$

F^{0} C^{0} G^{0} D^{0} A^{0}

A$\flat^{+1}$ E$\flat^{+1}$ B$\flat^{+1}$

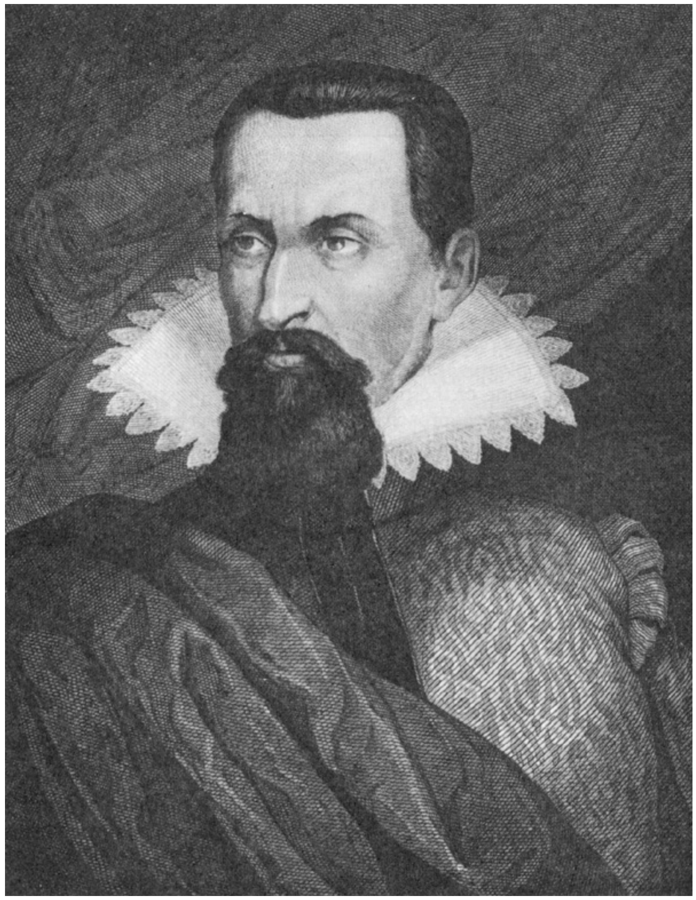
Mersenne's Spinet Tuning No. 1
(Marin Mersenne, *Harmonie universelle*, Paris, 1636–7)[13]

D^{-1} A^{-1} E^{-1} B^{-1}

B$\flat^{0}$ F^{0} C^{0} G^{0}

G$\flat^{+1}$ D$\flat^{+1}$ A$\flat^{+1}$ E$\flat^{+1}$

[13]See page 88 for a picture of Mersenne.



Johannes Kepler (1571–1630)

Mersenne's Spinet Tuning No. 2

$$\begin{array}{ccccc}
 & F\sharp^{-2} \quad C\sharp^{-2} & G\sharp^{-2} & D\sharp^{-2} & \\
 & A^{-1} & E^{-1} & B^{-1} & \\
B\flat^{0} & F^{0} & C^{0} & G^{0} & D^{0}
\end{array}$$

Mersenne's Lute Tuning No. 1

$$\begin{array}{ccccc}
 & D^{-1} & A^{-1} & E^{-1} & B^{-1} \\
 & F^{0} & C^{0} & G^{0} & \\
G\flat^{+1} & D\flat^{+1} & A\flat^{+1} & E\flat^{+1} & B\flat^{+1}
\end{array}$$

Mersenne's Lute Tuning No. 2

$$\begin{array}{ccccc}
 & A^{-1} & E^{-1} & B^{-1} & \\
 & F^{0} & C^{0} & G^{0} & D^{0} \\
G\flat^{+1} & D\flat^{+1} & A\flat^{+1} & E\flat^{+1} & B\flat^{+1}
\end{array}$$

Friedrich Wilhelm Marpurg
(1718–1795)

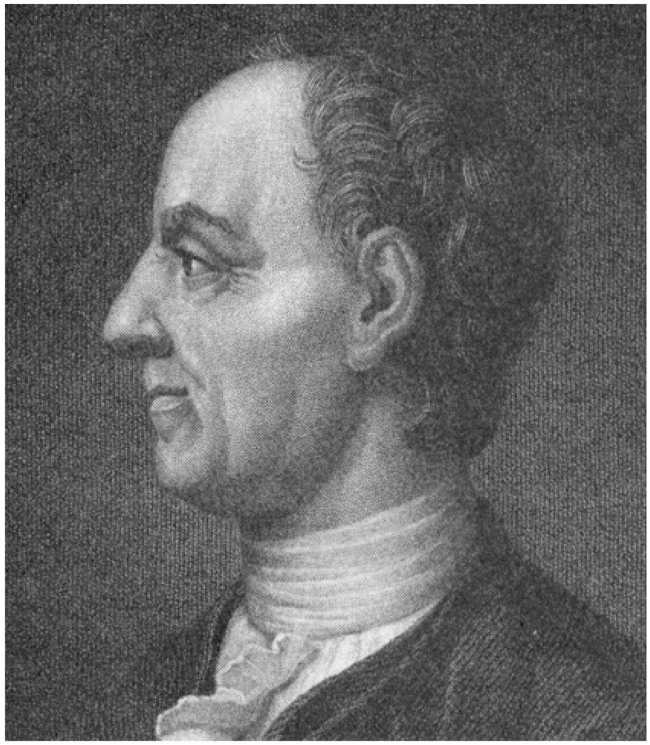
Marpurg's Monochord No. 1
(Friedrich Wilhelm Marpurg, *Versuch über die musikalische Temperatur*, Breslau, 1776)

$$
\begin{array}{ccccccc}
 & & \mathrm{C}\sharp^{-2} & & \mathrm{G}\sharp^{-2} & & \\
 & \mathrm{A}^{-1} & & \mathrm{E}^{-1} & & \mathrm{B}^{-1} & \mathrm{F}\sharp^{-1} \\
\mathrm{F}^{0} & & \mathrm{C}^{0} & & \mathrm{G}^{0} & & \mathrm{D}^{0} \\
 & & & \mathrm{E}\flat^{+1} & & \mathrm{B}\flat^{+1} &
\end{array}
$$

[Marpurg's Monochord No. 2 is the same as Kepler's monochord]

Marpurg's Monochord No. 3

$$
\begin{array}{cccccccc}
 & & & \mathrm{C}\sharp^{-2} & & \mathrm{G}\sharp^{-2} & & \\
 & & & & \mathrm{E}^{-1} & & \mathrm{B}^{-1} & \mathrm{F}\sharp^{-1} \\
\mathrm{B}\flat^{0} & \mathrm{F}^{0} & \mathrm{C}^{0} & & \mathrm{G}^{0} & & \mathrm{D}^{0} & \mathrm{A}^{0} \\
 & & & & \mathrm{E}\flat^{+1} & & &
\end{array}
$$



Leonhard Euler
(1707–1783)

Marpurg's Monochord No. 4

$$
\begin{array}{ccccccccc}
 & F\sharp^{-2} & & C\sharp^{-2} & & G\sharp^{-2} & & & \\
D^{-1} & & A^{-1} & & E^{-1} & & B^{-1} & & \\
 & F^{0} & & C^{0} & & G^{0} & & & \\
 & & & & E\flat^{+1} & & B\flat^{+1} & &
\end{array}
$$

Malcolm's Monochord
(Alexander Malcolm, *A Treatise of Musick*, Edinburgh, 1721)

$$
\begin{array}{ccccccccc}
 & & & A^{-1} & & E^{-1} & & B^{-1} & F\sharp^{-1} \\
B\flat^{0} & & F^{0} & & C^{0} & & G^{0} & & D^{0} \\
 & D\flat^{+1} & & A\flat^{+1} & & E\flat^{+1} & & &
\end{array}
$$

Euler's Monochord
(Leonhard Euler, *Tentamen novæ theoriæ musicæ*, St. Petersburg, 1739)

$$
\begin{array}{ccccccccc}
 & & C\sharp^{-2} & & G\sharp^{-2} & & D\sharp^{-2} & & A\sharp^{-2} \\
 & A^{-1} & & E^{-1} & & B^{-1} & & F\sharp^{-1} & \\
F^{0} & & C^{0} & & G^{0} & & D^{0} & &
\end{array}
$$

Montvallon's Monochord
(André Barrigue de Montvallon, *Nouveau système de musique sur les intervalles des tons et sur les proportions des accords*, Aix, 1742)

$$\begin{array}{ccccccc} & \mathrm{A}^{-1} & \mathrm{E}^{-1} & \mathrm{B}^{-1} & \mathrm{F}\sharp^{-1} & \mathrm{C}\sharp^{-1} & \mathrm{G}\sharp^{-1} \\ \mathrm{B}\flat^{0} \quad \mathrm{F}^{0} & \mathrm{C}^{0} & \mathrm{G}^{0} & \mathrm{D}^{0} & & & \\ & \mathrm{E}\flat^{+1} & & & & & \end{array}$$

Romieu's Monochord
(Jean Baptiste Romieu, *Mémoire théorique & pratique sur les systèmes tempérés de musique*, Mémoires de l'académie royale des sciences, 1758)

$$\begin{array}{cccccc} & & \mathrm{C}\sharp^{-2} & \mathrm{G}\sharp^{-2} & & \\ & \mathrm{A}^{-1} & \mathrm{E}^{-1} & \mathrm{B}^{-1} & \mathrm{F}\sharp^{-1} & \\ \mathrm{B}\flat^{0} & \mathrm{F}^{0} & \mathrm{C}^{0} & \mathrm{G}^{0} & \mathrm{D}^{0} & \\ & & \mathrm{E}\flat^{+1} & & & \end{array}$$

Kirnberger I
(Johann Phillip Kirnberger, *Construction der gleichschwebenden Temperatur*, Berlin, 1764)

$$\begin{array}{cccccccc} & & & & \mathrm{A}^{-1} & \mathrm{E}^{-1} & \mathrm{B}^{-1} & \mathrm{F}\sharp^{-1} \\ \mathrm{D}\flat^{0} & \mathrm{A}\flat^{0} & \mathrm{E}\flat^{0} & \mathrm{B}\flat^{0} & \mathrm{F}^{0} & \mathrm{C}^{0} & \mathrm{G}^{0} & \mathrm{D}^{0} \end{array}$$

Rousseau's Monochord
(Jean Jacques Rousseau, *Dictionnaire de musique*, Paris, 1768)

$$\begin{array}{cccc} \mathrm{F}\sharp^{-2} & \mathrm{C}\sharp^{-2} & & \\ \mathrm{A}^{-1} & \mathrm{E}^{-1} & \mathrm{B}^{-1} & \\ \mathrm{F}^{0} & \mathrm{C}^{0} & \mathrm{G}^{0} & \mathrm{D}^{0} \\ \mathrm{A}\flat^{+1} & \mathrm{E}\flat^{+1} & \mathrm{B}\flat^{+1} & \end{array}$$

We shall return to the discussion of just intonation in §6.1, where we consider scales built using primes higher than 5. In §6.8, we look at a way of systematizing the discussion by using lattices, and we interpret the above scales as periodicity blocks.

Exercises

1. Choose several of the just scales described in this section, and write down the values of the notes

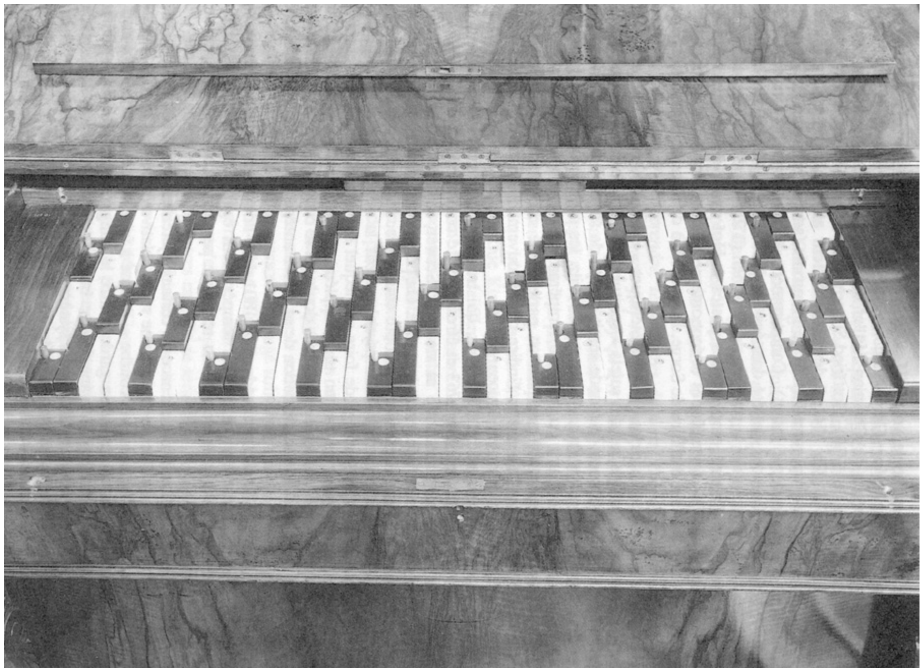
(i) in cents, and

(ii) as frequencies, giving the answers as multiples of the frequency for C.

2. Show that the Pythagorean scale with perfect fifths

$$\mathrm{G}\flat^{0} - \mathrm{D}\flat^{0} - \mathrm{A}\flat^{0} - \mathrm{E}\flat^{0} - \mathrm{B}\flat^{0} - \mathrm{F}^{0} - \mathrm{C}^{0} - \mathrm{G}^{0} - \mathrm{D}^{0} - \mathrm{A}^{0} - \mathrm{E}^{0} - \mathrm{B}^{0},$$

gives good approximations to just major triads on D, A and E, in the form $\mathrm{D}^{0} - \mathrm{G}\flat^{0} - \mathrm{A}^{0}$, $\mathrm{A}^{0} - \mathrm{D}\flat^{0} - \mathrm{E}^{0}$ and $\mathrm{E}^{0} - \mathrm{A}\flat^{0} - \mathrm{B}^{0}$. How far from just are the thirds of these chords (in cents)?



Colin Brown's voice harmonium
(Science Museum, London)

3. The voice harmonium of Colin Brown (1875) is shown above. A plan of a little more than one octave of the keyboard is shown at the top of the next page. Diagonal rows of black keys and white keys alternate, and each black key has a red peg sticking out of its upper left corner, represented by a small circle in the plan. The purpose of this keyboard is to be able to play in a number of different keys in just intonation. Locate examples of the following on this keyboard:

(i) A just major triad.

(ii) A just minor triad.

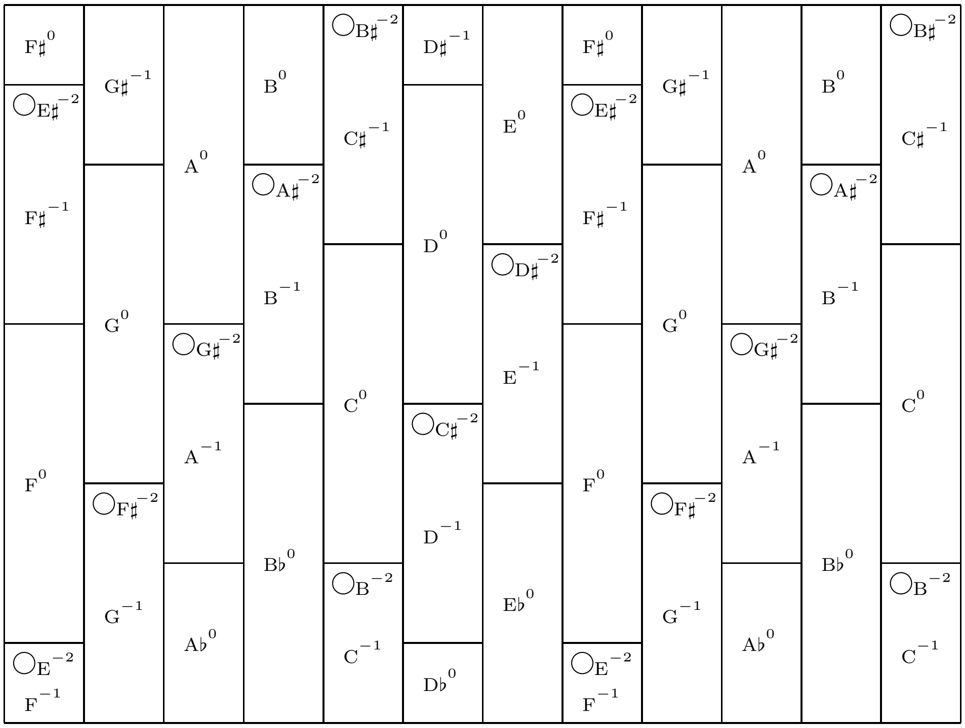
(iii) A just major scale.

(iv) Two notes differing by a syntonic comma.

(v) Two notes differing by a schisma.

(vi) Two notes differing by a diesis.

(vii) Two notes differing by an apotomē.



Keyboard diagram for Colin Brown's voice harmonium

5.11. Classical harmony

The main problem with the just major scale introduced in §5.5 is that certain harmonic progressions which form the basis of classical harmony don't quite work. This is because certain notes in the major scale are being given two different just interpretations, and switching from one to the other is a part of the progression. In this section, we discuss the progressions which form the basis for classical harmony,[14] and find where the problems lie.

We begin with the names of the triads. An upper case roman numeral denotes a major chord based on the given scale degree, whereas a lower case roman numeral denotes a minor chord. So for example the major chords I, IV and V form the basis for the just major scale in §5.5, namely C^0– E^{-1}– G^0, F^0– A^{-1}– $C^0$and G^0– B^{-1}– D^0 in the key of C major. The triads A^{-1}– C^0– E^{-1} and E^{-1}– G^0– B^{-1} are the minor triads vi and iii. The problem comes from the triad on the second note of the scale, D^0– F^0– A^{-1}. If we alter the D^0 to a D^{-1}, this is a just minor triad, which we would then call ii.

Classical harmony makes use of ii as a minor triad, so maybe we should have used D^{-1} instead of D^0 in our just major scale. But then the triad V becomes G^0– B^{-1}– D^{-1}, which doesn't quite work. We shall see that there

[14]The phrase "classical harmony" here is used in its widest sense, to include not only classical, romantic and baroque music, but also most of the rock, jazz and folk music of western culture.

is no choice of just major scale which makes all the required triads work. To understand this, we discuss classical harmonic practice.

We begin at the end. Most music in the western world imparts a sense of finality through the sequence V–I, or variations of it (V^7–I, vii^0–I).[15] It is not fully understood why V–I imparts such a feeling of finality, but it cannot be denied that it does. A great deal of music just consists of alternate triads V and I.

The progression V–I can stand on its own, or it can be approached in a number of ways. A sequence of fifths forms the basis for the commonest method, so that we can extend to ii–V–I, then to vi–ii–V–I, and even further to iii–vi–ii–V–I, each of these being less common than the previous ones. Here is a chart of the most common harmonic progressions in the music of the western world, in the major mode:

$$[\text{iii}] \to [\text{vi}] \to \begin{bmatrix} \text{IV} \\ \downarrow \\ \text{ii} \end{bmatrix} \begin{matrix} \searrow \\ \nearrow \end{matrix} \begin{bmatrix} \text{vii}^0 \\ \\ \text{V} \end{bmatrix} \begin{matrix} \searrow \\ \nearrow \end{matrix} \text{I}$$

and then either end the piece, or go back from I to any previous triad. Common exceptions are to jump from iii to IV, from IV to I and from V to vi.

Now take a typical progression from the above chart, such as

$$\text{I–vi–ii–V–I},$$

and let us try to interpret this in just intonation. Let us stipulate one simple rule, namely that if a note on the diatonic scale appears in two adjacent triads, it should be given the same just interpretation. So if I is $\text{C}^{0}\!-\text{E}^{-1}\!-\text{G}^{0}$ then vi must be interpreted as $\text{A}^{-1}\!-\text{C}^{0}\!-\text{E}^{-1}$, since the C and E are in common between the two triads. This means that the ii should be interpreted as $\text{D}^{-1}\!-\text{F}^{0}\!-\text{A}^{-1}$, with the A in common with vi. Then V needs to be interpreted as $\text{G}^{-1}\!-\text{B}^{-2}\!-\text{D}^{-1}$ because it has D in commion with ii. Finally, the I at the end is forced to be interpreted as $\text{C}^{-1}\!-\text{E}^{-2}\!-\text{G}^{-1}$ since it has G in common with V. We are now one syntonic comma lower than where we started.

To put the same problem in terms of ratios, in the second triad the A is $\frac{5}{3}$ of the frequency of the C, then in the third triad, D is $\frac{2}{3}$ of the frequency of A. In the fourth triad, G is $\frac{4}{3}$ of the frequency of D, and finally in the last triad, C is $\frac{2}{3}$ of the frequency of G. This means that the final C is

$$\frac{5}{3} \times \frac{2}{3} \times \frac{4}{3} \times \frac{2}{3} = \frac{80}{81}$$

of the frequency of the initial one.

A similar drift downward through a syntonic comma occurs in the sequences

$$\text{I–IV–ii–V–I}$$

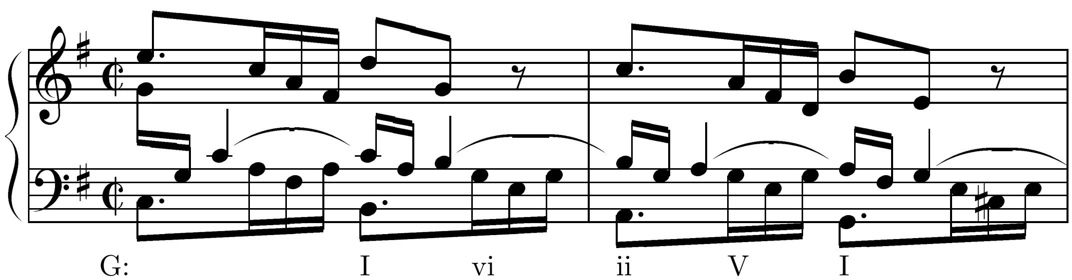
$$\text{I–iii–vi–ii–V–I}$$

[15] The superscript zero in the notation vii^0 denotes a *diminished triad* with two minor thirds as the intervals. It has nothing to do with the Eitz comma notation.

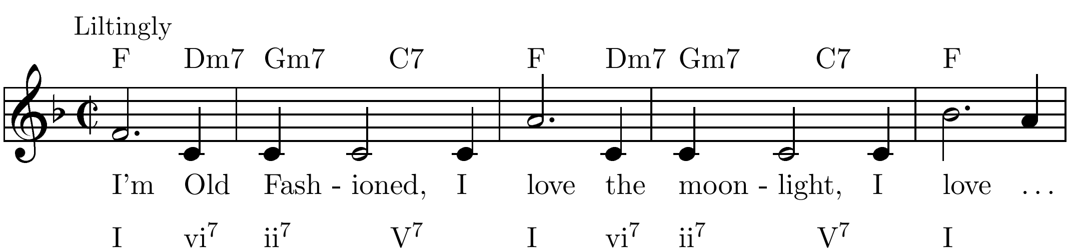
and so on. Here are some actual musical examples, chosen pretty much at random.

(i) W. A. Mozart, *Sonata* (K. 333), third movement, beginning.

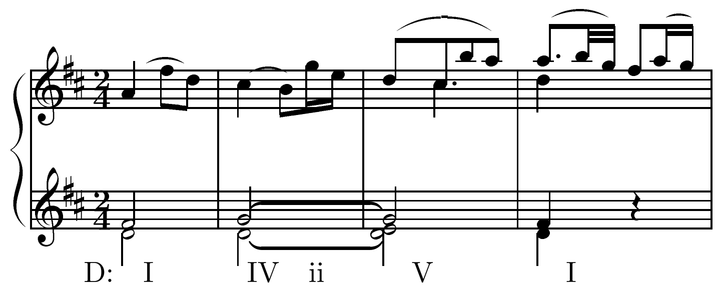
(ii) J. S. Bach, *Partita no. 5, Gigue,* bars 23–24.



(iii) *I'm Old Fashioned* (1942).
Music by Jerome Kern, words by Johnny Mercer.

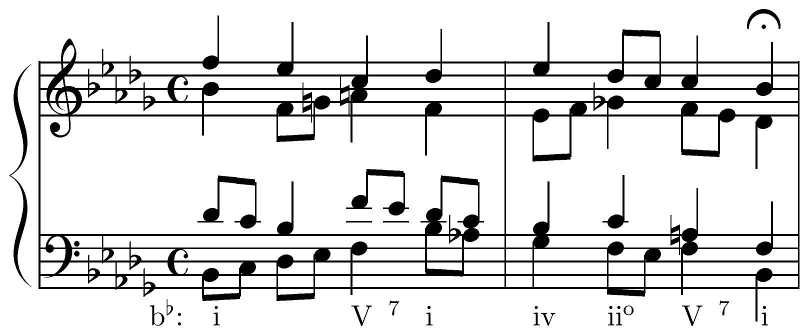


(iv) W. A. Mozart, *Fantasie* (Kv. 397), bars 55–59.



And a minor example:

(v) J. S. Bach, *Jesu, der du meine Seele.*



The meantone scale, which we shall discuss in the next section, solves the problem of the syntonic comma by deviating slightly from the just values of notes in such a way that the comma is spread equally between the four perfect fifths involved, shaving one quarter of a comma from each of them.

Harry Partch discusses th issue of the syntonic comma at length, towards the end of chapter 11 of [**98**]. He arrives at a different conclusion from the one adopted historically, namely that the progressions above sound fine, played in just intonation in such a way that the second note on the scale is played (in C major) as D^{-1} in ii, and as D^{0} in V. This means that these two versions of the "same" note are played in consecutive triads, but the sense of the harmonic progression is not lost.

5.12. Meantone scale

A *tempered scale* is a scale in which adjustments are made to the Pythagorean or just scale in order to spread around the problem caused by wishing to regard two notes differing by various commas as the same note, as in the example of §5.8, Exercise 2, and the discussion in §5.11.

The meantone scales are the tempered scales formed by making adjustments of a fraction of a (syntonic) comma to the fifths in order to make the major thirds better.

The commonest variant of the meantone scale, sometimes referred to as the classical meantone scale, or quarter-comma meantone scale, is the one in which the major thirds are made in the ratio 5:4 and then the remaining notes are interpolated as equally as possible. So C–D–E are in the ratios $1 : \sqrt{5}/2 : 5/4$, as are F–G–A and G–A–B. This leaves two semitones to decide, and they are made equal. Five tones of ratio $\sqrt{5}/2 : 1$ and two semitones make an octave 2:1, so the ratio for the semitone is

$$\sqrt{2/\left(\sqrt{5}/2\right)^5} : 1 = 8 : 5^{\frac{5}{4}}.$$

The table of ratios is therefore as follows:

note	do	re	mi	fa	so	la	ti	do
ratio	1:1	$\sqrt{5}$:2	5:4	2:$5^{\frac{1}{4}}$	$5^{\frac{1}{4}}$:1	$5^{\frac{3}{4}}$:2	$5^{\frac{5}{4}}$:4	2:1
cents	0.000	193.157	386.314	503.422	696.579	889.735	1082.892	1200.000

The fifths in this scale are no longer perfect.

Another, more enlightening way to describe the classical meantone scale is to temper each fifth by making it narrower than the Pythagorean value by exactly one quarter of a comma, in order for the major thirds to come out right. So working from C, the G is one quarter comma flat from its Pythagorean value, the D is one half comma flat, the A is three quarters of a comma flat, and finally, E is one comma flat from a Pythagorean major third, which makes it exactly equal to the just major third. Continuing in the same direction, this makes the B five quarters of a comma flatter than its Pythagorean value. Correspondingly, the F should be made one quarter comma sharper than the Pythagorean fourth.

Thus in Eitz's notation, the classical meantone scale can be written as

$$\mathrm{C}^0 - \mathrm{D}^{-\frac{1}{2}} - \mathrm{E}^{-1} - \mathrm{F}^{+\frac{1}{4}} - \mathrm{G}^{-\frac{1}{4}} - \mathrm{A}^{-\frac{3}{4}} - \mathrm{B}^{-\frac{5}{4}} - \mathrm{C}^0$$

Writing these notes in the usual array notation, we obtain

$$\begin{array}{ccccccccc} & \mathrm{E}^{-1} & & \mathrm{B}^{-\frac{5}{4}} & & & & & \\ \mathrm{C}^0 & & \mathrm{G}^{-\frac{1}{4}} & & \mathrm{D}^{-\frac{1}{2}} & & \mathrm{A}^{-\frac{3}{4}} & & \mathrm{E}^{-1} \\ & & & & & \mathrm{F}^{+\frac{1}{4}} & & \mathrm{C}^0 & \end{array}$$

The meantone scale can be completed by filling in the remaining notes of a twelve (or more) tone scale according to the same principles. The only question is how far to go in each direction with the quarter comma tempered fifths. Some examples, again taken from Barbour [**6**] follow.

Aaron's Meantone Temperament
(Pietro Aaron, *Toscanello in musica*, Venice, 1523)

$$C^0\ C\sharp^{-\frac{7}{4}}\ D^{-\frac{1}{2}}\ E\flat^{+\frac{3}{4}}\ E^{-1}\ F^{+\frac{1}{4}}\ F\sharp^{-\frac{3}{2}}\ G^{-\frac{1}{4}}\ A\flat^{+1}\ A^{-\frac{3}{4}}\ B\flat^{+\frac{1}{2}}\ B^{-\frac{5}{4}}\ C^0$$

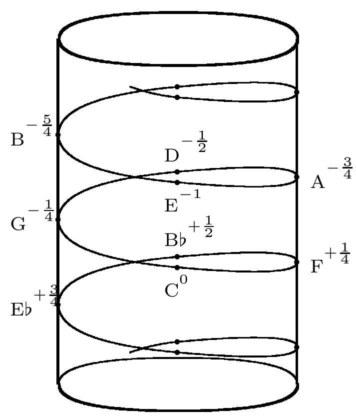
Gibelius' Monochord for Meantone Temperament
(Otto Gibelius, *Propositiones mathematico-musicæ*, Münden, 1666)
is the same, but with two extra notes

$$C^0\ C\sharp^{-\frac{7}{4}}\ D^{-\frac{1}{2}}\ D\sharp^{-\frac{9}{4}}\ E\flat^{+\frac{3}{4}}\ E^{-1}\ F^{+\frac{1}{4}}\ F\sharp^{-\frac{3}{2}}\ G^{-\frac{1}{4}}\ G\sharp^{-2}\ A\flat^{+1}\ A^{-\frac{3}{4}}\ B\flat^{+\frac{1}{2}}\ B^{-\frac{5}{4}}\ C^0$$

These meantone scales are represented in array notation as follows:

$$\begin{array}{ccccccccccc}
 & & & (G\sharp^{-2}) & & (D\sharp^{-\frac{9}{4}}) & & & & & \\
 & & E^{-1} & & B^{-\frac{5}{4}} & & F\sharp^{-\frac{3}{2}} & & C\sharp^{-\frac{7}{4}} & & (G\sharp^{-2}) \\
 & C^0 & & G^{-\frac{1}{4}} & & D^{-\frac{1}{2}} & & A^{-\frac{3}{4}} & & E^{-1} & \\
A\flat^{+1} & & E\flat^{+\frac{3}{4}} & & B\flat^{+\frac{1}{2}} & & F^{+\frac{1}{4}} & & C^0 & &
\end{array}$$

where the right hand edge is thought of as equal to the left hand edge. Thus the notes can be thought of as lying on a cylinder, with four quarter-comma adjustments taking us once round the cylinder.



So the syntonic comma has been taken care of, and modulations can be made to a reasonable number of keys. The Pythagorean comma has not been taken care of, so that modulation around an entire circle of fifths is still not feasible. Indeed, the difference between the enharmonic notes $A\flat^{+1}$ and $G\sharp^{-2}$ is three syntonic commas minus a Pythagorean comma, which is a ratio of 128:125, or a difference of 41.059 cents. This interval, called the great diesis is nearly half a semitone, and is very noticeable to the ear. The imperfect fifth between C♯ and A♭ (or wherever else it may happen to be placed) in the meantone scale is sometimes referred to as the *wolf*[16] interval of the scale. We shall see in §6.5 that one way of dealing with the wolf fifth is to use thirty-one tones to an octave instead of twelve.

[16]This has nothing to do with the "wolf" notes on a stringed instrument such as the cello, which has to do with the sympathetic resonance of the body of the instrument.

Although what we have described is the commonest form of meantone scale, there are others formed by taking different divisions of the comma. In general, the α-comma meantone temperament refers to the following temperament:

$$\begin{array}{ccccccccccc}
 & E^{-4\alpha} & & B^{-5\alpha} & & F\sharp^{-6\alpha} & & C\sharp^{-7\alpha} & & G\sharp^{-8\alpha} & \\
C^{0} & & G^{-\alpha} & & D^{-2\alpha} & & A^{-3\alpha} & & E^{-4\alpha} & & \\
 & E\flat^{+3\alpha} & & B\flat^{+2\alpha} & & F^{+\alpha} & & C^{0} & & &
\end{array}$$

Without any qualification, the phrase "meantone temperament" refers to the case $\alpha = \frac{1}{4}$. The following names are associated with various values of α:

0	Pythagoras	
$\frac{1}{6}$	Silbermann	Sorge, *Gespräch zwischen einem Musico theoretico und einem Studioso musices*, (Lobenstein, 1748), p. 20
$\frac{1}{5}$	Abraham Verheijen,	Simon Stevin, *Van de Spiegeling der Singconst*, c. 1600
	Lemme Rossi	*Sistema musico*, Perugia, 1666, p. 58
$\frac{2}{9}$	Lemme Rossi	*Sistema musico*, Perugia, 1666, p. 64
$\frac{1}{4}$	Aaron/Gibelius/Zarlino/...	
$\frac{2}{7}$	Gioseffo Zarlino	*Istitutioni armoniche*, Venice, 1558
$\frac{1}{3}$	Francisco de Salinas	*De musica libri VII*, Salamanca, 1577

So for example, Zarlino's $\frac{2}{7}$ comma meantone temperament is as follows:

$$\begin{array}{ccccccccccc}
 & E^{-\frac{8}{7}} & & B^{-\frac{10}{7}} & & F\sharp^{-\frac{12}{7}} & & C\sharp^{-2} & & G\sharp^{-\frac{16}{7}} & \\
C^{0} & & G^{-\frac{2}{7}} & & D^{-\frac{4}{7}} & & A^{-\frac{6}{7}} & & E^{-\frac{8}{7}} & & \\
 & E\flat^{+\frac{6}{7}} & & B\flat^{+\frac{4}{7}} & & F^{+\frac{2}{7}} & & C^{0} & & &
\end{array}$$

The value $\alpha = 0$ gives Pythagorean intonation, and a value close to $\alpha = \frac{1}{11}$ gives twelve tone equal temperament (see §5.14), so these can (at a pinch) be thought of as extreme forms of meantone. There is a diagram in Appendix J on page 367 which illustrates various meantone scales, and the extent to which the thirds and fifths deviate from their just values.

A useful way of thinking of meantone temperaments is that in order to name a meantone temperament, it is sufficient to name the size of the fifth. We have chosen to name this size as a narrowing of the perfect fifth by α commas. Knowing the size of the fifth, all other intervals are obtained by taking multiples of this size and reducing by octaves. So we say that the fifth *generates* a meantone temperament. In any meantone temperament, every key sounds just like every other key, until the wolf is reached.

Exercises

1. Show that the 1/3 comma meantone scale of Salinas gives pure minor thirds. Calculate the size of the wolf fifth.

2. What fraction of a comma should we use for a meantone system in order to minimize the mean square error of the fifth, the major third and the minor third from their just values?

3. Go to the web site

midiworld.com/mw_byrd.htm

and listen to some of John Sankey's MIDI files of keyboard music by William Byrd, sequenced in quarter comma meantone.

4. Charles Lucy is fond of a tuning system which he attributes to John Harrison (1693–1776) in which the fifths are tuned to a ratio of $2^{\frac{1}{2}+\frac{1}{4\pi}}:1$ and the major thirds $2^{\frac{1}{\pi}}:1$. Show that this can be considered as a meantone scale in which the fifths are tempered by about $\frac{3}{10}$ of a comma. Charles Lucy's web site can be found at

lehua.ilhawaii.net/∼lucy/index.html

5. In the meantone scale, the octave is taken to be perfect. Investigate the scale obtained by stretching the octave by $\frac{1}{6}$ of a comma, and shrinking the fifth by $\frac{1}{6}$ of a comma. How many cents away from just are the major third and minor third in this scale? Calculate the values in cents for notes of the major scale in this temperament.

Further listening: (See Appendix R)

The Katahn/Foote recording, *Six degrees of tonality* contains tracks comparing Mozart's *Fantasie* Kv. 397 in equal temperament, meantone, and an irregular temperament of Prelleur.

Edward Parmentier, *Seventeenth Century French Harpsichord Music*, recorded in $\frac{1}{3}$ comma meantone temperament.

Aldert Winkelman, *Works by Mattheson, Couperin and others.* This recording includes pieces by Louis Couperin and Gottlieb Muffat played on a spinet tuned in quarter comma meantone temperament.

Organs tuned in quarter comma meantone temperament are being built even today. The C. B. Fisk organ at Wellesley College, Massachusetts, USA is tuned in quarter comma meantone temperament. See

www.wellesley.edu/Music/organ.html

for a more detailed description of this organ. Bernard Lagacé has recorded a CD of music of various composers on this organ.

John Brombaugh apprenticed with the American organ builders Fritz Noack and Charles Fisk between 1964 and 1967, and has built a number of organs in quarter comma meantone temperament. These include the Brombaugh organs in the Duke University Chapel, Oberlin College, Southern College, and the Haga Church in Gothenburg, Sweden.

Another example of a modern organ tuned in meantone temperament is the Hellmuth Wollf organ of Knox College Chapel in Toronto University, Canada.

5.13. Irregular temperaments

The phrases *irregular temperament*, *circulating temperament* and *well tempered scale* all refer to a twelve tone scale in which the notes of the meantone scale have been bent so that the scale works more or less in all twelve possible key signatures. This means that the notes at the extremes of the circle of fifths, near the wolf fifth, have been changed in pitch so as to distribute the wolf between several fifths. The effect is that each of these fifths is more or less acceptable.

Historically, irregular temperaments superseded or lived alongside meantone temperament (§5.12) during the seventeenth century, and were in use for at least two centuries before equal temperament (§5.14) took hold.

Evidence from the 48 Preludes and Fugues of the *Well Tempered Clavier* suggests that rather than being written for meantone temperament, Bach intended a more irregular temperament in which all keys are more or less satisfactorily in tune.[17]

A typical example of such a temperament is Werckmeister's most frequently used temperament. This is usually referred to as Werckmeister III (although Barbour [**6**] refers to it as Werckmeister's Correct Temperament No. 1),[18] which is as follows.

Werckmeister III (Correct Temperament No. 1)
(Andreas Werckmeister, *Musicalische Temperatur* Frankfort and Leipzig, 1691; reprinted by Diapason Press, Utrecht, 1986, with commentary by Rudolph Rasch)

$$\begin{array}{cccccccccc}
 & \mathrm{E}^{-\frac{3}{4}p} & & \mathrm{B}^{-\frac{3}{4}p} & & \mathrm{F}\sharp^{-1p} & & \mathrm{C}\sharp^{-1p} & & \mathrm{G}\sharp^{-1p} \\
\mathrm{C}^{0} & & \mathrm{G}^{-\frac{1}{4}p} & & \mathrm{D}^{-\frac{1}{2}p} & & \mathrm{A}^{-\frac{3}{4}p} & & \mathrm{E}^{-\frac{3}{4}p} & \\
 & \mathrm{E}\flat^{0} & & \mathrm{B}\flat^{0} & & \mathrm{F}^{0} & & \mathrm{C}^{0} & &
\end{array}$$

In this temperament, the Pythagorean comma (*not* the syntonic comma) is distributed equally on the fifths from C–G–D–A and B–F♯. We use a modified version of Eitz's notation to denote this, in which "p" is used to denote the usage of the Pythagorean comma rather than the syntonic comma. A good way to think of this is to use the approximation discussed in §5.14 which says "$p = \frac{12}{11}$," so that for example $\mathrm{E}^{-\frac{3}{4}p}$ is essentially the same as $\mathrm{E}^{-\frac{9}{11}}$. Note that $\mathrm{A}\flat^{0}$ is equal to $\mathrm{G}\sharp^{-1p}$, so the circle of fifths does join up properly in this temperament. In fact, this was the first temperament to be widely adopted which has this property.

In this and other irregular temperaments, different key signatures have different characteristic sounds, with some keys sounding direct and others

---

[17] It is a common misconception that Bach intended the *Well Tempered Clavier* to be played in equal temperament. He certainly knew of equal temperament, but did not use it by preference, and it is historically much more likely that the 48 preludes and fugues were intended for an irregular temperament of the kind discussed in this section. (It should be mentioned that there is also evidence that Bach did intend equal temperament, see Rudolf A. Rasch, *Does 'Well-tempered' mean 'Equal-tempered'?*, in Williams (ed.), Bach, Händel, Scarlatti tercentenary Essays, Cambridge University Press, Cambridge, 1985, pp. 293-310.)

[18] *Werckmeister I* usually refers to just intonation, and *Werckmeister II* to classical meantone temperament. *Werckmeister IV* and *V* are described below. There is also a temperament known as *Werckmeister VI*, or "septenarius," which is based on a division of a string into 196 equal parts. This scale gives the ratios 1:1, 196:186, 196:176, 196:165, 196:156, 4:3, 196:139, 196:131, 196:124, 196:117, 196:110, 196:104, 2:1.

more remote. This may account for the modern myth that the same holds in equal temperament.[19]

An interesting example of the use of irregular temperaments in composition is J. S. Bach's Toccata in F♯ minor (BWV 910), bars 109ff, in which essentially the same musical phrase is repeated about twenty times in succession, transposed into different keys. In equal or meantone temperament this could get monotonous, but with an irregular temperament, each phrase would impart a subtly different feeling.

The point of distributing the comma unequally between the twelve fifths is so that in the most commonly used keys, the fifth and major third are very close to just. The price to be paid is that in the more "remote" keys the tuning of the major thirds is somewhat sharp. So for example in Werckmeister III, the thirds on C and F are about four cents sharp, while the thirds on C♯ and F♯ are about 22 cents sharp. Other examples of irregular temperaments with similar intentions include the following, taken from Asselin [**2**], Barbour [**6**] and Devie [**34**].

Mersenne's Improved Meantone Temperament, No. 1
(Marin Mersenne: Cogitata physico-mathematica, Paris, 1644)

$$\begin{array}{ccccccccccc}
 & \mathrm{E}^{-1p} & & \mathrm{B}^{-\frac{5}{4}p} & & \mathrm{F}\sharp^{-\frac{3}{2}p} & & \mathrm{C}\sharp^{-\frac{7}{4}p} & & \mathrm{G}\sharp^{-2p} \\
\mathrm{C}^{0} & & \mathrm{G}^{-\frac{1}{4}p} & & \mathrm{D}^{-\frac{1}{2}p} & & \mathrm{A}^{-\frac{3}{4}p} & & \mathrm{E}^{-1p} & \\
 & \mathrm{E}\flat^{+\frac{1}{4}p} & & \mathrm{B}\flat^{+\frac{1}{4}p} & & \mathrm{F}^{+\frac{1}{4}p} & & \mathrm{C}^{0} & &
\end{array}$$

Bendeler's Temperament, No. 1
(P. Bendeler, *Organopoeia*, Frankfurt, 1690; 2nd. ed. Frankfurt & Leipzig, 1739, p. 40)

$$\begin{array}{ccccccccccc}
 & \mathrm{E}^{-\frac{2}{3}p} & & \mathrm{B}^{-\frac{2}{3}p} & & \mathrm{F}\sharp^{-p} & & \mathrm{C}\sharp^{-p} & & \mathrm{G}\sharp^{-p} \\
\mathrm{C}^{0} & & \mathrm{G}^{-\frac{1}{3}p} & & \mathrm{D}^{-\frac{2}{3}p} & & \mathrm{A}^{-\frac{2}{3}p} & & \mathrm{E}^{-\frac{2}{3}p} & \\
 & \mathrm{E}\flat^{0} & & \mathrm{B}\flat^{0} & & \mathrm{F}^{0} & & \mathrm{C}^{0} & &
\end{array}$$

[19] If this were really true, then the shift of nearly a semitone in pitch between Mozart's time and our own would have resulted in a permutation of the resulting moods, which seems to be nonsense. Actually, this argument really only applies to keyboard instruments. It is still possible in equal temperament for string and wind instruments to give different characters to different keys. For example, a note on an open string on a violin sounds different in character from a stopped string. Mozart and others have made use of this difference with a technique called *scordatura*, (Italian *scordare*, to mistune) which involves unconventional retuning of stringed instruments. A well known example is his *Sinfonia Concertante*, in which all the strings of the solo viola are tuned a semitone sharp. The orchestra plays in E♭ for a softer sound, and the solo viola plays in D for a more brilliant sound.

A more shocking example (communicated to me by Marcus Linckelmann) is Schubert's *Impromptu No. 3* for piano in G♭ major. The same piece played in G major on a modern piano has a very different feel to it. It is possible that in this case, the mechanics of the fingering are responsible.

Bendeler's Temperament, No. 2
(P. Bendeler, 1690/1739, p. 42)

$$\begin{array}{cccccccccc}
 & E^{-\frac{2}{3}p} & & B^{-\frac{2}{3}p} & & F\sharp^{-\frac{2}{3}p} & & C\sharp^{-p} & & G\sharp^{-p} \\
C^{0} & & G^{-\frac{1}{3}p} & & D^{-\frac{1}{3}p} & & A^{-\frac{2}{3}p} & & E^{-\frac{2}{3}p} & \\
 & E\flat^{0} & & B\flat^{0} & & F^{0} & & C^{0} & &
\end{array}$$

Bendeler's Temperament, No. 3
(P. Bendeler, 1690/1739, p. 42)

$$\begin{array}{cccccccccc}
 & E^{-\frac{1}{2}p} & & B^{-\frac{3}{4}p} & & F\sharp^{-\frac{3}{4}p} & & C\sharp^{-\frac{3}{4}p} & & G\sharp^{-\frac{3}{4}p} \\
C^{0} & & G^{-\frac{1}{4}p} & & D^{-\frac{1}{2}p} & & A^{-\frac{1}{2}p} & & E^{-\frac{1}{2}p} & \\
 & E\flat^{0} & & B\flat^{0} & & F^{0} & & C^{0} & &
\end{array}$$

Werckmeister III (Correct Temperament No. 1) See page 173.

Werckmeister IV (Correct Temperament No. 2)
(Andreas Werckmeister, 1691; the least satisfactory of Werckmeister's temperaments)

$$\begin{array}{cccccccccc}
 & E^{-\frac{2}{3}p} & & B^{-1p} & & F\sharp^{-1p} & & C\sharp^{-\frac{4}{3}p} & & G\sharp^{-\frac{4}{3}p} \\
C^{0} & & G^{-\frac{1}{3}p} & & D^{-\frac{1}{3}p} & & A^{-\frac{2}{3}p} & & E^{-\frac{2}{3}p} & \\
 & E\flat^{0} & & B\flat^{+\frac{1}{3}p} & & F^{0} & & C^{0} & &
\end{array}$$

Werckmeister V (Correct Temperament No. 3)
(Andreas Werckmeister, 1691)

$$\begin{array}{cccccccccc}
 & E^{-\frac{1}{2}p} & & B^{-\frac{1}{2}p} & & F\sharp^{-\frac{1}{2}p} & & C\sharp^{-\frac{3}{4}p} & & G\sharp^{-1p} \\
C^{0} & & G^{0} & & D^{0} & & A^{-\frac{1}{4}p} & & E^{-\frac{1}{2}p} & \\
 & E\flat^{+\frac{1}{4}p} & & B\flat^{+\frac{1}{4}p} & & F^{+\frac{1}{4}p} & & C^{0} & &
\end{array}$$

Neidhardt's Circulating Temperament, No. 1 "für ein Dorf" (for a village)
(Johann Georg Neidhardt, *Sectio canonis harmonici*, Königsberg, 1724, 16–18)

$$\begin{array}{cccccccccc}
 & E^{-\frac{2}{3}p} & & B^{-\frac{3}{4}p} & & F\sharp^{-\frac{5}{6}p} & & C\sharp^{-\frac{5}{6}p} & & G\sharp^{-\frac{5}{6}p} \\
C^{0} & & G^{-\frac{1}{6}p} & & D^{-\frac{1}{3}p} & & A^{-\frac{1}{2}p} & & E^{-\frac{2}{3}p} & \\
 & E\flat^{0} & & B\flat^{0} & & F^{0} & & C^{0} & &
\end{array}$$

Neidhardt's Circulating Temperament, No. 2 "für ein kleine Stadt" (for a small town)
(Johann Georg Neidhardt, 1724)[20]

$$\begin{array}{cccccccccc}
 & E^{-\frac{7}{12}p} & & B^{-\frac{7}{12}p} & & F\sharp^{-\frac{2}{3}p} & & C\sharp^{-\frac{3}{4}p} & & G\sharp^{-\frac{5}{6}p} \\
C^{0} & & G^{-\frac{1}{6}p} & & D^{-\frac{1}{3}p} & & A^{-\frac{1}{2}p} & & E^{-\frac{2}{3}p} & \\
 & E\flat^{+\frac{1}{6}p} & & B\flat^{+\frac{1}{6}p} & & F^{+\frac{1}{12}p} & & C^{0} & &
\end{array}$$

Neidhardt's Circulating Temperament, No. 3 "für eine grosse Stadt" (for a large town)
(Johann Georg Neidhardt, 1724)

$$\begin{array}{cccccccccc}
 & E^{-\frac{7}{12}p} & & B^{-\frac{7}{12}p} & & F\sharp^{-\frac{2}{3}p} & & C\sharp^{-\frac{3}{4}p} & & G\sharp^{-\frac{5}{6}p} \\
C^{0} & & G^{-\frac{1}{6}p} & & D^{-\frac{1}{3}p} & & A^{-\frac{1}{2}p} & & E^{-\frac{2}{3}p} & \\
 & E\flat^{+\frac{1}{6}p} & & B\flat^{+\frac{1}{12}p} & & F^{0} & & C^{0} & &
\end{array}$$

---

[20] Barbour [**6**] has $E^{-\frac{1}{12}p}$, which is incorrect, although he gives the correct value in cents. This seems to be nothing more than a typographical error.

Neidhardt's Circulating Temperament, No. 4 "für den Hof" (for the court) is the same as twelve tone equal temperament.

Kirnberger II
(Johann Phillip Kirnberger, *Construction der gleichschwebenden Temperatur*,[21] Berlin, 1764)

$$\begin{array}{cccccccccc}
 & & \mathrm{E}^{-1} & & \mathrm{B}^{-1} & & \mathrm{F}\sharp^{-1} & & & \\
 & \mathrm{C}^{0} & & \mathrm{G}^{0} & & \mathrm{D}^{0} & & \mathrm{A}^{-\frac{1}{2}} & & \mathrm{E}^{-1} \\
\mathrm{A}\flat^{0} & & \mathrm{E}\flat^{0} & & \mathrm{B}\flat^{0} & & \mathrm{F}^{0} & & \mathrm{C}^{0} & \\
 & & & & & \mathrm{D}\flat^{0} & & \mathrm{A}\flat^{0} & &
\end{array}$$

Kirnberger III
(Johann Phillip Kirnberger, *Die Kunst des reinen Satzes in der Musik* 2nd part, 3rd division, Berlin, 1779)

$$\begin{array}{cccccccccc}
 & & \mathrm{E}^{-1} & & \mathrm{B}^{-1} & & \mathrm{F}\sharp^{-1} & & & \\
 & \mathrm{C}^{0} & & \mathrm{G}^{-\frac{1}{4}} & & \mathrm{D}^{-\frac{1}{2}} & & \mathrm{A}^{-\frac{3}{4}} & & \mathrm{E}^{-1} \\
\mathrm{A}\flat^{0} & & \mathrm{E}\flat^{0} & & \mathrm{B}\flat^{0} & & \mathrm{F}^{0} & & \mathrm{C}^{0} & \\
 & & & & & \mathrm{D}\flat^{0} & & \mathrm{A}\flat^{0} & &
\end{array}$$

Marpurg's Temperament I
(Friedrich Wilhelm Marpurg, *Versuch über die musikalische Temperatur*, Breslau, 1776)

$$\begin{array}{cccccccccc}
 & \mathrm{E}^{-\frac{1}{3}p} & & \mathrm{B}^{-\frac{1}{3}p} & & \mathrm{F}\sharp^{-\frac{1}{3}p} & & \mathrm{C}\sharp^{-\frac{1}{3}p} & & \mathrm{G}\sharp^{-\frac{2}{3}p} \\
\mathrm{C}^{0} & & \mathrm{G}^{0} & & \mathrm{D}^{0} & & \mathrm{A}^{0} & & \mathrm{E}^{-\frac{1}{3}p} & \\
 & \mathrm{E}\flat^{+\frac{1}{3}p} & & \mathrm{B}\flat^{+\frac{1}{3}p} & & \mathrm{F}^{+\frac{1}{3}p} & & \mathrm{C}^{0} & &
\end{array}$$

Barca's $\frac{1}{6}$-comma temperament
(Alessandro Barca, *Introduzione a una nuova teoria di musica, memoria prima* Accademia di scienze, lettere ed arti in Padova. Saggi scientifici e lettari (Padova, 1786), 365–418)

$$\begin{array}{cccccccccc}
 & \mathrm{E}^{-\frac{2}{3}} & & \mathrm{B}^{-\frac{5}{6}} & & \mathrm{F}\sharp^{-1} & & \mathrm{C}\sharp^{-1} & & \mathrm{G}\sharp^{-1} \\
\mathrm{C}^{0} & & \mathrm{G}^{-\frac{1}{6}} & & \mathrm{D}^{-\frac{1}{3}} & & \mathrm{A}^{-\frac{1}{2}} & & \mathrm{E}^{-\frac{2}{3}} & \\
 & \mathrm{E}\flat^{0} & & \mathrm{B}\flat^{0} & & \mathrm{F}^{0} & & \mathrm{C}^{0} & &
\end{array}$$

Young's Temperament, No. 1
(Thomas Young, *Outlines of experiments and inquiries respecting sound and light* Philosophical Transactions, XC (1800), 106–150)

$$\begin{array}{cccccccccc}
 & \mathrm{E}^{-\frac{3}{4}} & & \mathrm{B}^{-\frac{5}{6}} & & \mathrm{F}\sharp^{-\frac{11}{12}} & & \mathrm{C}\sharp^{-\frac{11}{12}} & & \mathrm{G}\sharp^{-\frac{11}{12}} \\
\mathrm{C}^{0} & & \mathrm{G}^{-\frac{3}{16}} & & \mathrm{D}^{-\frac{3}{8}} & & \mathrm{A}^{-\frac{9}{16}} & & \mathrm{E}^{-\frac{3}{4}} & \\
 & \mathrm{E}\flat^{+\frac{1}{6}} & & \mathrm{B}\flat^{+\frac{1}{6}} & & \mathrm{F}^{+\frac{1}{12}} & & \mathrm{C}^{0} & &
\end{array}$$

Vallotti and Young $\frac{1}{6}$-comma temperament (Young's Temperament, No. 2)
(Francescantonio Vallotti, *Trattato delle musica moderna*, 1780; Thomas Young, *Outlines of experiments and inquiries respecting sound and light* Philosophical Transactions, XC (1800), 106–150. Below

[21]It is a mistake to deduce from this title that Kirnberger was in favor of equal temperament. In fact, he explicitly states in an article dated 1779 that "die gleichschwebende Temperatur ist schlechterdings ganz verwerflich…" (equal beating temperament is simply completely objectionable). And actually, there's a difference between equal and equal beating—the latter refers to the method of tuning whereby the fifths in a given octave are all made to beat at the same speeds.

Francescantonio Vallotti (1697–1780)

is Young's version of this temperament. In Vallotti's version, the fifths which are narrow by $\frac{1}{6}$ Pythagorean commas are F–C–G–D–A–E–B instead of C–G–D–A–E–B–F♯)

$$\begin{array}{cccccccccc}
 & E^{-\frac{2}{3}p} & & B^{-\frac{5}{6}p} & & F\sharp^{-1p} & & C\sharp^{-1p} & & G\sharp^{-1p} \\
C^{0} & & G^{-\frac{1}{6}p} & & D^{-\frac{1}{3}p} & & A^{-\frac{1}{2}p} & & E^{-\frac{2}{3}p} & \\
 & E\flat^{0} & & B\flat^{0} & & F^{0} & & C^{0} & &
\end{array}$$

The temperament of Vallotti and Young is probably closest to the intentions of J. S. Bach for his well-tempered clavier. According to the researches of Barnes, it is possible that Bach preferred the F♯ to be one sixth of a Pythagorean comma sharper than in this temperament, so that the fifth from B to F♯ is pure. Barnes based his work on a statistical study of prominence of the different major thirds, and the mathematical procedure of Hall[22] for evaluating suitability of temperaments. Other authors, such as Kelletat

[22]D. E. Hall, *The objective measurement of goodness-of-fit for tuning and temperaments*, J. Music Theory **17** (2) (1973), 274–290; *Quantitative evaluation of musical scale tuning*, American J. of Physics **42** (1974), 543–552.

and Kellner have come to slightly different conclusions, and we will probably never find out who is right. Here are these reconstructions for comparison.

Kelletat's Bach reconstruction (1966)

Herbert Kelletat, *Zur musikalischen Temperatur insbesondere bei J. S. Bach.* Onkel Verlag, Kassel, 1960 and 1980.

$$\begin{array}{cccccccccc}
 & \mathrm{E}^{-\frac{5}{6}p} & & \mathrm{B}^{-1p} & & \mathrm{F}\sharp^{-1p} & & \mathrm{C}\sharp^{-1p} & & \mathrm{G}\sharp^{-1p} \\
\mathrm{C}^{0} & & \mathrm{G}^{-\frac{1}{12}p} & & \mathrm{D}^{-\frac{1}{3}p} & & \mathrm{A}^{-\frac{7}{12}p} & & \mathrm{E}^{-\frac{5}{6}p} & \\
 & \mathrm{E}\flat^{0} & & \mathrm{B}\flat^{0} & & \mathrm{F}^{0} & & \mathrm{C}^{0} & &
\end{array}$$

Kellner's Bach reconstruction (1975)

Herbert Anton Kellner, *Eine Rekonstruktion der wohltemperierten Stimmung von Johann Sebastian Bach.* Das Musikinstrument **26** (1977), 34–35; *Was Bach a mathematician?* English Harpsichord Magazine 2/2 April 1978, 32–36; *Comment Bach accordait-il son clavecin?* Flûte à Bec et instruments anciens 13–14, SDIA, Paris 1985.

$$\begin{array}{cccccccccc}
 & \mathrm{E}^{-\frac{4}{5}p} & & \mathrm{B}^{-\frac{4}{5}p} & & \mathrm{F}\sharp^{-1p} & & \mathrm{C}\sharp^{-1p} & & \mathrm{G}\sharp^{-1p} \\
\mathrm{C}^{0} & & \mathrm{G}^{-\frac{1}{5}p} & & \mathrm{D}^{-\frac{2}{5}p} & & \mathrm{A}^{-\frac{3}{5}p} & & \mathrm{E}^{-\frac{4}{5}p} & \\
 & \mathrm{E}\flat^{0} & & \mathrm{B}\flat^{0} & & \mathrm{F}^{0} & & \mathrm{C}^{0} & &
\end{array}$$

Barnes' Bach reconstruction (1979)

John Barnes, *Bach's Keyboard Temperament*, Early Music **7** (2) (1979), 236–249.

$$\begin{array}{cccccccccc}
 & \mathrm{E}^{-\frac{2}{3}p} & & \mathrm{B}^{-\frac{5}{6}p} & & \mathrm{F}\sharp^{-\frac{5}{6}p} & & \mathrm{C}\sharp^{-1p} & & \mathrm{G}\sharp^{-1p} \\
\mathrm{C}^{0} & & \mathrm{G}^{-\frac{1}{6}p} & & \mathrm{D}^{-\frac{1}{3}p} & & \mathrm{A}^{-\frac{1}{2}p} & & \mathrm{E}^{-\frac{2}{3}p} & \\
 & \mathrm{E}\flat^{0} & & \mathrm{B}\flat^{0} & & \mathrm{F}^{0} & & \mathrm{C}^{0} & &
\end{array}$$

Exercises

1. Take the information on various temperaments given in this section, and work out a table of values in cents for the notes of the scale.

2. If you have a synthesizer where each note of the scale can be retuned separately, retune it to some of the temperaments given in this section, using your answers to Exercise 1. Sequence some harpsichord music and play it through your synthesizer using these temperaments, and compare the results.

Further listening: (See Appendix R)

Johann Sebastian Bach, *The Complete Organ Music*, Volumes 6 and 8, recorded by Hans Fagius, using Neidhardt's Circulating Temperament No. 3 "für eine grosse Stadt" (for a large town).

The Katahn/Foote recording, *Six degrees of tonality* contains tracks comparing Mozart's *Fantasie* Kv. 397 in equal temperament, meantone, and an irregular temperament of Prelleur.

Johann Gottfried Walther, *Organ Works*, Volumes 1 and 2, played by Craig Cramer on the organ of St. Bonifacius, Tröchtelborn, Germany. This organ was restored in Kellner's reconstruction of Bach's temperament, shown above.

Aldert Winkelman, *Works by Mattheson, Couperin, and others.* The pieces by Johann Mattheson, François Couperin, Johann Jakob Froberger, Joannes de Gruytters and Jacques Duphly are played on a harpsichord tuned to Werckmeister III.

5.14. Equal temperament

Music is a science which should have definite rules; these rules should be drawn from an evident principle; and this principle cannot really be known to us without the aid of mathematics. Notwithstanding all the experience I may have acquired in music from being associated with it for so long, I must confess that only with the aid of mathematics did my ideas become clear and did light replace a certain obscurity of which I was unaware before.

Rameau [**108**], 1722.[23]

Each of the scales described in the previous sections has its advantages and disadvantages, but the one disadvantage of most of them is that they are designed to make one particular key signature or a few adjacent key signatures as good as possible, and leave the remaining ones to look after themselves.

Twelve tone equal temperament is a natural endpoint of these compromises. This is the scale that results when all twelve semitones are taken to have equal ratios. Since an octave is a ratio of 2:1, the ratios for the equal tempered scale give all semitones a ratio of $2^{\frac{1}{12}}$:1 and all tones a ratio of $2^{\frac{1}{6}}$:1. So the ratios come out as follows:

note	do	re	mi	fa	so	la	ti	do
ratio	1:1	$2^{\frac{1}{6}}$:1	$2^{\frac{1}{3}}$:1	$2^{\frac{5}{12}}$:1	$2^{\frac{7}{12}}$:1	$2^{\frac{3}{4}}$:1	$2^{\frac{11}{12}}$:1	2:1
cents	0.000	200.000	400.000	500.000	700.000	900.000	1100.000	1200.000

Equal tempered thirds are about 14 cents sharper than perfect thirds, and sound nervous and agitated. As a consequence, the just and meantone scales are more calm temperaments. To my ear, tonal polyphonic music played in meantone temperament has a clarity and sparkle that I do not hear on equal tempered instruments. The irregular temperaments described in the previous section have the property that each key retains its own characteristics and color; keys with few sharps and flats sound similar to meantone, while the ones with more sharps and flats have a more remote feel to them. Equal temperament makes all keys essentially equivalent.

A factor of

$$\left(\frac{81}{80}\right)^{\frac{12}{11}} \approx 1.013644082,$$

or 23.4614068 cents is an extremely good approximation to the Pythagorean comma of

$$\frac{531441}{524288} \approx 1.013643265,$$

or 23.4600104 cents. It follows that equal temperament is almost exactly equal to the $\frac{1}{11}$-comma meantone scale:

[23]Page xxxv of the preface, in the Dover edition.

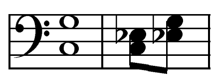
$$\begin{array}{ccccccccccc}
 & & \mathrm{E}^{-\frac{4}{11}} & & \mathrm{B}^{-\frac{5}{11}} & & \mathrm{F}\sharp^{-\frac{6}{11}} & & \mathrm{C}\sharp^{-\frac{7}{11}} & & \mathrm{G}\sharp^{-\frac{8}{11}} \\
 & \mathrm{C}^{0} & & \mathrm{G}^{-\frac{1}{11}} & & \mathrm{D}^{-\frac{2}{11}} & & \mathrm{A}^{-\frac{3}{11}} & & \mathrm{E}^{-\frac{4}{11}} & \\
\mathrm{A}\flat^{+\frac{4}{11}} & & \mathrm{E}\flat^{+\frac{3}{11}} & & \mathrm{B}\flat^{+\frac{2}{11}} & & \mathrm{F}^{+\frac{1}{11}} & & \mathrm{C}^{0} & &
\end{array}$$

where the difference between $\mathrm{A}\flat^{+\frac{4}{11}}$ and $\mathrm{G}\sharp^{-\frac{8}{11}}$ is 0.0013964 cents.

This observation was first made by Kirnberger[24] who used it as the basis for a recipe for tuning keyboard instruments in equal temperament. His recipe was to obtain an interval of an equal tempered fourth by tuning up three perfect fifths and one major third, and then down four perfect fourths. This corresponds to equating the equal tempered F with $\mathrm{E}\sharp^{-1}$. The disadvantage of this method is clear: in order to obtain one equal tempered interval, one must tune eight intervals by eliminating beats. The fifths and fourths are not so hard, but tuning a major third by eliminating beats is considered difficult. This method of tuning equal temperament was discovered independently by John Farey[25] nearly twenty years later.

Alexander Ellis in his Appendix XX (Section G, Article 11) to Helmholtz [**54**] gives an easier practical rule for tuning in equal temperament. Namely, tune the notes in the octave above middle C by tuning fifths upwards and fourths downwards. Make the fifths perfect and then flatten them (make them more narrow) by one beat per second (cf. §1.7). Make the fourths perfect and then flatten them (make them wider) by three beats every two seconds. The result will be accurate to within two cents on every note. Having tuned one octave using this rule, tuning out beats for octaves allows the entire piano to be tuned.

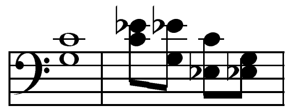
It is desirable to apply spot checks throughout the piano to ensure that the fifths remain slightly narrow and the fourths slightly wide. Ellis states at the end of Article 11 that there is no way of distinguishing slightly narrow fourths or fifths from slightly wide ones using beats. In fact, there is a method, which was not yet known in 1885, as follows (Jorgensen [**65**], §227).



For the fifth, say C3–G3, compare the intervals C3–E♭3 and E♭3–G3. If the fifth is narrow, as desired, the first interval will beat more frequently than the second. If perfect, the beat frequencies will be equal. If wide, the second interval will beat more frequently than the first.

[24]Johann Philipp Kirnberger, *Die Kunst des reinen Satzes in der Musik*, 2nd part 3rd division (Berlin, 1779), pp. 197f.

[25]John Farey, *On a new mode of equally tempering the musical scale*, Philosophical Magazine, XXVII (1807), 65–66.



For the fourth, say G3–C4, compare the intervals C4–E♭4 and G3–E♭4, or compare E♭3–C4 and E♭3–G3. If the fourth is wide, as desired, the first interval will beat more frequently than the second. If perfect, the beat frequencies will be equal. If narrow, the second interval will beat more frequently than the first. This method is based on the observation that in equal temperament, the major third is enough wider than a just major third, that gross errors would have to be made in order for it to have ended up narrower and spoil the test.

Exercises

1. Show that taking eleventh powers of the approximation of Kirnberger and Farey described in this section gives the approximation

$$2^{161} \approx 3^{84}\,5^{12}.$$

The ratio of these two numbers is roughly 1.000008873, and the eleventh root of this is roughly 1.0000008066.

2. Use the ideas of §4.6 to construct a spectrum which is close to the usual harmonic spectrum, but in such a way that the twelve tone equal tempered scale has consonant major thirds and fifths, as well as consonant seventh harmonics.

3. Calculate the accuracy of the method of Alexander Ellis for tuning equal temperament, described in this section.

4. Draw up a table of scale degrees in cents for the twelve notes in the Pythagorean, just, meantone and equal scales.

5. (Serge Cordier's equal temperament for piano with perfect fifths) Serge Cordier formalized a technique for piano tuning in the tradition of Pleyel (France). Cordier's recipe is as follows.[26] Make the interval F–C a perfect fifth, and divide it into seven equal semitones. Then use perfect fifths to tune from these eight notes to the entire piano.

Show that this results in octaves which are stretched by one seventh of a Pythagorean comma. This is of the same order of magnitude as the natural stretching of the octaves due to the inharmonicity of physical piano strings. Draw a diagram in Eitz's notation to demonstrate this temperament. This should consist of a horizontal strip with the top and bottom edges identified. Calculate the deviation of major and minor thirds from pure in this temperament.

Further reading:

Ian Stewart, *Another fine math you've got me into. . .*, W. H. Freeman & Co., 1992. Chapter 15 of this book, *The well tempered calculator*, contains a description of some

[26] Serge Cordier, *L'accordage des instruments à claviers.* Bulletin du Groupe Acoustique Musicale (G. A. M.) **75** (1974), Paris VII; *Piano bien tempéré et justesse orchestrale*, Buchet-Chastel, Paris 1982.

of the history of practical approximations to equal temperament. Particularly interesting is his description of Strähle's method of 1743.

5.15. Historical remarks

Ancient Greek music

The word *music* ($\mu o \upsilon \sigma \iota \kappa \acute{\eta}$) in ancient Greece had a wider meaning than it does for us, embracing the idea of ratios of integers as the key to understanding both the visible physical universe and the invisible spiritual universe.

It should not be supposed that the Pythagorean scale discussed in §5.2 was the main one used in ancient Greece in the form described there. Rather, this scale is the result of applying the Pythagorean ideal of using only the ratios 2:1 and 3:2 to build the intervals. The Pythagorean scale as we have presented it first occurs in Plato's *Timaeus*, and was used in medieval European music from about the eighth to the fourteenth century A.D.

The *diatonic syntonon* of Ptolemy is the same as the major scale of just intonation, with the exception that the classical Greek octave was usually taken to be made up of two Dorian[27] tetrachords, E–F–G–A and B–C–D–E, as described below, so that C was not the tonal center. It should be pointed out that Ptolemy recorded a long list of Greek diatonic tunings, and there is no reason to believe that he preferred the diatonic sytonic scale to any of the others he recorded.

The point of the Greek tunings was the construction of tetrachords, or sequences of four consecutive notes encompassing a perfect fourth; the ratio of 5:4 seems to have been an incidental consequence rather than representing a recognized consonant major third.

A Greek scale consisted of two tetrachords, either in *conjunction*, which means overlapping (for example two Dorian tetrachords B–C–D–E and E–F–G–A) or in *disjunction*, which means non-overlapping (for example E–F–G–A and B–C–D–E) with a whole tone as the gap. The tetrachords came in three types, called *genera* (plural of *genus*), and the two tetrachords in a scale belong to the same genus. The first genus is the *diatonic genus* in which the lowest interval is a semitone and the two upper ones are tones. The second is the *chromatic genus* in which the lowest two intervals are semitones and the upper one is a tone and a half. The third is the *enharmonic genus* in which the lowest two intervals are quarter tones and the upper one is two tones. The exact values of these intervals varied somewhat according to usage.[28] The interval between the lowest note and the higher of the two movable notes of a chromatic or enharmonic tetrachord is called the *pyknon*, and is always smaller than the remaining interval at the top of the tetrachord.

[27]Dorian tetrachords should not be confused with the Dorian mode of medieval church music, which is D–E–F–G–A–B–C–D. See Appendix M.

[28]For example, Archytas described tetrachords using the ratios 1:1, 28:27, 32:27, 4:3 (diatonic), 1:1, 28:27, 9:8, 4:3 (chromatic) and 1:1, 28:27, 16:15, 4:3 (enharmonic), in which the primes 2, 3, 5 and 7 appear. Plato, his contemporary, does not allow primes other than 2 and 3, in better keeping with the Pythagorean tradition.

Medieval to modern music

Little is known about the harmonic content, if any, of European music prior to the decline of the Roman Empire. The music of ancient Greece, for example, survives in a small handful of fragments, and is mostly melodic in nature. There is no evidence of continuity of musical practice from ancient Greece to medieval European music, although the theoretical writings had a great deal of impact.

Harmony, in a primitive form, seems to have first appeared in liturgical plainchant around 800 A.D., in the form of *parallel organum*, or melody in parallel fourths and fifths. Major thirds were not regarded as consonant, and a Pythagorean tuning system works perfectly for such music.

Polyphonic music started developing around the eleventh century A.D. Pythagorean intonation continued to be used for several centuries, and so the consonances in this system were the perfect fourths, fifths and octaves. The major third was still not regarded as a consonant interval, and it was something to be used only in passing.

The earliest known advocates of the 5:4 ratio as a consonant interval are the Englishmen Theinred of Dover (twelfth century) and Walter Odington (fl. 1298–1316),[29] in the context of early English polyphonic music. One of the earliest recorded uses of the major third in harmony is the four part vocal canon *sumer is icumen in*, of English origin, dating from around 1250. But for keyboard music, the question of tuning delayed its acceptance.

British folk music from the fourteenth and fifteenth centuries involved harmonizing around a melodic line by adding major thirds under it and perfect fourths over it to give parallel $\substack{6\\3}$ chords. The consonant major third traveled from England to the European continent in the early fifteenth century. But when the French imitated the sound of the parallel $\substack{6\\3}$ chords, they use the top line rather than the middle line as the melody, giving what is referred to as *Faux Bourdon*.

In more formal music, Dunstaple was one of the most well known British composers of the early fifteenth century to use the consonant major third. The story goes that the Duke of Bedford, who was Dunstaple's employer, inherited land in the north of France and moved there some time in the 1420's or 1430's. The French heard Dunstaple's consonant major thirds and latched onto the idea. Guillaume Dufay was the first major French composer to use it extensively. The accompanying transition from modality to tonality can be traced from Dufay through Ockegham, Josquin, Palestrina and Monteverdi during the fifteenth and sixteenth century.

The method for obtaining consonant major thirds in fourteenth and fifteenth century keyboard music is interesting. Starting with a series of Pythagorean fifths

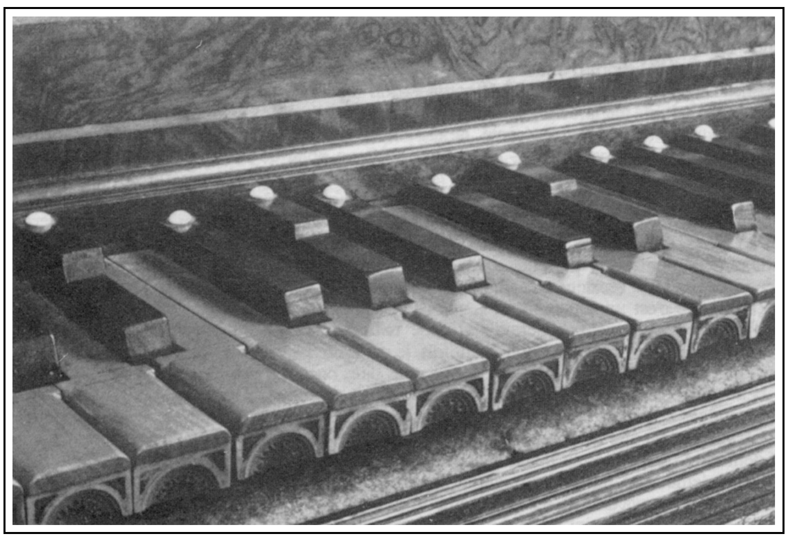
$$\mathrm{G}\flat^{0} - \mathrm{D}\flat^{0} - \mathrm{A}\flat^{0} - \mathrm{E}\flat^{0} - \mathrm{B}\flat^{0} - \mathrm{F}^{0} - \mathrm{C}^{0} - \mathrm{G}^{0} - \mathrm{D}^{0} - \mathrm{A}^{0} - \mathrm{E}^{0} - \mathrm{B}^{0},$$

[29] The "fl." indicates that these are the years in which he is known to have flourished.

the triad $\mathrm{D}^{0}-\mathrm{G}\flat^{0}-\mathrm{A}^{0}$ is used as a major triad. A just major triad would be $\mathrm{D}^{0}-\mathrm{F}\sharp^{-1}-\mathrm{A}^{0}$, and the difference between $\mathrm{F}\sharp^{-1}$ and $\mathrm{G}\flat^{0}$ is one schisma, or 1.953 cents. This is much more consonant than the modern equal temperament, in which the major thirds are impure by 13.686 cents. Other major triads available in this system are $\mathrm{A}^{0}-\mathrm{D}\flat^{0}-\mathrm{E}^{0}$ and $\mathrm{E}^{0}-\mathrm{A}\flat^{0}-\mathrm{B}^{0}$, but the system does not include a consonant C – E – G triad.

By the mid to late fifteenth century, especially in Italy, many aspects of the arts were reaching a new level of technical and mathematical precision. Leonardo da Vinci was integrating the visual arts with the sciences in revolutionary ways. In music, the meantone temperament was developed around this time, allowing the use of major and minor triads in a wide range of keys, and allowing harmonic progressions and modulations which had previously not been possible.

Many keyboard instruments from the sixteenth century have split keys for one or both of G♯/A♭ and D♯/E♭ to extend the range of usable key signatures. This was achieved by splitting the key across the middle, with the back part higher than the front part. The picture below shows the split keys of the Malamini organ in San Petronio, Bologna, Italy.



The practice for music of the sixteenth and seventeenth century was to choose a tonal center and gradually move further away. The furthest reaches were sparsely used, before gradually moving back to the tonal center.

Exact meantone tuning was not achieved in practice before the twentieth century, for lack of accurate prescriptions for tuning intervals. Keyboard instrument tuners tended to color the temperament, so that different keys had slightly different sounds to them. The irregular temperaments of §5.13 took this process further, and to some extent formalized it.

An early advocate for equal temperament for keyboard instruments was Rameau (1730). This helped it gain in popularity, until by the early nineteenth century it was fairly widely used, at least in theory. However, much of

Italian clavecin (1619) with split keys,
Musée Instrumental, Brussels, Belguim

Beethoven's piano music is best played with an irregular temperament (see §5.13), and Chopin was reluctant to compose in certain keys (notably D minor) because their characteristics did not suit him. In practice, equal temperament did not really take full hold until the end of the nineteenth century. Nineteenth century piano tuning practice often involved slight deviation from equal temperament in order to preserve, at least to some extent, the individual characteristics of the different keys. In the twentieth century, the dominance of chromaticism and the advent of twelve tone music have pretty much forced the abandonment of unequal temperaments, and piano tuning practice has reflected this.

Twelve tone music

Equal temperament is an essential ingredient in twentieth century twelve tone music, where combinatorics and chromaticism seem to supersede harmony. Some interesting evidence that harmonic content is irrelevant in Schoenberg's music is that the performance version of one of his most popular works, *Pierrot Lunaire*, contained many transcription errors confusing sharps, naturals and flats, until it was reedited for his collected works in the eighties.

The mathematics involved in twelve tone music of the twentieth century is different in nature to most of the mathematics we have described so far. It is more combinatorial in nature, and involves discussions of subsets of the twelve tones, and permutations, of the twelve tones of the chromatic scale. We shall have more to say on this subject in Chapter 9.

The role of the synthesizer

Before the days of digital synthesizers, we had a choice of several different versions of the tuning compromise. The just scales have perfect intervals, but do not allow us to modulate from the original key, and have problems with the triad on ii, and with syntonic commas interfering in fairly short harmonic sequences. Meantone scales sacrifice a little perfection in the fifths in order to remove the problem of the syntonic comma, but still have a problem with keys far removed from the original key, and with enharmonic modulations. Equal temperament works in all keys equally well, or rather, one might say equally badly. In particular, the equal tempered major third is nervous and agitated.

In these days of digitally synthesized and controlled music, there is very little reason to make do with the equal tempered compromise, because we can retune any note by any amount as we go along. It may still make sense to prefer a meantone scale to a just one on the grounds of interference of the syntonic comma, but it may also make sense to turn the situation around and use the syntonic comma for effect.

It seems that for most users of synthesizers the extra freedom has not had much effect, in the sense that most music involving synthesizers is written using the equal tempered twelve tone scale. A notable exception is Wendy Carlos, who has composed a great deal of music for synthesizers using many different scales. I particularly recommend *Beauty in the Beast*, which has been released on compact disc (SYNCD 200, Audion, 1986, Passport Records, Inc.). For example the fourth track, called *Just Imaginings*, uses a version of just intonation with harmonics all the way up to the nineteenth, and includes some deft modulations. Other tracks use other scales, including Carlos' *alpha* and *beta* scales and the Balinese gamelan *pelog* and *slendro*.

Wendy Carlos' earlier recordings, *Switched on Bach* and *The Well Tempered Synthesizer*, were recorded on a Moog synthesizer fixed in equal temperament. But when *Switched on Bach 2000* came out in 1992, twenty-five years after the original, it made use of a variety of meantone and unequal temperaments. It is not hard to hear from this recording the difference in clarity between these and equal temperament.

Further reading:

1) **History of music theory**

Thomas Christensen (ed.), *Cambridge history of music theory*, 2002 [**19**].

Leo Treitler, *Strunk's source readings in music history*, Revised Edition, Norton & Co., 1998. This 1552 page book, originally by Strunk but revised extensively by Treitler, contains translations of historical documents from ancient Greece to the twentieth century. It comes in seven sections, which are available in separate paperbacks.

2) **Ancient Greek music**

W. D. Anderson, *Music and musicians in ancient Greece*, Cornell University Press, 1994; paperback edition 1997.

Andrew Barker, *Greek Musical Writings, Vol. 2: Harmonic and acoustic theory*, Cambridge University Press, 1989. This 581 page book contains translations and commentaries on many of the most important ancient Greek sources, including Aristoxenus' *Elementa Harmonica*, the Euclidean *Sectio Canonis*, Nicomachus' *Enchiridion*, Ptolemy's *Harmonics*, and Aristides Quintilianus' *De Musica*.

Giovanni Comotti, *Music in Greek and Roman culture*, Johns Hopkins University Press, 1989; paperback edition 1991.

John G. Landels, *Music in ancient Greece and Rome*, Routledge, 1999; paperback edition 2001.

Thomas J. Mathiesen, *Apollo's lyre: Greek music and music theory in antiquity and the middle ages*, University of Nebraska Press, 1999.

M. L. West, *Ancient Greek music*, Oxford University Press, 1992; paperback edition 1994. Chapter 10 of this book reproduces all 51 known fragments of ancient Greek music.

R. P. Winnington-Ingram, *Mode in ancient Greek music*, Cambridge University Press, 1936. Reprinted by Hakkert, Amsterdam, 1968.

3) **Medieval to modern music**

Gustave Reese, *Music in the middle ages*, Norton, 1940, reprinted 1968. Despite the age of this text, it is still regarded as an invaluable source because of the quality of the scholarship. But the reader should bear in mind that much information has come to light since it appeared.

D. J. Grout and C. V. Palisca, *A history of western music*, fifth edition, Norton, 1996. Originally written in the 1950s by Grout, and updated a number of times by Palisca. This is a standard text used in many music history departments.

Owen H. Jorgensen, *Tuning* [**65**] contains an excellent discussion of the development of temperament, and argues that equal temperament was not commonplace in practice until the twentieth century.

4) **Twelve tone music**

Allen Forte, *The structure of atonal music* [**41**].

George Perle, *Twelve tone tonality* [**99**].

5) **The role of the synthesizer**

Easley Blackwood, *Discovering the microtonal resources of the synthesizer*, Keyboard, May 1982, 26–38.

Benjamin Frederick Denckla, *Dynamic intonation in synthesizer performance*, M.Sc. Thesis, MIT, 1997 (61 pp).

Henry Lowengard, *Computers, digital synthesizers and microtonality*, Pitch 1 (1) (1986), 6–7.

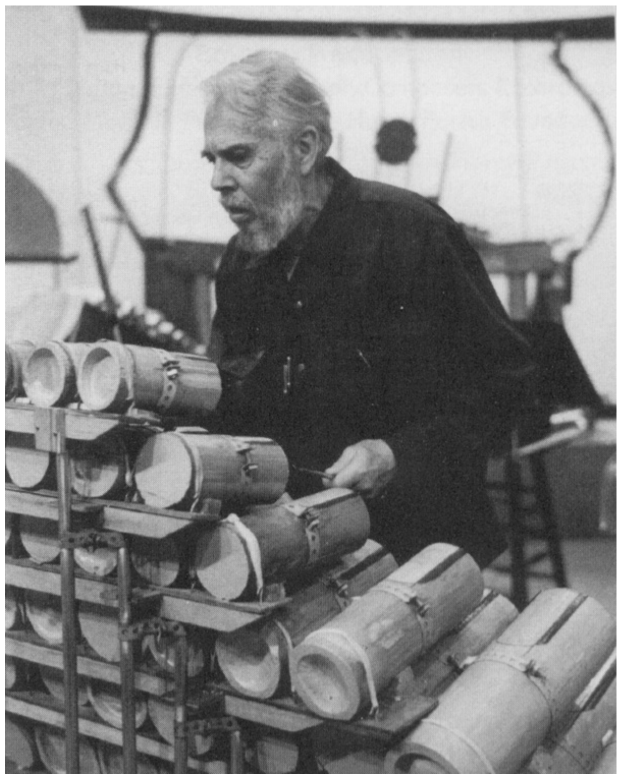
Robert Rich, *Just intonation for MIDI synthesizers*, Electronic Musician, Nov 1986, 32–45.

M. Yunik and G. W. Swift, *Tempered music scales for sound synthesis*, Computer Music Journal 4 (4) (1980), 60–65.

CHAPTER 6

More scales and temperaments

6.1. Harry Partch's 43 tone and other just scales



Harry Partch playing the
bamboo marimba (Boo I)

In §5.5, we talked about just intonation in its narrowest sense. This involved building up a scale using ratios only involving the primes 2, 3 and 5, to obtain a twelve tone scale. Just intonation can be extended far beyond this limitation. The phrase *super just* is sometimes used to denote a scale formed with exact rational multiples for the intervals, but using primes other than the 2, 3 and 5. Most of these come from the twentieth century.

Harry Partch developed a just scale of 43 notes which he used in a number of his compositions. The tonic for his scale is G^0. The scale is symmetric, in the sense that every interval upwards from G^0 is also an interval downwards from G^0.

The primes involved in Partch's scale are 2, 3, 5, 7 and 11. The terminology used by Partch to describe this is that his scale is based on the 11-limit, while the Pythagorean scale is based on the 3-limit and the just scales of §5.5 and §5.10 are based on the 5-limit. More generally, if p is a prime, then a p-limit scale only uses rational numbers whose denominators and numerators factor as products of prime numbers less than or equal to p (repetitions are allowed).

Harry Partch's 43 tone scale					
G^0	1:1	0.000		10:7	617.488
G^{+1}	81:80	21.506		16:11	648.682
	33:32	53.273	D^{-1}	40:27	680.449
	21:20	84.467	D^0	3:2	701.955
$A\flat^{+1}$	16:15	111.713		32:21	729.219
	12:11	150.637		14:9	764.916
	11:10	165.004		11:7	782.492
A^{-1}	10:9	182.404	$E\flat^{+1}$	8:5	813.686
A^0	9:8	203.910		18:11	852.592
	8:7	231.174	E^{-1}	5:3	884.359
	7:6	266.871	E^0	27:16	905.865
$B\flat^0$	32:27	294.135		12:7	933.129
$B\flat^{+1}$	6:5	315.641		7:4	968.826
	11:9	347.408	F^0	16:9	996.090
B^{-1}	5:4	386.314	F^{+1}	9:5	1017.596
	14:11	417.508		20:11	1034.996
	9:7	435.084		11:6	1049.363
	21:16	470.781	$F\sharp^{-1}$	15:8	1088.269
C^0	4:3	498.045		40:21	1115.533
C^{+1}	27:20	519.551		64:33	1146.727
	11:8	551.318	G^{-1}	160:81	1178.494
	7:5	582.512	G^0	2:1	1200.000

Here are some other just scales. The Chinese Lü scale by Huai-nan-dsi of the Han dynasty is the twelve tone just scale with ratios

1:1, 18:17, 9:8, 6:5, 54:43, 4:3, 27:19, 3:2, 27:17, 27:16, 9:5, 36:19, (2:1).

Wendy Carlos has developed several just scales. The "Wendy Carlos super just intonation" is the twelve tone scale with ratios

1:1, 17:16, 9:8, 6:5, 5:4, 4:3, 11:8, 3:2, 13:8, 5:3, 7:4, 15:8, (2:1).

The "Wendy Carlos harmonic scale" also has twelve tones, with ratios

1:1, 17:16, 9:8, 19:16, 5:4, 21:16, 11:8, 3:2, 13:8, 27:16, 7:4, 15:8, (2:1).

A better way of writing this might be to multiply all the entries by 16:

16, 17, 18, 19, 20, 21, 22, 24, 26, 27, 28, 30, (32).

Lou Harrison has a 16 tone just scale with ratios

1:1, 16:15, 10:9, 8:7, 7:6, 6:5, 5:4, 4:3, 17:12,
3:2, 8:5, 5:3, 12:7, 7:4, 9:5, 15:8, (2:1).

Wilfrid Perret[1] has a 19-tone 7-limit just scale with ratios

1:1, 21:20, 35:32, 9:8, 7:6, 6:5, 5:4, 21:16, 4:3, 7:5, 35:24,
3:2, 63:40, 8:5, 5:3, 7:4, 9:5, 15:8, 63:32, (2:1).

John Chalmers also has a 19 tone 7-limit just scale, differing from this in just two places. The ratios are

1:1, 21:20, 16:15, 9:8, 7:6, 6:5, 5:4, 21:16, 4:3, 7:5, 35:24,
3:2, 63:40, 8:5, 5:3, 7:4, 9:5, 28:15, 63:32, (2:1).

Michael Harrison has a 24 tone 7-limit just scale with ratios

1:1, 28:27, 135:128, 16:15, 243:224, 9:8, 8:7, 7:6, 32:27, 6:5, 135:112, 5:4,
81:64, 9:7, 21:16, 4:3, 112:81, 45:32, 64:45, 81:56, 3:2, 32:21, 14:9, 128:81,
8:5, 224:135, 5:3, 27:16, 12:7, 7:4, 16:9, 15:8, 243:128, 27:14, (2:1).

Harrison writes,

> Beginning in 1986, I spent two years extensively modifying a seven-foot Schimmel grand piano to create the *Harmonic Piano*. It is the first piano tuned in Just Intonation with the flexibility to modulate to multiple key centers at the press of a pedal. With its unique pedal mechanism, the Harmonic Piano can differentiate between notes usually shared by the same piano key (for example, C-sharp and D-flat). As a result, the Harmonic Piano is capable of playing 24 notes per octave. In contrast to the three unison strings per note of the standard piano, the Harmonic Piano uses only single strings, giving it a "harp-like" timbre. Special muting systems are employed to dampen unwanted resonances and to enhance the instrument's clarity of sound.[2]

The Indian Sruti scale,[3] commonly used to play ragas, is a 5-limit just scale with 22 tones, but has some large numerators and denominators:

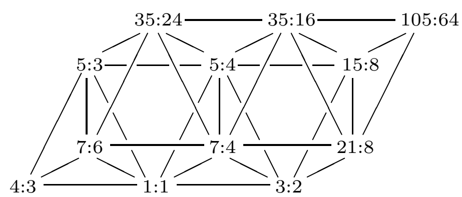
1:1, 256:243, 16:15, 10:9, 9:8, 32:27, 6:5, 5:4, 81:64, 4:3, 27:20, 45:32,
729:512, 3:2, 128:81, 8:5, 5:3, 27:16, 16:9, 9:5, 15:8, 243:128, (2:1).

[1] W. Perret, *Some questions of musical theory*, W. Heffer & Sons Ltd., Cambridge, 1926.

[2] From the liner notes to Harrison's CD *From Ancient Worlds, for Harmonic Piano*, see Appendix R.

[3] Taken from B. Chaitanya Deva, *The music of India* [**33**], Table 9.2. Note that the fractional value of note 5 given in this table should be 32/27, not 64/45, to match the other information given in this table. This also matches the value given in Tables 9.4 and 9.8 of the same work. Beware that the exact values of the intervals in Indian scales is a subject of much debate and historical controversy.

Various notations have been designed for describing just scales. For example, for 7-limit scales, a three-dimensional lattice of tetrahedra and octahedra can just about be drawn on paper. Here is an example of a twelve tone 7-limit just scale drawn three dimensionally in this way.[4]



The lines indicate major and minor thirds, perfect fifths, and three different septimal consonances 7:4, 7:5 and 7:6 (notes have been normalized to lie inside the octave 1:1 to 2:1). We return to the discussion of just intonation in §6.8, where we discuss unison vectors and periodicity blocks. We put the above diagram into context in §6.9.

Exercises

1. Taking 1:1 to be C^{0}, write the Indian Sruti scale described in this section as an array using Eitz's comma notation (like the scales in §5.10).

Further reading:

David B. Doty, *The just intonation primer* (1993), privately published and available from the Just Intonation Network at www.justintonation.net.

Harry Partch, *Genesis of a music* [**98**].

Joseph Yasser, *A theory of evolving tonality* [**142**].

Further listening: (See Appendix R)

Bill Alves, *Terrain of possibilities.*

Wendy Carlos, *Beauty in the Beast.*

Michael Harrison, *From Ancient Worlds.*

Harry Partch, *Bewitched.*

Robert Rich, *Rainforest*, *Gaudi.*

[4]This way of drawing the scale comes from Paul Erlich. According to Paul, the scale was probably first written down by Erv Wilson in the 1960's.

6.2. Continued fractions

$$e^{2\pi/5}\left(\sqrt{\frac{5+\sqrt{5}}{2}}-\frac{\sqrt{5}+1}{2}\right)=\frac{1}{1+}\,\frac{e^{-2\pi}}{1+}\,\frac{e^{-4\pi}}{1+}\,\frac{e^{-6\pi}}{1+}\cdots$$

Srinivasa Ramanujan

The modern twelve tone equal tempered scale is based around the fact that

$$7/12 = 0.58333\dots$$

is a good approximation to

$$\log_2(3/2) = 0.5849625007\dots,$$

so that if we divide the octave into twelve equal semitones, then seven semitones is a good approximation to a perfect fifth. This suggests the following question. Can $\log_2(3/2)$ be expressed as a ratio of two integers, m/n? In other words, is $\log_2(3/2)$ a rational number? Since $\log_2(3/2)$ and $\log_2(3)$ differ by one, this is the same as asking whether $\log_2(3)$ is rational.

LEMMA 6.2.1. *The number* $\log_2(3)$ *is irrational.*

PROOF. Suppose that $\log_2(3) = m/n$ with m and n integers. Then $3 = 2^{m/n}$, or $3^n = 2^m$. This is obviously impossible, as 3^n is odd while 2^m is even. □

So the best we can expect to do is to approximate $\log_2(3/2)$ by rational numbers such as $7/12$. There is a systematic theory of such rational approximations to irrational numbers, which is the theory of continued fractions.[5] A continued fraction is an expression of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \dots}}}$$

where $a_0, a_1, \dots$ are integers, usually taken to be positive for $i \geq 1$. The expression is allowed to stop at some finite stage, or it may go on for ever. For typographic convenience, we write the continued fraction in the form

$$a_0 + \frac{1}{a_1+}\,\frac{1}{a_2+}\,\frac{1}{a_3+}\dots$$

For even greater compression of notation, this is sometimes written as

$$[a_0; a_1, a_2, a_3, \dots].$$

Every real number has a unique continued fraction expansion, and it stops precisely when the number is rational. The easiest way to see this is as follows. If x is a real number, then the largest integer less than or equal to x

[5]The first mathematician known to have made use of continued fractions was Rafael Bombelli in 1572. The modern notation for them was introduced by P. A. Cataldi in 1613.

