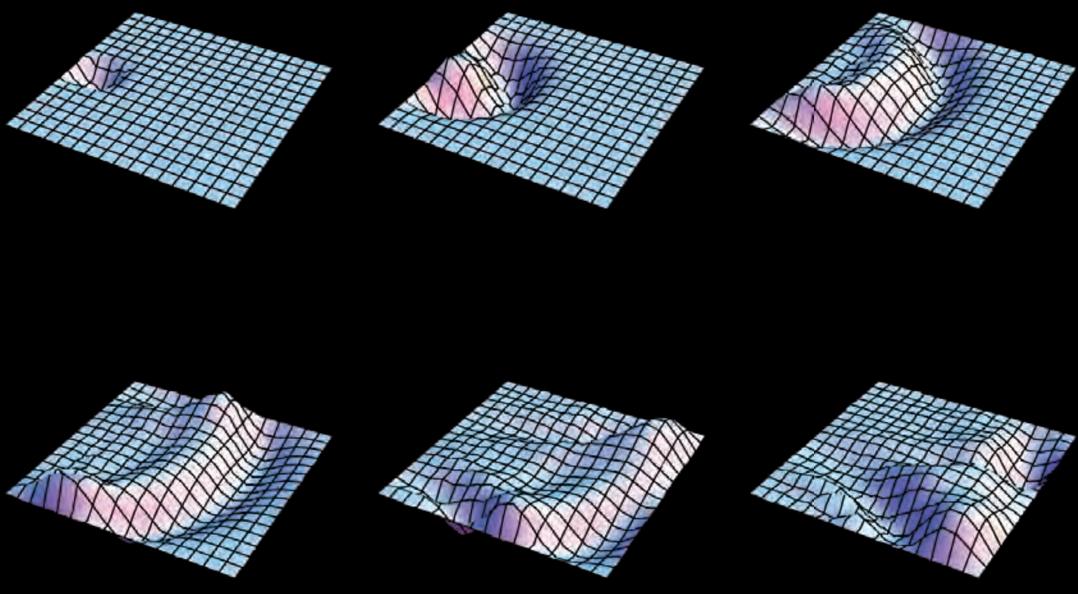
Signals, Systems and Sound Synthesis

Martin Neukom



Peter Lang

The subject of this book is the synthesis and treatment of sound by computer. Using illustrations, animations, sound examples and sample programs, it introduces the most important techniques of sound and score synthesis and explains the technical and mathematical principles necessary for understanding them. After reviewing fundamentals of acoustics, the author describes system and signal theory and introduces the programs and programming languages used in the book. The traditionally important synthesis techniques are described in detail, as are various nonlinear synthesis techniques and synthesis by physical modeling. The concluding chapters of the book deal with the projection of sound in space and with the use of algorithmic and stochastic procedures in computer music. The appendix contains a survey of basic mathematical principles, various tables for reference and a detailed index. The included CD contains the entire text of the book, as well as additional chapters and explanations, sound examples, animations illustrating dynamic processes and many sample computer programs.

Martin Neukom studied music theory at the Musikhochschule Zurich and musicology, mathematics and psychology at the University of Zurich, where he received a doctorate in musicology. He is a composer, professor of music theory at the Zurich University of the Arts and research associate at the Institute for Computer Music and Sound Technology ICST of the Zurich University of the Arts.

Signals, Systems and Sound Synthesis

Signals, Systems and Sound Synthesis

Martin Neukom

Translation from the German by
Gerald Bennett

ICST

Z
hdk
Zürcher Hochschule der Künste
Zurich University of the Arts

PETER LANG
Bern · Berlin · Bruxelles · Frankfurt am Main · New York · Oxford · Wien

Bibliographic information published by die Deutsche Nationalbibliothek
Die Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available on the Internet at ‹http://dnb.d-nb.de›.

British Library Cataloguing-in-Publication Data: A catalogue record for this book is available from The British Library, Great Britain

Library of Congress Cataloging-in-Publication Data

Neukom, Martin, author.
[Signale, Systeme und Klangsynthese. English]
Signals, systems and sound synthesis / Martin Neukom.
pages cm
Accompanied by CD.
Translation of: Neukom, Martin. Signale, Systeme und Klangsynthese. Bern : P. Lang, 2003.
Includes bibliographical references and index.
ISBN 978-3-0343-1428-2
1. Computer music–Instruction and study. 2. Computer sound processing.
3. Computer composition. I. Title.
MT723.N4813 2013
786.7–dc23

2013024500

Published with the support of the Swiss National Science Foundation.

Additional texts, emendations and programs for this book can be found at the following address: www.icst.net/Signals_Systems_and_Sound_Synthesis.

ISBN 978-3-0343-1428-2 pb. ISBN 978-3-0351-0609-1 eBook

Hochfeldstrasse 32, CH-3012 Bern, Switzerland
info@peterlang.com, www.peterlang.com

Printed in Hungary

Short Contents

Contents

Preface to the Original German Edition

The first digital sound synthesis dates back to 1957, when an IBM-709 Computer in New York calculated a 17-second long piece of music. The program that had produced these sounds was written by Max Mathews, who 12 years later would write the first book giving detailed information about the technology of digital sound synthesis. This book was of extreme importance. Interest in electroacoustic music, and particularly in music made by computer, had greatly increased since 1957, and Mathews' book contributed immeasurably to knowledge about digital sound synthesis and thus paved the way for further development in the field.

After 1969, great advances were made in the technology of sound synthesis and the digital treatment of sound: Frequency Modulation, Linear Predictive Synthesis, Granular Synthesis, Formant Synthesis, Synthesis in the Frequency Domain (for example the Phase Vocoder), to name only a few techniques. These developments took place outside the commercial world and were generally well documented in specialized journals and at conferences, so that the growing community of composers for whom digital sound synthesis had become important was able to remain abreast of the newest technologies.

The first international conference for computer music took place in 1974; the *Computer Music Journal* first appeared in 1977. By the middle of the 1980's there was enough known about digital sound synthesis that in 1985 Curtis Roads and John Strawn could publish *Foundations of Computer Music* [43], a collection of 36 technical articles, most of which had first appeared in the *Computer Music Journal*. In the same year, the first comprehensive book about techniques of computer music, *Computer Music: Synthesis, Composition, and Performance* by Charles Dodge and Thomas Jerse [2] was published. F. Richard Moore's book *Elements of Computer Music* [8] followed in 1990, Curtis Roads' 1234-page *Computer Music Tutorial* [11] in 1996, and a second, revised edition of Doge and Jerse's book was published in 1997. Each of these books added substantially to the body of acquired knowledge about the technology and aesthetics of computer music. Nowadays, courses in digital sound synthesis and the treatment of sound are offered in universities all over the world. These books, together with the *Computer Music Journal*, furnish the foundations for this teaching. Without the generosity of very many colleagues in the field who took the time to document their research and their achievements, and without the vision and magnanimity of the few companies that were willing to publish these writings, computer music as we know it today would surely not exist.

Signals, Systems and Sound Synthesis by Martin Neukom belongs in this group of definitive texts on computer music. Not only are all the important issues of sound synthesis discussed in the light of the most recent developments, but topics relevant to computer music but usually disregarded or mentioned only marginally in earlier books, like systems theory, room acoustics or programming for musicians, are treated at length here. The book is by no means simply a compendium of established techniques. Each theme is introduced and developed from a new point of view. The approach is often mathematical, and the resulting level of abstraction and change of perspective are refreshing and inspiring, even for mathematically less proficient readers. A CD-ROM is included, containing the entire text, many graphic animations and a great many sound examples, making of the book both an up-to-date documentation of what has been achieved in the field of computer and a handbook of areas of work for the future.

The publication of *Signals*, *Systems and Sound Synthesis* would have been a significant event in any country and language. It is our great good fortune that the book was published in German and thus became the first German-language textbook on digital sound synthesis. When I asked Martin Neukom in 1998 to write a short guide to sound synthesis so that we would not always have to direct our students to the English-language professional literature, I had no idea that I would receive such a rich, original and comprehensive text. The Zurich University of the Arts gave the project both moral and financial support, for which we are very grateful. With Signals, Systems and Sound Synthesis, Martin Neukom has given the computer music community a marvelous textbook. I am proud that the Zurich University of the Arts contributed to its realization.

Gerald Bennett

Preface to the English Edition

There is little to add to what I wrote nearly 10 years ago. In its German version, *Signals, Systems and Sound Synthesis* has become a standard work, the standard work in the field of computer music for German-language readers. The vision of a documentation which is both a conventional, albeit rather heavy, book and a complex interactive tool for learning has proven extremely fertile for teaching, and the book has been used for several years with success in courses at the Zurich University of the Arts. *Signals, Systems and Sound Synthesis* remains one of the few books to show in detail how to realize the techniques of contemporary computer music.

The time seemed ripe to extend the book's readership by issuing an English-language version. Martin Neukom has rewritten extensive sections of the book, bringing up to date both content and references. The enclosed CD-ROM contains much new additional material. Particularly the number of interactive programming examples has been greatly increased.

If I wrote the Foreword to the German Edition as supporter and defender of Martin Neukom's project, I now write as translator. Of course, the book's sheer bulk was a challenge, but not surprisingly I encountered the greatest difficulties with topics I knew little about. Despite 40 years' experience as a composer and researcher in computer music, I often found myself on paths which were new to me. I have tried to find factually and grammatically correct, felicitous English equivalents for the German text. I hope the reader will forgive me if she occasionally misses the specific jargon of a domain in which I sojourned for the first time.

Gerald Bennett

Acknowledgments

This book was commissioned by the Musikhochschule of Zurich, today the Department of Music of the Zurich University of the Arts (ZHdK) and was accepted by the University of Zurich as a doctoral dissertation in Musicology. The English translation was commissioned by the Institute for Computer Music and Sound Technology (ICST) of the Zurich University of the Arts. I wish to thank these institutions, and particularly Germán Toro-Pérez, Director of the ICST, for their trust and belief in the project and for their generous support. I am grateful to my colleagues at the ICST as well as to the students of the ZHdK for many interesting discussions about (and not about) electroacoustic music. Special thanks are due to Daniel Muzzulini, the editor of the English translation. The book has greatly profited from his thorough and meticulous reading and from his suggestions for the book's contents.

My particular thanks are reserved for Gerald Bennett, composer, teacher, founder of the ICST, and much more. He suggested that I write this book, and he accompanied its growth during the ten years from its first beginnings to the present final version. It has been great good fortune that Gerald Bennett agreed to translate the book into English, for he combines to an unusual degree professional knowledge, linguistic competence and generosity towards the ideas of others.

Martin Neukom

1 How to Use This Book

1.1 Getting Started

This book is a translation from the German original, *Systeme, Signale und Klangsynthese* (2003, 2005). The text for the English edition has been considerably revised, and whole sections have been rewritten. It has some specific features of format and layout that distinguish it from most other books, largely due to the fact that it is written using the computational software *Mathematica*. While graphically not as flexible as digital publishing software, Mathematica allows the text to be published simultaneously as a book and as a CD containing the complete text in computer-readable form, many directly executable programs as well as links to additional relevant passages within the text, example programs, animations and sound examples.

The book is stored on the CD in so-called *Notebooks* (see 1.3 below). It was decided that the printed text and the screen version of the Mathematica notebooks should look as much alike as possible. Because notebooks have no distinct pages, and hence no page numbers, reference within the book is not made to pages but to chapters and sub-chapters. The illustrations always follow immediately their explanations in the text and therefore have no captions. The mathematical formulas are not numbered as usual but instead are linked or repeated where necessary.

In the German original, Chapter 1 (this chapter) and the introductions to the programming languages Csound and C were quite long. In this edition these chapters have been drastically shortened since both programming languages are well documented in English. In Chapter 4 one will find short descriptions of the programming languages one will encounter in the text as well as explanations of some special applications and procedures used in the book.

The bibliography of the original edition contained a commented list of German and English books suitable for interested non-professional readers. In this edition, the books listed have been updated to their most recent editions, and most German references have been replaced by their English versions where these exist or by comparable English language literature. Most of the methods, formulas and algorithms used here belong to the common practice of their respective disciplines, and their origins cannot be determined with certainty. For this reason, no texts are quoted in the book and in general no mention is made of authors of formulas or techniques. In addition, reference is usually made to secondary literature and not to primary literature.

The CD contains:

– The complete text of the book as Mathematica-notebooks (*name*.nb) in the folder *Text*;
– Additional chapters, explanations and examples as hidden cells in the notebooks;
– C/C++ programs in source code (*name*.cpp) in the folder *CPP*;
– Java programs in source code (*name*.java) and Java classes (*name*.class) for the mxj-externals in the folder *Java*;
– Sound examples (*name*.wav) in the folder *Sounds*;
– Csound examples in the folder *Csound*;
– Max patches in the folder *Max*;
– Programs in the Processing language in the folder *Processing*.