(the *integer part* of x) is written $\lfloor x \rfloor$.[6] So $\lfloor x \rfloor$ is what we take for a_0. The remainder $x - \lfloor x \rfloor$ satisfies $0 \le x - \lfloor x \rfloor < 1$, so if it is nonzero, we now invert it to obtain a number $1/(x - \lfloor x \rfloor)$ which is strictly larger than one.

Writing $x_0 = x$, $a_0 = \lfloor x_0 \rfloor$ and $x_1 = 1/(x_0 - \lfloor x_0 \rfloor)$, we have

$$x = a_0 + \frac{1}{x_1}.$$

Now just carry on going. Let $a_1 = \lfloor x_1 \rfloor$, and $x_2 = 1/(x_1 - \lfloor x_1 \rfloor)$, so that

$$x = a_0 + \frac{1}{a_1+}\,\frac{1}{x_2}.$$

Inductively, we set $a_n = \lfloor x_n \rfloor$ and $x_{n+1} = 1/(x_n - \lfloor x_n \rfloor)$ so that

$$x = a_0 + \frac{1}{a_1+}\,\frac{1}{a_2+}\,\frac{1}{a_3+} \dots$$

This algorithm continues provided each $x_n \neq 0$, which happens exactly when x is irrational. Otherwise, if x is rational, the algorithm terminates to give a finite continued fraction. For irrational numbers the continued fraction expansion is unique. For rational numbers, we only have uniqueness if we stipulate that the last a_n is larger than one.

As an example, let us compute the continued fraction expansion of

$$\pi = 3.14159\,26535\,89793\,23846\,26433\,83279\,50288\,41971\,69399\,37510\,58209\,74944\,59230\,78164\dots$$

In this case, we have $a_0 = 3$ and

$$x_1 = 1/(\pi - 3) = 7.062513086\dots$$

So $a_1 = 7$, and

$$x_2 = 1/(x_1 - 7) = 15.99665\dots$$

Continuing this way, we obtain

$$\pi = 3 + \frac{1}{7+}\,\frac{1}{15+}\,\frac{1}{1+}\,\frac{1}{292+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{2+}\,\frac{1}{1+}\,\frac{1}{3+}\,\frac{1}{1+}\,\frac{1}{14+}\,\dots$$

In the more compressed (and tinier) notation, here are more terms:[7]

$$\begin{aligned}\pi = [3;\,&7, 15, 1, 292, 1, 1, 1, 2, 1, 3, 1, 14, 2, 1, 1, 2, 2, 2, 2, 1, 84, 2, 1, 1, 15, 3, 13,\\ &1, 4, 2, 6, 6, 99, 1, 2, 2, 6, 3, 5, 1, 1, 6, 8, 1, 7, 1, 2, 3, 7, 1, 2, 1, 1, 12, 1, 1, 1, 3,\\ &1, 1, 8, 1, 1, 2, 1, 6, 1, 1, 5, 2, 2, 3, 1, 2, 4, 4, 16, 1, 161, 45, 1, 22, 1, 2, 2, 1, 4,\\ &1, 2, 24, 1, 2, 1, 3, 1, 2, 1, 1, 10, 2, 5, 4, 1, 2, 2, 8, 1, 5, 2, 2, 26, 1, 4, 1, 1, 8, 2,\\ &42, 2, 1, 7, 3, 3, 1, 1, 7, 2, 4, 9, 7, 2, 3, 1, 57, 1, 18, 1, 9, 19, 1, 2, 18, 1, 3, 7, 30,\\ &1, 1, 1, 3, 3, 3, 1, 2, 8, 1, 1, 2, 1, 15, 1, 2, 13, 1, 2, 1, 4, 1, 12, 1, 1, 3, 3, 28, 1, 10,\\ &3, 2, 20, 1, 1, 1, 1, 4, 1, 1, 1, 5, 3, 2, 1, 6, 1, 4, 1, 120, 2, 1, 1, 3, 1, 23, 1, 15, 1, 3,\\ &7, 1, 16, 1, 2, 1, 21, 2, 1, 1, 2, 9, 1, 6, 4, 127, 14, 5, 1, 3, 13, 7, 9, 1, 1, 1, 1, 1, 5,\\ &4, 1, 1, 3, 1, 1, 29, 3, 1, 1, 2, 2, 1, 3, 1, 1, 1, 3, 1, 1, 10, 3, 1, 3, 1, 2, 1, 12, 1, 4, 1,\\ &1, 1, 1, 7, 1, 1, 2, 1, 11, 3, 1, 7, 1, 4, 1, 48, 16, 1, 4, 5, 2, 1, 1, 4, 3, 1, 2, 3, 1, 2, 2,\\ &1, 2, 5, 20, 1, 1, 5, 4, 1, 436, 8, 1, 2, 2, 1, 1, 1, 1, 1, 5, 1, 2, 1, 3, 6, 11, 4, 3, 1, 1, 1,\end{aligned}$$

---

[6] In some books, $[x]$ is used instead.

[7] Note that the values given in Hua [**58**], page 252, are erronious. The correct values for the first 20,000,000 terms in the continued fraction expansion of π can be downloaded from www.lacim.uqam.ca/piDATA/

2, 5, 4, 6, 9, 1, 5, 1, 5, 15, 1, 11, 24, 4, 4, 5, 2, 1, 4, 1, 6, 1, 1, 1, 4, 3, 2, 2, 1, 1, 2,
1, 58, 5, 1, 2, 1, 2, 1, 1, 2, 2, 7, 1, 15, 1, 4, 8, 1, 1, 4, 2, 1, 1, 1, 3, 1, 1, 1, 2, 1, 1, 1,
1, 1, 9, 1, 4, 3, 15, 1, 2, 1, 13, 1, 1, 1, 3, 24, 1, 2, 4, 10, 5, 12, 3, 3, 21, 1, 2, 1, 34,
1, 1, 1, 4, 15, 1, 4, 44, 1, 4, 20776, 1, 1, 1, 1, 1, 1, 1, 23, 1, 7, 2, 1, 94, 55, 1, 1, 2, . . .]

To get good rational approximations, we stop just before a large value of a_n. So for example, stopping just before the 15, we obtain the well known approximation $\pi \approx 22/7$.[8] Stopping just before the 292 gives us the extremely good approximation

$$\pi \approx 355/113 = 3.1415929\ldots$$

which was known to the Chinese mathematician Chao Jung-Tze (or Tsu Ch'ung-Chi, depending on how you transliterate the name) in 500 AD.

The rational approximations obtained by truncating the continued fraction expansion of a number are called the *convergents*. So the convergents for π are

$$\frac{3}{1}, \frac{22}{7}, \frac{333}{106}, \frac{355}{113}, \frac{103993}{33102}, \frac{104348}{33215}, \ldots$$

There is an extremely efficient way to calculate the convergents from the continued fraction.

THEOREM 6.2.2. *Define numbers p_n and q_n inductively as follows:*

$$p_0 = a_0, \qquad p_1 = a_1 a_0 + 1, \qquad p_n = a_n p_{n-1} + p_{n-2} \quad (n \geq 2) \tag{6.2.1}$$

$$q_0 = 1, \qquad q_1 = a_1, \qquad q_n = a_n q_{n-1} + q_{n-2} \quad (n \geq 2). \tag{6.2.2}$$

Then we have

$$a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_n} = \frac{p_n}{q_n}.$$

PROOF. (see Hardy and Wright [**52**], Theorem 149, or Hua [**58**], Theorem 10.1.1).

The proof goes by induction on n. It is easy enough to check the cases $n = 0$ and $n = 1$, so we assume that $n \geq 2$ and that the theorem holds for smaller values of n. Then we have

$$a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_{n-1}+}\frac{1}{a_n} = a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_{n-1} + \frac{1}{a_n}}.$$

So we can use the formula given by the theorem with $n-1$ in place of n to write this as

$$\begin{aligned} \frac{(a_{n-1} + \frac{1}{a_n})p_{n-2} + p_{n-3}}{(a_{n-1} + \frac{1}{a_n})q_{n-2} + q_{n-3}} &= \frac{a_n(a_{n-1}p_{n-2} + p_{n-3}) + p_{n-2}}{a_n(a_{n-1}q_{n-2} + q_{n-3}) + q_{n-2}} \\ &= \frac{a_n p_{n-1} + p_{n-2}}{a_n q_{n-1} + q_{n-2}} = \frac{p_n}{q_n}. \end{aligned}$$

So the theorem is true for n, and the induction is complete. □

[8] According to the bible, π is equal to 3. "Also, he made a molten sea of ten cubits from brim to brim, round in compass, and five cubits the height thereof; and a line of thirty cubits did compass it round about." I Kings 7:23.

So in the above example for π, we have $p_0 = a_0 = 3$, $q_0 = 1$, $p_1 = a_1 a_0 + 1 = 22$, $q_1 = a_1 = 7$, we get

$$\frac{p_2}{q_2} = \frac{p_0 + 15p_1}{q_0 + 15q_1} = \frac{333}{106}$$

so that $p_2 = 333$, $q_2 = 106$,

$$\frac{p_3}{q_3} = \frac{p_1 + p_2}{q_1 + q_2} = \frac{355}{113}$$

so that $p_3 = 355$, $q_3 = 113$, and so on.

Examining the value of x_2 in the case $x = \pi$ above, it may look as though it would be of advantage to allow negative as well as positive values for a_n. However, this doesn't really help, because if x_n is very slightly less than $a_n + 1$ then a_{n+1} will be equal to one, and from there on the sequence as it would have been. In other words, the rational approximations obtained this way are no better. A related observation is that if $a_{n+1} = 2$ then it is worth examining the approximation given by replacing a_n by $a_n + 1$ and stopping there.

The continued fraction expansion for the base of natural logarithms

$$\begin{aligned} e &= 2.71828\,18284\,59045\,23536\,02874\,71352\,66249\,77572\,47093\ldots \\ &= 2 + \frac{1}{1+}\,\frac{1}{2+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{4+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{6+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{8+}\,\frac{1}{1+}\,\frac{1}{1+}\ldots \end{aligned}$$

follows an easily described pattern, as was discovered by Leonhard Euler. The continued fraction expansion of the golden ratio is even easier to describe:

$$\tau = \tfrac{1}{2}(1 + \sqrt{5}) = 1 + \frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{1+}\ldots$$

Although the continued fraction expansion of π is not regular in this way, there is a closely related formula (Brouncker)

$$\frac{\pi}{4} = \frac{1}{1+}\,\frac{1}{3+}\,\frac{4}{5+}\,\frac{9}{7+}\,\frac{16}{9+}\ldots$$

which is a special case of the arctan formula

$$\tan^{-1} z = \frac{z}{1+}\,\frac{z^2}{3+}\,\frac{4z^2}{5+}\,\frac{9z^2}{7+}\,\frac{16z^2}{9+}\ldots$$

The tan formula

$$\tan z = \frac{z}{1+}\,\frac{-z^2}{3+}\,\frac{-z^2}{5+}\,\frac{-z^2}{7+}\ldots$$

can be used to show that π is irrational (Pringsheim).

How good are the rational approximations obtained from continued fractions? This is answered by the following theorems. Recall that $x_n = p_n/q_n$ denotes the nth convergent. In other words,

$$\frac{p_n}{q_n} = a_0 + \frac{1}{a_1+}\,\frac{1}{a_2+}\cdots\frac{1}{a_{n-1}+}\,\frac{1}{a_n}.$$

THEOREM 6.2.3. *The error in the nth convergent of the continued fraction expansion of a real number x is bounded by*

$$\left|\frac{p_n}{q_n} - x\right| < \frac{1}{q_n^2}.$$

PROOF. (see Hardy and Wright [**52**], Theorem 171, or Hua [**58**], Theorem 10.2.6).

First, we notice that $p_{n-1}q_n - p_nq_{n-1} = (-1)^n$. This is easiest to see by induction. For $n = 1$, we have $p_0 = a_0$, $q_0 = 1$, $p_1 = a_0a_1 + 1$, $q_1 = a_1$, so $p_0a_1 - p_1a_0 = -1$. For $n > 1$, using equations (6.2.1) and (6.2.2) we have

$$\begin{aligned}
p_{n-1}q_n - p_nq_{n-1} &= p_{n-1}(q_{n-2} + a_nq_{n-1}) - (p_{n-2} + a_np_{n-1})q_{n-1}\\
&= p_{n-1}q_{n-2} - p_{n-2}q_{n-1}\\
&= -(p_{n-2}q_{n-1} - p_{n-1}q_{n-2})\\
&= -(-1)^{n-1} = (-1)^n.
\end{aligned}$$

Now we use the fact that x lies between

$$\frac{p_{n-2} + a_np_{n-1}}{q_{n-2} + a_nq_{n-1}} \qquad \text{and} \qquad \frac{p_{n-2} + (a_n+1)p_{n-1}}{q_{n-2} + (a_n+1)q_{n-1}}$$

or in other words between $\dfrac{p_n}{q_n}$ and $\dfrac{p_n + p_{n-1}}{q_n + q_{n-1}}$. The distance between these two numbers is

$$\begin{aligned}
\left|\frac{p_n + p_{n-1}}{q_n + q_{n-1}} - \frac{p_n}{q_n}\right| &= \left|\frac{(p_n + p_{n-1})q_n - p_n(q_n + q_{n-1})}{(q_n + q_{n+1})q_n}\right|\\
&= \left|\frac{p_{n-1}q_n - p_nq_{n-1}}{q_n^2 + q_nq_{n-1}}\right| = \left|\frac{(-1)^n}{q_n^2 + q_nq_{n-1}}\right| < \frac{1}{q_n^2}.
\end{aligned}$$

□

Notice that if we choose a denominator q at random, then the intervals between the rational numbers of the form p/q are of size $1/q$. So by choosing p to minimize the error, we get $|p/q - x| \le 1/2q$. So the point of the above theorem is that the convergents in the continued fraction expansion are considerably better than random denominators. In fact, more is true.

THEOREM 6.2.4. *Among the fractions p/q with $q \le q_n$, the closest to x is p_n/q_n.*

PROOF. See Hardy and Wright [**52**], Theorem 181. □

It is not true that if p/q is a rational number satisfying $|p/q-x| < 1/q^2$ then p/q is a convergent in the continued fraction expansion of x. However, a theorem of Hurwitz (see Hua [**58**], Theorem 10.4.1) says that of any two consecutive convergents to x, at least one of them satisfies $|p/q - x| < 1/2q^2$. Moreover, if a rational number p/q satisfies this inequality then it is a convergent in the continued fraction expansion of x (see Hua [**58**], Theorem 10.7.2).

Distribution of the a_n

If we perform continued fractions on a transcendental number x, given an integer k, how likely is it that $a_n = k$? It seems plausible that $a_n = 1$ is the most likely, and that the probabilities decrease rapidly as k increases, but what is the exact distribution of probabilities?

Gauss answered this question in a letter addressed to Laplace, although he never published a proof.[9] Writing $\mu\{-\}$ for the measure of a set $\{-\}$, what he proved is the following. Given any t in the range $(0,1)$, in the limit the measure of the set of numbers x in the interval $(0,1)$ for which $x_n - \lfloor x_n \rfloor$ is at most t is given by[10]

$$\lim_{n\to\infty} \mu\{\, x \in (0,1) \mid x_n - \lfloor x_n \rfloor \le t \,\} = \log_2(1+t).$$

The continued fraction process says that we should then invert $x_n - \lfloor x_n \rfloor$. Writing u for $1/t$, we obtain

$$\lim_{n\to\infty} \mu\{\, x \in (0,1) \mid \frac{1}{x_n - \lfloor x_n \rfloor} \ge u \,\} = \log_2(1+1/u).$$

Now we need to take the integer part of $1/(x_n - \lfloor x_n \rfloor)$ to obtain a_{n+1}. So if k is an integer with $k \ge 1$ then

$$\begin{aligned}\lim_{n\to\infty} \mu\{\, x \in (0,1) \mid a_n = k \,\} &= \log_2(1+\frac{1}{k}) - \log_2(1+\frac{1}{k+1}) \\ &= \log_2\left(\frac{(k+1)^2}{k(k+2)}\right) = \log_2\left(1+\frac{1}{k(k+2)}\right).\end{aligned}$$

We now tabulate the probabilities given by this formula.

Value of k	Limiting probability that $a_n = k$ as $n \to \infty$
1	0.4150375
2	0.2223924
3	0.0931094
4	0.0588937
5	0.0406420
6	0.0297473
7	0.0227201
8	0.0179219
9	0.0144996
10	0.0119726

For large k, this decreases like $1/k^2$.

[9] According to A. Ya. Khinchin, *Continued Fractions*, Dover 1964, page 72, the first published proof was by Kuz'min in 1928.

[10] If you don't know what measure means in this context, think of this as giving the probability that a randomly chosen number in the given interval satisfies the hypothesis.

Multiple continued fractions

It is sometimes necessary to make simultaneous rational approximations for more than one irrational number. For example, in the equal tempered scale, not only do seven semitones approximate a perfect fifth with ratio 3:2, but also four semitones approximates a major third with ratio 5:4. So we have

$$\log_2(3/2) \approx 7/12; \qquad \log_2(5/4) \approx 4/12.$$

A theorem of Dirichlet tells us how closely we should expect to be able to approximate a set of k real numbers simultaneously.

THEOREM 6.2.5. *If α_1, α_2, $\ldots \alpha_k$ are real numbers, and at least one of them is irrational, then the there exist an infinite number of ways of choosing a denominator q and numerators p_1, p_2, $\ldots$, p_k in such a way that the approximations*

$$p_1/q \approx \alpha_1; \quad p_2/q \approx \alpha_2; \quad \ldots \quad p_k/q \approx \alpha_k$$

have the property that the errors are all less than $1/q^{1+\frac{1}{k}}$.

PROOF. See Hardy and Wright [**52**], Theorem 200. □

The case $k = 1$ of this theorem is just Theorem 6.2.3. There is no known method when $k \geq 2$ analogous to the method of continued fractions for obtaining the approximations whose existence is guaranteed by this theorem. Of course, we can just work through the possibilities for q one at a time, but this is much more tedious than one would like.

The power of q in the denominator in the above theorem (i.e., $1 + \frac{1}{k}$) is known to be the best possible. Notice that the error term remains better than the error term $1/2q$ which would result by choosing q randomly. But the extent to which it is better diminishes to insignificant as k grows large.

Exercises

1. Investigate the convergents for the continued fraction expansion of the golden ratio $\tau = \frac{1}{2}(1+\sqrt{5})$. What do these convergents have to do with the Fibonacci series?

Coupled oscillators have a tendency to seek frequency ratios which can be expressed as rational numbers with small numerators and denominators. For example, Mercury rotates on its axis exactly three times for every two rotations around the sun, so that one Mercurial day lasts two Mercurial years. In a similar way, the orbital times of Jupiter and the minor planet Pallas around the sun are locked in a ratio of 18 to 7 (Gauss calculated in 1812 that this would be true, and observation has confirmed it). This is also why the moon rotates once around its axis for each rotation around the earth, so that it always shows us the same face.

Among small frequency ratios for coupled oscillators, the golden ratio is the least likely to lock in to a nearby rational number. Why?

2. Find the continued fraction expansion of $\sqrt{2}$. Show that if a number has a periodic continued fraction expansion then it satisfies a quadratic equation with integer coefficients. In fact, the converse is also true: if a number satisfies a quadratic equation with integer coefficients then it has a periodic continued fraction expansion. See for example Hardy and Wright [**52**], §10.12.

3. (Hua [**58**]) The synodic month is the period of time between two new moons, and is 29.5306 days. When projected onto the star sphere, the path of the moon intersects the ecliptic (the path of the sun) at the ascending and the descending nodes. A draconic month is the period of time for the moon to return to the same node, and is 27.2123 days. Show that the solar and lunar eclipses occur in cycles with a period of 18 years 10 days.

4. In this problem, you will prove that π is not equal to $\frac{22}{7}$. This problem is not really relevant to the text, but it is interesting anyway.

Use partial fractions (actually, just the long division part of the algorithm) to prove that

$$\int_0^1 \frac{x^4(1-x)^4\,dx}{1+x^2} = \tfrac{22}{7} - \pi.$$

Deduce that $\pi < \frac{22}{7}$. Show that

$$\int_0^1 x^4(1-x)^4\,dx = \tfrac{1}{630},$$

and use this to deduce that

$$\tfrac{1}{1260} < \tfrac{22}{7} - \pi < \tfrac{1}{630}.$$

"This must be Fibonacci's."

What would this sentence be like if π were 3?

> If π were equal to 3, this sentence
> would look something like this.

(Scott Kim/Harold Cooper, quoted from Douglas Hofstadter's *Metamagical Themas*, Basic Books, 1985).

5. Show that if a and b have no common factor then $\log_a(b)$ is irrational. Show that if no pair among a, b and c has a common factor then $\log_a(b)$ and $\log_a(c)$ are rationally independent. In other words, there cannot exist nonzero integers n_1, n_2 and n_3 such that $n_1 \log_a(b) + n_2 \log_a(c) = n_3$.

6. Find the continued fraction expansion for the rational number 531441/524288 which represents the frequency ratio for the Pythagorean comma. Explain in terms of this example the relationship between the continued fraction expansion of a rational number and Euclid's algorithm for finding highest common factors (if you don't remember how Euclid's algorithm goes, it is described in Lemma 9.7.1).

7. The *Gaussian integers* are the complex numbers of the form $a + bi$ where a and b are in the rational integers $\mathbb{Z}$. Develop a theory of continued fractions for simultaneously approximating two real numbers α and β, by considering the complex number $\alpha + \beta i$. Explain why this method favors denominators which can be expressed as a sum of two squares, so that it does not always find the best approximations.

8. A certain number is known to be a ratio of two 3-digit integers. Its decimal expansion, to nine significant figures, is 0.137637028. What are the integers?

Further reading:

G. H. Hardy and E. M. Wright, *Number theory* [**52**], chapter X.

Hua, *Introduction to number theory* [**58**], chapter 10.

Hubert Stanley Wall, *Analytic theory of continued fractions.* Chelsea, New York, 1948. ISBN 0828402078.

J. Murray Barbour, *Music and ternary continued fractions*, American Mathematical Monthly 55 (9) (1948), 545–555.

Viggo Brun, *Music and ternary continued fractions*, Norske Vid. Selsk. Forh., Trondheim 23 (1950), 38–40. This article is a response to the above article of Murray Barbour.

Viggo Brun, *Music and Euclidean algorithms*, Nordisk Mat. Tidskr. 9 (1961), 29–36, 95.

J. B. Rosser, *Generalized ternary continued fractions*, American Mathematical Monthly 57 (8) (1950), 528–535. This is another response to Murray Barbour's article.

Murray Schechter, *Tempered scales and continued fractions*, Amer. Math. Monthly 87 (1) (1980), 40–42.

Bosanquet's harmonium

6.3. Fifty-three tone scale

The first continued fraction expansion of interest to us is the one for $\log_2(3/2)$. The first few terms are

$$\log_2(3/2) = \frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{2+}\,\frac{1}{2+}\,\frac{1}{3+}\,\frac{1}{1+}\,\frac{1}{5+}\,\frac{1}{2+}\,\frac{1}{23+}\,\frac{1}{2+}\,\frac{1}{2+}\,\frac{1}{1+}\cdots$$

The sequence of convergents for the continued fraction expansion of $\log_2(3/2)$ is

$$1, \frac{1}{2}, \frac{3}{5}, \frac{7}{12}, \frac{24}{41}, \frac{31}{53}, \frac{179}{306}, \frac{389}{665}, \frac{9126}{15601}, \ldots$$

The bottoms of these fractions tell us how many equal notes to divide an octave into, and the tops tell us how many of these notes make up one approximate fifth. The fourth of the above approximations give us our western scale. The next obvious places to stop are at 31/53 and 389/665, just before large denominators.

The fifty-three tone equally tempered scale is interesting enough to warrant some discussion. In 1876, Robert Bosanquet made a "generalized keyboard harmonium" with fifty-three notes to an octave.[11] A photograph of this

[11]Described in Bosanquet, *Musical intervals and temperaments*, Macmillan and Co., London, 1876. Reprinted with commentary by Rudolph Rasch, Diapason Press, Utrecht, 1986.

instrument can be found on page 202. A discussion of this harmonium can be found in the translator's appendix XX.F.8 (pages 479–481) in Helmholtz [**54**]. One way of thinking of the fifty-three note scale is that it is based around the approximation which makes the Pythagorean comma equal to one fifty-third of an octave, or $1200/53 = 22.642$ cents, rather than the true value of 23.460 cents. So if we go around a complete circle of fifths, we get from C to a note which we may call B♯ 22.642 cents higher. This corresponds to the equation

$$12 \times 31 - 7 \times 53 = 1,$$

which can be interpreted as saying that twelve 53-tone equal temperament fifths minus seven octaves equals one step in the 53-tone scale.

The following table shows the fifty-three tone equivalents of the notes on the Pythagorean scale:

note	C	B♯	D♭	C♯	D	E♭	D♯	E	F	G♭
degree	0	1	4	5	9	13	14	18	22	26
note	F♯	G	A♭	G♯	A	B♭	A♯	C♭	B	C
degree	27	31	35	36	40	44	45	48	49	53

Thus the fifty-three tone scale is made up of five whole tones each of nine scale degrees and two semitones each of four scale degrees, $5\times 9+2\times 4 = 53$. Flattening or sharpening a note changes it by five scale degrees. The perfect fifth is extremely closely approximated in this scale by the thirty-first degree, which is

$$\frac{31}{53} \times 1200 = 701.887$$

cents rather than the true value of 701.955.

The just major third is also closely approximated in this scale by the seventeenth degree, which is

$$\frac{17}{53} \times 1200 = 384.906$$

cents rather than the true value of 386.314 cents. In effect, what is happening is that we are approximating both the Pythagorean comma and the syntonic comma by a single scale degree in the 53 note scale, which is roughly half way between them. So in Eitz's notation, we are identifying the note $\mathrm{G}\sharp^{0}$ with $\mathrm{A}\flat^{+1}$, whose difference is one schisma. Similarly, we are identifying the note B^{-1} with $\mathrm{C}\flat^{0}$, $\mathrm{B}\sharp^{-1}$ with C^{0}, and so on. We are also identifying the note G^{+2} with $\mathrm{A}\flat^{-2}$, whose difference is a diesis minus four commas, or

$$\frac{256}{243}\left(\frac{80}{81}\right)^4 = \frac{2^{24}5^4}{3^{21}} = \frac{10485760000}{10460353203},$$

or about 4.200 cents. The effect of this is that the array notation introduced in §5.9 becomes periodic in both directions, so that we obtain the diagram on page 204. In this diagram, the top and bottom row are identified with

22	0	31	9	40	18	49	27	5	
5	36	14	45	23	1	32	10	41	19
19	50	28	6	37	15	46	24	2	33
33	11	42	20	51	29	7	38	16	47
47	25	3	34	12	43	21	52	30	8
8	39	17	48	26	4	35	13	44	22
22	0	31	9	40	18	49	27	5	

F^{0}	C^{0}	G^{0}	D^{0}	A^{0}	E^{0}	B^{0}	$F\sharp^{0}$	$C\sharp^{0}$	
$C\sharp^{0}$	$A\flat^{+1}$	$E\flat^{+1}$	$B\flat^{+1}$	F^{+1}	C^{+1}	G^{+1}	D^{+1}	A^{+1}	E^{+1}
E^{+1}	B^{+1}	$F\sharp^{+1}$	$C\sharp^{+1}$	$A\flat^{+2}$	$E\flat^{+2}$	$B\flat^{+2}$	F^{+2}	C^{+2}	G^{+2}
G^{+2}	D^{+2}	A^{+2}	E^{+2}	C^{-2}	G^{-2}	D^{-2}	A^{-2}	E^{-2}	B^{-2}
B^{-2}	$F\sharp^{-2}$	$C\sharp^{-2}$	$A\flat^{-1}$	$E\flat^{-1}$	$B\flat^{-1}$	F^{-1}	C^{-1}	G^{-1}	D^{-1}
D^{-1}	A^{-1}	E^{-1}	B^{-1}	$F\sharp^{-1}$	$C\sharp^{-1}$	$A\flat^{0}$	$E\flat^{0}$	$B\flat^{0}$	F^{0}
F^{0}	C^{0}	G^{0}	D^{0}	A^{0}	E^{0}	B^{0}	$F\sharp^{0}$	$C\sharp^{0}$	

Torus of thirds and fifths in 53 tone equal temperament

each other, and the left and right walls are identified with each other. The resulting geometric figure is called a torus, and it looks like a bagel, or a tire.

It appears that the Pythagoreans were aware of the 53 tone equally tempered scale. Philolaus, a disciple of Pythagoras, thought of the tone as being divided into two minor semitones and a Pythagorean comma, and took each minor semitone to be four commas. This makes nine commas to the whole tone and four commas to the minor semitone, for a total of 53 commas to the octave. The Chinese theorist King Fâng of the third century B.C.

also seems to have been aware that the 54th note in the Pythagorean system is almost identical to the first.

After 53, the next good denominator in the continued fraction expansion of $\log_2(3/2)$ is 665. The extra advantages obtained by going to an equally tempered 665 tone scale, which is gives a remarkably good approximation to the perfect fifth, are far outweighed by the fact that adjacent tones are so close together (1.805 cents) as to be almost indistinguishable. If 53 tone equal temperament is thought of as the scale of commas, then 665 tone equal temperament might be thought of as a scale of schismas.

6.4. Other equal tempered scales

Other divisions of the octave into equal intervals which have been used for experimental tunings have included 19, 24, 31 and 43. The 19 tone scale has the advantage of excellent approximations to the 6:5 minor third and the 5:3 major sixth as well as reasonable approximations to the 5:4 major third and the 8:5 minor sixth. The eleventh degree gives an approximation to the 3:2 perfect fifth which is somewhat worse than in twelve tone equal temperament, but still acceptable.

name	ratio	cents	19-tone degree	cents
fundamental	1:1	0.000	0	0.000
minor third	6:5	315.641	5	315.789
major third	5:4	386.314	6	378.947
perfect fifth	3:2	701.955	11	694.737
minor sixth	8:5	813.687	13	821.053
major sixth	5:3	884.359	14	884.211
octave	2:1	1200.000	19	1200.000

Christiaan Huygens, in the late 17th century, seems to have been the first to use the equally tempered 19 tone scale as a way of approximating just intonation in a way that allowed for modulation into other keys. Yasser[12] was an important twentieth century proponent. The properties of 19 tone equal temperament with respect to formation of a diatonic scale are very similar to those for 12 tones. But accidentals and chromatic scales behave very differently.

The main purpose I can see for the equally tempered 24 tone scale, usually referred to as the *quarter-tone scale*, is that it increases the number of tones available without throwing out the familiar twelve tones. It contains no better approximations to the ratios 3:2 and 5:4 than the twelve tone scale, but has a marginally better approximation to 7:4 and a significantly better approximation to 11:8. The two sets of twelve notes formed by taking every other note from the 24 tone scale can be alternated with interesting effect, but using notes from both sets of twelve at once has a strong tendency to make discords. Examples of works using the quarter-tone scale include

[12]Joseph Yasser, *A theory of evolving tonality*, American Library of Musicology, New York, 1932.

the German composer Richard Stein's *Zwei Konzertstücke* op. 26, 1906 for cello and piano and Alois Hába's Suite for String Orchestra, 1917.[13] Twentieth century American composers such as Howard Hanson and Charles Ives have composed music designed for two pianos tuned a quarter tone apart.

Appendix E contains a table of various equal tempered scales, quantifying how well they approximate perfect fifths, just major thirds, and seventh harmonics. An examination of this table reveals that the 31 tone scale is unusually good at approximating all three at once. We examine this scale in the next section.

Further reading:

Jim Aikin, *Discover 19-tone equal temperament*, Keyboard, March 1988, p. 74–80.

M. Yunik and G. W. Swift, *Tempered music scales for sound synthesis*, Computer Music Journal 4 (4) (1980), 60–65.

Further listening: (See Appendix R)

Between the Keys, *Microtonal masterpieces of the 20th century*. This CD contains recordings of Charles Ives' *Three quartertone pieces*, and a piece by Ivan Vyshnegradsky in 72 tone equal temperament.

Easley Blackwood *Microtonal Compositions*. This is a recording of a set of microtonal compositions in each of the equally tempered scales from 13 tone to 24 tone.

Clarence Barlow's "OTOdeBLU" is in 17 tone equal temperament, played on two pianos.

Neil Haverstick, *Acoustic stick*. Played on custum built acoustic guitars tuned in 19 and 34 tone equal temperament.

William Sethares, *Xentonality*, Music in 10-, 13-, 17- and 19-tone equal temperament, using spectrally adjusted instruments.

6.5. Thirty-one tone scale

The 31 tone equal tempered scale was first investigated by Nicola Vicentino[14] and also later by Christiaan Huygens.[15] It gives a better approximation to the perfect fifth than the 19 tone scale, but it is still worse than the 12 tone scale.

[13] It is said that Hába practiced to the point where he could accurately sing five divisions to a semitone, or sixty to an octave.

[14] Nicola Vicentino, *L'antica musica ridotta alla moderna pratica*, Rome, 1555. Translated as *Ancient music adapted to modern practice*, Yale University Press, 1996.

[15] Christiaan Huygens, *Lettre touchant le cycle harmonique*, Letter to the editor of the journal *Histoire des Ouvrage de Sçavans*, Rotterdam 1691. Reprinted with English and Dutch translation (ed. Rudolph Rasch), Diapason Press, Utrecht, 1986.

name	ratio	cents	31-tone degree	cents
fundamental	1:1	0.000	0	0.000
major third	5:4	386.314	10	387.097
perfect fifth	3:2	701.955	18	696.774
minor sixth	8:5	813.687	21	812.903
seventh harmonic	7:4	968.826	25	967.742

It also contains good approximations to the major third and minor sixth, as well as the seventh harmonic.

The main reason for interest in 31-tone equal temperament is that note 18 of this scale is an unexpectedly good approximation to the meantone fifth (696.579) rather than the perfect fifth. So the entire meantone scale can be approximated as shown in the table below. Fokker[16] was an important twentieth century proponent of the 31 tone scale.

note	meantone	31-tone	
C	0.000	0	0.000
C♯	76.049	2	77.419
D	193.157	5	193.548
E♭	310.265	8	309.677
E	386.314	10	387.097
F	503.422	13	503.226
F♯	579.471	15	580.645
G	696.579	18	696.774
A♭	813.686	21	812.903
A	889.735	23	890.323
B♭	1006.843	26	1006.452
B	1082.892	28	1083.871
C	1200.000	31	1200.000

The picture on page 208 shows a 31 tone equal tempered instrument, made by Vitus Trasuntinis in 1606. Each octave has seven keys as usual where the white keys would normally go, and five sets of four keys where the five black keys would normally go. Then there are two keys each between the

[16]See for example A. D. Fokker, *The qualities of the equal temperament by 31 fifths of a tone in the octave*, Report of the Fifth Congress of the International Society for Musical Research, Utrecht, 3–7 July 1952, Vereniging voor Nederlandse Muziekgeschiedenis, Amsterdam (1953), 191–192; *Equal temperament with 31 notes*, Organ Institute Quarterly 5 (1955), 41; *Equal temperament and the thirty-one-keyed organ*, Scientific Monthly 81 (1955), 161–166. Also M. Joel Mandelbaum, *31-Tone Temperament: The Dutch Legacy*, Ear Magazine East, New York, 1982/1983; Henk Badings, *A. D. Fokker: new music with 31 notes*, Zeitschrift für Musiktheorie 7 (1976), 46–48.

Trasuntius' 31 tone clavicord (1606),
State Museum, Bologna, Italy

white keys that would normally not be separated by black keys, for a total of $7 + 4 \times 5 + 2 \times 2 = 31$.

Let us examine the relationship between the meantone scale and 31 tone equal temperament in terms of continued fractions. Since the meantone scale is generated by the meantone fifth, which represents a ratio of $\sqrt[4]{5} : 1$, we should look at the continued fraction for $\log_2(\sqrt[4]{5})$. We obtain

$$\begin{aligned}\log_2(\sqrt[4]{5}) &= \tfrac{1}{4}\log_2(5) = 0.580482024\ldots\\ &= \frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{2+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{5+}\,\frac{1}{1+}\cdots\end{aligned}$$

with convergents

$$\frac{0}{1}, \frac{1}{1}, \frac{1}{2}, \frac{3}{5}, \frac{4}{7}, \frac{7}{12}, \frac{11}{19}, \frac{18}{31}, \frac{101}{174}, \frac{119}{205}, \ldots$$

Cutting off just before the denominator 5 gives the approximation 18/31, which gives rise to the 31 tone equal tempered scale described above.

Exercises

1. Draw a torus of thirds and fifths, analogous to the one on page 204, for the 31 tone equal tempered scale, regarded as an approximation to meantone tuning.

2. In the text, the 31 tone equal tempered scale was compared with the usual (quarter comma) meantone scale, using the observation that taking multiples of the fifth generates a meantone scale, and then applying the theory of continued fractions to approximate the fifth. Carry out the same process to make the following comparisons.

(i) Compare the 19 tone equal tempered scale with Salinas' $\frac{1}{3}$ comma meantone scale.

(ii) Compare the 43 tone equal tempered scale with the $\frac{1}{5}$ comma meantone scale of Verheijen and Rossi.

(iii) Compare the 50 tone equal tempered scale with Zarlino's $\frac{2}{7}$ comma meantone scale.

(iv) Compare the 55 tone equal tempered scale with Silbermann's $\frac{1}{6}$ comma meantone scale.

Appendix J has a diagram which is relevant to this question.

6.6. The scales of Wendy Carlos

The idea behind the alpha, beta and gamma scales of Wendy Carlos is to ignore the requirement that there are a whole number of notes to an octave, and try to find equal tempered scales which give good approximations to the just intervals 3:2 and 5:4 (perfect fifth and major third). Since $6/5 = 3/2 \div 5/4$, this automatically gives good approximations to the 6:5 minor third. This means that we need $\log_2(3/2)$ and $\log_2(5/4)$ to be close to integer multiples of the scale degree. So we must find rational approximations to the ratio of these quantities.

We investigate the continued fraction expansion of the ratio:

$$\frac{\log_2(3/2)}{\log_2(5/4)} = \frac{\ln(3/2)}{\ln(5/4)} = 1 + \frac{1}{1+}\,\frac{1}{4+}\,\frac{1}{2+}\,\frac{1}{6+}\,\frac{1}{1+}\,\frac{1}{10+}\,\frac{1}{135+}\cdots$$

The sequence of convergents obtained by truncating this continued fraction is:

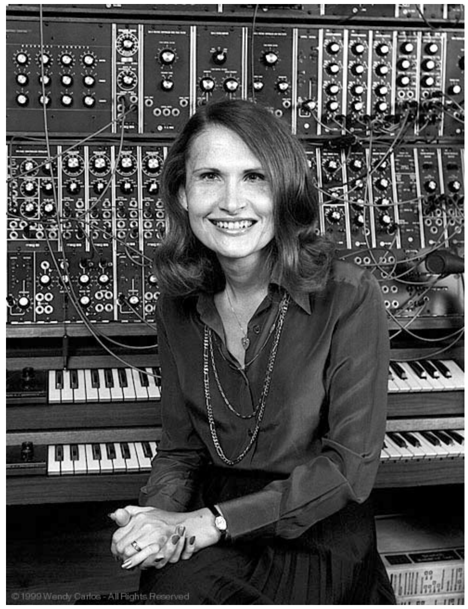
$$1,\ 2,\ \frac{9}{5},\ \frac{20}{11},\ \frac{129}{71},\ \frac{149}{82},\dots$$

Carlos' α (alpha) scale arises from the approximation 9/5 for the above ratio. This means taking a value for the scale degree so that nine of them approximate a 3:2 perfect fifth, five of them approximate a 5:4 major third, and four of them approximate a 6:5 minor third. In order to make the approximation as good as possible we minimize the mean square deviation. So if x denotes the scale degree (taking the octave as unit) then we must minimize

$$(9x - \log_2(3/2))^2 + (5x - \log_2(5/4))^2 + (4x - \log_2(6/5))^2.$$

Setting the derivative with respect to x of this quantity equal to zero, we obtain the equation

$$x = \frac{9\log_2(3/2) + 5\log_2(5/4) + 4\log_2(6/5)}{9^2 + 5^2 + 4^2} \approx 0.06497082462$$



Wendy Carlos, photo from her web site

Multiplying by 1200, we obtain a scale degree of 77.965 cents, and there are 15.3915 of them to the octave.[17]

Carlos also considers the scale α' obtained by doubling the number of notes in the octave. This gives the same approximations as before for the ratios 3:2, 5:4 and 6:5, but the twenty-fifth degree of the new scale (974.562 cents) is a good approximation to the seventh harmonic in the form of the ratio 7:4 (968.826 cents).

If instead we use the approximation

$$1+\frac{1}{1+}\,\frac{1}{5}=\frac{11}{6}$$

which we get by rounding up at the end instead of down, we obtain Carlos' β (beta) scale. We choose a value of the scale degree so that eleven of them approximate a 3:2 perfect fifth, six of them approximate a 5:4 major third, and five of them approximate a 6:5 minor third. Proceeding as before, we see that the proportion of an octave occupied by each scale degree is

$$\frac{11\log_2(3/2)+6\log_2(5/4)+5\log_2(6/5)}{11^2+6^2+5^2}\approx 0.05319411048.$$

Multiplying by 1200, we obtain a scale degree of 63.833 cents, and there are 18.7991 of them to the octave.[18] One advantage of the beta scale over the alpha scale is that the fifteenth scale degree (957.494 cents) is a reasonable

[17]This actually differs very slightly from Carlos' figure of 15.385 α-scale degrees to the octave. This is obtained by approximating the scale degree to 78.0 cents.

[18]Carlos has 18.809 β-scale degrees to the octave, corresponding to a scale degree of 63.8 cents.

approximation to the seventh harmonic in the form of the ratio 7:4 (968.826 cents). Indeed, it may be preferable to include this approximation into the above least squares calculation to get a scale in which the proportion of an octave occupied by each scale degree is

$$\frac{15\log_2(7/4) + 11\log_2(3/2) + 6\log_2(5/4) + 5\log_2(6/5)}{15^2 + 11^2 + 6^2 + 5^2} \approx 0.05354214235.$$

This gives a scale degree of 64.251 cents, and there are 18.677 of them to the octave. The fifteenth scale degree is then 963.759 cents.

Going one stage further, and using the approximation 20/11, we obtain Carlos' γ (gamma) scale. We choose a value of the scale degree so that twenty of them approximate a 3:2 perfect fifth, nine of them approximate a 5:4 major third, and eleven of them approximate a 4:3 minor third. The proportion of an octave occupied by each scale degree is

$$\frac{20\log_2(3/2) + 11\log_2(5/4) + 9\log_2(6/5)}{20^2 + 11^2 + 9^2} \approx 0.02924878523.$$

Multiplying by 1200, we obtain a scale degree of 35.099 cents, and there are 34.1895 of them to the octave.[19] This scale contains almost pure perfect fifths and major thirds, but it does not contain a good approximation to the ratio 7:4.

name	ratio	cents	α	cents	β	cents	γ	cents
fundamental	1:1	0.000	0	0.000	0	0.000	0	0.000
minor third	6:5	315.641	4	311.860	5	319.165	9	315.887
major third	5:4	386.314	5	389.825	6	382.998	11	386.084
perfect fifth	3:2	701.955	9	701.685	11	702.162	20	701.971
seventh harmonic	7:4	968.826	$12\frac{1}{2}$	974.562	15	957.494	–	— —

6.7. The Bohlen–Pierce scale

Jaja, unlike Stravinsky, has never been guilty of composing harmony in all his life. Jaja is pure absolute twelve tone. Never tempted, like some of the French composers, to write with thirteen tones. Oh no. This, says Jaja, is the baker's dozen, the "Nadir of Boulanger."

From Gerard Hoffnung's Interplanetary Music Festival, analysis by two "distinguished teutonic musicologists" of the work of a fictitious twelve tone composer, Bruno Heinz Jaja.

The Bohlen–Pierce scale is the thirteen tone scale described in the article of Mathews and Pierce, forming Chapter 13 of [**82**]. Like the scales of Wendy Carlos, it is not based around the octave as the basic interval. But whereas Carlos uses 3:2 and 5:4, Bohlen and Pierce replace the octave by an

[19] Carlos has 34.188 γ-scale degrees to the octave, corresponding to a scale degree of 35.1 cents.

octave and a perfect fifth (a ratio of 3:1). In the equal tempered version, this is divided into thirteen equal parts. This gives a good approximation to a "major" chord with ratios 3:5:7. The idea is that only odd multiples of frequencies are used. Music written using this scale works best if played on an instrument such as the clarinet, which involves predominantly odd harmonics, or using specially created synthetic voices with the same property. We shall prefix all words associated with the Bohlen–Pierce scale with the letters BP to save confusion with the corresponding notions based around the octave.

The basic interval of an octave and a perfect fifth, which is a ratio of exactly 3:1 or an interval of 1901.955 cents, is called a *BP-tritave*. In the equal tempered 13 tone scale, each scale degree is one thirteenth of this, or 146.304 cents. It may be felt that the scale of cents is inappropriate for calculations with reference to this scale, but we shall stick with it nonetheless for comparison with intervals in scales based around the octave.

The Pythagorean approach to the division of the tritave begins with a ratio of 7:3 as the analog of the fifth. We shall call this interval the perfect BP-tenth, since it will correspond to note ten in the BP-scale. The corresponding continued fraction is

$$\log_3(7/3) = \frac{1}{1+}\,\frac{1}{3+}\,\frac{1}{2+}\,\frac{1}{1+}\,\frac{1}{2+}\,\frac{1}{4+}\,\frac{1}{22+}\,\frac{1}{32+}\cdots,$$

whose convergents are

$$\frac{0}{1},\frac{1}{1},\frac{3}{4},\frac{7}{9},\frac{10}{13},\frac{27}{35},\frac{118}{153},\dots$$

If we perform the same calculation for the 5:3 ratio, we obtain the continued fraction

$$\log_3(5/3) = \frac{1}{2+}\,\frac{1}{6+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{3+}\,\frac{1}{7+}\cdots$$

with convergents

$$\frac{0}{1},\frac{1}{2},\frac{6}{13},\frac{7}{15},\frac{13}{28},\frac{20}{43},\frac{73}{157},\dots.$$

Comparing these continued fractions, it looks like a good idea to divide the tritave into 13 equal intervals, with note 10 approximating the ratio 7:3, and note 6 approximating the ratio 5:3.

note	degree	7/3-Pythag	Just
C	0	1:1	1:1
D	2	19683:16807	25:21
E	3	9:7	9:7
F	4	343:243	7:5
G	6	81:49	5:3
H	7	49:27	9:5
J	9	729:343	15:7
A	10	7:3	7:3
B	12	6561:2401	25:9
C	13	3:1	3:1

Basing a BP-Pythagorean scale around the ratio 7:3, we obtain a scale of 13 notes in which the circle of BP-tenths has a BP 7/3-comma given by a ratio of

$$\frac{7^{13}}{3^{23}} = \frac{96889010407}{94143178827}$$

or about 49.772 cents.

Using perfect BP-tenths to form a diatonic BP-Pythagorean scale, we obtain the third column of the table to the left. Following Bohlen, we name the notes of the scale using the letters A–H and J. Note that our choice of the second degree of the diatonic scale differs from the choice made by Mathews and Pierce, and gives what Bohlen calls the Lambda scale.

To obtain a major 3:5:7 triad, we introduce a just major BP-sixth with a ratio of 5:3. This is very close to the BP 7/3-Pythagorean G, which gives rise to an interval called the BP-minor diesis, expressing the difference between these two versions of G. This interval, namely the difference between 5:3 and 81:49, is a ratio of 245:243 or about 14.191 cents.

The BP version of Eitz's notation works in a similar way to the octave version. We start with the BP 7/3-Pythagorean values for the notes and then adjust by a number of BP-minor dieses indicated by a superscript. So G^{0} denotes the 81:49 version of G, while G^{+1} denotes the 5:3 version. The just scale given in the table above is then described by the following array:

$$\begin{array}{ccccc}
 & & \mathrm{D}^{+2} & & \mathrm{B}^{+2} \\
 & \mathrm{J}^{+1} & & \mathrm{G}^{+1} & \\
\mathrm{E}^{0} & & \mathrm{C}^{0} & & \mathrm{A}^{0} \\
 & \mathrm{H}^{-1} & & \mathrm{F}^{-1} &
\end{array}$$

A reasonable way to fill this in to a thirteen tone just scale is as follows:

BP Monochord

$F\sharp^{+2}$ D^{+2} B^{+2}

J^{+1} G^{+1}

E^{0} C^{0} A^{0}

H^{-1} F^{-1}

$D\flat^{-2}$ $B\flat^{-2}$ $J\flat^{-2}$

For comparison, here is a table of the scales discussed above, in cents to three decimal places, and also in the BP version of Eitz's notation. The column marked "discrepancy" gives the difference between the equal and just versions.

	BP 7/3-Pythag		BP-just		BP-equal	discrepancy
C	0.000	0	0.000	0	0.000	0.000
D♭	161.619	0	133.238	−2	146.304	+13.066
D	273.465	0	301.847	+2	292.608	−9.239
E	435.084	0	435.084	0	438.913	+3.829
F	596.703	0	582.512	−1	585.217	+2.705
F♯	708.550	0	736.931	+2	731.521	−5.410
G	870.168	0	884.359	+1	877.825	−6.534
H	1031.787	0	1017.596	−1	1024.130	+6.534
J♭	1193.405	0	1165.024	−2	1170.434	+5.410
J	1305.252	0	1319.443	+1	1316.738	−2.705
A	1466.871	0	1466.871	0	1463.042	−3.829
B♭	1628.490	0	1600.108	−2	1609.347	+9.239
B	1740.336	0	1768.717	+2	1755.651	−13.066
C	1901.955	0	1901.955	0	1901.955	0.000

A number of the intervals in the BP scale approximate intervals in the usual octave based scale, and some of these approximations are just far enough off to be disturbing to trained musicians. It is plausible that proper appreciation of music written in the BP scale would involve learning to "forget" the accumulated experience of the perpetual bombardment by octave based music which we receive from the world around us, even if we are not musicians. For this reason, it seems unlikely that such music will become popular. On the other hand, according to John Pierce (chapter 1 of [**32**]), Maureen Chowning, a coloratura soprano, has learned to sing in the BP scale, Richard Boulanger has composed a "considerable piece" using it, and two CDs by Charles Carpenter are available which make extensive use of scale (see below).

Exercises

1. (Paul Erlich) Investigate the refinement of the Bohlen–Pierce scale in which there are 39 tones to the BP-tritave. What relevant ratios are approximated by scale degrees 5, 7, 11, 13, 16, 22, 28 and 34?

Further reading:

Heinz Bohlen, *13 Tonstufen in der Duodezeme.* Acustica 39 (1978), 76–86.

M. V. Mathews and J. R. Pierce, *The Bohlen–Pierce scale.* Chapter 13 of [**82**].

M. V. Mathews, J. R. Pierce, A. Reeves and L. A. Roberts, *Theoretical and experimental explorations of the Bohlen–Pierce scale*. J. Acoust. Soc. Am. 84 (1988), 1214–1222.

M. V. Mathews, L. A. Roberts and J. R. Pierce, *Four new scales based on non-successive-integer-ratio chords*. J. Acoust. Soc. Am. 75 (1984), S10(A).

L. A. Roberts and M. V. Mathews, *Intonation sensitivity for traditional and non-traditional chords*. J. Acoust. Soc. Am. 75 (1984), 952–959.

Further listening: (See Appendix R)

Charles Carpenter, *Frog à la Pêche* and *Splat* are composed using the Bohlen–Pierce scale, and played in a progressive rock/jazz style.

On the CD of examples accompanying Cook [**21**], track 62 demonstrates the Bohlen–Pierce scale.

On the CD of examples accompanying Mathews and Pierce [**82**], tracks 71–74 demonstrate the Bohlen–Pierce scale.

6.8. Unison vectors and periodicity blocks

In this section, we return to just intonation, and we describe Fokker's periodicity blocks and unison vectors. The periodicity block corresponds to what a mathematician would call a *set of coset representatives*, or a *fundamental domain*. The starting point is octave equivalence; notes differing by a whole number of octaves are considered to be equivalent.

The Pythagorean scale is the one dimensional version of the theory. We place the notes of the Pythagorean scale along a one dimensional lattice, with the origin at C^0. We have labeled the vertices in three notations, namely note names, ratios and cents, for comparison. Because of octave equivalence, the value in cents is reduced or augmented by a multiple of 1200 as necessary to put it in the interval between zero and 1200.

792.180	294.135	996.090	498.045	0.000	701.955	203.910	905.865	407.820	1109.775	611.731	113.685
128/81	32/27	16/9	4/3	1/1	3/2	9/8	27/16	81/64	243/128	729/512	2187/2048
$\mathrm{A}\flat^0$	$\mathrm{E}\flat^0$	$\mathrm{B}\flat^0$	F^0	C^0	G^0	D^0	A^0	E^0	B^0	$\mathrm{F}\sharp^0$	$\mathrm{C}\sharp^0$

We write $\mathbb{Z}^1$ for this one dimensional lattice, to emphasize the fact that the points in the lattice may be indexed using the integers. In fact, given the choice of C^0 as the origin, there are two sensible ways to index with integers. One makes G^0 correspond to 1, D^0 to 2, F^0 to -1, etc., and the other makes F^0 correspond to 1, $\mathrm{B}\flat^0$ to 2, G^0 to -1, etc. For the sake of definiteness, we choose the former.

The twelve tone Pythagorean scale comes from observing that in this system, C^0 and $\mathrm{B}\sharp^0$ are close enough together in pitch that we may not want both of them in our scale. Since $\mathrm{B}\sharp^0$ is the twelfth note, we say that (12) is a unison vector. A periodicity block would then consist of a choice of 12 consecutive points on this lattice, to constitute a scale. Other choices of unison vector would include (53) and (665) (cf. §6.3).

Just intonation, at least the 5-limit version as we introduced it in §5.5, is really a 2-dimensional lattice, which we write as $\mathbb{Z}^2$. In Eitz's notation (see §5.9), here is a small part of the lattice with the origin circled.

$$\begin{gathered}
\mathrm{F}\sharp^{-2}\quad \mathrm{C}\sharp^{-2}\quad \mathrm{G}\sharp^{-2}\quad \mathrm{D}\sharp^{-2} \\
\mathrm{D}^{-1}\quad \mathrm{A}^{-1}\quad \mathrm{E}^{-1}\quad \mathrm{B}^{-1}\quad \mathrm{F}\sharp^{-1} \\
\mathrm{B}\flat^{0}\quad \mathrm{F}^{0}\quad \mathrm{C}^{0}\quad \mathrm{G}^{0}\quad \mathrm{D}^{0}\quad \mathrm{A}^{0} \\
\mathrm{D}\flat^{+1}\quad \mathrm{A}\flat^{+1}\quad \mathrm{E}\flat^{+1}\quad \mathrm{B}\flat^{+1}\quad \mathrm{F}^{+1}\quad \mathrm{C}^{+1} \\
\mathrm{F}\flat^{+2}\quad \mathrm{C}\flat^{+2}\quad \mathrm{G}\flat^{+2}\quad \mathrm{D}\flat^{+2}\quad \mathrm{A}\flat^{+2}
\end{gathered}$$

The same in ratio notation is as follows.

$$\begin{gathered}
\tfrac{25}{18}\quad \tfrac{25}{24}\quad \tfrac{25}{16}\quad \tfrac{75}{64} \\
\tfrac{10}{9}\quad \tfrac{5}{3}\quad \tfrac{5}{4}\quad \tfrac{15}{8}\quad \tfrac{45}{32} \\
\tfrac{16}{9}\quad \tfrac{4}{3}\quad \tfrac{1}{1}\quad \tfrac{3}{2}\quad \tfrac{9}{8}\quad \tfrac{27}{16} \\
\tfrac{16}{15}\quad \tfrac{8}{5}\quad \tfrac{6}{5}\quad \tfrac{9}{5}\quad \tfrac{27}{20}\quad \tfrac{81}{80} \\
\tfrac{32}{25}\quad \tfrac{48}{25}\quad \tfrac{36}{25}\quad \tfrac{27}{25}\quad \tfrac{81}{50}
\end{gathered}$$

We can choose a basis for this lattice, and write everything in terms of vectors with respect to this basis. This is the two dimensional version of our choice from two different ways of indexing the Pythagorean scale by the integers, but this time there are an infinite number of choices of basis.

For example, if our basis consists of G^0 and E^{-1} (i.e., $\frac{3}{2}$ and $\frac{5}{4}$) then here is the same part of the lattice in vector notation.

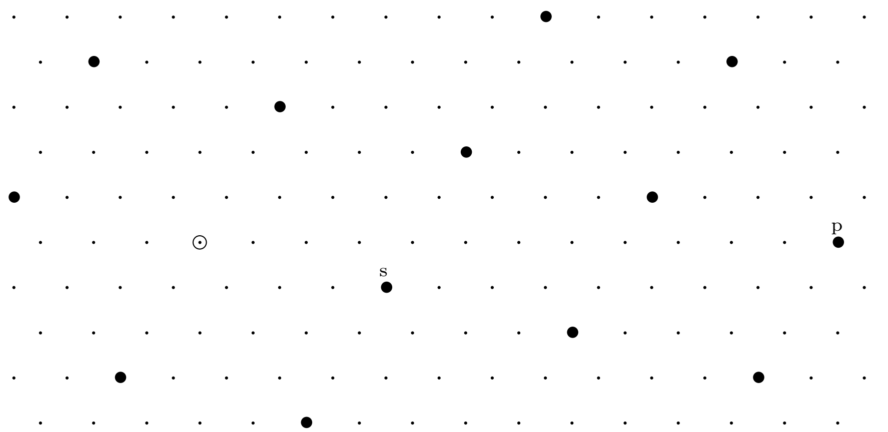
$$\begin{gathered}
(-2,\ 2)\quad (-1,\ 2)\quad (\ 0,\ 2)\quad (\ 1,\ 2) \\
(-2,\ 1)\quad (-1,\ 1)\quad (\ 0,\ 1)\quad (\ 1,\ 1)\quad (\ 2,\ 1) \\
(-2,\ 0)\quad (-1,\ 0)\quad (\ 0,\ 0)\quad (\ 1,\ 0)\quad (\ 2,\ 0)\quad (\ 3,\ 0) \\
(-1,-1)\quad (\ 0,-1)\quad (\ 1,-1)\quad (\ 2,-1)\quad (\ 3,-1)\quad (\ 4,-1) \\
(\ 0,-2)\quad (\ 1,-2)\quad (\ 2,-2)\quad (\ 3,-2)\quad (\ 4,-2)
\end{gathered}$$

The defining property of a basis is that every vector in the lattice has a unique expression as an integer combination of the basis vectors. The number of vectors in a basis is the dimension of the lattice.

Now we need to choose our unison vectors. The classical choice here is $(4,-1)$ and $(12,0)$, corresponding to the syntonic comma and the Pythagorean comma. The sublattice *generated* by these unison vectors consists of all linear combinations

$$m(4,-1)+n(12,0)=(4m+12n,-4m)$$

with $m,n\in\mathbb{Z}$. This is called the *unison sublattice*.



In this diagram, the syntonic comma and Pythagorean comma are marked with s and p respectively. Each vector (a,b) in the lattice may then be thought of as *equivalent* to the vectors

$$(a,b)+m(4,-1)+n(12,0)=(a+4m+12n,b-4m)$$

with $m,n\in\mathbb{Z}$, differing from it by vectors in the unison sublattice. So for example, taking $m=-3$ and $n=1$, we see that the vector $(0,3)$ is in the unison sublattice. This corresponds to the fact that three just major thirds approximately make one octave.

There are many ways of choosing unison vectors generating a given sublattice. In the above example, $(4,-1)$ and $(0,3)$ generate the same sublattice.

The set of vectors (or pitches) equivalent to a given vector is called a *coset*. The number of cosets is called the *index* of the unison sublattice in the lattice. It can be calculated by taking the *determinant* of the matrix formed from the unison vectors. So in our example, the index of the unison sublattice is

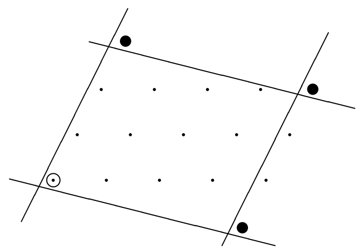
$$\begin{vmatrix} 4 & -1 \\ 12 & 0 \end{vmatrix} = 12.$$

The formula for the determinant of a 2×2 matrix is

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad-bc.$$

If the determinant comes out negative, the index is the corresponding positive quantity. If two rows of a matrix are swapped, then the determinant changes sign, so the sign of the determinant is irrelevant to the index. It has to do with *orientation*, and will not be discussed here.

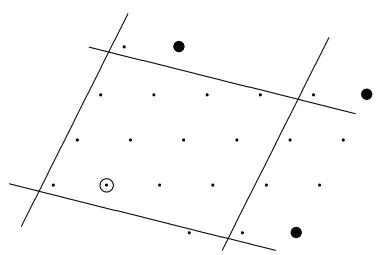
A *periodicity block* consists of a choice of one vector from each coset. In other words, we find a finite set of vectors with the property that each vector in the whole lattice is equivalent to a unique vector from the periodicity block. One way to do this is to draw a parallelogram using the unison vectors. We can then tile the plane using copies of this parallelogram, translated along unison vectors. In the above example, if we use the unison vectors $(4, -1)$ and $(0, 3)$ to generate the unison sublattice, then the parallelogram looks like this.



This choice of periodicity block leads to the following just scale with twelve tones.

$$\begin{array}{cccc} & \mathrm{G}\sharp^{-2} & \mathrm{D}\sharp^{-2} & \mathrm{A}\sharp^{-2} & \mathrm{E}\sharp^{-2} \\ \mathrm{E}^{-1} & \mathrm{B}^{-1} & \mathrm{F}\sharp^{-1} & \mathrm{C}\sharp^{-1} & \\ \mathrm{C}^{0} \; \mathrm{G}^{0} & \mathrm{D}^{0} & \mathrm{A}^{0} & & \end{array}$$

Of course, there are many other choices of periodicity block. For example, shifting this parallelogram one place to the left gives rise to Euler's monochord, described on page 162.



Periodicity blocks do not have to be parallelograms. For example, we can chop off a corner of the parallelogram, translate it through a unison vector, and stick it back on somewhere else to get a hexagon. Each of the just intonation scales in §5.10 may be interpreted as a periodicity block for the above choice of unison sublattice.

Of course, there are other choices of unison sublattices. If we choose the unison vectors $(4, -1)$ and $(-1, 5)$ for example, then we get a scale of

$$\begin{vmatrix} 4 & -1 \\ -1 & 5 \end{vmatrix} = 19$$

tones. This gives rise to just scales approximating the equal tempered scale described at the beginning of §6.4. The choice of $(4, 2)$ and $(-1, 5)$ gives the

Indian scale of 22 Srutis described in §6.1, corresponding to the calculation

$$\begin{vmatrix} 4 & 2 \\ -1 & 5 \end{vmatrix} = 22.$$

Taking the unison vectors $(4,-1)$ and $(3,7)$ gives rise to 31 tone scales approximating 31 tone equal temperament, whose relationship with meantone is described in §6.4. This corresponds to the calculation

$$\begin{vmatrix} 4 & -1 \\ 3 & 7 \end{vmatrix} = 31$$

The vectors $(8,1)$ and $(-5,6)$ correspond in the same way to just scales approximating the 53 tone equal tempered scale described in §6.3, corresponding to the calculation

$$\begin{vmatrix} 8 & 1 \\ -5 & 6 \end{vmatrix} = 53.$$

An example of a periodicity block for this choice of unison vectors can be found on page 204.

When we come to study groups and normal subgroups in §9.12, we shall make some more comments on how to interpret unison vectors and periodicity blocks in group theoretical language.

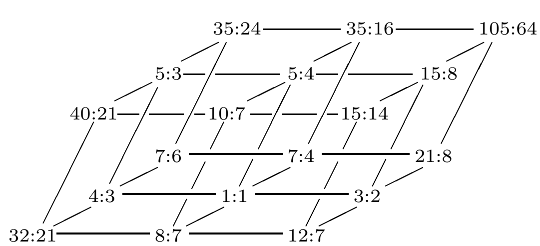
Further reading:

Paul Erlich, sonic-arts.org/td/erlich/intropblock1.htm
Some of the material in this and the next section expresses ideas from Paul's online article, together with the work of Fokker.

A. D. Fokker, *Selections from the harmonic lattice of perfect fifths and major thirds containing 12, 19, 22, 31, 41 or 53 notes*, Proc. Koninkl. Nederl. Akad. Wetenschappen, Series B, 71 (1968), 251–266.

6.9. Septimal harmony

Septimal harmony refers to 7-limit just intonation; in other words, just intonation involving the primes 2, 3, 5 and 7. Taking octave equivalence into account, this means that we need three dimensions, or $\mathbb{Z}^3$ to represent the septimal version of just intonation, to take account of the primes 3, 5 and 7. It is harder to draw a three dimensional lattice, but it can be done. In ratio notation, it will then look as follows.



We take as our basis vectors the ratios $\frac{3}{2}$, $\frac{5}{4}$ and $\frac{7}{4}$. So the vector (a, b, c) represents the ratio $3^a.5^b.7^c$, multiplied if necessary by a power of 2 so that it is between 1 and 2 (octave equivalence). The septimal comma introduced in §5.8 is a ratio of 64 : 63, which corresponds to the vector $(-2, 0, -1)$. So it would be reasonable to use the three commas $(4, -1, 0)$ (syntonic), $(12, 0, 0)$ (Pythagorean), and $(-2, 0, -1)$ (septimal) as unison vectors.

The determinant of a 3×3 matrix

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

is given by the formula

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - ceg - bdi - afh.$$

This can be visualized as three leading diagonals minus three trailing diagonals. If you have trouble visualizing these diagonals, it may help to think of the matrix as wrapped around a cylinder. So you should write the first two columns of the matrix again to the right of the matrix, and then the leading and trailing diagonals really look diagonal.

With the three commas as unison vectors, the determinant is

$$\begin{vmatrix} 4 & -1 & 0 \\ 12 & 0 & 0 \\ -2 & 0 & -1 \end{vmatrix} = 0 + 0 + 0 - 0 - 12 - 0 = -12.$$

Ignoring signs as usual, this tells us that we should expect the periodicity block to have 12 elements. One choice of periodicity block gives the 7-limit just intonation diagram on page 192.

There are many choices of unison vector in 7-limit just intonation. The table on page 221, adapted from Fokker, gives some of the most useful ones. Fokker also develops an elaborate system of notation for 7-limit just intonation, in which he ends up with notes such as $\diagdown f\!\downharpoonleft\!\downharpoonleft$.

Further reading:

A. D. Fokker, *Unison vectors and periodicity blocks in the three-dimensional (3-5-7-) harmonic lattice of notes*, Proc. Koninkl. Nederl. Akad. Wetenschappen, Series B, 72 (1969), 153–168.

vector	ratio	cents
(−7 4 1)	$\frac{4375}{4374}$	0.40
(−1 −2 4)	$\frac{2401}{2400}$	0.72
(−8 2 5)	$\frac{420175}{419904}$	1.12
(9 3 −4)	$\frac{2460375}{2458624}$	1.23
(8 1 0)	$\frac{32805}{32764}$	1.95
(1 5 1)	$\frac{65625}{65536}$	2.35
(0 3 5)	$\frac{2100875}{2097152}$	3.07
(−8 −6 2)	$\frac{102760448}{102515625}$	4.13
(1 −3 −2)	$\frac{6144}{6125}$	5.36
(0 −5 2)	$\frac{3136}{3125}$	6.08
(−7 −1 3)	$\frac{10976}{10935}$	6.48
(2 2 −1)	$\frac{225}{224}$	7.71
(8 −4 2)	$\frac{321489}{320000}$	8.04
(−5 6 0)	$\frac{15625}{15552}$	8.11
(1 0 3)	$\frac{1029}{1024}$	8.43
(3 7 0)	$\frac{2109375}{2083725}$	10.06
(−5 −2 −3)	$\frac{2097152}{2083725}$	11.12
(3 −1 −3)	$\frac{1728}{1715}$	13.07
(−4 3 −2)	$\frac{4000}{3969}$	13.47
(2 −3 1)	$\frac{126}{125}$	13.79
(−5 1 2)	$\frac{245}{243}$	14.19
(10 −2 1)	$\frac{413343}{409600}$	15.75
(3 2 2)	$\frac{33075}{32768}$	16.14
(−3 0 −4)	$\frac{65536}{64827}$	18.81
(3 −6 −1)	$\frac{110592}{109375}$	19.16
(−4 −2 0)	$\frac{2048}{2025}$	19.55
(5 1 −4)	$\frac{2430}{2401}$	20.79
(4 −1 0)	$\frac{81}{80}$	21.51
(−3 3 1)	$\frac{875}{864}$	21.90
(12 0 0)	$\frac{531441}{524288}$	23.46
(5 4 1)	$\frac{1063125}{1048576}$	23.86
(4 2 5)	$\frac{34034175}{33554432}$	24.58
(−3 −5 −2)	$\frac{4194302}{4134375}$	24.91
(−10 −1 −1)	$\frac{2097152}{2066715}$	25.31
(5 −4 −2)	$\frac{31104}{30625}$	26.87
(−2 0 −1)	$\frac{64}{63}$	27.26
(−3 −2 3)	$\frac{686}{675}$	27.99
(−1 5 0)	$\frac{3125}{3072}$	29.61
(−2 3 4)	$\frac{303125}{294912}$	30.33
(−1 −3 −3)	$\frac{131072}{128625}$	32.63
(−8 1 −2)	$\frac{327680}{321989}$	33.02
(−9 −1 2)	$\frac{100352}{98415}$	33.74
(0 2 −2)	$\frac{50}{49}$	34.98
(−1 0 2)	$\frac{49}{48}$	35.70
(1 7 −1)	$\frac{234375}{229376}$	37.33
(7 1 2)	$\frac{535815}{524288}$	37.65
(0 5 3)	$\frac{1071875}{1048576}$	38.05
(1 −1 −4)	$\frac{12278}{12005}$	40.33
(0 −3 0)	$\frac{128}{125}$	41.06
(−7 1 1)	$\frac{2240}{2187}$	41.45
(2 4 −1)	$\frac{5625}{5488}$	42.69
(1 2 1)	$\frac{525}{512}$	43.41
(0 0 5)	$\frac{16807}{16384}$	44.13
(1 −6 −2)	$\frac{786432}{765625}$	46.42
(−6 −2 −1)	$\frac{131072}{127575}$	46.81
(2 −1 −1)	$\frac{36}{35}$	48.77
(−6 1 4)	$\frac{12005}{11664}$	49.89
(2 2 4)	$\frac{540225}{524288}$	51.84
(2 −6 1)	$\frac{16128}{15625}$	54.85
(−5 −2 2)	$\frac{6272}{6075}$	55.25
(4 1 −2)	$\frac{405}{392}$	56.48
(3 −1 2)	$\frac{1323}{1280}$	57.20
(−4 3 3)	$\frac{42875}{42472}$	57.60
(4 −4 0)	$\frac{648}{625}$	62.57
(−3 0 1)	$\frac{28}{27}$	62.96
(−1 2 0)	$\frac{25}{24}$	70.72
(1 −1 1)	$\frac{21}{20}$	84.42
(3 1 0)	$\frac{135}{128}$	92.23
(−3 3 1)	$\frac{3584}{3575}$	104.02
(−1 4 −2)	$\frac{625}{588}$	105.65

CHAPTER 7

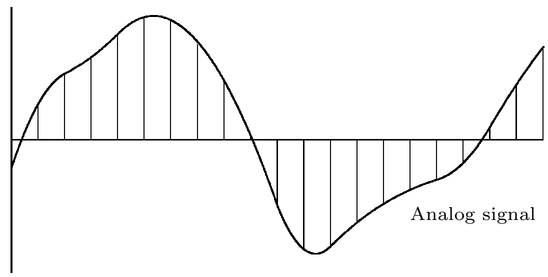
Digital music

7.1. Digital signals

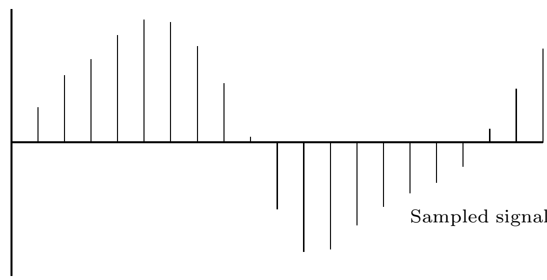
The commonest method of digital representation of sound is about as simple minded as you can get. To digitize an analog signal, the signal is sampled a large number of times a second, and a binary number represents the height of the signal at each sample time. Both of these processes are sometimes referred to as *quantization* (don't worry, there's no quantum mechanics involved here), but it is important to realize that the processes are separate, and need to be understood separately.

For example, the Compact Disc is based on a sample rate of 44.1 KHz, or 44,100 sample points per second.[1] At each sample point, a sixteen digit binary number represents the height of the waveform at that point. The computer WAV file format is another example, which we shall describe in detail in §7.3.

The following diagrams illustrate the process. The first picture shows the original analog signal.



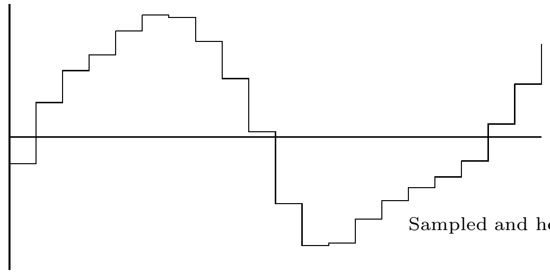
Next we have the sampled signal, but still with continuously variable amplitudes.



If we apply a “sample and hold” to the signal, we obtain a staircase waveform.

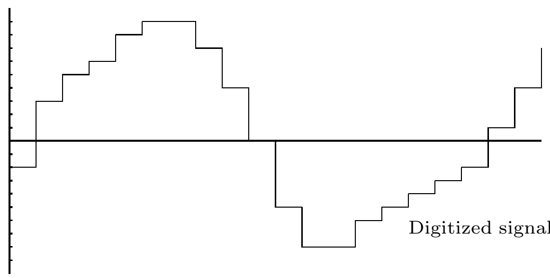
[1]It is annoying that the default sample rate for DAT (Digital Audio Tape) is 48 KHz, thereby making it difficult to make a digital copy on CD directly from DAT. This seems to be the result of industry paranoia at the idea that anyone might make a digital copy of music from a CD (DAT was originally designed as a consumer format, but never took off except among the music business professionals). The excuse that the higher sample rate for DAT gives a higher cutoff frequency and therefore better audio fidelity is easily seen through in light of the fact that the improvement is about three quarters of a tone, which is essentially insignificant.

Fortunately, the ratio $48,000/44,100$ can be written as a product of small fractions, $4/3 \times 8/7 \times 5/7$, which suggests an easy method of digital convertion. To multiply the sample rate by $4/3$, for example, we use linear interpolation to quadruple the sample rate and then omit two out of every three sample points. This gives much better fidelity than converting to an analog signal and then back to digital.



We shall see at the end of §7.6 that provided that the original analog signal has no frequency components at half the sample rate or above (this is achieved with a low pass filter), it may be reconstructed exactly from this sampled signal. This rather extraordinary statement is called the *sampling theorem*, and understanding it requires an understanding of what sampling does to the Fourier spectrum of a signal. We shall do this by systematically making use of Dirac delta functions, starting in §7.5.

Finally, if we digitize the samples, each sample value gets adjusted to the nearest allowed value.



This part of the process of turning an analog signal into a digital signal does entail some loss of information, even if the original signal contains no frequency component above half the sample rate. To see this, think what happens to a very low level signal. It will simply get reported as zero. There is a method for overcoming this limitation to a certain extent, called *dithering*, and it is described in §7.2.

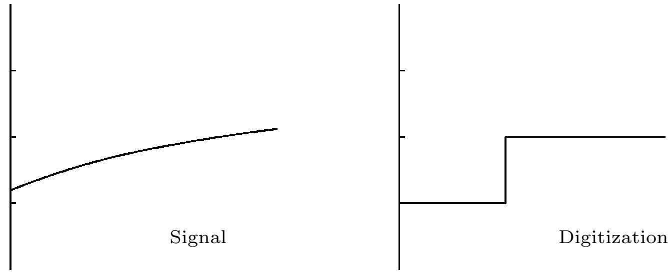
Further reading:

Ken C. Pohlmann, *The compact disc handbook, 2nd edition*, A-R Editions, Inc., Madison, Wisconsin, 1992.

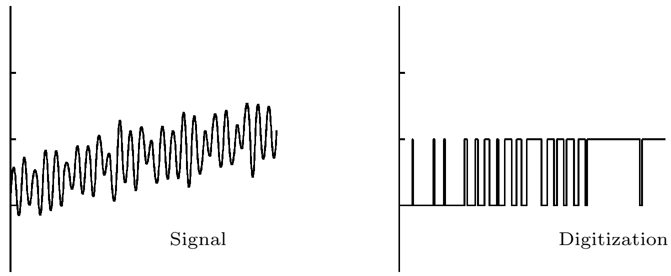
7.2. Dithering

Dithering is a method of decreasing the distortion of a low level signal due to digitization of signal level. This is based on the audacious proposition that adding a low level source of random noise to a signal can increase the signal resolution. This works best when the sample rate is high in comparison with the rate at which the signal is changing.

To see how this works, consider a slowly varying signal and its digitization.



Now if we add noise to the original signal at amplitude roughly one half the step size in the digitization process, here's what the signal looks like.



If the digitized signal is put through a resistor-capacitor circuit to smooth it out, some reasonable approximation to the original signal can be recovered. There is no theoretical limit to the accuracy possible with this method, as long as the sampling rate is high enough.

Further reading:

J. Vanderkooy and S. Lipshitz, *Resolution below the least significant bit in digital audio systems with dither*, J. Audio Eng. Soc. 32 (3) (1984), 106–113; *Correction*, J. Audio Eng. Soc. 32 (11) (1984), 889.

7.3. WAV and MP3 files

A common format for digital sound files on a computer is the WAV format. This is an example of Resource Interchange File Format (RIFF) for multimedia files; another example of RIFF is the AVI movie format. Here is what a WAV file looks like. The file begins with some header information, which comes in a 12 byte RIFF chunk and a 24 byte FORMAT chunk, and then the actual wave data, which comes in a DATA chunk occupying the rest of the file.

The binary numbers in a WAV file are always *little endian*, which means that the least significant byte comes first, so that the bytes are in the reverse of what might be thought of as the normal order. We shall represent binary numbers using *hexadecimal*, format, or base 16. Each hexadecimal digit encodes four binary digits, so that there are two hexadecimal digits in a byte. The sixteen symbols used are 0–9 and A–F. So for example in little endian format, 4E 02 00 00 would represent the hexadecimal number 24E, which is the binary number 0010 0100 1110, or in decimal, 590.

The first 12 bytes are called the RIFF chunk. Bytes 0–3 are 52 49 46 46, the ascii characters "RIFF". Bytes 4–7 give the total number of bytes in the remaining part of the entire WAV file (byte 8 onward), in little endian format as described above. Bytes 8–11 are 57 41 56 45, the ascii characters "WAVE" to indicate the RIFF file type.

The next 24 bytes are called the FORMAT chunk. Bytes 0–3 are 66 6D 74 20, the ascii characters "fmt ". Bytes 4–7 give the length of the remainder of the FORMAT chunk, which for a WAV file will always be 10 00 00 00 to indicate 16 bytes. Bytes 8–9 are always 01 00, don't ask me why. Bytes 10–11 indicate the number of channels, 01 00 for Mono and 02 00 for Stereo. Bytes 12–15 give the sample rate, which is measured in Hz. So for example 44,100 Hz comes out as 44 AC 00 00. Bytes 16–19 give the number of bytes per second. This can be found by multiplying the sample rate with the number of bytes representing each sample. Bytes 20–21 give the number of bytes per sample, so 01 00 for 8-bit mono, 02 00 for 8-bit stereo or 16-bit mono, and 04 00 for 16-bit stereo. Finally, bytes 22–23 give the number of bits per sample, which is 8 times as big as bytes 20–21.

So for 16-bit CD quality stereo audio, the number of bytes per second is 44,100 $\times$ 2 $\times$2 = 176,400, which in hexadecimal is 10 B1 02 00. So the RIFF and FORMAT chunk would be as follows.

```
52 49 46 46 xx xx xx xx-57 41 56 45 66 6D 74 20
10 00 00 00 01 00 02 00-44 AC 00 00 10 B1 02 00
04 00 20 00
```

Here, `xx xx xx xx` represents the total length of the WAV file after the first eight bytes.

Finally, for the DATA chunk, bytes 0–3 are 64 52 74 61 for ascii "data". Bytes 4–7 give the length of the remainder of the DATA chunk, in bytes. Bytes 8 onwards are the actual data samples, in little endian binary as always. The data come in pieces called *sample frames*, each representing the data to be played at a particular point in time. So for example for a 16-bit stereo signal, each sample frame would consist of two bytes for the left channel followed by two bytes for the right channel. Since both positive and negative numbers are to be encoded in the binary data, the format used is *two's complement*. So positive numbers from 0 to $32,767$ begin with a binary digit zero, and negative numbers from $-32,768$ to -1 are represented by adding $65,536$, so that they begin with a binary digit one. For example, the number -1 is represented by the bytes FF FF, $-32,768$ is represented by 00 80, and $32,767$ is represented by FF 7F.

Unfortunately, two's complement is only used when the samples are more than 8 bits long; 8-bit samples are represented using the numbers from 0 to 255, with no negative numbers. So 128 is the neutral position for the signal.

Other digital audio formats similar in nature to the WAV file include the AIFF format (Audio Interchange File Format), commonly used on Macintosh computers, and the AU format, developed by Sun Microsystems and commonly used on UNIX computers.

The MP3 format[2] is different from the WAV format in that it uses *data compression*. The file needs to be uncompressed as it is played.

There are two kinds of compression: lossless and lossy. Lossless compression gives rise to a shorter file from which the original may be reconstructed exactly. For example, the ZIP file format is a lossless compression format. This can only work with non-random data. The more random the data, the less it can be compressed. For example, if the data to be compressed consists of 10,000 consecutive copies of the binary string 01001000, then that information can be imparted in a lot less than 10,000 bytes. In information theory, this is captured by the concept of *entropy*. The entropy of a signal is defined to be the logarithm to base two of the number of different possibilities for the signal. The less random the signal, the fewer possibilities are allowed for the data in the signal, and hence the smaller the entropy. The entropy measures the smallest number of binary bits the signal could be compressed into.

Lossy compression retains the most essential features of the file, and allows some degeneration of the data. The kind of degeneration allowed must always depend on the context. For an audio file, for example, we can try to decide which aspects of the signal make little difference to the perception of the sound, and allow these aspects to change. This is precisely what the MP3 format does.

The actual algorithm is very complicated, and makes use of some subtle psychoacoustics. Here are some of the techniques used for encoding MP3 formats.

(i) The threshold of hearing depends on frequency, and the ear is most sensitive in the middle of the audio frequency range. This is described using the Fletcher–Munson curves, as explained on page 11. So low amplitude sounds at the extremes of the frequency range can be ignored unless there is no other sound present.

(ii) The phenomenon of masking means that some sounds will be present but will not be perceived because of the existence of some other component of the sound. These masked sounds are omitted from the compressed signal.

(iii) A system of borrowing is used, so that a passage which needs more bytes to represent the sound in a perceptually accurate way can use them at the expense of using fewer bytes to represent perceptually simpler passages.

(iv) A stereo signal often does not contain much more information than each channel alone, and joint stereo encoding makes use of this.

(v) The MP3 format also makes use of Huffman coding, in which strings of information which occur with higher probability are coded using a shorter string of bits.

---

[2]MP3 stands for "MPEG I/II Layer 3." MPEG is itself an acronym for "Motion Picture Experts Group," which is a family of standards for encoding audio-visual information such as movies and music.

7.4. MIDI

Most synthesizers these days talk to each other and to computers via MIDI cables. MIDI stands for "Musical Instrument Digital Interface." It is an internationally agreed data transmission protocol, introduced in 1982, for the transmission of musical information between different digital devices. It is important to realise that in general there is no waveform information present in MIDI, unless the message is a "sample dump." Instead, most MIDI messages give a short list of abstract parameters for an event.

There are three basic types of MIDI message: note messages, controller messages, and system exclusive messages. Note messages carry information about the starting time and stopping time of notes, which patch (or voice) should be used, the volume level, and so on. Controller messages change parameters like chorus, reverb, panning, master volume, etc. System exclusive messages are for transmitting information specific to a given instrument or device. They start with an identifier for the device, and the body can contain any kind of information in a format proprietary to that device. The commonest kind of system exclusive messages are for transmitting the data for setting up a patch on a synthesizer.

The MIDI standard also includes some hardware specifications. It specifies a baud rate of 31.25 KBaud. For modern machines this is very slow, but for the moment we are stuck with this standard. One of the results of this is that systems often suffer from MIDI "bottlenecks," which can cause audibly bad timing. The problem is especially bad with MIDI data involving continually changing values of a control variable such as volume or pitch.

Further reading:

S. de Furia and J. Scacciaferro, *MIDI programmer's handbook* [**42**].

F. Richard Moore, *The dysfunctions of MIDI*, Computer Music Journal 12 (1) (1988), 19–28.

Joseph Rothstein, *MIDI, A comprehensive introduction* [**124**].

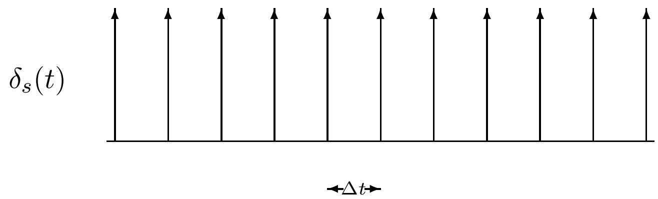
Eleanor Selfridge-Field (Editor), Donald Byrd (Contributor), David Bainbridge (Contributor), *Beyond MIDI: The Handbook of Musical Codes*, M. I. T. Press (1997).

7.5. Delta functions and sampling

One way to represent the process of sampling a signal is as multiplication by a stream of Dirac delta functions (see §2.17). Let N denote the sample rate, measured in samples per second, and let $\Delta t = 1/N$ denote the interval between sample times. So for example for compact disc recording we want $N = 44,100$ samples per second, and $\Delta t = 1/44,100$ seconds. We define the *sampling function* with spacing Δt to be

$$\delta_s(t) = \Delta t \sum_{n=-\infty}^{\infty} \delta(t - n\Delta t).$$

The reason for the factor of Δt in front of the summation is so that the integral of this function over an interval of time approximates the length of the interval.



If $f(t)$ represents an analog signal, then

$$f(t)\delta_s(t) = \Delta t \sum_{n=-\infty}^{\infty} f(t)\delta(t - n\Delta t) = \Delta t \sum_{n=-\infty}^{\infty} f(n\Delta t)\delta(t - n\Delta t)$$

represents the sampled signal. This has been digitized with respect to time, but not with respect to signal amplitude. The integral of the digitized signal $f(t)\delta_s(t)$ over any period of time approximates the integral of the analog signal $f(t)$ over the same time interval.

One of the keys to understanding the digitized signal is Poisson's summation formula from Fourier analysis.

Theorem 7.5.1.

$$\Delta t \sum_{n=-\infty}^{\infty} f(n\Delta t) = \sum_{n=-\infty}^{\infty} \hat{f}\left(\frac{n}{\Delta t}\right). \tag{7.5.1}$$

Proof. This follows from the Poisson summation formula (2.16.1), using Exercise 3 of §2.13. □

Corollary 7.5.2. *The Fourier transform of the sampling function* $\delta_s(t)$ *is another sampling function in the frequency domain,*

$$\widehat{\delta_s}(\nu) = \sum_{n=-\infty}^{\infty} \delta\left(\nu - \frac{n}{\Delta t}\right).$$

Proof. If $f(t)$ is a test function, then the definition of $\delta_s(t)$ gives

$$\int_{-\infty}^{\infty} f(t)\delta_s(t)\,dt = \Delta t \sum_{n=-\infty}^{\infty} f(n\Delta t).$$

Applying Parseval's identity (2.15.1) to the left hand side (and noting that the sampling function is real, so that $\delta_s(t) = \overline{\delta_s(t)}$) and applying formula (7.5.1) to the right hand side, we obtain

$$\int_{-\infty}^{\infty} \hat{f}(\nu)\overline{\widehat{\delta_s}(\nu)}\,d\nu = \sum_{n=-\infty}^{\infty} \hat{f}\left(\frac{n}{\Delta t}\right).$$

The required formula for $\widehat{\delta_s}(\nu)$ follows. □

Corollary 7.5.3. *The Fourier transform of a digital signal $f(t)\delta_s(t)$ is*

$$\widehat{f\delta_s}(\nu) = \sum_{n=-\infty}^{\infty} \hat{f}\left(\nu - \frac{n}{\Delta t}\right)$$

which is periodic in the frequency domain, with period equal to the sampling frequency $1/\Delta t$.

Proof. By Theorem 2.18.1(ii), we have

$$\widehat{f\delta_s}(\nu) = (\hat{f} * \hat{\delta}_s)(\nu),$$

and by Corollary 7.5.2, this is equal to

$$\int_{-\infty}^{\infty} \hat{f}(u) \sum_{n=-\infty}^{\infty} \delta\left(\nu - \frac{n}{\Delta t} - u\right) du = \sum_{n=-\infty}^{\infty} \hat{f}\left(\nu - \frac{n}{\Delta t}\right). \qquad \square$$

7.6. Nyquist's theorem

Nyquist's theorem[3] states that the maximum frequency that can be represented when digitizing an analog signal is exactly half the sampling rate. Frequencies above this limit will give rise to unwanted frequencies below the *Nyquist frequency* of half the sampling rate. What happens to signals at exactly the Nyquist frequency depends on the phase. If the entire frequency spectrum of the signal lies below the Nyquist frequency, then the *sampling theorem* states that the signal can be reconstructed exactly from its digitization.

To explain the reason for Nyquist's theorem, consider a pure sinusoidal wave with frequency ν, for example

$$f(t) = A\cos(2\pi\nu t).$$

Given a sample rate of $N = 1/\Delta t$ samples per second, the height of the function at the Mth sample is given by

$$f(M/N) = A\cos(2\pi\nu M/N).$$

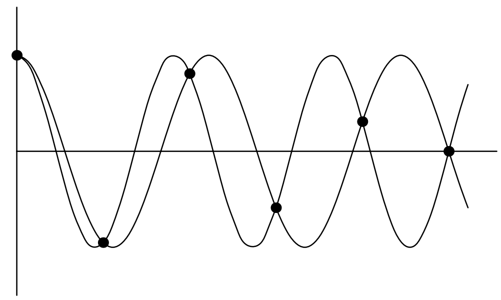
If ν is greater than $N/2$, say $\nu = N/2 + \alpha$, then

$$\begin{aligned} f(M/N) &= A\cos(2(N/2+\alpha)M\pi/N) \\ &= A\cos(M\pi + 2\alpha M\pi/N) \\ &= (-1)^M A\cos(2\alpha M\pi/N). \end{aligned}$$

Changing the sign of α makes no difference to the outcome of this calculation, so this gives exactly the same answer as the waveform with $\nu = N/2-\alpha$ instead of $\nu = N/2 + \alpha$. To put it another way, the sample points in

[3]Harold Nyquist, *Certain topics in telegraph transmission theory*, Transactions of the American Institute of Electrical Engineers, April 1928. Nyquist retired from Bell Labs in 1954 with about 150 patents to his name. He was renowned for his ability to take a complex problem and produce a simple minded solution that was far superior to other, more complicated approaches.

this calculation are exactly the points where the graphs of the functions $A\cos(2(N/2+\alpha)\pi t)$ and $A\cos(2(N/2-\alpha)\pi t)$ cross.



The result of this is that a frequency which is greater than half the sample frequency gets reflected through half the sample frequency, so that it sounds like a frequency of the corresponding amount less than half. This phenomenon is called *aliasing*. In the above diagram, the sample points are represented by black dots. The two waves have frequency slightly more and slightly less than half the sample frequency. It is easy to see from the diagram why the sample values are equal. Namely, the sample points are simply the points where the two graphs cross.

For waves at exactly half the sampling frequency, something interesting occurs. Cosine waves survive intact, but sine waves disappear altogether. This means that phase information is lost, and amplitude information is skewed.

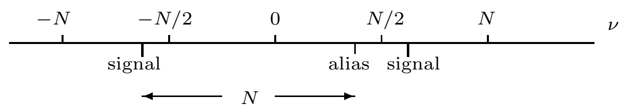
The upshot of Nyquist's theorem is that before digitizing an analog signal, it is essential to pass it through a low pass filter to cut off frequencies above half the sample frequency. Otherwise, each frequency will come paired with its reflection.

In the case of digital compact discs, the cutoff frequency is half of 44.1 KHz, or 22.05 KHz. Since the limit of human perception is usually below 20 KHz, this may be considered satisfactory, but only by a small margin.

We can also explain Nyquist's theorem in terms of Corollary 7.5.3. Namely, the Fourier transform

$$\widehat{f\delta_s}(\nu) = \sum_{n=-\infty}^{\infty} \hat{f}\left(\nu - \frac{n}{\Delta t}\right)$$

is periodic with period equal to the sampling frequency $N = 1/\Delta t$. The term with $n = 0$ in this sum is the Fourier transform of $f(t)$. The remaining terms with $n \neq 0$ appear as aliased artifacts, consisting of frequency components shifted in frequency by multiples of the sampling frequency $N = 1/\Delta t$. If $f(t)$ has a nonzero part of its spectrum at frequency greater than $N/2$, then its Fourier transform will be nonzero at plus and minus this quantity. Then adding or subtracting N will result in an artifact at the corresponding amount less than $N/2$, the other side of the origin.



Another remarkable fact comes out of Corollary 7.5.3, namely the *sampling theorem*. Provided the original signal $f(t)$ satisfies $\hat{f}(\nu) = 0$ for $\nu \geq N/2$, in other words, provided that the entire spectrum lies below the Nyquist frequency, the original signal can be reconstructed exactly from the sampled signal, without any loss of information. To reconstruct $\hat{f}(\nu)$, we begin by by truncating $\widehat{f\delta_s}(\nu)$, and then $f(t)$ is reconstructed using the inverse Fourier transform. Carrying this out in practice is a different matter, and requires very accurate analog filters.

7.7. The z-transform

For digital signals, it is often more convenient to use the z-transform instead of the Fourier transform. The point is that by Corollary 7.5.3, the Fourier transform of a digital signal is periodic, with period equal to the sampling frequency. So it contains a lot of redundant information. The idea of the z-transform is to wrap the Fourier transform round the unit circle in the complex plane. This is achieved by setting

$$\boxed{z = e^{2\pi i\nu\Delta t}}$$

so that as ν changes in value by $1/\Delta t$, z goes exactly once round the unit circle in the complex plane. Any periodic function of ν with period $1/\Delta t$ can then be written as a function of z. The Fourier transform of the sampled signal $f(t)\delta_s(t)$ is then

$$\begin{aligned}\int_{-\infty}^{\infty} f(t)\delta_s(t)e^{-2\pi i\nu t}\,dt &= \int_{-\infty}^{\infty}\left(\sum_{n=-\infty}^{\infty}\Delta t\, f(t)\delta(t-n\Delta t)\right) z^{-t/\Delta t}\,dt \\ &= \sum_{n=-\infty}^{\infty}\Delta t\, f(n\Delta t)z^{-n}.\end{aligned}$$

The factor of Δt is just an annoying constant, and so the z-transform of the digitized signal is simply defined as

$$\boxed{F(z) = \sum_{n=-\infty}^{\infty} f(n\Delta t)z^{-n}.} \tag{7.7.1}$$

The Fourier transform may be recovered as

$$\widehat{f\delta_s}(\nu) = \Delta t\, F(e^{2\pi i\nu\,\Delta t}).$$

Warning. It is necessary to exercise caution when manipulating expressions like equation (7.7.1), because of *Euler's joke*. Here's the joke. Consider a signal which is constant over all time,

$$\begin{aligned} F(z) &= \dots + z^2 + z + 1 + z^{-1} + z^{-2} + \dots \\ &= \sum_{n=-\infty}^{\infty} z^n. \end{aligned}$$

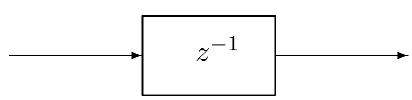
Divide this infinite sum up into two parts, and sum them separately.

$$\begin{aligned} F(z) &= (\dots + z^2 + z + 1) + (z^{-1} + z^{-2} + \dots) \\ &= \frac{1}{1-z} + \frac{z^{-1}}{1-z^{-1}} \\ &= \frac{1}{1-z} + \frac{1}{z-1} \\ &= 0. \end{aligned}$$

This is clearly nonsense. The problem is that the first parenthesized sum only converges for $|z| > 1$, while the second sum only converges for $|z| < 1$. So there is no value of z for which both sums make sense simultaneously.