1.2 Overview

This book gives an introduction to the techniques of digital sound synthesis and sound transformation. The relevant topics are presented using illustrations, animations of complex physical and mathematical relationships, sound examples and sample programs. Basic technical and mathematical principles will be explained where they are necessary for reading the specialized literature.
This first chapter gives an overview of the book's contents, the enclosed CD and the computer programs used in the book. In Chapters 2 to 4 the physical, mathematical and programming essentials for the rest of the book are developed. The short overview of acoustics in Chapter 2 is followed in Chapter 3 by a thorough presentation of signal theory and system theory. These theories provide the tools for a precise derivation of many techniques of sound synthesis, sound transformation and control theory. The text is written to give a clear and intuitive view of the material rather than a more abstract, general presentation. The chapters that follow are written so they can be understood in their broad lines without the theoretical material developed in Chapter 3. Traditionally, many programming languages and specific programs have been used in the treatment and synthesis of sound. In Chapter 4, only those languages are discussed which are later used in the book. Mathematica was chosen for the body of the text, because an editor, a high-level programming language and routines to generate illustrations, animations and sounds are integrated into the language. C/C++ was chosen as a general programming language because Csound, Max and many other programs are written in C and can be extended using additional routines written in C. Csound was chosen as the domain-specific language (DSL) for sound synthesis because it has a long history in computer music and because it has the flexibility of a general programming language. Max was chosen for interactive programming because it is the most frequently used language in live electronics. Max is intuitive to use and ideal for demonstrations because of its graphic interface. Several techniques of sound synthesis and sound treatment are introduced in Chapters 5 to 8. Nonlinear techniques and techniques that simulate the physical procedures of natural sound production are particularly emphasized because they give interesting results with only simple programming and modest computation times. Chapter 8 gives a comprehensive introduction to sound synthesis by physical modeling. At the same time it offers a first taste of digital filter theory. Chapter 9 discusses some of the many problems that arise in connection with the synthesis of sound in resonant spaces and presents suggestions for solutions. Chapter 10, finally, discusses techniques and aids for the composition of computer music.

1.3 Instructions for Using Specific Programs

The rest of this chapter contains instructions for using the data and the programs on the enclosed CD. Chapter 4 contains more detailed descriptions of the programs and programming languages Mathematica, C/C++, Csound, Max and Processing.

The text of this book is stored in Mathematica notebooks, which can be read with the *Wolfram CDF Player* or read and edited with the program *Mathematica*. The program Wolfram CDF Player can be downloaded from Wolfram Research (http://www.wolfram.com). The notebooks include closed *cells* containing additional text, the programs used to generate the illustrations, animations (Ø Animation) and sound examples (Ø Sound example) as well as additional cells. Longer sound examples are stored on the CD as WAVE files.

The *Csound program* (Version 5 or later) together with *Csound scores* and *orchestras* generates sounds. All the sounds synthesized by the Csound program in the course of the book can be found on the enclosed CD.

All the *C/C++ programs* can be compiled on a C++ compiler, but only some of the examples can be compiled on a classical C compiler.

The *Max programs* can be run using the freely available program *Max Runtime*. They can be executed, edited and extended using the Max program itself (http://cycling74.com).

The examples written in the language *Processing* can be executed, edited and extended using the freely available program Processing (http://processing.org).

1.3.1 Using the Mathematica Notebooks

After downloading and installing the Wolfram CDF Player, start the program. Now a notebook can be opened using the Menu item *File*. If the notebook *Contents.nb* or *Index.nb* is opened first, the other notebooks can be opened using hyperlinks (see below). A Mathematica notebook consists of cells that contain text, illustrations, sounds or other cells. The cells are indicated by brackets on the right edge of the screen. Nested cells are shown by nested brackets. By double-clicking on an outer bracket, whole groups of cells can be closed so that only the uppermost cell is visible. By double-clicking on an inner cell bracket, a group of cells can be closed so that only the selected cell is visible. Closed cells can be opend by double-clicking on a bracket with a hook. The text appears as in the printed book. The hidden cells contain the programs used to generate illustrations, animations and sounds as well as calculations and additional texts.

To start an animation, open the the group of cells whose visible cell has the comment (Ø Animation) by double-clicking on the corresponding bracket. Most animations are interactive. Their parameters can be changed by graphic elements in and sliders next to the illustrations. The sliders can be animated by clicking on the plus sign next to the slider. In this way continuous animations can be generated.

The CD version of the book contains hyperlinks which give access to other cells in the current notebook or to cells in other notebooks. These hyperlinks are colored blue. Clicking on a hyperlink retrieves the corresponding cell.

1.3.2 Using the Csound Programs

Csound is a command-line oriented programming language designed for synthesizing and manipulating sound. Its name comes from the fact that it is written in C. For the examples in this book *QuteCsound* was used as a development environment within which programs can be edited and executed. QuteCsound has a highlighting editor with autocompletion, interactive features and integrated help files. In the Csound folder the orchestra and score programs have been combined to form files in so-called *Csound unified file format* (*name.csd*).

1.3.3 Using the C/C++ Programs

C++ is an intermediate-level general-purpose programming language. It contains nearly all the features of the language C and adds object-oriented features, in particular, classes. This book does not always distinguish explicitly between programs in C and programs in C++. All the programs can be compiled with a C++ compiler, but only some can be compiled with a C compiler. In order to execute the programs that write sound files, the header file my_WAVE.h must be included in the program text using the command: #include “my_WAVE.h”.

1.3.4 Using the Max Patches

The following explanations can be tested interactively by double-clicking on the Max patch *Instructions.maxpat* in the folder Max. Max patches are made up of elements such as generators, faders, inputs and outputs, etc. They can also contain subpatches. The audio settings can be changed by clicking on the word “Audio”, which opens the window “DSP Status”. The settings used for the examples in the book are stored under “Presets”. Parameter values can be changed in the corresponding number fields and graphic elements and can then be stored as new presets. Clicking on the subpatch *p name* opens the subpatch and shows it in a new window. The subpatch *p comments* tells what the respective Max patch does and how to use it. The subpatch *p presets* explains the preset settings. Max objects can be programmed in C or Java as so-called *externals*. In this book only Java externals are used. Double-clicking on the object *mxj quickie name* displays the source code of the external *name*.

1.3.5 Using the Processing Programs

The examples written in the language Processing are stored as source code on the enclosed CD and must be executed using the freely available program Processing . Some examples require libraries that are not automatically installed with Processing (controlP5, oscP5, netP5) (http://processing.org). Instructions for and comments on the Processing examples can be found in the headers of the respective source code files.

2 Fundamentals of Acoustics

Further Reading: Musical Acoustics *by Donald E. Hall [20] gives a comprehensive introduction to the general topic.* Music Cognition and Computerized Sound *by Perry R. Cook [41] is especially useful because of the enclosed CD with sound examples. The fundamentals of acoustic are summarized in several books on computer music, among them* Computer Music *by Charles Dodge and Thomas A. Jerse [2].* Genealogie der Klangfarbe *by Daniel Muzzulini [77] presents a detailed discussion of the phenomenon of timbre.*

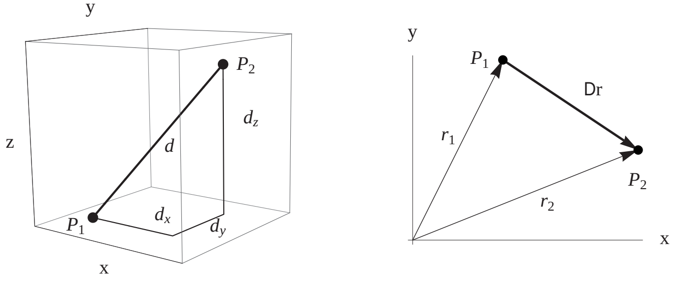
2.1 Basic Physical Principles and Units

2.1.1 Path, Velocity, Acceleration

The distance between the points P_1 and P_2 is $|x_2 - x_1| = |x_1 - x_2|$. In Cartesian coordinates (straight perpendicular axes with the same unit of measurement) in two or three dimensions, the Euclidean distance d is calculated from the difference of the coordinates by

$$d = \sqrt{d_x^{\,2} + d_y^{\,2}} \text{ or } d = \sqrt{d_x^{\,2} + d_y^{\,2} + d_z^{\,2}}$$

respectively (left figure below). If the points are defined by position vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ from the origin, then the *shortest path* $\Delta\boldsymbol{r} = \boldsymbol{r}_2 - \boldsymbol{r}_1$. The formulas above give the corresponding distance (right figure below).



If it takes time Δt for a point to traverse the path $\Delta\boldsymbol{r}$, then the point's mean *velocity* is $\bar{\boldsymbol{v}} = \Delta\boldsymbol{r}/\Delta t$. The direction of $\Delta\boldsymbol{r}$ and $\bar{\boldsymbol{v}}$ are the same. The instantaneous velocity $\boldsymbol{v}$ is equal to the limit of $\Delta\boldsymbol{r}/\Delta t$ when Δt approaches 0, that is, equal to the first derivative of the position with respect to time: $\boldsymbol{v} = \lim_{\Delta t \to 0}(\Delta\boldsymbol{r}/\Delta t) = d\boldsymbol{r}/dt$. In contrast to the vector velocity $\boldsymbol{v}$, we refer to the magnitude of the velocity as speed and write v or u. *Acceleration* is defined as the change in velocity over time: $\boldsymbol{a} = \Delta\boldsymbol{v}/\Delta t$. The instantaneous acceleration is defined as: $\boldsymbol{a} = \lim_{\Delta t \to 0}(\Delta\boldsymbol{v}/\Delta t) = d\boldsymbol{v}/dt$. Correspondingly, velocity is calculated by integrating acceleration over time, position by integrating velocity over time (A5.3).

In the following example a mass is dropped from a height of 100 m. We know the position $x = 100$ m at time $t = 0$ s, the speed $v = 0$ m/s at time $t = 0$ s and the acceleration given by gravitation g. If the x-axis is pointing upward (figure to the right), the acceleration is negative: $g \approx -10$ m/s^2. Since the acceleration is constant, we calculate from $a = dv/dt$ the speed $v = tg$. For the speed after 3 seconds, we have

$$v(3) = -3\text{ s}\cdot 10\text{ m/s}^2 = -30\text{ m/s}.$$

In general the speed is

$$v(t) = v_0 + \int_{t=0}^{t} g(t)\,dt.$$

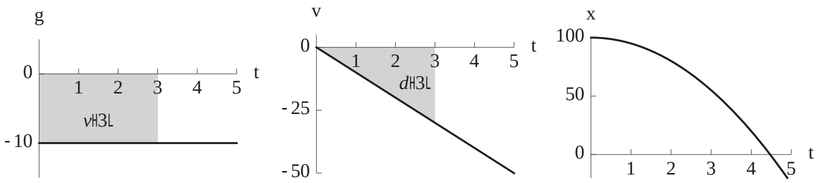
At time $t = 3$ s the speed is

$$v(3) = v_0 + \int_{t=0}^{3} g(t)\,dt = 0 + \int_{t=0}^{3} -10\ \mathrm{m/s^2}\,dt = -30\ \mathrm{m/s}, \text{ as above.}$$

We calculate the distance covered at time t by

$$d = \int_{t=0}^{t} v(t)\,dt = \int_{t=0}^{t} t(-10\mathrm{m/s^2})\,dt = -10t^2/2$$

and the position x at time t by $x(t) = x_0 + d = 100 - 10t^2/2$. So after 3 seconds the mass is still $(100 - 10\cdot 9/2)$ m = 55 m above the ground.

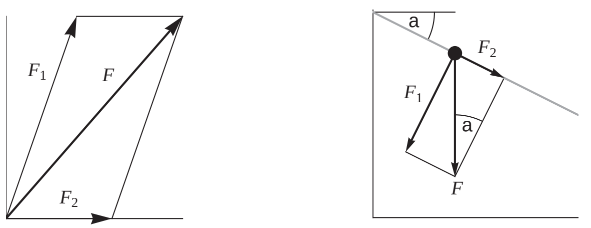


2.1.2 Mass and Force

Mass is a basic characteristic of matter. It is measured in kilograms (kg). Mass is different from weight, which arises from the action of a gravitational field on a body. The mass per volume unit of a body or material is called its density $\rho = m/V$ and is measured in kg/m^3 or in kg/Liter = kg/dm^3.

When a *force* acts upon a body, that body is either accelerated or deformed, or stress or pressure arises. Forces are defined by their magnitudes and directions; for this reason they are represented as vectors. If the forces $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ act simultaneously at the same point as in the illustration on the left below, the sum of the forces corresponds to the vector $\boldsymbol{F}$.

The accelerating force acts in the direction in which a body can move. If a mass is on an inclined plane, as in the illustration to the right, the force of gravitation $\boldsymbol{F}$ pulls the mass straight down, but acceleration can only take place in the direction of the plane's inclination, shown here by the vector $\boldsymbol{F}_2$.



Force is measured in Newtons. One Newton is equal to the force necessary to accelerate a mass of one kilogram by one meter per second within a second:

$$1\ \mathrm{N} = 1\ \mathrm{m\cdot kg/s^2}.$$

Newton's Second Law of motion states that force is proportional to a body's mass and to the acceleration acting upon that body.

$$\boxed{F = m\,a}$$

Example 1. The force of gravity at the earth's surface is equal to $\boldsymbol{F} = m\boldsymbol{g}$, where $\boldsymbol{g}$ is the acceleration due to gravity. Let the angle of inclination in the illustration to the right above be α, then the accelerating force $\boldsymbol{F}_2 = \boldsymbol{F}\cdot\sin(\alpha) = mg\cdot\sin(\alpha)$. For a mass of one kilogram, an angle of 30° and $g = 9.807\ \mathrm{m/s^2}$, that means:

$$\boldsymbol{F}_2 = 1\ \mathrm{kg}\cdot 9.807\ \mathrm{m/s^2}\cdot\sin(30°) = 4.9035\ \mathrm{N}.$$

Example 2. The force with which a stretched ideal spring pulls back is proportional to the elongation of the spring x, that is, $F = -Kx$. The minus sign indicates that the force F acts in the direction opposite that of the elongation. The constant K depends on the properties of the spring, such as elasticity, mass, etc., and is called the spring or force constant.

2.1.3 Momentum, Work, Power, Energy

The *momentum* $\boldsymbol{p}$ of an object is defined as the product of its mass and its velocity.

$$\boxed{\boldsymbol{p} = mv}$$

If no external force acts on an object, its momentum is conserved. If an external force $\boldsymbol{F}$ acts on an object, the object's momentum is changed according to:

$$\boxed{d\boldsymbol{p}/dt = \boldsymbol{F}}$$

The mechanical *work* W is the amount of energy transferred when a force $\boldsymbol{F}$ moves an object through the distance $\boldsymbol{s}$ and is given by:

$$\boxed{W = \int_a^b \boldsymbol{F}\cdot d\boldsymbol{s}}$$

Here $\boldsymbol{F}$ and $\boldsymbol{s}$ are vectors, whereas W is a scalar quantity, that is, a quantity without direction. If $\boldsymbol{F}$ and $\boldsymbol{s}$ are constant (with magnitude F and s) and include the angle β, then the work W is given by:

$$\boxed{W = Fs\cdot\cos(\beta)}$$

Work is measured in Joules: $1\ \mathrm{J} = 1\ \mathrm{m^2\cdot kg/s^2} = 1\ \mathrm{N\cdot m}$.

Power P is defined as work performed over time.

$$\boxed{P = W/t}$$

Power is measured in Watts: $1\ \mathrm{W} = 1\ \mathrm{m^2\cdot kg/s^3} = 1\ \mathrm{J/s}$. Specifications for electronic devices give some idea of how much a Watt is. Automobiles produce several tens of thousands of Watts of power (one horse-power is 735.3 W). The average overall performance of a human being in terms of power is about 100 Watts, whereas the sound power of a violin played fortissimo is only about 0.001 Watt.

Energy is a notion of great importance in physics. As work, energy is measured in Joules. In classical mechanics, the law of the conservation of energy states that the sum of all the various forms of energy in an isolated system remains constant. Forms of energy are mechanical energy, chemical energy and electrical energy as well as warmth.

Potential energy can be thought of as energy stored within a physical system. The potential energy E_P of an object results from the position of the object in space under the influence of certain forces. For objects in the earth's gravitational field the following relationship holds: $E_P = mgh$, where h is the object's height above a reference level. The potential energy E_P of a stretched spring is

$$\boxed{E_P = \frac{K}{2} x^2}$$

where x is the deviation of the spring from its initial position.

The *kinetic energy* E_k of an object of mass m and velocity v is:

$$\boxed{E_k = \frac{1}{2} mv^2}$$

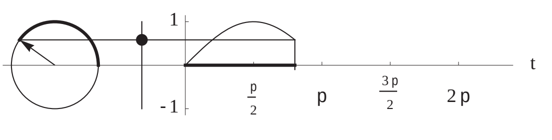
2.2 Vibration and Waves

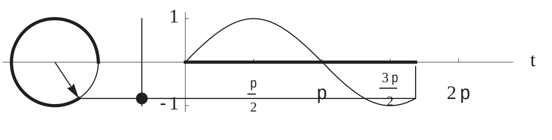
Further Reading: The book Waves *(Berkeley Physics Course, Vol. 3) by Frank S. Crawford [6] does not specifically treat musical acoustics, but it is nonetheless an excellent introduction to our subject. In addition, it describes numerous experiments that are easy to carry out.*

2.2.1 Harmonic Oscillation

2.2.1.1 Definition and Mathematical Representation

If the *excursion* of an oscillating point corresponds to a sine function in time, one speaks of *harmonic oscillation*. The figure below shows three snapshots taken from an animation and illustrates the relationship between the harmonic oscillation of a point, the rotation of a pointer on the left and a representation of the point's excursion as a function of time: $x = f(t) = \sin(t)$ on the right. (Ø Animation)

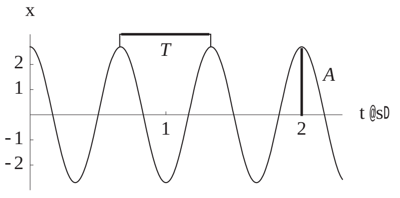




The duration of the cycle of an oscillation, that is the interval between two corresponding states of the oscillation, is its *period* T (in seconds), the number of cycles per second is its *frequency* f (in Hertz). The following relationships hold:

$$\boxed{f = \frac{1}{T} \text{ and } T = \frac{1}{f}}$$

In the illustration below, the period is 2/3 s, the frequency is therefore 3/2 Hz. The function $x = \sin(t)$ has the period 2π and the frequency $1/2\pi$. Hence, the so-called *angular frequency* of an oscillation is defined as $\omega = 2\pi f$, where f is the frequency. The maximum excursion from zero is the *amplitude*. The excursion is always a function of time; in the illustration below, the amplitude remains constant.



The instantaneous state of a vibration is described by the phase, or more precisely by the instantaneous phase. The *instantaneous phase* is defined either as the argument of a sine or cosine function (the *unwrapped phase*) or as this argument modulo 2π (the *wrapped phase*). Hence the phase is simply a number. In the representation of oscillation using a revolving vector, the phase corresponds to the angle α between the horizontal axis and the vector. So, for instance, the phase at a maximum point of a sine wave is 90° or $\pi/2$. Displacement along the time axis results from the zero phase angle or phase constant φ_0, that is, from the instantaneous phase at time $t = 0$. The oscillation above begins at a maximum point and thus has an initial phase $\varphi_0 = \pi/2$. A harmonic oscillation is determined by its amplitude A, its frequency f or angular frequency ω, and its initial phase φ_0. The instantaneous phase of the oscillation is $\omega t + \varphi_0$ or $(\omega t + \varphi_0)_{\mathrm{mod}\,2\pi}$.

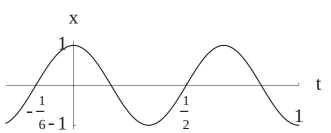
$$\boxed{x = A \cdot \sin(\omega t + \varphi_0)}$$

In the oscillation above these parameters are $A = 2.7$, $f = 3/2$ and $\varphi_0 = \pi/2$, which gives the equation: $x = A \cdot \sin(\omega t + \varphi_0) = 2.7 \cdot \sin(3\pi t + \pi/2)$.
From the time when the sine function is zero, or more precisely from a time t_0, at which the phase is zero, we obtain the zero phase angle as follows. By substituting $t - t_0$ for t the graph of a function is displaced along the positive x-axis by t_0, giving the function equation

$$x = A \cdot \sin(\omega(t - t_0)) = A \cdot \sin(\omega t - \omega t_0))$$

At time $t = 0$, therefore, we have for the phase $\varphi_0 = -\omega t_0$. In the figure below the frequency is 3/2, and we have for the phase constant

$\varphi_0 = -3/2 \cdot 2\pi(-1/6) = \pi/2$ or $\varphi_0 = -3/2 \cdot 2\pi \cdot 1/2 = -3/2 \cdot \pi \; (\approx \pi/2)$



The frequency of a periodic signal is defined as the number of oscillations per second. In general, *instantaneous angular frequency* $\omega(t)$ is defined as the time derivative of the *instantaneous phase* $\varphi(t)$ (2.2.3.5)

$$\omega(t) := \frac{d}{dt}\varphi(t).$$

Conversely, we obtain the instantaneous phase $\varphi(t)$ from the instantaneous frequency $f(t)$ by the relationship

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

Harmonic oscillation arises in systems in which reactive force is proportional to displacement and acts in opposition to it: $F = -Dx$. Newton's Second Law of movement (2.1.2.3) states that

$$F = ma = m\ddot{x}.$$

Hence, it follows that $\ddot{x} = -(D/m)x$. The solutions of these differential equations are harmonic oscillations. The total energy of an oscillating system consists of potential energy and kinetic energy. At the turning points of the oscillation, where the displacement is at its maximum and the velocity is zero, the potential energy

$$E_p = \frac{1}{2}Dx^2$$

is maximal (2.1.3.4) and the kinetic energy

$$E_k = \frac{1}{2}mv^2$$

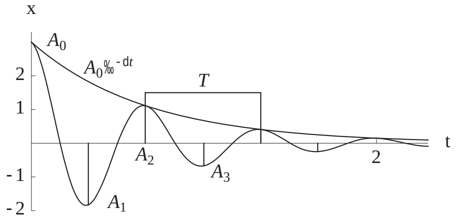
is zero. At zero crossings of the oscillation the displacement is zero and the velocity is at its maximum, hence the potential energy is zero and the kinetic energy is maximal.

Systems in which a reactive force F acts on a mass m proportionally to the displacement (i.e. $F = -Dx$) oscillate harmonically: $x = A\cdot\sin(\omega t + \varphi_0)$. The following relationships hold:

angular frequency $\omega = \sqrt{D/m}$
potential energy $E_p = \frac{1}{2} D x^2$
kinetic energy $E_k = \frac{1}{2} m v^2$

2.2.1.2 Damped Oscillation

Free vibration in mechanical systems is always damped due to friction. In the simplest case damping causes an exponential reduction of amplitude, which means that the amplitude diminishes by the same amount at each period of oscillation. The ratio $A_1/A_2 = A_2/A_3 = \ldots = c$ is called the *damping coefficient* and its natural logarithm $\Lambda = \ln(c)$ the *logarithmic decrement*.



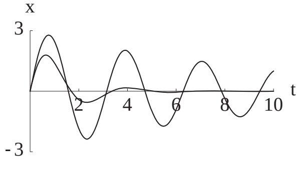
The damped oscillation can be described as the product of a sine or cosine function and a so-called *envelope*.

$$\boxed{x = A_0\, e^{-\delta t} \sin(\omega t + \varphi_0)}$$

The function of damped oscillation solves the equation of motion of a system in which a reactive force is proportional to displacement and additionally friction is present, proportionally to the instantaneous velocity of the oscillating object: $R = -rv = -r \cdot dx/dt$, where the factor r is known as the *coefficient of friction*. The equation of motion of the system is: $F = -Dx + R = -Dx - rv$. The frequency of the sinusoid is constant but depends upon the friction. If the mass of the oscillating body is given by m and $\delta = r/(2m)$, then

Frequency of the damped oscillation $f = \frac{1}{2\pi}\sqrt{(D/m - r^2/(4m^2))}$
Logarithmic decrement $\Lambda = \delta T = 2\pi\delta/\omega$
Energy $E = \frac{1}{2} D x^2 = \frac{1}{2} D A_0^2 e^{-\delta t}$

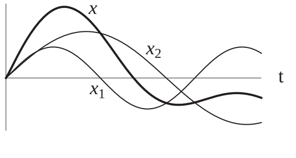
The following figure shows two oscillations ($A_0 = 3$, $D = 2$, $m = 0.5$) with different damping coefficients ($r_1 = 0.1$, $r_2 = 0.7$). From the formulas above it follows that the frequencies are $f_1 = 0.3179$ and $f_2 = 0.2982$, and the corresponding periods are $T_1 = 3.146$ and $T_2 = 3.354$. After 10 time units the amplitudes are 1.104 and 0.00274.



2.2.1.3 The Addition of Harmonic Oscillations

When two or more oscillations are superposed, their displacements are added together:

$$x = x_1 + x_2 = A_1 \sin(\omega_1 t + \varphi_{0,1}) + A_2 \sin(\omega_2 t + \varphi_{0,2}).$$



The addition can be simplified in straightforward cases. The sum of two oscillations of equal frequency and phase $x_1 = A_1 \sin(\omega t)$ and $x_2 = A_2 \sin(\omega t)$ is

$$x = x_1 + x_2 = A_1 \sin(\omega t) + A_2 \sin(\omega t) = (A_1 + A_2)\cdot\sin(\omega t).$$

The sum of two oscillations of equal frequency and amplitude but of different phase:

$$x = x_1 + x_2 = A\cdot\sin(\omega t + \varphi_{0,1}) + A\cdot\sin(\omega t + \varphi_{0,2})$$

can be written with the help of the formula $\sin(\alpha) + \sin(\beta) = 2\cdot\sin\left(\frac{\alpha+\beta}{2}\right)\cdot\cos\left(\frac{\alpha-\beta}{2}\right)$ as:

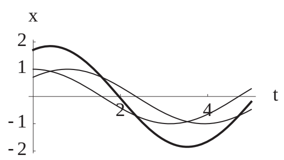
$$x = A\cdot(\sin(\omega t + \varphi_{0,1}) + \sin(\omega t + \varphi_{0,2})) = 2A\cdot\cos\left(\frac{\varphi_{0,1} - \varphi_{0,2}}{2}\right)\cdot\sin\left(\omega t + \frac{\varphi_{0,1} + \varphi_{0,2}}{2}\right).$$

The sum of the two oscillations is an harmonic oscillation of frequency ω whose initial phase is the arithmetic mean of the initial phases of the two partial oscillations

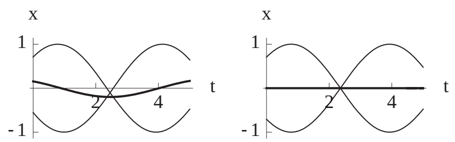
$\varphi_0 = \frac{\varphi_{0,1} + \varphi_{0,2}}{2}$ and whose amplitude is

$$A_{\mathrm{sum}} = 2A\cdot\cos\left(\frac{\varphi_{0,1} - \varphi_{0,2}}{2}\right).$$

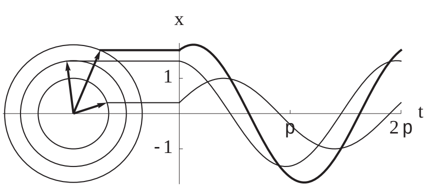
The difference $\Delta\varphi = \varphi_{0,1} - \varphi_{0,2}$ is called the *phase difference*. The following example uses the values $\varphi_{0,1} = \pi/4$, $\varphi_{0,2} = \pi/2$, $A = 1$ and $\omega = 1$.



As the phase difference of the two oscillations approaches π, the amplitude of their sum diminishes. When the phase difference is equal to π, the two oscillations cancel each other.