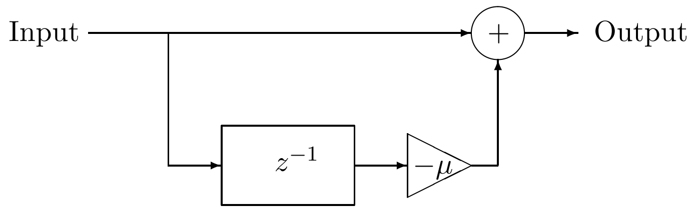
The resolution of this problem is only to allow signals with some finite starting point. So we assume that $f(n\Delta t) = 0$ for all large enough negative values of n.

In terms of the z-transform, delaying the signal by one sample corresponds to multiplication by z^{-1}. So in the literature, you will see the block diagram for such a digital delay drawn as follows. We shall use the same convention.



7.8. Digital filters

The subject of digital filters has a vast literature. We shall only touch the surface, in order to illustrate how the z-transform enters the picture. Let us begin with an example. Consider the following diagram.



If $f(n\Delta t)$ is the input and $g(n\Delta t)$ is the output, then the relation represented by the above diagram is

$$g(n\Delta t) = f(n\Delta t) - \mu f((n-1)\Delta t). \tag{7.8.1}$$

So the relation between the z-transforms is

$$G(z) = F(z) - \mu z^{-1}F(z) = (1 - \mu z^{-1})F(z).$$

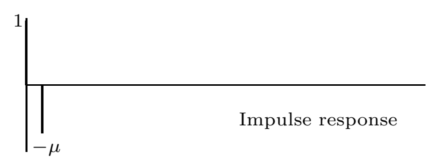
This tells us about the frequency response of the filter. A given frequency ν corresponds to the points $z = e^{\pm 2\pi i\nu\Delta t}$ on the unit circle in the complex plane, with half the sampling frequency corresponding to $e^{\pi i} = -1$.

At a particular point on the unit circle, the value of $1 - \mu z^{-1}$ gives the frequency response. Namely, the amplification is $|1 - \mu z^{-1}|$, and the phase shift is given by the argument of $1 - \mu z^{-1}$.

More generally, if the relationship between the z-transforms of the input and output signal, $F(z)$ and $G(z)$, is given by

$$G(z) = H(z)F(z)$$

then the function $H(z)$ is called the *transfer function* of the filter. The interpretation of the transfer function, for example $1 - \mu z^{-1}$ in the above filter, is that it is the z-transform of the *impulse response* of the filter.



The impulse response is defined to be the output resulting from an input which is zero except at the one sample point $t = 0$, where its value is one, namely

$$f(n\Delta t) = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0. \end{cases}$$

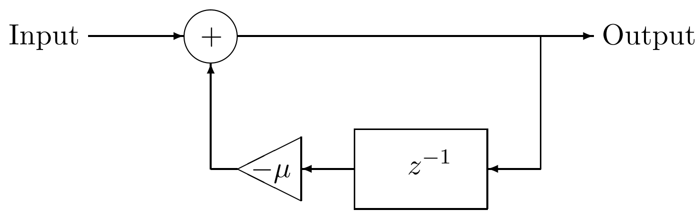
The sampled function $f\delta_s$ is then a Dirac delta function.

For digital signals, the convolution of f_1 and f_2 is defined to be

$$(f_1 * f_2)(n\Delta t) = \sum_{m=-\infty}^{\infty} f_1((n-m)\Delta t)f_2(m\Delta t).$$

Multiplication of z-transforms corresponds to convolution of the original signals. This is easy to see in terms of how power series in z^{-1} multiply. So in the above example, the impulse response is: 1 at $n = 0$, $-\mu$ at $n = 1$, and zero for $n \neq 0, 1$. Convolution of the input signal $f(n\Delta t)$ with the impulse response gives the output signal $g(n\Delta t)$ according to equation (7.8.1).

As a second example, consider a filter with feedback.



The relation between the input $f(n\Delta t)$ and the output $g(n\Delta t)$ is now given by

$$g(n\Delta t) = f(n\Delta t) - \mu g((n-1)\Delta t).$$

This time, the relation between the z-transforms is

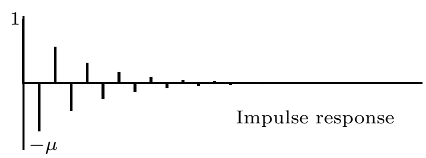
$$G(z) = F(z) - \mu z^{-1} G(z),$$

or

$$G(z) = \frac{1}{1 + \mu z^{-1}} F(z).$$

Notice that this is unstable when $|\mu| > 1$, in the sense that the signal grows without bound. Even when $|\mu| = 1$, the signal never dies away, so we say that this filter is *stable* provided $|\mu| < 1$. This is easiest to see in terms of the impulse response of this filter, which is

$$\frac{1}{1 + \mu z^{-1}} = 1 - \mu z^{-1} + \mu^2 z^{-2} - \mu^3 z^{-3} + \dots$$



Filters are usually designed in such a way that the output $g(n\Delta t)$ depends linearly on $f((n-m)\Delta t)$ for a finite set of values of $m \geq 0$ and on $g((n-m)\Delta t)$ for a finite set of values of $m > 0$. For such a filter, the z-transform of the impulse response is a rational function of z, which means that it is a ratio of two polynomials

$$\frac{p(z)}{q(z)} = a_0 + a_1 z^{-1} + a_2 z^{-2} + \dots$$

The coefficients $a_0, a_1, a_2, \ldots$ are the values of the impulse response at $t = 0$, $t = \Delta t$, $t = 2\Delta t$, ...

The coefficients a_n tend to zero as n tends to infinity, if and only if the poles μ of $p(z)/q(z)$ satisfy $|\mu| < 1$. This can be seen in terms of the complex partial fraction expansion of the function $p(z)/q(z)$.

The location of the poles inside the unit circle has a great deal of effect on the frequency response of the filter. If there is a pole near the boundary, it will cause a local maximum in the frequency response, which is called a *resonance*. The frequency is given in terms of the argument of the position of the pole by

$$\nu = (\text{sample rate}) \times (\text{argument})/2\pi.$$

Decay time. The decay time of a filter for a particular frequency is defined to be the time it takes for the amplitude of that frequency component to reach $1/e$ of its initial value. To understand the effect of the location of a pole on the decay time, we examine the transfer function

$$H(z) = \frac{1}{z-a} = \frac{z^{-1}}{1 - az^{-1}} = z^{-1} + az^{-2} + a^2 z^{-3} + \ldots.$$

So in a period of n sample times, the amplitude is multiplied by a factor of a^n. So we want $|a|^n = 1/e$, or $n = -1/\ln|a|$. So the formula for decay time is

$$\boxed{\text{Decay time} = \frac{-\Delta t}{\ln|a|} = \frac{-1}{N \ln|a|}} \qquad (7.8.2)$$

where $N = 1/\Delta t$ is the sample rate. So the decay time is inversely proportional to the logarithm of the absolute value of the location of the pole. The further the pole is inside the unit circle, the smaller the decay time, and the faster the decay. A pole near the unit circle gives rise to a slow decay.

Exercises

1. (a) Design a digital filter whose transfer function is $z^2/(z^2 + z + \frac{1}{2})$, using the symbol z^{-1} in a box to denote a delay of one sample time, as above.

(b) Compute the frequency response of this filter. Let N denote the number of sample points per second, so that the answer should be a function of ν for $-N/2 < \nu < N/2$.

(c) Is this filter stable?

Further reading:

R. W. Hamming, *Digital filters* [**51**].

Bernard Mulgrew, Peter Grant and John Thompson, *Digital signal processing* [**92**].

7.9. The discrete Fourier transform

The discrete Fourier transform describes the frequency components of a digitized signal of finite length. If the length of the signal is N, then the discrete Fourier transform is given by

$$F(k) = \frac{1}{N}\sum_{n=0}^{N-1} f(n)e^{-2\pi ink/N}$$
$$f(n) = \sum_{k=0}^{N-1} F(k)e^{2\pi ikn/N}.$$

For a long signal, the usual process is to choose for N a number that is used as a window size for a moving window in the signal. So the discrete Fourier transform is really a digitized version of the windowed Fourier transform.

The fast Fourier transform is a way to compute the discrete Fourier transform using $2N\log_2 N$ operations rather than N^2. The number of sample points N has to be a power of two for it to be this efficient, but the algorithm works for any highly composite value of N.

Further reading:

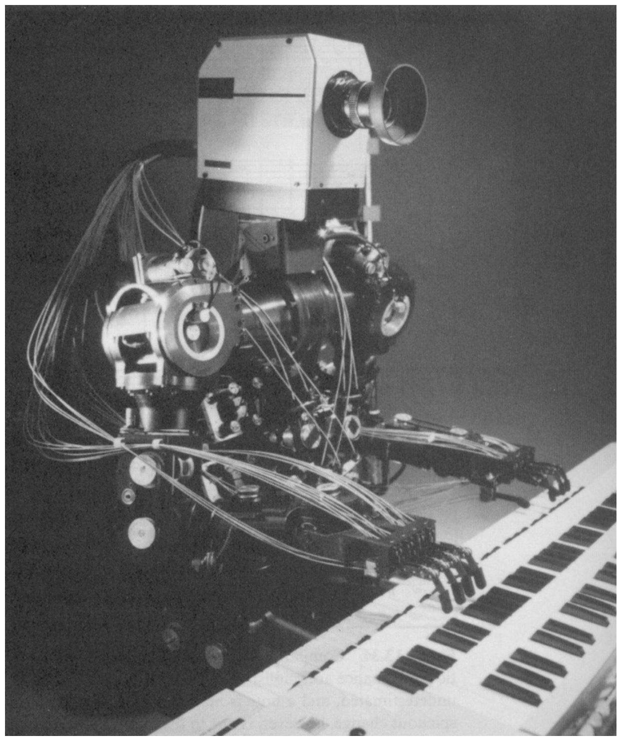
G. D. Bergland, *A guided tour of the fast Fourier transform*, IEEE Spectrum 6 (1969), 41–52.

James W. Cooley and John W. Tukey, *An algorithm for the machine calculation of complex Fourier series*, Math. of Computation 19 (1965), 297–301. This is usually regarded as the original article announcing the fast Fourier transform as a practical algorithm.

CHAPTER 8

Synthesis

8.1. Introduction



WABOT-2 (Waseda University and
Sumitomo Corp., Japan 1985)

In this chapter, we investigate synthesis of musical sounds. We pay special attention to Frequency Modulation (or FM) synthesis, not because it is a particularly important method of synthesis, but rather because it is easy to use FM synthesis as a vehicle for conveying general principles. We also discuss other aspects of digital music, such as aliasing and Nyquist's theorem, MIDI, and internet resources.

Interesting musical sounds do not in general have a static frequency spectrum. The development with time of the spectrum of a note can be understood to some extent by trying to mimic the sound of a conventional musical instrument synthetically. This exercise focuses our attention on what are usually referred to as the attack, decay, sustain and release parts of a note (ADSR). Not only does the amplitude change during these intervals, but also the frequency spectrum. Synthesizing sounds which do not sound mechanical and boring turns out to be harder than one might guess. The ear is very good at picking out the regular features produced by simple minded algorithms and identifying them as synthetic. This way, we are led to an appreciation of the complexity of even the simplest of sounds produced by conventional instruments.

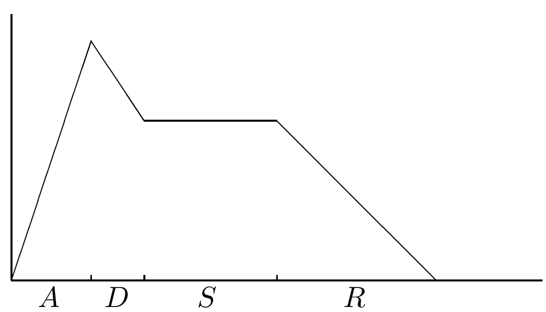
Of course, the real strength of synthesis is the ability to produce sounds not previously attainable, and to manipulate sounds in ways not previously possible. Most music, even in today's era of the availability of cheap and powerful digital synthesizers, seems to occupy only a very small corner of the available sonic pallette. The majority of musicians who use synthesizers just punch the presets until they find the ones they like, and then use them without modification. Exceptions to this rule stand out from the crowd; listening to a recording by the Japanese synthesist Tomita, for example, one is struck immediately by the skill expressed in the shaping of the sound.

Further listening: (See Appendix R)

Isao Tomita, *Pictures at an Exhibition* (Mussorgsky).

8.2. Envelopes and LFOs

Whatever method is used to synthesize sounds, attention has to be paid to envelopes, so we discuss these first. Very few sounds just consist of a spectrum, static in time. If we hear a note on almost any instrument, there is a clearly defined attack at the beginning of the sound, followed by a decay, then a sustained part in the middle, and finally a release. In any particular instrument, some of these may be missing, but the basic structure is there. Synthesis follows the same pattern. The commonly used abbreviation is ADSR envelope, for attack/decay/sustain/release envelope.



It was not really understood properly until the middle of the twentieth century, when electronic synthesis was taking its first tentative steps, that the attack portion of a note is the most vital to the human ear in identifying the instrument. The transients at the beginning are much more different from one instrument to another than the steady part of the note.

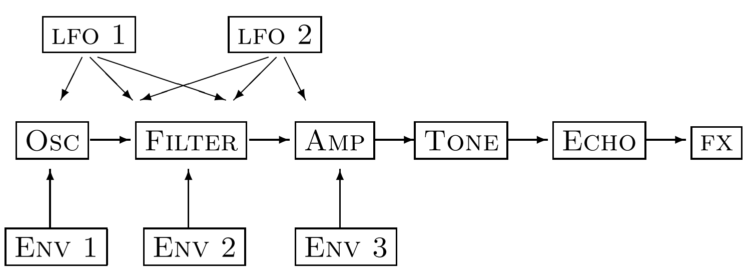
On a typical synthesizer, there are a number of envelope generators. Each one determines how the amplitude of the output of some component of the system varies with time. It is important to understand that amplitude of the final signal is not the only attribute which is assigned an envelope. For example, when a bell sounds, initially the frequency spectrum is very rich, but many of the partials die away very quickly leaving a purer sound. Mimicking this sort of behavior using FM synthesis turns out to be relatively easy, by assigning an envelope to a modulating signal, which controls timbre. We shall discuss this further when we discuss FM synthesis, but for the moment we note that aspects of timbre are often controlled with an envelope generator. When the synthesizer is controlled by a keyboard, as is often the case, it is usual to arrange that depressing a key initiates the attack, and releasing the key initiates the release portion of the envelope.

An envelope generator produces an envelope whose shape is determined by a number of programmable parameters. These parameters are usually given in terms of levels and rates. Here is an example of how an envelope might work in a typical keyboard synthesizer or other MIDI controlled environment. Level 0 is the level of the envelope at the “key on” event. Rate 1 then determines how fast the level changes, until it reaches level 1. Then it switches to rate 2 until level 2 is reached, and then rate 3 until level 3 is reached. Level 3 is then in effect until the “key off” event, when rate 4 takes effect until level 4 is reached. Finally, level 4 is the same as level 0, so that we are ready for the next “key on” event. In this example, there are two separate components to the decay phase of the envelope. Some synthesizers make do with only one, and some have even more.

Similar in concept to the envelope is the *low frequency oscillator* or LFO. This produces an output which is usually in the range 0.1–20 Hz, and whose waveform is usually something like triangle, sawtooth (up or down), sine, square or random. The LFO is used to produce repeating changes in some controllable parameter. Examples include pitch control for vibrato, and amplitude or timbre control for tremolo. The LFO can also be used to control less obvious parameters such as the cutoff and resonance of a filter, or the pulse width of a square wave (pulse width modulation, or PWM), see Exercise 6 in §2.4.

The parameters associated with an LFO are rate (or frequency), depth (or amplitude), waveform, and attack time. Attack time is used when the effect is to be introduced gradually at the beginning of the note.

Here is a block diagram for a typical analog synthesizer.

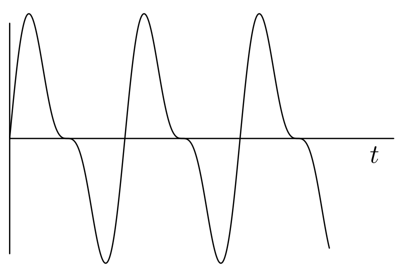


The oscillator (Osc) generates the basic waveform, which can be chosen from sine wave, square wave, triangular wave, sawtooth, noise, etc. The envelope (Env 1) specifies how the pitch changes with time. The filter specifies the "brightness" of the sound. It can be chosen from high pass, low pass and band pass. The envelope (Env 2) specifies how the brightness varies with time. Also, a *resonance* is specified, which determines the emphasis applied to the region at the cutoff frequency. The amplifier (Amp) specifies the volume, and the envelope (Env 3) specifies how the volume changes with time. The tone control (Tone) adjusts the overall tone, the delay unit (Echo) adds an echo effect, and the effects unit (fx) can be used to add reverberation, chorus, and so on. Low frequency oscillators (lfo 1 and lfo 2) are provided, which can be used to modulate the oscillator, filter or amplifier.

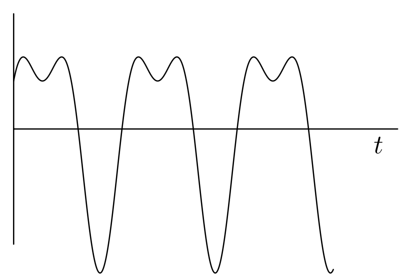
8.3. Additive Synthesis

The easiest form of synthesis to understand is additive synthesis, which is in effect the opposite of Fourier analysis of a signal. To synthesize a periodic wave, we generate its Fourier components at the correct amplitudes and mix them. This is a comparatively inefficient method of synthesis, because in order to produce a note with a large number of harmonics, a large number of sine waves will need to be mixed together. Each will be assigned a separate envelope in order to create the development of the note with time. This way, it is possible to control the development of timbre with time, as well as the amplitude. So for example, if it is desired to create a waveform whose attack phase is rich in harmonics and which then decays to a purer tone, then the components of higher frequency will have a more rapidly decaying envelope than the lower frequency components.

Phase is unimportant to the perception of steady sounds, but more important in the perception of transients. So for steady sounds, the graph representing the waveform is not very informative. For example, here are the graphs of the functions $\sin(\omega t) + \frac{1}{2}\sin(2\omega t)$ and $\sin(\omega t) + \frac{1}{2}\cos(2\omega t)$.

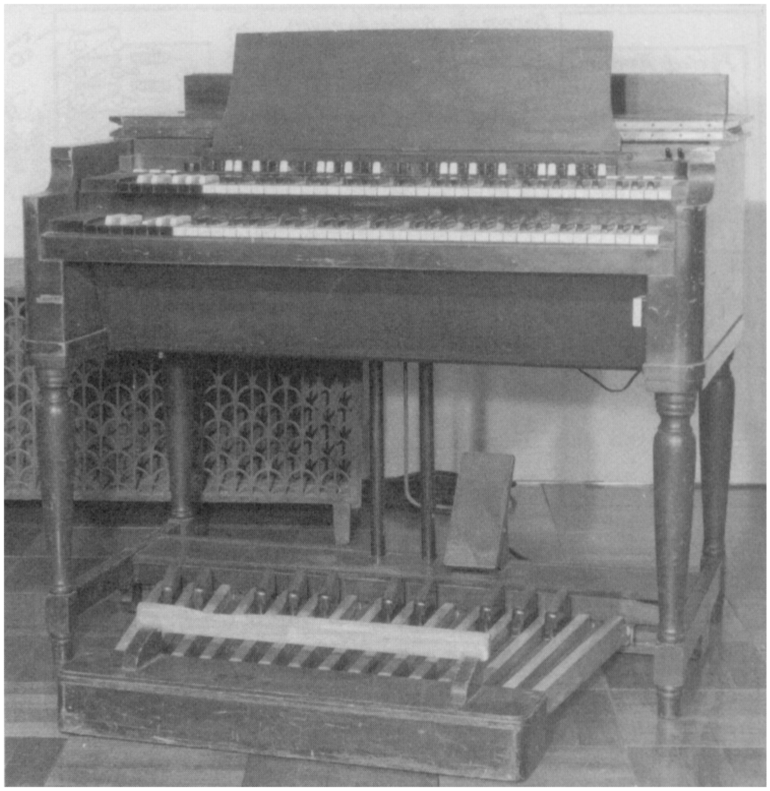


$\sin t + \frac{1}{2}\sin 2t$



$\sin t + \frac{1}{2}\cos 2t$

The only difference between these functions is that the second partial has had its phase changed by an angle of $\pi/2$, so as steady sounds, these will sound identical. With more partials, it becomes extremely hard to tell whether two waveforms represent the same steady sound. It is for this reason that the waveform is not a very useful way to represent the sound, whereas the spectrum, and its development with time, are much more useful.



Hammond B3 organ

In some ways, additive synthesis is a very old idea. A typical cathedral or church organ has a number of register stops, determining which sets of pipes are used for the production of the note. The effect of this is that depressing a single key can be made to activate a number of harmonically related pipes, typically a mixture of octaves and fifths. Early electronic instruments such as the Hammond organ operated on exactly the same principle.

More generally, additive synthesis may be used to construct sounds whose partials are not multiples of a given fundamental. This will give nonperiodic waveforms which nevertheless sound like steady tones.

Exercises

1. Explain how to use additive synthesis to construct a square wave out of pure sine waves.
[Hint: Look at §2.2]

2. Explain in terms of the human ear (§1.2) why the phases of the harmonic components of a steady waveform should not have a great effect on the way the sound is perceived.

Further reading:

F. de Bernardinis, R. Roncella, R. Saletti, P. Terreni and G. Bertini, *A new VLSI implementation of additive synthesis*, Computer Music Journal 22 (3) (1998), 49–61.

8.4. Physical modeling

The idea of physical modeling is to take a physical system such as a musical instrument, and to mimic it digitally. We give one simple example to illustrate the point. We examined the wave equation for the vibrating string in §3.2, and found d'Alembert's general solution

$$y = f(x + ct) + g(x - ct).$$

Given that time is quantized with sample points at spacing Δt, it makes sense to quantize the position along the string at intervals of $\Delta x = c\Delta t$. Then at time $n\Delta t$ and position $m\Delta x$, the value of y is

$$\begin{aligned} y &= f(m\Delta x + nc\Delta t) + g(m\Delta x - nc\Delta t) \\ &= f((m + n)c\Delta t) + g((m - n)c\Delta t). \end{aligned}$$

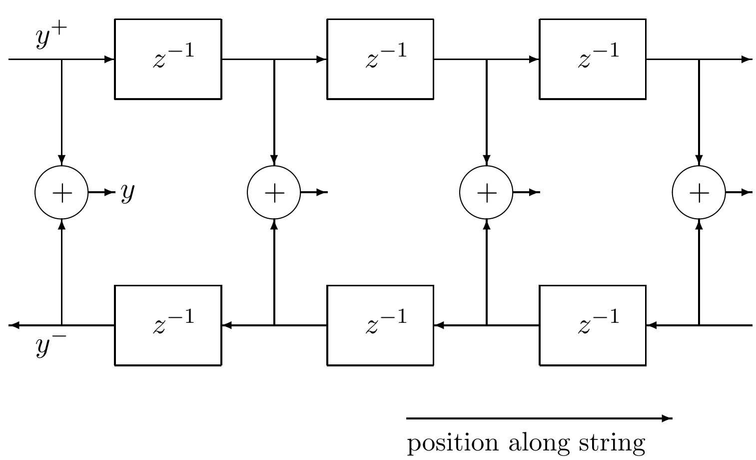
To simplify the notation, we write

$$y^-(n) = f(nc\Delta t), \qquad y^+(n) = g(nc\Delta t)$$

so that y^- and y^+ represent the parts of the wave travelling left, respectively right along the string. Then at time $n\Delta t$ and position $m\Delta x$ we have

$$y = y^-(m + n) + y^+(m - n).$$

This can be represented by two delay lines moving left and right:



It is a good idea to make the string an integer number of sample points long, let us say $l = L\Delta x$. Then the boundary conditions at $x = 0$ and $x = l$ (see equations (3.2.3) and (3.2.4)) say that

$$y^-(n) = -y^+(-n)$$

and that

$$y^+(n+2L) = y^+(n).$$

This means that at the ends of the string, the signal gets negated and passed round to the other set of delays. Then the initial pluck or strike is represented by setting the values of $y^-(n)$ and $y^+(n)$ suitably at $t=0$, for $0 \le n < 2L$.

Thinking in terms of digital filters, the z-transform of the y^+ signal

$$Y^+(z) = y^+(0) + y^+(1)z^{-1} + y^+(2)z^{-2} + \dots$$

satisfies

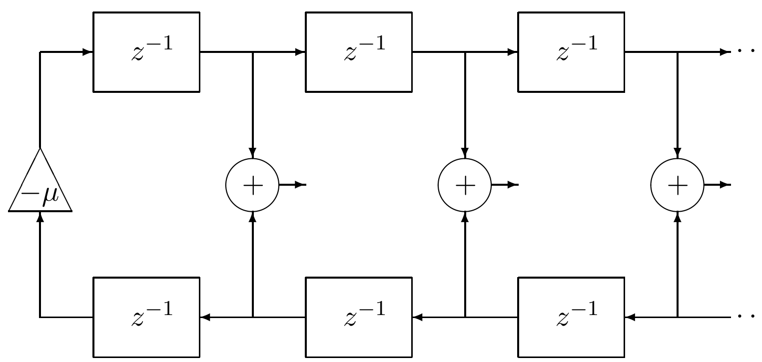
$$Y^+(z) = z^{-2L}Y^+(z) + (y^+(0) + y^+(1)z^{-1} + \cdots + y^+(2L-1))$$

or

$$Y^+(z) = \frac{y^+(0)z^{2L} + y^+(1)z^{2L-1} + \cdots + y^+(2L-1)z}{z^{2L}-1}.$$

The poles are equally spaced on the unit circle, so the resonant frequencies are multiples of $N/2L$, where N is the sample frequency. Since the poles are actually *on* the unit circle, the resonant frequencies never decay.

To make the string more realistic, we can put in energy loss at one end, represented by multiplication by a fixed constant factor $-\mu$ with $0 < \mu \le 1$, instead of just negating.



The effect of this on the filter analysis is to move the poles slightly inside the unit circle:

$$Y^+(z) = \frac{y^+(0)z^{2L} + y^+(1)z^{2L-1} + \cdots + y^+(2L-1)z}{z^{2L}-\mu}.$$

The absolute values of the location of the poles are all equal to $|\mu|^{\frac{1}{2L}}$. The decay time is given by equation (7.8.2) as

$$\text{Decay time} = \frac{-2L}{N\ln|\mu|}.$$

The above model is still not very sophisticated, because decay time is independent of frequency. But it is easy to modify by replacing the multiplication by μ by a more complicated digital filter. We shall see a particular example of this idea in the next section. Another easy modification is to have two or more strings cross-coupled, by adding a small multiple of the signal at the end of each into the end of the others. Adding a model of a sounding board is not so easy, but it can be done.

Further reading:

Eric Ducasse, *A physical model of a single-reed instrument, including actions of the player*, Computer Music Journal 27 (1) (2003), 59–70.

M. Laurson, C. Erkut, V. Välimäki and M. Kuuskankare, *Methods for modeling realistic playing in acoustic guitar synthesis*, Computer Music Journal 25 (3) (2001), 38–49.

Julius O. Smith III, *Physical modeling using digital waveguides*, Computer Music Journal 16 (4) (1992), 74–87.

Julius O. Smith III, *Acoustic modeling using digital waveguides*, appears as article 7 in Roads et al [**115**], pages 221–263.

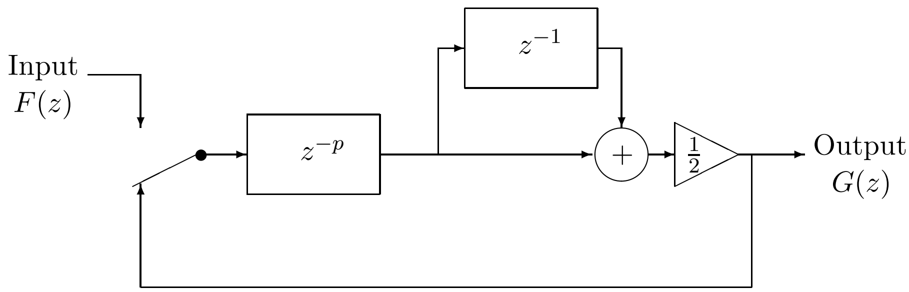
Vesa Välimämi, Mikael Laurson and Cumhur Erkut, *Commuted waveguide synthesis of the clavichord*, Computer Music Journal 27 (1) (2003), 71–82.

8.5. The Karplus–Strong algorithm

The Karplus–Strong algorithm gives very good plucked strings and percussion instruments. The basic technique is a modification of the technique described in the last section, and consists of a digital delay followed by an averaging process. Denote by $g(n\Delta t)$ the value of the nth sample point in the digital output signal for the algorithm. A positive integer p is chosen to represent the delay, and the recurrence relation

$$g(n\Delta t) = \tfrac{1}{2}(g((n-p)\Delta t) + g((n-p-1)\Delta t))$$

is used to define the signal after the first $p+1$ sample points. The first $p+1$ values to feed into the recurrence relation are usually chosen by some random algorithm, and then the feedback loop is switched in. This is represented by an input signal $f(n\Delta t)$ which is zero outside the range $0 \le n \le p$.



Computationally, this algorithm is very efficient. Each sample point requires one addition operation. Halving does not need a multiplication, only a shift of the binary digits.

Let us analyse the algorithm by regarding it as a digital filter, and using the z-transform, as described in §7.8. Let $G(z)$ be the z-transform of the signal $g(n\Delta t)$, and $F(z)$ be the z-transform of the signal given for the first $p+1$ sample points, $f(n\Delta t)$. We have

$$G(z) = \tfrac{1}{2}(1 + z^{-1})z^{-p}(F(z) + G(z)).$$

This gives

$$G(z) = \frac{z+1}{2z^{p+1} - z - 1}\, F(z),$$

and so the z-transform of the impulse response is $(z+1)/(2z^{p+1} - z - 1)$. The poles are the solutions of the equation

$$2z^{p+1} - z - 1 = 0.$$

These are roughly equally spaced around the unit circle, at amplitude just less than one. The solution with smallest argument corresponds to the fundamental of the vibration, with argument roughly $2\pi/(p + \frac{1}{2})$. A more precise analysis is given in §8.6.

The effect of this is a plucked string sound with pitch determined by the formula

$$\text{pitch} = (\text{sample rate})/(p + \tfrac{1}{2}).$$

Since p is constrained to be an integer, this restricts the possible frequencies of the resulting sound in terms of the sample rate. Changing the value of p without introducing a new inital values results in a slur, or tie between notes.

A simple modification of the algorithm gives drumlike sounds. Namely, a number b is chosen with $0 \le b \le 1$, and

$$g(n\Delta t) = \begin{cases} +\frac{1}{2}(g((n-p)\Delta t) + g((n-p-1)\Delta t) & \text{with probability } b \\ -\frac{1}{2}(g((n-p)\Delta t) + g((n-p-1)\Delta t)) & \text{with probability } 1-b. \end{cases}$$

The parameter b is called the *blend factor*. Taking $b = 1$ gives the original plucked string sound. The value $b = \frac{1}{2}$ gives a drumlike sound. With $b = 0$, the period is doubled and only odd harmonics result. This gives some interesting sounds, and at high pitches this gives what Karplus and Strong describe as a *plucked bottle* sound.

Another variation described by Karplus and Strong is what they call *decay stretching*. In this version, the recurrence relation

$$g(n\Delta t) = \begin{cases} g((n-p)\Delta t) & \text{with probability } 1-\alpha \\ \frac{1}{2}(g((n-p)\Delta t) + g((n-p-1)\Delta t)) & \text{with probability } \alpha. \end{cases}$$

The *stretch factor* for this version is $1/\alpha$, and the pitch is given by

$$\text{pitch} = (\text{sample rate})/(p + \tfrac{\alpha}{2}).$$

Setting $\alpha = 0$ gives a non-decaying periodic signal, while setting $\alpha = 1$ gives the original algorithm described above.

There are obviously a lot of variations on these algorithms, and many of them give interesting sounds.

8.6. Filter analysis for the Karplus–Strong algorithm

We saw in the last section that in order to understand the Karplus–Strong algorithm in its simplest form, we need to locate the zeros of the polynomial $2z^{p+1} - z - 1$, where p is a positive integer. In order to do this, we begin by rewriting the equation as

$$2z^{p+\frac{1}{2}} = z^{\frac{1}{2}} + z^{-\frac{1}{2}}.$$

Since we expect z to have absolute value close to one, the imaginary part of $z^{\frac{1}{2}} + z^{-\frac{1}{2}}$ will be very small. If we ignore this imaginary part, then the nth zero of the polynomial around the unit circle will have argument equal to $2n\pi/(p + \frac{1}{2})$. So we write

$$z = (1-\varepsilon)e^{2n\pi i/(p+\frac{1}{2})}$$

and calculate ε, ignoring terms in ε^2 and higher powers. Already from the form of this approximation, we see that the resonant frequency corresponding to the nth pole is equal to $nN/(p + \frac{1}{2})$, where N is the sample frequency. This means that the different resonant frequencies are at multiples of a fundamental frequency of $N/(p + \frac{1}{2})$.

We have

$$2z^{p+\frac{1}{2}} = 2(1-\varepsilon)^{p+\frac{1}{2}} \approx 2 - 2(p + \tfrac{1}{2})\varepsilon,$$

and

$$\begin{aligned} z^{\frac{1}{2}} + z^{-\frac{1}{2}} &= (1-\varepsilon)^{\frac{1}{2}} e^{n\pi i/(p+\frac{1}{2})} + (1-\varepsilon)^{-\frac{1}{2}} e^{-n\pi i/(p+\frac{1}{2})} \\ &\approx (1 - \tfrac{1}{2}\varepsilon)(1 + \tfrac{1}{2}i(\tfrac{2n\pi}{p+\frac{1}{2}}) - \tfrac{1}{8}(\tfrac{2n\pi}{p+\frac{1}{2}})^2) \\ &\qquad + (1 + \tfrac{1}{2}\varepsilon)(1 - \tfrac{1}{2}i(\tfrac{2n\pi}{p+\frac{1}{2}}) - \tfrac{1}{8}(\tfrac{2n\pi}{p+\frac{1}{2}})^2) \\ &\approx 2 - \left(\tfrac{n\pi}{p+\frac{1}{2}}\right)^2 + \tfrac{1}{2}i\varepsilon(\tfrac{2n\pi}{p+\frac{1}{2}}). \end{aligned}$$

So equating the real parts, we find that the approximate value of ε is

$$\varepsilon \approx \frac{n^2\pi^2}{2(p+\frac{1}{2})^3}.$$

Using the approximation $\ln(1-\varepsilon) \approx -\varepsilon$, equation (7.8.2) gives

$$\text{Decay time} \approx \frac{2(p+\frac{1}{2})^3}{Nn^2\pi^2}$$

where N is the sample rate. This means that the lower harmonics are decaying more slowly than the higher harmonics, in accordance with the behavior of a plucked string.

Further reading:

D. A. Jaffe and J. O. Smith III, *Extensions of the Karplus–Strong plucked string algorithm*, Computer Music Journal 7 (2) (1983), 56–69. Reprinted in Roads [**117**], 481–494.

M. Karjalainen, V. Välimäki and T. Tolonen, *Plucked-string models: From the Karplus–Strong algorithm to digital waveguides and beyond*, Computer Music Journal 22 (3) (1998), 17–32.

K. Karplus and A. Strong, *Digital synthesis of plucked string and drum timbres*, Computer Music Journal 7 (2) (1983), 43–55. Reprinted in Roads [**117**], 467–479.

F. Richard Moore, *Elements of computer music* [**89**], page 279.

Curtis Roads, *The computer music tutorial* [**113**], page 293.

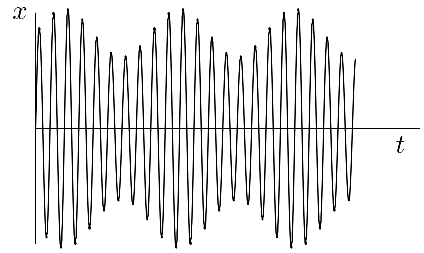
C. Sullivan, *Extending the Karplus–Strong plucked-string algorithm to synthesize electric guitar timbres with distortion and feedback*, Computer Music Journal 14 (3), 26–37.

8.7. Amplitude and frequency modulation

The familiar context for amplitude and frequency modulation is as a way of carrying audio signals on a radio frequency carrier (AM and FM radio). In the case of AM radio, the carrier frequency is usually in the range 500–2000 KHz, which is much greater than the frequency of the carried signal. The latter is encoded in the amplitude of the carrier. So for example a 700 KHz carrier signal modulated by a 440 Hz sine wave would be represented by the function

$$x = (A + B\sin(880\pi t))\sin(1400000\pi t),$$

where A is an offset to allow both positive and negative values of the waveform to be decoded.



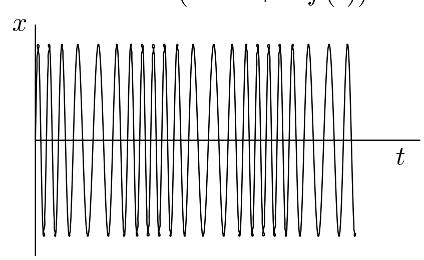
Decoding the received signal is easy. A diode is used to allow only the positive part of the wave through, and then a capacitor is used to smooth it out and remove the high frequency carrier wave. The resulting audio signal may then be amplified and put through a loudspeaker.

In the case of frequency modulation, the carrier frequency is normally around 90–120 MHz, which is even greater in comparison to the frequency of the carried signal. The latter is encoded in variations in the frequency of the carrier. So for example a 100 MHz carrier signal modulated by a 440 Hz sine wave would be represented by the function

$$x = A\sin(10^8.2\pi t + B\sin(880\pi t)).$$

The amplitude A is associated with the carrier wave, while the amplitude B is associated with the audio wave. More generally, an audio wave represented by $x = f(t)$, carried on a carrier of frequency ν and amplitude A, is represented by

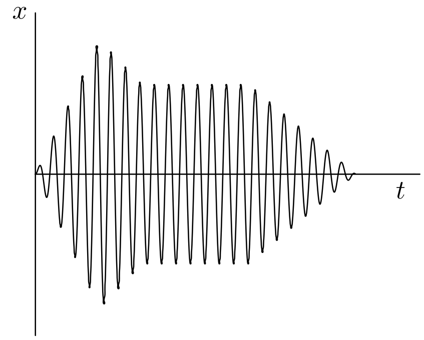
$$x = A\sin(2\pi\nu t + Bf(t)).$$



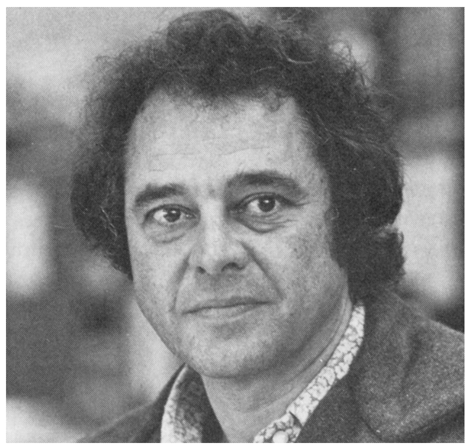
Decoding frequency modulated signals is harder than amplitude modulated signals, and will not be discussed here. But the big advantage is that it is less susceptible to noise, and so it gives cleaner radio reception.

An example of the use of amplitude modulation in the theory of synthesis is ring modulation. A ring modulator takes two inputs, and the output contains only the sum and difference frequencies of the partials of the inputs. This is generally used to construct waveforms with inharmonic partials, so as to impart a metallic or bell-like timbre. The method for constructing the sum and difference frequencies is to multiply the incoming amplitudes. Equations (1.7.7), (1.7.10) and (1.7.11) explain how this has the desired result. The origin of the term "ring modulation" is that in order to deal with both positive and negative amplitudes on the inputs and get the right sign for the outputs, four diodes were connected head to tail in a ring.

Another example of amplitude modulation is the application of envelopes, as discussed in §8.2. The waveform is multiplied by the function used to describe the envelope.



A great breakthrough in synthesis was achieved in the late nineteen sixties when John Chowning developed the idea of using frequency modulation instead of additive synthesis.

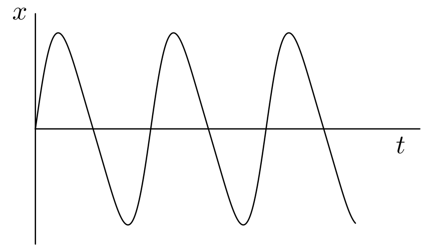


John Chowning

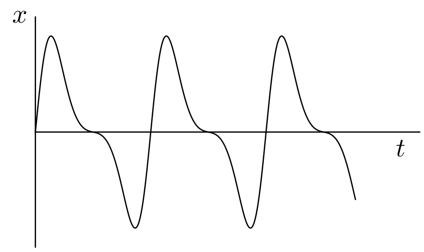
The idea behind FM synthesis or *frequency modulation synthesis* is similar to FM radio, but the carrier and the signal are both in the audio range, and usually related by a small rational frequency ratio. So for example, a 440 Hz carrier and 440 Hz modulator would be represented by the function

$$x = A\sin(880\pi t + B\sin(880\pi t)).$$

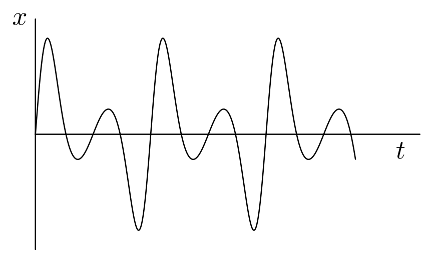
The resulting wave is still periodic with frequency 440 Hz, but has a richer harmonic spectrum than a pure sine wave. For small values of B, the wave is nearly a sine wave



whereas for larger values of B the harmonic content grows richer

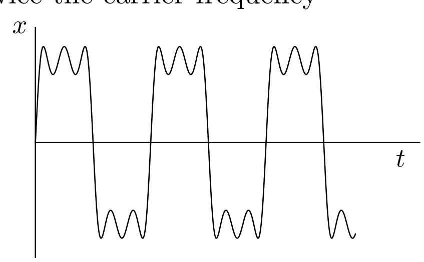


and richer.

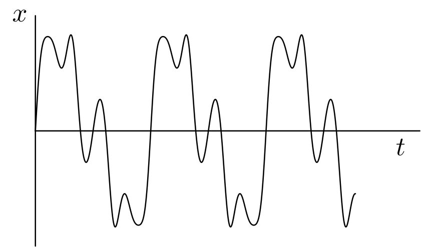


This gives a way of making an audio signal with a rich harmonic content relatively simply. If we wanted to synthesize the above wave using additive synthesis, it would be much harder.

Here are examples of frequency modulated waves in which the modulating frequency is twice the carrier frequency



and three times the carrier frequency.



In the next section, we discuss the Fourier series for a frequency modulated signal. The Fourier coefficients are called Bessel functions, for which the groundwork was laid in §2.8. We shall see that the Bessel functions may be interpreted as giving the amplitudes of side bands in a frequency modulated signal.

8.8. The Yamaha DX7 and FM synthesis

Yamaha DX7

The Yamaha DX7, which came out in the fall of 1983,[1] was the first affordable commercially available digital synthesizer. This instrument was the result of a long collaboration between John Chowning and Yamaha Corporation through the nineteen seventies. It works by FM synthesis, with six configurable "operators." An operator produces as output a frequency modulated sine wave, whose frequency is determined by the level of a modulating input, and whose envelope is determined by another input. The power of the method comes from hooking up the output of one such operator to the modulating input of another. In this section, we shall investigate FM synthesis in detail, using the Yamaha DX7 for the details of the examples. Most of the discussion translates easily to any other FM synthesizer. In Appendix B, there are tables which apply to various models of FM synthesizers. Later on, in §§8.11–8.12, we shall also investigate FM synthesis using the CSound computer music language.

The DX7 calculates the sine function in the simplest possible way. It has a digital lookup table of values of the function. This is much faster than any conceivable formula for calculating the function, but this is at the expense of having to commit a block of memory to this task.

[1]Original price US $2000; no longer manufactured but easy to obtain second hand for around US $250–$450.

Let us begin by examining a frequency modulated signal of the form

$$\sin(\omega_c t + I \sin \omega_m t). \tag{8.8.1}$$

Here, $\omega_c = 2\pi f_c$ where f_c denotes the carrier frequency, $\omega_m = 2\pi f_m$ where f_m denotes the modulating frequency, and I the index of modulation.

We first discuss the relationship between the index of modulation I, the maximal frequency deviation d of the signal, and the frequency f_m of the modulating wave. For this purpose, we make a linear approximation to the modulating signal at any particular time, and use this to determine the instantaneous frequency, to the extent that this makes sense. When $\sin \omega_m t$ is at a peak or a trough, namely when its derivative with respect to t vanishes, the linear approximation is a constant function, which then acts as a phase shift in the modulated signal. So at these points, the frequency is f_c. The maximal frequency deviation occurs when $\sin \omega_m t$ is varying most rapidly. This function *increases* most rapidly when $\omega_m t = 2n\pi$ for some integer n. Since the derivative of $\sin \omega_m t$ with respect to t is $\omega_m \cos \omega_m t$, which takes the value ω_m at these values of t, the linear approximation around these values of t is $\sin \omega_m t \simeq \omega_m t - 2n\pi$. So the function (8.8.1) approximates to

$$\sin(\omega_c t + I\omega_m(t - 2\pi)) = \sin((\omega_c + I\omega_m)t - 2\pi I\omega_m).$$

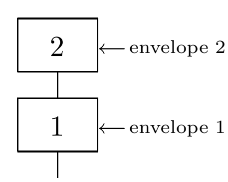
So the instantaneous frequency is $f_c + I f_m$. Similarly, $\sin \omega_m t$ *decreases* most rapidly when $\omega_m t = (2n+1)\pi$ for some integer n, and a similar calculation shows that the instantaneous frequency is $f_c - I f_m$. It follows that the maximal deviation in the frequency is given by

$$d = I f_m. \tag{8.8.2}$$

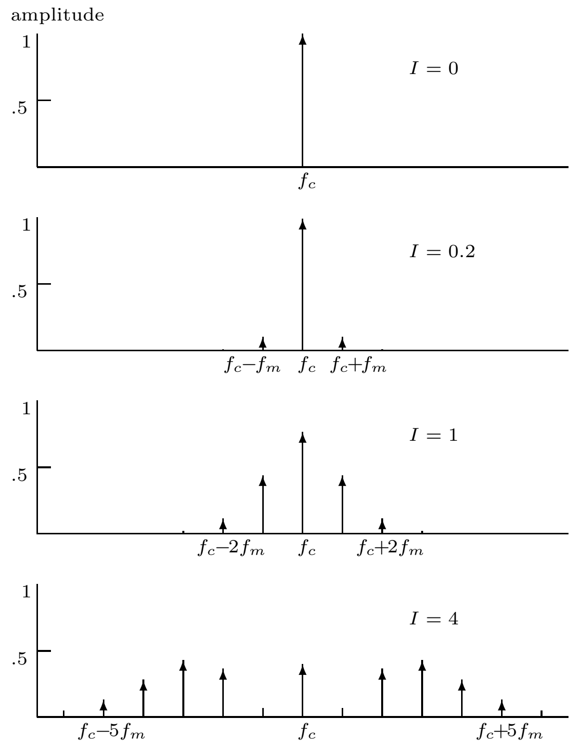
The Fourier series for functions of the form (8.8.1) were analysed in §2.8 in terms of the Bessel functions.

Putting $\phi = \omega_c t$, $z = I$ and $\theta = \omega_m t$ in equation (2.8.9), we obtain the fundamental equation for frequency modulation:

$$\sin(\omega_c t + I \sin \omega_m t) = \sum_{n=-\infty}^{\infty} J_n(I) \sin(\omega_c + n\omega_m)t. \tag{8.8.3}$$



The interpretation of this equation is that for a frequency modulated signal with carrier frequency f_c and modulating frequency f_m, the frequencies present in the modulated signal are $f_c + n f_m$. Notice that positive and negative values of n are allowed here. The component with frequency $f_c + n f_m$ is called the nth *side band* of the signal. Thus the Bessel function $J_n(I)$ is giving the amplitude of the nth side band in terms of the index of modulation. The block diagram on the DX7 for frequency modulating a sine wave in this fashion is as shown above. The box marked "1" represents the operator producing the carrier signal and the box marked "2" represents the operator producing the modulating signal.



Each operator has its own envelope, which determines how its amplitude develops with time. So envelope 1 determines how the amplitude of the final signal varies with time, but it is less obvious what envelope 2 is determining. Since the output of operator 2 is frequency modulating operator 1, the amplitude of the output can be interpreted as the index of modulation I. For small values of I, $J_0(I)$ is much larger than any other $J_n(I)$ (see the graphs in Section 2.8), and so operator 1 is producing an output which is nearly a pure sine wave, but with other frequencies present with small amplitudes. However, for larger values of I, the spectrum of the output of operator 1 grows richer in harmonics. For any particular value of I, as n gets larger, the amplitudes $J_n(I)$ eventually tend to zero. But the point is that for small values of I, this happens more quickly than for larger values of I, so the harmonic spectrum gives a purer note for small values of I and a richer sound for larger values of I. So envelope 2 is controlling the *timbre* of the output of operator 1.

Example. Suppose that we have a carrier frequency of 3ν and a modulating frequency of 2ν. Then the zeroth side band has frequency 3ν, the first 5ν, the second 7ν, and so on. But there are also side bands corresponding to negative values of n. The minus first side band has frequency ν. But there's no reason to stop there, just because the next side band has negative frequency $-\nu$. The point is that a sine wave with frequency $-\nu$ is just the same as a sine wave with frequency ν but with the amplitude negated. So really the way to think of it is that the side bands with negative frequency undergo reflection to make the corresponding positive frequency.

Notice also in this example that $3+2n$ is always an odd number, so only odd multiples of ν appear in the resulting frequency spectrum. In general, the frequency spectrum will depend in an interesting way on the ratio of f_m to f_c. If the ratio is a ratio of small integers, the resulting frequency spectrum will consist of multiples of a fundamental frequency. Otherwise, the spectrum is said to be *inharmonic*.

Let us calculate the spectrum in this example for various values of I. First we use a small value such as $I = 0.2$. Consulting Appendix B, we see that $J_0(I) \approx 0.9900$, $J_1(I) \approx 0.0995$, $J_2(I) \approx 0.0050$ and $J_n(I)$ is negligibly small for $n \geq 3$. Using equation (2.8.4) ($J_{-n}(I) = (-1)^n J_n(I)$), we see that $J_{-1}(I) \approx -0.0995$, $J_{-2}(I) \approx 0.0050$ and $J_{-n}(I)$ is negligibly small for $n \geq 3$. So the frequency modulated signal is approximately

$$\begin{aligned} 0.0050\sin(2\pi(-\nu)t) - 0.0995\sin(2\pi\nu t) + 0.9900\sin(2\pi(3\nu)t) \\ + 0.0995\sin(2\pi(5\nu)t) + 0.0050\sin(2\pi(7\nu)t). \end{aligned}$$

Since $\sin(-x) = -\sin(x)$, this is

$$-0.1045\sin(2\pi\nu t) + 0.9900\sin(6\pi\nu t) + 0.0995\sin(10\pi\nu t) + 0.0050\sin(14\pi\nu t).$$

This will be perceived as a note with fundamental frequency ν, but with very strong third harmonic.

Now let us carry out the same calculation with a larger value of I, say $I = 3$. Again consulting Appendix B, we see that $J_0(I) \approx -0.2601$, $J_1(I) \approx 0.3391$, $J_2(I) \approx 0.4861$, $J_3(I) \approx 0.3091$, $J_4(I) \approx 0.1320$, $J_5(I) \approx 0.0430$, $J_6(I) \approx 0.0114$, $J_7(I) \approx 0.0025$, $J_8(I) \approx 0.0005$, and only around $n \geq 8$ is $J_n(I)$ negligibly small. So the harmonic spectrum of the resulting frequency modulated signal is much richer, and the first few terms are given by

$$\begin{aligned} -\, 0.0430\sin(2\pi(-7\nu)t) + 0.1320\sin(2\pi(-5\nu)t) - 0.3091\sin(2\pi(-3\nu)t) \\ + 0.4861\sin(2\pi(-\nu)t) - 0.3991\sin(2\pi\nu t) - 0.2601\sin(2\pi(3\nu)t) \\ + 0.3391\sin(2\pi(5\nu)t) + 0.4861\sin(2\pi(7\nu)t) \end{aligned}$$

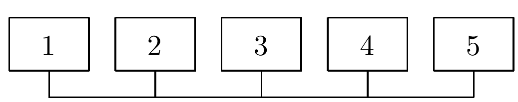
which makes

$$-0.8852\sin(2\pi\nu t) + 0.0490\sin(6\pi\nu t) + 0.2071\sin(10\pi\nu t) + 0.5291\sin(14\pi\nu t),$$

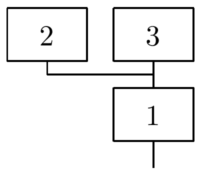
but it is clear that even higher harmonics than this are present with fairly large magnitude, up to about the seventeenth harmonic ($3 + 2 \times 7 = 17$), and then it starts tailing off. So the resulting note is very rich in harmonics. Notice also how we have conspired to choose I so that the amplitude of the third harmonic is now very small.

Suppose, for example, that operator 2 is assigned an envelope which starts at zero, peaks near the beginning, and then tails off to zero. Then the resulting frequency modulated signal will start off as a pure sine wave, fairly quickly attain a rich harmonic spectrum, and then tail off again into a fairly pure sine wave. It is easy to see that the possibilities opened up with even two operators is fairly wide.

In terms of block diagrams, additive synthesis for a waveform with five sinusoidal components is represented as follows.



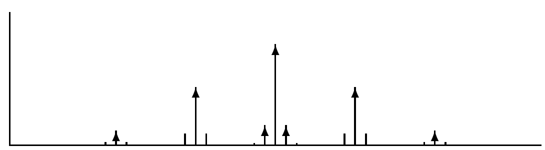
So in the above example, to synthesize the corresponding sound additively would require a large number of oscillators. The exact number would depend on where the cutoff for audibility occurs.



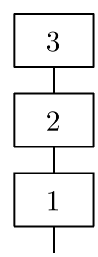
The DX7 allows a large number of different configurations or "algorithms" which mix additive and FM components. So for example if two sinusoidal waveforms of different frequencies are added together and the result used to modulate another sine wave, then the block diagram is as shown to the left. Oscillators labeled 2 and 3 are added together and used to modulate oscillator 1. The corresponding waveform is given by

$$\sin(\omega_1 t + I_2 \sin\omega_2 t + I_3 \sin\omega_3 t) = \sum_{n_2=-\infty}^{\infty} \sum_{n_3=-\infty}^{\infty} J_{n_2}(I_2) J_{n_3}(I_3) \sin(\omega_1 + n_2\omega_2 + n_3\omega_3)t.$$

So the side bands have frequencies given by adding positive and negative multiples of the two modulating frequencies to the carrier frequency in all possible ways. The amplitudes of these side bands are given by multiplying the corresponding values of the Bessel functions.

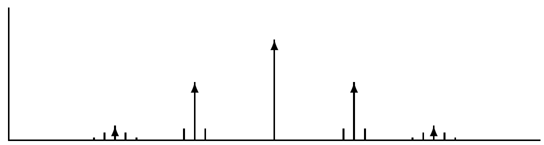


Another possible configuration is a cascade in which the modulating signal is also modulated. This should be thought of as equivalent to a larger number of added sine waves modulating a single sine wave, in an extension of the previous discussion. The block diagram for this configuration is shown to the right. The corresponding formula is obtained by feeding formula (8.8.3) into itself, giving



$$\begin{aligned}
&\sin(\omega_1 t + I_2 \sin(\omega_2 t + I_3 \sin\omega_3 t)) \\
&\quad = \sum_{n_2=-\infty}^{\infty} J_{n_2}(I_2) \sin(\omega_1 t + n_2\omega_2 t + n_2 I_3 \sin\omega_3 t) \\
&\quad = \sum_{n_2=-\infty}^{\infty} \sum_{n_3=-\infty}^{\infty} J_{n_2}(I_2) J_{n_3}(n_2 I_3) \sin(\omega_1 + n_2\omega_2 + n_3\omega_3)t.
\end{aligned}$$

Here, the subscripts 2 and 3 correspond to the numbering on the oscillators in the diagram. Again, the frequencies of the side bands are given by adding positive and negative multiples of the two modulating frequencies to the carrier frequency in all possible ways. But this time, the amplitudes of the side bands are given by the more complicated formula $J_{n_2}(I_2)J_{n_3}(n_2I_3)$. The effect of this is that the number of the side band on the second operator is used to scale the size of the index of modulation of the third operator. In particular, the original frequency has no side bands corresponding to the third operator, while the more remote side bands of the second are more heavily modulated.



Exercises

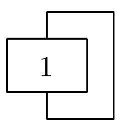
1. Find the amplitudes of the first few frequency components of the frequency modulated wave

$$y = \sin(440(2\pi t) + \tfrac{1}{10}\sin 660(2\pi t)).$$

Stop when the frequency components are attenuated by at least 100dB from the strongest one.

You will need to use the tables of Bessel functions in Appendix B. Also remember that power is proportional to square of amplitude, so that dividing the amplitude by 10 attenuates the signal by 20dB.

8.9. Feedback, or self-modulation



One final twist in FM synthesis is feedback, or self-modulation. This involves the output of an oscillator being wrapped back round and used to modulate the input of the same oscillator. This corresponds to the block diagram on the left, and the corresponding equation is given by

$$f(t) = \sin(\omega_c t + If(t)). \tag{8.9.1}$$

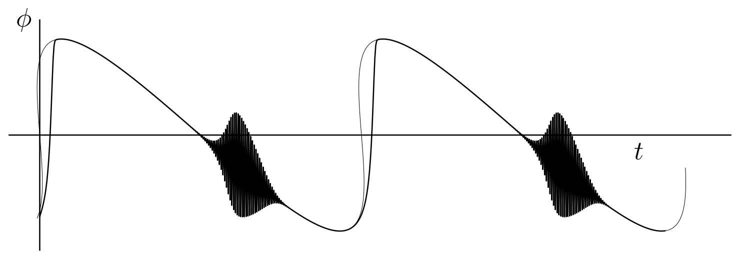
We saw in §2.11 that this equation only has a unique solution provided $|I| \le 1$, and that then it defines a periodic function of t. The Fourier series is given in equation (2.11.4) as

$$f(t) = \sum_{n=1}^{\infty} \frac{2J_n(nI)}{nI}\sin(n\omega_c t).$$

For values of I satisfying $|I| > 1$, equation (8.9.1) no longer has a single valued continuous solution (see §2.11), but it still makes sense in the form of a recursion defining the next value of $f(t)$ in terms of the previous one,

$$f(t_n) = \sin(\omega_c t_n + If(t_{n-1})). \tag{8.9.2}$$

Here, t_n is the nth sample time, and the sample times are usually taken to be equally spaced. The effect of this equation is not quite intuitively obvious. As might be expected, the graph of this function stays close to the solution to equation (8.9.1) when this is unique. When it is no longer unique, it continues going along the same branch of the function as long as it can, and then jumps suddenly to the one remaining branch when it no longer can. But the feature which it is easy to overlook is that there is a slightly delayed instability for small values of $f(t)$. Here is a graph of the solutions to equations (8.9.1) (red) and (8.9.2) (black) superimposed.



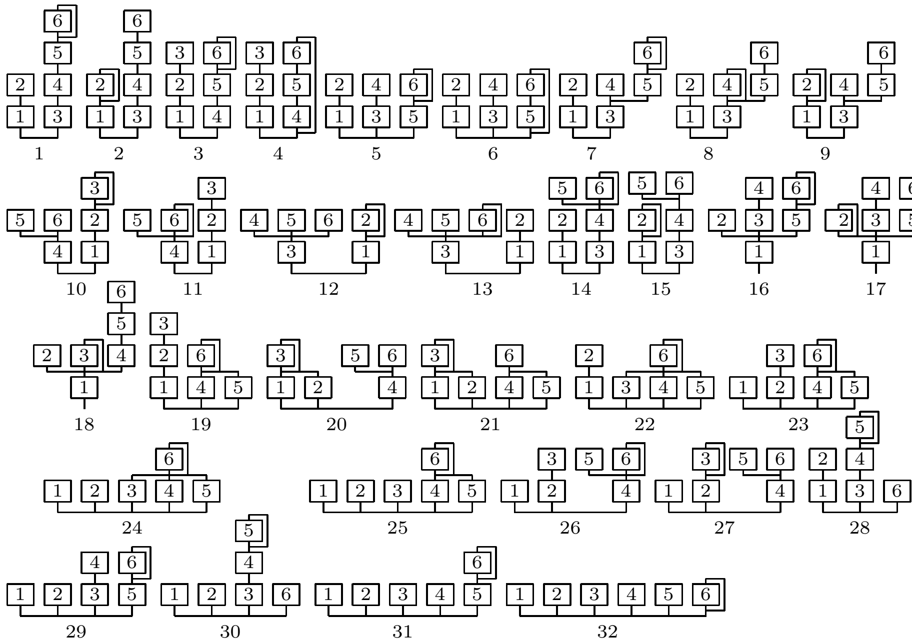
The effect of the instability is to introduce a wave packet whose frequency is roughly half the sampling frequency. Usually the sampling frequency is high enough that the effect is inaudible.

Feedback for a stack of two or more oscillators is also used. It seems hard to analyse this mathematically, and often the result is perceived as "noise." According to Slater (reference given on page 261), as the index of modulation increases, the behavior of a stack of two FM oscillators with different frequencies, each modulating the other, exhibits the kind of bifurcation that is characteristic of chaotic dynamical systems. This subject needs to be investigated further.

In the DX7, there are a total of six oscillators. The process of designing a patch[2] begins with a choice of one of 32 given configurations, or "algorithms" for these oscillators. Each oscillator is given an envelope whose parameters are determined by the patch, so that the amplitude of the output of each oscillator varies with time in a chosen manner. Here is a table of the 32 available algorithms.

---

[2] Yamaha uses the nonstandard terminology "voice" instead of the more usual "patch."



Not all the operators have to be used in a given patch. The operators which are not used can just be switched off. Output level is an integer in the range 0–99; index of modulation is not a linear function of output level, but rather there is a complicated recipe for causing an approximately exponential relationship. A table showing this relationship for various different FM synthesizers can be found in Appendix B.

We now start discussing how to use FM synthesis to produce various recognisable kinds of sounds. In order to sound like a brass instrument such as a trumpet, it is necessary for the very beginning of the note to be an almost pure sine wave. Then the harmonic spectrum grows rapidly richer, overshooting the steady spectrum by some way, and then returning to a reasonably rich spectrum. When the note stops, the spectrum decays rapidly to a pure note and then disappears altogether. This effect may be achieved with FM synthesis by using two operators, one modulating the other. The modulating operator is given an envelope looking like the one on page 239. The carrier operator uses a very similar envelope to control the amplitude.

Next, we discuss woodwind instruments such as the flute, as well as organ pipes. At the beginning of the note, in the attack phase, higher harmonics dominate. They then decrease in amplitude until in the steady state, the fundamental dominates and the higher harmonics are not very strong. This can be achieved either by making the modulating operator have an envelope looking like the one on page 239 only upside down, or by making the carrier frequency a small integer multiple of the modulating frequency so that for small values of the index of modulation, this higher frequency dominates. In any case, the decay phase for the modulating operator should be omitted for a more realistic sound. For some woodwind instruments such as the clarinet, it is necessary to make sure that predominantly odd harmonics are present. This can be achieved, as in the example on page 254, by setting $f_c = 3f$ and $f_m = 2f$, or some variation on this idea.

Percussive sounds have a very sharp attack and a roughly exponential decay. So an envelope looking like the graph of $x = e^{-t}$ is appropriate for the amplitude. Usually a percussion instrument will have an inharmonic spectrum, so that it is appropriate to make sure that f_c and f_m are not in a ratio which can be expressed as a ratio of small integers. We saw in Exercise 1 of §6.2 that the golden ratio is in some sense the number furthest from being able to be approximated well by ratios of small integers, so this is a good choice for producing spectra which will be perceived as inharmonic. Alternatively, the analysis carried out in §3.6 can be used to try to emulate the frequency spectrum of an actual drum.

Section §8.10 and the ones following it consist of an introduction to the public domain computer music language CSound. One of our goals will be to describe explicit implementations of two operator FM synthesis realizing the above descriptions.

Further reading on FM synthesis:

J. Bate, *The effect of modulator phase on timbres in FM synthesis*, Computer Music Journal 14 (3) (1990), 38–45.

John Chowning, *The synthesis of complex audio spectra by means of frequency modulation*, J. Audio Engineering Society 21 (7) (1973), 526–534. Reprinted as chapter 1 of Roads and Strawn [**116**], pages 6–29.

John Chowning, *Frequency modulation synthesis of the singing voice*, appeared in Mathews and Pierce [**82**], chapter 6, pages 57–63.

John Chowning and David Bristow, *FM theory and applications* [**18**].

L. Demany and K. I. McAnally, *The perception of frequency peaks and troughs in wide frequency modulations*, J. Acoust. Soc. Am. 96 (1994), 706–715.

L. Demany and S. Clément, *The perception of frequency peaks and troughs in wide frequency modulations, II. Effects of frequency register, stimulus uncertainty, and intensity*, J. Acoust. Soc. Am. 97 (1995), 2454–2459; *III. Complex carriers*, J. Acoust. Soc. Am. 98 (1995), 2515–2523; *IV. Effect of modulation waveform*, J. Acoust. Soc. Am. 102 (1997), 2935–2944.

A. Horner, *Double-modulator FM matching of instrument tones*, Computer Music Journal 20 (2) (1996), 57–71.

A. Horner, *A comparison of wavetable and FM parameter spaces*, Computer Music Journal 21 (4) (1997), 55–85.

A. Horner, J. Beauchamp and L. Haken, *FM matching synthesis with genetic algorithms*, Computer Music Journal 17 (4) (1993), 17–29.

M. LeBrun, *A derivation of the spectrum of FM with a complex modulating wave*, Computer Music Journal 1 (4) (1977), 51–52. Reprinted as chapter 5 of Roads and Strawn [**116**], pages 65–67.

F. Richard Moore, *Elements of computer music* [**89**], pages 316–332.

D. Morrill, *Trumpet algorithms for computer composition*, Computer Music Journal 1 (1) (1977), 46–52. Reprinted as chapter 2 of Roads and Strawn [**116**], pages 30–44.

C. Roads, *The computer music tutorial* [**113**], pages 224–250.

S. Saunders, *Improved FM audio synthesis methods for real-time digital music generation*, Computer Music Journal 1 (1) (1977), 53–55. Reprinted as chapter 3 of Roads and Strawn [**116**], pages 45–53.

W. G. Schottstaedt, *The simulation of natural instrument tones using frequency modulation with a complex modulating wave*, Computer Music Journal 1 (4) (1977), 46–50. Reprinted as chapter 4 of Roads and Strawn [**116**], pages 54–64.

D. Slater, *Chaotic sound synthesis*, Computer Music Journal 22 (2) (1998), 12–19.

B. Truax, *Organizational techniques for $c:m$ ratios in frequency modulation*, Computer Music Journal 1 (4) (1977), 39–45. Reprinted as chapter 6 of Roads and Strawn [**116**], pages 68–82.

8.10. CSound

CSound is a public domain synthesis program written by Barry Vercoe at the Media Lab in MIT in the C programming language. It has been compiled for various platform, and both source code and executables are freely available.

The program takes as input two files, called the *orchestra* file and the *score* file. The orchestra file contains the instrument definitions, or how to synthesize the desired sounds. It makes use of almost every known method of synthesis, including FM synthesis, the Karplus–Strong algorithm, phase vocoder, pitch envelopes, granular synthesis and so on, to define the instruments. The score file uses a language similar in conception to MIDI but different in execution, in order to describe the information for playing the instruments, such as amplitude, frequency, note durations and start times. The utility MIDI2CS mentioned in Appendix G provides a flexible way of turning MIDI files into CSound score files. The final output of the CSound program is a file in some chosen sound format, for example a WAV file or an AIFF file, which can be played through a computer sound card, downloaded into a synthesizer with sampling features, or written onto a CD.

We limit ourselves to a brief description of some of the main features of CSound, with the objective of getting as far as describing how to realise FM synthesis. The examples are adapted from the CSound manual.

Getting it. The source code and executables for a number of platforms, including Linux, Mac, MS-DOS and Windows can be obtained from

www.cs.bath.ac.uk/~jpff/dream.html

(files are at ftp://ftp.cs.bath.ac.uk/pub/dream/)

as can the manual and some example files. For a minimalist installation on an MS-DOS (or Windows) based machine, get the file `csound_new419.zip` from the subdirectory newest/ at the above site (or a later version if available; the above version was released on Mar 28, 2002). Unzip it[3] into a directory of your choice, and make sure the directory is in your path by editing the `autoexec.bat` file if necessary. If you are *really* short of space, delete everything except the files `csound.txt`, `csound.exe` and `dos4gw.exe` (total around 1 megabyte), and you will still be able to run all the examples described here. Make a new subdirectory for your orchestra and score files, and run CSound from that subdirectory. Instructions for running CSound can be found on page 264.

If you are running in an MS-DOS window under Windows 95 or higher, the above still works, but the file `csound_con419.zip` contains a smaller and more efficient version of just the `csound.exe` file and the `csound.txt` file; you won't need `dos4gw.exe`. The disadvantage is that the displays are in ascii instead of full screen graphics. There is also a Windows front end in `csound_win419.zip`. This is capable of realtime sound output and realtime MIDI handling, which the MS-DOS version is not, but apart from that, it is quite primitive. For example, the program needs to restart every time it is run, and cannot just replay the output.

The most up to date version of the manual is version 4.10 (March 24, 2001), which can be found at

lakewoodsound.com/csound/download.htm

This version does not seem to have made it onto the Bath ftp site mentioned above.

The orchestra file. This file has two main parts, namely the *header* section, which defines the sample rate, control rate, and number of output channels, and the *instrument* section which gives the instrument definitions. Each instrument is given its own number, which behaves like a patch number on a synthesizer.

The header section has the following format (everything after a semicolon is a comment):

```
sr = 44100 ; sample rate in samples per second
kr = 4410  ; control rate in control signals per second
ksmps = 10 ; ksmps = sr/kr must be an integer,
           ; samples per control period
nchnls = 1 ; number of channels                          (8.10.1)
```

An instrument definition consists of a collection of statements which generate or modify a digital signal. For example the statements

```
instr 1
   asig oscil 10000, 440, 1
```

[3]To unzip a file under *Windows*, get hold of *Winzip* from winzip.com. This is shareware, but can be used indefinitely without registration. If you prefer to use a free utility, get hold of Info-ZIP's free MS-DOS based program `unzip.exe` (138 kB) from www.cdrom.com/pub/infozip/UnZip.html

```
    out asig
endin
```
(8.10.2)

generate a 440 Hz wave with amplitude 10000, and send it to an output. The two lines of code representing the waveform generator are encased in a pair of statements which define this to be an instrument. For WAV file output, the possible range of amplitudes before clipping takes effect is from -32768 to $+32767$, for a total of 2^{15} possible values (see §7.3). The final argument `1` is a waveform number. This determines which waveform is taken from an `f` statement in the score file (see below). In our first example below, it will be a sine wave. The label `asig` is allowed to be any string beginning with `a` (for "audio signal"). So for example `a1` would have worked just as well. The `oscil` statement is one of CSound's many signal generators, and its effect is to output periodic signals made by repeating the values passed to it, appropriately scaled in amplitude and frequency. There is also another version called `oscili`, with the same syntax, which performs linear interpolation rather than truncation to find values at points between the sample points. This is slower by approximately a factor of two, but in some situations it can lead to better sounding output. In general, it seems to be better to use `oscil` for sound waves and `oscili` for envelopes (see page 266).

As it stands, the instrument (8.10.2) isn't very useful, because it can only play one pitch. To pass a pitch, or other attributes, as parameters from the score file to the orchestra file, an instrument uses variables named `p1`, `p2`, `p3`, and so on. The first three have fixed meanings, and then `p4`, `p5`, ... can be given other meanings. If we replace 440 by `p5`,

```
  asig oscil 10000, p5, 1
```

then the parameter `p5` will determine pitch.

The score file. Each line begins with a letter called an *opcode*, which determines how the line is to be interpreted. The rest of the line consists of numerical parameter fields `p1`, `p2`, `p3`, and so on. The possible opcodes are:

`f` (function table generator),

`i` (instrument statement; i.e., play a note),

`t` (tempo),

`a` (advance score time; i.e., skip parts),

`b` (offset score time),

`v` (local textual time variation),

`s` (section statement),

`r` (repeat sections),

`m` and `n` (repeat named sections),

`e` (end of score),

`c` (comment; semicolon is preferred).

If a line of the score file does not begin with an opcode, it is treated as a continuation line.

Each parameter field consists of a floating point number with optional sign and optional decimal point. Expressions are not permitted.

An `f` statement calls a subroutine to generate a set of numerical values describing a function. The set of values is intended for passing to the orchestra file for use by an instrument definition. The available subroutines are called `GEN01, GEN02, ....` Each takes some number of numerical arguments. The parameter fields of an `f` statement are as follows.

`p1` Waveform number

`p2` When to begin the table, in beats

`p3` Size of table; a power of 2, or one more, maximum 2^{24}

`p4` Number of `GEN` subroutine

`p5, p6, ...` Parameters for `GEN` subroutine

Beats are measured in seconds, unless there is an explicit `t` (tempo) statement; in our examples, `t` statements are omitted for simplicity.

So for example, the statement

```
f1 0 8192 10 1
```

uses `GEN10` to produce a sine wave, starting "now," of size 8192, and assigns it to waveform 1. The subroutine `GEN10` produces waveforms made up of weighted sums of sine waves, whose frequencies are integer multiples of the fundamental. So for example

```
f2 0 8192 10 1 0 0.5 0 0.333
```

produces the sum of the first five terms in the Fourier series for a square wave, and assigns it to waveform 2.

An `i` statement activates an instrument. This is the kind of statement used to "play a note." Its parameter fields are as follows.

`p1` Instrument number

`p2` Starting time in beats

`p3` Duration in beats

`p4, p5, ...` Parameters used by the instrument

An `e` statement denotes the end of a score. It consists of an `e` on a line on its own. Every score file must end in this way.

For example, if instrument `1` is given by (8.10.2) then the score file

```
f1 0 8192 10 1 ; use GEN10 to create a sine wave
i1 0 4         ; play instr 1 from time 0 for 4 secs
e
```
(8.10.3)

will play a 440Hz tone for 4 seconds.

Running CSound. The program CSound was designed as a command line program, and although various front ends have been designed for it, the command line remains the most convenient method. Having installed CSound according to the instructions that accompany the program, the procedure is to create an orchestra file called `<filename>.orc` and a score file called `<filename>.sco` using your favorite (ascii) text processor.[4] The basic syntax for running CSound is

```
csound <flags> <filename>.orc <filename>.sco
```

For example, if your files are called `ditty.orc` and `ditty.sco`, and you want a WAV file output, then use the `-W` flag (this is case sensitive).

```
csound -W ditty.orc ditty.sco
```

This will produce as output a file called `test.wav`. If you want some other name, it must be specified with the `-o` flag.

```
csound -W -o ditty.wav ditty.orc ditty.sco
```
(8.10.4)

If you want to suppress the graphical displays of the waveforms, which csound gives by default, this is achieved with the `-d` flag.

We are now ready to run our first example. Make two text files, one called `ditty.orc` containing the statements (8.10.1) followed by (8.10.2), and one called `ditty.sco` containing the statements (8.10.3). If the program is properly installed, then typing the command (8.10.4) at the command line should produce a file ditty.wav. Playing this file through a sound card or other audio device should then sound a pure sine wave at 440Hz for 4 seconds.

Warning. Both the orchestra and the score file are case sensitive. If you are having problems running CSound on the above orchestra and score files, check that you have typed everything in lower case.

There is also an annoying feature, which is that if the last line of text in the input file does not have a carriage return, then a wave file will be generated, but it will be unreadable. So it is best to leave a blank line at the end of each file.

Our "ditty" wasn't really very interesting, so let's modify it a bit. In order to be able to vary the amplitude and pitch, let us modify the instrument (8.10.2) to read

```
instr 1
   asig oscil p4, p5, 1 ; p4 = amplitude, p5 = frequency
      out asig
endin
```

[4]Word processors such as Word Perfect or Word by default save files with special formatting characters embedded in them. CSound will choke on these characters. In MS-DOS, the command

```
edit <filename>
```

will invoke a simple ascii text processor whose output will not choke CSound in this way. If you are running in an MS-DOS box inside Windows, the command

```
notepad <filename>
```

will start up the ascii text processor called `notepad` in a separate window, which is more convenient for switching between the editor and running CSound.

(8.10.5)

Now we can play the first ten notes of the harmonic series (see page 133) using the following score file.

```
f1 0 8192 10 1          ; sine wave
i1 0.0 0.4 32000  261.6 ; fundamental (C, to nearest tenth of a Hz)
i1 0.5 0.4 24000  523.2 ; second harmonic, octave
i1 1.0 0.4 16000  784.8 ; third harmonic, perfect fifth
i1 1.5 0.4 12000 1046.4 ; fourth harmonic, octave
i1 2.0 0.4  8000 1308.0 ; fifth harmonic, just major third
i1 2.5 0.4  6000 1569.6 ; sixth harmonic, perfect fifth
i1 3.0 0.4  4000 1831.2 ; seventh harmonic, listen carefully to this
i1 3.5 0.4  3000 2092.8 ; eighth harmonic, octave
i1 4.0 0.4  2000 2354.4 ; nineth harmonic, just major second
i1 4.5 0.4  1500 2616.0 ; tenth harmonic, just major third
e
```

(8.10.6)

This file plays a series of notes at half second intervals, each lasting 0.4 seconds, at successive integer multiples of 220Hz, and at steadily decreasing amplitudes. Make an orchestra file from (8.10.1) and (8.10.5), and a score file from (8.10.6), run CSound as before, and listen to the results.

Data rates. Recall from (8.10.1) that the header of the orchestra file defines two rates, namely the *sample rate* and the *control rate*. There are three different kinds of variables in CSound, which are distinguished by how often they get updated. `a`-rate variables, or audio rate variables, are updated at the sample rate, while the `k`-rate variables, or control rate variables, are updated at the control rate. Audio signals should be taken to be `a`-rate, while an envelope, for example, is usually assigned to a `k`-rate variable. It is possible to make use of audio rate signals for control, but this will increase the computational load. A third kind of variable, the `i`-rate variable, is updated just once when a note is played. These variables are used primarily for setting values to be used by the instrument. The first letter of the variable name (`a`, `k` or `i`) determines which kind of variable it is.

The variables discussed so far are all *local* variables. This means that they only have meaning within the given instrument. The same variable can be reused with a different meaning in a different instrument. There are also global versions of variables of each of these rates. These have names beginning with `ga`, `gk` and `gi`. Assignment of a global variable is done in the *header* section of the orchestra file.

Envelopes. One way to apply an envelope is to make an oscillator whose frequency is `1/p3`, the reciprocal of the duration, so that exactly one copy of the waveform is used each time the note is played. It is better to use `oscili` rather than `oscil` for envelopes, because many sample points of the envelope will be used in the course of the one period. So for example

```
kenv oscili p4, 1/p3, 2
```

uses waveform 2 to make an envelope. The first letter `k` of the variable name `kenv` means that this is a control rate variable. It would work just as well to make it an audio rate variable by using a name like `aenv`, but it would demand greater computation time, and result in no audible improvement.

The subroutine `GEN07`, which performs linear interpolation, is ideal for an envelope made from straight lines. The arguments `p4, p5, ...` of this subroutine alternate between numbers of points and values. So for example, the statement

```
f2 0 513 7 0 80 1 50 0.7 213 0.7 170 0 ; ADSR envelope
```

in the score file produces an envelope resembling the one on page 239 with ADSR sections of length 80, 50, 213, 170 samples, with heights varying linearly

$$0 \to 1 \to 0.7 \to 0.7 \to 0,$$

and assigns it to waveform 2. The numbers of sample points in the sections should always add up to the total length `p3`.

Recall that the total number of sample points must be either a power of two, or one more than a power of two. It is usual to use a power of two for repeating waveforms. For waveforms that will be used only once, such as an envelope, we use one more than a power of two so that the number of intervals between sample points is a power of two.

To apply the envelope to the instrument (8.10.5), we replace `p4` with `kenv` to make

```
instr 1
   kenv oscili p4, 1/p3, 2 ; envelope from waveform 2
                           ; p4 = amplitude
   asig oscil kenv, p5, 1  ; p5 = frequency
      out asig
endin
```

It would also be possible to replace the waveform number `2` in the definition of `kenv` with another variable, say `p6`, to give a more general purpose shaped sine wave.

Exercises

1. Make orchestra and score files to play a major scale using a sine wave with an ADSR envelope. Check that your files work by running CSound on them and listening to the result.

8.11. FM synthesis using CSound

Here is the most basic two operator FM instrument:

```
instr 1
   amod oscil p6 * p7, p6, 1     ; modulating wave
                                 ; p6 = modulating frequency
                                 ; p7 = index of modulation
   kenv oscili p4, 1/p3, 2       ; envelope, p4 = amplitude
   asig oscil kenv, p5 + amod, 1 ; p5 = carrier frequency
      out asig
endin
```

(8.11.1)

The parameter p7 here represents the index of modulation; the reason why it is multiplied by p6 in the definition of the modulating wave amod is that the modulation is taking place directly on the frequency rather than on the phase. According to equation (8.8.2), this means that the index of modulation must be multiplied by the frequency of the modulating wave before being applied. The argument p5 + amod in the definition of asig is the carrier frequency p5 plus the modulating wave amod. The wave has been given an envelope kenv.

For a score file to illustrate this simple instrument, we introduce some useful abbreviations available for repetitive scores. First, note that the i statements in a score do not have to be in order of time of execution. The score is sorted with respect to time before it is played. The *carry* feature works as follows. Within a group of consecutive i statements in the score file (not necessarily consecutive in time) whose p1 parameters are equal, empty parameter fields take their value from the previous statement. An empty parameter field is denoted by a dot, with spaces between consecutive fields. Intervening comments or blank lines do not affect the carry feature, but other non-i statements turn it off.

For the second parameter field p2 only, the symbol + gives the value of p2 + p3 from the previous i statement. This begins a note at the time the last one ended. The symbol + may also be carried using the carry feature described above. Liberal use of the carry and + features greatly simplify typing in and subsequent alteration of a score. Here, then, is a score illustrating simple FM synthesis with $f_m = f_c$, with gradually increasing index of modulation.

```
f1 0 8192 10 1                        ; sine wave
f2 0 513 7 0 80 1 50 0.7 213 0.7 170 0 ; ADSR
i1 1 1 10000 200 200 0                ; index = 0 (pure sine wave)
i1 + .   .     .   .  1               ; index = 1
i1 + .   .     .   .  2               ; index = 2
i1 + .   .     .   .  3               ; index = 3
i1 + .   .     .   .  4               ; index = 4
i1 + .   .     .   .  5               ; index = 5
e
```

Sections. An s statement consisting of a single s on a line by itself ends a section and starts a new one. Sorting of i and f statements (as well as a, which we haven't discussed) is done by section, and the timing starts again at the beginning for each section. Inactive instruments and data spaces are purged at the end of a section, and this frees up computer memory.

The following score, using the same instrument (8.11.1), has three sections with different ratios $f_m : f_c$ and with gradually increasing index of modulation.

```
f1 0 8192 10 1 ; sine wave
i1 1 1 10000 200 200 0 ; index = 0, fm:fc = 1:1
```

```
i1 + .   .    .   .  1 ; index = 1
i1 + .   .    .   .  2 ; index = 2
i1 + .   .    .   .  3 ; index = 3
i1 + .   .    .   .  4 ; index = 4
i1 + .   .    .   .  5 ; index = 5
s
i1 1 1 10000 200 400 0 ; index = 0, fm:fc = 1:2
i1 + .   .    .   .  1 ; index = 1
i1 + .   .    .   .  2 ; index = 2
i1 + .   .    .   .  3 ; index = 3
i1 + .   .    .   .  4 ; index = 4
i1 + .   .    .   .  5 ; index = 5
s
i1 1 1 10000 400 200 0 ; index = 0, fm:fc = 2:1
i1 + .   .    .   .  1 ; index = 1
i1 + .   .    .   .  2 ; index = 2
i1 + .   .    .   .  3 ; index = 3
i1 + .   .    .   .  4 ; index = 4
i1 + .   .    .   .  5 ; index = 5
e
```

Pitch classes. CSound has a function `cpspch` for converting octave and pitch class notation in twelve tone equal temperament into frequencies in Hertz. This function may be used in an instrument definition, so that the instrument can be fed notes from the score file in this notation.

The octave and pitch class notation consists of a whole number, representing octave, followed by a decimal point and then two digits representing pitch class. The pitch classes are taken to begin with .00 for C and end with .11 for B, although higher values will just overlap into the next octave. The octave numbering is such that 8.00 represents middle C, 9.00 represents the octave above middle C, and so on. So for example the A above middle C can be represented as 8.09, or as 7.21, so that

$$\texttt{cpspch(8.09)} = \texttt{cpspch(7.21)} = 440.$$

Notes between two pitches on the twelve tone equal tempered scale can be represented by using further digits. So if four digits are used after the decimal point then the value is interpreted in cents. For example, if 8.00 represents middle C, then a just major third above this would be 8.0386, taken to the nearest cent.

8.12. Simple FM instruments

The bell. In this section, we use CSound and FM synthesis to imitate some instruments. We begin with the sound of a bell.[5] For a typical bell sound, we need an inharmonic spectrum. We can obtain this by using simple two operator FM synthesis where f_c and f_m have a ratio which cannot be expressed

[5]The examples in this section are adapted from an article of Chowning, reprinted as chapter 1 of [**116**].

as a simple ratio of two integers. The golden ratio is particularly good in this regard, for reasons explained in Exercise 1 of §6.2, so we take f_m to be 1.618 times f_c.

The bell sound is most easily made using envelopes representing exponential decay for both amplitude and timbre. The subroutine `GEN05` is designed for this. It performs *exponential interpolation*, which is based on the fact that between any two points (x_1, y_1) and (x_2, y_2) in the plane, with y_1 and y_2 positive, there is a unique exponential curve. It is given by

$$y = y_1^{\frac{x-x_2}{x_1-x_2}} y_2^{\frac{x-x_1}{x_2-x_1}}.$$

If y_1 and y_2 are both negative, replace them by the corresponding positive number in the above formula and then negate the final answer.

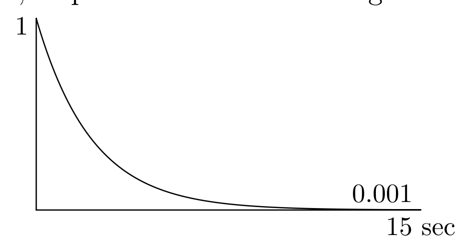
The fields for the `GEN05` subroutine are the same as for `GEN07` (see page 267), except that the values `p5, p7, ...` must all have the same sign. Referring back to the discussion of envelopes on page 266, we see that if we put

```
f2 0 513 5 1 513 .0001
```

in the score file and

```
   kenv oscili p4, 1/p3, 2
```

in the instrument definition, we will create an envelope with name `kenv` which decays exponentially from 1 to 0.0001. For a bell sound, we use an envelope like this for amplitude[6] and an envelope decaying exponentially from 1 to 0.001 scaled up by a factor of 10 for index of modulation. We also use a very long decay time, to permit the sound to linger.



This explains the following instrument definition. Pitches have been converted from octave and pitch class notation as explained above. In spite of the fact that lower frequency components are present, the perceived pitch of the note produced is equal to the carrier frequency.

```
instr 1 ; FM bell
ifc   = cpspch(p5)          ; carrier frequency
ifm   = cpspch(p5) * 1.618  ; modulating frequency
   kenv  oscili  p4, 1/p3, 2 ; envelope, p3 = duration, exp decay f2
                             ; p4 = amplitude
   ktmb oscili ifm * 10, 1/p3, 3 ; timbre envelope, max = 10,
                             ; exp decay f3
```

[6]Don't forget that amplitude is perceived logarithmically, so this sounds like a linear decrease, and indeed is a linear decrease when measured in decibels.

```
   amod oscil ktmb, ifm, 1   ; modulator
   asig oscil kenv, ifc + amod, 1 ; carrier
      out asig
endin
```

Here is the score file to play notes E, C, D, G for a chime, using this instrument.

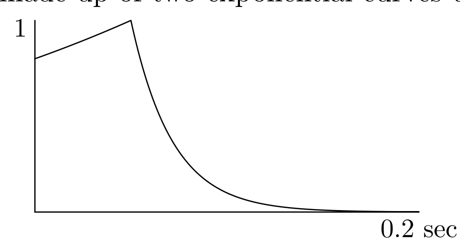
```
f1   0  8192    10  1
f2   0   513     5  1    513 .0001
f3   0   513     5  1    513 .001
i1   1    15  8000  8.04 ; 15 seconds at amplitude 8000 at middle C
i1   2.5   .     .  8.00
i1   4     .     .  8.02
i1   5.5   .     .  7.07
e
```

A general purpose instrument. It is not hard to modify the instrument described above to make a general purpose two operator FM synthesis instrument.

```
instr 1                         ; Two operator FM instrument
ifc = cpspch(p5) * p6           ; p6 = carrier frequency multiplier
ifm = cpspch(p5) * p7           ; p7 = modulator frequency multiplier
   kenv oscili p4, 1/p3, p8     ; p3 = duration
                                ; p4 = amplitude
                                ; p8 = carrier envelope
   ktmb oscili ifm * p10, 1/p3, p9 ; p9 = modulator envelope
                                ; p10 = maximum index of modulation
   amod oscil ktmb, ifm, 1      ; modulator
   asig oscil kenv, ifc + amod, 1  ; carrier
      out asig
endin
```

The rest of the examples in this section are described in terms of this setup.

The wood drum. To make a reasonably convincing wood drum, the amplitude envelope is made up of two exponential curves using GEN05,



while the envelope for the index of modulation is made up of two straight line segments, decreasing to zero and then staying there, using GEN07.

It turns out to be better to use a modulating frequency lower than the carrier frequency. So we use the reciprocal of the golden ratio, which is 0.618. We also use a large index of modulation, with a peak of 25, and a note duration of 0.2 seconds. This instrument works best in the octave going down from middle C. So the function table generators take the form

```
f1 0 8192  10  1                    ; sine wave
f2 0  513   5  .8 128  1 385 .0001 ; amplitude envelope
f3 0  513   7  1   64  0 449 0     ; modulating index envelope
```

and the instrument statements take the form

```
i1 <time> 0.2 <amplitude> <pitch> 1.0 0.618 2 3 25
```

Brass. For a brass instrument, we use a harmonic spectrum containing all multiples of the fundamental. This is easily achieved by taking $f_c = f_m$. The relative amplitude of higher harmonics is greater when the overall amplitude is greater, so the timbre and amplitude are given the same envelope. This is chosen to look like the ADSR curve on page 239, to represent an overshoot in intensity during the attack. The index of modulation does not want to be as great as in the above examples. A maximum index of 5 gives a reasonable sound. The envelope given below is suitable for a note of duration around 0.6 seconds. It would need to be modified slightly for other durations.

```
f1 0 8192 10 1 ; sine wave
f2 0 513 7 0 85 1 86 0.75 256 0.7 86 0 ; envelope for brass
```

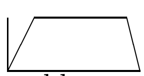
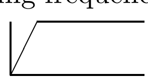
A typical note would then be represented by a statement of the form

```
i1 <time> 0.6 <amplitude> <pitch> 1.0 1.0 2 2 5
```

To improve the sound slightly on the brass tone presented here, we may wish to add a small deviation to the modulating frequency, so that there is a slight tremolo effect in the sound. If we replace the definition of the modulating frequency by the statement

```
ifm = cpspch(p5) * p7 + 0.5
```

then this will have the required effect.

Woodwind. For woodwind instruments, higher harmonics are present during the attack, and then the low frequencies enter. So we want the carrier frequency to be a multiple of the modulating frequency, and use an envelope of the form  for the carrier and  for the modulator. So the function table generators take the form

```
f1 0 8192 10 1                    ; sine wave
f2 0  513  7 0 50 1 443 1 20 0 ; amplitude envelope
```

```
f3 0  513  7 0 50 1 463 1      ; modulating index envelope
```

For a clarinet, where odd harmonics dominate, we take $f_c = 3f_m$ and a maximum index of 2. A basson sound is produced by giving the odd harmonics a more irregular distribution. This can be achieved by taking $f_c = 5f_m$ and a maximum index of 1.5.

8.13. Further techniques in CSound

The CSound language is vast. In this section, we cover just a few of the features which we have not touched on in the previous sections. For more information, see the CSound manual.

Tempo. The default tempo is 60 beats per minute, or one beat per second. To change this, a tempo statement is put in the score file. An example of the simplest form of tempo statement is

```
t 0 80
```

which sets the tempo to 80 beats per minute. The first argument (`p1`) of the tempo statement must always be zero. A tempo statement with more arguments causes accelerandos and ritardandos. The arguments are alternately times in beats (`p1` $= 0$, `p3`, `p5` $\ldots$) and tempi in beats per minute (`p2`, `p4`, `p6`, $\ldots$). The tempi between the specified times are calculate by making the durations of beats vary linearly. So for example the tempo statement

```
t 0 100 20 120 40 120
```

causes the initial tempo to be 100 beats per minute. By the twentieth beat, the tempo is 120 beats per minute. But the number of beats per minute is not linear between these values. Rather, the durations decrease linearly from 0.6 seconds to 0.5 seconds over the first twenty beats. The tempo is then constant from beat 20 until beat 40. By default, the tempo remains constant after the last beat where it is specified, so in this example the last two parameters are superfluous.

The tempo statement is only valid within the score section (cf. page 268) in which it is placed, and only one tempo statement may be used in each section. Its location within the section is irrelevant.

Stereo and Panning. For stereo output, we want to set `nchnls = 2` in the header of the orchestra file (8.10.1). In the instrument definition, instead of using `out`, we use `outs` with two arguments. So for example to do a simple pan from left to right, we might want the following lines in the instrument definition.

```
kpanleft lineseg 0, p3, 1
kpanright = 1 - kpanleft
   outs asig * kpanleft, asig * kpanright
```

The problem with this method of panning is that the total sound energy is proportional to the square of amplitude, summed over the two channels. So in the middle of the pan, the total energy is only $1/\sqrt{2}$ times the total enery on the left or right. So it sounds like there's a hole in the middle. The easiest way to correct this is to take the square root of the straight line produced by the signal generator `lineseg`. So for example we could have the following lines.

```
kpan lineseg 0, p3, 1
kpanleft  = sqrt(kpan)
kpanright = sqrt(1-kpan)
```

Since $\sin^2\theta + \cos^2\theta = 1$, another way to keep uniform total sound energy is as follows.

```
kpan lineseg 0, p3, 1
ipibytwo = 1.5708
kpanleft = sin(kpan * ipibytwo)
kpanright = sqrt((1 - kpan) * ipibytwo)
```

A good trick for obtaining what sounds like a wider sweep for the pan, especially when using headphones to listen to the output, is to make the angle go from $-\pi/4$ to $3\pi/4$ instead of 0 to $\pi/2$. This can be achieved by replacing the definition of `kpan` above with the following line.

```
kpan lineseg -0.5, p3, 1.5
```

A more realistic pan takes account of the fact that at the far reaches of the sweep, the sound should not be entirely concentrated in one channel. A slightly delayed version can be fed into the other channel, with delay varying up to about 0.7 seconds at the extreme end of the sweep.