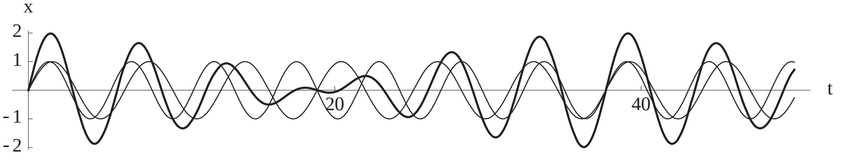


The sum of two oscillations of the same frequency but of different phase and amplitude is an harmonic oscillation whose phase and amplitude can be seen in a *pointer diagram*. If the oscillations are represented as circular motion (2.2.1.1), the pointer for their sum is the sum of the pointers of the two partial oscillations. In the illustration below, the bold curve represents the sum of two sine waves of different phase and amplitude.



2.2.1.4 Beats

When two oscillations of nearly equal frequency are superposed, so-called *beats* result. Because of the difference in frequency between the two oscillations, their phase difference gradually changes so that there are moments when they have the same phase and amplify each other and moments when the phase difference is equal to π and they cancel each other. The following illustration shows two oscillations with a frequency ratio of 6:7 and their sum (bold).



The sum of the oscillations

$$x = x_1 + x_2 = A\cdot\sin(\omega_1 t) + A\cdot\sin(\omega_2 t) = A(\sin(\omega_1 t) + \sin(\omega_2 t))$$

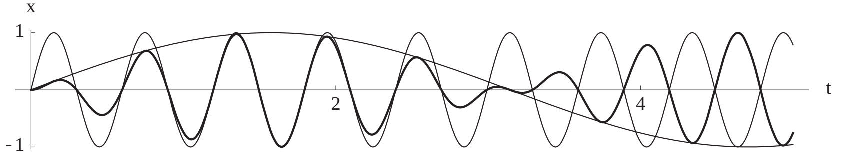
can be written as

$$x = 2A\cdot\cos\left(\frac{\omega_1 - \omega_2}{2}\,t\right)\cdot\sin\left(\frac{\omega_1 + \omega_2}{2}\,t\right)$$

or, since $\omega = 2\pi f$, as

$$x = 2A\cdot\cos\left(2\pi\frac{f_1 - f_2}{2}\,t\right)\cdot\sin\left(2\pi\frac{f_1 + f_2}{2}\,t\right).$$

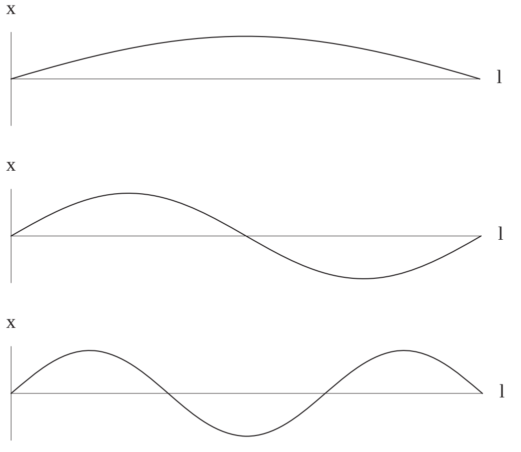
The resulting function corresponds to a sinusoidal oscillation of frequency $(f_1 + f_2)/2$ and amplitude $2A\cdot\cos(2\pi t(f_1 - f_2)/2)$, which varies slowly because $(f_1 - f_2)/2$ is small.
(Ø *Beats.maxpat*) (Ø *Beats_2.maxpat*) (Ø *Beats_Partials.maxpat*)



2.2.1.5 Natural Vibrations

Elastic bodies like strings, bars, gases, etc. oscillate if they are excited by an external force. If a body oscillates freely after the exciting force disappears, it does so only at certain frequencies. These oscillations are known as *natural oscillations* or *natural vibrations* of the body. When the body oscillates perpendicularly to the body's axis or surface, one speaks of *transversal* oscillation. Oscillation along the axis or surface is called *longitudinal*. Points whose displacement is always zero are called *nodes*, those whose displacement reaches a maximum are called *antinodes*. (Ø *natural_vibrations.pde, natural_vibrations_sum.pde*)

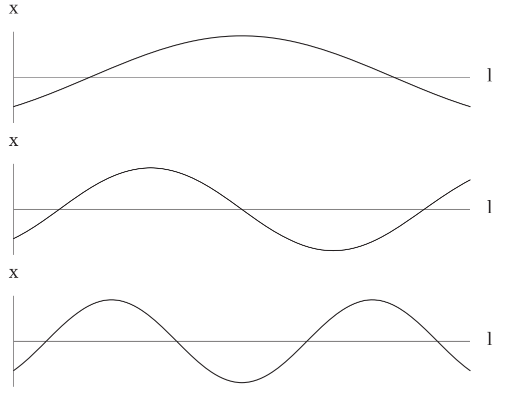
The following illustration shows the first three transverse natural oscillations of a string. Their frequencies are related as 1:2:3. (Ø Animation)



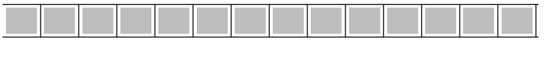
The attached ends of the string are always nodes. For the *n*th natural frequency there are $n - 1$ additional nodes which divide the string into *n* equal parts. The so-called *natural frequencies* or *eigenfrequencies* of the string can be calculated from the string's cross section *S*, from its normal stress s and from its density r:

$$f_n = \frac{n}{2l}\sqrt{\frac{s}{rS}} \,.$$

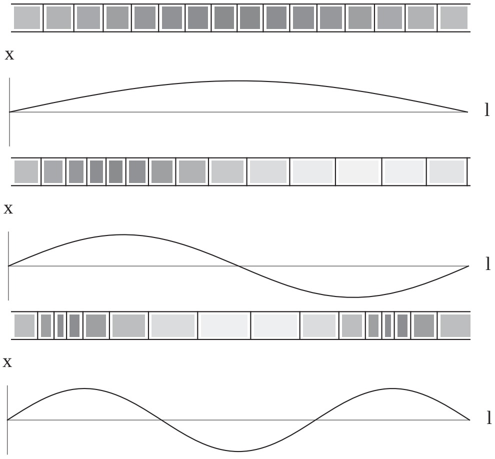
The following illustration shows the first three natural vibrations of a freely oscillating metal bar. The eigenfrequencies of metal bars are not harmonic (see [20]). (Ø Animation)



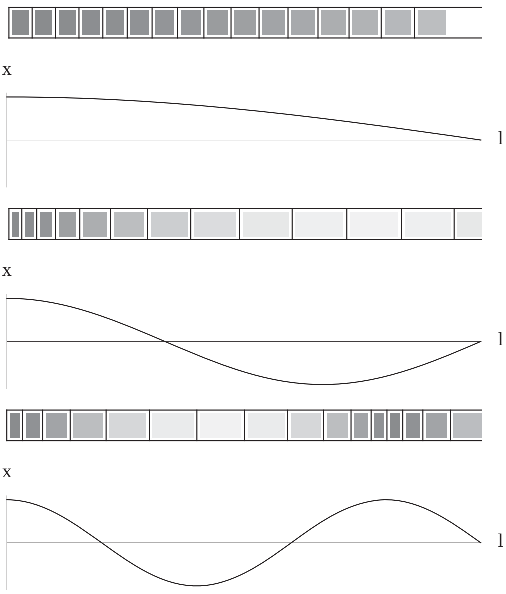
The oscillations which cause a column of air to sound are longitudinal. They can be described as variations of pressure or as displacement of the oscillating air molecules. In the animation below, an air column is modeled as a series of cells of equal volume. Air pressure is indicated by the intensity of the grey scale.



The next illustration shows the moment of maximum displacement for the air column's first three normal modes. If the distribution of pressure *p* is plotted as a function of position *l*, the same picture results as for the normal oscillations of the string shown above. (Ø Animation)



For tubes with one open and one closed end, there is always an antinode of the oscillation at the open end (that is, a pressure node) and an oscillation node (that is, pressure antinode) at the closed end. The following normal modes of oscillation are possible (Ø Animation):



This explains the typical properties of stopped organ pipes (Stopped Diapason, Gedackt, Bourdon). The first eigenfrequency is an octave lower than for a pipe of the same length open at both ends, and only odd-numbered partial tones are produced. (The frequency ratios of the oscillations are 1 : 3 : 5 ...)
(Ø *Natural_Modes_Open_Tube.maxpat*, *Natural_Modes_Closed_Tube.maxpat*)

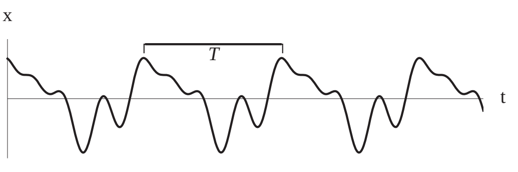
2.2.1.6 Driven Oscillation and Resonance

When a periodically oscillating force acts upon an object capable of oscillation, that object, after a certain settling time, effects *driven oscillation*, the frequency of which is equal to that of the driving force. If the driving force is a sinusoid $F = F_0\sin(\omega t)$, then the driven oscillation is $x = x_0\sin(\omega t + \alpha)$. The amplitude of the driven oscillation x_0 depends on the ratio between the frequency ω of the exciting force and the eigenfrequency ω_e of the system. If the ratio ω/ω_e is close to 1, x_0 increases greatly and the system begins to resonate.

2.2.2 Periodic Vibrations and their Spectrum

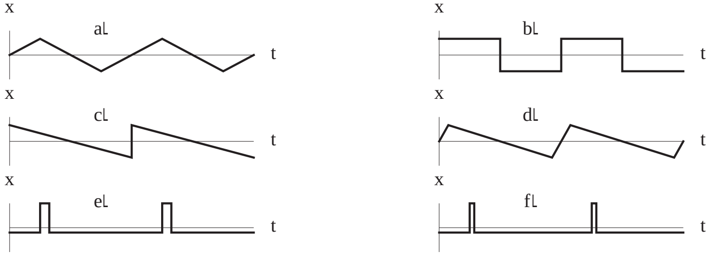
2.2.2.1 The Definition of Periodic Vibrations

A repeated vibration or oscillation is called *periodic*. A particular vibration is determined by the duration of its period T, or by its frequency f, and by the form of the vibration. The pitch of the heard tone depends on the frequency $f = 1/T$. The vibration's form or shape influences only the tone's timbre. The curve corresponding to the vibration's form cannot have any discontinuities, because the vibrating point of a physical object cannot suddenly be in a different place. The curve cannot have any corners, because that would correspond to a sudden change of velocity. Thus the function describing the curve must be continuous and differentiable.



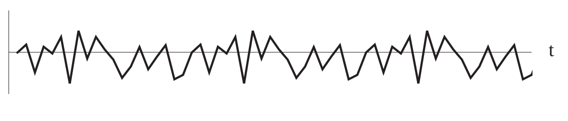
2.2.2.2 Standard Examples

The following waveforms, which are impossible in the real world because they have corners and discontinuities, are often approximated in electroacoustic music: triangle wave (a), square wave (b), sawtooth wave (c, d), impulses of different impulse length or duty cycle (e, f). The square wave corresponds to an impulse with a duty cycle of 50%.



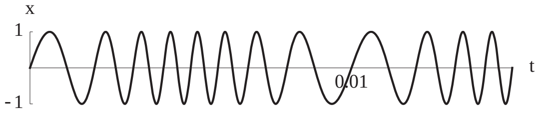
2.2.2.3 Other Examples

Periodic oscillation occurs whenever any waveform is repeated.



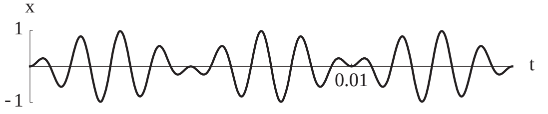
The following example shows a sine wave of varying period. Because the variation is regular and repeats every 0.01 s, the result is a periodic oscillation with a fundamental frequency of 100 Hz (cf. Frequency Modulation 6.1.2).

```
w=800*2*p;w2=100*2*p;p1=Plot[Sin[w*t+4*Sin[w2*t+p]]...]
```



In the following example the amplitude of a sine wave varies regularly, causing again a periodic oscillation of frequency 100 Hz (cf. Amplitude Modulation 6.1.1).