Display and spectral display. There is a facility for displaying either a waveform or its spectrum, in an instrument file. So for example the instrument

```
instr 1
   asig oscil 10000 440 1
      out asig
      display asig p3
endin
```

is the same as (8.10.2), except that the extra line causes the graph of `asig` (of length `p3`) to be displayed. If the flag `-d` (see page 265) is set, this line makes no difference at all. Replacing the display line with

```
      dispfft asig p3, 1024
```

causes a fast Fourier transform of `asig` to be displayed, using an input window size of 1024 points. The number of points must be a power of two between 16 and 4096.

Arithmetic. In the orchestra file, variables represent signed floating point real numbers. The standard arithmetic operations `+`, `-`, `*` (times) and `/` (divide) can be used, as well as parentheses to any depth. Powers are denoted `a^b`, but `b` is not allowed to be audio rate. The expression `a % b` returns `a` reduced modulo `b`. Among the available functions are

`int` (integer part)
`frac` (fractional part)
`abs` (absolute value)
`exp` (exponential function, raises e to the given power)
`log` and `log10` (natural and base ten logarithm; argument must be positive)
`sqrt` (square root)
`sin`, `cos` and `tan` (sine, cosine and tangent, argument in radians)

`sininv`, `cosinv`, `taninv` (arcsine, arccos and arctan, answer in radians)
`sinh`, `cosh` and `tanh` (hyperbolic sine, cosine and tangent)
`rnd` (random number between zero and the argument)
`birnd` (random number bewteen plus and minus the argument)

Conditional values can also be used. For example,

```
(ka > kb ? 3 : 4)
```

has value 3 if `ka` is greater than `kb`, and 4 otherwise. Comparisons may be made using

`>` (greater than)
`<` (less than)
`>=` (greater than or equal to)
`<=` (less than or equal to)
`==` (equal to)
`!=` (not equal to).

Expressions, as well as variables, may be compared in this way, but audio rate variables and expressions are not permitted.

Automatic score generation. There are a number of methods of avoiding the tedious process of writing a score file for CSound. One method is to use the *score translator* program `scot`. This takes a text file `<filename>.sc` written in a compressed score notation and writes out a score file `<filename>.sco`. Another is to use `Cscore`, which is a program for making and manipulating score files. The user writes a control program in the C language, which makes use of a set of function definitions contained in a header file `cscore.h`. Finally, there is MIDI2CS, a program which takes a MIDI file as input, and outputs a score file. There is also a considerable amount of support for MIDI within the CSound language.

DirectCSound is a realtime version of CSound for the PC, and can be obtained from Gabriel Maldonado's home page at

web.tiscalinet.it/G-Maldonado/home2.htm

I have not tried it out, so I cannot comment on how well it works, but it looks promising.

Further reading on CSound:

Richard Boulanger, *The CSound book* [**14**].

Electronic Musician, Feb 1998 issue.

Keyboard, Jan 1997 issue.

8.14. Other methods of synthesis

Sampling is not really a form of synthesis at all, but is often used in digital synthesizers. It is usual to sample sounds at only a small collection of pitches, and then to pitch shift by stretching or compressing the waveform, in order to fill in the gaps. Pitch shifting a digital signal introduces high frequency noise, related to the fact that the sample rate is not being shifted at the same time. This is removed using a low pass filter.

Wavetable synthesis is a method related to sampling, in which digitally recorded wave files are used as raw material to produce sounds which are a sort of hybrid between synthesis and sampling. It is usual to use one wave file for the attack portion of the sound, and another for the sustain portion. In the case of the sustain portion, a whole number of periods of the sound are used to form a loop which is repeated. An envelope is then applied to shape the sound, and then finally the result is pitch shifted and put through a low pass filter. An exception to this general procedure is "one shot" sounds such as short percussive sounds. These are usually just recorded as a single wavefile without looping.

Granular synthesis is a method where the sound comes in small packets called *grains*, whose duration is usually of the order of ten milliseconds. Thousands of these grains are used in each second, to create a sound texture. Usually, some algorithm is used for describing large quantities of grains at a time, so that each grain does not have to be described separately.

Further reading on granular synthesis:

S. Cavaliere and A. Piccialli, *Granular synthesis of musical signals*, appears as article 5 in Roads et al [**115**], pages 155–186.

John Duesenberry, *Square one: a world in a grain of sound*, Electronic Musician, November 1999.

Curtis Roads, *Granular synthesis of sound*, Computer Music Journal 2 (2) (1978), 61–62. A revised and updated version of this article appears as chapter 10 of Roads and Strawn [**116**], pages 145–159.

Curtis Roads, *Granular synthesis*, Keyboard, June 1997.

Curtis Roads, *Microsound* [**114**].

8.15. The phase vocoder

The phase vocoder is a method of sound analysis and manipulation. It is based on the technique of applying a discrete Fourier transform to small windows of the original sound. The transform may then be manipulated, and finally the sound may reconstructed from the manipulated transform. For example, it is not hard to stretch a sound without altering the pitch using this technique.

Further reading:

Mark Dolson, *The phase vocoder: a tutorial*, Computer Music Journal 10 (4) (1986), 14–27.

Marie-Hélène Serra, *Introducing the phase vocoder*, appears as article 2 in Roads et al [**115**], pages 31–90.

8.16. Chebyshev polynomials

Composition of functions in general is a good way of obtaining synthetic tones. For example, if we take a basic cosine wave $\cos \nu t$ and compose it with the function $f(x) = 2x^2 - 1$ then we obtain

$$2\cos^2 \nu t - 1 = \cos 2\nu t.$$

So composing with this function has the effect of doubling frequency. The corresponding functions for arbitrary integer multiples of frequency are called the *Chebyshev*[7] *polynomials of the first kind*, which we now investigate.

Let $T_n(x)$ be the polynomial defined inductively by $T_0(x) = 1$, $T_1(x) = x$, and for $n > 1$,

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x).$$

Thus for example we have

$$\begin{aligned}
T_0(x) &= 1 \\
T_1(x) &= x \\
T_2(x) &= 2x^2 - 1 \\
T_3(x) &= 4x^3 - 3x \\
T_4(x) &= 8x^4 - 8x^2 + 1 \\
T_5(x) &= 16x^5 - 20x^3 + 5x \\
T_6(x) &= 32x^6 - 48x^4 + 18x^2 - 1 \\
T_7(x) &= 64x^7 - 112x^5 + 56x^3 - 7x.
\end{aligned}$$

LEMMA 8.16.1. *For $n \geq 0$ we have $T_n(\cos \nu t) = \cos n\nu t$.*

PROOF. The proof is by induction on n. We begin by observing that

$$\begin{aligned}
\cos \nu t\, \cos(n-1)\nu t - \sin \nu t\, \sin(n-1)\nu t &= \cos n\nu t \\
\cos \nu t\, \cos(n-1)\nu t + \sin \nu t\, \sin(n-1)\nu t &= \cos(n-2)\nu t
\end{aligned}$$

(see §1.7), so that adding and rearranging, we have

$$\cos n\nu t = 2\cos \nu t\, \cos(n-1)\nu t - \cos(n-2)\nu t.$$

Now for $n = 0$ and $n = 1$, the statement of the lemma is obvious from the definition. For $n \geq 2$, assuming the statement to be true for smaller values of n, we have

$$\begin{aligned}
T_n(\cos \nu t) &= 2\cos \nu t\, T_{n-1}(\cos \nu t) - T_{n-2}(\cos \nu t) \\
&= 2\cos \nu t\, \cos(n-1)\nu t - \cos(n-2)\nu t \\
&= \cos n\nu t.
\end{aligned}$$

So by induction, the lemma is true for all $n \geq 0$. $\square$

[7] Other spellings for this name include Tchebycheff and Chebichev. These are all just transliterations of the Russian Чебышев.

Using a weighted sum of Chebyshev polynomials and composing, we can obtain a waveform with the corresponding weights for the harmonics. Changing the weighting with time will change the timbre of the resulting tone. So for example, if we apply the operation

$$T_1 + \tfrac{1}{3}T_3 + \tfrac{1}{5}T_5 + \tfrac{1}{7}T_7 + \tfrac{1}{9}T_9 + \tfrac{1}{11}T_{11}$$

to a cosine wave, we obtain an approximation to a square wave (see equation (2.2.10)). This operation will turn any mixture of cosine waves into the same mixture of square waves.

Exercises

1. Show that $y = T_n(x)$ satisfies *Chebyshev's differential equation*

$$(1 - x^2)\frac{d^2y}{dx^2} - x\frac{dy}{dx} + n^2y = 0.$$

2. Show that

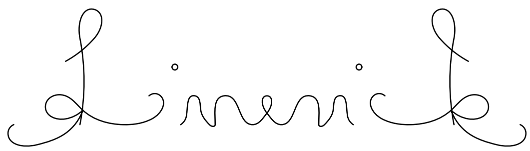
$$T_n(x) = x^n - \binom{n}{2}x^{n-2}(1 - x^2) + \binom{n}{4}x^{n-4}(1 - x^2)^2 - \dots$$

Hint: Use de Moivre's theorem (see Appendix C) and the binomial theorem.

3. Draw a graph of $y = T_n(x)$ for $-1 \le x \le 1$ and $0 \le n \le 5$.

CHAPTER 9

Symmetry in music

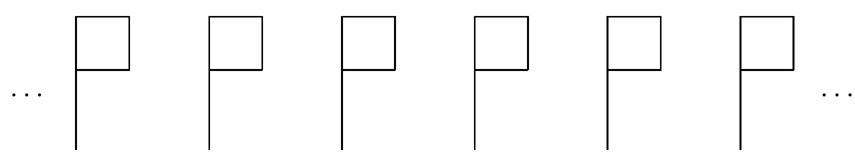


First, let me explain that I'm cursed;
I'm a poet whose time gets reversed.
    Reversed gets time
    Whose poet a I'm;
Cursed I'm that explain me let, first.

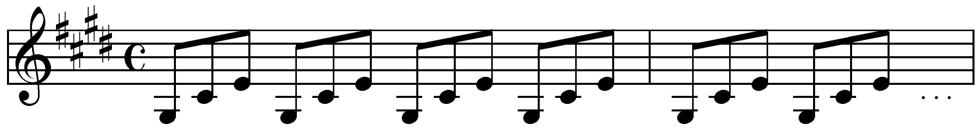
9.1. Symmetries

Music contains many examples of symmetry. In this chapter, we investigate the symmetries that appear in music, and the mathematical language of group theory for describing symmetry.

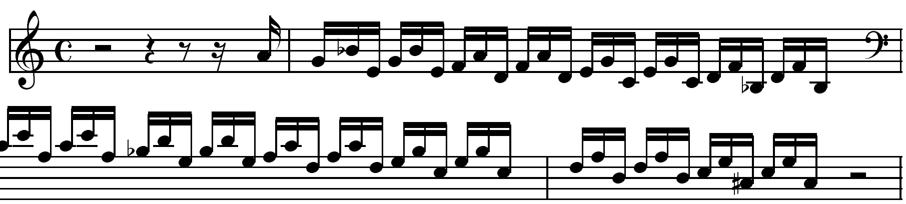
We begin with some examples. *Translational symmetry* looks like this:



In group theoretic language, which we explain in the next few sections, the symmetries form an *infinite cyclic group*. In music, this would just be represented by *repetition* of some rhythm, melody, or other pattern. Here is beginning of the right hand of Beethoven's *Moonlight* Sonata, Op. 27 No. 2.

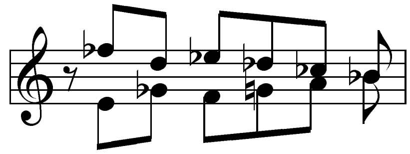


Of course, any actual piece of music only has finite length, so it cannot really have true translational symmetry. Indeed, in music, approximate symmetry is much more common than perfect symmetry. The musical notion of a *sequence* is a good example of this. A sequence consists of a pattern that is repeated with a shift; but the shift is usually not exact. The intervals are not the same, but rather they are modified to fit the harmony. For example, the sequence



comes from J. S. Bach's *Toccata and Fugue in D*, BWV 565, for organ. Although the general motion is downwards, the numbers of semitones between the notes in the triplets is constantly varying in order to give the appropriate harmonic structure.

Reflectional symmetry appears in music in the form of *inversion* of a figure or phrase. For example, the following bar from Béla Bartók's *Fifth string quartet* displays a reflectional symmetry whose horizontal axis is the note B♭.



The lower line is obtained by inverting the upper line. The symmetry group here is cyclic of order two.

Such symmetry can also be more global in character. For example, in Richard Strauss' *Elektra* (1906–1908), although symmetry plays little or no role in the choice of individual notes, its influence is apparent in the choice of keys. The introduction starts with Agamemnon's motive in in D minor. Then Elektra's motive consists of B minor and F minor triads, symmetrically placed around D. Then in Elektra's monologue, Agamemnon is associated with B♭ and Klytemnestra with F♯, again symmetrical around D. The opera continues this way, working either side of the initial D. The ending is in C major, with a prominent major third E in the last four bars. See pages 15–16 of Antokoletz, *The music of Béla Bartók*, for further details.

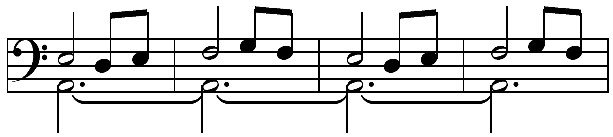
Der Spiegel (The Mirror) Duet

W.A. Mozart

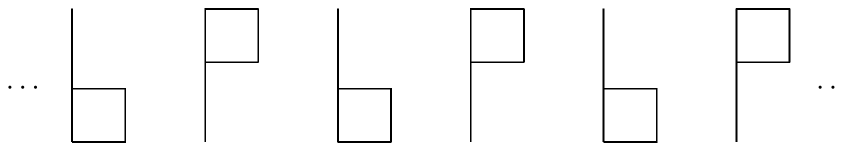
Public Domain. Sequenced by Fred Nachbaur using NoteWorthy
Confused? Try playing this from opposite sides of a table.

(Note: the attribution to Mozart is dubious)

It is more common for horizontal reflection to be combined with a displacement in time. For example, the left hand of Chopin's *Waltz*, Op. 34 No. 2, begins as follows.

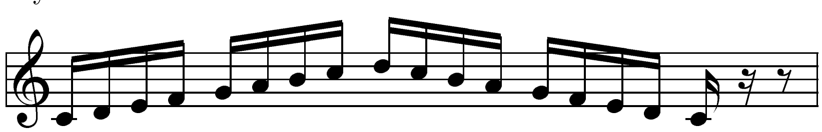


Each bar of the upper line of the left hand is inverted to form the next bar. Because of the displacement in time, this is really a *glide reflection*; namely a translation followed by a reflection about a mirror parallel to the direction of translation. In group theoretic terms, this is another manifestation of the infinite cyclic group.



The reason for the importance of symmetry in music is that regularity of pattern builds up expectations as to what is to come next. But it is important to break the expectations from time to time, to prevent boredom. Good music contains just the right balance of predictability and surprise.

In the above example, the mirror line for the reflectional symmetry was horizontal. It is also possible to have *temporal* reflectional symmetry with a vertical mirror line, so that the notes form a palindrome. For example, an ascending scale followed by a descending scale has this kind of reflectional symmetry, as in the following elementary vocal exercise. The symmetry group here is cyclic of order two.



This is the musical equivalent of the *palindrome*. One example of a musical form involving this kind of symmetry is the *retrograde canon* or *crab canon* (Cancrizans). This term denotes a work in the form of a canon and exhibiting temporal reflectional symmetry by means of playing the melody forwards and backwards at the same time. For example, the first canon of J. S. Bach's *Musical Offering* (BWV 1079) is a retrograde canon formed by playing Frederick the Great's royal theme, consisting of the following 18 bars

Doppelgänger

Entering the lonely house with my wife
   I saw him for the first time
Peering furtively from behind a bush—
      Blackness that moved,
   A shape amid the shadows,
A momentary glimpse of gleaming eyes
   Revealed in the ragged moon.
A closer look (he seemed to turn) might have
Put him to flight forever—
   I dared not
(For reasons that I failed to understand),
   Though I knew I should act at once.

I puzzled over it, hiding alone,
Watching the woman as she neared the gate.
   He came, and I saw him crouching
      Night after night.
      Night after night
  He came, and I saw him crouching,
Watching the woman as she neared the gate.

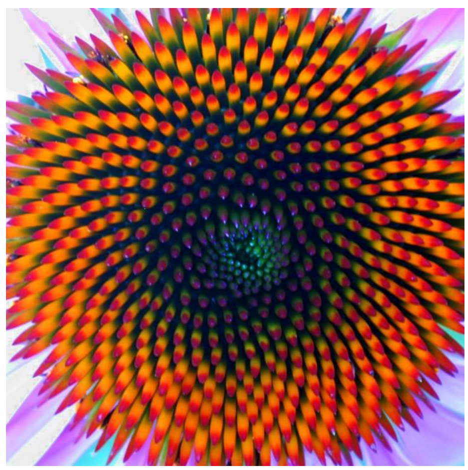
I puzzled over it, hiding alone—
   Though I knew I should act at once,
For reasons that I failed to understand
      I dared not
  Put him to flight forever.

A closer look (he seemed to turn) might have
   Revealed in the ragged moon
A momentary glimpse of gleaming eyes,
   A shape amid the shadows,
      Blackness that moved.

Peering furtively from behind a bush,
   I saw him, for the first time,
Entering the lonely house with my wife.

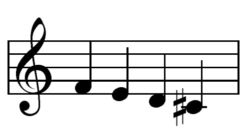
—by J. A. Lyndon,
from *Palindromes and Anagrams*,
H. W. Bergerson, Dover 1973.

simultaneously forwards and backwards in this way. The first voice starts at the beginning of the first bar and works forward to the end, while the second voice starts at the end of the last bar and works backwards to the beginning. Other examples can be found at the end of this section, under "further listening." The other parts of Bach's *Musical Offering* exhibit various other tricky ways of playing with symmetry and form.

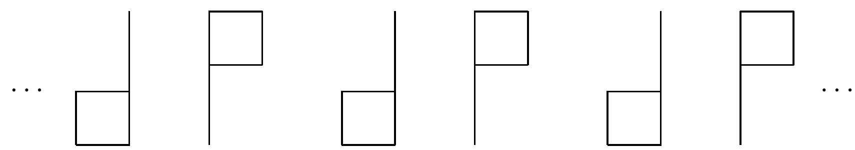


coneflower

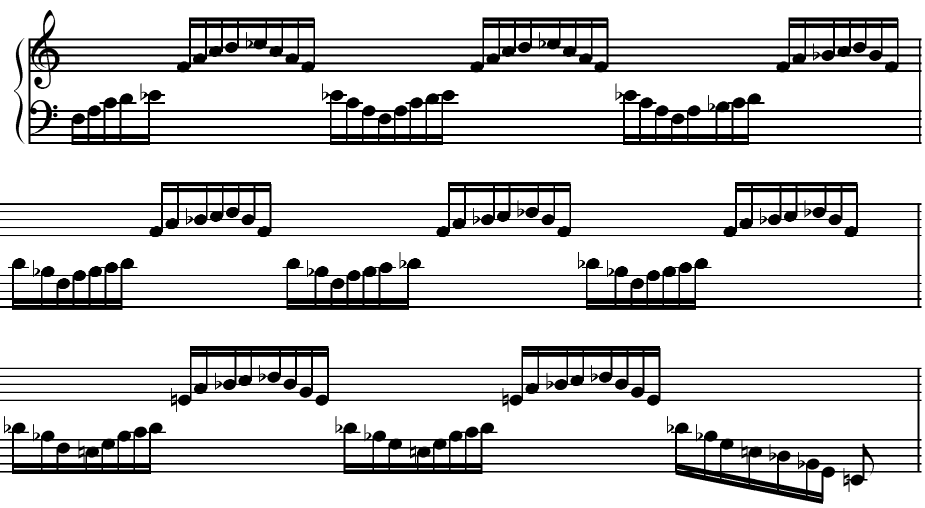
Examples of *rotational* symmetry can also be found in music. For example, the following four note phrase has perfect rotational symmetry, whose center is at the end of the second beat, at the pitch D♯.



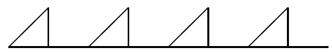
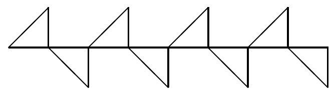
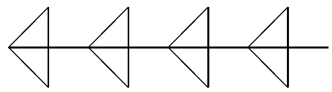
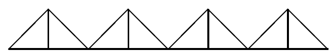
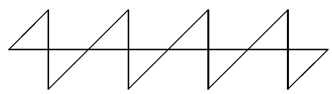
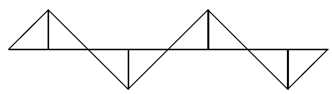
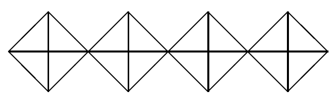
In Ravel's *Rhapsodie Espagnole* (1908), this four note phrase is repeated a large number of times. This really means that we have translations and rotations, as in the following diagram. In group theoretic language, the symmetries form an *infinite dihedral group*.



In the following example, from the middle of Mozart's *Capriccio*, KV 395 for piano, the symmetry is approximate. It is easy to observe that each beamed set of notes for the right hand has a gradual rise followed by a steeper descent, while those for the left hand have a steep descent followed by a more gradual rise. Each pair of beams is slightly different from the previous, so we do not get bored. Our expectations are finally thwarted in the last beam, where the descent continues all the way down to a low E♮.



Horizontally repeated patterns are sometimes known as *frieze patterns*, and they are classified into seven types. The numbering scheme shown below is the international one usually used by mathematicians and crystallographers, for reasons which are not likely to become clear any time soon (see for example pages 39 and 44 of Grünbaum and Shephard). The abstract groups are explained later on in this chapter.

Example	name	abstract group
	p111	$\mathbb{Z}$
	p1a1	$\mathbb{Z}$
	p1m1	$\mathbb{Z} \times \mathbb{Z}/2$
	pm11	D_∞
	p112	D_∞
	pma2	D_∞
	pmm2	$D_\infty \times \mathbb{Z}/2$

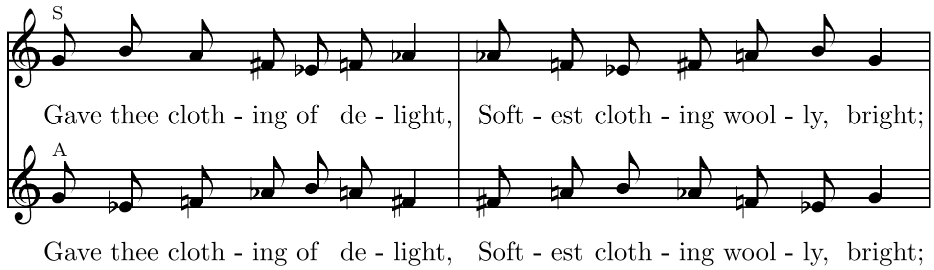
The seven frieze types

For example, the upper line of the left hand of the Chopin Waltz example on page 282 belongs to frieze type p1a1, while the Ravel example on page 284 belongs to frieze type p112.

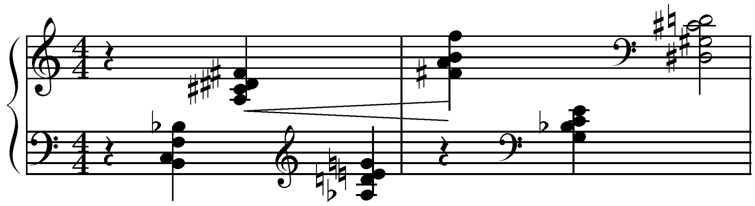
Exercises

1. What symmetry is present in the following extract from Béla Bartók's *Music for strings, percussion and celesta*? Is it exact or approximate?

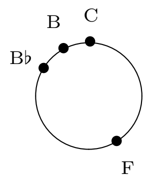
2. Find the symmetries in the following two bars from John Tavener's *The lamb* (words by William Blake). Are the symmetries exact or approximate?



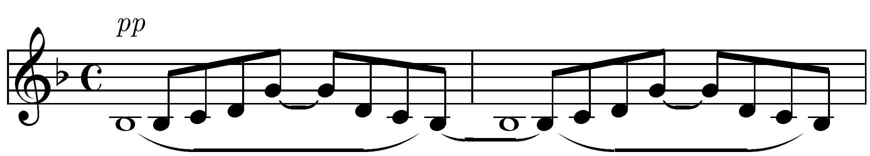
3. The symmetry in the first two bars of Schoenberg's *Klavierstück* Op. 33a is somewhat harder to see.



You may find it helpful to draw the chords on a circle; the first chord will come out as follows.



4. Which frieze pattern appears in the first few bars of Debussy's *Rêverie*, which are as follows?



Further reading:

Elliott Antokoletz, *The music of Béla Bartók*, University of California Press, 1984.

Bruce Archibald, *Some thoughts on symmetry in early Webern*, Perspectives in New Music 10 (1972), 159–163.

F. J. Budden, *The fascination of groups*, CUP, 1972. ISBN 0521080169. Chapter 23 is titled *Groups and music*.

Branko Grünbaum and G. C. Shephard, *Tilings and patterns, an introduction*. W. H. Freeman and Company, New York, 1989.

E. Lendvai, *Symmetries of music* [**74**].

George Perle, *Symmetric formations in the string quartets of Béla Bartók*, Music Review 16 (1955), 300–312.

Further listening: (see Appendix R)

William Byrd, *Diliges Dominum* exhibits temporal reflectional symmetry, making it a perfect palindrome.

In Joseph Haydn's *Sonata* 41 in A, the movement *Menuetto al rovescio* is also a perfect palindrome.

Guillaume de Machaut, *Ma fin est mon commencement* (My end is my beginning) is a retrograde canon in three voices, with a palindromic tenor line. The other two lines are exact temporal reflections of each other.

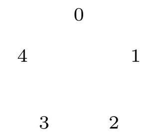
An unfinished keyboard piece employing invertible counterpoint.

From Prof. Peter Schickele, *The definite biography of P.D.Q. Bach* (1807–1742)?, Random House, New York, 1976.

9.2. The harp of the Nzakara

In this section, we take a look at an example taken from the article of Chemillier in [**1**]. The Nzakara and Zande people of the Central African Republic, Congo and Sudan have a musical tradition of the court which is now in a state of neglect. The music consists of poetry sung to the accompaniment of a five string harp. The harpist plays a formulaic repeating pattern of pairs of notes.

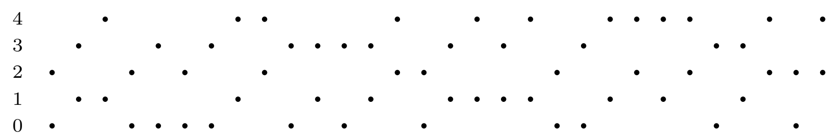
The five strings of the harp are tuned to notes which can be transcribed roughly as C, D, E, G, B♭. These five strings are regarded as having a cyclic order rather than a linear order, so that the lowest string is regarded as adjacent to the highest string.



The strings are plucked in pairs, and the two strings of a pair are *never* adjacent in the cycle. So there are only five possible pairs. The strings in the pair have a unique common neighbor, and we can label the pair using this common neighbor. So the five pairs are as follows.

label	strings	
0	1	4
1	0	2
2	1	3
3	2	4
4	0	3

The repeating harp patterns are divided into categories with names such as *ngbàkià*, *limanza* and *gitangi*. An example of a *limanza* line is given by repeating the following sequence of pairs.



Transcribing this using our labels, we obtain the sequence

$$120141403424231202014030342 2313.$$

At first sight, it is hard to see any pattern. But we divide it into groups of six as follows.

$$12\ 014140\ 342423\ 120201\ 403034\ 2313.$$

Since the pattern is supposed to repeat, the initial pair can be thought of as being at the end of the last group of four to make a group of six,

$$014140\ 342423\ 120201\ 403034\ 231312.$$

Now we can see that each group of six is obtained from the previous group by moving two places down the cycle of five strings. This forms a sort of twisted translational symmetry.

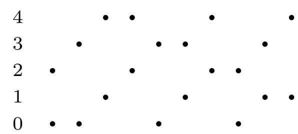
There is also a kind of rotational symmetry (this explains why we chose to move two time slots from the beginning to the end). We can reverse time, and simultaneously reverse the cyclic ordering of the five strings, replacing string x by string $4 - x$. This gives the sequence

$$231312\ 014140\ 342423\ 120201\ 403034,$$

which is the same sequence but with a different starting point.

Exercises

1. Here is a repeating *ngbàkià* harp line taken from the same article of Chemillier.



Find the symmetries in this pattern.

Further reading:

Marc Chemillier, *Mathématiques et musiques de tradition orale*, pages 133–143 of [**44**].

Marc Chemillier, *Ethnomusicology, ethnomathematics. The logic underlying orally transmitted artistic practices*, pages 161–183 of [**1**].

Further listening: (see Appendix R)

Marc Chemillier, *Central African Republic. Music of the former Bandia courts.*

9.3. Sets and groups

The mathematical structure which captures the notion of symmetry is the notion of a *group*. In this section, we give the basic axioms of group theory, and we describe how these axioms capture the notion of symmetry.

A *set* is just a collection of objects. The objects in the set are called the *elements* of the set. We write $x \in X$ to mean that an object x is an element of a set X, and we write $x \notin X$ to mean that x is not an element of X.

Strictly speaking, a set shouldn't be too big. For example, the collection of all sets is too big to be a set, and if we allow it to be a set then we run into Russel's paradox, which goes as follows. If the collection of all sets is regarded as a set, then it is possible for a set to be an element of itself: $X \in X$. Now form the set S consisting of all sets X such that $X \notin X$. If $S \notin S$ then S is one of the sets X satisfying the condition for being in S, and so $S \in S$. On the other hand, if $S \in S$ then S is not one of these sets X, and so $S \notin S$. This contradictary conclusion is Russel's paradox. Fortunately, finite and countably infinite collections are small enough to be sets, and we are mostly interested in such sets.[1] If a set X is finite, we write $|X|$ for the number of elements in X.

[1]For a reasonably modern and sophisticated introduction to set theory, I recommend W. Just and M. Weese, *Discovering modern set theory*, two volumes, published by the American Mathematical Society, 1995. None of the sophistication of modern set theory is necessary for music theory.

A *group* is a set G together with an operation which takes any two elements g and h of G and multiplies them to give again an element of G, written gh. For G to be a group, this multiplication must be defined for all pairs of elements g and h in G, and it must satisfy three axioms:

(i) (Associative law) Given any elements g, h and k in G (not necessarily different from each other), if we multiply gh by k we get the same answer as if we multiply g by hk:

$$(gh)k = g(hk).$$

(ii) (Identity) There is an element $e \in G$ called the *identity element*, which has the following property. For every element g in G, we have $eg = g$ and $ge = g$.

(iii) (Inverses) For each element $g \in G$, there is an *inverse* element written g^{-1}, with the property that $gg^{-1} = e$ and $g^{-1}g = e$.

It is worth noticing that a group does *not* necessarily satisfy the commutative law. An *abelian group* is a group satisfying the following axiom in addition to axioms (i)–(iii):

(iv) (Commutative law)[2] Given any elements g and h in G, we have $gh = hg$.

We can give a group by writing down a *multiplication table*. For example, here is the multiplication table for a group with three elements.

	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

To multiply elements g and h of a group using a multiplication table, we look in row g and column h, and the entry is gh. So for example, looking in the above table, we see that $ab = e$. The above example is an abelian group, because the table is symmetric about its diagonal. The following multiplication table describes a nonabelian group G with six elements.

	e	v	w	x	y	z
e	e	v	w	x	y	z
v	v	w	e	y	z	x
w	w	e	v	z	x	y
x	x	y	z	e	v	w
y	y	z	x	w	e	v
z	z	x	y	v	w	e

In this group, we have $xy = v$ but $yx = w$, which shows that the group is not abelian. We write $|G| = 6$ to indicate that the group G has six elements.

[2]In real life, as in group theory, operations seldom satisfy the commutative law. For example, if we put on our socks and then put on our shoes, we get a very different effect from doing it the other way round. The associative law is much more commonly satisfied.

Groups don't have to be finite of course. For example, the set $\mathbb{Z}$ of integers with operation of *addition* forms an abelian group. Usually, a group operation is only written additively if the group is abelian. The identity element for the operation of addition is 0, and the additive inverse of an integer n is $-n$.

It should by now be apparent that multiplication tables aren't a very good way of describing a group. Suppose we want to check that the above multiplication table satisfies the axioms (i)–(iii). We would have to make $6 \times 6 \times 6 = 216$ checks just for the associative law. Now try to imagine making the checks for a group with thousands of elements, or even millions.

Fortunately, there is a better way, based on permutation groups. A *permutation* of a set X is a function f from X to X such that each element x of X can be written as $f(y)$ for a unique $y \in X$. See page 297 for more discussion of this definition. This ensures that f has an *inverse function*, f^{-1} which takes x back to y. So we have $f^{-1}(f(y)) = f^{-1}(x) = y$, and $f(f^{-1}(x)) = f(y) = x$.

For example, if $X = \{1, 2, 3, 4, 5\}$, the function f defined by

$$f(1) = 3, \quad f(2) = 5, \quad f(3) = 4, \quad f(4) = 1, \quad f(5) = 2$$

is a permutation of X. There are two common notations for writing permutations on finite sets, both of which are useful. The first notation lists the elements of X and where they go. In this notation, the above permutation f would be written as follows.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix}$$

The other notation is called *cycle notation*. For the above example, we notice that 1 goes to 3 goes to 4 goes back to 1 again, and 2 goes to 5 goes back to 2. So we write the permutation as

$$f = (1, 3, 4)(2, 5).$$

This notation is based on the fact that if we apply a permutation repeatedly to an element of a finite set, it will eventually cycle back round to where it started. The entire set can be split up into disjoint cycles in this way, so that each element appears in one and only one cycle. If a permutation is written in cycle notation, to see its effect on an element, we locate the cycle containing the element. If the element is not at the end of the cycle, the permutation takes it to the next one in the cycle. If it is at the end, it takes it back to the beginning. The *length* of a cycle is the number of elements appearing in it. If a cycle has length one, then the element appearing in it is a *fixed point* of the permutation. Fixed points are often omitted when writing a permutation in cycle notation.

To multiply permutations, we compose functions. In the above example, suppose we have another permutation g of the same set X, given by

$$g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 1 & 4 & 3 \end{pmatrix}$$

or in cycle notation,

$$g = (1, 2, 5, 3)(4).$$

If we omit the fixed point 4 from the notation, this element is written $g = (1, 2, 5, 3)$. Then $f(g(1)) = f(2) = 5$. Continuing this way, fg is the following permutation,

$$fg = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 2 & 3 & 1 & 4 \end{pmatrix} = (1, 5, 4)$$

whereas gf is given by

$$gf = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 4 & 2 & 5 \end{pmatrix} = (2, 3, 4).$$

The *identity permutation* takes each element of X to itself. In the above example, the identity permutation is

$$e = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = (1)(2)(3)(4)(5).$$

Omitting fixed points from the identity permutation leaves us with a rather embarrassing empty space, which we fill with the sign e denoting the identity element. The *order* of a permutation is the number of times it has to be applied, to get back to the identity permutation. In the above example, f has order six, g has order four, and both fg and gf have order three. The order of an element g of any group is defined in the same way, as the least positive value of n such that $g^n = 1$. If there is no such n, then g is said to have *infinite order*. For example, the translation which began the chapter is a transformation of infinite order, whereas a reflection is a transformation of order two.

Notice how the commutative law is not at all built into the world of permutations, but the associative law certainly is. The inverse of a permutation is a permutation, and the composite of two permutations is also a permutation. So it is easy to check whether a collection of permutations forms a group. We just have to check that the identity is in the collection, and that the inverses and composites of permutations in the collection are still in the collection.

The set of *all* permutations of a set X forms a group which is called the *symmetric group* on the set X, with the multiplication given by composing permutations as above. We write the symmetric group on X as $\mathsf{Symm}(X)$. If $X = \{1, 2, \ldots, n\}$ is the set of integers from 1 to n, then we write S_n for $\mathsf{Symm}(X)$. Notice that the sets X and $\mathsf{Symm}(X)$ are quite different in size. If $X = \{1, 2, \ldots, n\}$ then X has n elements, but $\mathsf{Symm}(X)$ has $n!$ elements. To see this, if $f \in \mathsf{Symm}(X)$ then there are n possibilities for $f(1)$. Having chosen the value of $f(1)$, there are $n-1$ possibilities left for $f(2)$. Continuing this way, the total number of possibilities for f is $n(n-1)(n-2)\ldots 1 = n!$.

The definition of a *permutation group* is that it is a subgroup of $\mathsf{Symm}(X)$ for some set X. In general, a *subgroup* H of a group G is a subset of G which is a group in its own right, with multiplication inherited from G.

This is the same as saying that the identity element belongs to H, inverses of elements of H are also in H, and products of elements of H are in H. So to check that a set H of permutations of X is a group, we check these three properties so that H is a subgroup of $\mathsf{Symm}(X)$. Notice that the associative law is automatic for permutations, and does not need to be checked.

Exercises

1. If g and h are elements of a group, explain why gh and hg always have the same order.

Further reading:

Hans J. Zassenhaus, *The theory of groups.* Dover reprint, 1999. 276 pages, in print. ISBN 0486409228. This is a solid introduction to group theory, originally published in 1949 by Chelsea.

9.4. Change ringing

> The art of change ringing is peculiar to the English, and, like most English peculiarities, unintelligible to the rest of the world. To the musical Belgian, for example, it appears that the proper thing to do with a carefully tuned ring of bells is to play a tune upon it. By the English campanologist, the playing of tunes is considered to be a childish game, only fit for foreigners; the proper use of the bells is to work out mathematical permutations and combinations. When he speaks of the music of his bells, he does not mean musicians' music—still less what the ordinary man calls music. To the ordinary man, in fact, the pealing of bells is a monotonous jangle and a nuisance, tolerable only when mitigated by remote distance and sentimental association. The change-ringer does, indeed, distinguish musical differences between one method of producing his permutations and another; he avers, for instance, that where the hinder bells run 7, 5, 6, or 5, 6, 7, or 5, 7, 6, the music is always prettier, and can detect and approve, where they occur, the consecutive fifths of Tittums and the cascading thirds of the Queen's change. But what he really means is, that by the English method of ringing with rope and wheel, each several bell gives forth her fullest and her noblest note. His passion—and it is a passion—finds its satisfaction in mathematical completeness and mechanical perfection, and as his bell weaves her way rhythmically up from lead to hinder place and down again, he is filled with the solomn intoxication that comes of intricate ritual faultlessly performed.

Dorothy L. Sayers, *The Nine Tailors*, 1934

The symmetric group, described at the end of the last section, is essential to the understanding of *change ringing*, or *campanology*. This art began in England in the tenth century, and continues in thousands of English churches to this day. A set of swinging bells in the church tower is operated by pulling ropes. There are generally somewhere between six and twelve bells. The problem is that the bells are heavy, and so the timing of the peals of the bells is not easy to change. So for example, if there were eight bells, played in sequence as

$$1, 2, 3, 4, 5, 6, 7, 8,$$

then in the next round we might be able to change the positions of some adjacent bells in the sequence to produce

$$1, 3, 2, 4, 5, 7, 6, 8,$$

but we would not be able to move a bell more than one position in the sequence. So the general rules for change ringing state that a change ringing composition consists of a sequence of rows. Each row is a permutation of the set of bells, and the position of a bell in the row can differ by at most one from its previous position. It is also stipulated that a row is not repeated in a composition, except that the last row returns to the beginning. So for example Plain Bob on four bells goes as follows.

1	2	3	4
2	1	4	3
2	4	1	3
4	2	3	1
4	3	2	1
3	4	1	2
3	1	4	2
1	3	2	4
1	3	4	2
3	1	2	4
3	2	1	4
2	3	4	1
2	4	3	1
4	2	1	3
4	1	2	3
1	4	3	2
1	4	2	3
4	1	3	2
4	3	1	2
3	4	2	1
3	2	4	1
2	3	1	4
2	1	3	4
1	2	4	3
1	2	3	4

Plain Bob

This sequence of rows is really a walk around the symmetric group S_4. So the image of the first row under each of the $4! = 24$ elements of S_4 appears exactly once in the list, except that the first is repeated as the last.

In order to fix the notation, we think of a row as a function from the bells to the time slots. To go from one row to the next, we compose with a permutation of the set of time slots. The permutation is only allowed to fix a time slot, or to swap it with an adjacent time slot. So in the above example, the first few steps involve alternately applying the permutations $(1,2)(3,4)$ and $(1)(2,3)(4)$. Then when we reach the row 1 3 2 4, this prescription would take us back to the beginning. In order to avoid this, the permutation $(1)(2)(3,4)$ is applied, and then we may continue as before. At the line 1 4 3 2 we again have the problem that we would be taken to a previously used row, and we avert this by the same method. When we have exhausted all the permutations in S_4, we return to the beginning.

Exercises

1. The *Plain Hunt* consists of alternately applying the permutations

$$a = (1,2)(3,4)(5,6)\dots$$
$$b = (1)(2,3)(4,5)\dots$$

If the number of bells is n, how many rows are there before the return to the initial order?
[Hint: treat separately the cases n even and n odd.]

Further reading:

F. J. Budden, *The fascination of groups*, CUP, 1972. ISBN 0521080169. Chapter 24 is titled *Ringing the changes: groups and campanology.*

T. J. Fletcher, *Campanological groups*, Amer. Math. Monthly 63 (9) (1956), 619–626.

B. D. Price, *Mathematical groups in campanology*, Math. Gaz. 53 (1969), 129–133.

Ian Stewart, *Another fine math you've got me into...*, W. H. Freeman & Co., 1992. Chapter 13 of this book, *The group-theorist of Notre Dame*, is about change ringing.

Arthur T. White, *Ringing the changes*, Math. Proc. Camb. Phil. Soc. 94 (1983), 203–215.

Arthur T. White, *Ringing the changes II*, Ars Combinatorica 20–A (1985), 65–75.

Arthur T. White, *Ringing the cosets*, Amer. Math. Monthly 94 (8) (1987), 721–746.

Wilfred G. Wilson, *Change Ringing*, October House Inc., New York, 1965.

9.5. Cayley's theorem

Cayley's theorem explains why the axioms of group theory exactly capture the physical notion of symmetry. It says that any abstract group, in other words, any set with a multiplication satisfying the axioms described in Section 9.3, can be realised as a group of permutations of some set.

There is something mildly puzzling about this theorem. Where are we going to produce a set from? We're just given a group, and nothing else. So we do the obvious thing, and use the set of elements of the group itself as the set on which it will act as permutations. So before reading this, make very sure you have separated in your mind the set of elements of a permutation group and the set on which it acts by permutations. Because otherwise what follows will be very confusing.

Let G be a group. Then to each element $g \in G$, we assign the permutation in $\mathsf{Symm}(G)$ which sends an element $h \in G$ to $gh \in G$. We want to say that this displays a copy of the group G as a permutation group inside $\mathsf{Symm}(G)$. The best way to say this is to introduce the notion of a *homomorphism* of groups.

Recall that a *function* f from one set X to another set Y, written $f\colon X \to Y$, simply assigns an element $f(x)$ of Y to each element x of X in a well defined manner. Many elements of X are allowed to go to the same place in Y, and not every element of Y needs to be assigned. The *image* of f is the subset of Y consisting of the elements of the form $f(x)$. The function f is *injective* if no two elements of X go to the same place in Y. The function f is *surjective* if every element of Y is in the image of f. A function f which is both injective and surjective is said to be *bijective*. A bijective function is also called a one-one correspondence. A bijective function is the same thing as a function which has an *inverse*, namely a function $f' \colon Y \to X$ with the property that $f'(f(y)) = y$ for all $y \in Y$, and $f(f'(x)) = x$ for all $x \in X$. Namely, f' takes y to the unique x such that $y = f(x)$. In this language, a permutation of a set X is just a bijective function from X to itself.

If G and H are groups, then a *homomorphism* $f \colon G \to H$ is a function from the set G to the set H which "preserves the multiplication" in the sense that it sends the identity element of G to the identity element of H, and for elements g_1 and g_2 in G we have

$$f(g_1g_2) = f(g_1)f(g_2).$$

The image of a homomorphism f has the property that it is a subgroup of H. An injective homomorphism is called a *monomorphism*. A surjective homomorphism is called an *epimorphism*. A bijective homomorphism is called an *isomorphism*. If there is an isomorphism from G to H, we say that G and H are *isomorphic*. This means that they are "really" the same group, except that the elements happen to have different names. If f is a monomorphism, it can be regarded as identifying G with a subgroup of H. In other words, it induces an isomorphism between G and its image, which is a subgroup of H.

Example 9.5.1. Consider the group G of rotational symmetries of a cube. In other words, an element of G consists of a way of rotating a cube so that the faces are aligned in the same direction as they started. There are 24 elements of G, because we can put any one of six faces downwards, and four different ways round. Once we have decided which face to put downwards, and which way round to put it, the rotational symmetry is completely described. To multiply elements g and h of G to get gh is to do the rotational symmetry h followed by the rotational symmetry g, so that

$$gh\,(x) = g(h(x)).$$

The confusing order in which things happen is because we write our functions on the left of their arguments, so that $g(h(x))$ means first do h, then do g.

There is an isomorphism between this group G of symmetries of the cube and the group $\mathsf{Symm}\{a, b, c, d\}$ of permutations on a set of four objects. This may be visualized by labeling the four main diagonals of the cube with the symbols a, b, c, d and seeing the effect of a rotation on this labeling.

In the language of homomorphisms, we can describe Cayley's theorem as follows.

Theorem 9.5.2 (Cayley). *If G is a group, let f be the function from G to $\mathsf{Symm}(G)$ which is defined by $f(g)(h) = gh$. Then f is a monomorphism, and so G is isomorphic with a subgroup of $\mathsf{Symm}(G)$.*

Proof. First, we check that f does indeed take an element $g \in G$ to a permutation. In other words, we must check that $f(g)$ is a bijection. This is easy to check, because $f(g^{-1})$ is its inverse. Namely, for $h \in G$ we have

$$f(g^{-1})(f(g)(h)) = f(g^{-1})(gh) = g^{-1}(gh) = (g^{-1}g)h = h$$

and similarly $f(g)(f(g^{-1})(h)) = h$.

Clearly f takes the identity element of G to the identity permutation. The fact that f is a homomorphism is really a statement of the associative law in G. Namely,

$$\begin{aligned} f(g_1g_2)(h) = (g_1g_2)h = g_1(g_2h) &= f(g_1)(g_2h) \\ &= f(g_1)(f(g_2)(h)) = (f(g_1)f(g_2))(h). \end{aligned}$$

Finally, to prove that f is injective, if $f(g_1) = f(g_2)$ then for all $h \in G$, $f(g_1)(h) = f(g_2)(h)$. Taking for h the identity element of G, we see that $g_1 = g_2$. $\square$

9.6. Clock arithmetic and octave equivalence

Clock arithmetic is where we count up to twelve, and then start back again at one. So for example, to add $6 + 8$ in clock arithmetic, we count six up from 8 to get 9, 10, 11, 12, 1, 2, and so in this system we have $6 + 8 = 2$. It's probably better to write 0 instead of 12, so that we go from 11 back to 0 instead of 12 to 1. So here is the addition table for this clock arithmetic.

+	0	1	2	3	4	5	6	7	8	9	10	11
0	0	1	2	3	4	5	6	7	8	9	10	11
1	1	2	3	4	5	6	7	8	9	10	11	0
2	2	3	4	5	6	7	8	9	10	11	0	1
3	3	4	5	6	7	8	9	10	11	0	1	2
4	4	5	6	7	8	9	10	11	0	1	2	3
5	5	6	7	8	9	10	11	0	1	2	3	4
6	6	7	8	9	10	11	0	1	2	3	4	5
7	7	8	9	10	11	0	1	2	3	4	5	6
8	8	9	10	11	0	1	2	3	4	5	6	7
9	9	10	11	0	1	2	3	4	5	6	7	8
10	10	11	0	1	2	3	4	5	6	7	8	9
11	11	0	1	2	3	4	5	6	7	8	9	10

To emphasize that an addition is being done in clock arithmetic rather than ordinary arithmetic, it is often written using the *congruence* symbol "$\equiv$" rather than the equals sign, as in

$$6 + 8 \equiv 2 \pmod{12}.$$

More generally, $a \equiv b \pmod{n}$ means that $a - b$ is a multiple of n.

In terms of group theory, the above addition table makes the set $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$ into a group. The operation is written as addition; of course, clock arithmetic is abelian. The identity element is 0, and the inverse of i is either $-i$ or $12 - i$, depending which is in the range from 0 to 11. This group is written as $\mathbb{Z}/12$.

There is an obvious homomorphism from the group $\mathbb{Z}$ to $\mathbb{Z}/12$. It takes an integer to the unique integer in the range from 0 to 11 which differs from it by a multiple of 12.

In musical terms, we could think of the numbers from 0 to 11 as representing musical intervals in multiples of semitones, in the twelve tone equal tempered octave. So for example 1 is represented by the permutation which increases each note by one semitone, namely the permutation

$$\begin{pmatrix} \mathrm{C} & \mathrm{C}\sharp & \mathrm{D} & \mathrm{E}\flat & \mathrm{E} & \mathrm{F} & \mathrm{F}\sharp & \mathrm{G} & \mathrm{G}\sharp & \mathrm{A} & \mathrm{B}\flat & \mathrm{B} \\ \mathrm{C}\sharp & \mathrm{D} & \mathrm{E}\flat & \mathrm{E} & \mathrm{F} & \mathrm{F}\sharp & \mathrm{G} & \mathrm{G}\sharp & \mathrm{A} & \mathrm{B}\flat & \mathrm{B} & \mathrm{C} \end{pmatrix}$$

The circulating nature of clock arithmetic then becomes *octave equivalence* in the musical scale, where two notes belong to the same *pitch class* if they differ by a whole number of octaves. Each element of $\mathbb{Z}/12$ is then represented by a different permutation of the twelve pitch classes, with the number i representing an increase of i semitones. So for example the number 7 represents the permuation which makes each note higher by a fifth. Then addition has an obvious interpretation as addition of musical intervals.

This permutation representation looks like Cayley's theorem. But making this precise involves choosing a starting point somewhere in the octave. We choose to start by representing C as 0, so that the correspondence becomes

C	C♯	D	E♭	E	F	F♯	G	G♯	A	B♭	B
0	1	2	3	4	5	6	7	8	9	10	11

Under this correspondence, each element of $\mathbb{Z}/12$ is being represented by the permutation of the twelve notes of the octave given by Cayley's theorem.

Of course, there is nothing special about the number 12 in clock arithmetic. If n is any positive integer, we may form the group $\mathbb{Z}/n$ whose elements are the integers in the range from 0 to $n-1$. Addition is described by adding as integers, and then subtracting n if necessary to put the answer back in the right range. So for example, if we are interested in 31 tone equal temperament, which gives such a good approximation to quarter comma meantone (see Section 6.5), then we would use the group $\mathbb{Z}/31$.

Further reading:

Gerald J. Balzano, *The group-theoretic description of* 12*-fold and microtonal pitch systems*, Computer Music Journal 4 (4) (1980), 66–84.

Paul Isihara and M. Knapp, *Basic* $\mathbb{Z}_{12}$ *analysis of musical chords. With loose erratum*, UMAP J. 14 (1993), 319–348.

Paul F. Zweifel, *Generalized diatonic and pentatonic scales: a group-theoretic approach.* Perspectives of New Music 34 (1) (1996), 140–161.

9.7. Generators

If G is a group, a subset S of the set of elements of G is said to *generate* G if every element of G can be written as a product of elements of S

and their inverses.[3] We say that G is *cyclic* if it can be generated by a single element g. In this case, the elements of the group can all be written in the form g^n with $n \in \mathbb{Z}$. The case $n = 0$ corresponds to the identity element, while negative values of n are interpreted to give powers of the inverse of g.

There are two kinds of cyclic groups. If there is no nonzero value of n for which g^n is the identity element, then the elements g^n multiply the same way that the integers n add. In this case, the group is isomorphic to the additive group $\mathbb{Z}$ of integers. If there is a nonzero value of n for which g^n is the identity element, then by inverting if necessary, we can assume that n is positive. Then letting n be the smallest positive number with this property, it is easy to see that G is isomorphic to the group $\mathbb{Z}/n$ described in the last section.

How many generators does $\mathbb{Z}/n$ have? We can find out whether an integer i generates $\mathbb{Z}/n$ with the help of some elementary number theory.

LEMMA 9.7.1. *Let d be the greatest common divisor of n and i. Then there are integers r and s such that $d = rn + si$.*

PROOF. This follows from Euclid's algorithm for finding the greatest common divisor of two integers.

Let's just recall how Euclid's algorithm goes, and then we'll see how it enables you to write the greatest common divisor in this form. If we're given two integers, let's assume that they're positive (otherwise, just negate them) and that the second is bigger than the first (otherwise, swap them round). If the first is an exact divisor of the second, then it is the greatest common divisor. If it isn't, subtract as many of the first as you can from the second without going negative, and then swap them round. Now repeat.

For example, suppose we're given the integers 24 and 34. Since 24 is smaller than 34, we subtract 24 from 34 and swap them round, so our new numbers are 10 and 24. We can now subtract two 10's from 24 and swap them round to get 4 and 10. We subtract two 4's from 10 and swap to get 2 and 4. Now 2 is an exact divisor of 4, so 2 is the greatest common divisor.

If we keep track of the operations, it enables us to write 2 as $r \times 24 + s \times 34$:

$$\begin{aligned} 10 &= -24 + 34 \\ 4 &= 24 - 2 \times 10 = 24 - 2 \times (34 - 24) = 3 \times 24 - 2 \times 34 \\ 2 &= 10 - 2 \times 4 = (-24 + 34) - 2(3 \times 24 - 2 \times 34) = -7 \times 24 + 5 \times 34. \end{aligned}$$

So we have $r = -6$ and $s = 5$. $\square$

If i has no common factor with n, then $d = 1$, and the above equation says that s times i, considered as the sth power of i in the additive group $\mathbb{Z}/n$, is equal to 1. Since the element 1 is a generator of $\mathbb{Z}/n$, it follows that i is also a generator.

[3]To clarify, an empty product is considered to be the identity element. So if S is empty and G is the group with one element, then S does generate G.

On the other hand, if n and i have a common factor $d > 1$, then all powers of i in $\mathbb{Z}/n$ (i.e., all multiples of i when thinking additively) give numbers divisible by d, so the number 1 is not a power of i. So we have the following.

THEOREM 9.7.2. *The generators for $\mathbb{Z}/n$ are precisely the numbers i in the range $0 < i < n$ with the property that n and i have no common factor.* □

The number of possibilities for i in the above theorem is written $\phi(n)$, and called the *Euler phi function* of n.

For example, if $n = 12$, then the possibilities for i are 1, 5, 7 and 11, and so $\phi(12) = 4$. In terms of musical intervals, the fact that 7 is a generator for $\mathbb{Z}/12$ corresponds to the fact that all notes can be obtained from a given notes by repeatedly going up by a fifth. This is the *circle of fifths*. So it can be seen that apart from the circle of semitones upwards and downwards, the only other ways of generating all the musical intervals is via the circle of fifths, again upwards or downwards. This, together with the consonant nature of the fifth, goes some way toward explaining the importance of the circle of fifths in music.

It is interesting to see that if $n = p$ happens to be a prime number, for example $p = 31$, then *every* element of $\mathbb{Z}/p$ apart from zero is a generator. So $\phi(p) = p - 1$.

In fact, there is a recipe for finding $\phi(n)$ in general, which goes as follows. If $n = p^a$ is a power of a prime then $\phi(n) = p^{a-1}(p-1)$. If m and n are relatively prime (i.e., have no common factors greater than one), then $\phi(mn) = \phi(m)\phi(n)$. Any positive integer can be written as a product of prime powers for different primes, so this gives a recipe for calculating $\phi(n)$. For example,

$$\phi(72) = \phi(2^3.3^2) = \phi(2^3)\phi(3^2) = 2^2(2-1)3(3-1) = 24.$$

Here are the values of $\phi(n)$ for small values of n.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\phi(n)$	0	1	2	2	4	2	6	4	6	4	10	4	12	6	8	8

Exercises

1. Write down the generators for $\mathbb{Z}/24$. What is $\phi(24)$?

2. Show that each generator x of $\mathbb{Z}/n$ satisfies $x^2 \equiv 1 \pmod{n}$ if and only if n is a divisor of 24.

3. Find (a) $\phi(49)$, (b) $\phi(60)$, (c) $\phi(142)$, (d) $\phi(10000)$.

9.8. Tone rows

In twelve tone music, one begins with a *twelve tone row*, which consists of a sequence of twelve pitch classes in order, so that each of the twelve possible pitch classes appears just once.

If we want to be able to look at music which is not formally described as twelve tone as well, we should consider sequences of pitch classes of any length, and with possible repetitions.

A *transposition*[4] of a sequence $\mathbf{x}$ of pitch classes by n semitones is the sequence $\mathbf{T}^n(\mathbf{x})$ in which each of the pitch classes in $\mathbf{x}$ has been increased by n semitones. So for example if

$$\mathbf{x} = 3\ 0\ 8$$

then

$$\mathbf{T}^4(\mathbf{x}) = 7\ 4\ 0.$$

As another example, the first two bars of Chopin's *Étude*, Op. 25 No. 10 consist of the pitches

$$6\text{–}5\text{–}6\quad 7\text{–}8\text{–}9\quad 8\text{–}7\text{–}8\quad 9\text{–}10\text{–}11 \mid 10\text{–}9\text{–}10\quad 11\text{–}0\text{–}1\quad 0\text{–}11\text{–}0\quad 1\text{–}2\text{–}3$$

played as triplets, octave doubled, in both hands simultanously. The second half of the first bar is obtained by applying the transformation $\mathbf{T}^2$ to the first half. The transformation $\mathbf{T}^2$ is applied again to obtain the first half of the second bar, and again for the second half. So if $\mathbf{x}$ is the sequence 6 5 6 7 8 9 then these two bars can be written

$$\mathbf{x}\quad \mathbf{T}^2(\mathbf{x}) \mid \mathbf{T}^4(\mathbf{x})\ \mathbf{T}^6(\mathbf{x}).$$

Bars 3 and 4 of this piece go as follows.

$$2\text{–}3\text{–}4\quad 3\text{–}4\text{–}5\quad 4\text{–}5\text{–}6\quad 5\text{–}6\text{–}7 \mid 6\text{–}7\text{–}8\quad 7\text{–}8\text{–}9\quad 7\text{–}8\text{–}9\quad 8\text{–}9\text{–}10$$

Writing $\mathbf{y}$ for the sequence 2 3 4, we see that the last group in bar 2 is $\mathbf{T}^{-1}(\mathbf{y})$, while bars 3 and 4 can be written

$$\mathbf{y}\quad \mathbf{T}(\mathbf{y})\ \mathbf{T}^2(\mathbf{y})\ \mathbf{T}^3(\mathbf{y}) \mid \mathbf{T}^4(\mathbf{y})\ \mathbf{T}^5(\mathbf{y})\ \mathbf{T}^5(\mathbf{y})\ \mathbf{T}^6(\mathbf{y}).$$

Turning to the next operation, *inversion* $\mathbf{I}(\mathbf{x})$ of a sequence $\mathbf{x}$ just replaces each pitch class by its negative (in clock arithmetic). So in the first example above with $\mathbf{x} = 3\ 0\ 8$, we have

$$\mathbf{I}(\mathbf{x}) = 9\ 0\ 4.$$

The sequences $\mathbf{T}^n\mathbf{I}(\mathbf{x})$ are also regarded as inversions of $\mathbf{x}$. So for example

$$\mathbf{T}^6\mathbf{I}(\mathbf{x}) = 3\ 6\ 10$$

is an inversion of the above sequence $\mathbf{x}$.

The *retrograde* $\mathbf{R}(\mathbf{x})$ of $\mathbf{x}$ is just the same sequence in reverse order. So in the above example,

$$\mathbf{R}(\mathbf{x}) = 8\ 0\ 3.$$

The relations among the operations $\mathbf{T}$, $\mathbf{I}$ and $\mathbf{R}$ are

$$\mathbf{T}^{12} = e, \quad \mathbf{T}^n\mathbf{R} = \mathbf{R}\mathbf{T}^n, \quad \mathbf{T}^n\mathbf{I} = \mathbf{I}\mathbf{T}^{-n}, \quad \mathbf{R}\mathbf{I} = \mathbf{I}\mathbf{R}.$$

where e represents the identity operation which does nothing.

[4]Unfortunately, in group theory the word *transposition* is used to refer to a permutation which leaves all but two points fixed, and swaps those two points. These two usages from music and mathematics are not related, and this can be a source of confusion.

There are four forms of a tone row $\mathbf{x}$. The *prime* form is the original form $\mathbf{x}$ of the row, or any of its transpositions $\mathbf{T}^n(\mathbf{x})$. The *inversion* form is any one of the rows $\mathbf{T}^n\mathbf{I}(\mathbf{x})$. The *retrograde* form is any one of the rows $\mathbf{T}^n\mathbf{R}(\mathbf{x})$. Finally, the *retrograde inversion* form of the row is any one of the rows $\mathbf{T}^n\mathbf{RI}(\mathbf{x})$.

In group theoretic terms, the operations $\mathbf{T}^n$ $(0 \le n \le 11)$ form a cyclic group $\mathbb{Z}/12$. The operation $\mathbf{R}$ together with the identity operation form a cyclic group $\mathbb{Z}/2$. The operations $\mathbf{T}$ and $\mathbf{R}$ commute. The group theoretic way of describing a group with two types of operations which commute with each other is a Cartesian product, which we describe in §9.9. The relationship between $\mathbf{T}$ and $\mathbf{I}$ is more complicated, and is discussed in §9.10.

Further reading:

Allen Forte, *The structure of atonal music* [**41**].

George Perle, *Twelve-tone tonality* [**99**].

John Rahn, *Basic atonal theory*, Schirmer books, 1980.

9.9. Cartesian products

If G and H are groups, then the *Cartesian product*, or *direct product* $G \times H$ is the group whose elements are the ordered pairs (g, h) with $g \in G$ and $h \in H$. The multiplication is defined by

$$(g_1, h_1)(g_2, h_2) = (g_1 g_2, h_1 h_2).$$

The identity element is formed from the identity elements of G and H. The inverse of (g, h) is (g^{-1}, h^{-1}). The axioms of a group are easily verified, so that $G \times H$ with this multiplication does form a group.

Suppose that G and H are subgroups of a bigger group K, with the properties that each element of G commutes with each element of H, the only element which G and H have in common is the identity element (written $G \cap H = \{1\}$), and every element of K can be written as a product of an element of G and an element of H (written $K = GH$). Then there is an isomorphism from $G \times H$ to K given by sending (g, h) to gh. In this case, K is said to be an *internal direct product* of G and H.

For example, the group whose elements are the operations $\mathbf{T}^n$ and $\mathbf{T}^n\mathbf{R}$ of §9.8 is an internal direct product of the subgroup consisting of the operations $\mathbf{T}^n$ and the subgroup consisting of the identity and $\mathbf{R}$. So this group is isomorphic to $\mathbb{Z}/12 \times \mathbb{Z}/2$.

As another example, the lattice $\mathbb{Z}^2$ which we used in order to describe just intonation in §6.8 is really a direct product $\mathbb{Z} \times \mathbb{Z}$, where $\mathbb{Z}$ is the group of integers under addition, as usual. This can be viewed as an internal direct product, where the two copies of $\mathbb{Z}$ consist of the elements $(n, 0)$ and the elements $(0, n)$ for $n \in \mathbb{Z}$. Similarly, the lattice $\mathbb{Z}^3$ of §6.9 is $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$. This can be viewed as an internal direct product of three copies of $\mathbb{Z}$ consisting of the elements $(n, 0, 0)$, the elements $(0, n, 0)$ and the elements $(0, 0, n)$ with $n \in \mathbb{Z}$.

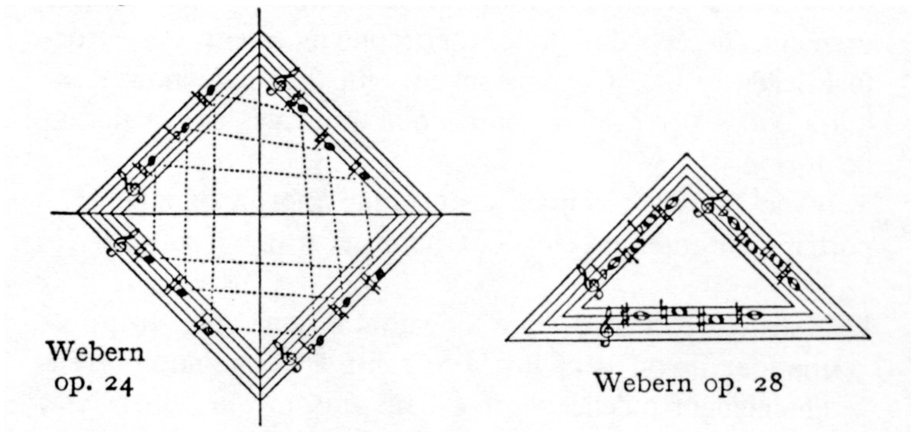
Exercises

1. Find an isomorphism between $\mathbb{Z}/3 \times \mathbb{Z}/4$ and $\mathbb{Z}/12$. Interpret this in terms of transpositions by major and minor thirds.

2. Show that there is no isomorphism between $\mathbb{Z}/12 \times \mathbb{Z}/2$ and $\mathbb{Z}/24$.
[Hint: how many elements of order two are there?]

3. The group $\mathbb{Z}/2\times\mathbb{Z}/2$ is called the *Klein four group*. Go back to Exercise **1** in §9.1 and explain what the Klein four group has to do with this example.

9.10. Dihedral groups



The operations $\mathbf{T}$ and $\mathbf{I}$ of §9.8 do not commute, but rather satisfy the relations $\mathbf{T}^n\mathbf{I} = \mathbf{I}\mathbf{T}^{-n}$. So we do not obtain a direct product in this case, but rather a more complicated construction, which in this case describes a *dihedral group.*

A dihedral group has two elements g and h such that $h^2 = 1$ and $gh = hg^{-1}$. Every element is either of the form g^i or of the form g^ih. The powers of g form a cyclic subgroup which is either $\mathbb{Z}/n$ or $\mathbb{Z}$. In the former case, the group has $2n$ elements and is written[5] D_{2n}. In the latter case, the group has infinitely many elements, and is written D_∞ and called the *infinite dihedral group.* This is one of the groups which appeared in Section 9.1.

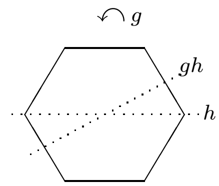
So the operations $\mathbf{T}^n$ and $\mathbf{T}^n\mathbf{I}$ form a group isomorphic to the dihedral group D_{24}. Finally, putting all this together, the group whose operations are

$$\mathbf{T}^n, \quad \mathbf{T}^n\mathbf{R}, \quad \mathbf{T}^n\mathbf{I}, \quad \mathbf{T}^n\mathbf{RI}$$

form a group which is isomorphic to $D_{24} \times \mathbb{Z}/2$.

The dihedral group D_{2n} has an obvious interpretation as the group of rigid symmetries of a regular polygon with n sides.

[5]Some authors write D_n for the dihedral group of order $2n$, just to confuse matters. Presumably these authors think that I'm confusing matters.



The element g corresponds to counterclockwise rotation through $1/n$ of a circle, while h corresponds to reflection about a horizontal axis. Then $g^i h$ corresponds to a reflection about an axis of symmetry which is rotated from the horizontal by i/n of a semicircle. The above diagram is for the case $n = 6$.

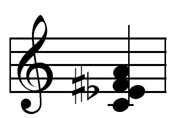
Exercises

1. Find an isomorphism between the dihedral group D_6 and the symmetric group S_3.

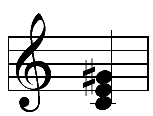
2. Find an isomorphism between D_{12} and $S_3 \times \mathbb{Z}/2$.

3. Show that D_{24} is not isomorphic to $S_3 \times \mathbb{Z}/4$.

4. Consider the group D_{24} generated by $\mathbf{T}$ and $\mathbf{I}$. Which elements fix the following *diminished seventh* chord setwise? What sort of a group do they form?



5. Repeat Exercise **4** with the following "augmented triad."



9.11. Orbits and cosets

If a group G acts on a set X, then we say that two elements x and x' of X are in the same *orbit* if there is an element $g \in G$ such that $g(x) = x'$. This partitions X into disjoint subsets, each consisting of elements related this way. These subsets are the orbits of G on X.

So for example, if G is a cyclic group generated by an element g, then the cycles of g as described in §9.3 are the orbits of G on X.

As another example, the group $\mathbb{Z}/12$ acts on the set of tone rows of a given length, via the operations $\mathbf{T}^n$. Two tone rows are in the same orbit exactly when one is a transposition of the other.

If there is only one orbit for the action of G on X, we say that G acts *transitively* on X. So for example $\mathbb{Z}/12$ acts transitively on the set of twelve pitch classes, but not on the set of tone rows of a given length bigger than one.

We discussed the related concept of cosets briefly in §6.8. Here we make the discussion more precise, and show how this concept is connected with permutations. If H is a subgroup of a group G, we can partition the elements of G into *left cosets* of H as follows. Two elements g and g' are in the same left coset of H in G if there exists some element $h \in H$ such that $gh = g'$. This partitions the group G into disjoint subsets, each consisting of elements related this way. These subsets are the left cosets of H in G. The notation for the left coset containing g is gH. So gH and $g'H$ are equal precisely when there exists an element $h \in H$ such that $gh = g'$; in other words, when $g^{-1}g'$ is an element of H. The coset gH consists of all the elements gh as h runs through the elements of H. The way of writing this is

$$gH = \{gh \mid h \in H\}.$$

The left cosets of H in G all have the same size as H does. So the number of left cosets, written $|G : H|$, is equal to $|G|/|H|$.

The example in §6.8 goes as follows. The group G is $\mathbb{Z}^2 = \mathbb{Z} \times \mathbb{Z}$. The subgroup H is the unison sublattice. Each coset consists of a set of vectors related by translation by the unison sublattice. The group theoretic notion corresponding to a periodicity block is a *set of coset representatives.* A set of left coset representatives for a subgroup H in a group G just consists of a choice of one element from each left coset.

If G acts as permutations on a set X, then there is a close connection between orbits and cosets of subgroups, which can be described in terms of *stabilizers.* If x is an element of X, then the stabilizer in G of x, written $\mathrm{Stab}_G(x)$, is the subgroup of G consisting of the elements h satisfying $h(x) = x$.

THEOREM 9.11.1. *Let $H = \mathrm{Stab}_G(x)$. Then the map sending the coset gH to the element $g(x) \in X$ is well defined, and establishes a bijective correspondence between the left cosets of H in G and the elements of X in the orbit containing x.*

PROOF. To say that the map is well defined is to say that if we are given another element g' such that $gH = g'H$, then $g(x) = g'(x)$. The reason why this is true is that there is an element $h \in H$ such that $gh = g'$, and then $g'(x) = gh(x) = g(h(x)) = g(x)$.

To see that the map is injective, if $g(x) = g'(x)$ then $x = g^{-1}g'(x)$ and so $g^{-1}g' \in H$, and $gH = g'H$. It is obviously surjective, by the definition of an orbit. □

A consequence of this theorem is that the size of an orbit is equal to the index of the stabilizer of one of its elements,

$$|\mathrm{Orbit}(x)| = |G : \mathrm{Stab}_G(x)|. \tag{9.11.1}$$

9.12. Normal subgroups and quotients

In the last section, we discussed left cosets of a subgroup. Of course, right cosets make just as much sense; the reason why left rather than right cosets made their appearance in understanding orbits was that we write functions on the left of their arguments. We write Hg for the right coset containing g, so that

$$Hg = \{hg \mid h \in H\}.$$

It does not always happen that the left and right cosets of H are the same. For example, if G is the symmetric group S_3, and H is the subgroup consisting of the identity and the permutation (12), then the left cosets are

$$\{e,\ (12)\}, \quad \{(123),\ (13)\}, \quad \{(132),\ (23)\}$$

while the right cosets are

$$\{e,\ (12)\}, \quad \{(123),\ (23)\}, \quad \{(132),\ (13)\}.$$

This is because $(123)(12) = (13)$ while $(12)(123) = (23)$.

A subgroup N of G is said to be *normal* if the left cosets and the right cosets agree. For example, if G is abelian, then every subgroup is normal.

THEOREM 9.12.1. *A subgroup N of G is normal if and only if, for each $g \in G$ we have $gNg^{-1} = N$.*

PROOF. To say that the subgroup N is normal means that for each $g \in G$ we have $gN = Ng$. Multiplying on the right by g^{-1}, and noticing that this can be undone by multiplication on the right by g, we see that this is equivalent to the condition that for each $g \in G$ we have $gNg^{-1} = N$. □

If N is normal in G, then the cosets can be made into a group as follows. If gN and $g'N$ are cosets then we multiply them to form the coset $gg'N$. If you check that this is well defined, in other words, that the product does not depend on which elements are used to define the cosets, you will discover that it works precisely when H is normal in G. To check the axioms for a group, we need an identity element, which is provided by the coset $eN = N$ containing the identity element e of G. The inverse of the coset gN is the coset $g^{-1}N$. It is an easy exercise to check the axioms with these definitions.

Clock arithmetic is a good example of a quotient group. Inside the additive group $\mathbb{Z}$ of integers, we have a (normal) subgroup $n\mathbb{Z}$ consisting of the integers divisible by n. The quotient group $\mathbb{Z}/n\mathbb{Z}$ is the clock arithmetic group, which we have been writing in the more usual notation $\mathbb{Z}/n$.

Another example is given by the unison vectors and periodicity blocks of §6.8. The quotient of $\mathbb{Z}^2$ (or more generally $\mathbb{Z}^n$) by the unison sublattice is a finite abelian group whose order is equal to the absolute value of the determinant of the matrix formed from the unison vectors.

There is a standard theorem of abstract algebra which says that every finite abelian group can be written in the form

$$\mathbb{Z}/n_1 \times \mathbb{Z}/n_2 \times \cdots \times \mathbb{Z}/n_r.$$

The positive integers $n_1, \dots, n_r$ are not uniquely determined; for example $\mathbb{Z}/12$ is isomorphic to $\mathbb{Z}/3 \times \mathbb{Z}/4$. However, they can be chosen in such a way that each one is a divisor of the next one. If they are chosen in this way, then they are uniquely determined, and then they are called the *elementary divisors* of the finite abelian group. There is a standard algorithm for finding the elementary divisors, which can be found in many books on abstract algebra. From the point of view of scales, it seems relevant to try to choose the unison sublattice so that the quotient group is cyclic, which corresponds to the case where there is just one elementary divisor.

There is an intimate relationship between normal subgroups and homomorphisms. If f is a homomorphism from G to H, then the *kernel* of f is defined to be the set of elements $g \in G$ for which $f(g)$ is equal to the identity element of H. Writing N for the kernel of f, it is not hard to check that N is a normal subgroup of G.

THEOREM 9.12.2 (First Isomorphism Theorem). *Let f be a homomorphism from G to H. Then there is an isomorphism between the quotient group G/N and the subgroup of H consisting of the image of the homomorphism f. This isomorphism takes a coset gN to $f(g)$.*

PROOF. There are a number of things to check here. We need to check that the function from G/N to the image of f which takes gN to $f(g)$ is well defined, that it is a group homomorphism, that it is injective, and that its image is the same as the image of f. These checks are all straightforward, and are left for the reader to fill in. $\square$

There are actually three isomorphism theorems in elementary group theory, but we shall not mention the second or third.

An example of the first isomorphism theorem is again provided by clock arithmetic. The homomorphism from $\mathbb{Z}$ to $\mathbb{Z}/12$ is surjective and has kernel $12\mathbb{Z}$, and so $\mathbb{Z}/12$ is isomorphic to the quotient of $\mathbb{Z}$ by $12\mathbb{Z}$, as we already knew.

9.13. Burnside's lemma

This section and the next are concerned with problems of counting. A typical example of the kind of problem we are interested in is as follows. Recall that a tone row consists of the twelve possible pitch classes in some order. The total number of tone rows is

$$12 \times 11 \times 10 \times 9 \times \cdots \times 3 \times 2 \times 1 = 12!$$

or 479001600.

We might wish to count the number of possible twelve tone rows, where two tone rows are considered to be the same if one can be obtained from the other by applying an operation of the form $\mathbf{T}^n$. In this case, each tone row has twelve distinct images under these operations. So the total number of tone rows up to this notion of equivalence is $1/12$ of the number of tone rows, or $11! = 39916800$.

If we want to complicate the situation further, we might consider two tone rows to be equivalent if one can be obtained from the other using the operations $\mathbf{T}^n$, $\mathbf{I}$ and $\mathbf{R}$. Now the problem is that some of the tone rows are fixed by some of the elements of the group. So the counting problem degenerates into a lot of special cases, unless we find a more clever way of counting. This is the kind of problem that can be solved using Burnside's counting lemma.

The abstract formulation of the problem is that we have a finite group acting as permutations on a finite set, and we want to know the number of orbits.

Burnside's lemma allows us to count the number of orbits of a finite group G on a finite set X, provided we know the number of fixed points of each element $g \in G$. It says that the number of orbits is the average number of fixed points.

LEMMA 9.13.1 (Burnside). *Let G be a finite group acting by permutations on a finite set X. For an element $g \in G$, write $n(g)$ for the number of fixed points of g on X. Then the number of orbits of G on X is equal to*

$$\frac{1}{|G|} \sum_{g \in G} n(g).$$

PROOF. We count in two different ways the number of pairs (g, x) consisting of an element $g \in G$ and a point $x \in X$ such that $g(x) = x$. If we count the elements of the group first, then for each element of the group we have to count the number of fixed points, and we get $\sum_{g \in G} n(g)$. On the other hand, if we count the elements of X first, then for each x, equation (9.11.1) shows that the number of elements $g \in G$ stabilizing it is equal to $|G|$ divided by the length of the orbit in which x lies. So each orbit contributes $|G|$ to the count. $\square$

So let us return to the problem of counting tone rows. Suppose that we wish to count the number of tone rows, and we wish to regard one tone row as equivalent to another if the first can be manipulated to the second using the operations $\mathbf{T}$, $\mathbf{I}$ and $\mathbf{R}$. In other words, we wish to count the number of orbits of the group $G = D_{24} \times \mathbb{Z}/2$ generated by $\mathbf{T}$, $\mathbf{I}$ and $\mathbf{R}$ on the set X of tone rows.

In order to apply Burnside's lemma, we should find the number of tone rows fixed by each operation in the group. The identity operation fixes all tone rows, so that one is easy. The operations $\mathbf{T}^n$ with $1 \leq n \leq 11$ don't fix any tone rows, so that's also easy. The operation $\mathbf{R}$ fixes the tone rows whose last six entries are the reverse of the first six; but then there are repetitions so these aren't allowed as tone rows. For the operation $\mathbf{T}^6\mathbf{R}$, the fixed tone rows are the ones where the last six entries are the reverse of the first six, but transposed by a tritone (half an octave). So the first six have to be chosen in a way that uses just one of each pair related by a tritone. The number of ways of doing this is

$$12 \times 10 \times 8 \times 6 \times 4 \times 2 = 46080.$$

For values of n other than zero or six, $\mathbf{T}^n\mathbf{R}$ does not fix any tone rows, because doing this operation twice gives $\mathbf{T}^{2n}$, which doesn't fix any tone rows.

Next, we need to consider inversions. The operation **I** fixes only those tone rows comprised of the entries 0 and 6; but then there must be repetitions, so these aren't tone rows. The same goes for any operation of the form $\mathbf{T}^n\mathbf{I}$; the entries come from a subset of size at most two, so we can't form a tone row this way.

Finally, for an operation $\mathbf{T}^n\mathbf{IR}$, the entries in a fixed tone row are again determined by the first six entries. So the tone row has the form

$$a_1,\ a_2,\ a_3,\ a_4,\ a_5,\ a_6,\ n-a_6,\ n-a_5,\ n-a_4,\ n-a_3,\ n-a_2,\ n-a_1.$$

If n is even, there is some tone fixed by $\mathbf{T}^n\mathbf{I}$, which forces us to repeat a tone, so there are no fixed tone rows. If n is odd, however, there are fixed tone rows, and there are

$$12 \times 10 \times 8 \times 6 \times 4 \times 2 = 46080$$

of them.

We summarize this information in the following table.

operation	how many in G	fixed points
identity	1	479001600
$\mathbf{T}^n$ $(1 \le n \le 11)$	11	0
$\mathbf{T}^6\mathbf{R}$	1	46080
$\mathbf{T}^n\mathbf{R}$ $(n \neq 6)$	11	0
$\mathbf{T}^n\mathbf{I}$	12	0
$\mathbf{T}^n\mathbf{IR}$ (n even)	6	0
$\mathbf{T}^n\mathbf{IR}$ (n odd)	6	46080

So the sum over $g \in G$ of the number of fixed points of g on X is

$$479001600 + 7 \times 46080 = 479324160.$$

Dividing by $|G| = 48$, the total number of orbits of G on tone rows is equal to 9985920. This proves the following theorem.

THEOREM 9.13.2 (David Reiner). *If two twelve tone rows are considered the same when one may be obtained from the other using the operations* **T**, **I** *and* **R**, *then the total number of tone rows is* 9985920. □

Further reading:

James A. Fill and Alan J. Izenman, *Invariance properties of Schoenberg's tone row system*, J. Austral. Math. Soc. Ser. B 21 (1979/80), 268–282.

James A. Fill and Alan J. Izenman, *The structure of RI-invariant twelve-tone rows*, J. Austral. Math. Soc. Ser. B 21 (1979/80), 402–417.

Colin D. Fox, *Alban Berg the mathematician*, Math. Sci. 4 (1979), 105–107.

David Reiner, *Enumeration in music theory*, Amer. Math. Monthly 92 (1) (1985), 51–54.

9.14. Pitch class sets

A *pitch class set* is defined to be a subset of the set of twelve pitch classes. For convenience, we number the pitch classes $\{0, 1, \ldots, 11\}$ as in §9.6.

Atonal theorists and composers such as Milton Babbitt, Allen Forte and Elliott Carter put an equivalence relation on pitch class sets. They say that two pitch class sets are *equivalent* if one can be obtained from the other using only transpositions $\mathbf{T}^n$ and inversion $\mathbf{I}$. In other words, the equivalence classes are the orbits of the dihedral group D_{24} generated by $\mathbf{T}$ and $\mathbf{I}$ on the collection of subsets of $\{0, 1, \ldots, 11\}$.

We can use Burnside's Lemma 9.13.1 to count how many equivalence classes there are of each size. For this purpose, we need to count the fixed points of the elements of D_{24} on the collection of sets of a given size. It is easy to verify the following table.

Group element	size of subset												
	0	1	2	3	4	5	6	7	8	9	10	11	12
Identity	1	12	66	220	495	792	924	792	495	220	66	12	1
$\mathbf{T}, \mathbf{T}^5, \mathbf{T}^7, \mathbf{T}^{11}$	1	0	0	0	0	0	0	0	0	0	0	0	1
$\mathbf{T}^2, \mathbf{T}^{10}$	1	0	0	0	0	0	2	0	0	0	0	0	1
$\mathbf{T}^3, \mathbf{T}^9$	1	0	0	0	3	0	0	0	3	0	0	0	1
$\mathbf{T}^4, \mathbf{T}^8$	1	0	0	4	0	0	6	0	0	4	0	0	1
$\mathbf{T}^6$	1	0	6	0	15	0	20	0	15	0	6	0	1
$\mathbf{T}^{2m}\mathbf{I}$	1	2	6	10	15	20	20	20	15	10	6	2	1
$\mathbf{T}^{2m+1}\mathbf{I}$	1	0	6	0	15	0	20	0	15	0	6	0	1

For example, the first row just consists of the binomial coefficients $\binom{12}{j}$, where j is the size of the subset. The remaining rows of the table for powers of $\mathbf{T}$ are also just binomial coefficients, but interspersed with zeros. The inversions $\mathbf{T}^n\mathbf{I}$ come in two varieties. If $n = 2m + 1$ is odd, then there are no fixed pitch classes. So the fixed subsets have even size and the numbers are again binomial coefficients $\binom{6}{j}$, where $2j$ is the size of the subset. If $n = 2m$ is even, then there are two fixed pitch classes, so there are $2\binom{5}{j}$ fixed subsets of odd size $2j + 1$.

We can now apply Burnside's Lemma 9.13.1 to find how many orbits of D_{24} there are on the subsets of various sizes. The answers are as follows.

size of subset	0	1	2	3	4	5	6	7	8	9	10	11	12
number of orbits	1	1	6	12	29	38	50	38	29	12	6	1	1

For example, to compute how many subsets there are of size 5, we compute

$$\tfrac{1}{24}(792 + 6 \times 20) = \tfrac{912}{24} = 38.$$

For reference, we can also compute the number of orbits under the group $\mathbb{Z}/12$ consisting of powers of $\mathbf{T}$ using the same data. The answers are as follows.

size of subset	0	1	2	3	4	5	6	7	8	9	10	11	12
number of orbits	1	1	6	19	43	66	80	66	43	19	6	1	1

Incidentally, the reason for the symmetry in the above tables is that complementation gives a one to one correspondence between subsets of size j and subsets of size $12 - j$, and this correspondence is preserved by the action of the group D_{24}.

Allen Forte describes the following method for choosing a preferred representative from each orbit, called the *prime form*.[6] When the elements of the subset are listed in increasing order, the first should be zero, and the last should be as small as possible. If there is more than one representative with the same last term, then the second should be as small as possible, then the third, and so on up to the next to last. In other words, the prime form is the earliest in the lexicographic order with respect to (first, last, second, third, ..., next to last).

For example, take the set $\{1, 7, 9\}$. We can use $\mathbf{T}^{11}$ to take it to a set containing zero, namely $\{0, 6, 8\}$. Or we could use $\mathbf{T}^5$ to take it to $\{0, 2, 6\}$, or $\mathbf{T}^3$ to take it to $\{0, 4, 10\}$. We also need to use $\mathbf{I}$ to get $\{3, 5, 11\}$, and then use powers of $\mathbf{T}$ to get $\{0, 2, 8\}$, $\{0, 4, 6\}$ and $\{0, 6, 10\}$. Of the six possibilities, the ones with the smallest last term are $\{0, 2, 6\}$ and $\{0, 4, 6\}$. To break the tie, we compare second terms, and we see that $\{0, 2, 6\}$ is the prime form.

There is an easy way to attach an invariant to each orbit, called the *interval vector*. This is computed as follows. To an unordered pair of pitch classes, we can assign a difference, in the range from 1 to 6, by going around the circle of pitch classes in the shorter of the two possible directions. Take all unordered pairs in the set, and to each pair find the difference in this way. Then record how many times one, two, up to six occur in a row vector of length six. For example, for the set $\{1, 7, 9\}$ the three differences are 2, 4 and 6. So the interval vector for this pitch class set is (0,1,0,1,0,1). It is clear that equivalent pitch class sets yield the same interval vectors. The converse is false; for example the sets $\{0, 1, 4, 6\}$ and $\{0, 1, 3, 7\}$ both have interval vector (1,1,1,1,1,1).

Here is a list of the prime forms of pitch class sets of size three, together with Allen Forte's name and Elliott Carter's numbering for them, and the interval vector.

Set	Forte	Carter	Vector
{0,1,2}	3-1(12)	4	(2,1,0,0,0,0)
{0,1,3}	3-2	12	(1,1,1,0,0,0)
{0,1,4}	3-3	11	(1,0,1,1,0,0)
{0,1,5}	3-4	9	(1,0,0,1,1,0)
{0,1,6}	3-5	7	(1,0,0,0,1,1)
{0,2,4}	3-6(12)	3	(0,2,0,1,0,0)
{0,2,5}	3-7	10	(0,1,1,0,1,0)
{0,2,6}	3-8	8	(0,1,0,1,0,1)
{0,2,7}	3-9(12)	5	(0,1,0,0,2,0)
{0,3,6}	3-10(12)	2	(0,0,2,0,0,1)
{0,3,7}	3-11	6	(0,0,1,1,1,0)
{0,4,8}	3-12(4)	1	(0,0,0,3,0,0)

[6] This should not be confused with the prime form of a tone row, described in §9.8.

Forte's number consists of the set size followed by a number indicating the placement with respect to lexicographical ordering on the interval vector, in backward order. The numbers in parentheses give the orbit size under the action of D_{24}, in case this is not 24. For reference, we give the corresponding information for sets of size four, five and six below. Sets of size greater than six are not named by Carter, and Forte uses the names of the complementary set, but with the initial number changed. So for example 9-3 is obtained by complementing 3-3 to obtain $\{2, 3, 5, 6, 7, 8, 9, 10, 11\}$, which is then put into prime form as $\{0, 1, 2, 3, 4, 5, 6, 8, 9\}$.

There is an easy way to obtain the interval vector for the complement of a set. For size three, add the vector (6,6,6,6,6,3); for size four add (4,4,4,4,4,2); and for size five add (2,2,2,2,2,1). The interval vector for the above three element set is (1,0,1,1,0,0), so for its nine element complement we get (7,6,7,7,6,3).

Set	Forte	Carter	Vector
{0,1,2,3}	4-1(12)	1	(3,2,1,0,0,0)
{0,1,2,4}	4-2	17	(2,2,1,1,0,0)
{0,1,3,4}	4-3(12)	9	(2,1,2,1,0,0)
{0,1,2,5}	4-4	20	(2,1,1,1,1,0)
{0,1,2,6}	4-5	22	(2,1,0,1,1,1)
{0,1,2,7}	4-6(12)	6	(2,1,0,0,2,1)
{0,1,4,5}	4-7(12)	8	(2,0,1,2,1,0)
{0,1,5,6}	4-8(12)	10	(2,0,0,1,2,1)
{0,1,6,7}	4-9(6)	2	(2,0,0,0,2,2)
{0,2,3,5}	4-10(12)	3	(1,2,2,0,1,0)
{0,1,3,5}	4-11	26	(1,2,1,1,1,0)
{0,2,3,6}	4-12	28	(1,1,2,1,0,1)
{0,1,3,6}	4-13	7	(1,1,2,0,1,1)
{0,2,3,7}	4-14	25	(1,1,1,1,2,0)
{0,1,4,6}	4-Z15	18	(1,1,1,1,1,1)

Set	Forte	Carter	Vector
{0,1,5,7}	4-16	19	(1,1,0,1,2,1)
{0,3,4,7}	4-17(12)	13	(1,0,2,2,1,0)
{0,1,4,7}	4-18	21	(1,0,2,1,1,1)
{0,1,4,8}	4-19	24	(1,0,1,3,1,0)
{0,1,5,8}	4-20(12)	15	(1,0,1,2,2,0)
{0,2,4,6}	4-21(12)	11	(0,3,0,2,0,1)
{0,2,4,7}	4-22	27	(0,2,1,1,2,0)
{0,2,5,7}	4-23(12)	4	(0,2,1,0,3,0)
{0,2,4,8}	4-24(12)	16	(0,2,0,3,0,1)
{0,2,6,8}	4-25(6)	12	(0,2,0,2,0,2)
{0,3,5,8}	4-26(12)	14	(0,1,2,1,2,0)
{0,2,5,8}	4-27	29	(0,1,2,1,1,1)
{0,3,6,9}	4-28(3)	5	(0,0,4,0,0,2)
{0,1,3,7}	4-Z29	23	(1,1,1,1,1,1)

The only extra thing to describe here is the meaning of the symbol Z in the Forte naming system. This indicates that there are two orbits with the same interval vector; the second one is listed at the end for some reason which he never explains. The same happens for sets of size five and six, but more often.

Set	Forte	Carter	Vector
{0,1,2,3,4}	5-1(12)	1	(4,3,2,1,0,0)
{0,1,2,3,5}	5-2	11	(3,3,2,1,1,0)
{0,1,2,4,5}	5-3	14	(3,2,2,2,1,0)
{0,1,2,3,6}	5-4	12	(3,2,2,1,1,1)
{0,1,2,3,7}	5-5	13	(3,2,1,1,2,1)
{0,1,2,5,6}	5-6	27	(3,1,1,2,2,1)
{0,1,2,6,7}	5-7	30	(3,1,0,1,3,2)
{0,2,3,4,6}	5-8(12)	2	(2,3,2,2,0,1)
{0,1,2,4,6}	5-9	15	(2,3,1,2,1,1)
{0,1,3,4,6}	5-10	19	(2,2,3,1,1,1)
{0,2,3,4,7}	5-11	18	(2,2,2,2,2,0)
{0,1,3,5,6}	5-Z12(12)	5	(2,2,2,1,2,1)
{0,1,2,4,8}	5-13	17	(2,2,1,3,1,1)
{0,1,2,5,7}	5-14	28	(2,2,1,1,3,1)
{0,1,2,6,8}	5-15(12)	4	(2,2,0,2,2,2)
{0,1,3,4,7}	5-16	20	(2,1,3,2,1,1)
{0,1,3,4,8}	5-Z17(12)	10	(2,1,2,3,2,0)
{0,1,4,5,7}	5-Z18	35	(2,1,2,2,2,1)
{0,1,3,6,7}	5-19	31	(2,1,2,1,2,2)

Set	Forte	Carter	Vector
{0,1,3,7,8}	5-20	34	(2,1,1,2,3,1)
{0,1,4,5,8}	5-21	21	(2,0,2,4,2,0)
{0,1,4,7,8}	5-22(12)	8	(2,0,2,3,2,1)
{0,2,3,5,7}	5-23	25	(1,3,2,1,3,0)
{0,1,3,5,7}	5-24	22	(1,3,1,2,2,1)
{0,2,3,5,8}	5-25	24	(1,2,3,1,2,1)
{0,2,4,5,8}	5-26	26	(1,2,2,3,1,1)
{0,1,3,5,8}	5-27	23	(1,2,2,2,3,0)
{0,2,3,6,8}	5-28	36	(1,2,2,2,1,2)
{0,1,3,6,8}	5-29	32	(1,2,2,1,3,1)
{0,1,4,6,8}	5-30	37	(1,2,1,3,2,1)
{0,1,3,6,9}	5-31	33	(1,1,4,1,1,2)
{0,1,4,6,9}	5-32	38	(1,1,3,2,2,1)
{0,2,4,6,8}	5-33(12)	6	(0,4,0,4,0,2)
{0,2,4,6,9}	5-34(12)	9	(0,3,2,2,2,1)
{0,2,4,7,9}	5-35(12)	7	(0,3,2,1,4,0)
{0,1,2,4,7}	5-Z36	16	(2,2,2,1,2,1)
{0,3,4,5,8}	5-Z37(12)	3	(2,1,2,3,2,0)
{0,1,2,5,8}	5-Z38	29	(2,1,2,2,2,1)

Finally, the six note pitch class sets, or *hexachords*.

Set	Forte	Carter	Vector
{0,1,2,3,4,5}	6-1(12)	4	(5,4,3,2,1,0)
{0,1,2,3,4,6}	6-2	19	(4,4,3,2,1,1)
{0,1,2,3,5,6}	6-Z3	49	(4,3,3,2,2,1)
{0,1,2,4,5,6}	6-Z4(12)	24	(4,3,2,3,2,1)
{0,1,2,3,6,7}	6-5	16	(4,2,2,2,3,2)
{0,1,2,5,6,7}	6-Z6(12)	33	(4,2,1,2,4,2)
{0,1,2,6,7,8}	6-7(6)	7	(4,2,0,2,4,3)
{0,2,3,4,5,7}	6-8(12)	5	(3,4,3,2,3,0)
{0,1,2,3,5,7}	6-9	20	(3,4,2,2,3,1)
{0,1,3,4,5,7}	6-Z10	42	(3,3,3,3,2,1)
{0,1,2,4,5,7}	6-Z11	47	(3,3,3,2,3,1)
{0,1,2,4,6,7}	6-Z12	46	(3,3,2,2,3,2)
{0,1,3,4,6,7}	6-Z13(12)	29	(3,2,4,2,2,2)
{0,1,3,4,5,8}	6-14	3	(3,2,3,4,3,0)
{0,1,2,4,5,8}	6-15	13	(3,2,3,4,2,1)
{0,1,4,5,6,8}	6-16	11	(3,2,2,4,3,1)
{0,1,2,4,7,8}	6-Z17	35	(3,2,2,3,3,2)
{0,1,2,5,7,8}	6-18	17	(3,2,2,2,4,2)
{0,1,3,4,7,8}	6-Z19	37	(3,1,3,4,3,1)
{0,1,4,5,8,9}	6-20(4)	2	(3,0,3,6,3,0)
{0,2,3,4,6,8}	6-21	12	(2,4,2,4,1,2)
{0,1,2,4,6,8}	6-22	10	(2,4,1,4,2,2)
{0,2,3,5,6,8}	6-Z23(12)	27	(2,3,4,2,2,2)
{0,1,3,4,6,8}	6-Z24	39	(2,3,3,3,3,1)
{0,1,3,5,6,8}	6-Z25	43	(2,3,3,2,4,1)

Set	Forte	Carter	Vector
{0,1,3,5,7,8}	6-Z26(12)	26	(2,3,2,3,4,1)
{0,1,3,4,6,9}	6-27	14	(2,2,5,2,2,2)
{0,1,3,5,6,9}	6-Z28(12)	21	(2,2,4,3,2,2)
{0,1,3,6,8,9}	6-Z29(12)	32	(2,2,4,2,3,2)
{0,1,3,6,7,9}	6-30(12)	15	(2,2,4,2,2,3)
{0,1,3,5,8,9}	6-31	8	(2,2,3,4,3,1)
{0,2,4,5,7,9}	6-32(12)	6	(1,4,3,2,5,0)
{0,2,3,5,7,9}	6-33	18	(1,4,3,2,4,1)
{0,1,3,5,7,9}	6-34	9	(1,4,2,4,2,2)
{0,2,4,6,8,10}	6-35(2)	1	(0,6,0,6,0,3)
{0,1,2,3,4,7}	6-Z36	50	(4,3,3,2,2,1)
{0,1,2,3,4,8}	6-Z37(12)	23	(4,3,2,3,2,1)
{0,1,2,3,7,8}	6-Z38(12)	34	(4,2,l,2,4,2)
{0,2,3,4,5,8}	6-Z39	41	(3,3,3,3,2,1)
{0,1,2,3,5,8}	6-Z40	48	(3,3,3,2,3,1)
{0,1,2,3,6,8}	6-Z41	45	(3,3,2,2,3,2)
{0,1,2,3,6,9}	6-Z42(12)	30	(3,2,4,2,2,2)
{0,1,2,5,6,8}	6-Z43	36	(3,2,2,3,3,2)
{0,1,2,5,6,9}	6-Z44	38	(3,1,3,4,3,1)
{0,2,3,4,6,9}	6-Z45(12)	28	(2,3,4,2,2,2)
{0,1,2,4,6,9}	6-Z46	40	(2,3,3,3,3,1)
{0,1,2,4,7,9}	6-Z47	44	(2,3,3,2,4,1)
{0,1,2,5,7,9}	6-Z48(12)	25	(2,3,2,3,4,1)
{0,1,3,4,7,9}	6-Z49(12)	22	(2,2,4,3,2,2)
{0,1,4,6,7,9}	6-Z50(12)	31	(2,2,4,2,3,2)

Complementation takes some hexachords to equivalent ones and some to inequivalent ones. Inequivalent pairs always share an interval vector, and these turn out to be the only coincidences of interval vectors for hexachords. The inequivalent pairs of complements are as follows:

6-Z3	6-Z36	6-Z12	6-Z41	6-Z24	6-Z46
6-Z4(12)	6-Z37(12)	6-Z13(12)	6-Z42	6-Z25	6-Z47
6-Z6(12)	6-Z38(12)	6-Z17	6-Z43	6-Z26(12)	6-Z48(12)
6-Z10	6-Z39	6-Z19	6-Z44	6-Z28(12)	6-Z49(12)
6-Z11	6-Z40	6-Z23(12)	6-Z45(12)	6-Z29(12)	6-Z50(12)

Further reading:

Allen Forte, *The structural function of atonal music* [**41**].

David Schiff, *The music of Elliott Carter*. Ernst Eulenberg Ltd, 1983. Reprinted by Faber and Faber, 1998.

9.15. Pólya's enumeration theorem

In this section, we show how to vamp up Burnside's Lemma 9.13.1 to address some more complicated counting problems. By way of illustration, we shall revisit the problem considered in §9.14. Suppose we want to know how many pitch class sets there are, consisting of three of the twelve possible pitch classes. Suppose further that we wish to consider two such sets to be equivalent if one can be obtained from the other by means of an operation $\mathbf{T}^n$ for some n. This is a typical kind of problem which can be solved using Pólya's enumeration theorem.

A lot of physical counting problems involving symmetry are of a similar nature. A typical example would involve counting how many different necklaces can be made from three red beads, two sepia beads and five turquoise beads. The symmetry group in this situation is a dihedral group whose order is twice the number of beads.

In the general form of the problem, the *configurations* being counted are regarded as functions from a set X to a set Y, and the symmetry group G acts on the set X. In the bead problem, the set X would consist of the places in the necklace where we wish to put the beads, and the set Y would consist of the possible colors. A function from X to Y then specifies for each place in the necklace what color bead to use. The group G acts on X by rotating and turning over the necklace.

In the pitch class set counting problem, the set X is the set of twelve pitches, and Y is taken to be the set $\{0, 1\}$. A pitch class set corresponds to a function taking the notes in the set to 1 and the remaining notes to 0. This gives a one-to-one correspondence between pitch class sets and functions from X to Y.

In the general setup, we write Y^X for the set of configurations, or functions from the set X to the set Y. The reason for this notation is that the number of elements of Y^X is equal to the number of elements of Y raised to the power of the number of elements of X ($|Y^X| = |Y|^{|X|}$). The action of G on the set Y^X of configurations is given by the formula

$$g(f)(x) = f(g^{-1}(x)).$$

The reason for the inverse sign is so that composition works right. For a group action, we need $g_1(g_2(f)) = (g_1g_2)(f)$. To see that this holds, we have

$$\begin{aligned}(g_1(g_2(f)))(x) &= (g_2(f))(g_1^{-1}(x)) = f(g_2^{-1}(g_1^{-1}(x))) = f((g_2^{-1}g_1^{-1})(x)) \\ &= f((g_1g_2)^{-1}(x)) = ((g_1g_2)(f))(x),\end{aligned}$$

whereas without the inverse sign the order of g_1 and g_2 would be reversed. The general problem is to find the number of orbits of G on configurations.

We begin by defining the *cycle index* of G on X as follows. We introduce variables $t_1, t_2, \dots$, and then the cycle index of an element g on X is

$$P_g(t_1, t_2, \dots) = t_1^{j_1(g)} t_2^{j_2(g)} \dots$$

where $j_k(g)$ denotes the number of cycles of length k in the action of G on X. We define the cycle index of the group to be the average cycle index of an element, namely

$$P_G(t_1, t_2, \dots) = \frac{1}{|G|} \sum_{g \in G} P_g(t_1, t_2, \dots) = \frac{1}{|G|} \sum_{g \in G} t_1^{j_1(g)} t_2^{j_2(g)} \dots \qquad (9.15.1)$$

For example, if G is a dihedral group of order eight acting on the set X consisting of the four corners of a square, then the cycle indices of the eight elements of G are as follows. The identity element has cycle index t_1^4, the two ninety degree rotations have cycle index t_4, the one hundred and eighty degree rotation and the reflections about the horizontal and vertical axes all have cycle index t_2^2, and the two diagonal reflections have cycle index $t_1^2 t_2$. So

$$P_G = \tfrac{1}{8}(t_1^4 + 2t_4 + 3t_2^2 + 2t_1^2 t_2).$$

Several standard examples of cycle index are worth writing out explicitly. If $G = \mathbb{Z}/n$, cycling a set X of n objects, we get

$$P_{\mathbb{Z}/n} = \frac{1}{n}\sum_{j|n} \phi(j) t_j^{n/j}. \tag{9.15.2}$$

Here, ϕ is the Euler phi function, described on page 301, and $j|n$ means j is a divisor of n. The formula is obvious, because there are $\phi(j)$ elements of $\mathbb{Z}/n$ having order j, and each one has n/j cycles of length j.

The next example generalizes the above dihedral calculation. For the dihedral group D_{2n} acting on the n vertices of a regular n-sided polygon, we have to divide into two cases according to whether n is even or odd. If $n = 2m + 1$ is odd, we get

$$P_{D_{4m+2}} = \tfrac{1}{2}P_{\mathbb{Z}/(2m+1)} + \tfrac{1}{2}t_1 t_2^m, \tag{9.15.3}$$

because each reflection has exactly one fixed point. If $n = 2m$ is even, we get

$$P_{D_{4m}} = \tfrac{1}{2}P_{\mathbb{Z}/2m} + \tfrac{1}{4}(t_2^m + t_1^2 t_2^{m-1}), \tag{9.15.4}$$

because half the reflections have no fixed points and half of them have two.

For the full symmetric group S_n on a set X of n elements, the formula is rather messy. But adding up the cycle indices of all the symmetric groups gives a much cleaner answer.

$$\sum_{n=0}^{\infty} P_{S_n} = \exp\left(\sum_{j=1}^{\infty}\frac{t_j}{j}\right) = \prod_{j=1}^{\infty}\sum_{i=0}^{\infty}\frac{1}{i!}\left(\frac{t_j}{j}\right)^i$$
$$= (1 + t_1 + \tfrac{1}{2!}t_1^2 + \tfrac{1}{3!}t_1^3 + \tfrac{1}{4!}t_1^4 + \dots)(1 + \tfrac{1}{2}t_2 + \tfrac{1}{2^2.2!}t_2^2 + \tfrac{1}{2^3.3!}t_2^3 + \dots)$$
$$(1 + \tfrac{1}{3}t_3 + \tfrac{1}{3^2.2!}t_3^2 + \tfrac{1}{3^3.3!}t_3^3 + \dots)(1 + \tfrac{1}{4}t_4 + \tfrac{1}{4^2.2!}t_4^2 + \tfrac{1}{4^3.3!}t_4^3 + \dots)\dots$$

The cycle index for an individual S_n can be extracted by taking the terms with total size n, where each t_j is regarded as having size j. So for example

$$P_{S_4} = \tfrac{1}{24}t_1^4 + \tfrac{1}{4}t_1^2 t_2 + \tfrac{1}{8}t_2^2 + \tfrac{1}{3}t_1 t_3 + \tfrac{1}{4}t_4.$$

The corresponding formula for the alternating group A_n (this is the group of even permutations; exactly half the elements of S_n are even) is

$$2 + 2t_1 + \sum_{n=2}^{\infty} P_{A_n} = \exp\left(\sum_{j=1}^{\infty}\frac{t_j}{j}\right) + \exp\left(\sum_{j=1}^{\infty}(-1)^{j+1}\frac{t_j}{j}\right).$$

Next, we assign a weight $w(y)$ to each of the elements y of Y. The weights can be any sorts of quantities which can be added and multiplied (the formal requirement is that the weights should belong to a *commutative ring*). For example, the weights can be independent formal variables, or one of them can be chosen to be 1 to simplify the algebra. The weight of a configuration is then defined to be the product over $x \in X$ of the weight of $f(x)$,

$$w(f) = \prod_{x\in X} w(f(x)).$$

The weights of two configurations in the same orbit of the action of G are clearly equal.

So for example if $Y = \{\text{red, sepia, turquoise}\}$ then we could assign variables $r = w(\text{red})$, $s = w(\text{sepia})$ and $t = w(\text{turquoise})$ for the weights.

We form a power series called the *configuration counting series* C using these weights. Namely, C is the sum, over all orbits of G on the set Y^X configurations, of the weight of a representative of the orbit. In the necklace example, the coefficient of $r^a s^b t^c$ in $C = C(r, s, t)$ gives the number of necklaces in which a beads are red, b are sepia and c are turquoise. So the coefficient of $r^3 s^2 t^5$ would give the number of necklaces in the original problem. Since $a+b+c$ is fixed, if we wanted to simplify the algebra, it would make sense to put $w(\text{turquoise}) = 1$ instead of t. Then the coefficient of $r^3 s^2$ would be the desired number of necklaces. In other words, once we know the number of red and sepia beads, the number of turquoise beads is also known by subtraction.

In the pitch class set example, where $Y = \{0, 1\}$, it would make sense to introduce just one variable z and set $w(0) = 1$ and $w(1) = z$. Then the coefficient of z^a would tell us about pitch class sets with a notes.

THEOREM 9.15.1 (Pólya). *The configuration counting series C is given in terms of the cycle index of G on X by*

$$C = P_G\left(\sum_{y\in Y} w(y), \sum_{y\in Y} w(y)^2, \sum_{y\in Y} w(y)^3, \dots\right)$$

We shall prove this theorem after seeing how to apply it.

Example. In the pitch class set example, we consider the cases $G = \mathbb{Z}/12$ and $G = D_{24}$, with X is the set of twelve pitch classes, $Y = \{0, 1\}$, $w(0) = 1$ and $w(1) = y$. Equations (9.15.2) and (9.15.4) give the cycle indices as

$$\begin{aligned} P_{\mathbb{Z}/12} &= \tfrac{1}{12}(t_1^{12} + t_2^6 + 2t_3^4 + 2t_4^3 + 2t_6^2 + 4t_{12}) \\ P_{D_{24}} &= \tfrac{1}{2}P_{\mathbb{Z}/12} + \tfrac{1}{4}(t_2^6 + t_1^2 t_2^5) \\ &= \tfrac{1}{24}(t_1^{12} + 6t_1^2 t_2^5 + 7t_2^6 + 2t_3^4 + 2t_4^3 + 2t_6^2 + 4t_{12}). \end{aligned}$$

Then Theorem 9.15.1 says that we should substitute $1 + z^n$ for t_n to give the configuration counting series C. This gives the following values.

(i) If $G = \mathbb{Z}/12$ then

$$C = 1+z+6z^2+19z^3+43z^4+66z^5+80z^6+66z^7+43z^8+19z^9+6z^{10}+z^{11}+z^{12}.$$

So for example there are 19 three note sets up to transposition.

(ii) If $G = D_{24}$ then

$$C = 1+z+6z^2+12z^3+29z^4+38z^5+50z^6+38z^7+29z^8+12z^9+6z^{10}+y^{11}+y^{12}.$$

So for example there are 12 three note sets and 50 hexachords, up to transposition and inversion. The reason why the coefficients in these polynomials are symmetric was described in §9.14. Namely, a set can be replaced by its complement, to give a natural correspondence between j note sets and $12-j$ note sets.

The advantage of using Pólya's enumeration theorem rather than just resorting to Burnside's Lemma 9.13.1 is that we do not have to do an explicit computation of numbers of fixed configurations, as we had to in §9.14. The disadvantage is that the machinery is harder to understand and remember.

The proof of Pólya's enumeration theorem depends on a weighted version of Burnside's Lemma 9.13.1.

LEMMA 9.15.2. *Let G be a finite group acting by permutations on a finite set X. Let w be a function on X which takes constant values on orbits, so that we can regard w as a function on the set of orbits of G on X. Then the sum of the weights of the orbits is equal to*

$$\frac{1}{|G|}\sum_{g\in G}\sum_{x=g(x)} w(x).$$

PROOF. Consider the set of pairs (g,x) where $g(x)=x$, and calculate in two different ways the sum over the elements of this set of the weights $w(x)$. If we sum over the elements of the group first, we obtain $\sum_{g\in G}\sum_{x=g(x)} w(x)$. On the other hand, if we sum over the elements of X first, then by equation (9.11.1), for each x, the number of elements of G is $|G|$ divided by the length of the orbit in which x lies. So the sum over the elements of the orbit in which x lies gives $|G|w(x)$. So the sum over all x gives $|G|$ times the sum of the weights of the orbits. □

PROOF OF PÓLYA'S ENUMERATION THEOREM. We are going to apply the above version of Burnside's lemma to the action of G on the set Y^X of configurations. It tells us that C is equal to

$$\frac{1}{|G|}\sum_{g\in G}\sum_{f=g(f)} w(f). \tag{9.15.5}$$

So we will be finished if we can prove that for each $g\in G$ we have

$$P_g\left(\sum_{y\in Y} w(y), \sum_{y\in Y} w(y)^2, \sum_{y\in Y} w(y)^3, \dots\right) = \sum_{f=g(f)} w(f),$$

because then, comparing (9.15.1) with (9.15.5), we see that averaging over the elements of G gives the formula in the theorem. Recalling that $j_k(g)$ denotes the number of cycles of length k in the action of g on X, by definition the left side of this equation is

$$\left(\sum_{y\in Y} w(y)\right)^{j_1(g)} \left(\sum_{y\in Y} w(y)^2\right)^{j_2(g)} \dots \tag{9.15.6}$$

The right hand side is

$$\sum_{f=g(f)}\prod_{x\in X} w(f(x)). \tag{9.15.7}$$

Now a configuration f satisfies $f=g(f)$ precisely when it is constant on orbits of g on X. So to pick such a configuration, we must assign an element of Y to each orbit of g on X. So when we multiply the weights of the $f(x)$, an orbit of length j with image $y \in Y$ corresponds to a factor of $w(y)^j$ in the product.

We regard (9.15.6) as being obtained by multiplying together a factor of $\sum_{y \in Y} w(y)^i$ for each orbit of g on X, where i is the length of the orbit. When these sums are all multiplied out, there will be one term for each way of assigning an element of Y to each orbit of g on X, and that term will exactly be the corresponding term in (9.15.7). $\square$

Further reading:

Harald Fripertinger, *Enumeration in music theory*, Séminaire Lotharingien de Combinatoire, 26 (1991), 29–42; also appeared in Beiträge zur Elektronischen Musik 1, 1992.

Harald Fripertinger, *Enumeration and construction in music theory*, Diderot Forum on Mathematics and Music Computational and Mathematical Methods in Music, Vienna, Austria, December 2-4, 1999. H. G. Feichtinger and M. Dörfler, editors. Österreichische Computergesellschaft (1999), 179–204.

Harald Fripertinger, *Enumeration of mosaics*, Discrete Math. 199 (1999), 49–60.

Harald Fripertinger, *Enumeration of non-isomorphic canons*, Tatra Mountains Math. Publ. 23 (2001).

Harald Fripertinger, *Classification of motives: a mathematical approach*, to appear in Musikometrika.

Michael Keith, *From polychords to Pólya; adventures in musical combinatorics* [**66**].

G. Pólya, *Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen*, Acta Math. 68 (1937), 145–254.

R. C. Read, *Combinatorial problems in the theory of music*, Discrete Mathematics 167/168 (1997), 543–551.

D. Reiner, *Enumeration in music theory*, Amer. Math. Monthly 92 (1) (1985), 51–54. Note that there is a typographical error in the formula for the cycle index of the dihedral group in this paper.

9.16. The Mathieu group M_{12}

The combinatorics of twelve tone music has given rise to a curious coincidence, which I find worth mentioning. Messiaen, in his *Ile de feu 2* for piano, nearly rediscovered the Mathieu group M_{12}. On pages 409–414 of Berry (reference at the end of the section), you can read about how Messiaen uses the permutations

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 7 & 6 & 8 & 5 & 9 & 4 & 10 & 3 & 11 & 2 & 12 & 1 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 6 & 7 & 5 & 8 & 4 & 9 & 3 & 10 & 2 & 11 & 1 & 12 \end{pmatrix}$$

to generate sequences of tones and sequences of durations. These permutations generate a group M_{12} of order $95,040$ discovered by Mathieu in the nineteenth century.[7]

A group is said to be *simple* if it has just two normal subgroups, namely the whole group and the subgroup consisting of just the identity element.[8] One of the outstanding achievements of twentieth century mathematics was the classification of the finite simple groups. Roughly speaking, the classification theorem says that the finite simple groups fall into certain infinite families which can be explicitly described, with the exception of 26 *sporadic* groups. Five of these 26 groups were discovered by Mathieu in the nineteenth century, and the remaining ones were discovered in the nineteen sixties and seventies.

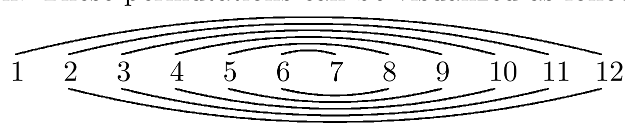
Diaconis, Graham and Kantor discovered that M_{12} was generated by the above two permutations, which they call *Mongean shuffles.* Start with a pack of twelve cards in your left hand, and transfer them to your right hand by placing them alternately under and over the stack you have so far. When you have finished, hand the pack back to your left hand. Since I did not tell you whether to start under or over, this describes two different permutations of the twelve cards. These are the permutations shown above. In cycle notation, these permutations are

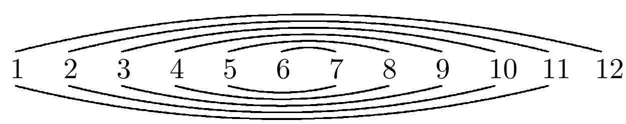
$$(1, 7, 10, 2, 6, 4, 5, 9, 11, 12)(3, 8)$$

of order ten, and

$$(1, 6, 9, 2, 7, 3, 5, 4, 8, 10, 11)(12)$$

of order eleven. These permutations can be visualized as follows.





Exercises

1. (Carl E. Linderholm [**76**]) If this book is read backwards (beginning at the last word of the last page), the last thing read is the introduction (reversed, of course). Thus the introduction acts as a sort of extraduction, and

[7] E. Mathieu, *Mémoire sur l'étude des fonctions de plusieurs quantités*, J. Math. Pures Appl. **6** (1861), 241–243; *Sur la fonction cinq fois transitive de* 24 *quantités*, J. Math. Pures Appl. **18** (1873), 25–46.

[8] So for example the group with only one element is not simple, because it has only one, not two, normal subgroups. Compare this with the definition of a prime number; 1 is not prime.

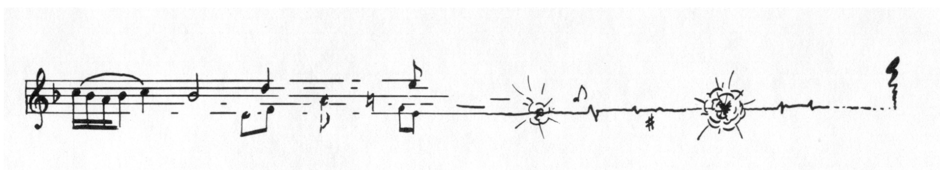
is suggested as a simple form of therapy, used in this way, if the reader gets stuck. Read this exercise backwards, and write an extraduction from it.

Further reading:

Wallace Berry, *Structural function in music*, Prentice-Hall, 1976. Reprinted by Dover, 1987. 447 pages, in print. ISBN 0486253848. This book contains a description of the Messiaen example referred to in this section.

J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups*, Grundlehren der mathematischen Wissenschaften 290, Springer-Verlag, Berlin/New York, 1988. This book contains a huge amount of information about the sporadic groups in general, and §11.17 contains more information on Mongean shuffles and the Mathieu group M_{12}.

P. Diaconis, R. L. Graham and W. M. Kantor, *The mathematics of perfect shuffles*, Adv. Appl. Math. **4** (1983), 175–196.



Unlike Mozart's *Requiem* and Bartok's *Third Piano Concerto*, the piece that P. D. Q. Bach was working on when he died has never been finished by anyone else.[9]

[9]Professor Peter Schickele, *The definitive biography of P. D. Q. Bach (1807–1742)?*, Random House, New York, 1976.

APPENDIX A

Answers to almost all exercises

§1.3 #**1.** The power has been quadrupled, so this represents a change of $\log_{10}(4)$ decibels, or approximately 6.02 dB.

§1.3 #**2.** (c) 73 dB. The power is doubled, so the number of decibels is increased by $\log_{10}(2)$.

§1.5 #**1.** We have

$$\frac{dy}{dt} = -A\sqrt{k/m}\sin(\sqrt{k/m}\,t) + B\sqrt{k/m}\cos(\sqrt{k/m}\,t)$$
$$\frac{d^2y}{dt^2} = -A(k/m)\cos(\sqrt{k/m}\,t) - B(k/m)\sin(\sqrt{k/m}\,t) = -(k/m)y.$$

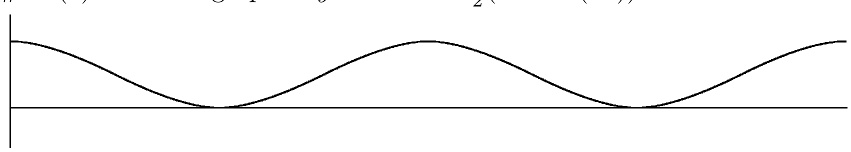
§1.5 #**2.** Take $c = \sqrt{A^2+B^2}$ and $B\tan\phi = A$. Beware that it is not correct to write $\phi = \tan^{-1} A/B$. This is only true when B is positive. When B is negative we have $\phi = \pi + \tan^{-1} A/B$. When $B = 0$, ϕ is either $\pi/2$ or $-\pi/2$, depending on the sign of A.

§1.7 #**1.** $\sin u + \cos v = 2\sin(\frac{\pi}{4} + \frac{u+v}{2})\sin(\frac{\pi}{4} + \frac{u-v}{2})$.

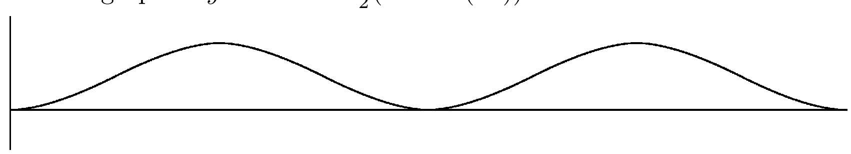
§1.7 #**2.** The frequency of vibration of the other string is either 435 Hz or 445 Hz.

§1.7 #**3.** $\int_0^{\pi/2}\sin(3x)\sin(4x)\,dx = \int_0^{\pi/2}\frac{1}{2}(\cos(x)-\cos(7x))\,dx$
$= [\frac{1}{2}\sin(x) - \frac{1}{14}\sin(7x)]_0^{\pi/2} = \frac{1}{2} + \frac{1}{14} = \frac{4}{7}$.

§1.7 #**4.** (a) Here is a graph of $y = \cos^2 x = \frac{1}{2}(1+\cos(2x))$:



(b) Here is a graph of $y = \sin^2 x = \frac{1}{2}(1-\cos(2x))$:



§1.7 #**5.** $\int_0^{\frac{2\pi}{\omega}}[c\sin(\omega t+\phi)]^2\,dt = \int_0^{\frac{2\pi}{\omega}}\frac{c^2}{2}(1-\cos 2(\omega t+\phi))\,dt$
$= \left[\frac{c^2}{2}(t - \frac{1}{2\omega}\sin 2(\omega t+\phi))\right]_0^{\frac{2\pi}{\omega}} = \frac{c^2}{2}\frac{2\pi}{\omega}$. Multiply both sides by $\frac{\omega}{2\pi}$ and then take the square root.

§1.7 #**6.** Put $A = kt$ and $B = \frac{1}{2}t$ in (1.7.11) to obtain the formula for $\sin kt \sin \frac{1}{2}t$. Then the sum on the left of equation (1.7.17), multiplied by $\sin\frac{1}{2}t$, can be rearranged to make a collapsing sum as follows:

$$\begin{aligned}
\sin\tfrac{1}{2}t\sum_{k=1}^{n}\sin kt &= \sum_{k=1}^{n}\sin kt\sin\tfrac{1}{2}t \\
&= \sum_{k=1}^{n}\tfrac{1}{2}(\cos(k-\tfrac{1}{2})t - \cos(k+\tfrac{1}{2})t) \\
&= (\tfrac{1}{2}\cos\tfrac{1}{2}t - \tfrac{1}{2}\cos\tfrac{3}{2}t) + (\tfrac{1}{2}\cos\tfrac{3}{2}t - \tfrac{1}{2}\cos\tfrac{5}{2}t) \\
&\qquad + \cdots + (\tfrac{1}{2}\cos(n-\tfrac{1}{2})t - \tfrac{1}{2}\cos(n+\tfrac{1}{2})t) \\
&= \tfrac{1}{2}\cos\tfrac{1}{2}t - \tfrac{1}{2}\cos(n+\tfrac{1}{2})t.
\end{aligned}$$

Now divide both sides by $\sin\frac{1}{2}t$ to obtain the first equality of equation (1.7.17). Finally, use equation (1.7.14) with $u = \frac{1}{2}t$ and $v = (n+\frac{1}{2})t$ to obtain the second equality in equation (1.7.17).

Equation (1.7.18) works the same way. We use equation (1.7.7) and the fact that $\sin(\frac{1}{2}-k)t = -\sin(k-\frac{1}{2})t$ to obtain

$$\cos kt\sin\tfrac{1}{2}t = \tfrac{1}{2}(\sin(k+\tfrac{1}{2})t - \sin(k-\tfrac{1}{2})t),$$

and then use a collapsing sum as before. The second equality of (1.7.18) uses equation (1.7.12).

§1.7 #**7.** Theoretically, no beats are heard in this situation. This is because if b is small then $\sin(a)+\sin(2a+b) = 2\sin\frac{1}{2}(3a+b)\cos\frac{1}{2}(a+b)$ does not give us a low frequency envelope. In the case of $\sin(a)+\sin(a+b) = 2\sin\frac{1}{2}(2a+b)\cos\frac{b}{2}$, the low frequency envelope function is $\cos\frac{b}{2}$.

§1.8 #**1.** (i) $\sin(2\pi t + \frac{\pi}{2})$,

(ii) $\sqrt{2}\sin(2\pi t + \pi/4)$,

(iii) Since the vectors $(2\cos\pi/6, 2\sin\pi/6) = (\sqrt{3},1)$ and $(-\cos\pi/2, -\sin\pi/2) = (0,-1)$ add up to $(\sqrt{3},0) = (\sqrt{3}\cos 0, \sqrt{3}\sin 0)$, the answer is $\sqrt{3}\sin(4\pi t)$.

§1.8 #**2.** Circular motion of the form $x = c\cos(\omega t+\phi)$, $y = c\sin(\omega t+\phi)$ can be written in terms of $z = x+iy$ as

$$z = c(\sin(\omega t+\phi) + i\cos(\omega t+\phi)) = ce^{i(\omega t+\phi)}.$$

Here, c is interpreted as the radius of the circular motion, ω is the angular velocity, and ϕ determines the starting phase.

§1.9 #**1.** Since $\Delta > 0$, the functions (1.9.3) are real and linearly independent. Since equation (1.9.1) is linear, we can check independently that the functions $e^{(-\mu+\sqrt{\Delta})t/2m}$ and $e^{(-\mu-\sqrt{\Delta})t/2m}$ are solutions. We'll check the first of these functions, as the second is essentially the same calculation. We have $\dot{y} = (-\mu+\sqrt{\Delta})y/2m$ and $\ddot{y} = (-\mu+\sqrt{\Delta})^2y/4m^2$. So

$$\begin{aligned}
m\ddot{y} + \mu\dot{y} + ky &= \{(-\mu+\sqrt{\Delta})^2/4m + \mu(-\mu+\sqrt{\Delta})/2m + k\}y \\
&= \{\mu^2/4m - \mu\sqrt{\Delta}/2m + \Delta/4m - \mu^2/2m + \mu\sqrt{\Delta}/2m + k\}y.
\end{aligned}$$

Using the fact that $\Delta = \mu^2 - 4mk$, all the terms cancel out to give zero, as required.

§2.2 #**2.** (i) Yes, period 8π. Four and five times the fundamental are present.

(ii) No. If τ is a period of $f(\theta) = \sin\theta + \sin\sqrt{2}\,\theta$ then τ is also a period of $f''(\theta) = -\sin\theta - 2\sin\sqrt{2}\,\theta$. So τ is also a period of $-f(\theta) - f''(\theta) = \sin\sqrt{2}\,\theta$ and of $2f(\theta) + f''(\theta) = \sin\theta$. So τ is a multiple of $\sqrt{2}\,\pi$ and also a multiple of 2π. This cannot happen, because $2\pi/\sqrt{2}\,\pi = \sqrt{2}$ is irrational.

(iii) Yes, period π. The identity $\sin^2\theta = \frac{1}{2}(1 - \cos 2\theta)$ shows that only the fundamental frequency is present, plus a constant offset.

(iv) No, because the intervals on the θ axis between the zeros of the function decrease as $|\theta|$ increases.

(v) Yes, period 2π. The identity $\sin\theta + \sin(\theta + \frac{\pi}{3}) = \sqrt{3}\sin(\theta + \frac{\pi}{6})$ shows that only the fundamental frequency is present; see equation (1.7.12).

§2.3 #**1.** We have $\sin(\sin(\theta+\pi)) = \sin(-\sin\theta) = -\sin(\sin\theta)$ and $\sin 2(\theta+\pi) = \sin(2\theta + 2\pi) = \sin 2\theta$. So the function $\sin(\sin\theta)\sin 2\theta$ is half-period antisymmetric. It follows that the integral is zero.

§2.3 #**2.** We have $\tan(-\theta) = -\tan\theta$, so the tangent function is odd, and so $a_m = 0$. We have $\tan(\theta+\pi) = \tan\theta$, so the tangent function is half-period symmetric, and so $b_{2m+1} = 0$. The only coefficients which can be nonzero are the coefficients b_{2m}. The first nonzero coefficient is

$$b_2 = \frac{1}{\pi}\int_0^{2\pi} \sin(2\theta)\tan\theta\,d\theta = \frac{1}{\pi}\int_0^{2\pi} 2\sin^2\theta = 2.$$

§2.4 #**1.** For $x \neq 0$,

$$\frac{dy}{dx} = 2x\sin(1/x^2) - (2/x)\cos(1/x^2),$$

which is unbounded for small values of x. For $x = 0$, we have

$$\frac{dy}{dx} = \lim_{h\to 0}(h^2\sin(1/h^2))/h = \lim_{h\to 0}(h\sin(1/h^2)) = 0$$

since $-h \le h\sin(1/h^2) \le h$.

§2.4 #**2.** The Fourier series is

$$|\sin\theta| = \frac{2}{\pi} - \frac{4}{\pi}\left(\frac{\cos 2\theta}{1\cdot 3} + \frac{\cos 4\theta}{3\cdot 5} + \cdots\right) = \frac{2}{\pi} - \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\cos 2n\theta}{(2n-1)(2n+1)}.$$

§2.4 #**3.** The Fourier series for the sawtooth function defined by $\phi(\theta) = (\pi - \theta)/2$ for $0 < \theta < 2\pi$ and $\phi(0) = \phi(2\pi) = 0$ is

$$\phi(\theta) = \frac{\sin\theta}{1} + \frac{\sin 2\theta}{2} + \frac{\sin 3\theta}{3} + \cdots = \sum_{n=1}^{\infty}\frac{\sin n\theta}{n}.$$

§2.4 #**4.** The Fourier series for the triangular function is

$$f(\theta) = \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\cos(2n+1)\theta}{(2n+1)^2}.$$

§2.4 #**5.** (a) If $|f(\theta)| \le M$ then

$$|a_m| = \frac{1}{\pi}\left|\int_0^{2\pi} f(\theta)\sin m\theta\,d\theta\right| \le \frac{1}{\pi}\int_0^{2\pi}|f(\theta)||\sin m\theta|\,d\theta \le \frac{1}{\pi}\int_0^{2\pi} M\,d\theta = 2M.$$

Similarly $|b_m| \le 2M$.

(b) $a_m(f') = -mb_m(f)$, $b_m(f') = ma_m(f)$.

(c) If $|f^{(k)}(\theta)| \le M$ then by (a), $|a_m(f^{(k)})| \le 2M$ and $|b_m(f^{(k)})| \le 2M$. So by (b), $|a_m(f)| \le 2M/m^k$ and $|b_m(f)| \le 2M/m^k$.

§2.4 #**6.** $a_0 = 2\pi^2/3$, and for $m > 0$, $a_m = 4(-1)^m/m^2$, $b_m = 0$. Since $f(0) = 0$, this gives $\frac{1}{2}(2\pi^2/3) + 4\sum_{m=1}^{\infty}(-1)^m/m^2 = 0$, or $\sum_{m=1}^{\infty}(-1)^m/m^2 = -\pi^2/12$. Since $f(\pi) = \pi^2$, we obtain $\frac{1}{2}(2\pi^2/3) + 4\sum_{m=1}^{\infty} 1/m^2 = \pi^2$, or $\sum_{m=1}^{\infty} 1/m^2 = \pi^2/6$.

§2.5 #**1.** We have $\frac{\sin\theta}{\theta} = 1 - \frac{1}{3!}\theta^2 + \frac{1}{5!}\theta^4 - \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n\theta^{2n}}{(2n+1)!}$. Since the series is absolutely convergent, we may integrate term by term to get the given power series formula for the integral. Putting in $x = \pi$ gives $\int_0^{\pi} \frac{\sin\theta}{\theta}\, d\theta = \pi - \frac{1}{3.3!}\pi^3 + \frac{1}{5.5!}\pi^5 - \cdots \approx 1.8519370$.

§2.6 #**1.** The square wave takes value one between $\theta = 0$ and $\theta = \pi$, and minus one between $\theta = \pi$ and $\theta = 2\pi$. So $\alpha_m = \frac{1}{2\pi}\left(\int_0^{\pi} e^{-im\theta}\, d\theta - \int_{\pi}^{2\pi} e^{-im\theta}\, d\theta\right) = \frac{1}{2\pi}\left(\left[\frac{-1}{im}e^{-im\theta}\right]_0^{\pi} - \left[\frac{-1}{im}e^{-im\theta}\right]_{\pi}^{2\pi}\right) = \frac{1}{2\pi}\frac{-1}{im}(((-1)^m - 1) - (1 - (-1)^m))$. If m is even, the terms in the parenthesis cancel to zero, whereas if m is odd, they add up to -4.

§2.7 #**1.** We can't use θ for the variable in both (2.6.2) and (2.7.1), so we use x instead in (2.6.2). This gives $s_m(\theta) = \sum_{n=-m}^{m}\left(\frac{1}{2\pi}\int_0^{2\pi} e^{-inx} f(x)\, dx\right) e^{in\theta} = \frac{1}{2\pi}\int_0^{2\pi} f(x)\left(\sum_{n=-m}^{m} e^{in(\theta-x)}\right) dx = \frac{1}{2\pi}\int_0^{2\pi} f(x) D_m(\theta - x)\, dx$.

§2.8 #**1.** $\sin(z\cos\theta) = 2\sum_{n=0}^{\infty}(-1)^n J_{2n+1}(z)\cos(2n+1)\theta$,
$\cos(z\cos\theta) = J_0(z) + 2\sum_{n=1}^{\infty}(-1)^n J_{2n}(z)\cos 2n\theta$.

§2.8 #**2.** Differentiate equation (2.8.9) with respect to ϕ, keeping z and θ constant.

§2.9 #**1.** Using equation (2.9.6), we have

$$\int_0^{\infty} J_1(z)\, dz = [-J_0(z)]_0^{\infty} = -\lim_{z\to\infty} J_0(z) + J_0(0) = 1.$$

§2.10 #**1.** If $y = J_n(\alpha x)$ then using equation (2.10.1) we have

$$\begin{aligned}\frac{dy}{dx} &= \alpha J_n'(\alpha x)\\ \frac{d^2y}{dx^2} &= \alpha^2 J_n''(\alpha x) = -\alpha^2\left(\frac{1}{\alpha x}J_n'(\alpha x) + \left(1 - \frac{n^2}{\alpha^2x^2}\right)J_n(\alpha x)\right)\\ &= -\frac{1}{x}\frac{dy}{dx} - \left(\alpha^2 - \frac{n^2}{x^2}\right)y.\end{aligned}$$

Since $Y_n(z)$ also satisfies equation (2.10.1), the same argument shows that $Y_n(\alpha x)$ is a solution of the given differential equation. Since the equation is linear, $y = AJ_n(\alpha x) + BY_n(\alpha x)$ is again a solution. The general theory of second order linear differential equations implies that the space of solutions is two dimensional, so we have found them all. Alternatively, we could argue that if $f(x)$ is any solution then $f(z/\alpha)$ has to be a solution of (2.10.1).

§2.10 #**2.** If $y = x^{\frac{1}{2}}J_n(x)$ then

$$\begin{aligned}\frac{dy}{dx} &= \tfrac{1}{2}x^{-\frac{1}{2}}J_n(x) + x^{\frac{1}{2}}J_n'(x)\\ \frac{d^2y}{dx^2} &= -\tfrac{1}{4}x^{-\frac{3}{2}}J_n(x) + x^{-\frac{1}{2}}J_n'(x) + x^{\frac{1}{2}}J_n''(x)\end{aligned}$$

$$= -\tfrac{1}{4}x^{-\frac{3}{2}}J_n(x) + x^{-\frac{1}{2}}J_n'(x) - x^{\frac{1}{2}}\left(\frac{1}{x}J_n'(x) + \left(1 - \frac{n^2}{x^2}\right)J_n(x)\right)$$
$$= -\left(1 + \frac{\frac{1}{4} - n^2}{x^2}\right)x^{\frac{1}{2}}J_n(x) = -\left(1 + \frac{\frac{1}{4} - n^2}{x^2}\right)y$$

and so y satisfies the given differential equation. The general solution is

$$y = \sqrt{x}\left(AJ_n(x) + BY_n(x)\right).$$

§2.10 #**3.** If $y = J_n(e^x)$ then

$$\begin{aligned}\frac{dy}{dx} &= e^x J_n'(e^x)\\ \frac{d^2y}{dx^2} &= e^{2x}J_n''(e^x) + e^x J_n'(e^x)\\ &= -e^{2x}\left(\frac{1}{e^x}J_n'(e^x) + \left(1 - \frac{n^2}{e^{2x}}\right)J_n(e^x)\right) + e^x J_n'(e^x)\\ &= -(e^{2x} - n^2)J_n(e^x) = -(e^{2x} - n^2)y\end{aligned}$$

and so y satisfies the given differential equation. The general solution is

$$y = AJ_n(e^x) + BY_n(e^x).$$

§2.10 #**4.** (a) $-\sin^2\theta\sin(\phi + z\sin\theta) = \displaystyle\sum_{n=-\infty}^{\infty} J_n''(z)\sin(\phi + n\theta)$.

(b) $-z\sin\theta\cos(\phi + z\sin\theta) - z^2\cos^2\theta\sin(\phi + z\sin\theta) = -\displaystyle\sum_{n=-\infty}^{\infty} n^2 J_n(z)\sin(\phi + n\theta)$.

§2.11 #**1.** We have

$$\begin{aligned}\frac{\partial\phi}{\partial z} &= \left(\phi + z\frac{\partial\phi}{\partial z}\right)\cos(\omega t + z\phi)\\ \frac{\partial\phi}{\partial t} &= \left(\omega + z\frac{\partial\phi}{\partial t}\right)\cos(\omega t + z\phi),\end{aligned}$$

and so

$$\left(\phi + z\frac{\partial\phi}{\partial z}\right)\frac{\partial\phi}{\partial t} = \left(\omega + z\frac{\partial\phi}{\partial t}\right)\frac{\partial\phi}{\partial z}.$$

This gives the partial differential equation for ϕ. If ϕ is replaced by ϕ/α and z is replaced by αz, then both sides of the differential equation get divided by α^2, and so the equation remains valid. The equations for ϕ become $\phi = \alpha\sin(\omega t + z\phi)$ and

$$\phi = \sum_{n=1}^{\infty}\frac{2J_n(n\alpha z)}{nz}\sin(n\omega t).$$

§2.13 #**1.** Set $I = \int_{-\infty}^{\infty} e^{-x^2}\,dx$. Then squaring and converting to polar coordinates gives

$$\begin{aligned}I^2 &= \left(\int_{-\infty}^{\infty} e^{-x^2}\,dx\right)\left(\int_{-\infty}^{\infty} e^{-y^2}\,dy\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2-y^2}\,dx\,dy\\ &= \int_0^{2\pi}\int_0^{\infty} e^{-r^2}\,r\,dr\,d\theta = 2\pi\int_0^{\infty} re^{-r^2}dr = 2\pi\left[-e^{-r^2}\right]_0^{\infty} = 2\pi.\end{aligned}$$

Since the integrand is positive, taking square roots gives $I = \sqrt{2\pi}$.

§2.13 #**4.** Substitute $\tau = t - a$ to get

$$\int_{-\infty}^{\infty} f(t-a)e^{-2\pi i\nu t}\,dt = \int_{-\infty}^{\infty} f(\tau)e^{-2\pi i\nu(\tau+a)}\,d\tau$$
$$= e^{-2\pi i\nu a}\int_{-\infty}^{\infty} f(\tau)e^{-2\pi i\nu\tau}\,d\tau = e^{-2\pi i\nu a}\hat{f}(\nu).$$

§2.13 #**5.** Using equation (C.3), we have

$$\int_{-\rho/2}^{\rho/2} e^{-2\pi i\nu t}\,dt = \left[-\frac{1}{2\pi i\nu}e^{-2\pi i\nu t}\right]_{-\rho/2}^{\rho/2} = \frac{e^{\pi i/nu\rho} - e^{-\pi i\nu\rho}}{2\pi i\nu} = \frac{\sin \pi\nu\rho}{\pi\nu}.$$

§2.17 #**1.** Using equation (C.3), we have $f(t) = \sin(2\pi\nu_0 t) = \frac{1}{2i}(e^{2\pi i\nu_0 t} - e^{-2\pi i\nu_0 t})$, and so

$$\hat{f}(\nu) = \frac{1}{2i}(\delta(\nu - \nu_0) - \delta(\nu + \nu_0)).$$

§2.17 #**2.** Given any test function $f(t)$, substituting $u = Ct$ gives

$$\int_{-\infty}^{\infty} f(t)\delta(Ct)\,dt = \frac{1}{|C|}\int_{-\infty}^{\infty} f(u/C)\delta(u)\,du = \frac{1}{|C|}f(0).$$

It follows that the values of the distributions $\delta(Ct)$ and $\frac{1}{|C|}\delta(t)$ agree on all test functions, and so they are equal as distributions. Note that if C is negative, the above substitution involves reversing the limits on the integral and negating.

§2.17 #**3.** Given any test function $f(t)$, integrating by parts gives

$$\int_{-\infty}^{\infty} \frac{dH(t)}{dt}f(t)\,dt = -\int_{-\infty}^{\infty} H(t)f'(t)\,dt = -\int_{0}^{\infty} f'(t)\,dt = [-f(t)]_0^{\infty} = f(0).$$

It follows that the values of the distributions $\frac{d}{dt}H(t)$ and $\delta(t)$ agree on all test functions, and so they are equal as distributions.

§2.17 #**4.** For any test function $f(t)$, $\int_{-\infty}^{\infty} t\delta(t)f(t)\,dt$ is the value of $tf(t)$ when $t = 0$, which always gives zero.

§3.2 #**1.** If the cross-sectional area is A then the tension is $T \approx 1.1 \times 10^9 A$ Newtons and the linear density is $\rho \approx 5900A$ kg/m. So the speed is $c = \sqrt{T/\rho} \approx 432$ m/s, and is independent of A. For a frequency of 262 Hz, the length would be given by $262 = c/2\ell$, or $\ell = c/524 \approx 0.824$ meters.

§3.2 #**2.** The square root of the tension should be increased by a factor of $3/2$, so the tension should be increased by a factor of $9/4$.

§3.2 #**3.** According to Mersenne's laws, the frequency is inversely proportional to the length of the string. Since the frequencies of the notes on a scale increase exponentially, the lengths of the strings decrease exponentially. Each octave halves the string length.

§3.6 #**1.** If we make the square from the interval $[0, a]$ on both the x and y axes, then the solutions to the wave equation are combinations of the functions

$$y = \sin\frac{m\pi}{a}x\,\sin\frac{n\pi}{a}y\,\sin(\omega t + \phi)$$

where

$$\omega = \frac{\pi c}{a}\sqrt{m^2 + n^2}$$

and m and n are positive integers.

§3.9. The answer to the challenge in the footnote on page 117 is that the series continues as follows. Set $z = (-1)^n e^{-(n+\frac{1}{2})\pi}$. Then

$$\begin{aligned}\lambda_n \approx (n+\tfrac{1}{2})\pi - z - 4z^2 - \tfrac{34}{3}z^3 - \tfrac{112}{3}z^4 - \tfrac{2006}{15}z^5 - \tfrac{1516}{3}z^6 - \tfrac{124834}{63}z^7 - \tfrac{502976}{63}z^8 \\ - \tfrac{2069150}{63}z^9 - \tfrac{389388268}{2835}z^{10} - \tfrac{518637298}{891}z^{11} - \tfrac{1728425360}{693}z^{12} - \tfrac{2623624535150}{243243}z^{13} \\ - \tfrac{879673454236}{18711}z^{14} - \tfrac{5004230870978}{24255}z^{15} - \tfrac{357875952715520}{392931}z^{16} - \tfrac{26997237726639718}{6679827}z^{17} \\ - \tfrac{12486057159188}{693}z^{18} - \tfrac{5419093013311552886}{67191201}z^{19} - \tfrac{1217363076852549595 04}{335956005}z^{20} - \cdots\end{aligned}$$

The corresponding series for the mbira in §3.10 is the same, but with $n+\frac{1}{2}$ replaced by $n-\frac{1}{2}$ in both the definition of z and in the first term of the formula for λ_n.

§5.3 #**1.** (a) G$\flat\flat$, (b) D$\flat\flat\flat$, (c) G$\sharp\sharp\sharp\sharp$ or G**xx**.

§5.4 #**1.** The Pythagorean comma, in cents, is $1200\ln(3^{12}/2^{19})/\ln(2)$, which works out to the figure of roughly 23.460 cents given in the text. In Savarts, we get $1000\log_{10}(3^{12}/2^{19})$ or roughly 5.8851.

§5.4 #**2.** To the nearest cent, the vibrational modes of the drum are as given in the following table, with respect to the lowest mode.

0	806	1313	1689	1989
1438	1854	2169	2425	2642
2217	2497	2727	2923	3095

§5.4 #**3.** E$\flat\flat \approx 180.450$ cents.

§5.8 #**1.** $1200\ln(81/80)/\ln 2 \approx 21.506$ cents;
$1200\ln(32805/32768)/\ln 2 \approx 1.953$ cents.

§5.10 #**1.** Here are some of the notes appearing in these scales, and their values in cents:

C^{0}, 0.000. $\mathrm{C}\sharp^{-2}$, 70.672. $\mathrm{D}\flat^{0}$, 90.225. $\mathrm{C}\sharp^{-1}$, 92.179. $\mathrm{D}\flat^{+1}$, 111.731. D^{-1}, 182.404. D^{0}, 203.910. $\mathrm{D}\sharp^{-2}$, 274.582. $\mathrm{E}\flat^{0}$, 294.135. $\mathrm{D}\sharp^{-1}$, 296.089. $\mathrm{E}\flat^{+1}$, 315.641. E^{-1}, 386.314. F^{0}, 498.045. $\mathrm{F}\sharp^{-2}$, 568.717. $\mathrm{F}\sharp^{-1}$, 590.224. G^{0}, 701.955. $\mathrm{G}\sharp^{-2}$, 772.627. $\mathrm{G}\sharp^{-1}$, 794.134. $\mathrm{A}\flat^{0}$, 792.180. $\mathrm{A}\flat^{+1}$, 813.687. A^{-1}, 884.359. $\mathrm{B}\flat^{0}$, 996.091. $\mathrm{B}\flat^{+1}$, 1017.596. B^{-1}, 1088.269. (C^{0}, 1200.)

§5.10 #**2.** In these triads, the fifths are perfect, and the major thirds are flat by one schisma, or 1.955 cents. This is much closer to just than, for example, the twelve tone equal tempered major triad.

§5.10 #**3.** (i) $\mathrm{C}^{0}-\mathrm{E}^{-1}-\mathrm{G}^{0}$, or many others.

(ii) $\mathrm{C}^{-1}-\mathrm{E}\flat^{0}-\mathrm{G}^{-1}$, or many others.

(iii) Horizontal cross-sections are designed to contain just major scales, for example $\mathrm{C}^{0}-\mathrm{D}^{0}-\mathrm{E}^{-1}-\mathrm{F}^{0}-\mathrm{G}^{0}-\mathrm{A}^{-1}-\mathrm{B}^{-1}-\mathrm{C}^{0}$.

(iv) Each black key is a syntonic comma lower than the white key above it, for example C^{-1} to C^{0}.

(v) C^{-1} to $\mathrm{B}\sharp^{-2}$, $\mathrm{E}\flat^{0}$ to $\mathrm{D}\sharp^{-1}$, and F^{-1} to $\mathrm{E}\sharp^{-2}$ are examples of pairs of notes on the diagram, differing by a schisma.

(vi) From a white note near the top of the keyboard, go to the right one column and down past the black note to the next white note to obtain a note one diesis higher. For example C^{0} to $\mathrm{D}\flat^{0}$ or E^{0} to F^{0}.

(vii) Each key is one apotomē higher than the corresponding key in the same position two notes lower down on the keyboard. For example C^{-1} to $\mathrm{C}\sharp^{-1}$ is an apotomē.

§5.12 #**2.** If we use α commas, then the fifth will be out by α commas, the major third by $4\alpha - 1$ commas, and the minor third by $3\alpha - 1$ commas. The total square deviation is then

$$\alpha^2 + (4\alpha - 1)^2 + (3\alpha - 1)^2.$$

Setting the derivative with respect to α of this expression equal to zero gives $\alpha = \frac{7}{26}$. The root mean square deviation for a $\frac{7}{26}$ comma meantone scale is $1/\sqrt{26}$ of a comma, or 4.218 cents. This compares with $1/\sqrt{24}$ of a comma, or 4.390 cents for the quarter tone meantone scale. This represents an improvement of about four percent.

If we make the fifth and major third three times as important as the minor third, then the quarter tone meantone scale exactly minimizes the mean square deviation. If we make the minor third twice as important as the fifth and major third, Zarlino's $\frac{2}{7}$-comma meantone scale minimizes the mean square deviation.

§5.12 #**4.** The tempering in this scale is by

$$\log_2(3/2) - (\tfrac{1}{2} + \tfrac{1}{4\pi})$$

of an octave, which works out at about 6.462 cents, or about 0.30047 commas.

§5.12 #**5.** The major thirds are just, and the minor thirds are narrow by one sixth of a comma. Thus the important intervals of octave, fifth, major and minor third, are all within one sixth of a comma, or 3.584 cents of the just values. The major scale for this temperament is given in cents as follows:

C^{0}, 0.000; $\mathrm{D}^{-\frac{1}{2}}$, 193.157; E^{-1}, 386.314; $\mathrm{F}^{+\frac{1}{3}}$, 505.214; $\mathrm{G}^{-\frac{1}{6}}$, 698.371; $\mathrm{A}^{-\frac{2}{3}}$, 891.527; $\mathrm{B}^{-\frac{7}{6}}$, 1084.684; $\mathrm{C}^{+\frac{1}{6}}$, 1203.584.

§5.13 #**1.** Here is a table of some of the scales discussed in this section, in cents to three decimal places, and also in Eitz's comma notation. The symbol p denotes the Pythagorean comma, which is almost exactly equal to 12/11 of a syntonic comma.

		Werckmeister III		Werckmeister IV		Werckmeister V		Vallotti–Young	
do	C	0.000	0	0.000	0	0.000	0	0.000	0
	C♯	90.225	$-1p$	82.405	$-\frac{4}{3}p$	96.090	$-\frac{3}{4}p$	90.225	$-1p$
re	D	192.180	$-\frac{1}{2}p$	196.090	$-\frac{1}{3}p$	203.910	0	196.090	$-\frac{1}{3}p$
	E♭	294.135	0	294.135	0	300.000	$+\frac{1}{4}p$	294.135	0
mi	E	390.225	$-\frac{3}{4}p$	392.180	$-\frac{2}{3}p$	396.090	$-\frac{1}{2}p$	392.180	$-\frac{2}{3}p$
fa	F	498.045	0	498.045	0	503.910	$+\frac{1}{4}p$	498.045	0
	F♯	588.270	$-1p$	588.270	$-1p$	600.000	$-\frac{1}{2}p$	588.270	$-1p$
so	G	696.090	$-\frac{1}{4}p$	694.135	$-\frac{1}{3}p$	701.955	0	698.045	$-\frac{1}{6}p$
	G♯	792.180	$-1p$	784.360	$-\frac{4}{3}p$	792.180	$-1p$	792.180	$-1p$
la	A	888.270	$-\frac{3}{4}p$	890.225	$-\frac{2}{3}p$	900.000	$-\frac{1}{4}p$	894.135	$-\frac{1}{2}p$
	B♭	996.090	0	1003.910	$+\frac{1}{3}p$	1001.955	$+\frac{1}{4}p$	996.090	0
ti	B	1092.180	$-\frac{3}{4}p$	1086.315	$-1p$	1098.045	$-\frac{1}{2}p$	1090.225	$-\frac{5}{6}p$
do	C	1200.000	0	1200.000	0	1200.000	0	1200.000	0

§5.13 #**2.** I use a Roland Sound Canvas SCC-1 card with my computer. Here are some system exclusives for the SCC-1 for various temperaments. These should also work with other versions of the Sound Canvas.

Just intonation in C:
F0 41 10 42 12 40 11 40 40 23 44 50 32 3E 36 42 25 30 52 34 35 F7

Just intonation in D:
F0 41 10 42 12 40 11 40 52 34 40 23 44 50 32 3E 36 42 25 30 35 F7

Meantone (with G♯):
F0 41 10 42 12 40 11 40 40 28 39 4A 32 43 2B 3D 25 36 47 2F 56 F7

Meantone (with A♭):
F0 41 10 42 12 40 11 40 40 28 39 4A 32 43 2B 3D 4E 36 47 2F 2D F7

Werckmeister III:
F0 41 10 42 12 40 11 40 40 36 38 3A 36 3E 34 3C 38 34 3C 38 43 F7

§5.14 #**1.** The approximation of Kirnberger and Farey is $\left(\frac{81}{80}\right)^{\frac{12}{11}} \approx \frac{531411}{524288}$, or $\left(\frac{3^4}{2^4.5}\right)^{\frac{12}{11}} \approx \frac{3^{12}}{2^{19}}$. Taking eleventh powers gives $\left(\frac{3^4}{2^4.5}\right)^{12} \approx \left(\frac{3^{12}}{2^{19}}\right)^{11}$, which can be written as $\frac{3^{48}}{2^{48}.5^{12}} \approx \frac{3^{132}}{2^{209}}$. Cross multiplying and cancelling gives $2^{161} \approx 3^{84}.5^{12}$.

§5.14 #**2.** A good spectrum to use for twelve tone equal temperament consists of the following multiples of the fundamental frequency:

1:1, 2:1, $2^{\frac{19}{12}}$:1, 4:1, $2^{\frac{7}{3}}$:1, $2^{\frac{31}{12}}$:1, $2^{\frac{17}{6}}$:1, 8:1.

These approximate the first eight harmonics in such a way as to make the equal tempered major thirds (C–E) and the equal tempered approximation to the seventh harmonic (C–B♭) consonant.

§5.14 #**4.** Here is a table of the Pythagorean, just, meantone and equal scales, in cents to three decimal places, and also in Eitz's comma notation. The symbol p denotes the Pythagorean comma, which is almost exactly equal to 12/11 of a syntonic comma.

		Pythagorean		Just		Meantone		Equal	
do	C	0.000	0	0.000	0	0.000	0	0.000	0
	C♯	113.685	0	70.672	-2	76.049	$-\frac{7}{4}$	100.000	$-\frac{7}{12}p$
re	D	203.910	0	203.910	0	193.157	$-\frac{1}{2}$	200.000	$-\frac{1}{6}p$
	E♭	294.135	0	315.641	$+1$	310.265	$+\frac{3}{4}$	300.000	$+\frac{1}{4}p$
mi	E	407.820	0	386.314	-1	386.314	-1	400.000	$-\frac{1}{3}p$
fa	F	498.045	0	498.045	0	503.422	$+\frac{1}{4}$	500.000	$+\frac{1}{12}p$
	F♯	611.730	0	590.224	-1	579.471	$-\frac{3}{2}$	600.000	$-\frac{1}{2}p$
so	G	701.955	0	701.955	0	696.579	$-\frac{1}{4}$	700.000	$-\frac{1}{12}p$
	G♯	815.640	0	772.627	-2	772.627	-2	800.000	$-\frac{2}{3}p$
	A♭	792.180	0	— —	–	813.686	$+1$	800.000	$+\frac{1}{3}p$
la	A	905.865	0	884.359	-1	889.735	$-\frac{3}{4}$	900.000	$-\frac{1}{4}p$
	B♭	996.090	0	1017.596	$+1$	1006.843	$+\frac{1}{2}$	1000.000	$+\frac{1}{6}p$
ti	B	1109.775	0	1088.269	-1	1082.892	$-\frac{5}{4}$	1100.000	$-\frac{5}{12}p$
do	C	1200.000	0	1200.000	0	1200.000	0	1200.000	0

§5.14 #**5.** In Cordier's equal temperament, every semitone is exactly one seventh of a perfect fifth, or a frequency ratio of $\left(\frac{3}{2}\right)^{\frac{1}{7}}$. So twelve such semitones give a stretched octave with frequency ratio of $\left(\frac{3}{2}\right)^{\frac{12}{7}}$. Seven such stretched octaves give a frequency ratio of $\left(\frac{3}{2}\right)^{12}$, which differs from seven pure octaves by a ratio of $\left(\frac{3}{2}\right)^{12}/2^7 = 3^{12}/2^{19}$, or one Pythagorean comma. So one octave is stretched by $\frac{1}{7}$ of a Pythagorean comma.

In Eitz's notation, this comes out as follows:

$$\begin{array}{cccccccccc}
\mathrm{C}^{-\frac{1}{7}p} & & \mathrm{G}^{-\frac{1}{7}p} & & \mathrm{D}^{-\frac{1}{7}p} & & \mathrm{A}^{-\frac{1}{7}p} & & \mathrm{E}^{-\frac{1}{7}p} & \\
& \mathrm{E}\flat^{+\frac{1}{7}p} & & \mathrm{B}\flat^{+\frac{1}{7}p} & & \mathrm{F}^{+\frac{1}{7}p} & & \mathrm{C}^{+\frac{1}{7}p} & & \mathrm{G}^{+\frac{1}{7}p} \\
\mathrm{B}^{-\frac{4}{7}p} & & \mathrm{F}\sharp^{-\frac{4}{7}p} & & \mathrm{D}\flat^{+\frac{3}{7}p} & & \mathrm{A}\flat^{+\frac{3}{7}p} & & \mathrm{E}\flat^{+\frac{3}{7}p} & \\
& \mathrm{D}^{-\frac{2}{7}p} & & \mathrm{A}^{-\frac{2}{7}p} & & \mathrm{E}^{-\frac{2}{7}p} & & \mathrm{B}^{-\frac{2}{7}p} & & \mathrm{F}\sharp^{-\frac{2}{7}p} \\
\mathrm{B}\flat^{0} & & \mathrm{F}^{0} & & \mathrm{C}^{0} & & \mathrm{G}^{0} & & \mathrm{D}^{0} & \\
& \mathrm{D}\flat^{+\frac{2}{7}p} & & \mathrm{A}\flat^{+\frac{2}{7}p} & & \mathrm{E}\flat^{+\frac{2}{7}p} & & \mathrm{B}\flat^{+\frac{2}{7}p} & & \mathrm{F}^{+\frac{2}{7}p} \\
\mathrm{A}^{-\frac{3}{7}p} & & \mathrm{E}^{-\frac{3}{7}p} & & \mathrm{B}^{-\frac{3}{7}p} & & \mathrm{F}\sharp^{-\frac{3}{7}p} & & \mathrm{D}\flat^{+\frac{4}{7}p} & \\
& \mathrm{C}^{-\frac{1}{7}p} & & \mathrm{G}^{-\frac{1}{7}p} & & \mathrm{D}^{-\frac{1}{7}p} & & \mathrm{A}^{-\frac{1}{7}p} & & \mathrm{E}^{-\frac{1}{7}p}
\end{array}$$

The top and bottom rows are identified to form a horizontal cylinder. Three major thirds, going diagonally upwards and to the right three spaces, correspond to the octave stretched by $\frac{1}{7}$ of a Pythagorean comma. Four minor thirds, going diagonally downwards and to the right four places, have the same effect.

The major thirds in this temperament are sharp by one syntonic comma minus $\frac{2}{7}$ of a Pythagorean comma, or 14.803 cents. This is very slightly worse than the already badly sharp major thirds of the usual equal temperament. The minor thirds are flat by the same amount, which is slightly better than in equal temperament.

§6.1 #**1.** The Indian Sruti scale comes out as

$$\begin{array}{cccccccccccccc}
& & & & \mathrm{D}^{-1} & & \mathrm{A}^{-1} & \mathrm{E}^{-1} & & \mathrm{B}^{-1} & \mathrm{F}\sharp^{-1} & & & \\
\mathrm{D}\flat^{0} & \mathrm{A}\flat^{0} & \mathrm{E}\flat^{0} & \mathrm{B}\flat^{0} & & \mathrm{F}^{0} & \mathrm{C}^{0} & & \mathrm{G}^{0} & \mathrm{D}^{0} & & \mathrm{A}^{0} & \mathrm{E}^{0} & \mathrm{B}^{0} \quad \mathrm{F}\sharp^{0} \\
& & & & \mathrm{D}\flat^{+1} & & \mathrm{A}\flat^{+1} & \mathrm{E}\flat^{+1} & & \mathrm{B}\flat^{+1} & \mathrm{F}^{+1} & & &
\end{array}$$

§6.2 #**6.** The continued fraction expansion for the frequency ratio which represents the Pythagorean comma is

$$\frac{531441}{524288} = 1 + \frac{1}{73+}\,\frac{1}{3+}\,\frac{1}{2+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{23+}\,\frac{1}{2+}\,\frac{1}{5}.$$

This corresponds to the following application of Euclid's algorithm to obtain 1 as the highest common factor of the numerator and denominator:

$$\begin{aligned}
531441 - 1 \times 524288 &= 7153 \\
524288 - 73 \times 7153 &= 2119 \\
7153 - 3 \times 2119 &= 796 \\
2119 - 2 \times 796 &= 527 \\
796 - 1 \times 527 &= 269 \\
527 - 1 \times 269 &= 258 \\
269 - 1 \times 258 &= 11 \\
258 - 23 \times 11 &= 5 \\
11 - 2 \times 5 &= 1 \\
[\,5 - 5 \times 1 &= 0\,]
\end{aligned}$$

The numbers $a_0, a_1, a_2, \ldots$ appear as the multiples to subtract in the application of Euclid's algorithm. This happens whether or not the fraction is in reduced form.

§6.2 #**8.** The fraction is 113/821.

§6.4 #**1.** The 31 tone scale amounts to identifying $\mathrm{F}\flat\flat^{+\frac{15}{4}}$ with $\mathrm{D}\sharp\sharp^{-4}$ in the extended meantone scale. The difference is 6.069 cents, so divided by 31, this makes each step out by 0.196 cents from the meantone equivalent. Here is the torus of thirds and fifths:

$$\mathrm{F}\flat\flat^{+\frac{15}{4}}\quad \mathrm{C}\flat\flat^{+\frac{7}{2}}\quad \mathrm{G}\flat\flat^{+\frac{13}{4}}\quad \mathrm{D}\flat\flat^{+3}$$
$$\mathrm{B}\sharp^{-3}\quad \mathrm{F}\sharp\sharp^{-\frac{13}{4}}\quad \mathrm{C}\sharp\sharp^{-\frac{7}{2}}\quad \mathrm{G}\sharp\sharp^{-\frac{15}{4}}\quad \mathrm{F}\flat\flat^{+\frac{15}{4}}$$
$$\mathrm{G}\sharp^{-2}\quad \mathrm{D}\sharp^{-\frac{9}{4}}\quad \mathrm{A}\sharp^{-\frac{5}{2}}\quad \mathrm{E}\sharp^{-\frac{11}{4}}\quad \mathrm{B}\sharp^{-3}$$
$$\mathrm{E}^{-1}\quad \mathrm{B}^{-\frac{5}{4}}\quad \mathrm{F}\sharp^{-\frac{3}{2}}\quad \mathrm{C}\sharp^{-\frac{7}{4}}\quad \mathrm{G}\sharp^{-2}$$
$$\mathrm{C}^{0}\quad \mathrm{G}^{-\frac{1}{4}}\quad \mathrm{D}^{-\frac{1}{2}}\quad \mathrm{A}^{-\frac{3}{4}}\quad \mathrm{E}^{-1}$$
$$\mathrm{A}\flat^{+1}\quad \mathrm{E}\flat^{+\frac{3}{4}}\quad \mathrm{B}\flat^{+\frac{1}{2}}\quad \mathrm{F}^{+\frac{1}{4}}\quad \mathrm{C}^{0}$$
$$\mathrm{F}\flat^{+2}\quad \mathrm{C}\flat^{+\frac{7}{4}}\quad \mathrm{G}\flat^{+\frac{3}{2}}\quad \mathrm{D}\flat^{+\frac{5}{4}}\quad \mathrm{A}\flat^{+1}$$
$$\mathrm{D}\flat\flat^{+3}\quad \mathrm{A}\flat\flat^{+\frac{11}{4}}\quad \mathrm{E}\flat\flat^{+\frac{5}{2}}\quad \mathrm{B}\flat\flat^{+\frac{9}{4}}\quad \mathrm{F}\flat^{+2}$$
$$\mathrm{F}\flat\flat^{+\frac{15}{4}}\quad \mathrm{C}\flat\flat^{+\frac{7}{2}}\quad \mathrm{G}\flat\flat^{+\frac{13}{4}}\quad \mathrm{D}\flat\flat^{+3}$$

§6.4 #**2.**

note	$\frac{1}{3}$-comma	19 tone		$\frac{1}{5}$-comma	43 tone	
C	0.000	0	0.000	0.000	0	0.000
D	189.572	3	189.474	195.307	7	195.349
E	379.145	6	378.947	390.615	14	390.698
F	505.214	8	505.263	502.346	18	502.326
G	694.786	11	694.737	697.654	25	697.674
A	884.359	14	884.211	892.961	32	893.023
B	1073.931	17	1073.684	1088.269	39	1088.372
C	1200.000	19	1200.000	1200.000	43	1200.000

note	$\frac{2}{7}$-comma	50 tone		$\frac{1}{6}$-comma	55 tone	
C	0.000	0	0.000	0.000	0	0.000
D	191.621	8	192.000	196.741	9	196.364
E	383.241	16	384.000	393.482	18	392.727
F	504.190	21	504.000	501.629	23	501.818
G	695.810	29	696.000	698.371	32	698.182
A	887.431	37	888.000	895.112	41	894.545
B	1079.052	45	1080.000	1091.853	50	1090.909
C	1200.000	50	1200.000	1200.000	55	1200.000

Comparing the $\frac{1}{3}$-comma meantone with 19 tone equal temperament, the fifths differ by 0.0493955 cents, or about 1/24294 of an octave. This is about 67.296 times as good as what is guaranteed by Theorem 6.2.3. This explains the second line of the following table. For comparison, quarter comma meantone is compared with 31 tone equal temperament, and $\frac{1}{11}$-comma meantone with 12 tone equal temperament.

commas	tones	cents	octaves	factor
$\frac{1}{11}$	12	0.000116436	1/10306055	71570
$\frac{1}{3}$	19	0.0493955	1/24294	67.296
$\frac{1}{4}$	31	0.1957651	1/6130	6.379
$\frac{1}{5}$	43	0.0206757	1/58039	31.389
$\frac{2}{7}$	50	0.1896534	1/6327	2.531
$\frac{1}{6}$	55	0.1880102	1/6356	2.101

It can be seen from this table that 12 tone equal temperament is a fantistically good approximation to $\frac{1}{11}$-comma meantone, while 19 tone equal temperament is a pretty good approximation to $\frac{1}{3}$-comma meantone. The 50 and 55 tone approximations come out worst in this comparison.

§6.7 #**1.** Scale degree 5 (243.8 cents) approximates the ratio 15/13 (247.7 cents), 7 (341.4 cents) approximates 11/9 (347.4 cents), 11 (536.5 cents) approximates 15/11 (536.9 cents), 13 (634.0 cents) approximates 13/9 (636.6 cents), 16 (780.3 cents) approximates 11/7 (782.5 cents), 22 (1072.9 cents) approximates 13/7 (1071.7 cents), 28 (1365.5 cents) approximates 11/5 (1365.0 cents), and 34 (1658.2 cents) approximates 13/5 (1654.2 cents).

§7.8 #**1.** (a) We have

$$\frac{z^2}{z^2+z+\frac{1}{2}} = \frac{1}{1+z^{-1}+\frac{1}{2}z^{-2}},$$

and so this transfer function can be written in the form

$$G(z) = F(z) - z^{-1}G(z) - \tfrac{1}{2}z^{-2}G(z).$$

(b) $(\frac{9}{4} + 3\cos 2\pi\nu/N + \cos 4\pi\nu/N)^{-\frac{1}{2}}$

(c) The poles of the transfer function are at $z = (-1 \pm i)/2$, which are inside the unit circle, so the filter is stable.

§8.8 #**1.** Working to five decimal places,

$$\begin{aligned}\sin(440(2\pi t) + \tfrac{1}{10}\sin 660(2\pi t)) &= 0.04994\sin 220(2\pi t) + 0.99750\sin 440(2\pi t)\\ &\quad - 0.00125\sin 880(2\pi t) + 0.04994\sin 1100(2\pi t)\\ &\quad + 0.00002\sin 1540(2\pi t) + 0.00125\sin 1760(2\pi t)\\ &\quad + 0.00002\sin 2420(2\pi t) + \dots\end{aligned}$$

§8.16 #**1.** Differentiate the equation $T_n(\cos t) - \cos nt = 0$ using the chain rule to get

$$-(\sin t)\,T_n'(\cos t) + n\sin nt = 0$$

and again to get

$$(\sin^2 t)\,T_n''(\cos t) - (\cos t)\,T_n'(\cos t) + n^2\cos nt = 0.$$

Now substitute $x = \cos t$, $y = T_n(x) = \cos nt$, and $1 - x^2 = \sin^2 t$.

§8.16 #**2.** De Moivre's theorem says that

$$\cos nt + i\sin nt = (\cos t + i\sin t)^n.$$

Expanding out the right hand side using the binomial theorem, we obtain

$$\begin{aligned}\cos nt + i\sin nt &= \cos^n t + in\cos^{n-1} t\,\sin t + i^2\binom{n}{2}\cos^{n-2} t\,\sin^2 t\\ &\quad + i^3\binom{n}{3}\cos^{n-3} t\,\sin^3 t + i^4\binom{n}{4}\cos^{n-4} t\,\sin^4 t + \cdots\end{aligned}$$

Taking real parts picks out every other term on the right,

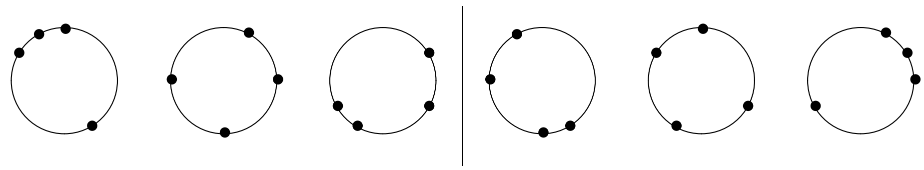
$$\cos nt = \cos^n t - \binom{n}{2}\cos^{n-2} t\,\sin^2 t + \binom{n}{4}\cos^{n-4} t\,\sin^4 t - \cdots$$

Now substitute $x = \cos t$, $T_n(x) = \cos nt$ and $1 - x^2 = \sin^2 t$.

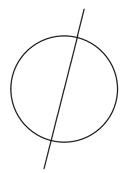
§9.1 #**1.** There is a horizontal axis of exact reflectional symmetry at the note A.

§9.1 #**2.** There is a vertical axis of reflectional symmetry in the barline. There is a horizontal axis of reflectional symmetry so that in the Alto line the pitches are a reflection of the pitches of the Soprano line. The line of symmetry is on the G of the treble clef. The composite of these two symmetries is a rotational symmetry around the middle of the piece. The symmetry in the pitches is exact, but the durations and the words do not display the temporal symmetry.

§9.1 #**3.** Here are the chords in the circle notation.



The second set of three chords has been obtained from the first by temporal reflection followed by a reflection of the chords about a mirror line which passes between C and C♯ and between F♯ and G.



§9.1 #**4.** The frieze pattern here is pm11.

§9.2 #**1.** The sequence transcribes to 1403423120. This can be divided into five pairs 14 03 42 31 20. Each pair is obtained from the previous one by moving one place down the cycle of five strings. Reversing time and the cyclic ordering of the strings, we get 42 31 20 14 03 which is the same sequence, but with a different starting point.

§9.3 #**1.** Write e for the identity element of G. If $(gh)^n = e$ then $(gh)^{n-1}g = h^{-1}$, so $h(gh)^{n-1}g = e$, i.e., $(hg)^n = e$. Using this both ways round, we see that gh and hg must have the same order.

§9.4 #**1.** If n is even, we have

$$ba = (1, 3, 5, \dots, n-3, n-1, n, n-2, n-4, \dots, 6, 4, 2),$$

of order n, so the total number of rows before returning to the beginning is $2n$.

If n is odd, we have

$$ba = (1, 3, 5, \dots, n-2, n, n-1, n-3, n-5, \dots, 6, 4, 2),$$

again of order n, so the number of rows is either n or $2n$. But $a(ba)^{(n-1)/2}$ is not the identity (for example, it doesn't fix 1), so the number of rows again has to be $2n$.

§9.7 #**1.** The numbers 1, 5, 7, 11, 13, 17, 19 and 23 are generators for $\mathbb{Z}/24$, so $\phi(24) = 8$.

§9.9 #**1.** An example of an isomorphism between $\mathbb{Z}/3 \times \mathbb{Z}/4$ and $\mathbb{Z}/12$ is the map taking (a, b) to $4a+3b$. We can interpret this as follows. We can find a copy of $\mathbb{Z}/3$ as the subgroup of $\mathbb{Z}/12$ by using the multiples of four. Similarly, a copy of $\mathbb{Z}/4$ is given by the multiples of three. If we look at $\mathbb{Z}/12$ as the group of transpositions in the twelve tone scale, then $\mathbb{Z}/3$ is the subgroup consisting of transpositions by a whole number of major thirds, while $\mathbb{Z}/4$ is the subgroup consisting of transpositions by a whole number of minor thirds. So every transposition by a whole number of semitones can be written as a combination of a number of major thirds and a number of minor thirds. These numbers are uniquely determined, up to octave equivalence.

§9.9 #**2.** The group $\mathbb{Z}/12\times\mathbb{Z}/2$ has three elements of order two, while $\mathbb{Z}/24$ has only one element of order two. So there cannot be any isomorphism between these groups.

§9.9 #**3.** The group $\mathbb{Z}/2 \times \mathbb{Z}/2$ can be regarded as the group of symmetries of the Taverner example. One way to do this is to make $(1, 0)$ correspond to the temporal symmetry about the bar line, and $(0, 1)$ correspond to the pitch symmetry about the G in the treble clef. Then $(1, 1)$ corresponds to the rotational symmetry around the midpoint of the piece.

§9.10 #**1.** The dihedral group of order six is the group of rigid symmetries of an equilateral triangle. The action on the three vertices of the triangle gives all permutations of this three element set. This action therefore induces an epimorphism from D_6 to S_3, and comparing orders, it must be an isomoporphism.

§9.10 #**2.** The dihedral group of order twelve is the group of rigid symmetries of a regular hexagon. The action on the three pairs of opposite vertices gives a homomorphism from D_{12} to S_3. There are two equilateral triangles formed by taking three equally spaces vertices, and the action on this set of two triangles gives a homomorphism from D_{12} to $\mathbb{Z}/2$. We can use these two homomorphisms to give the coordinates of a homomorphism from D_{12} to $S_3 \times \mathbb{Z}/2$. It is not hard to check that this is an isomorphism.

§9.10 #**3.** The group D_{24} has an element of order eight, whereas $S_3 \times \mathbb{Z}/4$ doesn't.

§9.10 #**4.** The subgroup fixing the chord setwise consists of the elements $\mathbf{T}^n$ and $\mathbf{IT}^n$ with n divisible by three. It is isomorphic to D_8.

§9.10 #**5.** The subgroup fixing the chord setwise consists of the elements $\mathbf{T}^n$ and $\mathbf{IT}^n$ with n divisible by four. It is isomorphic to S_3.

APPENDIX B

Bessel functions

z	$J_0(z)$	$J_1(z)$	$J_2(z)$	$J_3(z)$	$J_4(z)$
0	1	0	0	0	0
0.0001	0.99999 99975 00000	0.00005 00000	1.250×10^{-09}	2.083×10^{-14}	2.604×10^{-19}
0.0002	0.99999 99900 00000	0.00010 00000	5.000×10^{-09}	1.667×10^{-13}	4.167×10^{-18}
0.0005	0.99999 99375 00001	0.00025 00000	3.125×10^{-08}	2.604×10^{-12}	1.628×10^{-16}
0.001	0.99999 97500 00016	0.00049 99999	0.00000 01250	2.083×10^{-11}	2.604×10^{-15}
0.002	0.99999 90000 00250	0.00099 99995	0.00000 05000	1.667×10^{-10}	4.167×10^{-14}
0.005	0.99999 37500 09766	0.00249 99922	0.00000 31250	2.604×10^{-09}	1.628×10^{-12}
0.01	0.99997 50001 56250	0.00499 99375	0.00001 24999	2.083×10^{-08}	2.604×10^{-11}
0.02	0.99990 00024 99972	0.00999 95000	0.00004 99983	0.00000 01667	4.167×10^{-10}
0.03	0.99977 50126 55934	0.01499 83126	0.00011 24916	0.00000 05625	2.109×10^{-09}
0.04	0.99960 00399 98222	0.01999 60003	0.00019 99733	0.00000 13332	6.666×10^{-09}
0.05	0.99937 50976 49468	0.02499 21883	0.00031 24349	0.00000 26038	1.628×10^{-08}
0.06	0.99910 02024 79751	0.02998 65020	0.00044 98650	0.00000 44990	3.374×10^{-08}
0.07	0.99877 53751 05191	0.03497 85669	0.00061 22499	0.00000 71436	6.251×10^{-08}
0.08	0.99840 06398 86234	0.03996 80085	0.00079 95734	0.00001 06624	0.00000 01066
0.09	0.99797 60249 25619	0.04495 44529	0.00101 18167	0.00001 51798	0.00000 01708
0.10	0.99750 15620 66040	0.04993 75260	0.00124 89587	0.00002 08203	0.00000 02603
0.15	0.99438 29052 14140	0.07478 92602	0.00280 72303	0.00007 02137	0.00000 13169
0.20	0.99002 49722 39576	0.09950 08326	0.00498 33542	0.00016 62504	0.00000 41583
0.25	0.98443 59292 95853	0.12402 59773	0.00777 18893	0.00032 42513	0.00001 01408
0.30	0.97762 62465 38296	0.14831 88163	0.01116 58619	0.00055 93430	0.00002 09990
0.35	0.96960 86763 23187	0.17233 39552	0.01515 67821	0.00088 64113	0.00003 88400
0.4	0.96039 82266 59563	0.19602 65780	0.01973 46631	0.00132 00532	0.00006 61351
0.5	0.93846 98072 40813	0.24226 84577	0.03060 40235	0.00256 37300	0.00016 07365
0.6	0.91200 48634 97211	0.28670 09881	0.04366 50967	0.00439 96567	0.00033 14704
0.7	0.88120 08886 07405	0.32899 57415	0.05878 69444	0.00692 96548	0.00061 00970
0.8	0.84628 73527 50480	0.36884 20461	0.07581 77625	0.01024 67663	0.00103 29850
0.9	0.80752 37981 22545	0.40594 95461	0.09458 63043	0.01443 40285	0.00164 05522
1.0	0.76519 76865 57967	0.44005 05857	0.11490 34849	0.01956 33540	0.00247 66390
1.1	0.71962 20185 27511	0.47090 23949	0.13656 41540	0.02569 45286	0.00358 78203
1.2	0.67113 27442 64363	0.49828 90576	0.15934 90183	0.03287 43369	0.00502 26663
1.3	0.62008 59895 61509	0.52202 32474	0.18302 66988	0.04113 58257	0.00683 09584
1.4	0.56695 51203 74289	0.54194 77139	0.20735 58995	0.05049 77133	0.00906 28717
1.5	0.51182 76717 35918	0.55793 65079	0.23208 76721	0.06096 39511	0.01176 81324
1.6	0.45540 21676 39381	0.56989 59353	0.25696 77514	0.07252 34433	0.01499 51611
1.7	0.39798 48594 46109	0.57776 52315	0.28173 89424	0.08514 99269	0.01879 02116
1.8	0.33998 64110 42558	0.58151 69517	0.30614 35353	0.09880 20157	0.02319 65169
1.9	0.28181 85593 74385	0.58115 70727	0.32992 57277	0.11342 34066	0.02825 34512
2.0	0.22389 07791 41236	0.57672 48078	0.35283 40286	0.12894 32495	0.03399 57198
2.1	0.16660 69803 31990	0.56829 21358	0.37462 36252	0.14527 66741	0.04045 25864
2.2	0.11036 22669 22174	0.55596 30498	0.39505 86875	0.16232 54728	0.04764 71475
2.3	0.05553 97844 45602	0.53987 25326	0.41391 45917	0.17997 89313	0.05559 56638
2.4	0.00250 76832 97244	0.52018 52682	0.43098 00402	0.19811 47988	0.06430 69568
2.5	-0.04838 37764 68198	0.49709 41025	0.44605 90584	0.21660 03910	0.07378 18801
2.6	-0.09680 49543 97038	0.47081 82665	0.45897 28517	0.23529 38130	0.08401 28707
2.7	-0.14244 93700 46012	0.44160 13791	0.46956 15027	0.25404 52916	0.09498 35897
2.8	-0.18503 60333 64387	0.40970 92469	0.47768 54954	0.27269 86037	0.10666 86554
2.9	-0.24431 15457 91968	0.37542 74818	0.48322 70505	0.29109 25878	0.11903 34761

z	$J_0(z)$	$J_1(z)$	$J_2(z)$	$J_3(z)$	$J_4(z)$
3.0	-0.26005 19549 01933	0.33905 89585	0.48609 12606	0.30906 27223	0.13203 41839
3.1	-0.29206 43476 50698	0.30092 11331	0.48620 70142	0.32644 27561	0.14561 76751
3.2	-0.32018 81696 57123	0.26134 32488	0.48352 77001	0.34306 63764	0.15972 17556
3.3	-0.34429 62603 98885	0.22066 34530	0.47803 16865	0.35876 88942	0.17427 53940
3.4	-0.36429 55967 62000	0.17922 58517	0.46972 25683	0.37338 89346	0.18919 90810
3.5	-0.38012 77399 87263	0.13737 75274	0.45862 91842	0.38677 01117	0.20440 52930
3.6	-0.39176 89837 00798	0.09546 55472	0.44480 53988	0.39876 26737	0.21979 90574
3.7	-0.39923 02033 71191	0.05383 39877	0.42832 96562	0.40922 51000	0.23527 86141
3.8	-0.40255 64101 78564	0.01282 10029	0.40930 43065	0.41802 56354	0.25073 61706
3.9	-0.40182 60148 87640	-0.02724 40396	0.38785 47125	0.42504 37448	0.26605 87410
4.0	-0.39714 98098 63847	-0.06604 33280	0.36412 81459	0.43017 14739	0.28112 90650
4.1	-0.38866 96798 35854	-0.10327 32577	0.33829 24809	0.43331 47026	0.29582 65960
4.2	-0.37655 70543 67568	-0.13864 69421	0.31053 47010	0.43439 42764	0.31002 85510
4.3	-0.36101 11172 36535	-0.17189 65602	0.28105 92288	0.43334 70056	0.32361 10116
4.4	-0.34225 67900 03886	-0.20277 55219	0.25008 50982	0.43012 65203	0.33645 00658
4.5	-0.32054 25089 85121	-0.23106 04319	0.21784 89837	0.42470 39730	0.34842 29803
4.6	-0.29613 78165 74141	-0.25655 28361	0.18459 31051	0.41706 85798	0.35940 93901
4.7	-0.26933 07894 19753	-0.27908 07358	0.15057 30295	0.40722 79950	0.36929 24960
4.8	-0.24042 53272 91183	-0.29849 98581	0.11605 03864	0.39520 85134	0.37796 02554
4.9	-0.20973 83275 85326	-0.31469 46710	0.08129 15231	0.38105 50980	0.38530 65561
5.0	-0.17759 67713 14338	-0.32757 91376	0.04656 51163	0.36483 12306	0.39123 23605
5.1	-0.14433 47470 60501	-0.33709 72020	0.01213 97659	0.34661 85870	0.39564 68071
5.2	-0.11029 04397 90987	-0.34322 30059	-0.02171 84086	0.32651 65377	0.39846 82598
5.3	-0.07580 31115 85584	-0.34596 08338	-0.05474 81465	0.30464 14780	0.39962 52913
5.4	-0.04121 01012 44991	-0.34534 47908	-0.08669 53768	0.28112 59931	0.39905 75914
5.5	-0.00684 38694 17819	-0.34143 82154	-0.11731 54816	0.25611 78651	0.39671 67891
5.6	0.02697 08846 85114	-0.33433 28363	-0.14637 54691	0.22977 89298	0.39256 71796
5.7	0.05992 00097 24037	-0.32414 76802	-0.17365 60379	0.20228 37940	0.38658 63473
5.8	0.09170 25675 74816	-0.31102 77443	-0.19895 35139	0.17381 84244	0.37876 56770
5.9	0.12203 33545 92823	-0.29514 24447	-0.22208 16409	0.14457 86204	0.36911 07464
6.0	0.15064 52572 50997	-0.27668 38581	-0.24287 32100	0.11476 83848	0.35764 15948
6.1	0.17729 14222 42744	-0.25586 47726	-0.26118 15116	0.08459 82076	0.34439 28633
6.2	0.20174 72229 48904	-0.23291 65671	-0.27688 15994	0.05428 32771	0.32941 38031
6.3	0.22381 20061 32191	-0.20808 69402	-0.28987 13522	0.02404 16372	0.31276 81496
6.4	0.24331 06048 23407	-0.18163 75090	-0.30007 23264	-0.00590 76950	0.29453 38623
6.5	0.26009 46055 81606	-0.15384 13014	-0.30743 03906	-0.03534 66313	0.27480 27310
6.6	0.27404 33606 24146	-0.12498 01652	-0.31191 61379	-0.06405 99184	0.25367 98485
6.7	0.28506 47377 10576	-0.09534 21180	-0.31352 50715	-0.09183 70291	0.23128 29558
6.8	0.29309 56031 04273	-0.06521 86634	-0.31227 75629	-0.11847 40207	0.20774 16623
6.9	0.29810 20354 04820	-0.03490 20961	-0.30821 85850	-0.14377 53445	0.18319 65463
7.0	0.30007 92705 19556	-0.00468 28235	-0.30141 72201	-0.16755 55880	0.15779 81447
7.1	0.29905 13805 01550	0.02515 32743	-0.29196 59511	-0.18964 11340	0.13170 58379
7.2	0.29507 06914 00958	0.05432 74202	-0.27997 97413	-0.20987 17210	0.10508 66405
7.3	0.28821 69476 35014	0.08257 04305	-0.26559 49119	-0.22810 18891	0.07811 39072
7.4	0.27859 62326 57478	0.10962 50949	-0.24896 78286	-0.24420 22995	0.05096 59642
7.5	0.26633 96578 80378	0.13524 84276	-0.23027 34105	-0.25806 09132	0.02382 46800
7.6	0.25160 18338 49976	0.15921 37684	-0.20970 34737	-0.26958 40177	-0.00312 60139
7.7	0.23455 91395 86464	0.18131 27153	-0.18746 49278	-0.27869 70934	-0.02970 16385
7.8	0.21540 78077 46263	0.20135 68728	-0.16377 78404	-0.28534 55088	-0.05571 87049
7.9	0.19436 18448 41278	0.21917 93999	-0.13887 33892	-0.28949 50400	-0.08099 62615
8.0	0.17165 08071 37554	0.23463 63469	-0.11299 17204	-0.29113 22071	-0.10535 74349
8.1	0.14751 74540 44378	0.24760 77670	-0.08637 97338	-0.29026 44256	-0.12863 09519
8.2	0.12221 53017 84138	0.25799 85976	-0.05928 88146	-0.28691 99706	-0.15065 26274
8.3	0.09600 61008 95010	0.26573 93020	-0.03197 25341	-0.28114 77522	-0.17126 68048
8.4	0.06915 72616 56985	0.27078 62683	-0.00468 43406	-0.27301 69067	-0.19032 77356
8.5	0.04193 92518 42935	0.27312 19637	0.02232 47396	-0.26261 62039	-0.20770 08835
8.6	0.01462 29912 78741	0.27275 48445	0.48808 36792	-0.25005 32781	-0.22326 41433
8.7	-0.01252 27324 49665	0.26971 90241	0.07452 71058	-0.23545 36881	-0.23690 89597
8.8	-0.03923 38031 76542	0.26407 37032	0.09925 05539	-0.21895 98151	-0.24854 13369

z	$J_0(z)$	$J_1(z)$	$J_2(z)$	$J_3(z)$	$J_4(z)$
8.9	-0.06525 32468 51244	0.25590 23714	0.12275 93977	-0.20072 96084	-0.25808 27293
9.0	-0.09033 36111 82876	0.24531 17866	0.14484 73415	-0.18093 51903	-0.26547 08018
9.1	-0.11423 92326 83199	0.23243 07450	0.16532 29129	-0.15976 13327	-0.27066 00554
9.2	-0.13674 83707 64864	0.21740 86550	0.18401 11218	-0.13740 38194	-0.27362 23084
9.3	-0.15765 51899 43403	0.20041 39278	0.20075 49594	-0.11406 77088	-0.27434 70295
9.4	-0.17677 15727 51508	0.18163 22040	0.21541 67225	-0.08996 55136	-0.27284 15184
9.5	-0.19392 87476 87422	0.16126 44308	0.22787 91542	-0.06531 53132	-0.26913 09309
9.6	-0.20897 87183 68872	0.13952 48117	0.23804 63875	-0.04033 88170	-0.26325 81481
9.7	-0.22179 54820 31723	0.11663 86479	0.24584 46878	-0.01529 39520	-0.25528 34889
9.8	-0.23227 60275 79367	0.09284 00911	0.25122 29849	0.00969 99027	-0.24528 42690
9.9	-0.24034 11055 34760	0.06836 98323	0.25415 31929	0.03431 83264	-0.23335 42071
10.0	-0.24593 57644 51348	0.04347 27462	0.25463 03137	0.05837 93793	-0.21960 26861
10.5	-0.23664 81944 62347	-0.07885 00142	0.22162 91441	0.16328 01644	-0.12832 61931
11.0	-0.17119 03004 07196	-0.17678 52990	0.13904 75188	0.22734 80331	-0.01503 95007
11.5	-0.06765 39481 11665	-0.22837 86207	0.02793 59271	0.23809 54649	0.09628 77937
12.0	0.04768 93107 96834	-0.22344 71045	-0.08493 04949	0.19513 69395	0.18249 89646
12.5	0.14688 40547 00421	-0.16548 38046	-0.17336 14634	0.11000 81363	0.22616 53689
13.0	0.20692 61023 77068	-0.07031 80521	-0.21774 42642	0.00331 98170	0.21927 64875
13.5	0.21498 91658 80401	0.03804 92921	-0.20935 22337	-0.10007 95836	0.16487 24188
14.0	0.17107 34761 10459	0.13337 51547	-0.15201 98826	-0.17680 94069	0.07624 44225
14.5	0.08754 48680 10376	0.19342 94636	-0.06086 49420	-0.21021 97924	-0.02612 25583
15.0	-0.01422 44728 26781	0.20510 40386	0.04157 16780	-0.19401 82578	-0.11917 89811
15.5	-0.10923 06509 00050	0.16721 31804	0.13080 65451	-0.13345 66526	-0.18246 71848
16.0	-0.17489 90739 83629	0.09039 71757	0.18619 87209	-0.04384 74954	-0.20264 15317
16.5	-0.19638 06929 36861	-0.00576 42137	0.19568 20004	0.05320 22744	-0.17633 57188
17.0	-0.16985 42521 51184	-0.09766 84928	0.15836 38412	0.13493 05730	-0.11074 12860
17.5	-0.10311 03982 28686	-0.16341 99694	0.08443 38303	0.18271 91306	-0.02178 72712
18.0	-0.01335 58057 21984	-0.18799 48855	-0.00753 25149	0.18632 09933	0.06963 95127
18.5	0.07716 48214 22555	-0.16663 36400	-0.09517 92690	0.14605 43386	0.14254 82437
19.0	0.14662 94396 59651	-0.10570 14311	-0.15775 59061	0.07248 96614	0.18064 73781
19.5	0.17885 38270 40173	-0.02087 70701	-0.18099 50650	-0.01625 01227	0.17599 50273
20.0	0.16702 46643 40583	0.06683 31242	-0.16034 13519	-0.09890 13946	0.13067 09336