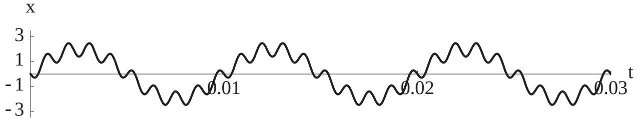
```
w=100*2*p;p1=Plot[Sin[w*t]*Sin[8*w*t]...]
```



2.2.2.4 Constructing Periodic Oscillation from Harmonic Waveforms

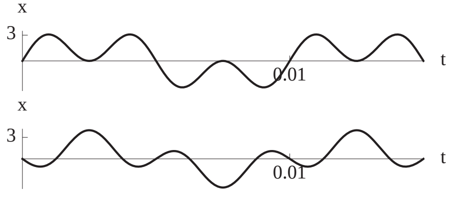
When one adds together harmonic oscillations whose frequencies are all multiples of a fixed *fundamental frequency* f_1, one obtains periodic oscillation, regardless of the amplitudes and initial phase angles of the components. Because the sine function is symmetrical and repeats after 2p, any sinusoidal oscillation can be written with positive amplitude and phase constant between 0 and 2p, or with positive or negative amplitude and a phase constant between 0 and p. In the first example, we add to a sine oscillation of 100 Hz (period = 0.01 s) and amplitude 2 a second sine oscillation of 900 Hz and amplitude –0.6.

```
w=100*2*p;Plot[2*Sin[w*t]-.6*Sin[9*w*t]...]
```



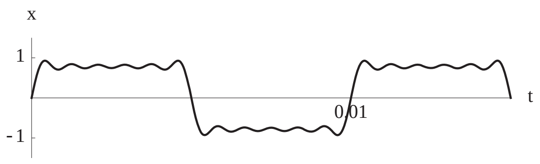
If one varies the phase constant of a partial oscillation, the waveform of the periodic oscillation changes. In the second waveform the phase of the third harmonic is displaced by p.

```
Plot[2*Sin[w*t]+2*Sin[3*w*t]...]
Plot[2*Sin[w*t]+2*Sin[3*w*t+p]...]
```



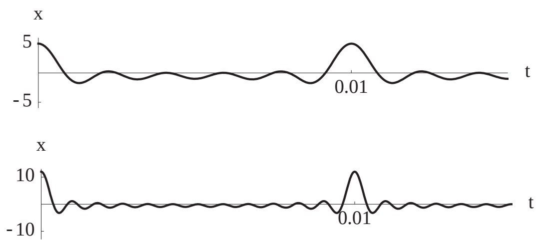
One can approximate a square wave by summing the odd partials with amplitudes inversely proportional to the partial number (1, 1/3, 1/5, ...). (Ø Animation)

```
Plot@Sin@w * tD + Sin@3 * w * tD ê 3 + Sin@5 * w * tD ê 5 +
   Sin@7 * w * tD ê 7 + Sin@9 * w * tD ê 9 + Sin@11 * w * tD ê 11 ...D
```



One can approximate a pulse train by adding together several harmonic cosine waves of the same amplitude. As the number of partials increases, the width of the pulse becomes narrower and the curve between the pulses becomes flatter. (Ø Animation)

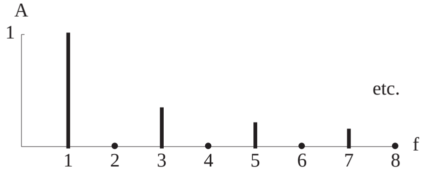
```
Plot@Cos@w * tD + Cos@2 * w * tD + Cos@3 * w * tD ...D
```



(Ø *Timbre_and_Spectrum.maxpat*)

2.2.2.5 The Spectrum of Periodic Oscillations

At the beginning of the 19th century, the French mathematician and physicist Jean Baptiste Joseph Fourier showed that any waveform with the period T could be expanded into a series of sine waves of frequency $f_1 = 1/T$, $f_2 = 2f_1$, $f_3 = 3f_1$, ... having suitable amplitudes A_1, A_2, A_3, ... and phase constants f_1, f_2, f_3, ... The sinusoidal oscillations f_1, f_2, f_3, ... are known as the partials of the waveform; the overtones of a waveform are the partials without the fundamental frequency. The representation of the amplitudes A_1, A_2, A_3, ... as a function of frequency is called the amplitude spectrum of the periodic oscillation or sound. Although the phase constants are essential for determining the shape of the waveform of a sound, they are usually disregarded in the spectrum. The spectrum shows the partials present in a sound and hence informs one about the sound's timbre. The missing phase information is generally not a problem, because the phase constants do not as a rule influence the timbre. The following illustration shows the spectrum of a square wave.



Any harmonic sound of period T and corresponding angular frequency $\omega = 2\pi f = 2\pi/T$ can be produced by summing sine waves.

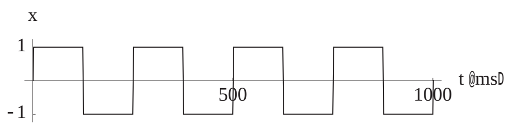
$$\boxed{x(t) = \sum_{n=1}^{\infty} A_n \sin(n\omega t + \varphi_n)}$$

The indentity $\sin(\alpha + \varphi) = \sin(\alpha)\cdot\cos(\varphi) + \cos(\alpha)\cdot\sin(\varphi)$ implies that a sine wave of an arbitrary phase constant φ can be represented as a weighted sum of a sine wave and a cosine wave of the same frequency without phase constant:

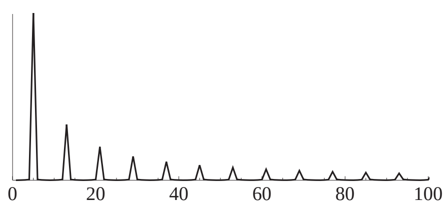
$$\boxed{\begin{array}{c} x(t) = \sum_{n=1}^{\infty} (B_n \sin(n\omega t) + C_n \cos(n\omega t)) \\ \text{with } B_n = \cos(\varphi_n) \text{ and } C_n = \sin(\varphi_n) \end{array}}$$

The calculation of the coefficients A_n (or B_n and C_n) in the formula above can be carried out exactly and without computer only in simple cases, like that of the square wave. There are, however, programs that can calculate the spectrum of a given sound using the so-called Fast Fourier Transform (cf. 3.2.2.5) and show the spectrum graphically. In this book, we will use the Fourier function in Mathematica to evaluate the spectrum of a list of values representing a waveform. In the following example, we calculate a square wave of 4 Hz and store 1000 points of this function in a list *list1*. Mathematica then evaluates the spectrum of the list.

```
list1=Table[Sign[Sin[ω*t/1000]]...]
ListPlot[list1]
```

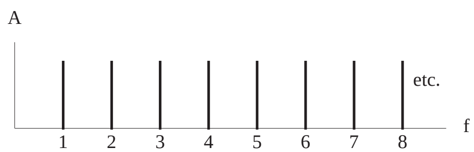



```
ListPlot[Fourier[list1]]
```



There are only odd partials (4, 12, 20, 28, 36, 44 Hz ...) whose amplitudes are inversely proportional to their frequencies.

In Chapter 2.2.2.4 a pulse train was approximated by adding several harmonic cosine functions of the same amplitude. To produce ideal pulses, infinitely many partials would have to be summed. In this case, the waveform of the pulse train and its spectrum would be identical.

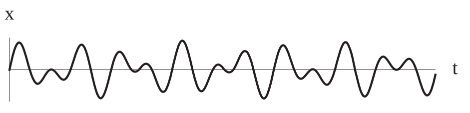


2.2.3 Aperiodic Oscillation

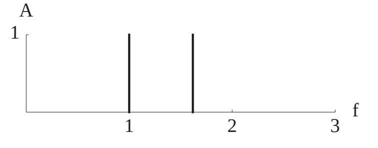
2.2.3.1 Non-harmonic Partials

If we add two sine waves the ratio of whose frequencies is irrational, the resulting oscillation will not contain any periods. In the following example, two sine waves with the frequency ratio of the golden ratio 1 : 1.61803 ... (10.3.2) are combined. Although the resulting waveform looks simple and sounds smooth, in fact no section of it, taken at any order of magnitude, is ever repeated.

```
Plot[Sin[w*t]+Sin[1.61803*w*t]
```

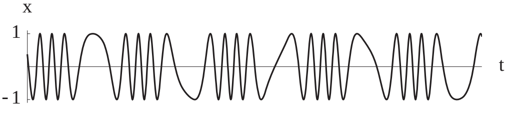


The spectrum does not indicate that the oscillation is aperiodic, because one cannot tell whether the frequency ratio is rational or irrational.



If we change the second example of Chapter 2.2.2.3 so that the ratio of frequency variation to the sine wave's nominal frequency is irrational, an aperiodic oscillation results.

```
w=437.32171*2*p;w2=100*2*p;
Plot[Sin[w*t+3.5*Cos[w2*t+p]],...]
```

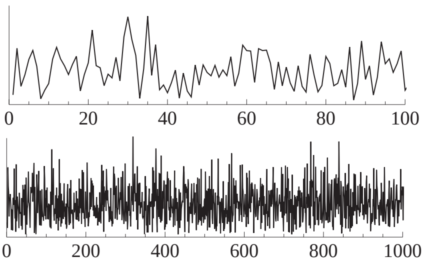


2.2.3.2 Noise

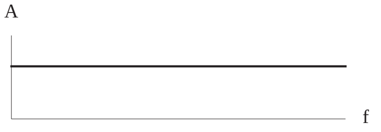
Random oscillation yields *noise*. The waveform in the illustration below was generated using a list (sequence) of random numbers between –1 and 1.

```
l=RandomReal[{-1,1},100];
```

The following figure shows two spectra, the first of a list of 200 random numbers, the second of a list of 2000 random numbers. Clearly, the first series will have a maximum of 100 oscillations, the second of 1000 oscillations. The spectra show partials in all frequency ranges. The amplitudes of the partials are randomly scattered around a certain mean value. In the case of light, this superposition of frequencies corresponds to the superposition of all the colors of the spectrum and gives rise to white light. Noise having these spectral characteristics is therefore known as *white noise*.



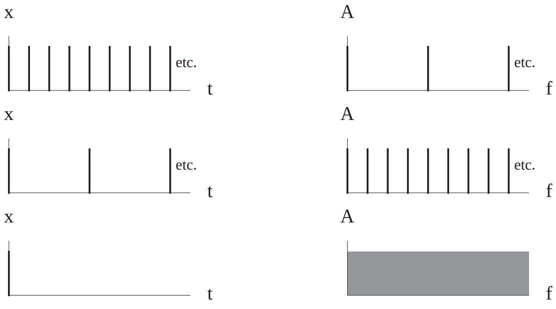
The spectrum of ideal white noise is a straight line.



If some frequencies are more strongly represented than others, one speaks of *colored noise* (5.3.2.1). White noise contains as much high-frequency as low-frequency energy, which is not the case for most natural sounds. Noise whose spectrum decreases exponentially with frequency is generally felt to be more natural. It is called pink noise and is often found in commercial synthesis devices.

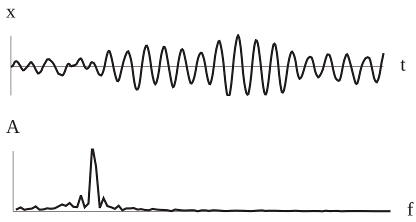
2.2.3.3 Pulses

In Chapter 2.2.2.5 we saw how pulse trains can be made by summing cosine functions whose frequencies are multiples of a fundamental frequency. A single pulse can be considered to be the sum of cosine functions of all frequencies, since the functions $A \cdot \cos(\omega t)$ all take the value A at time $t = 0$, while at time $t \neq 0$ they have different values and cancel each other in the limit. The spectrum can be explained as follows. The spectrum of a pulse train is a line spectrum of infinitely many components of equal amplitude, separated by the frequency of the pulse train. As the pulse train's frequency decreases (and hence the length of its period increases), the spectral lines move closer to one another, blending together to form a single rectangle when the period T becomes infinitely long. Thus a single pulse has the same spectrum as does white noise.



2.2.3.4 Quasi-periodic Oscillation

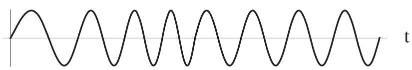
Various kinds of oscillation exist which are not strictly periodic but can be described as slowly varying or mildly perturbed oscillation. Beats, for instance, can be characterized as periodic oscillation with slowly varying amplitude, particularly when the ratio of the beat frequency to the oscillation’s frequency is small (2.2.1.4). Whistle tones have a small noise component which makes them quasi-periodic, but to the ear they have a clearly defined pitch. The quasi-periodicity is easily recognizable in the waveform of a whistle tone below, and the tone’s spectrum shows a frequency line which is only slightly broadened.



2.2.3.5 Variable Frequencies

The multiplier w in the function sin(w*t*) should not be confused with the frequency if it is not constant, i.e., if w = w(*t*). Using the Mathematica command below, we produce a function w(*t*) which begins at 2·2p, goes in one second to 4·2p and remains at 4·2p for another second.

```
Plot[Sin[If[t<1,(2+2*t)*2p*t,4*2p*t]]...]
```

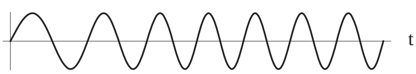


At the end of the glissando, that is in the middle of the depicted waveform, the frequency is noticeably higher than 4 Hz. (Ø Sound Example)

Using the formula above, one can only produce tones of constant frequency correctly. Because the frequency corresponds to the velocity of the changing argument of the sine function, that is, the phase, the instantaneous frequency is defined as the derivative of the phase with respect to time. When the argument *c* of sin(*c*) is constant, the frequency is *dc*/*dt* = 0; the function sin(w*t*) has the constant frequency *d*(w*t*)/*dt* = w. We calculate the phase by integrating the instantaneous time-dependent frequency w(*t*). Thus in the first example, we have for the phase, instead of (2 + 2*t*)·2p*t* = 4p*t* + 4pt^2, the function

```
ÿH2 + 2 * tL * 2 p „ t      = 4 p t + 2 p t^2

PlotASinAIfAt < 1, 4 p t + 2 p t^2, 4 * 2 p * tEE ...E
```



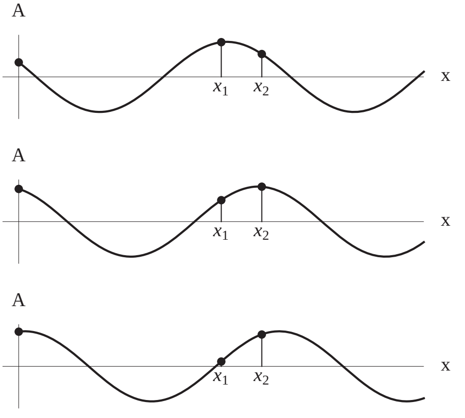
Often the frequency w(*t*) cannot be integrated analytically, or else the variation of frequency is not known. In such cases, integration has to be done numerically during the calculation of the waveform. (Ø Sound Example) (Ø *Variable Frequency.maxpat*)

2.2.4 Waves

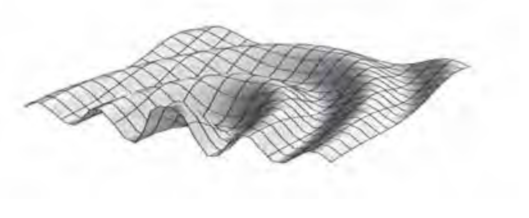
2.2.4.1 Definition and Examples

The word "wave" originally referred to changes on the surface of a liquid. More generally, a wave is the propagation of a physical state through a medium. In an undisturbed medium, all parts of the medium are at rest. If one particle of the medium is moved out of its rest position by excitation, interaction among neighboring particles causes the excitation to be transmitted through the medium, forming a wave. One distinguishes between *transverse waves*, where the displacement of the particles is perpendicular to the direction of propagation, and *longitudinal waves*, where the displacement is in the same direction as the wave's propagation. The phase of a wave refers to its position in the vibration cycle. Points on a wave having the same phase form a *wave front*.