z	$J_5(z)$	$J_6(z)$	$J_7(z)$	$J_8(z)$	$J_9(z)$
0	0	0	0	0	0
0.1	2.603×10^{-09}	2.169×10^{-11}	1.550×10^{-13}	9.685×10^{-16}	5.380×10^{-18}
0.2	8.319×10^{-08}	1.387×10^{-09}	1.982×10^{-11}	2.477×10^{-13}	2.753×10^{-15}
0.3	0.00000 06304	1.577×10^{-08}	3.381×10^{-10}	6.341×10^{-12}	1.057×10^{-13}
0.4	0.00000 26489	8.838×10^{-08}	2.527×10^{-09}	6.321×10^{-11}	1.405×10^{-12}
0.5	0.00000 80536	0.00000 03361	1.202×10^{-08}	3.758×10^{-10}	1.045×10^{-11}
0.6	0.00001 99482	0.00000 09996	4.291×10^{-08}	1.611×10^{-09}	5.375×10^{-11}
0.7	0.00004 28824	0.00000 25088	0.00000 01257	5.509×10^{-09}	2.145×10^{-10}
0.8	0.00008 30836	0.00000 55601	0.00000 03186	1.597×10^{-08}	7.109×10^{-10}
0.9	0.00014 86580	0.00001 12036	0.00000 07229	4.077×10^{-08}	2.043×10^{-09}
1.0	0.00024 97577	0.00002 09383	0.00000 15023	9.422×10^{-08}	5.249×10^{-09}
1.1	0.00039 87099	0.00003 68150	0.00000 29084	0.00000 02008	1.231×10^{-08}
1.2	0.00061 01049	0.00006 15414	0.00000 53093	0.00000 04002	2.679×10^{-08}
1.3	0.00090 08414	0.00009 85905	0.00000 92248	0.00000 07540	5.471×10^{-08}
1.4	0.00129 01251	0.00015 23073	0.00001 53661	0.00000 13538	0.00000 01059
1.5	0.00179 94218	0.00022 80127	0.00002 46798	0.00000 23321	0.00000 01956
1.6	0.00245 23620	0.00033 21012	0.00003 83972	0.00000 38744	0.00000 03469
1.7	0.00327 45981	0.00047 21304	0.00005 80872	0.00000 62348	0.00000 05936
1.8	0.00429 36149	0.00065 68991	0.00008 57125	0.00000 97534	0.00000 09843
1.9	0.00553 84930	0.00089 65121	0.00012 36884	0.00001 48764	0.00000 15863
2.0	0.00703 96298	0.00120 24290	0.00017 49441	0.00002 21796	0.00000 24923
2.1	0.00882 84171	0.00158 74951	0.00024 29833	0.00003 23938	0.00000 38266
2.2	0.01093 68819	0.00206 59518	0.00033 19463	0.00004 64337	0.00000 57535
2.3	0.01339 72905	0.00265 34256	0.00044 66689	0.00006 54286	0.00000 84866
2.4	0.01624 17239	0.00336 68927	0.00059 27398	0.00009 07560	0.00001 23002

z	$J_5(z)$	$J_6(z)$	$J_7(z)$	$J_8(z)$	$J_9(z)$
2.5	0.01950 16251	0.00422 46205	0.00077 65532	0.00012 40774	0.00001 75420
2.6	0.02320 73276	0.00524 60815	0.00100 53563	0.00016 73755	0.00002 46466
2.7	0.02738 75668	0.00645 18427	0.00128 72898	0.00022 29934	0.00003 41524
2.8	0.03206 89832	0.00786 34275	0.00163 14204	0.00029 36744	0.00004 67189
2.9	0.03727 56220	0.00950 31514	0.00204 77633	0.00038 26023	0.00006 31459
3.0	0.04302 84349	0.01139 39323	0.00254 72945	0.00049 34418	0.00008 43950
3.1	0.04934 47926	0.01355 90753	0.00314 19503	0.00063 03778	0.00011 16123
3.2	0.05623 80126	0.01602 20338	0.00384 46142	0.00079 81533	0.00014 61522
3.3	0.06371 69093	0.01880 61494	0.00466 90886	0.00100 21053	0.00018 96036
3.4	0.07178 53735	0.02193 43706	0.00563 00521	0.00124 81970	0.00024 38159
3.5	0.08044 19866	0.02542 89545	0.00674 30003	0.00154 30467	0.00031 09276
3.6	0.08967 96760	0.02931 11538	0.00802 41700	0.00189 39518	0.00039 33937
3.7	0.09948 54170	0.03360 08913	0.00949 00447	0.00230 89068	0.00049 40152
3.8	0.10983 99868	0.03831 64263	0.01115 92541	0.00279 66150	0.00061 59670
3.9	0.12071 77752	0.04347 40159	0.01304 84275	0.00336 64932	0.00076 28267
4.0	0.13208 66560	0.04908 75752	0.01517 60694	0.00402 86678	0.00093 86019
4.1	0.14390 79237	0.05516 83400	0.01756 03884	0.00479 39619	0.00114 77557
4.2	0.15613 62970	0.06172 45370	0.02021 95230	0.00567 38731	0.00139 52316
4.5	0.19471 46586	0.08427 62611	0.03002 20377	0.00912 56340	0.00242 46609
5.0	0.26114 05461	0.13104 87318	0.05337 64102	0.01840 52167	0.00552 02831
5.5	0.32092 47371	0.18678 27330	0.08660 12258	0.03365 67508	0.01130 93220
6.0	0.36208 70749	0.24583 68634	0.12958 66518	0.05653 19909	0.02116 53240
6.5	0.37356 53771	0.29991 32338	0.18012 05930	0.08803 88126	0.03659 03304
7.0	0.34789 63248	0.33919 66050	0.23358 35695	0.12797 05340	0.05892 05083
7.5	0.28347 39052	0.35414 05269	0.28315 09379	0.17440 78905	0.08891 92285
8.0	0.18577 47722	0.33757 59001	0.32058 90780	0.22345 49864	0.12632 08947
8.5	0.06713 30194	0.28668 09063	0.33759 29660	0.26935 45671	0.16942 73956
9.0	-0.05503 88557	0.20431 65177	0.32746 08792	0.30506 70723	0.21488 05825
9.5	-0.16132 12602	0.09931 90781	0.28677 69378	0.32329 95671	0.25772 75962
10.0	-0.23406 15282	-0.01445 88421	0.21671 09177	0.31785 41268	0.29185 56853
10.5	-0.26105 25019	-0.12029 52374	0.12357 22307	0.28505 82116	0.31080 21870
11.0	-0.23828 58518	-0.20158 40009	0.01837 60326	0.22497 16788	0.30885 55001
11.5	-0.17111 26519	-0.24508 14040	-0.08462 44654	0.14206 03158	0.28227 36003
12.0	-0.07347 09631	-0.24372 47672	-0.17025 38041	0.04509 53291	0.23038 09096
12.5	0.03473 76998	-0.19837 52091	-0.22517 79005	-0.05382 40395	0.15628 31300
13.0	0.13161 95599	-0.11803 06721	-0.24057 09496	-0.14104 57351	0.06697 61987
13.5	0.19778 17577	-0.01836 74131	-0.21410 83471	-0.21410 83471	-0.20367 08728
14.0	0.22037 76483	0.08116 81834	-0.15080 49196	-0.23197 31031	-0.11430 71981
14.5	0.19580 73465	0.16116 21076	-0.06243 18091	-0.22144 10957	-0.18191 69861
15.0	0.13045 61346	0.20614 97375	0.03446 36554	-0.17398 36591	-0.22004 62251
15.5	0.03928 00410	0.20780 91468	0.12160 44597	-0.09797 28606	-0.22273 77352
16.0	-0.05747 32704	0.16672 07377	0.18251 38237	-0.00702 11420	-0.18953 49657
16.5	-0.13869 83805	0.09227 60942	0.20580 82672	0.08234 91022	-0.12595 45923
17.0	-0.18704 41194	0.00071 53334	0.18754 90607	0.15373 68342	-0.04285 55697
17.5	-0.19267 90261	-0.08831 50294	0.13212 01488	0.19401 11484	0.04526 14726
18.0	-0.15537 00988	-0.15595 62342	0.05139 92760	0.19593 34488	0.12276 37897
18.5	-0.08441 18549	-0.18817 62733	-0.03764 84305	0.15968 55691	0.17575 48687
19.0	0.00357 23925	-0.17876 71715	-0.11647 79745	0.09294 12956	0.19474 43287
19.5	0.08845 32108	-0.13063 44063	-0.16884 36147	0.00941 33496	0.17656 73888
20.0	0.15116 97680	-0.05508 60496	-0.18422 13977	-0.07386 89288	0.12512 62546

z	$J_{10}(z)$	$J_{11}(z)$	$J_{12}(z)$	$J_{13}(z)$	$J_{14}(z)$
0	0	0	0	0	0
0.1	2.691×10^{-20}	1.223×10^{-22}	5.096×10^{-25}	1.960×10^{-27}	7.000×10^{-30}
0.2	2.753×10^{-17}	2.503×10^{-19}	2.086×10^{-21}	1.605×10^{-23}	1.146×10^{-25}
0.3	1.586×10^{-15}	2.163×10^{-17}	2.704×10^{-19}	3.120×10^{-21}	3.344×10^{-23}
0.4	2.812×10^{-14}	5.114×10^{-16}	8.525×10^{-18}	1.312×10^{-19}	1.874×10^{-21}
0.5	2.613×10^{-13}	5.942×10^{-15}	1.238×10^{-16}	2.382×10^{-18}	4.255×10^{-20}
0.6	1.614×10^{-12}	4.405×10^{-14}	1.102×10^{-15}	2.544×10^{-17}	5.454×10^{-19}

z	$J_{10}(z)$	$J_{11}(z)$	$J_{12}(z)$	$J_{13}(z)$	$J_{14}(z)$
0.7	7.518×10^{-12}	2.394×10^{-13}	6.989×10^{-15}	1.883×10^{-16}	4.710×10^{-18}
0.8	2.848×10^{-11}	1.037×10^{-12}	3.460×10^{-14}	1.065×10^{-15}	3.046×10^{-17}
0.9	9.212×10^{-11}	3.774×10^{-12}	1.417×10^{-13}	4.911×10^{-15}	1.580×10^{-16}
1.0	2.631×10^{-10}	1.198×10^{-11}	5.000×10^{-13}	1.925×10^{-14}	6.885×10^{-16}
1.1	6.791×10^{-10}	3.403×10^{-11}	1.563×10^{-12}	6.623×10^{-14}	2.606×10^{-15}
1.2	1.613×10^{-09}	8.820×10^{-11}	4.420×10^{-12}	2.044×10^{-13}	8.776×10^{-15}
1.3	3.570×10^{-09}	2.116×10^{-10}	1.149×10^{-11}	5.761×10^{-13}	2.680×10^{-14}
1.4	7.444×10^{-09}	4.755×10^{-10}	2.783×10^{-11}	1.502×10^{-12}	7.529×10^{-14}
1.5	1.474×10^{-08}	1.010×10^{-09}	6.333×10^{-11}	3.665×10^{-12}	1.969×10^{-13}
1.6	2.791×10^{-08}	2.040×10^{-09}	1.366×10^{-10}	8.433×10^{-12}	4.834×10^{-13}
1.7	5.080×10^{-08}	3.947×10^{-09}	2.809×10^{-10}	1.844×10^{-11}	1.123×10^{-12}
1.8	8.924×10^{-08}	7.347×10^{-09}	5.539×10^{-10}	3.852×10^{-11}	2.486×10^{-12}
1.9	0.00000 01520	1.321×10^{-08}	1.052×10^{-09}	7.728×10^{-11}	5.267×10^{-12}
2.0	0.00000 02515	2.304×10^{-08}	1.933×10^{-09}	1.495×10^{-10}	1.073×10^{-11}
2.1	0.00000 04059	3.907×10^{-08}	3.443×10^{-09}	2.798×10^{-10}	2.110×10^{-11}
2.2	0.00000 06400	6.460×10^{-08}	5.968×10^{-09}	5.084×10^{-10}	4.018×10^{-11}
2.3	0.00000 09880	0.00000 01043	1.009×10^{-08}	8.987×10^{-10}	7.430×10^{-11}
2.4	0.00000 14958	0.00000 01650	1.665×10^{-08}	1.550×10^{-10}	1.338×10^{-10}
2.5	0.00000 22247	0.00000 02559	2.693×10^{-08}	2.612×10^{-09}	2.349×10^{-10}
2.6	0.00000 32547	0.00000 03897	4.268×10^{-08}	4.309×10^{-09}	4.034×10^{-10}
2.7	0.00000 46894	0.00000 05837	6.645×10^{-08}	6.971×10^{-09}	6.781×10^{-10}
2.8	0.00000 66611	0.00000 08607	0.00000 01017	1.107×10^{-08}	1.118×10^{-09}
2.9	0.00000 93376	0.00000 12511	0.00000 01533	1.729×10^{-08}	1.810×10^{-09}
3.0	0.00001 29284	0.00000 17940	0.00000 02276	2.659×10^{-08}	2.880×10^{-09}
3.1	0.00001 76936	0.00000 25402	0.00000 03333	4.028×10^{-08}	4.512×10^{-09}
3.2	0.00002 39530	0.00000 35542	0.00000 04819	6.017×10^{-08}	6.962×10^{-09}
3.3	0.00003 20960	0.00000 49177	0.00000 06884	8.872×10^{-08}	1.059×10^{-08}
3.4	0.00004 25933	0.00000 67328	0.00000 09721	0.00000 01292	1.591×10^{-08}
3.5	0.00005 60095	0.00000 91267	0.00000 13581	0.00000 01860	2.360×10^{-08}
4.0	0.00019 50406	0.00003 66009	0.00000 62645	0.00000 09859	0.00000 01436
4.5	0.00057 30098	0.00012 20492	0.00002 36751	0.00000 42179	0.00000 06950
5.0	0.00146 78026	0.00035 09274	0.00007 62781	0.00001 52076	0.00000 28013
5.5	0.00335 55759	0.00089 27721	0.00021 55123	0.00004 76455	0.00000 97207
6.0	0.00696 39810	0.00204 79460	0.00054 51544	0.00013 26717	0.00002 97564
6.5	0.01328 82562	0.00429 66118	0.00125 41220	0.00033 39927	0.00008 18487
7.0	0.02353 93444	0.00833 47614	0.00265 56200	0.00077 02216	0.00020 52029
7.5	0.03899 82579	0.01507 61259	0.00522 50447	0.00164 40171	0.00047 42147
8.0	0.06076 70268	0.02559 66722	0.00962 38218	0.00327 47932	0.00101 92562
8.5	0.08943 28589	0.04100 28606	0.01669 21921	0.00612 80346	0.00205 23844
9.0	0.12469 40928	0.06221 74015	0.02739 28887	0.01083 03016	0.00398 46493
9.5	0.16502 64047	0.08969 64137	0.04269 16060	0.01815 60646	0.00699 86761
10.0	0.20748 61066	0.12311 65280	0.06337 02550	0.02897 20839	0.01195 71632
10.5	0.24774 55375	0.16109 40750	0.08978 49053	0.04412 85657	0.01948 58287
11.0	0.28042 82305	0.20101 40099	0.12159 97893	0.06429 46213	0.03036 93155
11.5	0.29975 92326	0.23904 68041	0.15754 76971	0.08974 83898	0.04536 17059
12.0	0.30047 60353	0.27041 24826	0.19528 01827	0.12014 78829	0.06504 02303
12.5	0.27887 17466	0.28991 16646	0.23137 27831	0.15432 40789	0.08962 13011
13.0	0.23378 20102	0.29268 84324	0.26153 68754	0.19014 88760	0.11876 08767
13.5	0.16729 84008	0.27512 88367	0.28105 97034	0.22453 28582	0.15137 39495
14.0	0.08500 67054	0.23574 53488	0.28545 02712	0.25359 79733	0.18551 73935
14.5	-0.00438 68871	0.17586 61074	0.27121 82225	0.27304 68125	0.21838 29586
15.0	-0.09007 18110	0.09995 04771	0.23666 58441	0.27871 48734	0.24643 99366
15.5	-0.16069 03157	0.01539 53923	0.18254 18403	0.26725 00378	0.26574 85457
16.0	-0.20620 56944	-0.06822 21524	0.11240 02349	0.23682 25048	0.27243 63353
16.5	-0.21975 41120	-0.14041 40283	0.03253 54076	0.18773 82576	0.26329 45740
17.0	-0.19911 33197	-0.19139 53947	-0.04857 48381	0.12281 91527	0.23641 58951
17.5	-0.14745 64908	-0.21378 31764	-0.12129 95024	0.04742 95731	0.19176 62968
18.0	-0.07316 96592	-0.20406 34110	-0.17624 11765	-0.03092 48243	0.13157 19858
18.5	0.11319 16799	-0.16351 79303	-0.20577 29230	-0.10343 07265	0.06041 08209

z	$J_{10}(z)$	$J_{11}(z)$	$J_{12}(z)$	$J_{13}(z)$	$J_{14}(z)$
19.0	0.09155 33316	-0.09837 24007	-0.20545 82166	-0.16115 37677	-0.01506 79918
19.5	0.15357 19323	-0.01905 77146	-0.17507 29436	-0.19641 66776	-0.08681 59598
20.0	0.18648 25580	0.06135 63034	-0.11899 06243	-0.20414 50525	-0.14639 79440

Table of zeros of Bessel functions

Note: The kth zero of J_n is denoted $j_{n,k}$.

k	J_0	J_1	J_2	J_3	J_4	J_5	J_6	J_7
1	2.40482 55577	3.831706	5.135622	6.380162	7.588342	8.771484	9.936110	11.08637
2	5.52007 81103	7.015587	8.417244	9.761023	11.06471	12.33860	13.58929	14.82127
3	8.65372 79129	10.17347	11.61984	13.01520	14.37254	15.70017	17.00382	18.28758
4	11.79153 44391	13.32369	14.79595	16.22347	17.61597	18.98013	20.32079	21.64154
5	14.93091 77086	16.47063	17.95982	19.40942	20.82693	22.21780	23.58608	24.93493
6	18.07106 39679	19.61586	21.11700	22.58273	24.01902	25.43034	26.82015	28.19119
7	21.21163 66299	22.76008	24.27011	25.74817	27.19909	28.62662	30.03372	31.42279
8	24.35247 15308	25.90367	27.42057	28.90835	30.37101	31.81172	33.23304	34.63709
9	27.49347 91320	29.04683	30.56920	32.06485	33.53714	34.98878	36.42202	37.83872
10	30.63460 64684	32.18968	33.71652	35.21867	36.69900	38.15987	39.60324	41.03077
11	33.77582 02136	35.33231	36.86286	38.37047	39.85763	41.32638	42.77848	44.21541
12	36.91709 83537	38.47477	40.00845	41.52072	43.01374	44.48932	45.94902	47.39417
13	40.05842 57646	41.61709	43.15345	44.66974	46.16785	47.64940	49.11577	50.56818
14	43.19979 17132	44.75932	46.29800	47.81779	49.32036	50.80717	52.27945	53.73833
15	46.34118 83717	47.90146	49.44216	50.96503	52.47155	53.96303	55.44059	56.90525

k	J_8	J_9	J_{10}	J_{11}	J_{12}	J_{13}	J_{14}	J_{15}
1	12.22509	13.35430	14.47550	15.58985	16.69825	17.80144	18.90000	19.99443
2	16.03777	17.24122	18.43346	19.61597	20.78991	21.95624	23.11578	24.26918
3	19.55454	20.80705	22.04699	23.27585	24.49489	25.70510	26.90737	28.10242
4	22.94517	24.23389	25.50945	26.77332	28.02671	29.27063	30.50595	31.73341
5	26.26681	27.58375	28.88738	30.17906	31.45996	32.73105	33.99318	35.24709
6	29.54566	30.88538	32.21186	33.52636	34.82999	36.12366	37.40819	38.68428
7	32.79580	34.15438	35.49991	36.83357	38.15638	39.46921	40.77283	42.06792
8	36.02562	37.40010	38.76181	40.11182	41.45109	42.78044	44.10059	45.41219

k	J_{16}	J_{17}	J_{18}	J_{19}	J_{20}	J_{21}	J_{22}	J_{23}
1	21.08515	22.17249	23.25678	24.33825	25.41714	26.49365	27.56794	28.64019
2	25.41701	26.55979	27.69790	28.83173	29.96160	31.08780	32.21059	33.33018
3	29.29087	30.47328	31.65012	32.82180	33.98870	35.15115	36.30943	37.46381
4	32.95366	34.16727	35.37472	36.57645	37.77286	38.96429	40.15105	41.33343
5	36.49340	37.73268	38.96543	40.19210	41.41307	42.62870	43.83932	45.04521

k	J_{24}	J_{25}	J_{26}	J_{27}	J_{28}	J_{29}	J_{30}	J_{31}
1	29.71051	30.77904	31.84589	32.91115	33.97493	35.03730	36.09834	37.15811
2	34.44678	35.56057	36.67173	37.78040	38.88671	39.99080	41.09278	42.19275
3	38.61452	39.76179	40.90580	42.04674	43.18477	44.32003	45.45267	46.58280

k	J_{32}	J_{33}	J_{34}	J_{35}	J_{36}	J_{37}	J_{38}	J_{39}
1	38.21669	39.27413	40.33048	41.38580	42.44014	43.49352	44.54601	45.59762
2	43.29082	44.38706	45.48156	46.57441	47.66568	48.75542	49.84371	50.93060

Fourier series

$$\sin(z\sin\theta) = 2\sum_{n=0}^{\infty} J_{2n+1}(z)\sin(2n+1)\theta$$

$$\cos(z\sin\theta) = J_0(z) + 2\sum_{n=1}^{\infty} J_{2n}(z)\cos 2n\theta$$

$$J_n(z) = \frac{1}{\pi}\int_0^{\pi}\cos(n\theta - z\sin\theta)\,d\theta.$$

Differential equation

$$J_n''(z) + \frac{1}{z}J_n'(z) + \left(1 - \frac{n^2}{z^2}\right)J_n(z) = 0$$

Power series

$$J_n(z) = \sum_{k=0}^{\infty}\frac{(-1)^k(\frac{z}{2})^{n+2k}}{k!(n+k)!}$$

Generating function

$$e^{\frac{1}{2}z(t-\frac{1}{t})} = \sum_{n=-\infty}^{\infty} J_n(z)t^n$$

Limiting values

If n is constant, z is real and $|z| \to \infty$,

$$J_n(z) = \sqrt{\tfrac{2}{\pi z}}\cos(z - \tfrac{1}{2}(n+\tfrac{1}{2})\pi) + O(|z|^{-3/2}).$$

[Here, $O(|z|^{-3/2})$ represents an error term which is bounded by some constant multiple of $|z|^{-3/2}$]

If z is constant and $n \to \infty$, $J_n(z) \sim \frac{1}{\sqrt{2n\pi}}\left(\frac{ez}{2n}\right)^n$.

For n fixed, as $k \to \infty$, $j_{n,k} \sim (k + \frac{1}{2}n - \frac{1}{4})\pi$.

Other formulas

$$J_{-n}(z) = (-1)^n J_n(z)$$

$$J_n'(z) = \tfrac{1}{2}(J_{n-1}(z) - J_{n+1}(z))$$

$$J_n(z) = \tfrac{z}{2n}(J_{n-1}(z) + J_{n+1}(z))$$

$$\frac{d}{dz}(z^n J_n(z)) = z^n J_{n-1}(z)$$

$$1 = \sum_{n=-\infty}^{\infty} J_n(z) = J_0(z) + 2J_2(z) + 2J_4(z) + 2J_6(z) + \dots$$

$$1 = \sum_{n=-\infty}^{\infty} J_n(z)^2 = J_0(z)^2 + 2J_1(z)^2 + 2J_2(z)^2 + 2J_3(z)^2 + \dots$$

In particular, $|J_n(z)| \leq 1$ for all n and z, and if $n \neq 0$ then $|J_n(z)| \leq \frac{1}{\sqrt{2}}$.

Computation

Although the power series converges very quickly for small values of z, and converges for all values of z, rounding errors tend to accumulate for larger z because a small number is resulting from addition and subtraction of very large numbers.

Instead, a computer program for calculating the Bessel functions can be based on the recurrence relation $J_n(z) = (2(n+1)/z)J_{n+1}(z) - J_{n+2}(z)$ and normalizing via the relation $J_0(z) + 2J_2(z) + 2J_4(z) + \cdots = 1$. This is called *Miller's backwards recurrence algorithm* (J. C. P. Miller, *The Airy integral*, 1946). Build an array indexed by n and make the last two entries 1 and 0, use the recurrence relation to calculate the remaining entries, and then normalize. An array containing 100 entries gives reasonable accuracy, and does not consume much memory. Here is a simple C++ program which implements this method. I haven't put in any exception checking.

```
/* file bessel.cpp */
#include <iostream.h>
#include <stdio.h>
#define length 100
void main() {
  long double X[length], z, sum;
  int n=0, j=0;
  X[length - 2]=1; X[length - 1]=0;
  while (1)
  {
    printf("\n\nOrder (integer); -1 to exit: ");
    cin>>n;
    if (n<0)
      break;
    printf("Argument (real): ");
    cin>>z;
    if (z==0)        // prevent divide by zero
      {printf("J_0(0)=1; J_n(0)=0 (n>0)");}
    else
      {for(j=length - 3; j>=0; --j)
        {X[j]=(2*(j+1)/z)*X[j+1] - X[j+2];}
      sum=X[0];
      for(j=2; j < length; j=j+2)
        {sum+=2*X[j];}
      printf("J_%d(%Lf)= %11.10Lf",n,z,X[n]/sum);
    }
  }
}
```

I compiled this program using Borland C++. It prints out the answer to 10 decimal places, and at least for reasonably small values of n and z, up to about 50, the answers it gives agree with published tables to this accuracy. If you need more accuracy, I recommend the standard Unix multiple precision arithmetic utility bc. If invoked with the option -l (which loads the library mathlib of mathematical functions), it recognizes the syntax j(n,z) and calculates $J_n(z)$ using the above algorithm. The number of digits after the decimal point is set to 50, for example, by using the command scale=50. Windows users can use bc in the free Unix environment Cygwin (www.cygwin.com); there is also a (free) version compiled for MS-DOS in `UnxUtils.zip` (unxutils.sourceforge.net). Here is a sample session:

```
$ bc -l
j(1,1)
.44005058574493351595
scale=50
for (n=0;n<5;n++) {j(n,1)}
.76519768655796655144971752610266322090927428975532
.44005058574493351595968220371891491312737230199276
.11490348493190048046964688133516660534547031423020
.01956335398266840591890532162175150825450895492805
.00247663896410995504378504839534244418158341533812
quit
$
```

FM Synthesis

$$\sin(\phi + z\sin\theta) = \sum_{n=-\infty}^{\infty} J_n(z)\sin(\phi + n\theta)$$

The following table shows how index of modulation (z) varies as a function of operator output level (an integer in the range 0–99) on the Yamaha six operator synthesizers DX7, DX7IID, DX7IIFD, DX7S, DX5, DX1, TX7, TX816, TX216, TX802 and TF1:

	0	1	2	3	4	5	6	7	8	9
0	0.0002	0.0003	0.0005	0.0007	0.0010	0.0012	0.0016	0.0019	0.0023	0.0027
10	0.0032	0.0038	0.0045	0.0054	0.0064	0.0076	0.0083	0.0091	0.0108	0.0118
20	0.0140	0.0152	0.0166	0.0181	0.0198	0.0216	0.0235	0.0256	0.0280	0.0305
30	0.0332	0.0362	0.0395	0.0431	0.0470	0.0513	0.0559	0.0610	0.0665	0.0725
40	0.0791	0.0862	0.0940	0.1025	0.1118	0.1219	0.1330	0.1450	0.1581	0.1724
50	0.1880	0.2050	0.2236	0.2438	0.2659	0.2900	0.3162	0.3448	0.3760	0.4101
60	0.4472	0.4877	0.5318	0.5799	0.6324	0.6897	0.7521	0.8202	0.8944	0.9754
70	1.0636	1.1599	1.2649	1.3794	1.5042	1.6403	1.7888	1.9507	2.1273	2.3198
80	2.5298	2.7587	3.0084	3.2807	3.5776	3.9014	4.2545	4.6396	5.0595	5.5174
90	6.0168	6.5614	7.1552	7.8028	8.5090	9.2792	10.119	11.035	12.034	13.123

The following table shows how index of modulation (z) varies as a function of operator output level (an integer in the range 0–99) on the Yamaha four operator synthesizers DX11, DX21, DX27, DX27S, DX100 and TX81Z:

	0	1	2	3	4	5	6	7	8	9
0	0.0004	0.0006	0.0009	0.0013	0.0018	0.0024	0.0031	0.0036	0.0043	0.0052
10	0.0061	0.0073	0.0087	0.0103	0.0123	0.0146	0.0159	0.0174	0.0206	0.0225
20	0.0268	0.0292	0.0318	0.0347	0.0379	0.0413	0.0450	0.0491	0.0535	0.0584
30	0.0637	0.0694	0.0757	0.0826	0.0900	0.0982	0.1071	0.1168	0.1273	0.1388
40	0.1514	0.1651	0.1801	0.1963	0.2141	0.2335	0.2546	0.2777	0.3028	0.3302
50	0.3601	0.3927	0.4282	0.4670	0.5093	0.5554	0.6056	0.6604	0.7202	0.7854
60	0.8565	0.9340	1.0185	1.1107	1.2112	1.3209	1.4404	1.5708	1.7130	1.8680
70	2.0371	2.2214	2.4225	2.6418	2.8809	3.1416	3.4259	3.7360	4.0741	4.4429
80	4.8450	5.2835	5.7617	6.2832	6.8519	7.4720	8.1483	8.8858	9.6900	10.567
90	11.523	12.566	13.704	14.944	16.297	17.772	19.380	21.134	23.047	25.133

APPENDIX C

Complex numbers

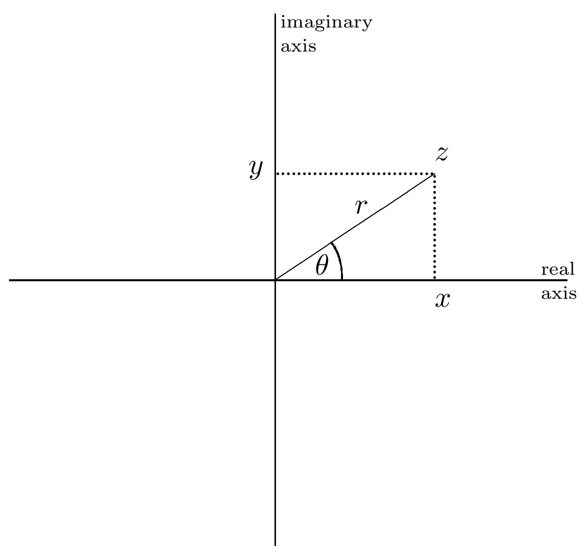
We use i to denote $\sqrt{-1}$, and the general complex number is of the form $a+ib$ where a and b are real numbers. Addition and multiplication are given by

$$(a_1+ib_1)+(a_2+ib_2)=(a_1+a_2)+i(b_1+b_2)$$
$$(a_1+ib_1)(a_2+ib_2)=(a_1a_2-b_1b_2)+i(a_1b_2+b_1a_2).$$

These formulas follow from the equation $i^2=-1$ and the usual rules of multiplication and addition, such as the distributivity of multiplication over addition.

The real numbers a and b can be thought of as the Cartesian coordinates of the complex number $a+ib$, so that complex numbers correspond to points on the plane. In this language, the real numbers are contained in the complex numbers as the x axis, and the points on the y axis are called pure imaginary numbers.

For the purpose of multiplication, it is easier to work in polar coordinates. If $z=x+iy$ is a complex number, we define the *absolute value* of z to be $|z|=\sqrt{x^2+y^2}$. The *argument* of z is the angle θ formed by the line from zero to z. Angle is measured counterclockwise from the x axis.



The *complex conjugate* of $z = x + iy$ is defined to be $\bar{z} = x - iy$, so that

$$z\bar{z} = |z|^2 = x^2 + y^2.$$

So division by a nonzero complex number z is achieved by multiplying by

$$\frac{\bar{z}}{|z|^2} = \frac{x}{x^2 + y^2} - i\,\frac{y}{x^2 + y^2},$$

which is the *multiplicative inverse* of z.

The *exponential function* is defined for a complex argument $z = x + iy$ by

$$e^z = e^x(\cos y + i \sin y).$$

This means that convertion from Cartesian coordinates to polar coordinates is given by

$$z = x + iy = re^{i\theta},$$

where $r = \sqrt{x^2 + y^2}$ and $\tan\theta = y/x$. Translation in the other direction is given by $x = r\cos\theta$ and $y = r\sin\theta$. The trigonometric identities

$$\sin(A + B) = \sin A \cos B + \cos A \sin B$$
$$\cos(A + B) = \cos A \cos B - \sin A \sin B.$$

are equivalent to the statement that if z_1 and z_2 are complex numbers then

$$e^{z_1} e^{z_2} = e^{z_1 + z_2}.$$

So we have Euler's formula

$$e^{i\theta} = \cos\theta + i \sin\theta \tag{C.1}$$

and

$$\cos\theta = \tfrac{1}{2}(e^{i\theta} + e^{-i\theta}) \tag{C.2}$$
$$\sin\theta = \tfrac{1}{2i}(e^{i\theta} - e^{-i\theta}). \tag{C.3}$$

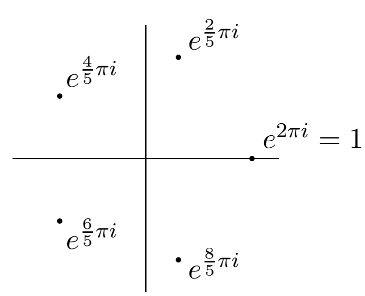
Using (C.1), the relation $(e^{i\theta})^n = e^{in\theta}$ translates as de Moivre's Theorem

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i \sin n\theta.$$

The complex nth roots of unity (i.e., of the number one) are the numbers

$$e^{2\pi i m/n} = \cos 2\pi m/n + i \sin 2\pi m/n$$

for $0 \le m \le n-1$. These are equally spaced around the unit circle in the complex plane. For example, here is a picture of the complex fifth roots of unity.



Remark. Engineers use the letter j instead of i.

Hyperbolic functions: In Section 3.9 the analysis of the xylophone involves the *hyperbolic functions* $\cosh x$ and $\sinh x$. These are defined by analogy with equations (C.2) and (C.3) via

$$\cosh x = \tfrac{1}{2}(e^x + e^{-x}) \tag{C.4}$$

$$\sinh x = \tfrac{1}{2}(e^x - e^{-x}). \tag{C.5}$$

The standard identities for these functions are

$$\cosh^2 x - \sinh^2 x = 1,$$

and

$$\sinh(A+B) = \sinh A \cosh B + \cosh A \sinh B$$

$$\cosh(A+B) = \cosh A \cosh B + \sinh A \sinh B.$$

The values at zero are given by

$$\sinh(0) = 0, \qquad \cosh(0) = 1.$$

The derivatives are given by

$$\frac{d}{dx}\sinh x = \cosh x, \qquad \frac{d}{dx}\cosh x = \sinh x.$$

Note the changes in sign from the corresponding trigonometric formulas.

APPENDIX D

Dictionary

As an aide to reading the literature on the subject in French, German, Italian, Latin and Spanish, as well as the literature on ancient Greek music, here is a dictionary of common terms. I have tried to avoid including words whose meaning is obvious.

abaissé (Fr.), *lowered*
abdämpfen (G.), *to damp, mute*
Abklingen (G.), *decay*
Abgeleiteter Akkord (G.), *inversion of a chord*
Absatz (G.), *cadence*
Abstimmung (G.), *tuning*
accord (Fr.), *chord*
accordage (Fr.), accordatura (It.), *tuning, intonation*
accordo (It.), *chord*
Achtelnote (G.), *eighth note* (USA), *quaver* (GB)
acorde (Sp.), *chord*
afinación (Sp.), *tuning*
affaiblissement (Fr.), *decay*
aigu (Fr.), *acute, high*
Akkord (G.), *chord*
allgemein (G.), *general*
alma (Sp.), âme (Fr.), *sound post*
anima (It.), *sound post*
Anklang (G.), *tune, harmony, accord*
archet (Fr.), arco (It., Sp.), *bow*
armoneggiare (It.), *to harmonize*
armonica (It.), armónico (Sp.), *harmonic*
armure (Fr.), *key signature*
atenuamiento (Sp.), attenuazione (It.), *decay*
audición (Sp.), audition (Fr.), *hearing*
auferions (archaic Eng.), *wire strings*
Aufhaltung (G.), *suspension (harmony)*
aufzählen (G.), *to enumerate*
aulos (Gk.), *ancient Greek reed instrument*
Ausdruck (G.), *expression*

B (G.), *B*♭ *(in German H denotes B)*
barre (Fr.), *bar line*
battements (Fr.), battimenti (It.), *beats*
battuta (It.), *beat*
bec (Fr.), becco (It.), *mouthpiece*
bécarre (Fr.), becuardo (Sp.), *natural* (♮)
Bedingung (G.), *condition*
Beispiel (G.), *example*
beliebig (G.), *arbitrary*
bémol (Fr.), bemol (Sp.), bemolle (It.), *flat* (♭)
bequadro (It.), *natural* (♮)
beweisen (G.), *to prove*
Beziehung (G.), *relation*
blanche (Fr.), *half note* (USA), *minim* (GB)
Blasinstrument (G.), *wind instrument*
Bogen (G.), *bow*
bois (Fr.), *wood, (pl.) woodwind*
bruit (Fr.), *noise*
Bund (G.), *fret*

cadenza d'inganno (It.), *deceptive cadence*
caisse (Fr.), *drum*
canon (Gk.), *monochord*
Canonici, *followers of the Pythagorean system of music, where consonance is based on ratios, see also* Musici
chevalet (Fr.), *bridge of stringed instrument*
cheville (Fr.), *peg, pin*
chiave (It.), clave (Sp.), clavis (L.), *clef, key*
chiffrage (Fr.), *time signature*
chiuso (It.), *closed*
clavecin (Fr.), *harpsichord*

cloche (Fr.), *bell*
comma enharmonique (Fr.), *great diesis*
concento (It.), concentus (L.), *harmony*
controreazione (It.), *feedback*
conversio (L.), *inversion*
cor (Fr.), *horn*
corde (Fr.), *string*
crotchet (GB), *quarter note* (USA)
cuarta (Sp.), *fourth*
cuerda (Sp.), *string*

Dach (G.), *sounding board*
daher (G.), *hence*
Darstellung (G.), *representation*
demi-ton (Fr.), *semitone*
denarius (L.), *numbers* 1–10
diapason (Fr., It.), diapasón (Sp.), *pitch*
diapason (Gk.), *octave*
diapente (Gk.), *fifth*
diastema (Gk.), *interval*
diatessaron (Gk.), *fourth*
diazeuxis (Gk.), *separation of two tetrachords by a tone*
dièse (Fr.), diesis (It.), *sharp* (♯)
disdiapason (Gk.), *two octaves*
dodécaphonique (Fr.), *twelve tone*
Doppelbee (G.), *double flat* (♭♭)
Doppelkreuz (G.), *double sharp* (𝄪)
Dreiklang (G.), *triad*
Dur (G.), *major*
durchgehend (G.), *transient*

échantilloneur (Fr.), *sampler*
échelle (Fr.), *scale*
écouter (Fr.), *to hear*
égale (Fr.), *equal*
eighth note (USA), *quaver* (GB)
einfach (G.), *simple*
Einführung (G.), *introduction*
Einheit (G.), *unity*
Einklang (G.), *consonance*
Einselement (G.), *identity element*
emmeleia (Gk.), *consonance*
enmascaramiento (Sp.), *masking*
ensemble (Fr.), *set*
entier (Fr.), *integer*
entonación (Sp.), *intonation*
entsprechen (G.), *to correspond to*
epimoric, *ratio* $n+1:n$
erhöhen (G.), *to raise, increase*
erweitern (G.), *to extend, augment*
escala (Sp.), *scale*
espectro (Sp.), *spectrum*
estribo (Sp.), étrier (Fr.), *stapes*
étroit (Fr.), *narrow*

faux (Fr.), *false, out of tune*
feinte brisée (Fr.), *split key*
fistula (L.), *pipe, flute*
Folge (G.), *sequence, series*

gama (Sp.), gamma (It.), gamme (Fr.), *scale*
ganancia (Sp.), *gain*
ganze Note (G.), *whole note* (USA), *semibreve* (GB)
ganze Zahl (G.), *integer*
ganzer Ton (G.), *whole tone*
Gegenpunkt (G.), *counterpoint*
Geige (G.), *violin*
gerade (G.), *even, just, exactly*
Geräusch (G.), *noise*
Gesetz (G.), *law, rule*
giusto (It.), *just, precise*
gleichschwebende (G.), *equal beating*
gleichstufige (G.), *equal (temperament)*
Gleichung (G.), *equation*
gleichzeitig (G.), *simultaneous*
Glied (G.), *term*
Grundlage (G.), *foundation*
Grundton (G.), *fundamental*
guadagno (It.), *gain*

H (G.), *B (in German B denotes B♭)*
Halbton (G.), *semitone*
half note (USA), *minim* (GB)
hautbois (Fr.), *oboe*
hauteur (Fr.), *pitch*
helicon (Gk.), *instrument used for calculating ratios*
hemiolios (Gk.), *ratio* 3:2
Höhe (G.), *pitch*
Hörbar (G.), *audible*
Hören (G.), *hearing*

impair (Fr.), *odd*
inégale (Fr.), *unequal*

Kettenbruch (G.), *continued fractions*
Klang(farbe) (G.), *timbre*
Klangstufe (G.), *degree of scale*
Klappe (G.), *key (wind instruments)*

klein (G.), *small, minor*
Kombinationston (G.), *combination tone*
Komma (G.), *comma*
Kraft (G.), *energy*
Kreuz (G.), *sharp* (♯)

laud (Sp.), Laute (G.), *lute*
Leistung (G.), *power*
leiten (G.), *to derive, deduce*
Leiter (G.), *scale*
ley (Sp.), *law*
limaçon (Fr.), *cochlea*
llave (Sp.), *key (wind instruments)*
Lösung (G.), *solution*
loup (Fr.), *wolf*

maggiore (It.), majeur (Fr.), mayor (Sp.), *major*
marche d'harmonie (Fr.), *harmonic sequence*
Menge (G.), *set*
menor (Sp.), *minor*
mehrstimmig (G.), *polyphonic*
mesolabium, *mechanical means for producing ratio* 18:17, *approximation to equal tempered semitone for lutes*
mésotonique (Fr.), *meantone*
minim (GB), *half note* (USA)
minore (It.), *minor*
mitteltönig (G.), *meantone*
Moll (G.), *flat* (♭), *minor*
Mundstück (G.), *mouthpiece*
Musici, *followers of the Aristoxenian system of music, in which the ear is the judge of consonance, see also* Canonici
Muster (G.), *pattern*

Nachhall (G.), *reverberation*
Naturseptime (G.), *natural seventh*
Nebendreiklang (G.), *secondary triad (not* I, IV *or* V*)*
Nenner (G.), *denominator*
neuvième (Fr.), *ninth*
niveau (Fr.), *level*
nœud (Fr.), *node (vibration)*
None (G.), *nineth (interval)*
Notenschlussel (G.), *clef*
numérique (Fr.), *digital*

Oberwelle (G.), *harmonic*
offen (G.), *open*
Ohr (G.), *ear*
Ohrmuschel (G.), *auricle*
oído (Sp.), *ear*
onda (It., Sp.), *wave*
onda portante (It.), onda portadora (Sp.), *carrier*
onde (Fr.), *wave*
ordinateur (Fr.), *computer*
orecchio (It.), oreille (Fr.), *ear*
organo (It)., órgano (Sp.), Orgel (G.), orgue (Fr.), *organ*
ouïe (Fr.), *hearing; sound-hole*

padigione (It.), *auricle*
pair (Fr.), par (Sp.), *even*
paraphonia (Gk., L.), *Intervals of fourth and fifth*
parfait (Fr.), *perfect*
pavillon (Fr.), *auricle*
plagal cadence, *the cadence* IV–I
point d'orgue (Fr.), *fermata*
portée (Fr.), *staff, stave*
porteuse (Fr.), *carrier*
potencia (Sp.), potenza (It.), *power*
profondeur (Fr.), *depth*
puissance (Fr.), *power*
pulsaciones (Sp.), *beats*

Quadrat (G.), *natural* (♮)
quadrivium (L.), *The four disciplines: arithmetic, geometry, astronomy and music*
quarta (It., L.), quarte (Fr.), Quarte (G.), *fourth*
quarter note (USA), *crotchet* (GB)
quaternarius (L.), *numbers* 1–4
quaver (GB), *eighth note* (USA)
quinta (It., L., Sp.), quinte (Fr.), Quinte (G.), *fifth*

réaction (Fr.), *feedback*
reine (G.), *pure*
renversement (Fr.), *inversion*
résoudre (Fr.), *to resolve*
retard (Fr.), *delay*
retroalimentación (Sp.), *feedback*
ronde (Fr.), *whole note* (USA), *semibreve* (GB)
Rückkopplung (G.), *feedback*

Saite (G.), *string*
Satz (G.), *theorem; movement*
Schall (G.), *sound*
Scheibe (G.), *disc*
Schlag (G.), *beat*
Schlüssel (G.), *clef*
Schnecke (G.), *cochlea*
Schwebungen (G.), *beats*
Schwelle (G.), *threshold, limen*
Schwingungen (G.), *vibrations*
semibreve (GB), *whole note* (USA)
semiquaver (GB), *sixteenth note* (USA)
senarius (L.), *numbers* 1–6
sensible (Fr.), *leading note*
septenarius (L.), *numbers* 1–7
septième (Fr.), septima (L.), Septime (G.), *seventh*
Septimenakkord (G.), *seventh chord*
série de hauteurs (Fr.), *tone row*
sesquialtera (L.), *ratio* 3:2
sesquitertia (L.), *ratio* 4:3
settima (It.), *seventh*
seuil (Fr.), *threshold, limen*
Sext (G.), sexta (L.), *sixth*
sibilo (It.), *hiss*
siècle (Fr.), *century*
sifflement (Fr.), silbo (Sp.), *hiss*
sillet (Fr.), *bridge*
sixteenth note (USA), *semiquaver* (GB)
Skala (G.), *scale*
soglia (It.), *threshold, limen*
son (Fr.), *sound*
son combiné (Fr.), *combination tone*
son différentiel (Fr.), *difference tone*
sonido (Sp.), *sound*
sonido de combinación (Sp.), *combination tone*
sonorità (It.), *harmony, resonance*
sonus (L.), *sound*
sostenido (Sp.), *sharp* (♯)
spectre (Fr.), *spectrum*
staffa (It.), *stapes*
stanghetta (It.), *bar line*
stark (G.), *loud*
Stege (G.), *bridge*
Steigbügel (G.), *stapes*
Stimmstock (G.), *sound post*
Stimmung (G.), *tuning, key, pitch*
Stufe (G.), *scale degree*
subsemitonia (L.), *split keys*
suono (It.), *sound*
suono di combinazione (It.), *combination tone*
superparticular, *ratio* $n+1:n$
synaphe (Gk.), *conjunction, or overlapping of two tetrachords*
système incomplet (Fr.), just intonation

Takt (G.), *time, measure, bar*
Taktstrich (G.), *bar line*
tambour (Fr.), tamburo (It.), tambor (Sp.), *drum*
Tastame (It.), Tastatur, Tastenbrett, Tastenleiter (G.), Tastatura, Tastiera (It.), *keyboard of piano or organ*
tasto (It.), tecla (Sp.), *fret*
teilbar (G.), *divisible*
Teilmenge (G.), *subset*
Teilung (G.), *division*
Temperatur (G.), *temperament*
temperiert (G.), *tempered*
temps (Fr.), *time, beat, measure*
tercera (Sp.), tertia (L.), Terz (G.), terza (It.), *third*
tiempo (Sp.), *beat*
tierce (Fr.), *third*
ton (Fr.), *pitch, tone, key*
tonalité (Fr.), Tonart (G.), *key*
Tonausweichung (G.), *modulation*
Tonhöhe (G.), *pitch*
tono medio (It., Sp.), *meantone*
Tonschluss (G.), *cadence*
Tonstufe (G.), *scale degree*
touche (Fr.), *fret, key*
Träger (G.), *carrier*
traité (Fr.), *treatise*
tripla (L.), *ratio* 3:1
Trommel (G.), *drum*
tuyau (Fr.), *pipe*
tuyau à bouche (Fr.), *open pipe*
tuyau d'orgue (Fr.), *organ pipe*
tympan (Fr.), *eardrum*

überblasen (G.), *to overblow*
Übereinstimmung (G.), *consonance, harmony*
übermässig (G.), *augmented*
udibile (It.), *audible*

udito (It.), *hearing*
uguale (It.), *equal*
umbral (S.), *threshold, limen*
Umkehrung (G.), *inversion*
Unterdominant (G.), *subdominant*
Unterhalbton (G.), *leading note*
Unterleitton (G.), *dominant seventh*
Untergruppe (G.), *subgroup*
Untertaste (G.), *white key*

valeur propre (Fr.), *eigenvalue*
vent (Fr.), *wind*
Ventil (G.), ventile (It.), *valve (wind instruments)*
ventre (Fr.), *antinode (vibration)*
vents (Fr.), *wind instruments*
Verbindung (G.), *combination, union*
Verdeckung (G.), *masking*
vergleichen (G.), *to compare*
Verhältnis (G.), *ratio, proportion*
Verknüpfung (G.), *operation*
verlängertes Intervall (G.), *augmented interval*
vermindert (G.), *diminished*
versetzen (G.), *to transpose*
Versetzungszeichen (G.), *accidentals*
Verspätung (G.), *delay*
Verstärker (G.), *amplifier*
Verstärkung (G.), *gain*
verstimmt (G.), *out of tune*
verwandt (G.), *related*
Verzerrung (G.), *distortion*
Viertel (G.), *quarter*
voix (Fr.), *voice*
Vollkommenheit (G.), *perfection*

Welle (G.), *wave*
wenig (G.), *little, slightly*
whole note (USA), *semibreve* (GB)
wohltemperirte (G.), *well tempered*

Zahl (G.), *number*
Zählzeit (G.), *beat*
Zeichen (G.), *sign, note*
Zeit (G.), *time*
Zischen (G.), *hiss*
Zuklang (G.), *unison, consonance*

APPENDIX E

Equal tempered scales

q	p_3	e_3	p_5	e_5	p_7	e_7	e_{35}	e_{357}	$e_5 . q^2$	$e_{35} . q^{\frac{3}{2}}$	$e_{357} . q^{\frac{4}{3}}$
2	1	+213.686	1	−101.955	2	+231.174	166.245	190.365	392	470	480
3	1	+13.686	2	+98.045	2	−168.826	70.000	112.993	882	**364**	489
4	1	−86.314	2	−101.955	3	−68.826	94.459	86.760	1631	756	551
5	2	+93.686	3	+18.045	4	−8.826	67.464	55.319	451	754	473
6	2	+13.686	4	+98.045	5	+31.174	70.000	59.922	3530	1029	653
7	2	−43.457	4	−16.241	6	+59.746	32.804	43.672	796	608	585
8	3	+63.686	5	+48.045	6	−68.826	56.410	60.831	3075	1276	973
9	3	+13.686	5	−35.288	7	−35.493	26.764	23.104	2858	723	433
10	3	−26.314	6	+18.045	8	−8.826	22.561	19.113	1804	713	412
11	4	+50.050	6	−47.410	9	+12.992	48.748	40.503	5737	1778	991
12	4	+13.686	7	−1.955	10	+31.174	9.776	19.689	**282**	**406**	541
13	4	−17.083	8	+36.507	10	−45.749	28.500	35.202	6170	1336	1076
14	5	+42.258	8	−16.241	11	−25.969	32.012	30.132	3183	1677	1017
15	5	+13.686	9	+18.045	12	−8.826	16.015	14.034	4060	930	519
16	5	−11.314	9	−26.955	13	+6.174	20.671	17.250	6900	1323	695
17	5	−33.373	10	+3.927	14	+19.409	23.761	22.404	1135	1665	979
18	6	+13.686	11	+31.378	15	+31.174	24.207	26.732	10167	1849	1261
19	6	−7.366	11	−7.218	15	−23.457	7.293	13.745	2606	604	697
20	6	−26.314	12	+18.045	16	−8.826	22.561	19.113	7218	2018	1038
21	7	+13.686	12	−16.241	17	+2.603	15.018	12.354	7162	1445	716
22	7	−4.496	13	+7.136	18	+12.992	5.964	8.943	3454	615	551
31	10	+0.783	18	−5.181	25	−1.084	3.705	3.089	4979	639	**301**
41	13	−5.826	24	+0.484	33	−2.972	4.134	3.786	814	1085	535
53	17	−1.408	31	−0.068	43	+4.759	0.997	2.866	**192**	**385**	570
65	21	+1.379	38	−0.417	52	−8.826	1.018	5.163	1760	534	1349
68	22	+1.922	40	+3.927	55	+1.762	3.092	2.722	18160	1734	755
72	23	−2.980	42	−1.955	58	−2.159	2.520	2.406	10135	1540	721
84	27	−0.599	49	−1.955	68	+2.603	1.446	1.911	13794	1113	703
99	32	+1.565	58	+1.075	80	−0.871	1.343	1.206	10539	1323	552
118	38	+0.127	69	−0.260	95	−2.724	0.205	1.582	3621	**262**	915
130	42	+1.379	76	−0.417	105	+0.405	1.018	0.864	7040	1509	569
140	45	−0.599	82	+0.902	113	−0.254	0.766	0.642	17682	1269	467
171	55	−0.349	100	−0.201	138	−0.405	0.285	0.330	5866	636	**313**
441	142	+0.081	258	+0.086	356	−0.118	0.083	0.096	16689	772	**324**
494	159	−0.079	289	+0.069	399	+0.405	0.074	0.241	16909	815	943
612	197	−0.039	358	+0.006	494	−0.198	0.028	0.117	2166	**424**	607
665	214	−0.148	389	−0.0001	537	+0.197	0.105	0.142	**50**	1798	825

This table shows how well the scales based around equal divisions of the octave approximate the 5:4 major third, the 3:2 perfect fifth and the 7:4 seventh harmonic. The first column (q) gives the number of divisions to the octave. The second column (p_3) shows the scale degree closest to the 5:4 major third (counting from zero for the tonic), and the next column (e_3) shows the error in cents:

$$e_3 = 1200\left(\frac{p_3}{q} - \log_2\left(\frac{5}{4}\right)\right).$$

Similarly, the next two columns (p_5 and e_5) show the scale degree closest to the 3:2 perfect fifth and the error in cents:

$$e_5 = 1200\left(\frac{p_5}{q} - \log_2\left(\frac{3}{2}\right)\right).$$

The two columns after that (p_7 and e_7) show the scale degree closest to the 7:4 seventh harmonic and the error in cents:

$$e_7 = 1200\left(\frac{p_7}{q} - \log_2\left(\frac{7}{4}\right)\right).$$

We write e_{35} for the root mean square (RMS) error of the major third and perfect fifth:

$$e_{35} = \sqrt{(e_3^2 + e_5^2)/2}$$

and e_{357} for the RMS error for the major third, perfect fifth and seventh harmonic:

$$e_{357} = \sqrt{(e_3^2 + e_5^2 + e_7^2)/3}.$$

Theorem 6.2.3 shows that the quantity $e_5.q^2$ is a good measure of how well the perfect fifth is approximated by p_5/q of an octave, with respect to the number of notes in the scale. This theorem shows that there are infinitely many values of q for which $e_5.q^2 < 1200$, while on average we should expect this quantity to grow linearly with q.

Similarly, Theorem 6.2.5 with $k = 2$ shows that the quantity $e_{35}.q^{\frac{3}{2}}$ is a good measure of how well the major third and perfect fifth are simultaneously approximated, and shows that there are infinitely many values of q for which $e_{35}.q^{\frac{3}{2}} < 1200$, while on average we should expect this quantity to grow like the square root of q. Theorem 6.2.5 with $k = 3$ shows that the quantity $e_{357}.q^{\frac{4}{3}}$ is a good measure of how well all three intervals: major third, perfect fifth and seventh harmonic are simultaneously approximated, and shows that there are infinitely many values of q for which $e_{357}.q^{\frac{4}{3}} < 1200$, while on average we should expect this quantity to grow like the cube root of q.

Particularly good values of $e_5.q^2$, $e_{35}.q^{\frac{3}{2}}$ and $e_{357}.q^{\frac{4}{3}}$ are indicated in bold face in the last three columns of the table.

APPENDIX F

Frequency and MIDI chart

This table shows the frequencies and MIDI numbers of the notes in the standard equal tempered scale, based on the standard A4 = 440 Hz.

	MIDI	Hz	USA	Eur
piano ↑	108	4186.01	C8	c'''''
violin ↑	107	3951.07	B7	
	106	3729.31		
	105	3520.00	A7	
	104	3322.44		
	103	3135.96	G7	
	102	2959.96		
	101	2793.83	F7	
	100	2637.02	E7	
	99	2489.02		
flute ↑	98	2349.32	D7	
	97	2217.46		
	96	2093.00	C7	c''''
	95	1975.53	B6	
	94	1864.66		
	93	1760.00	A6	
	92	1661.22		
–	91	1567.98	G6	
	90	1479.98		
	89	1396.91	F6	
–	88	1318.51	E6	
leger	87	1244.51		
lines	86	1174.66	D6	
	85	1108.73		
–	84	1046.50	C6	c'''
	83	987.767	B5	
	82	932.328		
–	81	880.000	A5	
	80	830.609		
	79	783.991	G5	
	78	739.989		
┌	77	698.456	F5	
|	76	659.255	E5	
|	75	622.254		
⊢	74	587.330	D5	
|	73	554.365		
| treble	72	523.251	C5	c''
⊢ clef	71	493.883	B4	
|	70	466.164		
|	69	440.000	A4	
|	68	415.305		
⊢	67	391.995	G4	
|	66	369.994		
|	65	349.228	F4	
└	64	329.628	E4	
	63	311.127		
	62	293.665	D4	
	61	277.183		
middle c	60	261.626	C4	c'

	MIDI	Hz	USA	Eur
flute ↓	59	246.942	B3	
	58	233.082		
┌	57	220.000	A3	
|	56	207.652		
violin ↓	55	195.998	G3	
|	54	184.997		
⊢	53	174.614	F3	
|	52	164.814	E3	
| bass	51	155.563		
⊢ clef	50	146.832	D3	
|	49	138.591		
|	48	130.813	C3	c
⊢	47	123.471	B2	
|	46	116.541		
|	45	110.000	A2	
|	44	103.826		
└	43	97.9989	G2	
	42	92.4986		
	41	87.3071	F2	
–	40	82.4069	E2	
	39	77.7817		
	38	73.4162	D2	
	37	69.2957		
–	36	65.4064	C2	C
leger	35	61.7354	B1	
lines	34	58.2705		
–	33	55.0000	A1	
	32	51.9131		
	31	48.9994	G1	
	30	46.2493		
–	29	43.6535	F1	
	28	41.2034	E1	
	27	38.8909		
	26	36.7081	D1	
	25	34.6478		
	24	32.7032	C1	C_1
	23	30.8677	B0	
	22	29.1352		
piano ↓	21	27.5000	A0	
	20	25.9565		
	19	24.4997	G0	
	18	23.1247		
	17	21.8268	F0	
	16	20.6017	E0	
	15	19.4454		
	14	18.3540	D0	
	13	17.3239		
	12	16.3516	C0	C_2
	11	15.4339		

APPENDIX G

Getting stuff from the internet

This appendix is about software and other resources which may be found online. The information is, of course, very volatile. So it is likely that by the time you are reading this, a lot of the information will already be out of date.

Scales and Temperaments: The best internet resource on the subject of scales, temperaments and tunings is

www.xs4all.nl/∼huygensf/doc/bib.html

This is part of the Huygens-Fokker Foundation website, maintained by Manuel Op de Coul, and consists of a giant bibliography together with links to other internet resources on the subject. The front page of the website is at

www.xs4all.nl/∼huygensf/english/

Also on the same website, a discography of microtonal music can be found at

www.xs4all.nl/∼huygensf/doc/discs.html

A large collection of scales and temperaments can be found at

www.xs4all.nl/∼huygensf/doc/scales.zip

and the Scala scales and temperaments software can be found at

www.xs4all.nl/∼huygensf/scala/

To subscribe to the alternate tunings email discussion group, send an empty email message to tuning-subscribe@onelist.com.

Just Intonation Network: www.dnai.com/∼jinetwk/

Bohlen–Pierce scale: members.aol.com/bpsite/index.html

Music Theory: Sites offering free music theory tuition online include

Easy Music Theory (Gary Ewer): www.musictheory.halifax.ns.ca/

Java Music Theory: academics.hamilton.edu/music/spellman/JavaMusic/

Online Music Instruction Page (Ken Fansler):
orathost.cfa.ilstu.edu/∼kwfansle/onlinemusicpage.htm

Practical Music Theory: www.teoria.com/java/eng/java.htm

Sound editors: There are some good shareware sound editors. Among the best are:

Cool Edit: www.syntrillium.com/cooledit/index.html

Goldwave: www.goldwave.com/

Acid Wav: www.polyhedric.com/software/acid/

There is a freeware audio frequency analyzer (complete with C++ source-code) for the PC called

Sound Frequency Analyzer: www.relisoft.com/freeware/index.htm

CSound: This free software is described in §8.10. Versions for various platforms (PC, Mac, Unix, Atari, NeXT) are available from

ftp://ftp.maths.bath.ac.uk/pub/dream/

To subscribe to the email discussion group for CSound, send an empty message to csound-subscribe@lists.bath.ac.uk. Further information about CSound can be found at the following www pages:

www.mitpress.com/e-books/csound/frontpage.html
(the CSound front page, MIT Press)

www.bright.net/~dlphilp/dp_csound.html
(Dave Phillips' PC CSound page)

www.bright.net/~dlphilp/linux_csound.html
(Dave Phillips' Linux CSound page)

music.dartmouth.edu/~dupras/wCsound/csoundpage.html
(Martin Dupras' CSound page)

A utility for PC and Unix called MIDI2CS, written by Rudiger Borrmann, converts MIDI files to Csound scores. It is available from

www.snafu.de/~rubo/songlab/midi2cs/csound.html

A utility for emulating the Yamaha DX7 with CSound can be found at Jeff Harrington's site

www.parnasse.com/dx72csnd.shtml

Other synthesis software: This is a rapidly expanding field, and new products turn up almost every week. The free ones I know of are as follows.

CLM (Common Lisp Music, freeware):
www-ccrma.stanford.edu/CCRMA/Software/clm/clm.html

CMix (Next, Linux, Sparc, SGI, PowerMac; freeware):
www.music.princeton.edu/winham/cmix.html

Nyquist (freeware):
www.cs.cmu.edu/afs/cs.cmu.edu/project/music/web/music.software.html

SoundForum Synthesizer (freeware, Mac/Windows):
www.keyboardmag.com/features/soundforum/index.shtml

SynFactory (nice freeware Windows synthesizer):
www.syntiac.com/synfactory.html

Synthesis Toolkit (C++ code):
www-ccrma.stanford.edu/CCRMA/Software/STK/

Synthesizers and patches: The best general websites for synthesizers and patches are

Synthesizer and Midi Links Page:
www.interlog.com/~spinner/lbquirke/synthesis/links/

Synth Site: www.sonicstate.com/bbsonic/synth/index.cfm

At the anonymous ftp site ftp.ucsd.edu, in the subdirectory /midi/patches, there are patches for Casio CZ-1, CZ-2, Ensoniq ESQ1, SQ1, Kawai K1, K4, K5, XD-5, Korg M1, T3, WS (Wavestation), Roland D10, D5, D50, D70, SC55, U20, and Yamaha DX7, FB01, TX81Z, SY22, SY55, SY77, SY85.

For the Yamaha DX7, there is a web page which I maintain at

www.math.uga.edu/~djb/dx7.html

which contains, among other things, a patch archive and instructions for joining the email discussion group.

Typesetting software:

CMN (Common Music Notation, freeware for NeXT and SGI machines):
ccrma-www.stanford.edu/CCRMA/Software/cmn/cmn.html

Finale is a commercial music notation package for the Mac and Windows (current version Finale 2002), and is available from Coda Music Software. Their web site

www.codamusic.com/

has more information. A free demonstration version of the program is available on this web site. Without academic discount, Finale is very expensive, but with academic discount it costs about $200–$250. To subscribe to the email discussion group for Finale, send an email message to listserv@shsu.edu with the phrase "subscribe Finale" or "subscribe Finale-Digest" in the body of the message. To be removed from the list, send "signoff Finale" or "signoff Finale-Digest" to the same address.

Finale forum (not sanctioned by Coda Music): www.cmp.net/finale/

Finale Resource Page: www.peabody.jhu.edu/~skot/finale/fin_home.html

Ftp site for Finale users: ftp://ftp.shsu.edu/pub/finale/

Keynote is a public domain textual, graphical and algorithmic music editor for the Unix X Window system, the Mac or the Amiga, available from

ftp://xcf.berkeley.edu

LilyPond is a GNU project (and hence free) music typesetter for Unix systems. It is available from

www.cs.uu.nl/~hanwen/lilypond/index.html

Lime (Mac, Windows): www.cerlsoundgroup.org/

Mozart: www.mozart.co.uk/

Muzika 3 is a public domain (freeware) music notation package for Windows, available from

ftp://garbo.uwasa.fi/windows/sound/muzika3.zip or from

ftp://ftp.cica.indiana.edu/ftp/pub/win3/sounds/muzika3.zip

Nutation (NeXT, freeware): ftp://ccrma-ftp.stanford.edu/pub/Nu.pkg.tar

Overture is a Mac based commercial music notation package.

Score: ace.acadiau.ca/score/links3.htm

Sibelius is a notation package for the PC: www.sibelius.com/

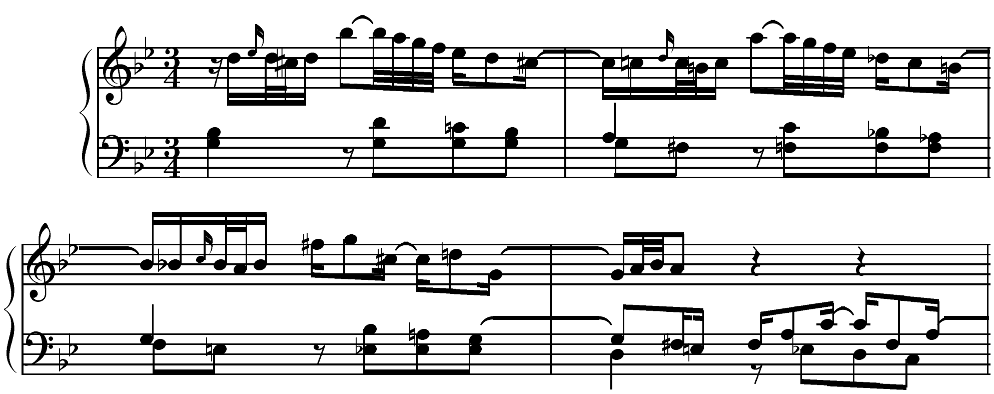
MusicTEX: MusicTEX, written by the french organist Daniel Taupin, and its successor MusixTEX are public domain music typesetting packages to run under Donald Knuth's TEX program. The necessary files may be found on

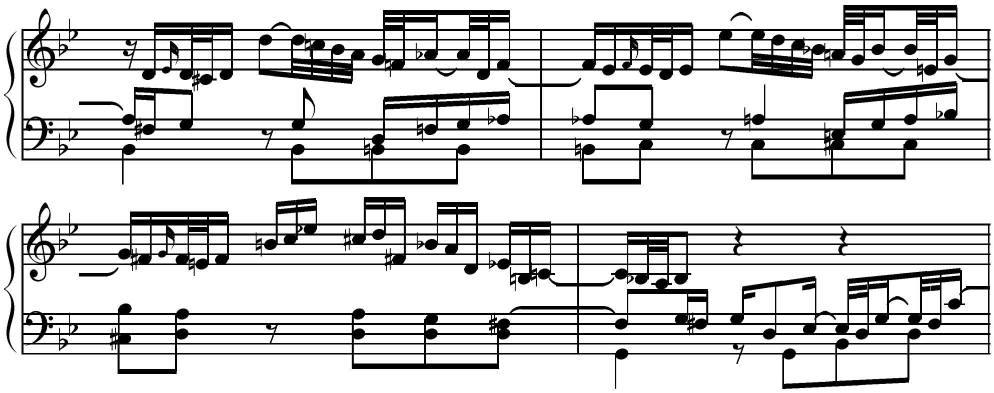
ftp://rsovax.ups.circe.fr/TeX/musictex/

See also: www.gmd.de/Misc/Music/

A public domain version of TEX for Windows 95 or higher, called MikTeX, and can be found at www.miktex.de. Versions for all platforms are available from CTAN at ftp.tex.ac.uk, ftp.dante.de or ctan.tug.org. See also TUG (the TEX user's group) at tug.org.

Goldberg Variation 25, J. S. Bach





Example of Output from MusicTEX

MuTEX is the precursor of MusicTEX, written by Andrea Steinbach and Angelica Schofer. It is in the public domain, and is available by anonymous ftp from ymir.claremont.edu in [anonymous.tex.music.mtex] (VMS).

MIDI2TEX is a program written by Hans Kuykens for converting MIDI files into MusicTEX files. The latest version can be found on CTAN (see page 359).

ABC2MTEX is a program for converting tunes from its own text-based format into MusicTEX files. It is designed primarily for folk and traditional music of Western European origin written on one stave in standard classical notation. It can be obtained directly from its author, Chris Walshaw, via email: C.Walshaw@gre.ac.uk, or from

ftp://celtic.stanford.edu/pub/tunes/abc2mtex/

Sequencers: Cakewalk and Cubase are competing commercial Windows based sequencers, neither of which is cheap, but both of which are packed with features. To subscribe to the Cakewalk users' group, send a message to listserv@lists.colorado.edu with the phrase "subscribe cakewalk" in the body of the message. To subscribe to the Cubase users' group, send a message to cubase-users-request@nessie.mcc.ac.uk. Messages for the group should be sent to cubase-users@mcc.ac.uk.

Power Tracks Pro Audio is a very cheap, but fully functional commercial Windows based sequencer, available from PG Music for $29. More information can be found at

www.pgmusic.com/

Rosegarden is an integrated MIDI sequencer and musical notation editor. It is free software for Unix and X, and it may be found at

www.bath.ac.uk/~masjpf/rose.html

WinJammer is a shareware Windows based sequencer, which may be found at

ftp://ftp.cnr.it/pub/msdos/win3/sounds/wjmr23.zip

WinJammer Pro (I'm not sure what the difference is) is in the same directory, as wjpro.zip.

Random music: There are a number of freeware/shareware probabilistic music programs designed to run under Windows.

Aleatoric composer (shareware):
ftp://oak.oakland.edu/msdos/music/alcomp11.zip

Art Song 4.5 (shareware): www.artsong.org/

FMusic 1.9 (freeware): www.fractal-vibes.com/fm/index.html

FractMus 2.3 (freeware): ftp://ftp.cdrom.com/pub/win95/music/frctmu25.zip

Fractal Tune Smithy (freeware/shareware):
matrix.crosswinds.net/∼fractalmelody/index.htm

Improvise 1.2 (shareware):
ftp://ftp.cnr.it/pub/msdos/win3/sounds/impvz120.zip

Make-Prime-Music (freeware):
members.tripod.de/Latrodectus98/index.html

Mandelbrot Music (freeware): www.fin.ne.jp/∼yokubota/mandele.shtml

MusiNum 2.08 (freeware):
www.forwiss.uni-erlangen.de/∼kinderma/musinum/musinum.html

QuasiFractalComposer 2.01 (freeware):
members.tripod.com/∼paulwhalley/

Tangent (free/shareware): www.randomtunes.com/

The Well Tempered Fractal 3.0 (freeware):
www-ks.rus.uni-stuttgart.de/people/schulz/fmusic/wtf/wtf30.zip

MIDI: The MIDI specification can be obtained via email by sending a message with the phrase GET MIDISPEC PACKAGE in the message body, to listserv@auvm.american.edu. There are archives of MIDI files available at

ftp://ftp.cs.ruu.nl/MIDI/DOC/archives/

ftp://ftp.waldorf-gmbh.de/pub/midi/

There are two programs called mf2t and t2mf which convert standard MIDI files into human readable ASCII text and back again. The MIDI home page on the WWW is

www.eeb.ele.tue.nl/midi/index.html

A good starting point for information about MIDI is the Northwestern University site

nuinfo.nwu.edu/musicschool/links/projects/midi/expmidiindex.html

Academic Computer Music: The following departments in American universities have programs in computer music. CalArts (David Rosenboom, Morton Subotnick), Carnegie Mellon (Roger Dannenberg), MIT (Tod Machover, Barry Vercoe), Princeton (Paul Lansky), Stanford (John Chowning, Chris Chaffe, Perry Cook, etc.), SUNY Buffalo (David Felder, Cort Lippe), UC Berkeley (David Wessel), UCSD (Miller Puckett, F. Richard Moore, George Lewis, Peter Otto).

IRCAM is an institution in Paris for computer music, which has an anonymous ftp site at ftp.ircam.fr. In particular, the music/programming environment MAX can be found there.

Music Theory Online (the Online Journal of the Society for Music Theory) can be found at

boethius.music.ucsb.edu/mto/mtohome.html

Other resources:

Everyone seems to want to know more about the infamous "Mozart effect." Volume VII, Issue 1 (Winter 2000) of MuSICA Research Notes is devoted to this much overpublicized and misunderstood topic, and can be found at

www.musica.uci.edu/mm/V7I1W00.html

Online papers: See Appendix O for a selection of relevant papers which can be downloaded from academic journals.

APPENDIX I

Intervals

This is a table of intervals not exceeding one octave (or a tritave in the case of the Bohlen–Pierce, or BP scale). A much more extensive table may be found in Appendix XX to Helmholtz [**54**] (page 453), which was added by the translator, Alexander Ellis. Names of notes in the BP scale are denoted with a subscript BP, to save confusion with notes which may have the same name in the octave based scale.

The first column is equal to 1200 times the logarithm to base two of the ratio given in the second column. Logarithms to base two can be calculated by taking the natural logarithm and dividing by $\ln 2$. So the first column is equal to

$$\frac{1200}{\ln 2} \approx 1731.234$$

times the natural logarithm of the second column.

We have given all intervals to three decimal places for theoretical purposes. While intervals of less than a few cents are imperceptible to the human ear in a melodic context, in harmony very small changes can cause large changes in beats and roughness of chords. Three decimal places gives great enough accuracy that errors accumulated over several calculations should not give rise to perceptible discrepancies.

If more accuracy is needed, I recommend using the multiple precision package bc (see page 343) with the -l option. The following lines can be made into a file to define some standard intervals in cents. For example, if the file is called music.bc then the command "bc -l music.bc" will load them at startup.

```
scale=50 /* fifty decimal places - seems like plenty but you never know */
octave=1200
savart=1.2*l(10)/l(2)
syntoniccomma=octave*l(81/80)/l(2)
pythagoreancomma=octave*l(3^12/2^19)/l(2)
septimalcomma=octave*l(64/63)/l(2)
schisma=pythagoreancomma-syntoniccomma
diaschisma=syntoniccomma-schisma
perfectfifth=octave*l(3/2)/l(2)
equalfifth=700
meantonefifth=octave*l(5)/(4*l(2))
perfectfourth=octave*l(4/3)/l(2)
justmajorthird=octave*l(5/4)/l(2)
justminorthird=octave*l(6/5)/l(2)
justmajortone=octave*l(9/8)/l(2)
justminortone=octave*l(10/9)/l(2)
```

Cents	Interval ratio	Eitz	Name, etc.	Ref
0.000	1:1	C^0, C^0_{BP}	Fundamental	§4.1
1.000	$2^{\frac{1}{1200}}$:1		Cent	§5.4
1.805	$2^{\frac{1}{665}}$:1		Degree of 665 tone scale	§6.4
1.953	32805:32768	$B\sharp^{-1}$	Schisma	§5.8
3.986	$10^{\frac{1}{1000}}$:1		Savart	§5.4
14.191	245:243	C^{+1}_{BP}	BP-minor diesis	§6.7
19.553	2048:2025	$D\flat\flat^{+2}$	Diaschisma	§5.8
21.506	81:80	C^{+1}	Syntonic, or ordinary comma	§5.5
22.642	$2^{\frac{1}{53}}$:1		Degree of 53 tone scale	§6.3
23.460	3^{12}:2^{19}	$B\sharp^0$	Pythagorean comma	§5.2
27.264	64:63		Septimal comma	§5.8
35.099			Carlos' γ scale degree	§6.6
41.059	128:125	$D\flat\flat^{+3}$	Great diesis	§5.12
49.772	7^{13}:3^{23}	$D\flat\flat^0_{BP}$	BP 7/3 comma	§6.7
63.833			Carlos' β scale degree	§6.6
70.672	25:24	$C\sharp^{-2}$	Small (just) semitone	§5.5
77.965			Carlos' α scale degree	§6.6
90.225	256:243	$D\flat^0$	Diesis or Limma	§5.2
100.000	$2^{\frac{1}{12}}$:1	$\approx C\sharp^{-\frac{7}{11}}$	Equal semitone	§5.14
111.731	16:15	$D\flat^{+1}$	Just minor semitone (ti–do, mi–fa)	§5.5
113.685	2187:2048	$C\sharp^0$	Pythagorean apotomē	§5.2
133.238	27:25	$D\flat^{-2}_{BP}$		§6.7
146.304	$3^{\frac{1}{13}}$:1		BP-equal semitone	§6.7
182.404	10:9	D^{-1}	Just minor tone (re–mi, so–la)	§5.5
193.157	$\sqrt{5}$:2	$D^{-\frac{1}{2}}$	Meantone whole tone	§5.12
200.000	$2^{\frac{1}{6}}$:1	$\approx D^{-\frac{2}{11}}$	Equal whole tone	§5.14
203.910	9:8	D^0	Just major tone (do–re, fa–so, la–ti);	§5.5
			Pythagorean major tone;	§5.2
			Nineth harmonic	§4.1
294.135	32:27	$E\flat^0$	Pythagorean minor third	§5.2
300.000	$2^{\frac{1}{4}}$:1	$\approx E\flat^{+\frac{3}{11}}$	Equal minor third	§5.14
315.641	6:5	$E\flat^{+1}$	Just minor third (mi–so, la–do, ti–re)	§5.5
386.314	5:4	E^{-1}	Just major third (do–mi, fa–la, so–ti);	§5.5
			Meantone major third;	§5.12
			Fifth harmonic	§4.1
400.000	$2^{\frac{1}{3}}$:1	$\approx E^{-\frac{4}{11}}$	Equal major third	§5.14
407.820	81:64	E^0	Pythagorean major third	§5.2
498.045	4:3	F^0	Perfect fourth	§5.2

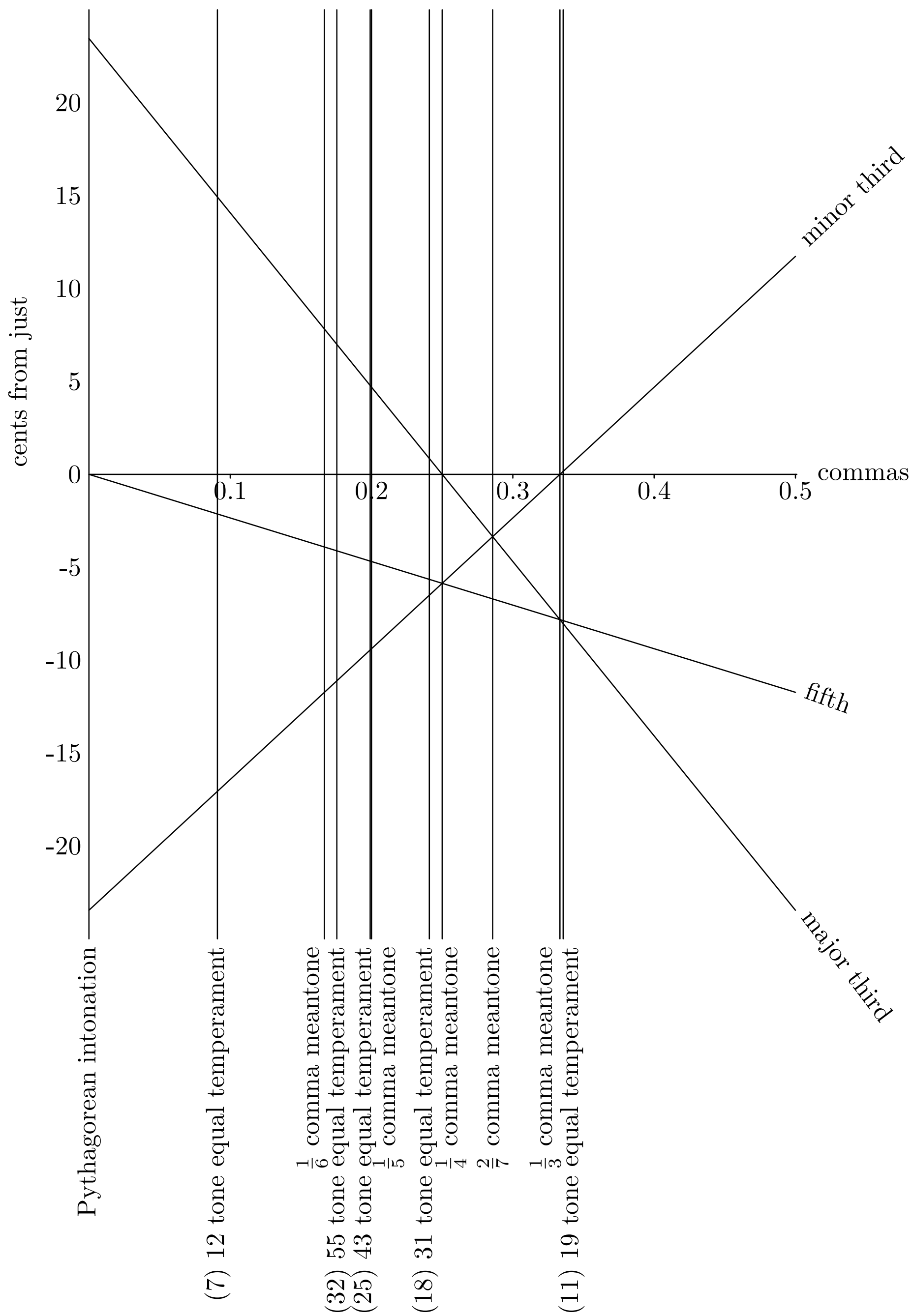
Cents	Interval ratio	Eitz	Name, etc.	Ref
500.000	$2^{\frac{5}{12}}:1$	$\approx \mathrm{F}^{+\frac{1}{11}}$	Equal fourth	§5.14
503.422	$2:5^{\frac{1}{4}}$	$\mathrm{F}^{+\frac{1}{4}}$	Meantone fourth	§5.12
551.318	11:8		Eleventh harmonic	§4.1
600.000	$\sqrt{2}:1$	$\approx \mathrm{F}\sharp^{-\frac{6}{11}}$	Equal tritone	§5.14
611.731	729:512	$\mathrm{F}\sharp^{0}$	Pythagorean tritone	§5.2
696.579	$5^{\frac{1}{4}}:1$	$\mathrm{G}^{-\frac{1}{4}}$	Meantone fifth	§5.12
700.000	$2^{\frac{7}{12}}:1$	$\approx \mathrm{G}^{-\frac{1}{11}}$	Equal fifth	§5.14
701.955	3:2	G^{0}	Just and Pythagorean (perfect) fifth;	§5.2
			Third harmonic	§4.1
792.180	128:81	$\mathrm{A}\flat^{0}$	Pythagorean minor sixth	§5.2
800.000	$2^{\frac{2}{3}}:1$	$\approx \mathrm{A}\flat^{+\frac{4}{11}}$	Equal minor sixth	§5.14
813.687	8:5	$\mathrm{A}\flat^{+1}$	Just minor sixth	§5.5
840.528	13:8		Thirteenth harmonic	§4.1
884.359	5:3	A^{-1}	Just major sixth	§5.5
889.735	$5^{\frac{3}{4}}:2$	$\mathrm{A}^{-\frac{3}{4}}$	Meantone major sixth	§5.12
900.000	$2^{\frac{3}{4}}:1$	$\approx \mathrm{A}^{-\frac{3}{11}}$	Equal major sixth	§5.14
905.865	27:16	A^{0}	Pythagorean major sixth	§5.2
968.826	7:4		Seventh harmonic	§4.1
996.091	16:9	$\mathrm{B}\flat^{0}$	Pythagorean minor seventh	§5.2
1000.000	$2^{\frac{5}{6}}:1$	$\approx \mathrm{B}\flat^{+\frac{2}{11}}$	Equal minor seventh	§5.14
1082.892	$5^{\frac{5}{4}}:4$	$\mathrm{B}^{-\frac{5}{4}}$	Meantone major seventh	§5.12
1088.269	15:8	B^{-1}	Just major seventh;	§5.5
			Fifteenth harmonic	§4.1
1100.000	$2^{\frac{11}{12}}:1$	$\approx \mathrm{B}^{-\frac{5}{11}}$	Equal major seventh	§5.14
1109.775	243:128	B^{0}	Pythagorean major seventh	§5.2
1200.000	2:1	C^{0}	Octave; Second harmonic	§4.1
1466.871	7:3	$\mathrm{A}^{0}_{\mathrm{BP}}$	BP-tenth	§6.7
1901.955	3:1	$\mathrm{C}^{0}_{\mathrm{BP}}$	BP-Tritave	§6.7

APPENDIX J

Just, equal and meantone scales compared

The figure on the next page has its horizontal axis measured in multiples of the (syntonic) comma, and the vertical axis measured in cents. Each vertical line represents a regular scale, generated by its fifth. The size of the fifth in the scale is equal to the Pythagorean fifth (ratio of 3:2, or 701.955 cents) minus the multiple of the comma given by the position along the horizontal axis. The three sloping lines show how far from the just values the fifth, major third and minor third are in these scales. This figure is relevant to Exercise **2** in §6.4.

It is worth noting that if $\frac{1}{11}$ comma meantone were drawn on this diagram, it would be indistinguishable from 12 tone equal temperament; see §5.14.



Regular scales and their deviations from just intonation

APPENDIX L

Logarithms

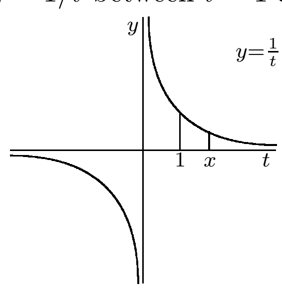
The purpose of this appendix is to give a quick review of the definition and standard properties of logarithms, since they are so important to the theory of scales and temperaments. A commonly used definition of logarithm is that $b = \log_a(c)$ means the same as $a^b = c$.

The main problem in understanding the above definition is understanding what the notation a^b means. If b is rational, this can be explained in terms of multiplication and extraction of roots. But what on earth does 2^π mean? How do we multiply 2 by itself π times? It turns out that logically, the easiest way to develop exponentials and logarithms begins with the logarithm as a definite integral and proceeds in the reverse of the order in which these concepts are usually learned.

The definition of the natural logarithm is

$$\ln(x) = \int_1^x \frac{1}{t}\,dt,$$

which makes sense provided $x > 0$. In other words, $\ln(x)$ is the area under the graph of the function $y = 1/t$ between $t = 1$ and $t = x$.



According to the usual conventions of calculus, if x lies between zero and one, this area is interpreted as negative, while for $x > 1$ it is positive. It is immediately apparent from the definition that

$$\ln(1) = 0.$$

The fundamental theorem of calculus implies that

$$\frac{d}{dx}\ln(x) = \frac{1}{x}.$$

Applying the chain rule, if a is a constant then

$$\frac{d}{dx}\ln(ax) = \frac{a}{ax} = \frac{1}{x}.$$

One of the consequences of the mean value theorem is that two functions with the same derivative differ by a constant. We apply this to $\ln(ax)$ and $\ln(x)$, and find out the value of the constant by setting $x = 1$, to get $\ln(ax) - \ln(x) = \ln(a) - \ln(0) = \ln(a)$. If b is another constant, then evaluating at $x = b$ gives

$$\ln(ab) = \ln(a) + \ln(b).$$

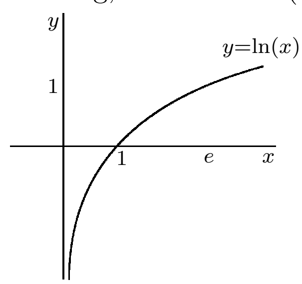
The particular case where $a = 1/b$ gives us

$$\ln(1/b) = -\ln(b).$$

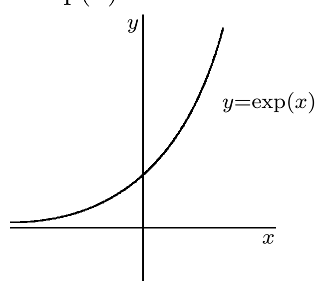
Combining these formulas gives

$$\ln(a/b) = \ln(a) - \ln(b).$$

From these properties and the definition, it easily follows that the logarithm function is monotonically increasing, with domain $(0, \infty)$ and range $(-\infty, \infty)$.



The *exponential function* $\exp(x)$ is defined to be the inverse function of $\ln(x)$. In other words, $y = \exp(x)$ means the same as $x = \ln(y)$.



So the area under the graph of $y = 1/t$ between $t = 1$ and $t = \exp(x)$ is equal to x. The above properties of the logarithm translate into the following properties of the exponential function:

$$\begin{aligned}
&\exp(0) = 1 \\
&\exp(a + b) = \exp(a)\exp(b) \\
&\exp(-b) = 1/\exp(b) \\
&\exp(a - b) = \exp(a)/\exp(b).
\end{aligned}$$

The number e is defined to be $\exp(1)$, and it is an irrational number whose approximate value is 2.71828. The domain of the exponential function is $(-\infty, \infty)$, and its range is $(0, \infty)$.

We define a^b to mean $\exp(b\ln(a))$ $(a>0)$. So the area under the graph of $y=1/t$ between $t=1$ and $t=a^b$ is exactly b times as big as the area between $t=1$ and $t=a$. It follows immediately from this definition that

$$\ln(a^b)=b\ln(a) \qquad (a>0).$$

If $b=m/n$ is rational, it is not hard to check using the above properties of the exponential and logarithm function that this definition agrees with the more usual one with powers and roots ($a^{m/n}$ is the unique positive number whose nth power equals the mth power of a). But this definition gets us around the problem of trying to understand what it means to multiply a by itself an irrational number of times! Thus for example

$$e^x=\exp(x\ln(e))=\exp(x)$$

so that the exponential function can be written as e^x. With these definitions, it is easy to prove the usual laws of indices:

$$a^0=1, \qquad a^1=a, \qquad a^{-1}=1/a, \qquad a^{-b}=1/a^b, \qquad a^{b+c}=a^ba^c,$$
$$a^{b-c}=a^b/a^c, \qquad a^cb^c=(ab)^c, \qquad (a^b)^c=a^{bc}, \qquad a^{\frac{1}{n}}=\sqrt[n]{a}$$

We define

$$\log_a(b)=\frac{\ln(b)}{\ln(a)} \qquad (a>0,\ b>0).$$

Thus $c=\log_a(b)$ is equivalent to $c\ln(a)=\ln(b)$, or $\exp(c\ln(a))=b$, or $a^c=b$. So $c=\log_a(b)$ means that c is the power to which a has to be raised to obtain b. For example, $\log_e(b)$ is the same as $\ln(b)$, the natural logarithm of b, because $\ln(e)=1$.

If we write out what it means for the derivative of $\ln(t)$ to be $\frac{1}{t}$, we get

$$\frac{1}{t}=\lim_{h\to 0}\frac{\ln(t+h)-\ln(t)}{h}=\lim_{h\to 0}\ln\left(\frac{t+h}{t}\right)^{\frac{1}{h}}.$$

The exponential function is continuous, so we can exponentiate both sides to get

$$e^{\frac{1}{t}}=\lim_{h\to 0}\left(\frac{t+h}{t}\right)^{\frac{1}{h}}.$$

Substituting x for $1/t$ and n for $1/h$, we get

$$\boxed{e^x=\lim_{n\to\infty}(1+\tfrac{x}{n})^n.}$$

Expand out using Pascal's triangle to get

$$\begin{aligned}
e^x &= \lim_{n\to\infty}(1+n\tfrac{x}{n}+\tfrac{n(n-1)}{2!}\tfrac{x^2}{n^2}+\tfrac{n(n-1)(n-2)}{3!}\tfrac{x^3}{n^3}+\cdots)\\
&= \lim_{n\to\infty}(1+x+(1-\tfrac{1}{n})\tfrac{x^2}{2!}+(1-\tfrac{1}{n})(1-\tfrac{2}{n})\tfrac{x^3}{3!}+\cdots)\\
&= 1+x+\tfrac{x^2}{2!}+\tfrac{x^3}{3!}+\cdots
\end{aligned}$$

In particular, putting $x = 1$ gives

$$e = 1 + 1 + \tfrac{1}{2!} + \tfrac{1}{3!} + \cdots = 2.71828\ldots$$

The scale of cents in music theory is defined in such a way that a frequency ratio of $f{:}1$ is represented as an interval of

$$1200 \log_2(f) \text{ cents} = \frac{1200 \ln(f)}{\ln(2)} \text{ cents.}$$

Thus one octave, or a frequency ratio of 2:1, is an interval of 1200 cents. In the 12 tone equal tempered scale, this is divided into 12 equal semitones of 100 cents each. For more details, see §5.4.