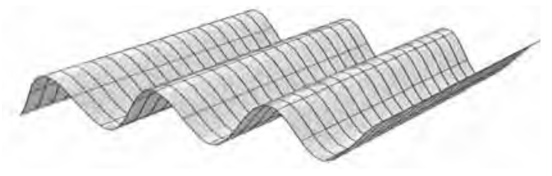
The first example shows a one-dimensional transverse wave caused by excitation at $x = 0$. One speaks of a *sine wave* or an *harmonic wave* if the waveform can be described by a sine function. Note that the horizontal axis in the graphic represents a spatial dimension and not time. Therefore, the following figures do not represent, as in case of oscillation, behavior over time. Rather, they are snapshots of instantaneous states. To illustrate the behavior of a wave over time, an animation can be made or several successive snapshots can be presented. The figure below shows three images from an animation. Considering two neighboring points, one can see that they go through the same motion at different times. Hence, a wave can be described as a set of coordinated oscillations of points in space. In the propagation of waves, no particle of the medium itself is transported. (Ø Animation)



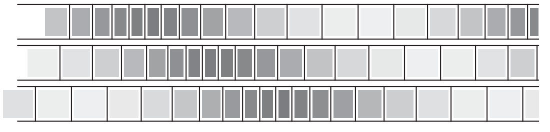
In the two-dimensional wave below, circular wavefronts move away from the point of excitation. Their amplitude diminishes with the distance from the excitation because the total energy is distributed over an ever-larger circle. (Ø Animation)



If the excitation takes place along a straight line instead of at a single point, two-dimensional waves with straight wavefronts result. Waves excited at a single point also exhibit nearly straight wavefronts at considerable distance from the point of excitation. (Ø Animation)



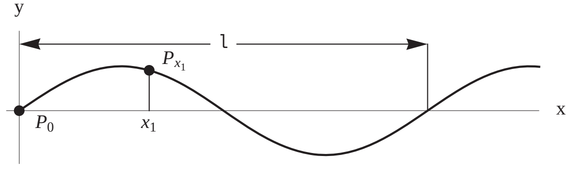
Longitudinal waves are more difficult to imagine and visualize. Experimentation with spiral springs can help the imagination here. In the following animation a longitudinal wave is created by a sinusoidal excitation in the direction of the wave's propagation. (Ø Animation)



Three-dimensional waves cannot be illustrated even with animations. Sound waves are three-dimensional longitudinal waves of changing air pressure which propagate in all directions away from the sound source, giving rise to spherical wavefronts. All points on a wavefront have the same air pressure.

2.2.4.2 Mathematical Description

The displacement $y(t, x)$ of the oscillating particles of a wave depends on time t and position x of the particle. It is meaningful, for one-dimensional undamped waves, to speak of the wave's *amplitude* and *frequency*, because the oscillations of all the particles always have the same frequency and amplitude. The distance between the nearest particles of different fronts oscillating with the same phase is called the *wavelength* λ. The wavelength must not be confused with the period of an oscillation, which corresponds to the reciprocal of the frequency. The period is a duration in time, the wavelength a distance. The wave number k indicates the number of waves per unit of length, hence $k = 1/\lambda$ and $\lambda = 1/k$. The speed with which an excitation propagates in a medium is known as the *wave velocity* c (or more usually the *phase velocity*, referred to a specific phase of the wave). In acoustics, one speaks of the *speed of sound*. At the frequency f, a wavefront covers a distance of $f\lambda$ in one second. For the wave velocity we write: $c = f\lambda = \lambda/T$, where T is the wave's period.



To derive the mathematical description of a wave, we begin by considering the point P_0 at $x = 0$. Its movement is described by the sinusoidal oscillation $y = A \cdot \sin(\omega t)$. Any point on the wave oscillates with the same amplitude and frequency but with a time delay dependent upon its distance from P_0 and upon the wave velocity c. This difference of phase is equal to $2\pi/\lambda \cdot x$.

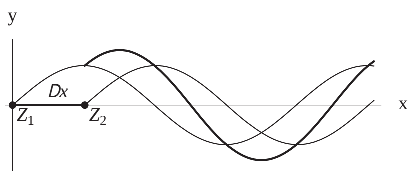
Hence, the equation for the harmonic wave giving the displacement of the particles as a function of their positions and of time is

$$\boxed{\begin{aligned} &y = A\cdot\sin\left(\omega t - \frac{2\pi}{\lambda}x\right) \quad \text{or} \\ &y = A\cdot\sin\left(2\pi\frac{t}{T} - 2\pi\frac{x}{\lambda}\right) \quad \text{or} \\ &y = A\cdot\sin\left(2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right)\right) \end{aligned}}$$

Here y = displacement, A = amplitude, T = period of the oscillation, λ = wavelength, t = time, x = distance from the origin $x = 0$.
If t is made constant, then $a = t/T$ is also constant, and the equation for the waveform becomes $y(x) = A\cdot\sin(2\pi(a - x/\lambda))$. If x is made constant, then $b = x/\lambda$ is also constant, and the equation for the oscillation of the point P_x becomes $y(t) = A\cdot\sin(2\pi(t/T - b))$.

2.2.4.3 The Superposition of Waves

The superposition of waves is given by adding their instantaneous displacements. If two waves of the same vibration direction and of equal frequency are superposed, one speaks of *interference*. The following figure shows two sine waves with the origins Z_1 and Z_2 and the wave resulting from the superposition of the two. The distance between the two origins Δx is called the phase difference.



The equations for the two waves are

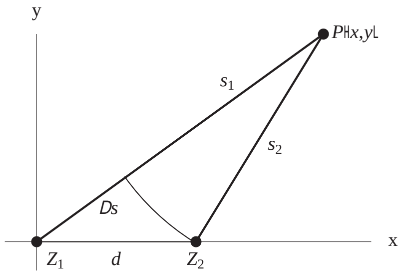
$$y_1 = A\cdot\sin\left(2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right)\right) \text{ and } y_2 = A\cdot\sin\left(2\pi\left(\frac{t}{T} - \frac{x+\Delta x}{\lambda}\right)\right).$$

Their sum is

$$\begin{aligned} y = y_1 + y_2 &= A\cdot\sin\left(2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right)\right) + A\cdot\sin\left(2\pi\left(\frac{t}{T} - \frac{x+\Delta x}{\lambda}\right)\right) \\ &= A\left[\sin\left(2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right)\right) + \sin\left(2\pi\left(\frac{t}{T} - \frac{x+\Delta x}{\lambda}\right)\right)\right] \\ &= 2A\cdot\cos\left(\pi\frac{\Delta x}{\lambda}\right)\cdot\sin\left(2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right) - \pi\frac{\Delta x}{\lambda}\right) \qquad \text{(see 2.2.1.3)} \\ &= 2A\cdot\cos\left(\pi\frac{\Delta x}{\lambda}\right)\cdot\sin\left(2\pi\left(\frac{t}{T} - \frac{x+\Delta x/2}{\lambda}\right)\right). \end{aligned}$$

Only the term sin(...) depends on t and x, and so this equation describes a sinusoidal oscillation having the frequency of the two waves and the constant amplitude $2A\cdot\cos(\pi\cdot\Delta x/\lambda)$. The amplitude depends on the phase difference. If the phase difference is a multiple of the wavelength, the amplitude of the sum is maximal, if the ratio of phase difference to wavelength is 0.5, 1.5, 2.5, etc., the amplitude is minimal.

When circular waves are superposed, there are areas where the waves amplify each other and areas where they cancel each other. To calculate the points where the waves amplify each other maximally, the distance between the two origins Z_1 and Z_2 is called d, the distances of a point $P(x, y)$ to the two origins are called s_1 and s_2, and the difference between these distances Δs.



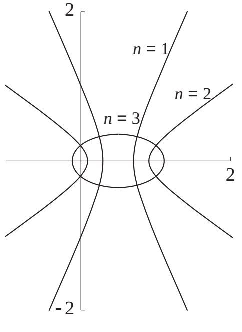
The greatest amplification occurs where Δs is a multiple of the wavelength λ. If we take the origin Z_1 as the origin of our coordinate system, we have

$$\sqrt{x^2 + y^2} - \sqrt{(x - d)^2 + y^2} = n + \lambda.$$

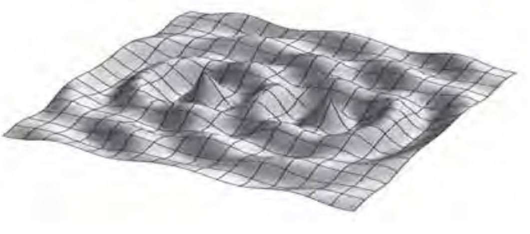
We solve the equation and show the solutions for $n = 1$, $n = 2$ and $n = 3$ graphically ($\lambda = 0.41$, $d = 1$). Here we solve the equation for x, because, as the figure below shows, the curves for $n = 1$ and $n = 2$ can only be written as unique functions in terms of y.

```
Solve[√(x² + y²) - √((x - d)² + y²) == λ, x]
```

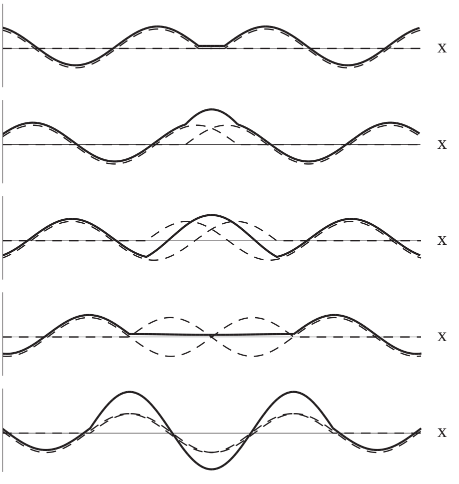
$$\to \frac{d^3 - d\,\lambda^2 \pm \lambda \sqrt{d^2 - \lambda^2}\,\sqrt{d^2 + 4y^2 - \lambda^2}}{2d^2 - 2\lambda^2}$$



The following figure shows a snapshot of a corresponding animation.
(→ Animation) (→ *Interference.maxpat*)

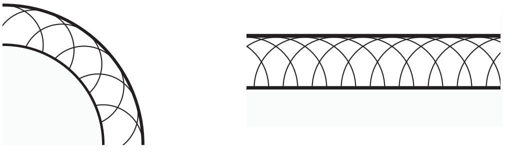


When two waves of the same frequency meet each other traveling in exactly opposite directions, a phenomenon occurs which is of particular importance for understanding the behavior of strings and air columns. The superposition creates points (in one-dimensional waves) or lines (in two-dimensional waves) where the waves' displacements are cancelled and points or lines where they are amplified. This phenomenon gives rise to *standing waves*, which correspond to the natural oscillations of strings and air columns. The following excerpts from an animation show two waves meeting each other head-on (dashed lines) and the resulting wave (bold line). (Ø Animation)

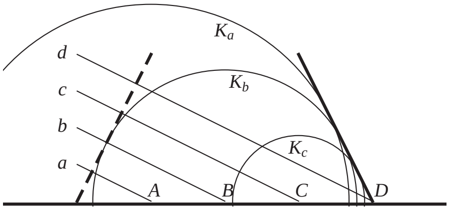


2.2.4.4 The Propagation of Waves

Many features of the behavior of waves can be explained by the *Principle of Huyghens*, named after the Dutch mathematician, astronomer and physicist Christiaan Huyghens (1629–1695). The Principle states that at any instant the wavefront of a propagating wave corresponds to the envelope of the spherical so-called wavelets or elementary waves emanating from every point on the wavefront at the prior instant. In the figures below, the inside (figure left) and the lower (figure right) lines show the prior wavefront and the outside and upper lines show the present wavefront as the envelope of the more finely drawn wavelets.



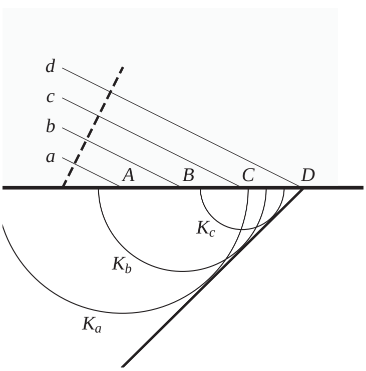
From the Principle of Huyghens we can derive the *law of reflection*, which says that the angle with which an incident wave strikes a reflective surface is the same as that of the reflected wave. The following figure shows a straight wavefront coming from the upper left (heavy dashed line) which is reflected on a horizontal surface. Four so-called normals, *a*, *b*, *c*, and *d*, are drawn perpendicular to the incident wavefront, meeting the surface at *A*, *B*, *C* and *D* respectively. By the time the wave reaches *D*, it has already reached *A*, *B* and *C* and provoked the wavelets K_a, K_b and K_c. The envelope of the wavelets is the straight line (heavy line) which includes the same angle to the surface as the incident wavefront. That the angles are the same can be seen from the normal *b*, which provokes the wavelet K_b at *B*, whose current radius is the same length as the normal from the incident wavefront to the point *B*.

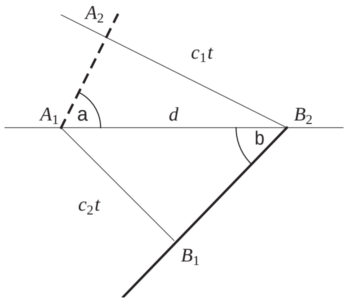


The displacement of reflected waves either keep their original direction or they are multiplied by –1. A wave in a wire which is fixed at one end turns negative when reflected at the fixed end and remains positive when reflected at the loose end (8.2.2.2).

In the discussion of reflection above, it may seem that the Principle of Huyghens offers a complicated description of a simple situation, but it provides a very elegant explanation for the phenomenon of *refraction*. Refraction is familiar from the change in direction of light when it passes from air to water. However, waves are always refracted when they pass from one medium to another having a different wave velocity. In the following figures the horizontal line represents the boundary between two media. At the bounding surface part of the wave is reflected, while another part continues in the new medium but in a different direction. In the example below, the phase velocity of the lower medium is lower than in the upper medium, hence the radius of the wavelet K_b is smaller than the length of the normal *b* between the wavefront (dashed line) and the point *B*. The simplified figure to the right will make the calculation of the relationship between the angle of the incident wave and that of the refracted wave easier. During time *t* the incident wave passes through the first medium with phase velocity c_1 from A_2 to B_2. During the same time, the refracted wave moves through the second medium with phase velocity c_2 from A_1 to B_1. From the illustration it follows that

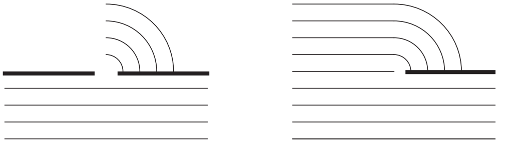
$$\frac{\sin(\alpha)}{\sin(\beta)} = \frac{c_1\, t/d}{c_2\, t/d} = \frac{c_1}{c_2}.$$





Refraction occurs not only at bounding surfaces but happens continuously in non-homogenous media. Sound waves, for instance, are skewed towards the ground when the air near the ground is colder than the air above, lowering the speed of sound close to the ground.

If a wavefront strikes a surface with a slit-like aperture, new circular waves, corresponding to the wavelets described above, propagate behind the aperture (figure left). By the same reason, there are no sharp shadows behind an obstacle struck by a wave: part of the wave is bent around the obstacle (figure right). This bending is called *diffraction*; its degree depends on the size of the obstacle and on the wavelength. The sound waves of low tones are more strongly diffracted than those of high tones.

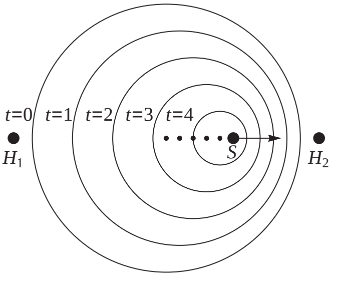


The propagation speed of certain waves depends on their wavelength. The small ripples caused by the wind on the surface of water, for instance, move more slowly than the bow wave of a ship. The dependence of the propagation speed of a wave on its wavelength is known as *dispersion*. Non-sinusoidal waves suffer from dispersion, because the various spectral components have different wave velocities.

In the passage from one medium to another, but also in the propagation of waves in a homogeneous medium, energy is lost because of *absorption*. Absorption at bounding surfaces plays an important role in room acoustics, because the reverberation of a room depends on the absorptivity (9.2.1.2). Absorption by a medium not only contributes to a decrease of amplitude in circular and spherical waves but also causes a frequency-dependent change of spectrum (9.1.2.1). The decrease of energy caused by absorption is exponential: $W = W_0 \cdot e^{-kx}$, where W is the wave's energy after traversing the distance x, W_0 is the energy at the wave's origin ($x = 0$) and k is the *absorption coefficient*, which is dependent on the properties of the medium, the kind of wave and its frequency.

2.2.4.5 The Doppler Effect

If a sound source and a listener move relative to one another, the frequency of the sound generated and the frequency of the sound heard will not be the same. The heard frequency will be higher than the generated frequency when source and listener approach each other, lower when they move apart. The figure below illustrates the so-called *Doppler-Effect* (named after the Austrian physicist Christian Doppler, who described it in 1842) with the example of a moving sound source S and two stationary listeners H_1 and H_2. At time $t = 0$, the sound source was halfway between the two listeners and then began moving to the right. The circles represent the sound waves produced at different times, the dots indicate the position of the source at those times. The wavelength of the waves arriving at listener H_1 is greater than that produced by the source, and the perceived frequency is correspondingly lower. Conversely, the wavelength of the waves arriving at listener H_2 is smaller than that produced by the source, and the perceived frequency is correspondingly higher.



Let the speed of sound c = 340 m/s and the frequency f_0 be that of a sound source. Then there are f_0 cycles of sound distributed along a distance of 340 m = c·1 s. If the sound source moves with a velocity v_s towards the listener, the cycles of sound will be distributed along $(c - v)$·1 s meters. The wavelength of the compressed wave is therefore

$$\lambda = \frac{c - v_s}{f_0},$$

and since $c = \lambda \cdot f$, the resulting frequency is

$$f = \frac{c}{\lambda} = \frac{c}{\frac{c - v_s}{f_0}} = \frac{cf_0}{c - v_s} = \frac{f_0}{\frac{c - v_s}{c}} = \frac{f_0}{1 - v_s/c}.$$

If the sound source moves away from the listener, its velocity is negative. If we consider the case of a stationary source and a moving listener, we will see that $f = f_0(1 \pm v_e/c)$. These formulas can be summarized as

$$\boxed{f = f_0 \frac{1 \pm v_e/c}{1 \mp v_s/c}}$$

Here f_0 represents the original frequency of the source, v_s the velocity of the source and v_e the velocity of the listener.

Chapter 9.1.3.1 shows how the Doppler effect can easily be simulated, even when the sound source is not moving straight in the direction of the listener. (Ø *Doppler-Effect.maxpat*, *doppler_effect.pde*)

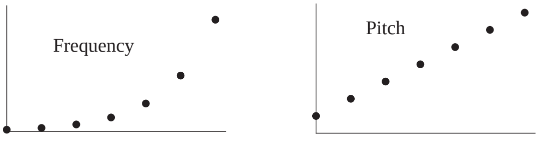
2.3 Sound and Hearing

Further Reading: Musical Acoustics *by Donald E. Hall [20],* Music Cognition and Computerized Sound *by Perry R. Cook (including a CD),* Spatial Hearing – The Psychophysics of Human Sound Localization *by Jens Blauert [45].*

2.3.1 Pitch

2.3.1.1 Frequency Range and Octaves

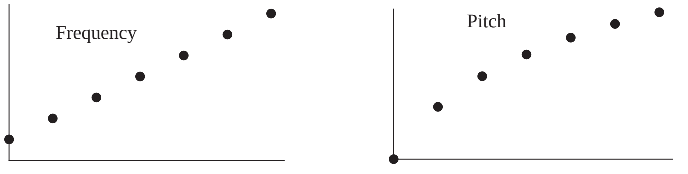
Human beings perceive vibrations with frequencies between about 20 and 20,000 Hz as tones. The doubling of the frequency of a tone raises its perceived pitch by an octave. That is why an exponential increase of frequency means a linear increase in pitch. In Sound Example 2-3-1a one hears the frequencies 250, 500, 1000, 2000, 4000, 8000 and 16,000 Hz.



Conversely, halving a tone's frequency makes its pitch drop by an octave. In Sound Example 2-3-1b, we hear the frequencies 800, 400, 200, 100, 50, 25, 12.5 and 6.25 Hz. The last two frequencies are not audible, but they can be made visible by holding a flame in front of the loudspeaker. If we generate sequences of periodic impulses instead of sine waves (Sound Example 2-3-1c), the tone does not just become softer and finally disappear: even at less than 20 Hz, the impulses remain audible. The frequency at which single events are perceived as fusing into continuity is similar for audition and vision. The individual images of a film can be perceived singly only up to a frequency of about 20 images per second.

2.3.1.2 The Harmonic Series and Pure Intervals

In Sound Example 2-3-1d, one hears a sequence of many tones. The frequency of each tone is 100 Hz higher than that of the preceding one (100, 200, 300, 400, ... 4000 Hz). The distances between the tones in perceived pitch become smaller. A linear increase in frequency corresponds to a logarithmic increase in pitch.



This succession of tones is called the *harmonic series*, or in a more general sense, an *overtone* or *partial series*. The first (lowest) tone is called the *fundamental*, the following tones the first, second, etc. overtone, or, rather confusingly, since the fundamental is considered the first partial tone, the second, third, etc. partial tone. Hence, the first overtone corresponds to the second partial tone, the second overtone to the third partial tone, etc. Starting from C^1 (or Pedal C, ca. 32.7 Hz), we have the following harmonic series (the seventh, 11th and 14 partial tones are noticeably lower than BŸ3, F¤4 and BŸ4, and the 13th partial is noticeably higher than AŸ4)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
C1	C2	G2	C3	E3	G3	BŸ3	C4	D4	E4	F¤4	G4	AŸ4	BŸ4	B4	C5

The distance between two tones, known as an *interval* in music, is defined by the relationship of the frequencies of the tones. Those intervals appearing in the harmonic series are known as *pure intervals*, because they correspond to simple frequency relationships. From the table above we can read out the frequency relationships of the following pure intervals: octave 2:1, fifth 3:2, fourth 4:3, major third 5:4, minor third 6:5, major second or whole tone (major tone) 9:8, and the major second or whole tone (minor tone) 10:9. In Sound Example 2-3-1e, we hear tones having the frequencies 200, 400, 600, ... 3200 Hz. Notice the pure intervals, the two different major seconds (major and minor tone) and the tones we do not use in our musical system corresponding to the seventh and 11th partial tones.

Just Intervals	Octave	Fifth	Fourth	Major Third	Minor Third	Just Major Tone	Just Minor Tone
Frequency Ratio	1 : 2	2 : 3	3 : 4	4 : 5	5 : 6	8 : 9	9 : 10

The significance of the harmonic series lies in the fact that a tone's or a pitched sound's timbre is mainly determined by the relative strengths of fundamental and partials (2.3.2). The

harmonic series is also the basis for many elements of traditional music theory, like octave identity, the major triad, etc. The harmonics of string instruments and the intervals in which wind instruments can be overblown correspond to the intervals of the harmonic series.
(Ø *Harmonic_Series.maxpat*)

2.3.1.3 Intervals

Because intervals are defined as frequency ratios, we add them by multiplying their ratios and subtract them by dividing their ratios. If we add a fifth and a fourth, we get the frequency ratio 2/1 from the product of 3/2·4/3. If we subtract a minor third from an octave, we get the major sixth 2/(6/5) = 5/3. If one tries to construct a circle of fifths using pure fifths, the circle does not quite close, because 12 pure fifths are a little more than seven octaves. This difference is called the *Pythagorean comma*. Twelve fifths, e.g. C1-G1-D2-...-B♯4 make an interval equal to $(3/2)^{12}$, seven octaves make an interval equal to 2^7. The difference in pitch between B♯4 and C5 is $(3/2)^{12}/2^7$ = 129.746 : 128 = 1.013643... Nor do three just major thirds give a just octave: $(5/4)^3$ = 1.95313. This discrepancy is called a *diesis*. Neither six major tones nor six minor tones give an octave. The difference between the just major and the just minor tone is known as the *syntonic comma* and corresponds to a frequency ratio of (9/8)/(10/9) = (81/80), or about 1/4 semitone. These discrepancies are the reason for the centuries-old search for a “good” tuning of the musical intervals. The discrepancies should be so distributed that they either are not noticeable or have a specific effect on the music played. Sound Example 2-3-1f shows that it is impossible to tune all the intervals even in a single key purely. After playing the degrees I–VI–II–V–I (5/3·2/3·4/3·2/3 = 81/80), we reach a note a syntonic comma below the beginning note. In so-called equal temperament, in which the octave is divided into 12 semitones of equal size, no interval except the octave is pure but rather all are defined by irrational ratios. If we call the frequency ratio of the equal tempered semitone x, then x^{12} must be equal to 2, which gives a frequency ratio of $2^{1/12}$ = 1.05946... Then we have for the major third $2^{1/3}$ = 1.25992..., for the whole tone $2^{1/6}$ = 1.12246... and for the fifth (seven semitones) $2^{7/12}$ = 1.49831. Small intervals are often indicated by cents. One *cent* is 1/100 of a tempered semitone and corresponds to the frequency ratio $2^{1/1200}$ = 1.00057779... The size of an interval in cents can be calculated from its frequency ratio using the following formula:

$$\boxed{x\,(\text{cents}) = 1200 \log_2 \frac{f_1}{f_2} = \frac{1200}{\log 2} \log \frac{f_1}{f_2}}$$

Whether or not the pitches of two tones can be discriminated depends not only on the ratio of their frequencies, but also on the frequency range in which they occur and on the ability of the listener. This ability can be trained to some degree. Under ideal circumstances, most wind and string players can distinguish between tones whose frequencies vary by only a few cents. The literature indicates that the smallest Just Noticeable Difference JND for an average listener at frequencies around 100 Hz is 3% (50 cents) and is .8% around 2000 Hz (8.6 cents) (see [2] p. 33). Another representation (see [20] p. 113) shows a JND of ca. 1 Hz for frequencies below 1000 Hz (17 cents at 100 Hz, 1.7 cents at 1000 Hz) with a relatively rapid increase above 1000 Hz.

The frequencies of the chromatic scale can be calculated by using a pitch reference from which to start. Beginning with A = 440 Hz, we get the frequencies listed in Table B.1.

(Ø Sound Examples) (Ø *Intervals_1.maxpat*, *Intervals_2.maxpat*)

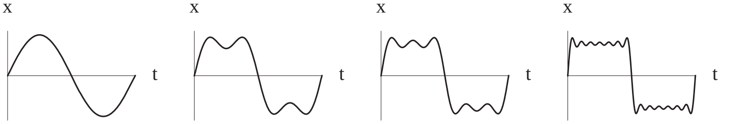
2.3.2 Timbre

Timbre is considerably more difficult to define than pitch or loudness. That is because timbre is determined by physical attributes which cannot be measured with a simple scale, as is the case for pitch and loudness. By the same token, there is no vocabulary in everyday language for timbre; depending on the sound in question, words for timbre can be borrowed from the visual world, like shrill or dull, or words with tactile associations are used, like rough or smooth.