The scale of decibels (dB) for loudness is also logarithmic. Adding 10 decibels multiplies the signal power by 10. So an acoustic signal power ratio of $b{:}1$ is represented as a difference of

$$10 \log_{10}(b) \text{ dB} = \frac{10 \ln(b)}{\ln(10)} \text{ dB.}$$

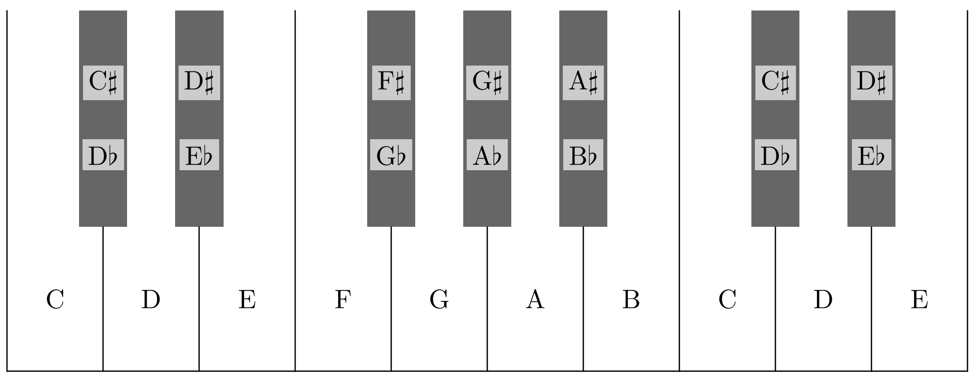
Since power is proportional to the square of amplitude, an acoustic signal amplitude ratio of $a{:}1$ is represented by a difference of

$$10 \log_{10}(a^2) \text{ dB} = 20 \log_{10}(a) \text{ dB} = \frac{20 \ln(a)}{\ln(10)} \text{ dB.}$$

APPENDIX M

Music theory

This appendix consists of the background in elementary music theory needed to understand the main text. The emphasis is slightly different than that of a standard music text. We begin with the piano keyboard, as a convenient way to represent the modern scale (see also Appendix F).



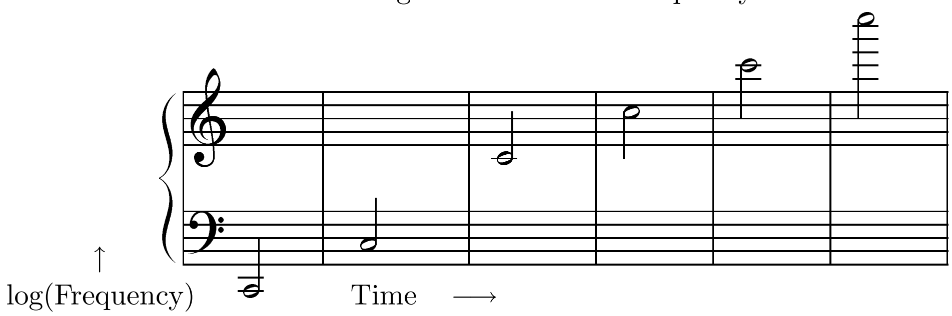
Both the black and the white keys represent notes. This keyboard is periodic in the horizontal direction, in the sense that it repeats after seven white notes and five black notes. The period is one *octave*, which represents doubling the frequency corresponding to the note. The principle of *octave equivalence* says that notes differing by a whole number of octaves are regarded as playing equivalent roles in harmony. In practice, this is almost but not quite completely true.

On a modern keyboard, each of the twelve intervals making up an octave represents the same frequency ratio, called a *semitone*. The name comes from the fact that two semitones make a *tone*. The twelfth power of the semitone's frequency ratio is a factor of 2:1, so a semitone represents a frequency ratio of $2^{\frac{1}{12}}$:1. The arrangement where all the semitones are equal in this way is called *equal temperament*. Frequency is an exponential function of position on the keyboard, and so the keyboard is really a *logarithmic* representation of frequency.

Because of this logarithmic scale, we talk about *adding* intervals when we want to *multiply* the frequency ratios. So when we add a semitone to another semitone, for example, we get a tone with a frequency ratio of $2^{\frac{1}{12}}\times 2^{\frac{1}{12}}$:1

or $2^{\frac{1}{6}}$:1. This transition between additive and multiplicative notation can be a source of great confusion.

Staff notation works in a similar way, except that the logarithmic frequency is represented vertically, and the horizontal direction represents time. So music notation paper can be regarded as graph paper with a linear horizontal time axis and a logarithmic vertical frequency axis.

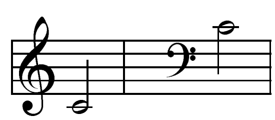


In the above diagram, each note is twice the frequency of the previous one, so they are equally spaced on the logarithmic frequency scale (except for the break between the bass and treble clefs). The gap between adjacent notes is one octave, so the gap between the lowest and highest note is described *additively* as five octaves, representing a *multiplicative* frequency ratio of 2^5:1.

There are two clefs on this diagram. The upper one is called the *treble clef*, with lines representing the notes E, G, B, D, F, beginning with the E two white notes above middle C and working up the lines. The spaces between them represent the notes F, A, C, E between them, so that this takes care of all the white notes between the E above middle C and the F an octave and a semitone above that. The black notes are represented in by using the line or space with the likewise lettered white note with a sharp ($\sharp$) or flat ($\flat$) sign in front.

The lower clef is called the bass clef, with lines representing the notes G, B, D, F, A, with the last note representing the A two white notes below middle C and the first note representing the G an octave and a tone below that.

Middle C itself is represented using a *leger line*, either below the treble clef or above the bass clef.



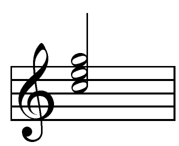
The frequency ratio represented by seven semitones, for example the interval from C to the G above it, is called a *perfect fifth*. Well, actually, this isn't quite true. A perfect fifth is supposed to be a frequency ratio of 3:2, or 1.5:1, whereas seven semitones on our modern equal tempered scale produce a frequency ratio of $2^{\frac{7}{12}}$:1 or roughly 1.4983:1. The perfect fifth is a consonant interval, just as the octave is, for reasons described in Chapter 4. So seven semitones is very close to a consonant interval. It is very difficult to discern the difference between a perfect fifth and an equal tempered fifth except by listening for beats; the difference is about one fiftieth of a semitone.

The perfect fourth represents the interval of 4:3, which is also consonant. The difference between a perfect fourth and the equal tempered fourth of five semitones is exactly the same as the difference between the perfect fifth and the equal tempered fifth, because they are obtained from the corresponding versions of a fifth by subtracting from an octave.

The frequency ratio represented by four semitones, for example the interval from C to the E above it, is called a *major third*. This represents a frequency ratio of $2^{\frac{4}{12}}$:1 or $\sqrt[3]{2}$:1, or roughly 1.25992:1. The *just major third* is defined to be the frequency ratio of 5:4 or 1.25:1. Again it is the just major third which represents the consonant interval, and the major third on our modern equal tempered scale is an approximation to it. The approximation is quite a bit worse than it was for the perfect fifth. The difference between a just major third and an equal tempered major third is quite audible; the difference is about one seventh of a semitone.

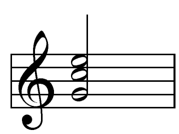
The frequency ratio represented by three semitones, for example the interval from E to the G above it, is called a *minor third*. This represents a frequency ratio of $2^{\frac{3}{12}}$:1 or $\sqrt[4]{2}$:1, or roughly 1.1892:1. The consonant *just minor third* is defined to be the frequency ratio of 6:5 or 1.2:1. The equal tempered minor third again differs from it by about a seventh of a semitone.

A major third plus a minor third makes up a fifth, either in the just/perfect versions or the equal tempered versions. So the intervals C to E (major third) plus E to G (minor third) make C to G (fifth). In the just/perfect versions, this gives ratios 4:5:6 for a *just major triad* C—E—G. We refer to C as the *root* of this chord. The chord is named after its root, so that this is a C major chord.



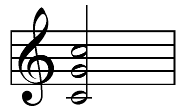
4:5:6

If we used the frequency ratios 3:4:5, it would just give an *inversion* of this chord, which is regarded as a variant form of the C major chord, because of the principle of octave equivalence.



3:4:5

while the frequency ratios 2:3:4 give a much simpler chord with a fifth and an octave.



2:3:4

So the just major triad 4:5:6 is the chord that is basic to the western system of musical harmony. On an equal tempered keyboard, this is approximated with the chord $1 : 2^{\frac{4}{12}} : 2^{\frac{7}{12}}$, which is a good approximation except for the somewhat sharp major third.

The *major scale* is formed by taking three major triads on three notes separated by intervals of a fifth. So for example the scale of C major is formed from the notes of the F major, C major and G major triads. Between them, these account for the white notes on the keyboard, which make up the scale of C major. So in just intonation, the C major scale would have the following frequency ratios.

C	D	E	F	G	A	B	C	D
$\frac{1}{1}$	$\frac{9}{8}$	$\frac{5}{4}$	$\frac{4}{3}$	$\frac{3}{2}$	$\frac{5}{3}$	$\frac{15}{8}$	$\frac{2}{1}$	$\frac{9}{4}$
4	:	5	:	6		:	(8)	
			4	:	5	:	6	
	(3)		:	4	:	5	:	6

Here, we have made use of 2:1 octaves to transfer ratios between the right and left end of the diagram.

The basic problem with this scale is that the interval from D to A is almost, but not quite equal to a perfect fifth. It is just close enough that it sounds like a nasty, out of tune fifth. It is short of a perfect fifth by a ratio of 81:80. This interval is called a *syntonic comma*. In this text, when we use the word comma without further qualification, it will always mean the syntonic comma. This and other commas are investigated in Section 5.8.

The *meantone* scale addresses this problem by distributing the syntonic comma equally between the four fifths C–G–D–A–E. So in the meantone scale, the fifths are one quarter of a comma smaller than the perfect fifth, and the major thirds are just. In the meantone scale, a number of different keys work well, but the more remote keys do not. For further details, see Section 5.12.

To make all keys work well, the meantone scale must be bent to meet around the back. A number of different versions of this compromise have been used historically, the first ones being due to Werckmeister. Some of these well tempered scales are described in Section 5.13. Meantone and well tempered scales were in common use for about four centuries before equal temperament became widespread in the late nineteenth and early twentieth century.

A *minor triad* is obtained by inverting the order of the intervals in a major triad. So for example the minor triad on the note C consists of C, E♭ and G. In just intonation, the frequency ratios are 5:6 for C–E♭ and 4:5 for E♭–G, so that C–G still makes a perfect fifth. So the ratios are 10:12:15. See §5.6 for a discussion of the role of the minor triad. A *minor scale* can be built out of three minor triads in the same way as we did for the major scale, to give the following frequency ratios.

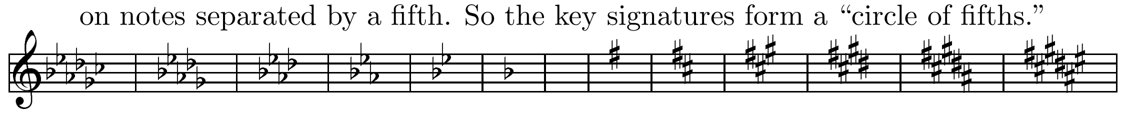
C	D	E♭	F	G	A♭	B♭	C	D
$\frac{1}{1}$	$\frac{9}{8}$	$\frac{6}{5}$	$\frac{4}{3}$	$\frac{3}{2}$	$\frac{8}{5}$	$\frac{9}{5}$	$\frac{2}{1}$	$\frac{9}{4}$
10	:	12	:	15				
			10	:	12	:	15	
				10	:	12	:	15

This is called the *natural minor* scale. Other forms of the minor scale occur because the sixth and seventh notes can be varied by moving one or both of them up a semitone to their major equivalents.

The concept of *key signature* arises from the following observation. If we look at major scales which start on notes separated by the interval of a fifth, then the two scales have all but one of the notes in common. For example, in C major, the notes are C–D–E–F–G–A–B–C, while in G major, the notes are G–A–B–C–D–E–F♯–G. The only difference, apart from a cyclic rearrangement of the notes, is that F♯ appears instead of F. So to indicate that we are in G major rather than C major, we write a sharp sign on the F at the beginning of each stave.

Similarly, the key of F major uses the notes F–G–A–B♭–C–D–E–F, which only differs from C major in the use of B♭ instead of B.

This means that key signatures are regarded as "adjacent" if they begin on notes separated by a fifth. So the key signatures form a "circle of fifths."



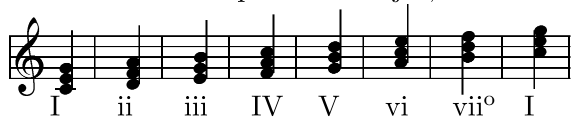
G♭ D♭ A♭ E♭ B♭ F C G D A E B F♯

In the above sequence of key signatures, the first and last are *enharmonic* versions of the same key. This means that in equal temperament, they are just different ways of writing the same keys, but in other systems such as meantone, the actual pitches may differ.

There is an easy way to memorize the correspondence between key signatures and the names of the major keys. For key signatures with sharps, the last sharp in the signature is the leading note of the key (i.e., a semitone below the note describing the key signature). So for example with four sharps, the last sharp is D♯ and so the key is E major. For key signatures with flats, the second to last flat gives the key signature. So for example with four flats, the second to last flat is A♭, so the key is A♭ major. The only case where this fails is if there is only one flat, and this is such a familiar key signature that most people find it easy to remember that it's F major.

The notes which occur in a natural minor scale are the same as the notes which occur in the major scale starting three semitones higher. For example, the notes of A minor are A–B–C–D–E–F–G–A. So the same key signature is used for A minor as for C major, and we say that A minor is the *relative minor* of C major.

The note on which a scale starts is called the *tonic*. The word *dominant* refers to the fifth above the tonic. The *roman numeral notation* is a device for naming triads relative to the tonic. So for example the major triad on the dominant is written V. Upper case roman numerals refer to major triads and lower case to minor. So for example in C major, the chords are as follows.



In D major, each chord would be a whole tone higher; so V would refer to the chord of A major instead of G major. So the roman numeral refers to the harmonic function of the chord within the key signature, rather than giving the absolute pitches.

The only triad here which is neither major nor minor is the *diminished* triad on the seventh note of the scale. This is denoted vii°, and consists of two intervals of a minor third with no major thirds.

Mode. The word *mode* refers to an arrangement of tones and semitones, with the tones approximately twice the size of the semitones (exact size depending on choice of scale), to form an octave. The naming of the modes can be a source of considerable confusion. The problem is that the names of the medieval church modes conflict with the names of the ancient Greek *tonoi*, because of a misreading of the ancient literature by some tenth century authors. The two definitions of *Hypodorian* agree, but then the medieval church modes go the wrong way around the circle.

Each mode can be considered to be the set of white keys on the piano, for a given choice of starting point. So for example *Hypodorian* goes from A to A, so that the arrangement of tones and semitones, from bottom to top, is TSTTSTT, like the minor scale. Of course, it should be realized that the pitches in a mode are not absolute, so the entire discussion can be transposed into any other key signature. For convenience, we stick to the "white note" key signature of C.

The medieval church modes also come with a choice of *finalis* or *final note*, which would normally be used as the last note of the melody. The *authentic* modes start and end with the *finalis*, while the *plagal* mode has its *finalis* on the fourth note of the scale. The four choices of *finalis* were D, E, F, G, corresponding to the authentic modes *Dorian*, *Phrygian*, *Lydian* and *Mixolydian*. The prefix *Hypo-* then turns it into the plagal mode with the same *finalis*.

To add to the confusion, the sixteenth century Swiss theorist Glareanus added four more modes with *finalis* A and C, whose authentic forms he called *Aeolian* and *Ionian*. He did not consider B to be a valid choice of *finalis*, because the fifth above it has the wrong size. More information can be found in the excellent discussion of mode in Grout and Palisca, *A history of western music*.

We summarize with a table. The first column gives the pattern of semitones and tones, from the bottom to the top of the scale. The *finalis* column only refers to the medieval church modes, not to the Greek *tonoi*. The numbers 1 to 8 are used in most medieval treatises rather than the names, and 9 to 12 are from Glareanus. Modern books on music theory often use the names for numbers 1, 3, 5, 7, 9, 4 and 11 in the following table as their names of the modes.

Intervals	Greek *tonoi*	Medieval church modes		White keys	*finalis*
TSTTTST	Phrygian	1.	Dorian	D → D	D
STTTSTT	Dorian	3.	Phrygian	E → E	E
TTTSTTS	Hypolydian	5.	Lydian	F → F	F
TTSTTST	Hypophrygian	7.	Mixolydian	G → G	G
TSTTSTT	Hypodorian	2.	Hypodorian	A → A	D
STTSTTT	Mixolydian	4.	Hypophrygian	B → B	E
TTSTTTS	Lydian	6.	Hypolydian	C → C	F
TSTTTST		8.	Hypomixolydian	D → D	G
TSTTSTT		9.	Aeolian	A → A	A
STTTSTT		10.	Hypoaeolian	E → E	A
TTSTTTS		11.	Ionian	C → C	C
TTSTTST		12.	Hypoionian	G → G	C

To put it briefly, the reason for the ascendence of the Ionian mode to the role of the modern major scale is that it is the only mode where there are three major chords for use in harmony.

APPENDIX O

Online papers

Several journals have good selections of papers available online. Access usually requires you to be logged on from an academic establishment which subscribes to the journal in question. Here is a selection of what is available from a typlical academic institution.

From www.jstor.org you can obtain online copies of papers from the American Mathematical Monthly, a publication which concentrates on undergraduate level mathematics. Papers include the following, in chronological order.

J. M. Barbour, *Synthetic musical scales*, Amer. Math. Monthly 36 (3) (1929), 155–160.

J. M. Barbour, *A sixteenth century Chinese approximation for π*, Amer. Math. Monthly 40 (2) (1933), 69–73.

J. M. Barbour, *Music and ternary continued fractions*, Amer. Math. Monthly 55 (9) (1948), 545–555.

J. B. Rosser, *Generalized ternary continued fractions*, Amer. Math. Monthly 57 (8) (1950), 528–535. This article is a reply to the above article of Barbour.

T. J. Fletcher, *Campanological groups*, Amer. Math. Monthly 63 (9) (1956), 619–626.

J. M. Barbour, *A geometrical approximation to the roots of numbers*, Amer. Math. Monthly 64 (1) (1957), 1–9. This article discusses an eighteenth century geometric method of Strähle for constructing a very good approximation to equal temperament for the frets of a guitar.

Mark Kac, *Can one hear the shape of a drum?* Amer. Math. Monthly 73 (4) (1966), 1–23.

John Rogers and Bary Mitchell, *A problem in mathematics and music*, Amer. Math. Monthly 75 (8) (1968), 871–873.

A. L. Leigh Silver, *Musimatics, or the nun's fiddle*, Amer. Math. Monthly 78 (4) (1971), 351–357.

G. D. Hasley and Edwin Hewitt, *More on the superparticular ratios in music*, Amer. Math. Monthly 79 (10) (1972), 1096–1100.

I. J. Schoenberg, *On the location of the frets on the guitar*, Amer. Math. Monthly 83 (7) (1976), 550–552. Schoenberg was the referee of the 1957 article of Barbour on Strähle's method referred to above, and this article expands on his footnotes to Barbour's article.

David Gale, *Tone perception and decomposition of periodic function*, Amer. Math. Monthly 86 (1) (1979), 36–42.

Murray Schechter, *Tempered scales and continued fractions*, Amer. Math. Monthly 87 (1)

(1980), 40–42.

David L. Reiner, *Enumeration in music theory*, Amer. Math. Monthly 92 (1) (1985), 51–54.

John Clough and Gerald Myerson, *Musical scales and the generalized circle of fifths*, Amer. Math. Monthly 93 (9) (1986), 695–701.

Arthur T. White, *Ringing the cosets*, Amer. Math. Monthly 94 (8) (1987), 721–746.

S. J. Chapman, *Drums that sound the same*, Amer. Math. Monthly 102 (2) (1995), 124–138.

Rachel W. Hall and Krešimir Josić, *The mathematics of musical instruments*, Amer. Math. Monthly 108 (4) (2001), 347–357.

There are occasionally relevant articles in the SIAM[1] journals, also available from www.jstor.org. Examples include the following.

A. A. Goldstein, *Optimal temperament*, SIAM Review 19 (3) (1977), 554–562.

A. Inselberg, *Cochlear dynamics: the evolution of a mathematical model*, SIAM Review 20 (2) (1978), 301–351.

Robert Burridge, Jay Kappraff and Christine Mordeshi, *The Sitar string, a vibrating string with a one-sided inelastic constraint*, SIAM J. Appl. Math. 42 (6) (1982), 1231–1251.

M. H. Protter, *Can one hear the shape of a drum? Revisited*, SIAM Review 29 (2) (1987), 185–197.

Tobin A. Driscoll, *Eigenmodes of isospectral drums*, SIAM Review 39 (1) (1997), 1–17.

From ojps.aip.org/jasa/ (then hit "browse html" or "search") you can obtain online copies of articles from the Journal of the Acoustical Society of America (JASA) from 1997 to the current issue. Here is a selection of some relevant articles that can be downloaded.

Donald L. Sullivan, *Accurate frequency tracking of timpani spectral lines*, JASA 101 (1) (1997), 530–538.

Antoine Chaigne and Vincent Doutaut, *Numerical simulations of xylophones. I. Time-domain modeling of the vibrating bars*, JASA 101 (1) (1997), 539–557.

Hugh J. McDermott and Colette M. McKay, *Musical pitch perception with electrical stimulation of the cochlea*, JASA 101 (3) (1997), 1622–1631.

John Sankey and William A. Sethares, *A consonance-based approach to the harpsichord tuning of Domenico Scarlatti*, JASA 101 (4) (1997), 2332–2337.

Knut Guettler and Anders Askenfelt, *Acceptance limits for the duration of pre-Helmholtz transients in bowed string attacks*, JASA 101 (5) (1997), 2903–2913.

Marc-Pierre Verge, Benoit Fabre, A. Hirschberg and A. P. J. Wijnands, *Sound production in recorderlike instruments. I. Dimensionless amplitude of the internal acoustic field*, JASA 101 (5) (1997), 2914–2924.

M. P. Verge, A. Hirschberg and R. Caussé, *Sound production in recorderlike instruments. II. A simulation model*, JASA 101 (5) (1997), 2925–2939.

[1]Society for Industrial and Applied Mathematics

David M. Mills, *Interpretation of distortion product otoacoustic emission measurements. I. Two stimulus tones*, JASA 102 (1) (1997), 413–429.

Eric Prame, *Vibrato extent and intonation in professional Western lyric singing*, JASA 102 (1) (1997), 616–621.

Guy Vandegrift and Eccles Wall, *The spatial inhomogeneity of pressure inside a violin at main air resonance*, JASA 102 (1) (1997), 622–627.

Harold A. Conklin, Jr., *Piano strings and "phantom" partials*, JASA 102 (1) (1997), 659.

I. Winkler, M. Tervaniemi and R. Näätänen, *Two separate codes for missing-fundamental pitch in the human auditory cortex*, JASA 102 (2) (1997), 1072–1082.

Alain de Cheveigné, *Harmonic fusion and pitch shifts of mistuned partials*, JASA 102 (2) (1997), 1083–1087.

Robert P. Carlyon, *The effects of two temporal cues on pitch judgments*, JASA 102 (2) (1997), 1097–1105.

N. Giordano, *Simple model of a piano soundboard*, JASA 102 (2) (1997), 1159–1168.

Ray Meddis and Lowel O'Mard, *A unitary model of pitch perception*, JASA 102 (3) (1997), 1811–1820.

Bruno H. Repp, *Acoustics, perception, and production of* legato *articulation on a computer-controlled grand piano*, JASA 102 (3) (1997), 1878–1890.

William A. Sethares, *Specifying spectra for musical scales*, JASA 102 (4) (1997), 2422–2431.

Laurent Demany and Sylvain Clément, *The perception of frequency peaks and troughs in wide frequency modulations. IV. Effect of modulation waveform*, JASA 102 (5) (1997), 2935–2944.

Ana Barjau, Vincent Gibiat and Noël Grand, *Study of woodwind-like systems through nonlinear differential equations. Part I. Simple geometry*, JASA 102 (5) (1997), 3023–3031.

Ana Barjau and Vincent Gibiat, *Study of woodwind-like systems through nonlinear differential equations. Part II. Real geometry*, JASA 102 (5) (1997), 3032–3037.

Eric D. Scheirer, *Tempo and beat analysis of acoustic musical signals*, JASA 103 (1) (1998), 588–601.

Myeong-Hwa Lee, Jeong-No Lee and Kwang-Sup Soh, *Chaos in segments from Korean traditional singing and Western singing*, JASA 103 (2) (1998), 1175–1182.

Alain de Cheveigné, *Cancellation model of pitch perception*, JASA 103 (3) (1998), 1261–1271.

Louise J. White and Christopher J. Plack, *Temporal processing of the pitch of complex tones*, JASA 103 (4) (1998), 2051–2063.

N. Giordano, *Mechanical impedance of a piano soundboard*, JASA 103 (4) (1998), 2128–2133.

Henry T. Bahnson, James F. Antaki and Quinter C. Beery, *Acoustical and physical dynamics of the diatonic harmonica*, JASA 103 (4) (1998), 2134–2144.

Jian-Yu Lin and William M. Hartmann, *The pitch of a mistuned harmonic: evidence for a* *template model*, JASA 103 (5) (1998), 2608–2617.

Shigeru Yoshikawa, *Jet-wave amplification in organ pipes*, JASA 103 (5) (1998), 2706–2717.

Teresa D. Wilson and Douglas H. Keefe, *Characterizing the clarinet tone: measurements of Lyapunov exponents, correlation dimension, and unsteadiness*, JASA 104 (1) (1998), 550–561.

Bruno H. Repp, *A microcosm of musical expression. I. Quantitative analysis of pianists' timing in the initial measures of Chopin's Etude in E major*, JASA 104 (2) (1998), 1085–1100.

Cornelis J. Nederveen, *Influence of a toroidal bend on wind instrument tuning*, JASA 104 (3) (1998), 1616–1626.

Joël Gilbert, Sylvie Ponthus and Jean-François Petiot, *Artificial buzzing lips and brass instruments: Experimental results*, JASA 104 (3) (1998), 1627–1632.

Vincent Doutant, Denis Matignon and Antoine Chaigne, *Numerical simulations of xylophones. II. Time-domain modeling of the resonator and of the radiated sound pressure*, JASA 104 (3) (1998), 1633–1647.

N. Giordano, *Sound production by a vibrating piano soundboard: Experiment*, JASA 104 (3) (1998), 1648–1653.

Jeffrey M. Brunstrom and Brian Roberts, *Profiling the perceptual suppression of partials in periodic complex tones: Further evidence for a harmonic template*, JASA 104 (6) (1998), 3511–3519.

George Bissinger, *A0 and A1 coupling, arching, rib height, and f-hole geometry dependence in the 2 degree-of-freedom network model of violin cavity modes*, JASA 104 (6) (1998), 3608–3615.

Harold A. Conklin, Jr., *Generation of partials due to nonlinear mixing in a stringed instrument*, JASA 105 (1) (1999), 536–545.

N. H. Fletcher and A. Tarnopolsky, *Blowing pressure, power, and spectrum in trumpet playing*, JASA 105 (2) (1999), 874–881.

Stephen McAdams, James W. Beauchamp and Suzanna Meneguzzi, *Discrimination of musical instrument sounds resynthesized with simplified spectrotemporal parameters*, JASA 105 (2) (1999), 882–897.

Judith C. Brown, *Computer identification of musical instruments using pattern recognition with cepstral coefficients as features*, JASA 105 (3) (1999), 1933–1941.

J. Bretos, C. Santamaría and J. Alonso Moral, *Vibrational patterns and frequency responses of the free plates and box of a violin obtained by finite element analysis*, JASA 105 (3) (1999), 1942–1950.

Daniel Pressnitzer and Stephen McAdams, *Two phase effects in roughness perception*, JASA 105 (5) (1999),2773–2782.

Seiji Adachi and Masashi Yamada, *An acoustical study of sound production in biphonic singing Xöömij*, JASA 105 (5) (1999), 2920–2932.

Xavier Boutillon and Gabriel Weinreich, *Three-dimensional mechanical admittance: Theory* *and new measurement method applied to the violin bridge*, JASA 105 (6) (1999), 3524–3533.

Eiji Hayashi, Masami Yamane and Hajime Mori, *Behavior of piano-action in a grand piano. I. Analysis of the motion of the hammer prior to string contact*, JASA 105 (6) (1999), 3534–3544.

Leïla Rhaouti, Antoine Chaigne and Patrick Joly, *Time-domain modeling and numerical simulation of a kettledrum*, JASA 105 (6) (1999), 3545–3562.

Sten Ternström, *Preferred self-to-other ratios in choir singing*, JASA 105 (6) (1999), 3563–3574.

Howard F. Pollard, *Tonal portrait of a pipe organ*, JASA 106 (1) (1999), 360–370.

Bruno H. Repp, *A microcosm of musical expression. III. Contributions of timing and dynamics to the aesthetic impression of pianists' performances of the initial measures of Chopin's Etude in E major*, JASA 106 (1) (1999), 469–478.

Alain de Cheveigné, *Pitch shifts of mistuned partials: A time-domain model*, JASA 106 (2) (1999), 887–897.

E. Obataya and M. Norimoto, *Acoustic properties of a reed (*Arundo donax *L.) used for the vibrating plate of a clarinet*, JASA 106 (2) (1999), 1106–1110.

George R. Plitnik and Bruce A. Lawson, *An investigation of correlations between geometry, acoustic variables, and psychoacoustic parameters for French horn mouthpieces*, JASA 106 (2) (1999), 1111–1125.

Valter Ciocca, *Evidence against an effect of grouping by spectral regularity on the perception of virtual pitch*, JASA 106 (5) (1999), 2746–2751.

Thomas D. Rossing and Gila Eban, *Normal modes of a radially braced guitar determined by electronic TV holography*, JASA 106 (5) (1999), 2991–2996.

Edward M. Burns and Adrianus J. M. Houtsma, *The influence of musical training on the perception of sequentially presented mistuned harmonics*, JASA 106 (6) (1999), 3564–3570.

Maureen Mellody and Gregory H. Wakefield, *The time-frequency characteristics of violin vibrato: modal distribution analysis and synthesis*, JASA 107 (1) (2000), 598–611.

Alpar Sevgen, *A principle of least complexity for musical scales*, JASA 107 (1) (2000), 665–667.

Huanping Dai, *On the relative influence of individual harmonics on pitch judgment*, JASA 107 (2) (2000), 953–959.

Jeffrey M. Brunstrom and Brian Roberts, *Separate mechanisms govern the selection of spectral components for perceptual fusion and for the computation of global pitch*, JASA 107 (3) (2000), 1566–1577.

N. Giordano and J. P. Winans II, *Piano hammers and their force compression characteristics: Does a power law make sense?*, JASA 107 (4) (2000), 2248–2255.

Richard J. Krantz and Jack Douthett, *Construction and interpretation of equal-tempered scales using frequency ratios, maximally even sets, and P-cycles*, JASA 107 (5) (2000), 2725–2734.

Anna Runnemalm, Nils-Erik Molin and Erik Jansson, *On operating deflection shapes of the violin body including in-plane motions*, JASA 107 (6) (2000), 3452–3459.

G. R. Plitnik, *Vibration characteristics of pipe organ reed tongues and the effect of the shallot, resonator, and reed curvature*, JASA 107 (6) (2000), 3460–3473.

Robert P. Carlyon, Brian C. J. Moore and Christophe Micheyl, *The effect of modulation rate on the detection of frequency modulation and mistuning of complex tones*, JASA 108 (1) (2000), 304–315.

J. Woodhouse, R. T. Schumacher and S. Garoff, *Reconstruction of bowing point friction force in a bowed string*, JASA 108 (1) (2000), 357–368.

M. J. Elejabarrieta, A. Ezcurra and C. Santamaría, *Evolution of the vibrational behavior of a guitar soundboard along successive construction phases by means of the modal analysis technique*, JASA 108 (1) (2000), 369–378.

Georg Essl and Perry R. Cook, *Measurements and efficient simulations of bowed bars*, JASA 108 (1) (2000), 379–388.

J. M. Harrison and N. Thompson-Allen, *Constancy of loudness of pipe organ sounds at different locations in an auditorium*, JASA 108 (1) (2000), 389–399.

A. Z. Tarnopolsky, N. H. Fletcher and J. C. S. Lai, *Oscillating reed valves—An experimental study*, JASA 108 (1) (2000), 400–406.

Thomas D. Rossing, Uwe J. Hansen and D. Scott Hampton, *Vibrational mode shapes in Caribbean steelpans. I. Tenor and double second*, JASA 108 (2) (2000), 803–812.

N. H. Fletcher, *A class of chaotic bird calls?*, JASA 108 (2) (2000), 821–826.

Alberto Recio and William S. Rhode, *Basilar membrane responses to broadband stimuli*, JASA 108 (5) (2000), 2281–2298.

Gabriel Weinreich, Colin Holmes and Maureen Mellody, *Air-wood coupling and the Swiss-cheese violin*, JASA 108 (5) (2000), 2389–2402.

Robert P. Carlyon, Laurent Demany and John Deeks, *Temporal pitch perception and the binaural system*, JASA 109 (2) (2000), 686–700.

Hedwig Gockel, Brian C. J. Moore and Robert P. Carlyon, *Influence of rate of change of frequency on the overall pitch of frequency-modulated tones*, JASA 109 (2) (2000), 701–712.

Daniel Pressnitzer, Roy D. Patterson and Katrin Krumbholz, *The lower limit of melodic pitch*, JASA 109 (5) (2000), 2074–2084.

R. Ranvaud, W. F. Thompson, L. Silveira-Moriyama and L.-L. Balkwill, *The speed of pitch resolution in a musical context*, JASA 109 (6) (2001), 3021–3030.

Jeffrey M. Brunstrom and Brian Roberts, *Effects of asynchrony and ear of presentation on the pitch of mistuned partials in harmonic and frequency-shifted complex tones*, JASA 110 (1) (2001), 391–401.

Lily M. Wang and Courtney B. Burroughs, *Acoustic radiation from bowed violins*, JASA 110 (1) (2001), 543–555.

Michael W. Thompson and William J. Strong, *Inclusion of wave steepening in a frequency-domain model of trombone sound reproduction*, JASA 110 (1) (2001), 556–562.

Werner Goebl, *Melody lead in piano performance: Expressive device or artifact?*, JASA 110 (1) (2001), 563–572.

Michael A. Akeroyd, Brian C. J. Moore and Geoffrey A. Moore, *Melody recognition using three types of dichotic-pitch stimulus*, JASA 110 (3) (2001), 1498-1504.

Alexander Galembo, Anders Askenfelt, Lola L. Cuddy and Frank A. Russo, *Effects of relative phases on pitch and timbre in the piano bass range*, JASA 110 (3) (2001), 1649–1666.

L. Rossi and G. Girolami, *Instantaneous frequency and short term Fourier transforms: Applications to piano sounds*, JASA 110 (5) (2001), 2412–2420.

Laurent Demany and Catherine Semal, *Learning to perceive pitch differences*, JASA 111 (3) (2002), 1377–1388.

N. H. Fletcher, W. T. McGee and A. Z. Tarnopolsky, *Bell clapper impact dynamics and the voicing of a carillon*, JASA 111 (3) (2002), 1437–1444.

From ojps.aip.org/chaos/ you can obtain online copies of papers from the journal “Chaos” from 1991 to the current issue. The relevant articles I’ve found are the following.

Jean-Pierre Boon and Oliver Decroly, *Dynamical systems theory for music dynamics*, Chaos 5 (3) (1995), 501–508.

R. T. Schumacher and J. Woodhouse, *The transient behaviour of models of bowed-string motion*, Chaos 5 (3) (1995), 509–523.

Diana S. Dabby, *Musical variations from a chaotic mapping*, Chaos 6 (2) (1996), 95–107.

From www.elsevier.com you can download the following papers.

R. C. Read, *Combinatorial problems in the theory of music*, Discrete Mathematics 167/168 (1997), 543–551.

Ján Haluška, *Equal temperament and Pythagorean tuning: a geometrical interpretation in the plane*, Fuzzy Sets and Systems 114 (2000), 261–269.

Jeong Seop Sim, Costas S. Iliopoulos, Kunsoo Park and W. F. Smyth, *Approximate periods of strings*, Theoretical Computer Science 262 (2001), 557–568.

From www.idealibrary.com, you can obtain online copies of papers from a number of journals; for example, the following papers come from the Journal of Sound and Vibration.

F. Gautier and N. Tahani, *Vibroacoustic behaviour of a simplified musical wind instrument*, Journal of Sound and Vibration 213 (1) (1998), 107–125.

S. Gaudet, C. Gauthier and V. G. LeBlanc, *On the vibrations of an N-string*, Journal of Sound and Vibration 238 (1) (2000), 147–169.

From www.emis.de/journals/SLC, you can obtain online copies of papers from the Séminaire Lotharingien de Combinatoire. The following paper is relevant to §9.15.

Harald Fripertinger, *Enumeration in musical theory*, Séminaire Lotharingien de Combinatoire 26 (1991), 29–42.

From www.combinatorics.org, you can obtain online copies of papers from the Electronic Journal of Combinatorics. The only relevant paper I've found in this journal is the following.

Maxime Crochemore, Costas S. Iliopoulos and Yoan J. Pinzon, *Computing Evolutionary Chains in Musical Sequences*, Electronic J. Comb. 8 (2) (2001), #R5.

Guerino Mazzola keeps some of his papers on mathematics and music available online at

www.ifi.unizh.ch/mml/musicmedia/publications.php4

Harald Fripertinger's papers on music and combinatorics can be downloaded from

www-ang.kfunigraz.ac.at/∼fripert/publications.html

You can download Julius O. Smith III, *Mathematics of the discrete Fourier transform* (237 pages of lecture notes, pdf or compressed postscript format) from

ccrma-www.stanford.edu/∼jos/r320/

The School of Music at Indiana University has made a large number of original documents in Latin available for anyone to download from the *Thesaurus Musicarum Latinarum.* This contains, for example, the works of Boethius, Gaffurius, Odington and Ramis de Pareja.

www.music.indiana.edu/tml/

A companion database of Italian documents, the *saggi musicali italiani*, contains for example the works of Zarlino. It is available at

www.music.indiana.edu/smi/

APPENDIX P

Partial derivatives

Partial derivatives are what happens when we differentiate a function of more than one variable. For example, a geographical map which indicates height above sea level, by some device such as coloration or contours, can be regarded as describing a function $z = f(x, y)$. Here, x and y represent the two coordinates of the map, and z denotes height above sea level. If we move due east, which we take to be the direction of the x axis, then we are keeping y constant and changing x. So the slope in this direction would be the derivative of $z = f(x, y)$ with respect to x, regarding y as a constant. This derivative is denoted $\dfrac{\partial z}{\partial x}$. More formally,

$$\frac{\partial z}{\partial x} = \lim_{h \to 0} \frac{f(x+h, y) - f(x, y)}{h}.$$

Similarly, $\dfrac{\partial z}{\partial y}$ is the derivative of z with respect to y, regarding x as a constant. As an example, let $z = x^4 + x^2 y - 2y^2$. Then we have $\dfrac{\partial z}{\partial x} = 4x^3 + 2xy$, because $x^2 y$ is being regarded as a constant multiple of x^2, and $-2y^2$ is just a constant. Similarly, $\dfrac{\partial z}{\partial y} = x^2 - 4y$, because x^4 is a constant and $x^2 y$ is a constant multiple of y.

Second partial derivatives are defined similarly, but we now find that we can mix the variables. As well as $\dfrac{\partial^2 z}{\partial x^2}$ and $\dfrac{\partial^2 z}{\partial y^2}$, we can now form $\dfrac{\partial^2 z}{\partial x \partial y}$ by taking the partial derivative of $\dfrac{\partial z}{\partial y}$ with respect to x, regarding y as constant, and we can also form $\dfrac{\partial^2 z}{\partial y \partial x}$ by taking partial derivatives in the opposite order. So in the above example, we have

$$\frac{\partial^2 z}{\partial x^2} = 12x^2 + 2y, \qquad \frac{\partial^2 z}{\partial y^2} = -4, \qquad \frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x} = 2x.$$

In fact, the two mixed partial derivatives agree under some fairly mild hypotheses.

THEOREM P.1. *Suppose that the partial derivatives* $\dfrac{\partial^2 z}{\partial x \partial y}$ *and* $\dfrac{\partial^2 z}{\partial y \partial x}$ *both exist and are both continuous at some point (i.e., for some chosen values of x and y). Then they are equal at that point.*

PROOF. See any book on elementary analysis; for example, J. C. Burkhill, *A first course in mathematical analysis*, CUP, 1962, theorem 8.3. □

Partial derivatives work in exactly the same way for functions of more variables. So for example if $f(x, y, z) = xy^2 \sin z$ then we have $\dfrac{\partial f}{\partial x} = y^2 \sin z$, $\dfrac{\partial f}{\partial y} = 2xy \sin z$, and $\dfrac{\partial f}{\partial z} = xy^2 \cos z$. For each pair of variables, the two mixed partial derivatives with respect to those variables agree provided they are both continuous.

The chain rule for partial derivatives needs some care. Suppose, by way of example, that z is a function of u, v and w, and that each of u, v and w is a function of x and y. Then z can also be regarded as a function of x and y. A change in the value of x, keeping y constant, will result in a change of all of u, v and w, and each of these changes will result in a change in the value of z. These changes have to be added as follows:

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial z}{\partial w}\frac{\partial w}{\partial x}.$$

Similarly, we have

$$\frac{\partial z}{\partial y} = \frac{\partial z}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial y} + \frac{\partial z}{\partial w}\frac{\partial w}{\partial y}.$$

It is essential to keep track of which variables are independent, intermediate, and dependent. In this example, the independent variables are x and y, the intermediate ones are u, v and w, and the dependent variable is z.

A good illustration of the chain rule for partial derivatives is given by the conversion from Cartesian to polar coordinates. If z is a function of x and y then it can also be regarded as a function of r and θ. To convert from polar to Cartesian coordinates, we use $x = r \cos \theta$ and $y = r \sin \theta$, and to convert back we use $r = \sqrt{x^2 + y^2}$ and $\tan \theta = y/x$. Let us convert the quantity

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2},$$

into polar coordinates, assuming that all mixed second partial derivatives are continuous, so that the above theorem applies. This calculation will be needed in §3.6, where we investigate the vibrational modes of the drum. For this purpose, it is actually technically slightly easier to regard x and y as the intermediate variables and r and θ as the independent variables, although it would be quite permissible to interchange their roles. The dependent variable is z. We have

$$\frac{\partial z}{\partial r} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial r} = \cos \theta \frac{\partial z}{\partial x} + \sin \theta \frac{\partial z}{\partial y}. \tag{P.1}$$

To take the second derivative, we do the same again.

$$\begin{aligned}\frac{\partial^2 z}{\partial r^2} &= \cos\theta\frac{\partial}{\partial r}\left(\frac{\partial z}{\partial x}\right) + \sin\theta\frac{\partial}{\partial r}\left(\frac{\partial z}{\partial y}\right)\\ &= \cos\theta\left(\cos\theta\frac{\partial^2 z}{\partial x^2} + \sin\theta\frac{\partial^2 z}{\partial y\partial x}\right) + \sin\theta\left(\cos\theta\frac{\partial^2 z}{\partial x\partial y} + \sin\theta\frac{\partial^2 z}{\partial y^2}\right)\\ &= \cos^2\theta\frac{\partial^2 z}{\partial x^2} + 2\sin\theta\cos\theta\frac{\partial^2 z}{\partial x\partial y} + \sin^2\theta\frac{\partial^2 z}{\partial y^2}. \qquad \text{(P.2)}\end{aligned}$$

Similarly, we have

$$\frac{\partial z}{\partial\theta} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial\theta} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial\theta} = (-r\sin\theta)\frac{\partial z}{\partial x} + (r\cos\theta)\frac{\partial z}{\partial y},$$

and

$$\begin{aligned}\frac{\partial^2 z}{\partial\theta^2} &= (-r\sin\theta)\frac{\partial}{\partial\theta}\left(\frac{\partial z}{\partial x}\right) + (-r\cos\theta)\frac{\partial z}{\partial x}\\ &\qquad + (r\cos\theta)\frac{\partial}{\partial\theta}\left(\frac{\partial z}{\partial y}\right) + (-r\sin\theta)\frac{\partial z}{\partial y}\\ &= (-r\sin\theta)\left((-r\sin\theta)\frac{\partial^2 z}{\partial x^2} + (r\cos\theta)\frac{\partial^2 z}{\partial y\partial x}\right) + (-r\cos\theta)\frac{\partial z}{\partial x}\\ &\qquad + (r\cos\theta)\left((-r\sin\theta)\frac{\partial^2 z}{\partial x\partial y} + (r\cos\theta)\frac{\partial^2 z}{\partial y^2}\right) + (-r\cos\theta)\frac{\partial z}{\partial y}\\ &= r^2\left(\sin^2\theta\frac{\partial^2 z}{\partial x^2} - 2\sin\theta\cos\theta\frac{\partial^2 z}{\partial x\partial y} + \cos^2\theta\frac{\partial^2 z}{\partial y^2}\right)\\ &\qquad - r\left(\cos\theta\frac{\partial z}{\partial x} + \sin\theta\frac{\partial z}{\partial y}\right). \qquad \text{(P.3)}\end{aligned}$$

Comparing the formula (P.2) for $\dfrac{\partial^2 z}{\partial r^2}$ with the formula (P.3) for $\dfrac{\partial^2 z}{\partial\theta^2}$, and using the fact that $\sin^2\theta + \cos^2\theta = 1$, we see that

$$\frac{\partial^2 z}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 z}{\partial\theta^2} = \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} - \frac{1}{r}\left(\cos\theta\frac{\partial z}{\partial x} + \sin\theta\frac{\partial z}{\partial y}\right).$$

Finally, looking back at equation (P.1) for $\dfrac{\partial z}{\partial r}$, we obtain the formula we were looking for, namely

$$\frac{\partial^2 z}{\partial r^2} + \frac{1}{r}\frac{\partial z}{\partial r} + \frac{1}{r^2}\frac{\partial^2 z}{\partial\theta^2} = \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}. \qquad \text{(P.4)}$$

APPENDIX R

Recordings

Go to the entry "compact discs" in the index to find the points in the text which refer to these recordings.

Bill Alves, *Terrain of possibilities*, Emf media #2, 2000. Music made with Synclavier and CSound using just intonation.

Johann Sebastian Bach, *The Complete Organ Music*, recorded by Hans Fagius, Volumes 6 and 8, BIS-CD-397/398 (1989) and BIS-CD-443/444 (1989 & 1990). These recordings are played on the reconstructed 1764 Wahlberg organ, Fredrikskyrkan, Karlskrona, Sweden. This organ was reconstructed using the original temperament, which was Neidhardt's Circulating Temperament No. 3 "für eine grosse Stadt" (for a large town).

Clarence Barlow's "OTOdeBLU" is in 17 tone equal temperament, played on two pianos. This piece was composed in celebration of John Pierce's eightieth birthday, and appeared as track 15 on the Computer Music Journal's Sound Anthology CD, 1995, to accompany volumes 15–19 of the journal. The CD can be obtained from MIT press for $15.

Between the Keys, *Microtonal masterpieces of the 20th century*, Newport Classic CD #85526, 1992. This CD contains recordings of Charles Ives' *Three quartertone pieces*, and a piece by Ivan Vyshnegradsky in 72 tone equal temperament. Unfortunately, this CD seems to have gone out of print.

Easley Blackwood has composed a set of microtonal compositions in each of the equally tempered scales from 13 tone to 24 tone, as part of a research project funded by the National Endowment for the Humanities to explore the tonal and modal behavior of these temperaments. He devised notations for each tuning, and his compositions were designed to illustrate chord progressions and practical application of his notations. The results are available on compact disc as Cedille Records CDR 90000 018, Easley Blackwood: *Microtonal Compositions* (1994). Copies of the scores of the works can be obtained from Blackwood Enterprises, 5300 South Shore Drive, Chicago, IL 60615, USA for a nominal cost.

Dietrich Buxtehude, *Orgelwerke*, Volumes 1–7, recorded by Harald Vogel, published by Dabringhaus and Grimm. These works are recorded on a variety of European organs in different temperaments. Extensive details are given in the liner notes.

CD1 Tracks 1–8: Norden – St. Jakobi/Kleine organ in Werckmeister III;

Tracks 9–15: Norden – St. Ludgeri organ in modified $\frac{1}{5}$ Pythagorean comma meantone with $\mathrm{C}\sharp^{-\frac{6}{5}p}$, $\mathrm{G}\sharp^{-\frac{6}{5}p}$, $\mathrm{B}\flat^{+\frac{1}{5}p}$ and $\mathrm{E}\flat^{0}$;

CD2 Tracks 1–6: Stade – St. Cosmae organ in modified quarter comma meantone with[1] $C\sharp^{-\frac{3}{2}}$, $G\sharp^{-\frac{3}{2}}$, F^{0}, $B\flat^{0}$, $E\flat^{-\frac{1}{5}}$;

Tracks 7–15: Weener – Georgskirche organ in Werckmeister III;

CD3 Tracks 1–10: Grasberg organ in Neidhardt No. 3;

Tracks 11–14: Damp – Herrenhaus organ in modified meantone with pitches taken from original pipe lengths;

CD4 Tracks 1–8: Noordbroeck organ in Werckmeister III;

Tracks 9–15: Groningen – Aa-Kerk organ in (almost) equal temperament;

CD5 Tracks 1–5: Pilsum organ in modified $\frac{1}{5}$ Pythagorean comma meantone (the same as the Norden – St. Ludgeri organ described above);

Tracks 6–7: Buttforde organ;

Tracks 8–10: Langwarden organ in modified quarter comma meantone with $G\sharp^{-\frac{7}{4}}$, $B\flat^{-\frac{1}{4}}$, $E\flat^{-\frac{1}{4}}$;

Tracks 11–13: Basedow organ in quarter comma meantone;

Tracks 14–15: Groß Eichsen organ in quarter comma meantone;

CD6 Tracks 1–10: Roskilde organ in Neidhardt (no. 3?);

Track 11: Helsingør organ (unspecified temperament);

Tracks 12–15: Torrlösa organ (unspecified temperament);

CD7 Tracks 1–10 modified $\frac{1}{5}$ comma meantone with[2] $C\sharp^{-\frac{6}{5}}$, $G\sharp^{-\frac{6}{5}}$, $B\flat^{+\frac{1}{5}}$ and $E\flat^{\frac{1}{5}-\frac{1}{10}p}$.

William Byrd, *Cantones Sacrae 1575, The Cardinall's Music*, conducted by David Skinner. Track 12, *Diliges Dominum*, exhibits temporal reflectional symmetry, so that it is a perfect palindrome (see §9.1).

Wendy Carlos, *Beauty in the Beast*, Audion, 1986, Passport Records, Inc., SYNCD 200. Tracks 4 and 5 make use of Carlos' just scales described in §6.1.

Wendy Carlos, *Switched-On Bach 2000*, 1992. Telarc CD-80323. Carlos' original "Switched-On Bach" recording was performed on a Moog analog synthesizer, back in the late 1960s. The Moog is only capable of playing in equal temperament. Improvements in technology inspired her to release this new recording, using a variety of temperaments and modern methods of digital synthesis. The temperaments used are $\frac{1}{5}$ and $\frac{1}{4}$ comma meantone, and various circular (irregular) temperaments.

Wendy Carlos, *Tales of Heaven and Hell*, 1998. East Side Digital, ESD 81352. The third track, *Clockwork Black*, uses $\frac{1}{5}$th comma meantone temperament. The sixth track, *Afterlife*, uses 15 tone equal temperament, alternating with another more *ad hoc* scale. The seventh and final track uses a variation of Werckmeister III.

Charles Carpenter has two CDs, titled *Frog à la Pêche* (Caterwaul Records, CAT8221, 1994) and *Splat* (Caterwaul Records, CAT4969, 1996), composed using the Bohlen–Pierce scale, and played in a progressive rock/jazz style. These recordings can be ordered directly from www.kspace.com/carpenter for $13.95 each. Although Carpenter does not restrict himself to sounds composed mainly of odd harmonics, his compositions are nonetheless compelling.

---

[1]The liner notes are written as though $G\sharp^{-\frac{3}{2}}$ were equal to $A\flat^{-\frac{2}{5}}$, which is not quite true. But the discrepancy is only about 0.2 cents.

[2]The liner notes identify $A\flat^{-\frac{1}{10}p}$ with $G\sharp^{-\frac{6}{5}}$, in accordance with the approximation of Kirnberger and Farey described in §5.14.

Marc Chemillier and E. de Dampierre, *Central African Republic. Music of the former Bandia courts*, CNRS/Musée de l'Homme, Le Chant du Monde, CNR 2741009, Paris, 1996.

Perry Cook (ed.), *Music, congnition and computerized sound. An introduction to psychoacoustics* [**21**] comes with an accompanying CD full of sound examples.

Michael Harrison, *From Ancient Worlds*, for Harmonic Piano, New Albion Records, Inc., 1992. NA 042 CD. The pieces on this recording all make use of his 24 tone just scale, described in §6.1.

Michael Harrison has also just released another CD using his Harmonic Piano, *Revelation*, recorded live in the Lincoln Center in October 2001 and issued in January 2002. In this recording, the harmonic piano is tuned to a just scale using only the primes 2, 3 and 7 (not 5). The 12 notes in the octave have ratios

$$1{:}1,\ 63{:}64,\ 9{:}8,\ 567{:}512,\ 81{:}64,\ 21{:}16,\ 729{:}512,\ 3{:}2,\\ 189{:}128,\ 27{:}16,\ 7{:}4,\ 243{:}128,\ (2{:}1).$$

The scale begins on F, and has the peculiarity that ♯ lowers a note by a septimal comma.

Jonathan Harvey, *Mead: Ritual melodies*, Sargasso CD #28029, 1999. Track two on this CD, *Mortuos Plango, Vivos Voco*, makes use of a scale derived from a spectral analysis of the Great Bell of Winchester Cathedral.

Neil Haverstick, *Acoustic stick*, Hapi Skratch, 1998. The pieces on this CD are played on custom made guitars using 19 and 34 tone equal temperament.

In Joseph Haydn's *Sonata 41* in A (Hob. XVI:26), the movement *Menuetto al rovescio* is a perfect palindrome (see §9.1). This piece can be found as track 16 on the Naxos CD number 8.553127, Haydn, *Piano sonatas, Vol. 4*, with Jenõ Jandó at the piano.

A. J. M. Houtsma and T. D. Rossing and W. M. Wagenaars, *Auditory Demonstrations*, Audio CD and accompanying booklet, Philips, 1987. This classic collection of sound examples illustrates a number of acoustic and psychoacoustic phenomena. It can be obtained from the Acoustical Society of America at asa.aip.org/discs.html for $26 + shipping.

Ben Johnson, *Music for piano*, played by Phillip Bush, Koch International Classics CD #7369. Pieces for piano in a microtonal just scale.

Enid Katahn, *Beethoven in the Temperaments* (Gasparo GSCD-332, 1997). Katahn plays Beethoven's Sonatas Op. 13, *Pathétique* and Op. 14 Nr. 1 using the Prinz temperament, and Sonatas Op. 27 Nr. 2, *Moonlight* and Op. 53 *Waldstein* in Thomas Young's temperament. The instrument is a modern Steinway concert grand rather than a period instrument. The tuning and liner notes are by Edward Foote.

Enid Katahn and Edward Foote have also brought out a recording, *Six degrees of tonality* (Gasparo GSCD-344, 2000). This begins with Scarlatti's Sonata K. 96 in quarter comma meantone, followed by Mozart's *Fantasie* Kv. 397 in Prelleur temperament, a Haydn sonata in Kirnberger III, a Beethoven sonata in Young temperament, Chopin's *Fantaisie-Impromptu* in DeMorgan temperament, and Grieg's *Glochengeläute* in Coleman 11 temperament. Finally, and in many ways the most interesting part of this recording, the Mozart *Fantasie* is played in quarter comma meantone, Prelleur temperament and equal temperament in succession, which allows a very direct comparison to be made. Unfortunately, the tempi are slightly different, which makes this recording not very useful for a blind test.

Bernard Lagacé has recorded a CD of music of various composers on the C. B. Fisk organ at Wellesley College, Massachusetts, USA, tuned in quarter comma meantone temperament. This recording is available from Titanic Records Ti-207, 1991.

Guillaume de Machaut (1300–1377), *Messe de Notre Dame* and other works. The Hilliard Ensemble, Hyperíon, 1989, CDA66358. This recording is sung in Pythagorean intonation throughout. The mass alternates polyphonic with monophonic sections. The double leading-note cadences at the end of each polyphonic section are particularly striking in Pythagorean intonation. Track 19 of this recording is *Ma fin est mon commencement* (My end is my beginning). This is an example of retrograde canon, meaning that it exhibits temporal reflectional symmetry (see §9.1).

Mathews and Pierce, *Current directions in computer music research* [**82**] comes with a companion CD containing numerous examples; note that track 76 is erroneous, cf. Pierce [**103**], page 257.

Microtonal works, Mode CD #18, contains microtonal works of Joan la Barbara, John Cage, Dean Drummond and Harry Partch.

Edward Parmentier, *Seventeenth Century French Harpsichord Music*, Wildboar, 1985, WLBR 8502. This collection contains pieces by Johann Jakob Froberger, Louis Couperin, Jacques Champion de Chambonnières, and Jean-Henri d'Anglebert. The recording was made using a Keith Hill copy of a 1640 harpsichord by Joannes Couchet, tuned in $\frac{1}{3}$ comma meantone temperament.

Many of Harry Partch's compositions have been rereleased on CD by Composers Recordings Inc., 73 Spring Street, Suite 506, New York, NY 10012-5800. As a starting point, I would recommend *The Bewitched*, CRI CD 7001, originally released on Partch's own label, Gate 5. This piece makes extensive use of his 43 tone just scale, described in §6.1.

A number of Robert Rich's recordings are in some form of just scale. His basic scale is mostly 5-limit with a 7:5 tritone:

$$1{:}1,\ 16{:}15,\ 9{:}8,\ 6{:}5,\ 5{:}4,\ 4{:}3,\ 7{:}5,\ 3{:}2,\ 8{:}5,\ 5{:}3,\ 9{:}5,\ 15{:}8.$$

This appears throughout the CDs *Numena*, *Geometry*, *Rainforest*, and others. One of the nicest examples of this tuning is *The Raining Room* on the CD *Rainforest*, Hearts of Space HS11014-2. He also uses the 7-limit scale

$$1{:}1,\ 15{:}14,\ 9{:}8,\ 7{:}6,\ 5{:}4,\ 4{:}3,\ 7{:}5,\ 3{:}2,\ 14{:}9,\ 5{:}3,\ 7{:}4,\ 15{:}8.$$

This appears on *Sagrada Familia* on the CD *Gaudi*, Hearts of Space HS11028-2. See www.amoeba.com for a more complete discography of Robert Rich's work.

William Sethares, *Xentonality*, Music in 10-, 13-, 17- and 19-tone equal temperament using spectrally adjusted instruments. Frog Peak Music www.frogpeak.org, 1997.

William Sethares, *Tuning, timbre, spectrum, scale* [**129**] comes with a CD full of examples.

Isao Tomita, *Pictures at an Exhibition* (Mussorgsky), BMG 60576-2-RG. This recording was made on analog synthesizers in 1974, and is remarkably sophisticated for that era.

Johann Gottfried Walther, *Organ Works*, Volume 1, played by Craig Cramer on the organ of St. Bonifacius, Tröchtelborn, Germany. Naxos CD number 8.554316. This organ was restored in Kellner's reconstruction of Bach's temperament, see §5.13. For more information about the organ (details are not given in the CD liner notes), see www.gdo.de/neurest/troechtelborn.html.

Aldert Winkelman, *Works by Mattheson, Couperin, and others*. Clavigram VRS 1735-2. This recording is hard to obtain. The pieces by Johann Mattheson, François Couperin, Johann Jakob Froberger, Joannes de Gruytters and Jacques Duphly are played on a harpsichord tuned to Werckmeister III. The pieces by Louis Couperin and Gottlieb Muffat are played on a spinet tuned in quarter comma meantone.

APPENDIX W

The wave equation

This appendix is a supplement to Section 3.7. Its purpose is to justify the method of separation of variables for the wave equation, to show that a drum has "enough" eigenvalues, and to explain the construction of two different drums with the same Dirichlet spectrum. The account of the solution of the wave equation given here is deliberately much more compressed than the account usually given in books on partial differential equations, to emphasize the shape of the reasoning rather than the more computational aspects usually considered. The level of mathematical sophistication needed to follow this appendix is rather greater than for the rest of the book. The reader eager to understand how two different drums can have the same Dirichlet spectrum should jump straight to page 417 and examine the correspondence of eigenfunctions described there.

We discuss solutions z of the two dimensional wave equation

$$\frac{\partial^2 z}{\partial t^2} = c^2 \nabla^2 z, \tag{W.1}$$

on a closed, bounded domain Ω. For boundary conditions, we assume that z is identically zero on the boundary S (Dirichlet boundary conditions). Initial conditions are given by specifying the values of z and $\frac{\partial z}{\partial t}$ at $t = 0$.

Throughout this appendix, Ω is a closed, bounded, simply connected domain in $\mathbb{R}^2$ with piecewise twice continuously differentiable boundary S. We write $\mathbf{x}$ for the position vector (x, y) on Ω, and $d\mathbf{x}$ for the element $dx\,dy$ of area on Ω. We write $\mathbf{n}$ for the outward normal vector to S, and $d\sigma$ denotes the element of length on S. With this notation, the divergence theorem states that if $f(\mathbf{x})$ is a continuously differentiable function on Ω then

$$\int_S f\,.\,\mathbf{n}\,d\sigma = \int_\Omega \nabla f\,d\mathbf{x}. \tag{W.2}$$

In order to solve the wave equation, we begin with a study of Laplace's equation

$$\nabla^2 \phi = 0$$

on Ω, with Dirichlet boundary conditions, in other words with given value of ϕ on the boundary S. We then use this to construct Green's functions, which we in turn use in order to find an integral operator which is an inverse for ∇^2. This integral operator $\mathbf{K}$ will turn out to be a compact positive self-adjoint operator, which is what allows us to get information about its eigenvalues.

Green's identities

Let Ω be a closed bounded region with boundary S. Suppose that $f(\mathbf{x})$ and $g(\mathbf{x})$ are functions on Ω. Then we have

$$\nabla.(f\nabla g) = f\nabla^2 g + \nabla f\,.\,\nabla g. \tag{W.3}$$

If Ω is a closed bounded region with boundary S, then integrating over Ω and using the divergence theorem (W.2), we get Green's first identity.

THEOREM W.1 (Green's First Identity). *Let $f(\mathbf{x})$ be continuously differentiable, and $g(\mathbf{x})$ be twice continuously differentiable on Ω. Then*

$$\int_S (f\nabla g)\,.\,\mathbf{n}\,d\sigma = \int_\Omega (f\nabla^2 g + \nabla f\,.\,\nabla g)\,d\mathbf{x}. \tag{W.4}$$

Reversing the roles of f and g and subtracting gives Green's second identity.

THEOREM W.2 (Green's Second Identity). *Let $f(\mathbf{x})$ and $g(\mathbf{x})$ be twice continuously differentiable on Ω. Then*

$$\int_S (f\nabla g - g\nabla f)\,.\,\mathbf{n}\,d\sigma = \int_\Omega (f\nabla^2 g - g\nabla^2 f)\,d\mathbf{x}. \tag{W.5}$$

The following is a useful consequence of Green's second identity.

LEMMA W.3. *For twice continuously differentiable functions f and g on Ω vanishing on the boundary S, we have*

$$\int_\Omega f\nabla^2 g\,d\mathbf{x} = \int_\Omega g\nabla^2 f\,d\mathbf{x}. \qquad \square$$

Gauss' formula

We start with the function of two variables $\mathbf{x}$ and $\mathbf{x}'$ in Ω given by $z = \ln|\mathbf{x}-\mathbf{x}'|$. For functions of two variables, it makes sense to apply ∇ with respect to $\mathbf{x}$ keeping $\mathbf{x}'$ constant, or vice versa. These are analogs of partial differentiation. To distinguish between these two options, we write $\nabla_{\mathbf{x}}$ or $\nabla_{\mathbf{x}'}$.

An easy calculation in terms of coordinates shows that as long as $\mathbf{x} \neq \mathbf{x}'$, we have

$$\nabla_{\mathbf{x}'} \ln|\mathbf{x}-\mathbf{x}'| = -\frac{\mathbf{x}-\mathbf{x}'}{|\mathbf{x}-\mathbf{x}'|^2} \tag{W.6}$$

and

$$\nabla^2_{\mathbf{x}'} \ln|\mathbf{x}-\mathbf{x}'| = 0. \tag{W.7}$$

For $\mathbf{x} = \mathbf{x}'$, the quantity $\nabla^2_{\mathbf{x}'} \ln|\mathbf{x}-\mathbf{x}'|$ doesn't make sense, because the logarithm isn't defined. But if we pretend that it is continuously differentiable, and integrate using the divergence theorem (W.2) we get

$$\int_\Omega \nabla^2_{\mathbf{x}'} \ln|\mathbf{x}-\mathbf{x}'|\,d\mathbf{x}' = \int_S \nabla_{\mathbf{x}'} \ln|\mathbf{x}-\mathbf{x}'|\,.\,\mathbf{n}'\,d\sigma' = -\int_S \frac{\mathbf{x}-\mathbf{x}'}{|\mathbf{x}-\mathbf{x}'|^2}\,.\,\mathbf{n}'\,d\sigma', \tag{W.8}$$

where $\mathbf{n}'$ and σ' are with respect to $\mathbf{x}'$. The shape of the region Ω doesn't matter in this calculation, as long as $\mathbf{x}'$ is in the interior, because of equation (W.7). If we measure using $\mathbf{x}$ as the origin and make the region a unit disk centered at the origin, then the calculation reduces to $\int_S \mathbf{x}'.\mathbf{n}'\,d\sigma'$. But in this case $\mathbf{x}'$ and $\mathbf{n}'$ are unit vectors in the same direction, so $\mathbf{x}'.\mathbf{n}' = 1$. Since the circumference of the unit circle is 2π, the integral gives 2π,

$$\int_S \nabla_{\mathbf{x}'} \ln|\mathbf{x}-\mathbf{x}'|\,.\,\mathbf{n}'\,d\sigma' = 2\pi. \tag{W.9}$$

The interpretation of this calculation is that although $\ln|\mathbf{x}-\mathbf{x}'|$ is not differentiable with respect to $\mathbf{x}'$ at $\mathbf{x}'=\mathbf{x}$, we can think of $\nabla^2_{\mathbf{x}'}\ln|\mathbf{x}-\mathbf{x}'|$ as a distribution, in the sense in which we introduced the term in Section 2.17. We have to replace $\int_{-\infty}^{\infty}$ with $\int_\Omega$, so that the delta function $\delta(\mathbf{x})$ is defined to be zero for $\mathbf{x}\neq\mathbf{0}$, and $\int_\Omega \delta(\mathbf{x})\,d\mathbf{x} = 1$. In terms of this delta function, the above calculation can be expressed as saying that

$$\nabla^2_{\mathbf{x}'}\ln|\mathbf{x}-\mathbf{x}'| = 2\pi\delta(\mathbf{x}-\mathbf{x}'). \tag{W.10}$$

So far, we have assumed that $\mathbf{x}'$ is in the interior of Ω. For a point $\mathbf{x}'$ outside Ω, the integrand in equation (W.8) is zero so the integral is zero. If $\mathbf{x}'$ is on the boundary S, and it is a point where S is continuously differentiable, then instead of a circle, in the above calculation we have to integrate over a semicircle. So the integral is π instead of 2π. At a corner with angle θ, we are integrating over a sector of a circle with angle θ, so the integral is θ. So we define a function $p(\mathbf{x})$ on $\mathbb{R}^2$ by

$$p(\mathbf{x}) = \begin{cases} 2\pi & \text{if } \mathbf{x} \text{ is in the interior of } \Omega, \\ 0 & \text{if } \mathbf{x} \text{ is not in } \Omega, \\ \pi & \text{if } \mathbf{x} \text{ is a continuously differentiable point on } S, \\ \theta & \text{if } \mathbf{x} \text{ is a corner of } S \text{ with interior angle } \theta. \end{cases}$$

Then the extension of equation (W.9) to the plane is **Gauss' formula**

$$\int_S \nabla_{\mathbf{x}'} \ln|\mathbf{x}-\mathbf{x}'|\,.\,\mathbf{n}'\,d\sigma' = p(\mathbf{x}). \tag{W.11}$$

If $f(\mathbf{x})$ is any continuous function on Ω, then we have

$$\int_\Omega f(\mathbf{x}')\nabla^2_{\mathbf{x}'}\ln|\mathbf{x}-\mathbf{x}'|\,d\mathbf{x}' = p(\mathbf{x})f(\mathbf{x}). \tag{W.12}$$

This is because the integrand is zero except near $\mathbf{x}=\mathbf{x}'$, so $f(\mathbf{x}')$ may as well be replaced by $f(\mathbf{x})$ and taken out of the integral before applying the divergence theorem.

Remark. The above calculation was performed in two dimensions. The corresponding calculation in three dimensions uses the function $1/|\mathbf{x}-\mathbf{x}'|$ instead of $\ln|\mathbf{x}-\mathbf{x}'|$. The unit circle is replaced by the unit sphere, of surface area 4π, and the analog of equation (W.9) is

$$\int_S \nabla_{\mathbf{x}'}\frac{1}{|\mathbf{x}-\mathbf{x}'|}\,.\,\mathbf{n}'\,d\sigma' = 4\pi.$$

The definition of $h(\mathbf{x},\mathbf{x}')$ and $G(\mathbf{x},\mathbf{x}')$ below are adjusted accordingly.

Similarly, in n dimensions ($n \geq 3$), the corresponding formula is

$$\int_S \nabla_{\mathbf{x}'} \frac{1}{|\mathbf{x}-\mathbf{x}'|^{n-2}} \,.\, \mathbf{n}'\, d\sigma' = n(n-2)\alpha(n)$$

where $\alpha(n)$ denotes the $(n-1)$-dimensional volume of the surface of the n-dimensional sphere.

Green's functions

Equation (W.10) is an important property of the function $\ln|\mathbf{x}-\mathbf{x}'|$. But the main problem with this function is that it doesn't vanish on the boundary S of Ω. To remedy this, we adjust it as follows. Suppose that we can find a solution $h(\mathbf{x},\mathbf{x}')$ to Laplace's equation

$$\nabla^2_{\mathbf{x}'} h(\mathbf{x},\mathbf{x}') = 0 \tag{W.13}$$

on Ω, with boundary conditions

$$h(\mathbf{x},\mathbf{x}') = \frac{1}{2\pi} \ln|\mathbf{x}-\mathbf{x}'| \tag{W.14}$$

for $\mathbf{x}'$ on S. That is, we insist that $h(\mathbf{x},\mathbf{x}')$ is defined even when $\mathbf{x} = \mathbf{x}'$ (in the interior of Ω). Then the function

$$G(\mathbf{x},\mathbf{x}') = h(\mathbf{x},\mathbf{x}') - \frac{1}{2\pi} \ln|\mathbf{x}-\mathbf{x}'|$$

still satisfies

$$\nabla^2_{\mathbf{x}'} G(\mathbf{x},\mathbf{x}') = \delta(\mathbf{x}-\mathbf{x}') \tag{W.15}$$

for $\mathbf{x}'$ in the interior of Ω, but it now also satisfies $G(\mathbf{x},\mathbf{x}') = 0$ for $\mathbf{x}'$ on S. The function $G(\mathbf{x},\mathbf{x}')$ defined this way is called the Green's function for the Laplace operator ∇^2.

Lemma W.4. *The Green function, if it exists, satisfies the symmetry relation* $G(\mathbf{x},\mathbf{x}') = G(\mathbf{x}',\mathbf{x})$.

Proof. Since $G(\mathbf{x},\mathbf{x}') = 0$ for $\mathbf{x}'$ on S, Lemma W.5 shows that

$$\int_\Omega G(\mathbf{x},\mathbf{x}'')\nabla^2_{\mathbf{x}''} G(\mathbf{x}',\mathbf{x}'')\, d\mathbf{x}'' = \int_\Omega G(\mathbf{x}',\mathbf{x}'')\nabla^2_{\mathbf{x}''} G(\mathbf{x},\mathbf{x}'')\, d\mathbf{x}''.$$

Since $\nabla^2_{\mathbf{x}'} G(\mathbf{x},\mathbf{x}') = \delta(\mathbf{x}-\mathbf{x}')$, this gives

$$\int_\Omega G(\mathbf{x},\mathbf{x}'')\delta(\mathbf{x}'-\mathbf{x}'')\, d\mathbf{x}'' = \int_\Omega G(\mathbf{x}',\mathbf{x}'')\delta(\mathbf{x}-\mathbf{x}'')\, d\mathbf{x}'',$$

so that $G(\mathbf{x},\mathbf{x}') = G(\mathbf{x}',\mathbf{x})$. □

The construction of the Green's function $G(\mathbf{x},\mathbf{x}')$ depends on solving Laplace's equation (W.13) with boundary conditions (W.14). We do this using Fredholm theory.

Hilbert space

A *Hilbert space* V is a (usually infinite dimensional) complex vector space with inner product $\langle\ ,\ \rangle$ satisfying

(i) $\langle x, \lambda y_1 + \mu y_2\rangle = \lambda\langle x, y_1\rangle + \mu\langle x, y_2\rangle$,

(ii) $\langle x, y\rangle = \overline{\langle y, x\rangle}$ (and in particular $\langle x, x\rangle$ is real), and

(iii) $\langle x, x\rangle \geq 0$, and $\langle x, x\rangle = 0$ if and only if $x = 0$,

(iv) Writing $|x|$ for $\sqrt{\langle x, x\rangle}$, the metric with distance function $|x - y|$ is complete. In other words, every Cauchy sequence has a limit.

For example, if D is a compact domain in $\mathbb{R}^n$ then the space $L^2(D)$ of square integrable functions on D is a Hilbert space, with inner product

$$\langle f, g\rangle = \int_D \bar{f}\, g\, d\mathbf{x}.$$

In this example, the completeness is a standard fact from Lebesgue integration theory. In order to satisfy (iii), we stipulate that two functions are identified if they agree except on a set of measure zero. Of course, this never identifies two different continuous functions.

In terms of this inner product, we can write Lemma W.3 (with $\bar{f}$ in place of f) as follows.

LEMMA W.5. *Let $f(\mathbf{x})$ and $g(\mathbf{x})$ be twice continuously differentiable functions on Ω. Then $\langle f, \nabla^2 g\rangle = \langle \nabla^2 f, g\rangle$.* □

We shall often need to make use of the following inequality.

LEMMA W.6 (Schwartz's inequality). *For vectors x and y in Hilbert space, we have $|\langle x, y\rangle| \leq |x||y|$.*

PROOF. Consider the quantity

$$\langle x - ty, x - ty\rangle = |x|^2 - t\langle x, y\rangle - \bar{t}\langle y, x\rangle + t^2|y|^2 \geq 0.$$

Setting $t = \langle y, x\rangle/|y|^2$, we get

$$|x|^2 - 2|\langle x, y\rangle|^2/|y|^2 + |\langle x, y\rangle|^2/|y|^2 \geq 0,$$

or $|\langle x, y\rangle|^2/|y|^2 \leq |x|^2$. Now multiply by $|y|^2$ and take the square root to get $|\langle x, y\rangle| \leq |x||y|$. □

Elements x and y satisfying $\langle x, y\rangle = 0$ are said to be *orthogonal*. If W is a subspace of V, we write $W^\perp$ for the subspace consisting of vectors v such that for all $w \in W$ we have $\langle v, w\rangle = 0$. If W is finite dimensional, then any vector v in V can be written in a unique way as $v = w + x$ with w in W and x in $W^\perp$. So we have

$$V = W \oplus W^\perp.$$

If $\mathbf{K}$ is a linear operator on V, its *image* is

$$\mathrm{Im}\,(\mathbf{K}) = \{\mathbf{K}v,\ v \in V\}$$

and its *kernel* is

$$\mathrm{Ker}\,(\mathbf{K}) = \{v \in V \mid Kv = 0\}.$$

Operators $\mathbf{K}$ and $\mathbf{K}^*$ on V are said to be *adjoint* (to each other) if for all x and y in V we have

$$\langle \mathbf{K}^* x, y\rangle = \langle x, \mathbf{K}y\rangle.$$

LEMMA W.7. *If* $\mathbf{K}$ *and* $\mathbf{K}^*$ *are adjoint linear operators on* V *and the image of* $\mathbf{K}$ *is finite dimensional, then*

(i) $V = \mathrm{Im}\,\mathbf{K} \oplus \mathrm{Ker}\,\mathbf{K}^*$, *and*

(ii) $V = \mathrm{Im}\,\mathbf{K}^* \oplus \mathrm{Ker}\,\mathbf{K}$

are orthogonal direct sum decompositions of V, *and*

$$\dim \mathrm{Im}\,(\mathbf{K}) = \dim \mathrm{Im}\,(\mathbf{K}^*).$$

PROOF. If $\mathbf{K}^* x \in \mathrm{Im}\,(\mathbf{K}^*)$ and $y \in \mathrm{Ker}\,(\mathbf{K})$ then

$$\langle \mathbf{K}^* x, y\rangle = \langle x, \mathbf{K}y\rangle = 0$$

so $\mathrm{Im}\,(\mathbf{K}^*) \perp \mathrm{Ker}\,(\mathbf{K})$. If $x \in \mathrm{Im}\,(\mathbf{K}^*) \cap \mathrm{Ker}\,(\mathbf{K})$ then $\langle x, x\rangle = 0$ and so $x = 0$. Thus

$$\mathrm{Im}\,(\mathbf{K}^*) \oplus \mathrm{Ker}\,(\mathbf{K}) \leq V. \tag{W.16}$$

so we have

$$\dim \mathrm{Im}\,(\mathbf{K}) = \dim(V/\mathrm{Ker}\,(\mathbf{K})) \geq \dim \mathrm{Im}\,(\mathbf{K}^*), \tag{W.17}$$

with equality if and only if (W.16) is an equality. In particular, it follows that $\mathrm{Im}\,(\mathbf{K}^*)$ is also finite dimensional. So we may repeat the above argument with the roles of $\mathbf{K}$ and $\mathbf{K}^*$ reversed, so that

$$\mathrm{Im}\,(\mathbf{K}) \oplus \mathrm{Ker}\,(\mathbf{K}^*) \leq V \tag{W.18}$$

and

$$\dim \mathrm{Im}\,(\mathbf{K}^*) \geq \dim \mathrm{Im}\,(\mathbf{K}) \tag{W.19}$$

with equality if and only if (W.18) is an equality. Comparing (W.17) with (W.19), we see that both must be equalities, so (W.16) and (W.18) are equalities. □

LEMMA W.8. *If* $\mathbf{K}$ *and* $\mathbf{K}^*$ *are adjoint operators and* $\mathrm{Im}\,(\mathbf{K})$ *is finite dimensional then*

(i) $V = \mathrm{Im}\,(\mathbf{I} - \mathbf{K}) \oplus \mathrm{Ker}\,(\mathbf{I} - \mathbf{K}^*)$ *and*

(ii) $V = \mathrm{Im}\,(\mathbf{I} - \mathbf{K}^*) \oplus \mathrm{Ker}\,(\mathbf{I} - \mathbf{K})$

are orthogonal decompositions of V, *and* $\dim \mathrm{Ker}\,(\mathbf{I} - \mathbf{K}) = \dim \mathrm{Ker}\,(\mathbf{I} - \mathbf{K}^*)$ *is finite.*

PROOF. By Lemma W.7, $\mathrm{Im}\,(\mathbf{K}^*)$ is finite dimensional, so setting $V_1 = \mathrm{Im}\,(\mathbf{K}) + \mathrm{Im}\,(\mathbf{K}^*) \leq V$, we see that V_1 is also finite dimensional. So $V = V_1 \oplus V_2$ where

$$V_2 = V_1^{\perp} = \mathrm{Ker}\,(\mathbf{K}) \cap \mathrm{Ker}\,(\mathbf{K}^*).$$

So $\mathbf{I} - \mathbf{K}$ and $\mathbf{I} - \mathbf{K}^*$ send V_1 into V_1 and act as the identity map on V_2. Applying Lemma W.7 with $\mathbf{I} - \mathbf{K}$ instead of $\mathbf{K}$ and V_1 in place of V, we see that V_1 decomposes in the way described in the lemma. Since $\mathbf{I}-\mathbf{K}$ and $\mathbf{I}-\mathbf{K}^*$ act as the identity on V_2, this just contributes another summand to $\mathrm{Im}\,(\mathbf{I}-\mathbf{K})$ and $\mathrm{Im}\,(\mathbf{I}-\mathbf{K}^*)$, so the decomposition holds for V.

Since the dimensions of $\mathrm{Im}\,(\mathbf{I}-\mathbf{K})$ and $\mathrm{Im}\,(\mathbf{I}-\mathbf{K}^*)$ on V_1 are equal, and V_1 is finite dimensional, the dimensions of $\mathrm{Ker}\,(\mathbf{I}-\mathbf{K})$ and $\mathrm{Ker}\,(\mathbf{I}-\mathbf{K}^*)$ on V_1 must also be equal. But the kernels of these operators are contained in V_1, so this proves the last statement of the lemma. □

The Fredholm alternative

Now let V be the vector space $L^2(D)$ of Lebesgue square integrable functions on a compact domain D in $\mathbb{R}^n$. Suppose that $K(\mathbf{x},\mathbf{x}')$ is a continuous complex valued function of two variables $\mathbf{x}$ and $\mathbf{x}'$ in D. We are interested in the operator $\mathbf{K}$ on $L^2(D)$ given by

$$\mathbf{K}\psi(\mathbf{x}) = \int_D \psi(\mathbf{x}')K(\mathbf{x},\mathbf{x}')\,d\mathbf{x}'. \tag{W.20}$$

Such an operator is called a *Fredholm operator*, and the function $K(\mathbf{x},\mathbf{x}')$ is called the *kernel function*. The adjoint of $\mathbf{K}$ is given by

$$\mathbf{K}^*\psi(\mathbf{x}) = \int_D \psi(\mathbf{x}')\overline{K(\mathbf{x}',\mathbf{x})}\,d\mathbf{x}', \tag{W.21}$$

because

$$\langle\psi,\mathbf{K}\phi\rangle = \int_D\int_D \phi(\mathbf{x})\overline{\psi(\mathbf{x}')}K(\mathbf{x},\mathbf{x}')\,d\mathbf{x}\,d\mathbf{x}' = \langle\mathbf{K}^*\psi,\phi\rangle$$

(reverse the roles of $\mathbf{x}$ and $\mathbf{x}'$!). In general, the image of a Fredholm operator is not finite dimensional, so we can't apply Lemma W.8 directly. However, a *separable* function, namely one of the form $K(\mathbf{x},\mathbf{x}') = g(\mathbf{x})h(\mathbf{x}')$, gives rise to an operator $\mathbf{K}$ with one dimensional image spanned by $g(\mathbf{x})$. Any polynomial function of $\mathbf{x}$ and $\mathbf{x}'$ can be written as a finite sum of monomials, each of which has this form. So if $K(\mathbf{x},\mathbf{x}')$ is a polynomial function, we may apply Lemma W.8.

The Weierstrass approximation theorem states that any continuous function on a compact domain in $\mathbb{R}^n$ may be uniformly approximated by polynomial functions. Applying this to $K(\mathbf{x},\mathbf{x}')$ on $D\times D$, we may write $K = K_1 + K_2$ where K_1 is a polynomial function and K_2 satisfies $B < 1$, where B is defined by

$$B = \iint_D |K_2(\mathbf{x},\mathbf{x}')|^2\,d\mathbf{x}\,d\mathbf{x}'. \tag{W.22}$$

For any function $\psi(\mathbf{x})$ in $L^2(D)$, Schwartz's inequality (Lemma W.6) implies that for any $\mathbf{x}$ in D we have

$$|\mathbf{K}_2\psi(\mathbf{x})|^2 = \left|\int_D \psi(\mathbf{x}')K_2(\mathbf{x},\mathbf{x}')\,d\mathbf{x}'\right|^2 \le \langle\psi,\psi\rangle\int_D |K_2(\mathbf{x},\mathbf{x}')|^2\,d\mathbf{x}'.$$

Integrating with respect to $\mathbf{x}$ gives

$$\langle \mathbf{K}_2\psi, \mathbf{K}_2\psi\rangle \le B\langle\psi,\psi\rangle.$$

It follows by comparing with the geometric series

$$1 + B + B^2 + B^3 + \dots$$

that the sequence whose nth term is

$$\sum_{i=0}^{n} \mathbf{K}_2^i\psi$$

forms a Cauchy sequence in $L^2(D)$. Since $L^2(D)$ is complete, it follows that this Cauchy sequence has a limit; in other words, the infinite sum

$$\sum_{i=0}^{\infty} \mathbf{K}_2^i\psi = \psi + \mathbf{K}_2\psi + \mathbf{K}_2^2\psi + \mathbf{K}_2^3\psi + \cdots$$

converges in $L^2(D)$. It is now easy to check that the operator

$$\mathbf{I} + \mathbf{K}_2 + \mathbf{K}_2^2 + \mathbf{K}_2^3 + \dots$$

is an inverse to $\mathbf{I} - \mathbf{K}_2$ on $L^2(D)$. So we write $(\mathbf{I} - \mathbf{K}_2)^{-1}$ for this inverse. Similarly, $\mathbf{I} - \mathbf{K}_2^*$ is invertible, with inverse $\mathbf{I} + \mathbf{K}_2^* + (\mathbf{K}_2^*)^2 + (\mathbf{K}_2^*)^3 + \dots$

We use this to prove the following theorem, which is known as the *Fredholm alternative.*

THEOREM W.9. *With* $\mathbf{K}$ *and* $\mathbf{K}^*$ *defined by equations* (W.20) *and* (W.21)*, the kernels of* $\mathbf{I}-\mathbf{K}$ *and* $\mathbf{I}-\mathbf{K}^*$ *are finite dimensional, and have the same dimension. If this dimension is zero, then* $\mathbf{I}-\mathbf{K}$ *is invertible, so that the equation*

$$\psi - \mathbf{K}\psi = f$$

has a unique solution ψ *for any given element* f *of* $L^2(D)$.

PROOF. The idea is to make use of the identity

$$\mathbf{I} - \mathbf{K} = \mathbf{I} - (\mathbf{K}_1 + \mathbf{K}_2) = (\mathbf{I} - \mathbf{K}_2)(\mathbf{I} - (\mathbf{I} - \mathbf{K}_2)^{-1}\mathbf{K}_1).$$

Since K_1 is a polynomial, and hence a finite sum of separable functions, the image of $\mathbf{K}_1$ is finite dimensional. It follows that the image of $(\mathbf{I}-\mathbf{K}_2)^{-1}\mathbf{K}_1$ is also finite dimensional. So by Lemma W.8, $L^2(D)$ decomposes as a direct sum of the kernel of $\mathbf{I}-(\mathbf{I}-\mathbf{K}_2)^{-1}\mathbf{K}_1$, which has finite dimension, say d, and the image of $\mathbf{I}-((\mathbf{I}-\mathbf{K}_2)^{-1}\mathbf{K}_1)^*$. Since $\mathbf{I}-\mathbf{K}_2$ is invertible, the kernel of $\mathbf{I}-\mathbf{K}$ is the same as the kernel of $\mathbf{I}-(\mathbf{I}-\mathbf{K}_2)^{-1}\mathbf{K}_1$, and therefore has dimension d.

The adjoint of $\mathbf{I} - \mathbf{K}$ is

$$\mathbf{I} - \mathbf{K}^* = (\mathbf{I} - (\mathbf{I} - \mathbf{K}_2)^{-1}\mathbf{K}_1)^*(\mathbf{I} - \mathbf{K}_2)^*.$$

Since $(\mathbf{I}-\mathbf{K}_2)^* = \mathbf{I}-\mathbf{K}_2^*$ is also invertible, the kernel of $\mathbf{I}-\mathbf{K}^*$ has the same dimension as the kernel of $(\mathbf{I} - (\mathbf{I} - \mathbf{K}_2)^{-1}\mathbf{K}_1)^*$, which by Lemma W.8 is equal to d.

If the kernel of $\mathbf{I}-\mathbf{K}^*$ is zero then so is the kernel of $(\mathbf{I}-(\mathbf{I}-\mathbf{K}_2)^{-1}\mathbf{K}_1)^*$. So again applying Lemma W.8, it follows that the image of $\mathbf{I}-(\mathbf{I}-\mathbf{K}_2)^{-1}\mathbf{K}_1$

is the whole of $L^2(D)$. Since $\mathbf{I}-\mathbf{K}_2$ is invertible, it follows that the image of $\mathbf{I}-\mathbf{K}$ is also the whole of $L^2(D)$. In other words, the equation $\psi - \mathbf{K}\psi = f$ has a solution for every value of f. The solution is unique because the difference of two solutions is in the kernel of $\mathbf{I}-\mathbf{K}$, which is zero. $\square$

We have proved the Fredholm alternative under the condition that $K(\mathbf{x},\mathbf{x}')$ is continuous. Actually, we are going to want to use the theory for kernel functions K with singularities along $\mathbf{x}=\mathbf{x}'$ which are not too bad. The definition of "not too bad" depends on the dimension of D. In n dimensions, we allow kernel functions of the form $K(\mathbf{x},\mathbf{x}')=\kappa(\mathbf{x},\mathbf{x}')/|\mathbf{x}-\mathbf{x}'|^\alpha$ with $0\le\alpha<n$ and $\kappa(\mathbf{x},\mathbf{x}')$ continuous on $D\times D$. The point is that if Σ is a disc of radius ε around $\mathbf{x}$, then $\int_\Sigma |K(\mathbf{x},\mathbf{x}')|\,d\mathbf{x}'$ tends to zero as ε tends to zero. So we can approximate the value of K by a polynomial K_1 on the closed subset of $D\times D$ consisting of the points with $|\mathbf{x}-\mathbf{x}'|\ge\varepsilon$, and let K_2 absorb the singularity. In this way, we can arrange for ε to be small enough so that $B<1$, where B is defined in equation (W.22), and the arguments go through exactly as above.

Solving Laplace's equation

In the section on Green's functions (page 398), we saw that if we can solve Laplace's equation (W.13) with boundary conditions (W.14) then we can construct a Green's function $G(\mathbf{x},\mathbf{x}')$ satisfying equation (W.15) and zero on the boundary S. In this section we use Fredholm theory to solve Laplace's equation

$$\nabla^2\phi(\mathbf{x})=0 \tag{W.23}$$

subject to twice continuously differentiable boundary conditions $\phi(\mathbf{x})=f(\mathbf{x})$ on S.

We begin with uniqueness. We define the *potential energy* of a continuously differentiable function ϕ on Ω by

$$E=\rho c^2\int_\Omega \nabla\phi\,.\,\nabla\phi\,d\mathbf{x}.$$

So $E\ge 0$, and if $E=0$ then $\nabla\phi=0$, so that ϕ is constant. If ϕ_1 and ϕ_2 are solutions of (W.23) satisfying the same boundary conditions, then $\phi=\phi_1-\phi_2$ satisfies (W.23) and is zero on the boundary. By Green's first identity (W.4) with $f=g=\phi$, we see that we have $E=0$, so ϕ is constant; since $\phi=0$ on the boundary, this constant is zero. We conclude that if a solution to Laplace's equation (W.23) with given values on the boundary exists, then it is unique.

The same method can also be used for solutions of Laplace's equation (W.23) for the unbounded region Ω' obtained by removing the interior of Ω from $\mathbb{R}^2$, but we need to be careful about the behavior of ϕ as $\mathbf{x}$ goes off to infinity. The point is that we need to apply Green's first identity (W.4) for a region with a hole, bounded by S and a large circle S' of radius R surrounding Ω, and then let $R \to \infty$. The extra term we get from the second boundary component is $\int_{S'} \phi \, \nabla\phi \, . \big(\frac{\mathbf{x}}{R}\big)\, d\sigma$, because the unit normal vector is $\mathbf{x}/R$. The length of S' is $2\pi R$, so we need to check that $2\pi R|\phi\,\nabla\phi\,.\big(\frac{\mathbf{x}}{R}\big)| \to 0$ as $|\mathbf{x}| \to 0$. So we have proved the following theorem.

Theorem W.10. (i) *If $\nabla^2\phi = 0$ has a solution on Ω with specified values on S, then the solution is unique.*

(ii) *If $\nabla^2\phi = 0$ has a solution on Ω' with specified values on S, and satisfying*

$$\lim_{|\mathbf{x}|\to\infty} |\phi\,\nabla\phi\,.\,\mathbf{x}| = 0$$

then that solution is unique. □

We now examine the question of existence of solutions. To this end, we look for solutions of equation (W.23) of the form

$$\phi(\mathbf{x}) = \int_S \psi(\mathbf{x}')\nabla_{\mathbf{x}'} \ln|\mathbf{x}-\mathbf{x}'|\,.\,\mathbf{n}'\,d\sigma', \tag{W.24}$$

with ψ a twice continuously differentiable function defined on S.

Any twice continuously differentiable function ψ on S can be extended to a twice continuously differentiable function on Ω,[1] which we also denote by ψ. So we can use Green's first identity (W.4) with $f(\mathbf{x}') = \psi(\mathbf{x}')$ and $g(\mathbf{x}') = \ln|\mathbf{x}-\mathbf{x}'|$ to write

$$\phi(\mathbf{x}) = \int_\Omega (\psi(\mathbf{x}')\nabla^2_{\mathbf{x}'} \ln|\mathbf{x}-\mathbf{x}'| + \nabla\psi(\mathbf{x}')\,.\,\nabla_{\mathbf{x}'} \ln|\mathbf{x}-\mathbf{x}'|)\,d\mathbf{x}'.$$

By equation (W.12), we have

$$\phi(\mathbf{x}) = p(\mathbf{x})\psi(\mathbf{x}) + \int_\Omega \nabla\psi(\mathbf{x}')\,.\,\nabla_{\mathbf{x}'} \ln|\mathbf{x}-\mathbf{x}'|\,d\mathbf{x}'. \tag{W.25}$$

Now if Σ is a disc of radius ε around $\mathbf{x}$ then using (W.6) and changing variables to polar coordinates around $\mathbf{x}$, we have

$$\int_\Sigma |\nabla_{\mathbf{x}'} \ln|\mathbf{x}-\mathbf{x}'||\,d\mathbf{x}' = \int_0^{2\pi}\int_0^{\varepsilon} \frac{|\mathbf{r}|}{r^2}\,r\,dr\,d\theta = 2\pi\varepsilon. \tag{W.26}$$

Since $\nabla\psi(\mathbf{x}')$ is continuous, the singularity of the logarithm function can be excised with as small an effect as we please on the integral in equation (W.25). It follows that the integral term is continuous as $\mathbf{x}$ crosses the boundary S. Now $p(\mathbf{x})$ is discontinuous at S, so $\phi(\mathbf{x})$ is also discontinuous at S, and to solve Laplace's equation (W.23) using ϕ, we should use the limiting value at

[1] The function we're going to use for $\psi(\mathbf{x})$ is the logarithmic function $h(\mathbf{x},\mathbf{x}')$ of equation (W.14), which obviously extends to an open neighborhood of S, and therefore can be adjusted to extend in this manner over the whole of Ω.

the boundary rather than the actual value. Namely, for $\mathbf{x}_0$ in S and $\mathbf{x}$ in Ω but not in S, we have

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \phi(\mathbf{x}) = 2\pi\psi(\mathbf{x}_0) + \int_\Omega \nabla\psi(\mathbf{x}') \,.\, \nabla_{\mathbf{x}'} \ln|\mathbf{x}_0 - \mathbf{x}'|\, d\mathbf{x}',$$

whereas except at the corners, the value of ϕ on S is given by

$$\phi(\mathbf{x}_0) = \pi\psi(\mathbf{x}_0) + \int_\Omega \nabla\psi(\mathbf{x}') \,.\, \nabla_{\mathbf{x}'} \ln|\mathbf{x}_0 - \mathbf{x}'|\, d\mathbf{x}'.$$

So we have

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \phi(\mathbf{x}) = \phi(\mathbf{x}_0) + \pi\psi(\mathbf{x}_0).$$

In order to satisfy the boundary condition we want

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \phi(\mathbf{x}) = f(\mathbf{x}_0).$$

So we must solve the equation

$$\phi(\mathbf{x}) + \pi\psi(\mathbf{x}) = f(\mathbf{x}) \tag{W.27}$$

on S. Notice that the value of ψ at corners is irrelevant to the integral (W.24), so we just ignore the anomalous values of ϕ at corners and solve (W.27) for all $\mathbf{x}$ in S.