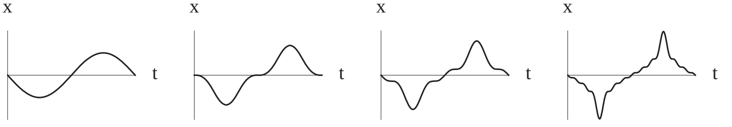
Further Reading: Daniel Muzzulini, Genealogie der Klangfarbe *[77]*.

2.3.2.1 Periodic Vibration

We saw in Chapter 2.2.2 that every strictly periodic oscillation can be described by its spectrum. In Sound Example 2-3-2a, overtones are gradually added to a fundamental. As each overtone enters, one first hears it as a separate sound, but it soon blends with the fundamental, enriching the tone's timbre. In Sound Example 2-3-2b, only the odd overtones are added, so that the result gradually approaches a square wave (2.2.2.5). The following figure shows the waveform of this example in four snapshots (fundamental alone, fundamental and third partial, fundamental, third and fifth partial, and fundamental with third, fifth, ninth, 11th and 13th partials).



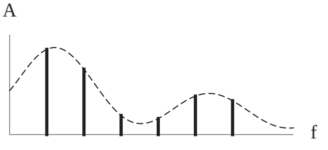
The next Sound Example shows that the phase constant hardly plays an important role for the timbre. The sound consists of the same odd partials at equal relative strength as before, but here every other partial has negative amplitude. Subtracting a partial is equivalent to adding that partial with a phase constant of p: $\sin(a + p) = -\sin(a)$ (Sound Example 2-3-2c). The four figures below show the same combination of partials as above.

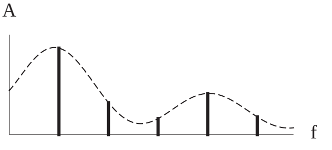


As the example shows, it is difficult to make predictions about how a given waveform will sound. Often it is not even possible to know whether the sounds corresponding to two different waveforms will have the same timbre. That is why sounds are usually described by their spectra and not by their waveforms. But even the spectrum permits one to draw only fairly general conclusions about a sound's timbre. Sounds with many partials are in principle brighter and brilliant, sounds with few partials are flat and dull, sounds with only odd partials are hollow and have a nasal character, etc. (Ø *Timbre_and_Spectrum.maxpat*)

2.3.2.2 Formants

Various studies indicate, especially for the vowels of speech, that timbre is not determined by the relationship of the strengths of individual partials, but rather by so-called *formants* (B2), regions of frequency in which the partials are particularly strong. We notice especially that slowing down or speeding up the playback of a recording of the human voice not only changes the voice's pitch but also its timbre, although the spectral relationships are the same as for the original recording. In Sound Example 2-3-2d, a recording of the sung vowel “a” is played back first sped up and then slowed down, both times by a factor of 1.5, transposing it up and down respectively by a fifth. Conversely, the analysis of a vowel sung at various pitches shows the regions of strong partials essentially fixed at the same frequencies, regardless of the fundamental. This means that the relationships between the strengths of the partials differ at every pitch. If one draws a line connecting the amplitudes of the partials, one obtains an envelope that for a given timbre remains about the same for various fundamentals. The amplitudes of the partials in the first sound shown below have the relationships .85 : .65 : .19 : .16 : .38 ..., those in the second .87 : .32 : .16 : .41 : .17.

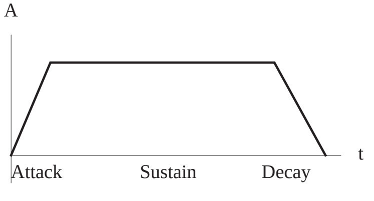


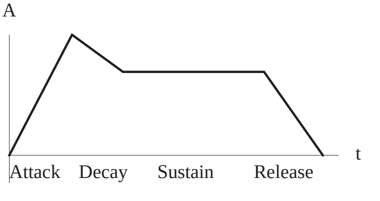


Sounds having the same spectral envelope arise from vibrations radiated by the same resonators. The shape of the spectral envelope reproduces the resonances of the sounding body. It is often possible to indicate which formant was produced by a particular part of the resonator. Our audition tries to draw conclusions about the object producing a sound from the sound's timbre. (Ø *Formants.maxpat*)

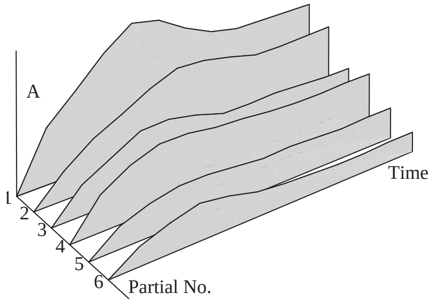
2.3.2.3 Spectra of Natural Sounds

The properties of natural sounds are usually not constant, but rather are in constant flux. Even in sounds that seem to be held without change, loudness, frequency and spectrum change slightly. The changes are especially noticeable during the sound's attack (which can be between a few milliseconds and about 0.2 second long) and its release. Normally we can identify sounds within a fraction of a second, because the attack contains considerable information about how the sound was produced. Even the simplest models for the synthesis of instrumental sounds take into account the physical behavior of the sounding medium by distinguishing for amplitude and spectrum, at a minimum, the phases attack, sustain and decay.

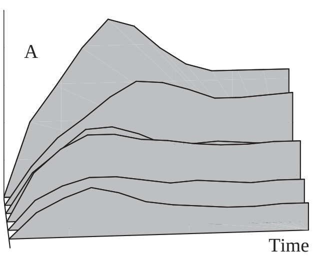


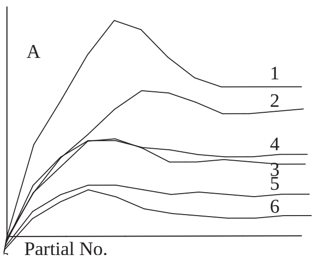


An evolving spectrum can be shown by displaying spectral snapshots made at regular intervals so that a three-dimensional image results, graphing the change in the amplitudes of the partials against time. The following figure shows typical behavior of the spectrum during the attack phase (t = 0 to t = .12).

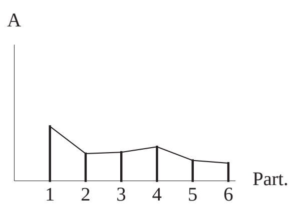


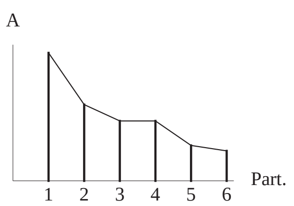
If one looks at this figure perpendicularly to the time-amplitude plane, one has a two-dimensional representation:





A temporal cross-section gives the spectrum at a specific moment (below t = .02 and t = .05).





If one looks at the original figure from above, one has a spectrogram (or sonogram), in which the amplitudes of the of the partials can be estimated by the width or color of the lines.
(Ø figure from above)

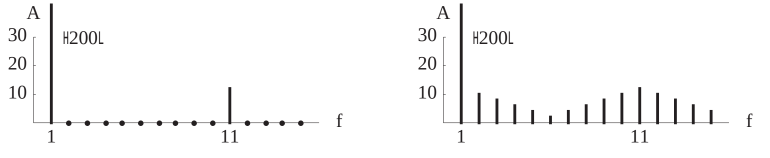
2.3.2.4 Non-Harmonic Spectra

Even the simplest bodies, like strings and air columns, produce sounds whose spectra are only approximately harmonic. For example, the frequencies of the upper partials of a loudly played string are somewhat high because of the string's stiffness. In pipes, the frequencies of the partials are not quite in integral relationships to each other because the length of the actually vibrating air column is not quite the same as the length of the pipe itself and is not quite the same for all partials. Objects in which vibrations can propagate in any direction, like plates and bells, have natural resonances whose frequencies can be in arbitrary proportions to each other.

2.3.2.5 Fusion

Whether or not sine waves fuse into one sound, depends on many factors. Although the audition essentially makes a Fourier analysis of the sounds it hears, as a rule it cannot perceive the individual partials of natural sounds. On the other hand, even the smallest differences in temporal patterns of sounds, or the smallest distance between sound sources, suffice to let us distinguish very similar sounds, for instance two violins playing in unison. Therefore, in the real world, we rarely have to ask which tones belong to which sounds. In electroacoustic music, especially when the sounds are relatively static, the situation is not always so clear.

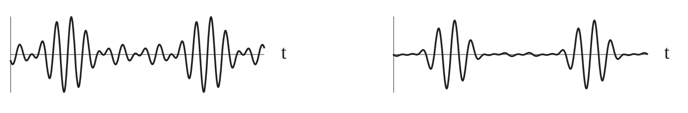
The more isolated a partial is, the more likely it will be heard as a separate tone. In Sound Example 2-3-2e, we first hear the fundamental with amplitude 200 and then the eleventh partial with amplitude 12. Then other partials enter corresponding to the pattern in the figure below right. The eleventh partial only fuses with the fundamental when it becomes part of an amplitude envelope.



Normally the partials occur simultaneously, and changes of timbre are caused by changes in the envelope and not by changes in the strength of individual partials. That is why in Sound Example 2-3-2c in Chapter 2.3.2.1 the new tones are heard separately and only fuse when other new tones enter. If one interrupts playback for a moment and then resumes it, all the tones sound simultaneously, and it is considerably more difficult to hear the most recent partial. (See Chapter 5.2.1.3 for further examples.) (Ø *Timbre_and_Spectrum.maxpat*)

2.3.2.6 Missing Fundamentals and Residue Pitch

When sine waves sound whose frequencies are multiples of a *missing fundamental* of frequency *f*, we perceive a tone with the frequency *f*, the so-called *residue pitch*, and we perceive the tones which actually sound as its overtones. In Sound Example 2-3-2f we first hear four equally strong tones with the frequencies 1600, 1800, 2000 and 2200 Hz, and then six tones with the frequencies 1200, 1400, 1600, 1800, 2000 and 2200 Hz and the amplitudes 1 : 2 : 3 : 3 : 2 : 1. The following figure shows the waveforms of the two sounds. It is easy to see a period in both waveforms which corresponds to the least common multiple of the individual period durations.



In the example above, it is difficult to decide whether the tone one hears is a phantom fundamental or a difference tone generated by the partials, because the frequency of the difference tone of two successive partials is also 200 Hz. Important features of combination tones are missing, however. In the first place, the fundamental can be heard even at low amplitudes, and in the second, there is no sensation of slight pressure in the ear (2.3.4.3). In Sound Example 2-3-2g we hear three tones with glissando (1600-1630 Hz, 1800-1830 Hz, and 2000-2030 Hz). The difference tones of the successive partials remain constant (200 Hz), while the missing fundamental rises slightly (from 200 to approximately 204 Hz), even though the frequencies are no longer integer multiples of the fundamental. If one listens to the example softly, one hears principally the slightly rising missing fundamental. If one increases the volume, the combination tones get louder and may even beat with the missing fundamental (see [20]). (Ø *Timbre_and_Spectrum.maxpat*)

2.3.3 Loudness

2.3.3.1 Sound Power and Sound Intensity

Sound waves carry energy. The total energy radiated from a sound source per second is the source's *sound power* or *acoustic power*. It is measured in *watts* (W). The sound power of speaking is about .00001 W, that of a violin playing very loudly about .001 W and that of a grand piano playing loudly about 2 W. This relatively small amount of energy is dispersed in space so that very little of it actually reaches the listener's ear. The flow of sound power through a surface perpendicular to the flow is *sound intensity*. It is measured in watts per square meter, or more usually in watts per square centimeter. Even as small a sound intensity as 10^{-16} W/cm^2 can elicit in humans an auditory response. The threshold of pain is 10^{13} times higher at .001 W/cm^2.
Sound intensity is proportional to the square of the sound's amplitude. Doubling a sound's amplitude increases its sound intensity by a factor of four. The energy per surface (the sound intensity) carried by a spherical wave decreases inversely proportionally to the square of the distance traveled, because the sphere's surface increases with the square of its radius. Since the energy is proportional to the square of the amplitude, the amplitude decreases inversely proportionally to distance.

2.3.3.2 Decibels

When speaking about sound intensity, one is usually interested in comparing values. Because the comparisons can involve both very large and very small numbers, one uses a logarithmic unit of measure, the *decibel* (dB). The decibel measurement of the relationship between a given sound intensity J and a reference value J_0 is called the *sound intensity level* L and is derived from the following formula:

$$\boxed{L = 10 \log_{10} \left(\frac{J}{J_0}\right) \text{dB}}$$

For the proportion J:J_0 = 1000 we get L = 10·log$_{10}$(1000) dB = 10·log$_{10}$(10^3) dB = 30 dB. For the proportion J:J_0 = 1/1000 we get L = 10·log$_{10}$(1/1000) dB = 10·log$_{10}$(10^{-3}) dB = –30 dB.

J ê J_0	...	1ê1000	1ê1000	1ê100	1ê10	1	10	100	1000	1000	...
L in dB	...	-40	-30	-20	-10	0	10	20	30	40	...

Doubling the sound intensity raises the sound intensity level by 10·log$_{10}$(2) dB = 3.0103... dB. Since the sound intensity is proportional to the square of the amplitude, a doubling of the amplitude raises the sound intensity level by 10·log$_{10}$(2^2) dB = 10·2·log$_{10}$(2) dB = 6.02... dB.

$$\boxed{L = 20 \log_{10} \left(\frac{A}{A_0}\right) \text{dB}}$$