We rewrite equation (W.27) as

$$\psi(\mathbf{x}) + \frac{1}{\pi}\int_S \psi(\mathbf{x}')\nabla_{\mathbf{x}'} \ln|\mathbf{x} - \mathbf{x}'|.\mathbf{n}'\, d\sigma' = \frac{1}{\pi} f(\mathbf{x}). \tag{W.28}$$

Setting

$$K(\mathbf{x}, \mathbf{x}') = -\frac{1}{\pi}\nabla_{\mathbf{x}'} \ln|\mathbf{x} - \mathbf{x}'|.\mathbf{n}' = \frac{(\mathbf{x} - \mathbf{x}').\mathbf{n}'}{\pi|\mathbf{x} - \mathbf{x}'|^2}$$

and $D = S$, we use equation (W.20) to obtain an operator $\mathbf{K}$ on $L^2(S)$ given by

$$\mathbf{K}\psi(\mathbf{x}) = -\frac{1}{\pi}\int_S \psi(\mathbf{x}')\nabla_{\mathbf{x}'} \ln|\mathbf{x} - \mathbf{x}'|.\mathbf{n}'\, d\sigma'.$$

Equation (W.28) then becomes

$$\psi - \mathbf{K}\psi = \frac{1}{\pi} f. \tag{W.29}$$

The kernel function $K(\mathbf{x}, \mathbf{x}')$ has a singularity at $\mathbf{x} = \mathbf{x}'$; it is of the form $\kappa(\mathbf{x}, \mathbf{x}')/|\mathbf{x} - \mathbf{x}'|$, where κ is continuous. The Fredholm alternative (Theorem W.9) therefore applies for this function, by the argument described in the paragraph following the theorem. So equation (W.29) always has a solution provided we can prove that the only solution of the equation

$$\psi - \mathbf{K}\psi = 0$$

is the zero function. So assume that ψ satisfies this equation, and define $\phi(\mathbf{x})$ by equation (W.24). Then $\nabla^2\phi = 0$, and $\phi(\mathbf{x}) \to 0$ as $\mathbf{x}$ approaches the boundary from inside Ω. So by Theorem W.10 (i), we have $\phi(\mathbf{x}) = 0$ for $\mathbf{x}$ in Ω. Similarly, we define $\phi(\mathbf{x})$ by equation (W.24) on the unbounded region Ω'. Then using equation (W.6) we find that $|\phi\,\nabla\phi\,.\,\mathbf{x}| \to 0$ as $R \to \infty$. So by

Theorem W.10 (ii), we have $\phi(\mathbf{x}) = 0$ in Ω'. Now since $p(\mathbf{x})$ changes value by 2π as we cross from one side of the boundary to the other, it follows from equation (W.25) that for a point $\mathbf{x}_0$ on S

$$\lim_{\substack{\mathbf{x}\to\mathbf{x}_0\\ \text{in }\Omega}} \phi(\mathbf{x}) - \lim_{\substack{\mathbf{x}\to\mathbf{x}_0\\ \text{in }\Omega'}} \phi(\mathbf{x}) = 2\pi\psi(\mathbf{x}_0).$$

Since we've just shown that the left hand side is zero, it follows that $\psi(\mathbf{x}_0) = 0$. This completes the proof that the only solution of $\psi - \mathbf{K}\psi = 0$ is $\psi = 0$. Applying Fredholm theory as mentioned above, this completes the proof of existence of solutions of Laplace's equation. We summarize what we have proved in the following theorem.

THEOREM W.11. *Given any twice continuously differentiable function ψ on S, there exists a unique twice continuously differentiable function ϕ on Ω satisfying $\nabla^2\phi = 0$ and $\phi(\mathbf{x}) = \psi(\mathbf{x})$ on S.* □

Applying this theorem to equation (W.13) with boundary conditions (W.14) as promised, we obtain the existence of Green's functions. The following theorem summarizes the properties of Green's functions.

THEOREM W.12. *There exists a Green's function $G(\mathbf{x},\mathbf{x}')$, a function of two variables $\mathbf{x}$ and $\mathbf{x}'$ in Ω, satisfying*

(i) $G(\mathbf{x},\mathbf{x}') + \frac{1}{2\pi}\ln|\mathbf{x}-\mathbf{x}'|$ *is twice continuously differentiable,*

(i) $\nabla^2_{\mathbf{x}'}G(\mathbf{x},\mathbf{x}') = \delta(\mathbf{x},\mathbf{x}')$,

(ii) $G(\mathbf{x},\mathbf{x}') = G(\mathbf{x}',\mathbf{x})$, *and*

(iii) $G(\mathbf{x},\mathbf{x}') = 0$ *for $\mathbf{x}'$ on the boundary S of the region Ω.* □

Conservation of energy

We are now ready to begin proving existence and uniqueness for solutions of the wave equation (W.1). The basic tool for proving uniqueness of solutions is the conservation of energy. We define the *energy* $E(t)$ of a continuously differentiable function z of $\mathbf{x}$ and t to be the quantity

$$E(t) = \rho\int_\Omega \left(\left(\frac{\partial z}{\partial t}\right)^2 + c^2\nabla z.\nabla z\right)d\mathbf{x}. \tag{W.30}$$

The two terms in this integral correspond to kinetic and potential energy respectively. Since $E(t)$ is obtained by integrating a sum of squares, it satisfies $E(t) \geq 0$. Furthermore, $E(t) = 0$ can only occur if the integrand is zero; namely if $\frac{\partial z}{\partial t}$ and ∇z are zero.

Suppose that z satisfies the wave equation (W.1). Differentiating, and using the divergence theorem (W.2), we get

$$\begin{aligned}\frac{dE}{dt} &= \int_\Omega \rho\left(2\frac{\partial z}{\partial t}\frac{\partial^2 z}{\partial t^2} + 2c^2\nabla z\,.\frac{\partial\nabla z}{\partial t}\right)d\mathbf{x}\\ &= \int_\Omega \rho\left(2\frac{\partial z}{\partial t}c^2\nabla^2 z + 2c^2\nabla z\,.\,\nabla\frac{\partial z}{\partial t}\right)d\mathbf{x}\end{aligned}$$

$$= \int_\Omega 2\rho c^2 \nabla . \left(\frac{\partial z}{\partial t} \nabla z \right) d\mathbf{x}$$
$$= \int_S 2\rho c^2 \left(\frac{\partial z}{\partial t} \nabla z \right) . \mathbf{n}\, d\sigma.$$

Since $\frac{\partial z}{\partial t} = 0$ on S, we obtain

$$\frac{dE}{dt} = 0$$

so that E is a constant, independent of t. This is the statement of the conservation of energy for solutions of the wave equation.

Uniqueness of solutions

We now prove the uniqueness theorem for solutions to the wave equation. Suppose that z_1 and z_2 are solutions to the wave equation (W.1) on Ω, with the same initial conditions (i.e., the same values of z and $\frac{\partial z}{\partial t}$ for $t = 0$), and both vanishing on S. Then $z = z_1 - z_2$ satisfies the initial conditions $z = 0$ and $\frac{\partial z}{\partial t} = 0$ at $t = 0$. Equation (W.30) then shows that $E(0) = 0$. Conservation of energy implies that $E(t) = 0$ for all t. So $\frac{\partial z}{\partial t} = 0$ for all t, which implies that z is independent of t. Since it is zero at $t = 0$, we deduce that $z = 0$ for all values of t. Thus z_1 and z_2 are equal. It follows that there is at most one solution to the wave equation (W.1) for a given set of initial conditions for z and $\frac{\partial z}{\partial t}$.

It is less easy to prove existence of solutions. For this, we use the eigenvalue method. This will occupy the rest of the appendix.

Eigenvalues are nonnegative and real

We now prove that the eigenvalues of the Laplace operator $-\nabla^2$ are nonnegative and real—even if we allow f to take complex values (for real valued functions, ignore the bars in the proof of the lemma).

LEMMA W.13. *If f is a nonzero (complex valued) twice differentiable function on Ω satisfying $f = 0$ on the boundary S, then the quantity $\langle f, -\nabla^2 f\rangle$ is a nonnegative real number.*

PROOF. Let $\bar f$ be the complex conjugate of f. Then using Green's first identity (W.4), we have

$$\int_S (\bar f\, \nabla f) . \mathbf{n}\, d\sigma = \int_\Omega \nabla \bar f . \nabla f\, d\mathbf{x} + \int_\Omega \bar f (\nabla^2 f)\, d\mathbf{x}.$$

The left hand side vanishes because $f = 0$ on S, and the right hand side is $|\nabla f|^2 - \langle f, -\nabla^2 f\rangle$. So we have

$$\langle f, -\nabla^2 f\rangle = |\nabla f|^2, \tag{W.31}$$

which is nonnegative and real. □

In particular, if f is an eigenfunction of $-\nabla^2$ with eigenvalue λ, in other words, if

$$\nabla^2 f = -\lambda f,$$

then λ is a nonnegative real number. In fact, equation (W.31) shows that

$$\lambda = \frac{|\nabla f|^2}{|f|^2}.$$

This expression for λ is called *Rayleigh's quotient*.

Orthogonality

The relationship between ∇^2 and the inner product for functions on Ω is expressed in Lemma W.5, which says that for functions f and g vanishing on the boundary, ∇^2 is *self-adjoint* with respect to the inner product:

$$\langle f, \nabla^2 g\rangle = \langle \nabla^2 f, g\rangle.$$

This allows us to see easily why the eigenvalues of ∇^2 are real numbers (Lemma W.13). Namely if $\nabla^2 f = -\lambda f$, and $f(\mathbf{x}) = 0$ on the boundary S, then we have

$$\bar{\lambda}\langle f, f\rangle = \langle \lambda f, f\rangle = -\langle \nabla^2 f, f\rangle = -\langle f, \nabla^2 f\rangle = \langle f, \lambda f\rangle = \lambda\langle f, f\rangle.$$

Since $\langle f, f\rangle \neq 0$, we have $\lambda = \bar{\lambda}$. However, positivity is less easy to see from this point of view.

A similar argument shows that eigenfunctions with distinct eigenvalues are orthogonal, as in the following lemma.

LEMMA W.14. *Let f and g be Dirichlet eigenfunctions on Ω with eigenvalues λ and μ respectively. If $\lambda \neq \mu$ Then*

$$\langle f, g\rangle = 0.$$

PROOF. Using the fact that ∇^2 is self-adjoint (see Lemma W.5), we have

$$\lambda\langle f, g\rangle = \langle \nabla^2 f, g\rangle = \langle f, \nabla^2 g\rangle = \mu\langle f, g\rangle,$$

and so $(\lambda - \mu)\langle f, g\rangle = 0$. If $\lambda \neq \mu$, it follows that $\langle f, g\rangle = 0$. □

Inverting ∇^2

The key to understanding the eigenvalues and eigenfunctions of ∇^2 is to find an inverse $\mathbf{K}$ for the operator $-\nabla^2$ using Green's functions. The inverse is an integral operator with a wider domain of definition, and whose eigenvalues are the reciprocals of those for $-\nabla^2$. The operator $\mathbf{K}$ is an example of a *compact operator*, which is what makes the eigenvalue theory easier.

The construction of the inverse goes as follows. If $f(\mathbf{x})$ satisfies

$$\nabla^2 f(\mathbf{x}) = g(\mathbf{x})$$

on Ω and $f(\mathbf{x}) = 0$ on S, then using equation (W.15) and Green's second identity (W.5), for $\mathbf{x}$ in Ω but not on S, we have

$$\begin{aligned} f(\mathbf{x}) &= \int_\Omega f(\mathbf{x}')\delta(\mathbf{x}-\mathbf{x}')\,d\mathbf{x}' = \int_\Omega f(\mathbf{x}')\nabla^2 G(\mathbf{x},\mathbf{x}')\,d\mathbf{x}' \\ &= \int_\Omega G(\mathbf{x},\mathbf{x}')\nabla^2 f(\mathbf{x}')\,d\mathbf{x}' = \int_\Omega g(\mathbf{x}')G(\mathbf{x},\mathbf{x}')\,d\mathbf{x}'. \end{aligned}$$

So the operator sending $g(\mathbf{x})$ to $\int_\Omega g(\mathbf{x}')G(\mathbf{x},\mathbf{x}')\,d\mathbf{x}'$ undoes the effect of ∇^2.

We write $\mathbf{K}$ for the operator defined by

$$\mathbf{K}f(\mathbf{x}) = -\int_\Omega f(\mathbf{x}')G(\mathbf{x},\mathbf{x}')\,d\mathbf{x}'. \tag{W.32}$$

for $\mathbf{x}$ in Ω but not in S, and $\mathbf{K}f(\mathbf{x}) = 0$ for $\mathbf{x}$ in S. Then the above calculation shows that for twice continuously differentiable functions $f(\mathbf{x})$ which vanish on S, we have

$$f(\mathbf{x}) = -\mathbf{K}\nabla^2 f(\mathbf{x}). \tag{W.33}$$

Also, differentiating under the integral sign and using equation (W.15) shows that for any continuous function f on Ω, $\mathbf{K}f(\mathbf{x})$ is twice continuously differentiable, and we have

$$f(\mathbf{x}) = -\nabla^2\mathbf{K}f(\mathbf{x}). \tag{W.34}$$

So $\mathbf{K}$ and $-\nabla^2$ are inverse operators.

If $f(\mathbf{x})$ satisfies

$$\nabla^2 f(\mathbf{x}) = -\lambda f(\mathbf{x}) \tag{W.35}$$

on Ω and $f(\mathbf{x}) = 0$ on S, then we have

$$f(\mathbf{x}) = \lambda\mathbf{K}f(\mathbf{x}).$$

In particular, $f(\mathbf{x}) \neq 0$ implies $\lambda \neq 0$, so zero is not an eigenvalue of ∇^2. So if $f(\mathbf{x})$ satisfies (W.35) then

$$\mathbf{K}f(\mathbf{x}) = \frac{1}{\lambda}f(\mathbf{x}).$$

It follows that $f(\mathbf{x})$ is an eigenfunction of $\mathbf{K}$ with eigenvalue $1/\lambda$.

Conversely, if $f(\mathbf{x})$ is an eigenfunction of $\mathbf{K}$ then equation (W.34) shows that it has *nonzero* eigenvalue μ, and that it is also an eigenfunction of $-\nabla^2$ with eigenvalue $\lambda = 1/\mu$. Applying the equation repeatedly shows that any such eigenfunction $f(\mathbf{x})$ is *infinitely* differentiable.

LEMMA W.15. *If f is a continuous function on Ω then $\langle\mathbf{K}f, f\rangle$ is a nonnegative real number.*

PROOF. It follows from equation (W.34) and Lemma W.13 that

$$\langle\mathbf{K}f, f\rangle = \langle\mathbf{K}f, -\nabla^2\mathbf{K}f\rangle$$

is nonnegative and real. □

A nonzero self-adjoint operator $\mathbf{K}$ satisfying $\langle\mathbf{K}f, f\rangle \geq 0$ for all f is said to be *positive*.

LEMMA W.16. *If* $\mathbf{K}$ *is a self-adjoint operator on a Hilbert space* V*, and* $\langle \mathbf{K}x, x\rangle = 0$ *for all* x *in* V*, then* $\mathbf{K} = 0$.

PROOF. For all x and y in V we have

$$\begin{aligned} 0 = \langle \mathbf{K}(x+y), x+y\rangle &= \langle \mathbf{K}x, x\rangle + \langle \mathbf{K}x, y\rangle + \langle \mathbf{K}y, x\rangle + \langle \mathbf{K}y, y\rangle \\ &= \langle \mathbf{K}x, y\rangle + \langle x, \mathbf{K}y\rangle \\ &= 2\langle \mathbf{K}x, y\rangle. \end{aligned}$$

Given x in V, the fact that this holds for all y in V shows that $\mathbf{K}x = 0$. This is true for all x in V, so $\mathbf{K} = 0$. □

Compact operators

Let V be a Hilbert space. We say that a sequence of elements $x_1, x_2, \ldots$ of elements of V is *bounded* if there is some positive constant M such that all the x_i satisfy $|x_i| \le M$. A continuous operator $\mathbf{K}$ on V is said to be *compact* if, given any bounded sequence $x_1, x_2, \ldots$, the sequence of images $\mathbf{K}x_1, \mathbf{K}x_2, \ldots$ has a convergent subsequence.

Example. If the image of $\mathbf{K}$ is finite dimensional then the Bolzano–Weierstrass theorem implies that $\mathbf{K}$ is compact. More generally, the Fredholm alternative can be expressed in terms of compact operators.

THEOREM W.17. *If* $\mathbf{K}$ *is a compact positive self-adjoint operator then* $\mathbf{K}$ *has an eigenvalue* $\mu > 0$.

PROOF. There is an upper bound to the values of $\langle \mathbf{K}x, x\rangle$ as x runs over the elements of V satisfying $|x| = 1$. This is because otherwise, there would be a sequence $x_1, x_2, \ldots$ such that $\langle \mathbf{K}x_i, x_i\rangle > i$, and then by Schwartz's inequality (Lemma W.6), $\langle \mathbf{K}x_i, \mathbf{K}x_i\rangle > i^2$, so that there could not exist a convergent subsequence; this would contradict the fact that $\mathbf{K}$ is compact. Writing μ for the least upper bound of the values for $\langle \mathbf{K}x, x\rangle$ for $|x| = 1$, Lemma W.16 shows that $\mu > 0$.

We can find a sequence $x_1, x_2, \ldots$ of elements with $|x_i| = 1$, such that $\langle \mathbf{K}x_1, x_1\rangle, \langle \mathbf{K}x_2, x_2\rangle, \ldots$ converges to μ. Using Schwartz's inequality again, we have

$$\begin{aligned} \langle \mathbf{K}x_i - \mu x_i, \mathbf{K}x_i - \mu x_i\rangle &= \langle \mathbf{K}x_i, \mathbf{K}x_i\rangle - 2\mu\langle \mathbf{K}x_i, x_i\rangle + \mu^2 \\ &\le \langle \mathbf{K}x_i, x_i\rangle^2 - 2\mu\langle \mathbf{K}x_i, x_i\rangle + \mu^2 \\ &\le 2\mu^2 - 2\mu\langle \mathbf{K}x_i, x_i\rangle \\ &= 2\mu(\mu - \langle \mathbf{K}x_i, x_i\rangle) \\ &\to 0 \quad \text{as} \quad i \to \infty, \end{aligned}$$

and so $\mathbf{K}x_i - \mu x_i \to 0$ as $i \to \infty$.

Since $\mathbf{K}$ is compact, we can replace $x_1, x_2, \ldots$ by a subsequence with the property that $\mathbf{K}x_1, \mathbf{K}x_2, \ldots$ converges. So $\mu x_1, \mu x_2, \ldots$ converges, and since

$\mu \neq 0$, this implies that $x_1, x_2, \dots$ also converges. Setting $x = \lim_{i\to\infty} x_i$, the continuity of $\mathbf{K}$ implies that $\mathbf{K}x = \lim_{i\to\infty} \mathbf{K}x_i$, so we have

$$\mathbf{K}x = \mu x.$$

In other words, x is an eigenvector of $\mathbf{K}$ with eigenvalue μ. □

Remark. The method of proof of the above theorem finds the *largest* eigenvalue of $\mathbf{K}$. This is because if $\mu' \geq 0$ is any eigenvalue then an eigenvector x chosen with $|x| = 1$ will satisfy $\mu \geq \langle \mathbf{K}x, x\rangle = \mu'\langle x, x\rangle = \mu'$.

LEMMA W.18. *Let* $\mathbf{K}$ *be a compact operator. Then given any* $\varepsilon > 0$, *all but a finite number of the eigenvalues* μ *of* $\mathbf{K}$ *satisfy* $|\mu| < \varepsilon$. *The linear span of the eigenvectors with eigenvalue* $\geq \varepsilon$ *is finite dimensional.*

PROOF. If not, then there is an infinite sequence of orthogonal eigenvectors $x_1, x_2, \dots$ with $|x_i| = 1$, with eigenvalues μ_i satisfying $|\mu_i| \geq \varepsilon$. But then the sequence $\mathbf{K}x_1, \mathbf{K}x_2, \dots$ has the property that every pair of terms has distance $\geq \varepsilon$, and so it does not have a convergent subsequence, contradicting the definition of a compact operator. □

The inverse of ∇^2 is compact

THEOREM W.19. *The operator* $\mathbf{K}$ *defined in equation* (W.32), *which is inverse to* $-\nabla^2$, *is compact.*

PROOF. The argument is essentially due to Arzelà and Ascoli.[2] We are given a sequence of functions $f_1, f_2, \dots$, and we must show that the sequence $\mathbf{K}f_1, \mathbf{K}f_2, \dots$ has a convergent subsequence.

For this purpose, we begin by choosing a sequence of points $\mathbf{x}_1$, $\mathbf{x}_2, \mathbf{x}_3, \dots$ which are dense in Ω. Using Schwartz's inequality, we have

$$|\mathbf{K}f_i(\mathbf{x})|^2 \leq \int_\Omega |G(\mathbf{x}, \mathbf{x}')|^2 \, d\mathbf{x}'.$$

So $\max_{\mathbf{x}\in\Omega} |\mathbf{K}f_i(\mathbf{x})|$ is bounded, independent of i. It follows that we can choose a subsequence $f_{1,1}, f_{1,2}, f_{1,3}, \dots$ of the sequence $f_1, f_2, f_3, \dots$ such that $\mathbf{K}f_{1,1}(\mathbf{x}_1), \mathbf{K}f_{1,2}(\mathbf{x}_1), \mathbf{K}f_{1,3}(\mathbf{x}_1), \dots$ converges. Repeating this argument, we choose a subsequence $f_{2,1}, f_{2,2}, f_{2,3}, \dots$ of the sequence $f_{1,1}, f_{1,2}, f_{1,3}, \dots$ such that $\mathbf{K}f_{2,1}(\mathbf{x}_1), \mathbf{K}f_{2,2}(\mathbf{x}_1), \mathbf{K}f_{2,3}(\mathbf{x}_1), \dots$ converges. Continue this way, and then take the diagonal subsequence $f_{1,1}, f_{2,2}, f_{3,3}, \dots$ We claim that the sequence $\mathbf{K}f_{1,1}, \mathbf{K}f_{2,2}, \mathbf{K}f_{3,3}, \dots$ converges.

[2] What is usually referred to as the Arzelà-Ascoli theorem states that if a sequence of continuous functions on a compact set is uniformly bounded and equicontinuous then it has a uniformly convergent subsequence. This is the statement that is really being proved in this section. For further details, see Theorem IV.6.7 and the notes and remarks at the end of Chapter IV of Dunford and Schwartz, *Linear Operators, Part I*, Wiley Interscience, 1967; or Theorem 43 in Section 5.4 of Colton [**20**].

To prove this, we argue as follows. Using Schwartz's inequality again, for $\mathbf{y}$ and $\mathbf{z}$ in Ω we have

$$|\mathbf{K}f_i(\mathbf{y}) - \mathbf{K}f_i(\mathbf{z})| \le \int_\Omega |G(\mathbf{y},\mathbf{x}') - G(\mathbf{z},\mathbf{x}')|\,d\mathbf{x}'.$$

So given $\varepsilon > 0$, we can choose $\delta > 0$ (independent of i) such that if $|\mathbf{y}-\mathbf{z}| < \delta$ then $|\mathbf{K}f_i(\mathbf{y}) - \mathbf{K}f_i(\mathbf{z})| < \varepsilon$.

Now choose M large enough so that every point of Ω is within δ of one of the points $\mathbf{x}_1, \dots, \mathbf{x}_M$. Choose N large enough so that

$$|\mathbf{K}f_{m,m}(\mathbf{x}_i) - \mathbf{K}f_{n,n}(\mathbf{x}_i)| < \varepsilon$$

for $m, n \ge N$ and $1 \le i \le M$. Then for $\mathbf{x} \in \Omega$, choose $\mathbf{x}_i$ within δ of $\mathbf{x}$. We have

$$\begin{aligned} |\mathbf{K}f_{m,m}(\mathbf{x}) - \mathbf{K}f_{n,n}(\mathbf{x})| &\le |\mathbf{K}f_{m,m}(\mathbf{x}) - \mathbf{K}f_{m,m}(\mathbf{x}_i)| \\ &\quad + |\mathbf{K}f_{m,m}(\mathbf{x}_i) - \mathbf{K}f_{n,n}(\mathbf{x}_i)| \\ &\quad + |\mathbf{K}f_{n,n}(\mathbf{x}_i) - \mathbf{K}f_{n,n}(\mathbf{x})| \\ &< 3\varepsilon. \end{aligned}$$

This proves that the sequence $\mathbf{K}f_{n,n}$ converges, as claimed, and completes the proof that $\mathbf{K}$ is compact. □

Eigenvalue stripping

Let $\mathbf{K}$ be a compact positive self-adjoint operator on an infinite dimensional Hilbert space V. We have an inductive procedure for finding eigenvalues, which goes as follows. Suppose that we have found orthogonal eigenvectors $x_1, \dots, x_n$ of $\mathbf{K}$ with eigenvalues $\mu_1 \ge \mu_2 \ge \cdots \ge \mu_n$, and that for all $x \in V$, $\langle \mathbf{K}x, x\rangle \le \mu_n$. Then we define

$$\mathbf{K}_n x = \mathbf{K}x - \sum_{i=1}^{n} \mu_i \langle x, x_i\rangle x_i.$$

Then $\mathbf{K}_n x_i = 0$ for $1 \le i \le n$, and if x is orthogonal to x_i for all $1 \le i \le n$ then $\mathbf{K}_n x = \mathbf{K}x$. So the eigenvalues of $\mathbf{K}_n$ are the same as those of $\mathbf{K}$, except that $\mu_1, \dots, \mu_n$ have been replaced by zero. It is easy to check that $\mathbf{K}_n$ is either a compact positive self-adjoint operator or it is the zero operator. Now we apply Theorem W.17 to the operator $\mathbf{K}_n$, provided it is nonzero, to find an eigenvector x_{n+1} for its largest eigenvalue μ_{n+1}, which is necessarily $\le \mu_n$, and form the operator $\mathbf{K}_{n+1}$ as above.

This process either stops at some finite stage with $\mathbf{K}_n = 0$, in which case $\mathbf{K}$ has zero as an eigenvalue, or we find an infinite sequence of eigenvalues $\mu_1 \ge \mu_2 \ge \cdots$. By Lemma W.18, we have

$$\lim_{n\to\infty} \mu_n = 0.$$

The convergence of the sum

$$\sum_{i=1}^{\infty} \mu_i \langle x, x_i \rangle x_i$$

is a consequence of the fact that the μ_i are bounded, together with *Bessel's inequality*, which is as follows.

LEMMA W.20 (Bessel's inequality). *If $x_1, x_2, \ldots$ are orthogonal elements of a Hilbert space V with $|x_i| = 1$, then for any $x \in V$ we have*

$$\sum_{i=1}^{\infty} |\langle x, x_i \rangle|^2 \leq |x|^2.$$

PROOF. Set $y_n = \sum_{i=1}^{n} \langle x, x_i \rangle x_i$, $z_n = x - y_n$. Then

$$\begin{aligned}\langle y_n, z_n \rangle &= \left\langle \sum_{i=1}^{n} \langle x, x_i \rangle x_i, x \right\rangle - \left\langle \sum_{i=1}^{n} \langle x, x_i \rangle x_i, \sum_{i=1}^{n} \langle x, x_i \rangle x_i \right\rangle \\ &= \sum_{i=1}^{n} |\langle x, x_i \rangle|^2 - \sum_{i=1}^{n} |\langle x, x_i \rangle|^2 = 0.\end{aligned}$$

So $|x|^2 = |y_n|^2 + |z_n|^2 \geq |y_n|^2 = \sum_{i=1}^{n} |\langle x, x_i \rangle|^2$. This holds for all $n \geq 1$, and so the lemma is proved. $\square$

Now set

$$\mathbf{K}_\infty x = \mathbf{K}x - \sum_{i=1}^{\infty} \mu_i \langle x, x_i \rangle x_i.$$

Then $\mathbf{K}_\infty$ is either zero or compact, positive and self-adjoint. By Lemma W.18, given any $\varepsilon > 0$, all its eigenvalues are bounded above by ε. So applying Theorem W.17, we see that the only possibility is that $\mathbf{K}_\infty = 0$. So we have the following equation.

$$\mathbf{K}x = \sum_{i=1}^{\infty} \mu_i \langle x, x_i \rangle x_i. \tag{W.36}$$

To summarize, if $\mathbf{K}$ is a compact positive self-adjoint operator on an infinite dimensional Hilbert space V, then either equation (W.36) holds, where x_i are eigenvectors with strictly positive real eigenvalues $\mu_1 \geq \mu_2 \geq \cdots$ satisfyint $\lim_{n\to\infty} \mu_n = 0$, or a similar equation holds with just a finite sum. In the latter case, $\mathbf{K}$ has zero as an eigenvalue.

Solving the wave equation

We are finally ready to show existence of solutions of the wave equations with given initial conditions. Let $\mathbf{K}$ be defined by equation (W.32), so that $\mathbf{K}$ and $-\nabla^2$ are inverse operators by equations (W.33) and (W.34). By Theorem W.19, $\mathbf{K}$ is compact. Since it is inverse to $-\nabla^2$, it does not have zero as an eigenvalue. So equation (W.36) applies to $\mathbf{K}$. Namely, there is a sequence of infinitely differentiable orthogonal eigenfunctions $f_1, f_2, \ldots$ of $\mathbf{K}$

with strictly positive eigenvalues $\mu_1 \geq \mu_2 \geq \dots$ satisfying $\lim_{n\to\infty} \mu_n = 0$. In particular, for any $f \in L^2(\Omega)$, the sum

$$\sum_{i=1}^{\infty} \langle f, f_i \rangle f_i$$

converges in $L^2(\Omega)$ by Bessel's inequality, and the function

$$f_\infty = f - \sum_{i=1}^{\infty} \langle f, f_i \rangle f_i$$

has the property that $\mathbf{K} f_\infty = 0$, so $f_\infty = 0$. It follows that we have

$$f = \sum_{i=1}^{\infty} \langle f, f_i \rangle f_i, \qquad \mathbf{K} f = \sum_{i=1}^{\infty} \mu_i \langle f, f_i \rangle f_i,$$

and so

$$-\nabla^2 f = \sum_{i=1}^{\infty} \lambda_i \langle f, f_i \rangle f_i$$

where $\lambda_i = 1/\mu_i$ are the eigenvalues of $-\nabla^2$, with the same eigenfunctions f_i as $\mathbf{K}$.

Now suppose that we wish to solve the wave equation (W.1) on Ω with initial conditions $z(\mathbf{x}, 0) = f(\mathbf{x})$ and $\frac{\partial z}{\partial t}(\mathbf{x}, 0) = g(\mathbf{x})$. Set

$$z(\mathbf{x}, t) = \sum_{i=1}^{\infty} f_i(\mathbf{x}) \left(\langle f, f_i \rangle \cos(c\sqrt{\lambda_i}\, t) + \frac{\langle g, f_i \rangle}{c\sqrt{\lambda_i}} \sin(c\sqrt{\lambda_i}\, t) \right). \tag{W.37}$$

Then $z(\mathbf{x}, 0) = \sum_{i=1}^{\infty} \langle f, f_i \rangle f_i(\mathbf{x}) = f(\mathbf{x})$ and $\frac{\partial z}{\partial t}(\mathbf{x}, 0) = \sum_{i=1}^{\infty} \langle g, f_i \rangle f_i(\mathbf{x}) = g(\mathbf{x})$, so the initial conditions are satisfied. It is an easy exercise to show that z also satisfies the wave equation (W.1). We proved uniqueness on page 407, and so this is the unique function with these properties.

Polyhedra and finite groups

In this section, we consider what happens if we allow ourselves to take a finite set of polygonal regions in $\mathbb{R}^2$ and glue them together using distance preserving linear maps along the edges, to form a polyhedron Ω. We allow at most two faces to meet at an edge, so that Ω is a 2-dimensional manifold, possibly with boundary. The operator ∇^2 on this manifold comes from the individual faces, matched along the edges. We also assume that we have a finite group G acting on Ω in such a way that each group element takes each face isometrically to the same face or another face of Ω, and that if it is taken to the same face then the isometry is the identity map. If H is a subgroup of G, then the quotient Ω/H is also a polyhedron in which the faces are orbits of H on the faces of Ω.

In order to deal with the possibility that an element $g \in G$ takes a face to an adjacent face, we give each face an orientation in such a way that adjacent faces have opposite orientations, and we assume that the action of G

preserves orientation. The effect of this is that if there is an element $g \in G$ which swaps two faces glued along an edge, then G-invariant functions vanish along that edge. So H-invariant functions on Ω vanishing on the boundary correspond to functions on Ω/H vanishing along the boundary.

Imagine that we have already found the Dirichlet eigenspaces of ∇^2 on Ω. We write V_λ for the eigenspace corresponding to the eigenvalue λ. So V_λ is a finite dimensional complex vector space. Then each element $g \in G$ transports eigenfunctions of ∇^2 on Ω to eigenfunctions with the same eigenvalue, and induces a linear map from V_λ to itself. This way, we get a linear representation of G on V_λ; namely a homomorphism $\phi\colon G \to GL(V_\lambda)$, where $GL(V_\lambda)$ is the *general linear* group of invertible linear transformations on V_λ.

If H is a subgroup of G, then the eigenfunctions of ∇^2 on Ω/H are the H-invariant elements of V_λ, denoted V_λ^H. Now $\frac{1}{|H|}\sum_{h\in H}\phi(H)$ is a matrix which sends each element of V_λ to an H-invariant element, and which acts as the identity map on the H-invariant elements. So its trace is the dimension of V_λ^H,

$$\dim V_\lambda^H = \frac{1}{|H|}\sum_{h\in H}\operatorname{Tr}(h, V_\lambda).$$

Now conjugate elements of G have the same trace on V_λ, so we can divide up the above sum into contributions from the conjugacy classes of G.

$$\dim V_\lambda^H = \frac{1}{|H|}\sum_{\substack{\text{conj. classes}\\ \mathcal{C}_g \text{ of elements of } G}} |\mathcal{C}_g \cap H|\operatorname{Tr}(g, V_\lambda).$$

The upshot of this computation is that if H_1 and H_2 are two subgroups of G with the property that for each conjugacy class $\mathcal{C}$ in G we have

$$|\mathcal{C}\cap H_1| = |\mathcal{C}\cap H_2|$$

then for all λ we have $\dim V_\lambda^{H_1} = \dim V_\lambda^{H_2}$. We summarize this in the following theorem, essentially due to Sunada.[3]

THEOREM W.21. *Let H_1 and H_2 be subgroups of G such that for each conjugacy class $\mathcal{C}$ of elements of G we have*

$$|\mathcal{C}\cap H_1| = |\mathcal{C}\cap H_2|.$$

Then the Dirichlet eigenvalues of ∇^2 and their multiplicities on Ω/H_1 and Ω/H_2 coincide. □

An example

To find inequivalent drums with the same resonant frequencies (see Section 3.7), we apply Theorem W.21 to construct planar regions with the

[3] T. Sunada, *Riemannian coverings and isospectral manifolds*, Ann. Math. 121 (1985), 169–186.

same Dirichlet spectrum.[4] We need to begin by choosing a finite group G with subgroups H_1 and H_2 which are not conjugate in G, but which satisfy the hypothesis of the theorem. An example is $G = GL(3, \mathbb{F}_2)$, the general linear group of invertible matrices with entries in the field of two elements $\mathbb{F}_2 = \{0, 1\}$. This group has 168 elements, and it has subgroups H_1 and H_2 of order 24 consisting of the matrices of the form $\left(\begin{smallmatrix} * & * & * \\ * & * & * \\ 0 & 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & * & * \\ 0 & * & * \\ 0 & * & * \end{smallmatrix}\right)$ respectively. The left cosets of H_1 and H_2 in G correspond to nonzero row vectors and column vectors of length three respectively.

Let T be a triangle in $\mathbb{R}^2$ with acute angles and three edges of different lengths, colored red, blue and yellow. We construct Ω from 168 triangles T_g, one for each $g \in G$, each one of which is a copy of T. Let r, b and y be the following elements of G:

$$r = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad b = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \qquad y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

It is easy to check that these matrices satisfy the following relations:

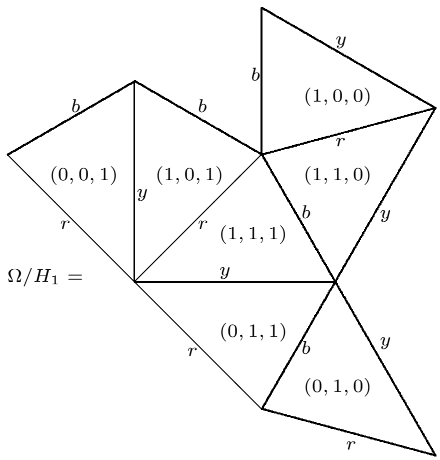
$$r^2 = b^2 = y^2 = 1, \qquad (rb)^4 = (by)^4 = (yr)^4 = 1.$$

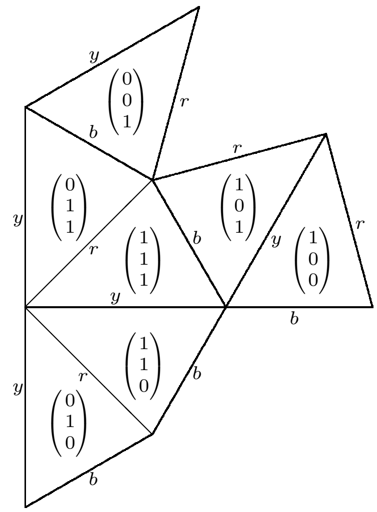
We glue a triangle T_g along its red edge to T_{gr}, along its blue edge to T_{gb}, and along its yellow edge to T_{gy}, in such a way that adjacent triangles have opposite orientations. The above relations between r, b and y imply that there are eight triangles around each vertex. The resulting polyhedron Ω has 168 faces, $\frac{3}{2} \times 168 = 252$ edges and $\frac{3}{8} \times 168 = 63$ vertices.[5] The action of G on Ω is given by the formula $h(T_g) = T_{hg}$. It is easy to check that this action preserves the way that the faces are glued along the edges.

Each of Ω/H_1 and Ω/H_2 has $168/24 = 7$ triangular faces, and each of them embeds in the plane, but the configuration of faces is different. So these are examples of inequivalent drums with the same Dirichlet spectrum.

[4]The example described in this section is an elaboration of an example taken from Peter Buser, John Conway, Peter Doyle and Dieter Semmler, *Some planar isospectral domains*, International Mathematics Research Notices (1994), 391–400.

[5]In particular, the Euler characteristic of Ω is $168 - 252 + 63 = -21$, which is odd. So Ω is not orientable; it is a connected sum of 23 real projective planes.

$\Omega/H_1 =$ 

$\Omega/H_2 =$ 

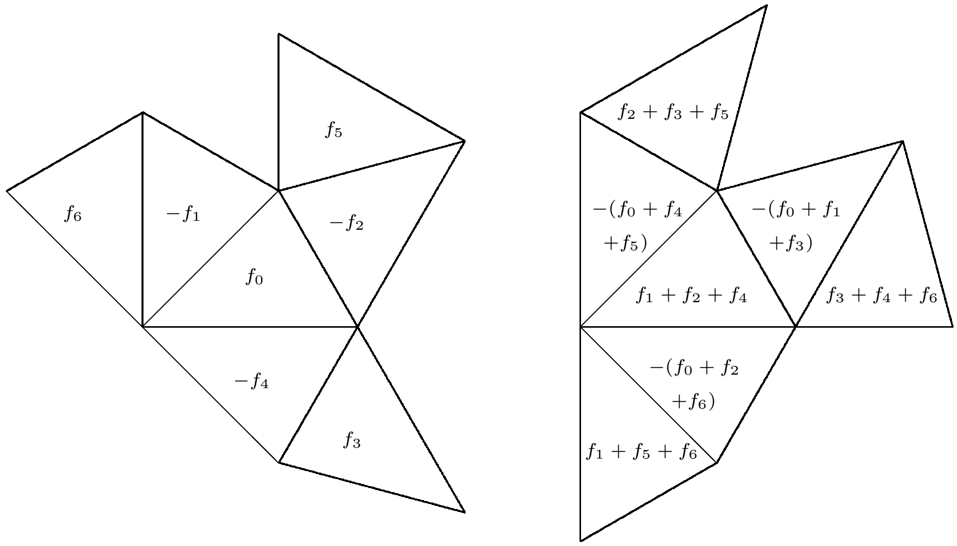
The method described above can even be used to give an explicit correspondence between eigenfunctions of ∇^2 on Ω/H_1 and Ω/H_2 (Bérard). Take a vector space $\mathbb{C}[G/H_1]$ whose basis elements are the left cosets of H_1 in G, and let G permute these basis elements by left multiplication. This gives a matrix representation of G on $\mathbb{C}[G/H_1]$ in which the matrices have the property that each row and each column have one entry equal to 1 and the rest equal to zero. Doing the same with H_2, we obtain representations $\phi_1\colon G \to GL(\mathbb{C}[G/H_1])$ and $\phi_2\colon G \to GL(\mathbb{C}[G/H_2]$. The hypothesis of Theorem W.21 can be expressed by saying that for each group element $g \in G$, we have $\mathrm{Tr}(g, \mathbb{C}[G/H_1]) = \mathrm{Tr}(g, \mathbb{C}[G/H_2])$. Character theory of finite groups[6] implies that there is an invertible linear map $\psi\colon \mathbb{C}[G/H_1] \to \mathbb{C}[G/H_2]$ such that for all $g \in G$ and $v \in \mathbb{C}[G/H_1]$ we have $\phi_2(g)(\psi(v)) = \psi(\phi_1(g)(v))$. Such a map ψ can be used to create eigenfunctions on Ω/H_2 out of eigenfunctions on Ω/H_1. One way of explaining this is that Frobenius reciprocity gives an isomorphism $V_\lambda^{H_1} \cong \mathrm{Hom}_G(\mathbb{C}[G/H_1], V_\lambda)$ (and similarly for H_2) so that

$$V_\lambda^{H_2} \cong \mathrm{Hom}_G(\mathbb{C}[G/H_2], V_\lambda) \cong \mathrm{Hom}_G(\mathbb{C}[G/H_1], V_\lambda) \cong V_\lambda^{H_1},$$

where the middle isomorphism is given by composition with ψ.

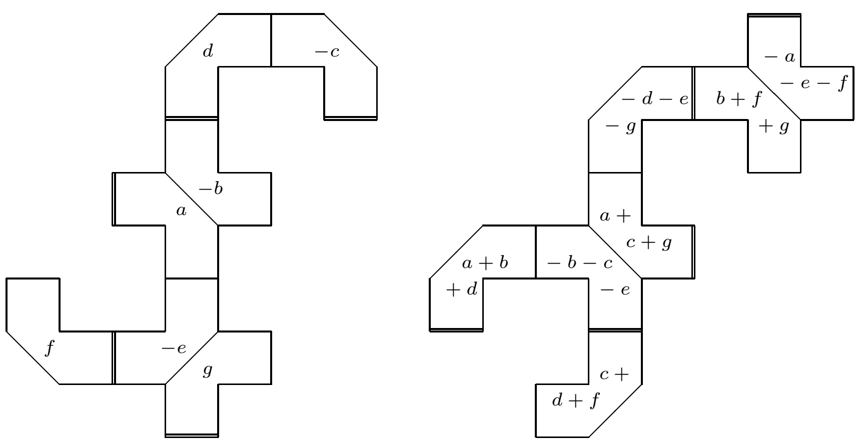
In the example above, one possible choice for ψ takes the basis element of $\mathbb{C}[G/H_1]$ corresponding to a length three row vector (α, β, γ) to the sum of the three basis elements of $\mathbb{C}[G/H_2]$ corresponding to the three column vectors (u, v, w) satisfying $\alpha u + \beta v + \gamma w = 0$. So taking the orientations into account, the correspondence between eigenfunctions is given by the following diagram.

[6] See for example G. D. James and M. Liebeck, *Representations and characters of groups*, 2nd edition, Cambridge University Press, 2001.



Even without knowing how this example was constructed, it is easy to check that this recipe works. It is necessary to notice that if an eigenfunction which is zero on the boundary were continued beyond the boundary, it would get negated and reflected. So for example, as we go from the middle region of Ω/H_2 to the neighbor below it, f_1 gets replaces by $-f_6$, f_2 gets replaced by $-f_2$, and f_4 gets replaced by $-f_0$. This matches with the given value.

This kind of check can be used for the example of Gordon, Webb and Wolpert in Section 3.7 too. Here is the recipe for transporting eigenfunctions.



This example is based on the same group and subgroups, but with a different choice of elements of order two for the gluing of faces.

Other choices of G with pairs of nonconjugate subgroups H_1 and H_2 satisfying the condition of Theorem W.21 include the following.

(i) G is the semidirect product $\mathbb{Z}/8 \rtimes (\mathbb{Z}/8)^{\times}$ where $(\mathbb{Z}/8)^{\times}$ is the multiplicative group $\{1, 3, 5, 7\}$ of the invertible numbers modulo eight, which acts as the automorphism group of $\mathbb{Z}/8$ by multiplication. The subgroups are $H_1 = \{(0, 1), (0, 3), (0, 5), (0, 7)\}$ and $H_2 = \{(0, 1), (4, 3), (4, 5), (0, 7)\}$.

More generally, we can let $G = K \rtimes H$, any semidirect product with nonconjugate complements H_1 and H_2 for K in G, but where each element of H_1 is conjugate to the corresponding element of H_2.

(ii) G is the symmetric group on six letters, a group of order 720, $H_1 = \{(12)(34), (13)(24), (14, 23)\}$ and $H_2 = \{(12)(34), (12)(56), (34)(56)\}$. This example works with the same choice of H_1 and H_2, with G equal to the alternating group of degree six.

More generally, if H_1 and H_2 are two nonisomorphic groups of order n with the same number of elements of each order, then the regular permutation representation embeds H_1 and H_2 as subgroups of the symmetric group on n letters, which is the choice for G.

(iii) $G = PSL(3, \mathbb{F}_3)$, H_1 and H_2 representatives of the two conjugacy classes of subgroups of index 13.

(iv) $G = GL(4, \mathbb{F}_2)$, H_1 and H_2 representatives of the two conjugacy classes of subgroups of index 15.

(v) $G = PSL(3, \mathbb{F}_4)$, H_1 and H_2 representatives of the two conjugacy classes of subgroups of index 21.

Further reading:

P. Bérard, *Transplantation et isospectralité, I*, Math. Ann. 292 (1992), 547–559.

P. Buser, J. H. Conway, P. Doyle and K.-D. Semmler, *Some planar isospectral domains*, International Mathematics Research Notices (1994), 391–400.

D. Colton, *Partial differential equations, an introduction* [**20**].

R. Courant and D. Hilbert, *Methods of mathematical physics, I*, Chapters III and V, Interscience, 1953.

C. Gordon, D. Webb and S. Wolpert, *Isospectral plane domains and surfaces via Riemannian orbifolds*, Invent. Math. 110 (1992), 1–22.

R. Guralnick, *Subgroups inducing the same permutation representation*, J. Algebra 81 (1983), 312–319.

T. Sunada, *Riemannian coverings and isospectral manifolds*, Ann. of Math. 121 (1985), 169–186.

Bibliography

1. G. Assayag, H. G. Feichtinger, and J. F. Rodrigues (eds.), *Mathematics and music, a Diderot mathematical forum*, Springer-Verlag, Berlin/New York, 2002. 288 pages, in print. ISBN 3540437274.

   This collection of essays comes from the Fourth Diderot Mathematical Forum, held under the auspices of the European Mathematical Society in 1999. The essays are as follows. 1. M. P. Ferreira, *Proportions in ancient and medieval music*. 2. E. Knoblauch, *The sounding algebra: relations between combinatorics and music from Mersenne to Euler*. 3. B. Scimemi, *The use of mechanical devices and numerical algorithms in the 18th century for the equal temperament of the musical scale*. 4. J. Dhombres, *Lagrange,* "Working Mathematician" *on music considered as a source for science*. 5. W. Hodges and R. J. Wilson, *Musical patterns*. 6. F. Nicholas, *Questions of logic: writing, dialectics and musical strategies*. 7. M.-J. Durand-Richard, *The formalization of logic and the issue of meaning*. 8. L. Fichet, *Musical analysis using mathematical proceedings in the XXth century*. 9. S. Dubnov and G. Assayag, *Universal prediction applied to stylistic music generation*. 10. M. Chemillier, *Ethnomusicology, ethnomathematics. The logic underlying orally transmitted artistic practices*. 11. M. Leman, *Expressing coherence of musical perception in formal logic*. 12. G. Mazzola, *The topos geometry of musical logic*. 13. J.-C. Risset, *Computing musical sound*. 14. E. Neuwirth, *The mathematics of tuning musical instruments – a simple toolkit for experiments*. 15. X. Serra, *The musical communication chain and its modeling*. 16. G. De Poli and D. Rocchesso, *Computational models for musical sound sources*.

2. Pierre-Yves Asselin, *Musique et tempérament*, Éditions Costallat Paris, 1985; reprinted by Jobert, 1997. 236 pages. ISBN 2905335009.

   "Music and temperament." Pierre-Yves Asselin is a canadian organist, who studied music at McGill University and later with Marie-Claire Alain in Paris. This book, written in French, starts with a few pages of explanation of harmonics, intervals and beats. The rest of the book describes various scales and temperaments, and gives instructions for how to tune them. The last chapter gives historical examples of pieces intended for various temperaments. The appendices give extensive tables of various scales in both cents and savarts. You can obtain the reprinted version of this book directly from the publisher by emailing info@jobert.fr.

3. John Backus, *The acoustical foundations of music*, W. W. Norton & Co., 1969. Reprinted 1977. 384 pages, in print. ISBN 0393090965.

   This book gives a non-technical discussion of the physical basis for acoustics, the ear, and the production of sound in musical instruments. Very readable.

4. Denis Baggi, *Readings in computer generated music*, IEEE Computer Society Press Tutorial, 1992. 248 pages, in print. ISBN 0818627476

   Available from www.barnesandnoble.com..

5. Patrice Bailhache, *Une histoire de l'acoustique musicale*, CNRS Éditions, 2001. 195 pages, in print. ISBN 2271058406.

   "A history of musical acoustics." This French book can be ordered from www.amazon.fr.

The six chapters cover Greek antiquity, the renaissance, the classical age, the enlightenment, Helmholtz, and the twentieth century.

6. J. Murray Barbour, *Tuning and temperament, a historical survey*, Michigan State College Press, E. Lansing, 1951. Reprinted by Da Capo Press, New York, 1972. 228 pages, out of print. ISBN 0306704226.

   Setting a high standard for academic excellence, this book is a standard source on scales and temperaments, and their history. It compares and contrasts Pythagorean tuning, just intonation, meantone, irregular temperaments, and finally equal temperament. Barbour displays a strong predisposition towards twelve tone equal temperament in this work, and interprets the history of scales and temperaments as an inexorable march towards equal temperament.

7. Scott Beall, *Functional melodies: finding mathematical relationships in music*, Key Curriculum Press, 2000. 170 pages, in print. ISBN 1559533781.

   This is one of the few books on mathematics and music aimed at high school level; the only other one that I'm aware of is Garland and Kahn [**43**]. It comes with a CD of musical examples.

8. James Beament, *The violin explained: components, mechanism, and sound*, Oxford University Press, 1997. 245 pages, in print. ISBN 0198167393 (pbk), 0198166230 (hbk).

9. ———, *How we hear music: the relationship between music and the hearing mechanism*, The Boydell Press, Woodbridge, Suffolk, 2001. 174 pages, in print. ISBN 0851159400 (pbk), 0851158137 (hbk).

10. Georg von Békésy, *Experiments in hearing*, McGraw-Hill, New York/Toronto/London, 1960. 745 pages, in print.

    Von Békésy is responsible for the classical experiments in the functioning of the cochlea. This book may be obtained directly from the Acoustical Society of America at http://asa.aip.org.

11. Arthur H. Benade, *Fundamentals of musical acoustics*, Oxford Univ. Press, 1976. Reprinted by Dover, 1990. 596 pages, in print. ISBN 048626484X.

    Arthur Benade (1925–1987) made numerous contributions to the physics of music. This classic book mostly concerns the physics of musical instruments as well as the human voice.

12. Richard E. Berg and David G. Stork, *The physics of sound*, Prentice-Hall, 1982. Second edition, 1995. 416 pages, in print. ISBN 0131830473.

    A nicely presented textbook on elementary acoustics, musical instruments, and the human ear and voice.

13. Easley Blackwood, *The structure of recognizable diatonic tunings*, Princeton University Press, 1985. 318 pages, out of print. ISBN 0691091293.

    This book discusses various just, meantone and equal temperaments, and tries a little too hard to be mathematical about it. Example: THEOREM 16. *The number of greater* $(j+1)$*ths occurring in the diatonic scale is* h *where* $5j \equiv h \pmod 7$ *and* $0 \le h \le 6$.

14. Richard Charles Boulanger (ed.), *The CSound book: perspectives in software synthesis, sound design, signal procesing, and programming*, MIT Press, 2000. 782 pages, in print. ISBN 0262522616.

    CSound is a multiplatform *free* software synthesis program. It's hard to use at first, but is the most powerful thing around. In other words, CSound is to synthesis as TEX is to mathematical typesetting. Almost every synthesis technique you've ever heard of is implemented in a very flexible fashion. The first version came out in 1985, and it has been developing steadily since. This book contains separate articles by many authors, so there is something of a lack of overall coherence to the work. It comes with 2

CD-ROMs containing software for Mac, Linux and PC, hundreds of musical compositions, more than 3000 working instruments, and much more. There is a third CD-ROM available separately, called "The Csound Catalog with Audio," available from http://www.csounds.com. This CD-ROM contains over 2000 orchestra and score text files, and the corresponding audio files in mp3 format. It is also possible to order separately an updated version of the 2 CD-ROMs that came with the book, from the same web site, whether or not you own the book.

15. Pierre Buser and Michel Imbert, *Audition*, MIT Press, 1992. 406 pages, in print. ISBN 0262023318.

16. Murray Campbell and Clive Greated, *The musician's guide to acoustics*, OUP, 1986, reprinted 1998. 613 pages, in print. ISBN 0198165056.

    A well written account of acoustics for the musician, requiring essentially no mathematical background. This book is in print in the UK but not the USA, so try for example www.amazon.co.uk.

17. Peter Castine, *Set theory objects: abstractions for computer-aided analysis and composition of serial and atonal music*, European University Studies, vol. 36, Peter Lang Publishing, 1994. In print. ISBN 3631478976.

18. John Chowning and David Bristow, *FM Theory and applications*, Yamaha Music Foundation, 1986. 195 pages, out of print. ISBN 4636174828.

    This short book came out a couple of years after the Yamaha DX7 became available. It describes FM synthesis using the DX7 for the details of the examples. Note that the graphs for the Bessel functions J_{10} and J_{11} on page 176 have apparently been accidentally interchanged.

19. Thomas Christensen (ed.), *The Cambridge history of western music theory*, Cambridge University Press, 2002. 998 pages, in print. ISBN 0521623715.

    This is a very nice collection of essays on various historical aspects of music theory. It includes (among others) the following essays: Thomas Mathiesen, *Greek music theory*; Rudolph Rasch, *Tuning and temperament*; Penelope Gouk, *The role of harmonics in the scientific revolution*; Catherine Nolan, *Music theory and mathematics*.

20. David Colton, *Partial differential equations, an introduction*, Random House, 1988. 308 pages. ISBN 0394358279.

    This book contains a good treatment of the solution of the wave equation, complete with the background from functional analysis necessary for the proof. The existence of a complete set of eigenfunctions can be found on page 233. A C^2 boundary is assumed, but only in order to solve Laplace's equation with logarithmic boundary conditions, for the construction of Green's functions.

21. Perry R. Cook (ed.), *Music, cognition, and computerized sound. An introduction to psychoacoustics*, MIT Press, 1999. 392 pages, in print. ISBN 0262032562.

    This is an excellent collection of essays on various aspects of psychoacoustics, written by some of the leading figures in the area of computer music. It comes with a CD full of sound examples.

    Chapter headings: 1. Max Mathews, *The ear and how it works.* 2. Max Mathews, *The auditory brain.* 3. Roger Shepard, *Cognitive psychology and music.* 4. John Pierce, *Sound waves and sine waves.* 5. John Pierce, *Introduction to pitch perception.* 6. Max Mathews, *What is loudness?* 7. Max Mathews, *Introduction to timbre.* 8. John Pierce, *Hearing in time and space.* 9. Perry R. Cook, *Voice physics and neurology.* 10. Roger Shepard, *Stream segregation and ambiguity in audition.* 11. Perry R. Cook, *Formant peaks and spectral valleys.* 12. Perry R. Cook, *Articulation in speech and sound.* 13. Roger Shepard, *Pitch perception and measurement.* 14. John Pierce, *Consonance and scales.* 15. Roger Shepard, *Tonal structure and scales.* 16. Perry R. Cook, *Pitch, periodicity, and noise in the voice.* 17. Daniel J. Levitin, *Memory for musical attributes.*

18. Brent Gillespie, *Haptics*. 19. Brent Gillespie, *Haptics in manipulation*. 20. John Chowning, *Perceptual fusion and auditory perspective*. 21. John Pierce, *Passive nonlinearities in acoustics*. 22. John Pierce, *Storage and reproduction of music*. 23. Daniel J. Levitin, *Experimental design in psychoacoustic research*.

22. Deryck Cooke, *The language of music*, Oxford Univ. Press, 1959, reprinted in paperback, 1990. 289 pages, in print. ISBN 0198161808

    This wonderful little book explains how the basic elements of musical expression communicate emotional content, both locally and on a larger scale. Highly recommended to anyone trying to understand how music works. Deryck Cooke is the person who orchestrated Mahler's tenth symphony, starting with Mahler's original draft. Take a listen to the excellent Bournemouth Symphony/Simon Rattle recording.

23. David H. Cope, *New directions in music*, Wm. C. Brown Publishers, Dubuque, Iowa, Fifth edition, 1989. Sixth edition, Waveland Press, 1998. 439 pages, in print. ISBN 0697033422.

    An introduction to computers and the avant-garde in twentieth century music. Reads a bit like a scrapbook of ideas, pictures and music.

24. ———, *Computers and musical style*, Oxford University Press, 1991. 246 pages, in print. ISBN 019816274X.

    David Cope is well known for his attempts to induce computers to compose music in the style of various famous composers such as Bach and Mozart. Unsurprisingly, the compositions are not an unqualified success, but the account of the process presented in this book is interesting.

25. ———, *Experiments in musical intelligence*, Computer Music and Digital Audio, vol. 12, A-R Editions, Madison, Wisconsin, 1996. 263 pages, in print. ISBN 0895793148/0895793377.

    This book is a continuation of the project described in Cope's 1991 book, and comes with a CD-ROM full of examples for the Macintosh platform. I have not seen a copy, but from the review in Computer Music Journal 21 (3) (1997), it seems that the subject has progressed a good deal since [**24**] appeared in 1991. Artificial intelligence is still in a very primitive stage of development, and it will probably take another generation to produce a computational model which convincingly simulates one of the great composers. And then another generation after that, to compose with real originality. I think the real core of the problem is that when a human being composes, a hugely complex world view is invoked, which has taken a lifetime to accumulate. We'll end up teaching a baby computer how to talk before it grows up to be a real composer! But I'm glad that someone of the calibre of Cope is battling with these problems.

26. ———, *Virtual music*, MIT press, 2001. 565 pages, in print. ISBN 026203283X

    The saga continues....

27. Lothar Cremer, *The physics of the violin*, MIT Press, 1984. 450 pages, in print. ISBN 0262031027.

    Translation of *Physik der Geige*, S. Hirzel Verlag, Stuttgart, 1981. This book is the standard reference on the physics of the violin. The technical standard is high and the writing is clear. Strongly recommended. The other book to look at is Beament [**8**].

28. Malcolm J. Crocker (ed.), *Handbook of acoustics*, Wiley Interscience, 1998. 1461 pages, large format, in print. ISBN 047125293X.

    This enormous volume consists of 114 chapters by various experts, arranged in parts by subject. The subjects are: I. General linear acoustics, II. Nonlinear acoustics and cavitation, III. Aeroacoustics and atmospheric sound, IV. Underwater sound, V. Ultrasonics, quantum acoustics, and physical effects of sound, VI. Mechanical vibrations and shock, VII. Statistical methods in acoustics, VIII. Noise: its effects and control,

IX. Architectural acoustics, X. Acoustic signal processing, XI. Physiological acoustics, XII. Psychological acoustics, XIII. Speech communication, XIV. Music and musical acoustics, XV. Acoustic measurements and instrumentation, XVI. Transducers. Part XIV is particularly relevant, and consists of an introduction by Thomas Rossing; *Stringed instruments: bowed*, by J. Woodhouse; *Woodwind instruments*, by Neville H. Fletcher; *Brass instruments*, by J. M. Bowsher; and *Pianos and other stringed keyboard instruments*, by Gabriel Weinreich.

29. Alain Daniélou, *Sémantique musicale. Essai de psycho-physiologie auditive*, Hermann, Paris, 1967. Reprinted 1978, 131 pages, in print. ISBN 270561334X.

    "Musical semantics. Essay on auditive psycho-physiology." This French book can be obtained from www.amazon.fr.

30. ———, *Music and the power of sound*, Inner Traditions, Rochester, Vermont, 1995, revised from a 1943 publication. 172 pages, in print. ISBN 0892813369.

    This is a book about tuning and scales in different cultures, especially Chinese, Indian and Greek, and their effect on the emotional content of music. The original 1943 version was entitled *Introduction to the study of musical scales*, and published by the India Society, London. This original version has been reprinted by Munshiram Manoharlal Publishers Pvt. Ltd., New Delhi, 1999, 279 pages, in print. ISBN 8121509203.

31. Peter Desain and Henkjan Honig, *Music, mind and machine: Studies in computer music, music cognition, and artificial intelligence (Kennistechnologie)*, Thesis Publishers, 1992. 330 pages, in print. ISBN 9051701497.

32. Diana Deutsch (ed.), *The psychology of music*, Academic Press, 1982; 2nd ed., 1999. 807 pages, in print. ISBN 0122135652 (pbk), 0122135644 (hbk).

    This is an excellent collection of essays on various aspects of the psychology of music, by some of the leading figures in the field. The second edition has been completely revised to reflect recent progress in the subject. It is interesting to compare this collection of essays with Perry Cook's [**21**], which have a slightly different purpose.

    Chapter headings: 1. John R. Pierce, *The nature of musical sound.* 2. Manfred R. Schroeder, *Concert halls: from magic to number theory.* 3. Norman M. Weinberger, *Music and the auditory system.* 4. Rudolf Rasch and Reinier Plomp, *The perception of musical tones.* 5. Jean-Claude Risset and David L. Wessel, *Exploration of timbre by analysis and synthesis.* 6. Johan Sundberg, *The perception of singing.* 7. Edward M. Burns, *Intervals, scales and tuning.* 8. W. Dixon Ward, *Absolute pitch.* 9. Diana Deutsch, *Grouping mechanisms in music.* 10. Diana Deutsch, *The processing of pitch combinations.* 11. Jamshed J. Bharucha, *Neural nets, temporal composites, and tonality.* 12. Eugene Narmour, *Hierarchical expectation and musical style.* 13. Eric F. Clarke, *Rhythm and timing in music.* 14. Alf Gabrielson, *The performance of music.* 15. W. Jay Dowling, *The development of music perception and cognition.* 16. Rosamund Shuter-Dyson, *Musical ability.* 17. Oscar S. M. Marin and David W. Perry, *Neurological aspects of music perception and performance.* 18. Edward C. Carterette and Roger A. Kendall, *Comparative music perception and cognition.*

33. B. Chaitanya Deva, *The music of India: A scientific study*, Munshiram Manoharlal Publishers Pvt. Ltd., 1981. 278 pages, in print.

    This, and most other books on Indian music, are hard to get hold of. But there is a wonderful little bookstore called "Bazaar of India" at 1810 University Avenue in Berkeley, California which keeps copies of a dozen or more of them, including this one, in stock at very reasonable prices. Call them at 510-548-4110.

34. Dominique Devie, *Le tempérament musical: philosophie, histoire, théorie et practique*, Société de musicologie du Languedoc Béziers, 1990. 540 pages, out of print. ISBN 2905400528.

    "Musical temperament: philosophy, history, theory and practise." This French book is an extensive discussion of scales and temperaments, with a great deal of historical information and philosophical discussion.

35. Charles Dodge and Thomas A. Jerse, *Computer music: synthesis, composition, and performance*, Simon & Schuster, Second ed., 1997. 453 pages, in print. ISBN 0028646827 (pbk), 002873100X (hbk).
36. W. Jay Dowling and Dane L. Harwood, *Music cognition*, Academic Press Series in Cognition and Perception, 1986. 258 pages. ISBN 0122214307.
37. William C. Elmore and Mark A. Heald, *Physics of waves*, McGraw-Hill, 1969. Reprinted by Dover, 1985. 477 pages, in print. ISBN 0486649261.

    This book contains a useful discussion of waves on strings, rods and membranes.
38. Hans G. Feichtinger and Gerard Assayag (eds.), *Mathematics and music: a Diderot mathematical forum*, Springer/Verlag Berlin/New York, 2002. 288 pages, in print. ISBN 3540437274.

    This is a collection of essays on mathematics and music.
39. Laurent Fichet, *Les théories scientifiques de la musique aux XIX^e et XX^e siècles*, Librairie J. Vrin, 1996. 382 pages, in print. ISBN 2711642844.

    "Nineteenth and twentieth century scientific theories of music." This French book may be obtained from www.amazon.fr.
40. Neville H. Fletcher and Thomas D. Rossing, *The physics of musical instruments*, Springer-Verlag, Berlin/New York, 1991. ISBN 3540941517 (pbk), 3540969470 (hbk).

    This book is at a high technical level, and contains a wealth of interesting material. A difficult read, but worth the effort.
41. Allen Forte, *The structure of atonal music*, Yale Univ. Press, 1973. ISBN 0300021208.

    This book is about 12-tone music, and goes into a great deal of technical detail about the theory of pitch class sets, relations and complexes.
42. Steve De Furia and Joe Scacciaferro, *MIDI programmer's handbook*, M & T Publishing, Inc., 1989.
43. Trudi Hammel Garland and Charity Vaughan Kahn, *Math and music: harmonious connections*, Dale Seymore Publications, 1995. ISBN 0866518290.

    This book is aimed at high school level, and avoids technical material. It looks as though it would make good classroom material at the intended level. The only other book with this aim on the market that I am aware of is Beall [**7**].
44. H. Genevois and Y. Orlarey, *Musique & mathématiques*, Aléas–Grame, 1997. 194 pages, in print. ISBN 2908016834.

    "Music and mathematics." A collection of essays in French on various aspects of the connections between music and mathematics, coming out of the Rencontres Musicales Pluridisciplinaires at Lyons, 1996. This book can be ordered from www.amazon.fr.
45. Ben Gold and Nelson Morgan, *Speech and audio signal processing: processing and perception of speech and music*, Wiley & Sons, 2000. 537 pages, in print. ISBN 0471351547.

    The basic purpose of this book is to understand sound well enough to be able to perform speech recognition, but it contains a lot of material relevant to music recognition and synthesis. By some quirk of international pricing, the price of this book in the UK is about half what it is in the USA, so it may be worth your while checking out UK online bookstores such as amazon.co.uk or the UK branch of bol.com for this one.
46. Heinz Götze and Rudolf Wille (eds.), *Musik und Mathematik. Salzburger Musikgespräch 1984 unter Vorsitz von Herbert von Karajan*, Springer-Verlag, Berlin/New York, 1995. ISBN 3540154078

"Music and mathematics. Musical dialogue, Salzburg 1984, under the direction of Herbert von Karajan." A collection of essays, mostly in german.

47. Penelope Gouk, *Music, science and natural magic in seventeenth-century England*, Yale University Press, New Haven, 1999. 308 pages, in print. ISBN 0300073836.

48. Karl F. Graff, *Wave motion in elastic solids*, Oxford University Press, 1975. Reprinted by Dover, 1991. ISBN 0486667456.

    This book contains a lot of information about wave motion in strings, bars and plates, relevant to Chapter 3.

49. Niall Griffith and Peter M. Todd (eds.), *Musical networks: parallel distributed perception and performance*, MIT Press, 1999. 350 pages, in print. ISBN 0262071819.

50. Donald E. Hall, *Musical acoustics*, Wadsworth Publishing Company, Belmont, California, 1980. ISBN 0534007589.

    This book has some good chapters on the physics of musical instruments, as well as briefer acounts of room acoustics and of tuning and temperament.

51. R. W. Hamming, *Digital filters*, Prentice Hall, 1989. Reprinted by Dover Publications. 296 pages, in print. ISBN 048665088X

    Hamming is one of the pioneers of twentieth century communications and coding theory. This book on digital filters is a classic.

52. G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, Oxford University Press, Fifth edition, 1980. 426 pages, in print. ISBN 0198531710.

    This classic contains a good section on the theory of continued fractions, which may be used as a reference for the material presented in §6.2.

53. W. M. Hartmann, *Signals, sound and sensation*, Springer-Verlag, Berlin/New York, 1998. 647 pages, in print. ISBN 1563962837

    This book contains a very nice discussion of psychoacoustics, Fourier theory and digital signal processing, and the relationships between these subjects.

54. Hermann Helmholtz, *Die Lehre von den Tonempfindungen*, Longmans & Co., Fourth German edition, 1877. Translated by Alexander Ellis as *On the sensations of tone*, Dover, 1954 (and reprinted many times). 576 pages, in print. ISBN 0486607534.

    For anyone interested in scales and temperaments, or the history of acoustics and psychoacoustics, this book is an absolute gold mine. The appendices by the translator are also full of fascinating material. Strongly recommended.

55. Michael Hewitt, *The tonal Phoenix; a study of tonal progression through the prime numbers three, five and seven*, Verlag für systematische Musikwissenschaft GmbH, Bonn, 2000. 495 pages, in print. ISBN 3922626963.

    This German book (in English) should be available from www.amazon.de, but it doesn't yet seem to be listed.

56. Douglas R. Hofstadter, *Gödel, Escher, Bach*, Harvester Press, 1979. Reprinted by Basic Books, 1999. 777 pages, in print. ISBN 0465026567.

    A nice popularized account of the connections between mathematical logic, cognitive science, Escher's art and the music of J. S. Bach. A bit too longwinded to make a particularly good read, but fun for the occasional dip.

57. David M. Howard and James Angus, *Acoustics and psychoacoustics*, Focal Press, 1996. 365 pages, in print. ISBN 0240514289.

58. Hua, *Introduction to number theory*, Springer-Verlag, Berlin/New York, 1982. ISBN 3540108181.

    This book contains a good section on continued fractions, which may be used as a supplement to §6.2. Be warned that the continued fraction for π given on page 252 of Hua is erronious. The correct continued fraction can be found here on page 194.

59. Stuart M. Isacoff, *Temperament: The idea that solved music's greatest riddle*, Knopf, 2001; paperback 2003. 288 pages in small format, in print. ISBN 0375403558 (hbk), 0375703306 (pbk).

This is a chatty popularized account of the history of musical temperament. The style is very readable, and the information density is low.

60. Sir James Jeans, *Science & music*, Cambridge Univ. Press, 1937. Reprinted by Dover, 1968. 273 pages, in print. ISBN 0486619648.

Somewhat old fashioned, but still makes an interesting read.

61. Franck Jedrzejewski, *Mathématiques des systèmes acoustiques*, L'Harmattan, Paris, 2002. 367 pages, in print. ISBN 2747521966.

"Mathematics of acoustic systems." In spite of its title, this book is about musical temperament.

62. Jeffrey Johnson, *Graph theoretical methods of abstract musical transformation*, Greenwood Publishing Group, 1997. 216 pages, in print. ISBN 0313301581.

63. Tom Johnson, *Self-similar melodies*, Editions 75, 75 rue de la Roquette, 75011 Paris, 1996. 291 pages, ring-bound, in print. ISBN 2907200011.

Tom Johnson is a minimalist composer, whose work uses mathematical techniques such as the theory of automata to assist in the compositional process. Copies of this book may be obtained by writing to: Two Eighteen Press, PO Box 218, Village Station, New York, NY 10014, USA.

64. Ian Johnston, *Measured tones: The interplay of physics and music*, Institute of Physics Publishing, Bristol and Philadelphia, 1989. Reprinted 1997. 408 pages, in print. ISBN 0852742363.

This very readable book is about acoustics and the physics of musical instruments, from a historical perspective, and with essentially no equations.

65. Owen H. Jorgensen, *Tuning*, Michigan State University Press, 1991. 798 pages, large format, out of print. ISBN 0870132903.

This enormous book is subtitled: "Containing The Perfection of Eighteenth-Century Temperament, The Lost Art of Nineteenth-Century Temperament, and The Science of Equal Temperament, Complete With Instructions for Aural and Electronic Tuning." It is a mixture of history of tunings and temperaments, and explicit tuning instructions for various temperaments. An interesting thread running through the book is a detailed argument to the effect that equal temperament was not commonplace until the twentieth century.

66. Michael Keith, *From polychords to Pólya; adventures in musical combinatorics*, Vinculum Press, Princeton, New Jersey, 1991. 166 pages, in print. ISBN 0963009702.

This book describes the applications of Pólya's enumeration theorem to the combinatorics of chords, scales and keys, as in §9.15. Throughout, the author deals with the cyclic group consisting of the twelve musical transpositions. Unfortunately, atonal music theorists such as Allen Forte and Elliott Carter all seem to use the dihedral group of order 24 obtained by allowing inversion. Nonetheless, the ideas presented in the book can be applied just as easily in this case.

67. Lawrence E. Kinsler, Austin R. Frey, Alan B. Coppens, and James V. Sanders, *Fundamentals of acoustics*, John Wiley & Sons, Fourth edition, 2000. 548 pages, in print. ISBN 0471847895.

This is an excellent technical book on acoustics, and deservedly popular. The two original authors of the first (1950) edition were Kinsler and Frey, both now deceased. The book has gone through many print runs and editions. Coppens and Sanders have updated the book and added new material for the fourth edition. This is another book whose price in the UK is about half what it is in the USA, so it may be worth your while checking out the UK online bookstores for this one.

68. T. W. Körner, *Fourier analysis*, Cambridge Univ. Press, 1988, reprinted 1990. 591 pages, in print. ISBN 0521389917.

    This book makes great reading. There is a fair amount of high level mathematics, but also a number of sections of a more historical or narrative nature, and a wonderful sense of humor pervades the work. The account of the laying of the transatlantic cable in the nineteenth century and the technical problems associated with it is priceless. Several sections are devoted to the life of Fourier. There is also a companion volume entitled *Exercises for Fourier analysis*, ISBN 0521438497, in print.

69. Patricia Kruth and Henry Stobart (eds.), *Sound*, Cambridge Univ. Press, 2000. 235 pages, in print. ISBN 0521572096.

    A nice collection of nontechnical essays on the nature of sound. I particularly like Jonathan Ashmore's contribution. Contents: 1. Philip Peek, *Re-sounding Silences*. 2. Charles Taylor, *The Physics of Sound*. 3. Jonathan Ashmore, *Hearing*. 4. Peter Slater, *Sounds Natural: The Song of Birds*. 5. Peter Ladefoged, *The Sounds of Speech*. 6. Christopher Page, *Ancestral Voices*. 7. Brian Ferneyhough, *Shaping Sound*. 8. Steven Feld, *Sound Worlds*. 9. Michel Chion, *Audio-Vision and Sound*.

70. Albino Lanciani, *Mathématiques et musique. Les Labyrinthes de la phénoménologie*, Éditions Jérôme Millon, Grenoble, 2001. 275 pages, in print. ISBN 2841371131.

    "Mathematics and music. The labyrinths of phenomenology." This French book can be obtained from www.amazon.fr. It is an extended essay based around Bach's *Musical Offering* and mathematical logic, among other subjects. There are some obvious parallels between this book and Hofstadter's [**56**].

71. J. Lattard, *Gammes et tempéraments musicaux*, Masson, Paris, 1988. 130 pages, in print. ISBN 2225812187.

    "Scales and musical temperaments." This French book can be obtained from www.amazon.fr.

72. Marc Leman, *Music and schema theory: cognitive foundations of systematic musicology*, Springer Series on Information Science, vol. 31, Springer-Verlag, Berlin/New York, 1995. In print. ISBN 3540600213.

73. ______, *Music, Gestalt, and computing; studies in cognitive and systematic musicology*, Lecture Notes in Computer Science, vol. 1317, Springer-Verlag, Berlin/New York, 1997. 524 pages, in print. ISBN 3540635262.

    This book of conference proceedings comprises a collection of essays about the interactions between music, psychoacoustics, cognitive science and computer science. There is an accompanying CD of sound examples.

74. Ernő Lendvai, *Symmetries of music*, Kodály Institute, Kecskemét, 1993. 155 pages, in print. ISBN 9637295100.

    This book is a translation of a Hungarian book with the title *Szimmetria a zenében*. It seems to be quite hard to get hold of. I suggest going to the Kodály Institute web site at www.kodaly-inst.hu and emailing them.

75. David Lewin, *Generalized musical intervals and transformations*, Yale University Press, New Haven/London, 1987. ISBN 0300034938.

    This book discusses twelve tone music from a mathematical point of view, using some elementary group theory.

76. Carl E. Linderholm, *Mathematics made difficult*, Wolfe Publishing, Ltd., London, 1971. 207 pages, out of print.

    This book isn't relevant to the subject of the text, but is well worth digging out to pass a happy evening. The humor gets slightly heavy-handed at times, but this is balanced by some priceless moments.

77. Mark Lindley and Ronald Turner-Smith, *Mathematical models of musical scales*, Verlag für systematische Musikwissenschaft GmbH, Bonn, 1993. 308 pages, out of print. ISBN 3922626661.

78. Llewelyn S. Lloyd and Hugh Boyle, *Intervals, scales and temperaments*, Macdonald, London, 1963. 246 pages, out of print.

    An extensive discussion of just intonation, meantone and equal temperament.

79. R. Duncan Luce, *Sound and hearing, a conceptual introduction*, Lawrence Erlbaum Associates, Inc., 1993. 322 pages, in print. ISBN 0805813896.

    The book is available with or without the CD of psychoacoustic examples, which is also available separately. Most of these examples are taken from *Auditory Demonstrations*, by Houtsma, Rossing and Wagenaars, see Appendix R.

80. Charles Madden, *Fractals in music—Introductory mathematics for musical analysis*, High Art Press, 1999. ISBN 0967172756.

    This book has a promising title, but both the mathematics and the musical examples could do with some improvement. There is certainly an interesting area here to be investigated, and maybe the real point of the book will be to make us more aware of the possibilities.

81. Max V. Mathews, *The technology of computer music*, MIT Press, 1969. 188 pages, out of print. ISBN 0262130505.

    This book appeared early in the game, and was at one stage a standard reference. Although much of the material is now outdated, it is still worth looking at for its description of the Music V computer music language, one of the antecedants of CSound.

82. Max V. Mathews and John R. Pierce, *Current directions in computer music research*, MIT Press, 1989. Reprinted 1991. 432 pages, in print. ISBN 0262132419.

    A nice collection of articles on computer music, including an article by Pierce describing the Bohlen–Pierce scale. There is a companion CD, see Appendix R.

83. W. A. Mathieu, *Harmonic experience*, Inner Traditions International, Rochester, Vermont, 1997. 563 pages, large format, in print. ISBN 0892815604.

    You would not guess it from the title, but this book is about the conceptual transition from just intonation to equal temperament, and the parallel development of harmonic vocabulary. The writing is down to earth and easy to understand.

84. Guerino Mazzola, *Gruppen und Kategorien in der Musik*, Heldermann-Verlag, Berlin, 1985. 205 pages, out of print. ISBN 3885382105.

    "Groups and categories in music." The next item, by the same author, is much easier to get hold of.

85. ———, *Geometrie der Töne: Elemente der Mathematischen Musiktheorie*, Birkhäuser, 1990. 364 pages, in print. ISBN 3764323531.

    "Geometry of tones: elements of mathematical music theory." This is a book in German about music and mathematics, almost completely disjoint in content from these course notes. The author was a graduate student under the direction of the mathematician Peter Gabriel in Zürich, and the influence is clear. I was rather surprised, for example, to see the appearance of Yoneda's lemma from category theory. This book can be ordered from www.amazon.de.

86. Guerino Mazzola, with contributions by Stefan Göller, Stefan Müller, and Karin Ireland, *The topos of music: geometric logic of concepts, theory, and performance*, Birkhäuser, Basel, 2002. 1368 pages, in print. ISBN 0817657312.

    This huge book is a much expanded English version of [**85**]. Even in English, I still find most of the contents of this book incomprehensible and unenlightening. You can download the preface and table of contents for free as a 2.8MB pdf file from www.encyclospace.org/tom/tom_preface_toc.pdf.

87. Ernest G. McClain, *The myth of invariance: The origin of the gods, mathematics and music from the Ṛg Veda to Plato*, Nicolas-Hays, Inc., York Beach, Maine, 1976. Paperback edition, 1984. 216 pages, in print. ISBN 0892540125.

    A strange mixture of mysticism and theory of scales and temperaments. If you take this book too seriously, you will go completely insane.

88. Brian C. J. Moore, *Psychology of hearing*, Academic Press, 1997. ISBN 0125056273.

    A standard work on psychoacoustics. Highly recommended.

89. F. Richard Moore, *Elements of computer music*, Prentice Hall, 1990. 560 pages, out of print. ISBN 0132525526.

    A very readable work by an expert in the field. The book is written in terms of the computer music language CMusic, which was a precursor of CSound.

90. Joseph Morgan, *The physical basis of musical sounds*, Robert E. Krieger Publishing Company, Huntington, New York, 1980. 145 pages, in print. ISBN 0882756567.

91. Philip M. Morse and K. Uno Ingard, *Theoretical acoustics*, McGraw Hill, 1968. Reprinted with corrections by Princeton University Press, 1986, ISBN 0691084254 (hbk), 0691024014 (pbk).

    This book is the best textbook on acoustics that I have found, for an audience with a good mathematical background.

92. Bernard Mulgrew, Peter Grant, and John Thompson, *Digital signal processing*, Macmillan Press, 1999. 356 pages, in print. ISBN 0333745310.

    A number of books have recently appeared on the subject of digital signal processing. This is a good readable one.

93. Cornelius Johannes Nederveen, *Acoustical aspects of woodwind instruments*, Northern Illinois Press, 1998. ISBN 0875805779.

94. Erich Neuwirth, *Musical temperaments*, Springer-Verlag, Berlin/New York, 1997. 70 pages, in print. ISBN 3211830405.

    This very slim, overpriced volume explains the basics of scales and temperaments. It comes with a CD-ROM full of examples to go with the text.

95. Harry F. Olson, *Musical engineering*, McGraw Hill, 1952. Revised and enlarged version, Dover, 1967, with new title: *Music, physics and engineering*. ISBN 0486217698.

    This work was a classic in its time, although it is now somewhat outdated.

96. Jack Orbach, *Sound and music*, University Press of America, 1999. 409 pages, in print. ISBN 0761813764.

97. Charles A. Padgham, *The well-tempered organ*, Positif Press, Oxford, 1986. ISBN 0906894131.

    This book is hard to get hold of, but has a wealth of information about the usage of temperaments in organs.

98. Harry Partch, *Genesis of a music*, Second edition, enlarged. Da Capo Press, New York, 1974 (hbk), 1979 (pbk). 518 pages, in print. ISBN 030680106X.

    Harry Partch is one of the twentieth century's most innovative experimental composers. This well written book explains the origins of his 43 tone scale, and its applications in his compositions, and puts it into historical context with some unusual insights. The book also contains descriptions and photos of many musical instruments invented and constructed by Partch using this scale.

99. George Perle, *Twelve-tone tonality*, University of California Press, 1977. Second edition, 1996. 256 pages, in print. ISBN 0520033876.

100. Hermann Pfrogner, *Lebendige Tonwelt*, Langen Müller, 1976. 680 pages, out of print. ISBN 3784415776.

"Living world of tone." This German book contains a discussion of musical scales in India, China, Greece and Arabia, followed by a discussion of the development of western tonality, and then a third section on the music of Arnold Schönberg.

101. Dave Phillips, *Linux music and sound*, Linux Journal Press, 2000. 408 pages, in print. ISBN 1886411344

This book describes a number of different music and sound programs for the Linux operating system. It comes with a CD-ROM containing the software described in the text, to the extent that it is freely distributable. A book like this quickly becomes out of date, but is nonetheless a useful guide to what is avaiable to the Linux user.

102. James O. Pickles, *An introduction to the physiology of hearing*, Academic Press, London/San Diego, second edition, 1988. Out of print. ISBN 0125547544 (pbk).

103. John Robinson Pierce, *The science of musical sound*, Scientific American Books, 1983; 2nd ed., W. H. Freeman & Co, 1992. 270 pages, in print. ISBN 0716760053.

A classic by an expert in the field. Well worth reading. The second edition has been updated and expanded.

104. Ken C. Pohlmann, *Principals of digital audio*, McGraw-Hill, fourth edition, 2000. 736 pages, in print. ISBN 0071348190.

This is a standard work on digital audio. The fourth edition has been brought completely up to date, with sections on the newest technologies.

105. Giovanni De Poli, Aldo Piccialli, and Curtis Roads (eds.), *Representations of musical signals*, MIT Press, 1991. 494 pages, in print. ISBN 0262041138.

A collection of fourteen essays by various experts in the field. Topics include granular synthesis, wavelets, physical modeling, user interfaces, artificial intelligence and adaptive neural networks.

106. Stephen Travis Pope (ed.), *The well-tempered object: Musical applications of object-oriented software technology*, MIT Press, 1991. 203 pages, in print. ISBN 0262161265.

An edited collection of articles from the Computer Music Journal on applications of object oriented programming to music technology.

107. Daniel R. Raichel, *The science and applications of acoustics*, Amer. Inst. of Physics, 2000. 598 pages, in print. ISBN 0387989072.

A general interdisciplinary textbook on modern acoustics, containing a discussion of musical instruments, as well as music and voice synthesis, and psychoacoustics.

108. Jean-Philippe Rameau, *Traité de l'harmonie*, Ballard, Paris, 1722. Reprinted as "Treatise on Harmony" in English translation by Dover, 1971. 444 pages, in print. ISBN 0486224619.

109. J. W. S. Rayleigh, *The theory of sound (2 vols)*, Second edition, Macmillan, 1896. Dover, 1945. 480/504 pages, in print. ISBN 0486602923/0486602931.

This book revolutionized the field when it came out. It is now mostly of historical interest, because the subject has advanced a great deal during the twentieth century.

110. Joan Reinthaler, *Mathematics and music: some intersections*, Mu Alpha Theta, 1990. 47 pages, out of print. ISBN 0940790084.

This slim volume examines various topics such as the Pythagorean scale, equal temperament, the shape of the grand piano, change ringing and symmetry in music.

111. Geza Révész, *Einführung in die Musikpsychologie*, Amsterdam, 1946. Translated by G. I. C. de Courcy as *Introduction to the psychology of music*, University of Oklahoma Press, 1954, and reprinted by Dover, 2001. 265 pages, in print. ISBN 048641678X.

This book contains an interesting discussion (pages 160–167) of the question of whether mathematicians are more musically gifted than exponents of other special branches and professions. The author gives evidence for a negative answer to this question, in sharp contrast with widely held views on the subject.

112. John S. Rigden, *Physics and the sound of music*, Wiley & Sons, 1977. 286 pages. ISBN 0471024333. Second edition, 1985. 368 pages, in print. ISBN 0471874124.

113. Curtis Roads, *The computer music tutorial*, MIT Press, 1996. 1234 pages, large format, in print. ISBN 0262181584 (hbk), 0262680823 (pbk).

 This is a huge work by a renowned expert. It contains an excellent section on various methods of synthesis, but surprisingly, doesn't go far enough with technical aspects of the subject.

114. ———, *Microsound*, MIT Press, 2001. 392 pages, in print. ISBN 0262182157.

 This book discusses sound particles and granular synthesis, and comes with a CD full of examples.

115. Curtis Roads, Stephen Travis Pope, Aldo Piccialli, and Giovanni De Poli (eds.), *Musical signal processing*, Swets & Zeitlinger Publishers, 1997. 477 pages, in print. ISBN 9026514824 (hbk), 9026514832 (pbk).

 A collection of articles by various authors, in four sections: I, Foundations of musical signal processing. II, Innovations in musical signal processing. III, Musical signal macrostructures. IV, Composition and musical signal processing.

116. Curtis Roads and John Strawn (eds.), *Foundations of computer music. Selected readings from Computer Music Journal*, MIT Press, 1985. ISBN 0262181142 (hbk), 0262680513 (pbk).

117. Curtis Roads, *The music machine. Selected readings from Computer Music Journal*, MIT Press, 1989. 725 pages. ISBN 0262680785.

118. Juan G. Roederer, *The physics and psychophysics of music*, Springer-Verlag, Berlin/New York, 1995. 219 pages, in print. ISBN 3540943668.

119. Thomas D. Rossing (ed.), *Acoustics of bells*, Van Nostrand Reinhold, 1984. Out of print. ISBN 0442278179.

120. ———, *Musical acoustics, selected reprints*, American Association of Physics Teachers, 1988. 227 pages, in print. ISBN 091785330.

121. ———, *The science of sound*, Addison-Wesley, Reading, Mass., Second edition, 1990. 686 pages, in print. ISBN 0201157276.

 A very nicely written book by an expert in the field, explaining sound, hearing, musical instruments, acoustics, and electronic music. Highly recommended. I understand that a new edition has just come out (November 2001), ISBN 0805385657.

122. ———, *Science of percussion instruments*, World Scientific, 2000. 208 pages, in print. ISBN 9810241585 (Hbk), 9810241593 (Pbk).

123. Thomas D. Rossing and Neville H. Fletcher (contributor), *Principles of vibration and sound*, Springer-Verlag, Berlin/New York, 1995. 247 pages, in print. ISBN 0387943048.

124. Joseph Rothstein, *MIDI, A comprehensive introduction*, Oxford Univ. Press, 1992. 226 pages, in print. ISBN 0198162936.

 Rothstein is one of the editors of the Computer Music Journal.

125. Heiner Ruland, *Expanding tonal awareness*, Rudolf Steiner Press, London, 1992. 187 pages, out of print. ISBN 1855841703.

 A somewhat ideosynchratic account of the history of scales and temperaments.

126. Joseph Schillinger, *The Schillinger system of musical composition, two volumes*, Carl Fischer, Inc, 1941. Reprinted by Da Capo Press, 1978. 878/? pages, out of print. ISBN 0306775522 and 0306775220.

Schillinger uses many mathematical concepts in describing his theory of musical composition.

127. Albrecht Schneider, *Tonhöhe, Skala, Klang: Akustiche, tonometrische und psychoakustische Studien auf vergleichender Grundlage*, Verlag für systematische Musikwissenschaft, Bonn, 1997. 597 pages, in print. ISBN 3922626890.

 "Pitch, scale, timbre: Acoustic, tonometric and psychoacoustic studies on comparative foundations." This German book can be ordered from www.amazon.de.

128. Günter Schnitzler, *Musik und Zahl*, Verlag für systematische Musikwissenschaft, 1976. 297 pages, out of print.

 "Music and number." A collection of essays on music and mathematics, in German, by a number of different authors.

129. William A. Sethares, *Tuning, timbre, spectrum, scale*, Springer-Verlag, Berlin/New York, 1998. 345 pages, in print. ISBN 354076173X.

 The basic thesis of this book is the idea, first put forward by John Pierce, that the harmonic spectrum or timbre of an instrument determines the most appropriate scales and temperaments to be used, and therefore we could start with a prescribed scale and design a harmonic spectrum for which it would be appropriate. The author also introduces *adaptive tunings*, where the pitches of notes are allowed to be modified by a few cents from one chord to the next, in order to keep the chords better in tune. The book comes with a CD of examples. Beware that in section 3.6, the syntonic comma and the Pythagorean comma are confused.

130. Ken Steiglitz, *A digital signal processing primer: with applications to digital audio and computer music*, Addison-Wesley, 1996. 314 pages, in print. ISBN 0805316841.

131. Reinhard Steinberg (ed.), *Music and the mind machine. The psychophysiology and psychopathology of the sense of music*, Springer-Verlag, Berlin/New York, 1995. Out of print. ISBN 3540585281.

132. Charles Taylor, *Exploring music: The science and technology of tones and tunes*, Institute of Physics Publishing, Bristol and Philadelphia, 1992. Reprinted 1994. 255 pages, in print. ISBN 0750302135.

133. Stan Tempelaars, *Signal processing, speech and music*, Swets & Zeitlinger Publishers, 1996. 360 pages, in print. ISBN 9026514816.

134. David Temperley, *The cognition of basic musical structures*, MIT Press, 2001. 360 pages, in print. ISBN 0262201348.

135. Martin Vogel, *Die Lehre von den Tonbeziehungen*, Verlag für systematische Musikwissenschaft, 1975. 480 pages, in print. ISBN 3922626092.

 "The study of tonal relationships." This German book is one of the standard works on scales and temperaments. It can be ordered from www.amazon.de.

136. ———, *Anleitung zur harmonischen Analyse und zu reiner Intonation*, Verlag für systematische Musikwissenschaft, 1984. 210 pages, in print. ISBN 3992626246.

 "Introduction to harmonic analysis and just intonation." This book can be ordered from www.amazon.de.

137. ———, *Die Naturseptime*, Verlag für systematische Musikwissenschaft, 1991. 507 pages, out of print. ISBN 3922626610.

 "The natural seventh." The title refers to the seventh harmonic 7:4, as a just musical interval.

138. G. N. Watson, *A treatise on the theory of Bessel functions*, Cambridge Univ. Press, 1922. Reprinted, 1996. 804 pages, in print. ISBN 0521483913.

 This encyclopedic tome contains more than you will ever want to know about Bessel functions. After all these years, it is still the standard work on the subject.

139. Scott R. Wilkinson, *Tuning in: Microtonality in electronic music*, Hal Leonard Books, Milwaukee, 1988. ISBN 0881886335.

This is a nice short popularized account of the theory of scales and temperaments, designed for synthesists. Beware, though, that the tables have some inaccuracies. This applies especially to the values in cents given in the tables for the Werckmeister III and Vallotti-Young temperaments.

140. Fritz Winckel, *Music, sound and sensation, a modern exposition*, Dover, 1967. 189 pages, in print. ISBN 0486217647.

141. Iannis Xenakis, *Formalized music: Thought and mathematics in composition*, Indiana University Press, Bloomington/London, 1971. ISBN 0253323789. Pendragon revised edition (four new chapters and a new appendix), Pendragon Press, New York, 1992. 387 pages, in print. ISBN 0945193246. Paperback edition, 2001, ISBN 1576470792.

Xenakis is one of the twentieth century's leading composers and theorists of music using aleatoric, or stochastic processes. This book is about the theory behind these processes. It is hard to read, and the mathematics is suspect in places. Nonetheless, the book is full of interesting ideas.

142. Joseph Yasser, *A theory of evolving tonality*, American Library of Musicology, Inc., 1932. 381 pages, out of print.

This book is a classic work on just intonation. Somebody should reprint it.

143. William A. Yost, *Fundamentals of hearing. An introduction*, Academic Press, San Diego, 1977. 326 pages, in print. ISBN 012772690X.

This is a nice, well written introduction to acoustics and psychoacoustics, written in textbook style, and very accessible.

144. Eberhard Zwicker and H. Fastl, *Psychoacoustics: facts and models*, Springer-Verlag, Berlin/New York, Second edition, 1999. 380 pages, in print. ISBN 3540650636.

Eberhard Zwicker was one of the great names in psychoacoustics. This book, written with his student Hugo Fastl, is an excellent introduction to psychoacoustics, and presents a modern account with a great deal of qualitative and quantitative information. The second edition has been updated by Fastl.

Mobile instrument, Arthur Frick

Index