FM Theory & Applications

By Musicians for Musicians

by

Dr. John Chowning
and David Bristow
# CONTENTS

FOREWORD ........................................................................................................ V
1. AN INTRODUCTION ....................................................................................... 7
2. SOME BASIC IDEAS ..................................................................................... 17
3. WHAT IS FM? ............................................................................................... 41
4. SIMPLE FM — The Theory ........................................................................ 55
5. SIMPLE FM — The Practice ....................................................................... 83
6. COMPLEX FM ............................................................................................. 113
7. APPLICATIONS ........................................................................................... 139

APPENDIX 1 — Logarithmic Representation and "Pitch Frequency" .............. 160
APPENDIX 2 — "X" Synth Comparisons by Index vs. Op. Output Level ............ 165
APPENDIX 3 — Bessel Functions Graphs .......................................................... 170
APPENDIX 4 — Bessel Function Tables ............................................................ 179
APPENDIX 5 — A Short Bibliography ............................................................... 187
APPENDIX 6 — A Glossary of Terms ............................................................... 190
APPENDIX 7 — The Sampling Rate of the DX7 ............................................. 194
ACKNOWLEDGEMENTS

This book grew out of an original intention to prepare a short tutorial on the fundamental theory of FM synthesis. As you can see, it didn't stay short, and we would like to thank those people and institutions who have supported us in achieving a complete text. At IRCAM in Paris, where Pierre Boulez has generously encouraged this work, our thanks to David Wessel for his continuous support and Stephen McAdams for his useful background information on acoustics; to Emmanuel Favreau and Dan Timis for their help with some of the figures, and to Robert Gross and Michelle Dell-Prane for their assistance in mastering the main-frame computer and not least in maintaining our computer link between IRCAM and CCRMA, Stanford in the latter stages of our work.

In the USA, thanks to Gary Leuenberger for his programming work and support and to the staff and colleagues at CCRMA, Stanford for their help over the years. Finally, our thanks to the research and development team at LM Division, Nippon Gakki Ltd. in Japan, who took time and time and trouble to answer our requests for information regarding the details of the "X" Series synthesizers, to Phil Brimble, who took care of the graphics, and to Jerry Uwins, Jean-Claude Risset, and Doug Keislar for the final editing of this work.
FOREWORD

John Chowning — the inventor of FM Sound Synthesis Technique — and David Bristow — the gifted DX Series FM Voice Maker — have collaborated to produce this extremely useful "theory meets practice" book.

Their emphasis is on putting FM theory to work, enabling musicians to obtain more expressively potent control over their FM instruments. One of the really wonderful things about this book is that otherwise inaccessible information from acoustics and psychoacoustics is clearly explained and applied to musical voicing practice.

David Wessel
IRCAM, Paris
CHAPTER 1

AN INTRODUCTION
The remarkable acceptance of the "X" Series musical instruments produced by Yamaha has led to an equally remarkable need for understanding how the theory of frequency modulation (FM) synthesis can be effectively used in the creation of musical sounds. In this book we have gleaned from the experience of many musicians and composers over many years what we believe to be the most effective means of understanding and use. There is much to learn that we can only touch upon in this text — about acoustics, psychoacoustics (auditory perception), basic mathematics and indeed music itself. But touch upon them we must, for if we are to gain a useful understanding of FM synthesis we must broaden our scope. The required additional effort will surely be worth it, since the knowledge acquired will be of more general use than simply understanding basic FM synthesis. We will rely on your diligence to read the book intelligently and seek additional sources, for which direction will be given.
In our belief that it is of utmost importance always to rely upon our ears for a practical and comprehensive understanding, we encourage you to read this book with an "X" Series synthesizer at hand. As a tool we have found it to be a most extraordinary aid in elucidating acoustic and psychoacoustic phenomena, as well as providing the basis of understanding FM. You will see that our strategy is rather different from other texts on FM in that we are not presenting a recipe book of "patches," but rather a means to understanding, creating, and modifying them. If we are to be positive in our approach to synthesis and FM, a consideration of the whole musical process is important; an awareness of the nature of the transmission and perception of musical sounds will be of great value when considering the details of FM sound synthesis. The picture below is to remind us that the total concept we call

Fig. 1.1
music covers a large range of topics, and the first chapters of this book are designed to lay a broad foundation upon which we can build our understanding of FM synthesis.

In talking about music, we are talking mainly of an aural experience, yet this book on sound synthesis can only be visual. That is the nature of books! However, some steps can be taken to augment the visual nature of the book by following the sound examples outlined for "X" Series instruments from Yamaha. All the exercises are easy to follow and relate directly to the text, so following them will help to strengthen the connection between sound and the written word or spectral diagram (see Fig. 1.2), as well as lead to some good "patches." Try to get in the habit of doing the sound exercises while following the text. There are some simple "X"-ample boxes here in the introduction with easily followed instructions. You may find it useful to store these examples, as they may be recalled during the text or compared with one another. Although parameter values are given in terms of the DX7 synthesizer, you will find conversion tables at the back of the book to allow you to make the same experiments on other "X" Series instruments. We assume some working knowledge of the synthesizer being used. It is our intention that, with the use of some simple (very simple) arithmetic, you will arrive at a clear understanding of the practical theoretical basis of FM as well as the musical basis. Furthermore, throughout the text there will be intended but useful redundancy, in other words, we will often explain the same thing in different ways.

![Fig. 1.2](image)

This shows a spectrum taken from a sound on the DX7 synthesizer. The vertical lines show at what frequency simple components exist (frequency is measured in cycles per second, or more commonly called Hertz – Hz). The height of the line represents its energy.
And what of "spectral diagram" mentioned before? The spectral diagram is important because it can provide a link between aural experience and a visual representation of the same, but what exactly does it mean? Well, it helps to describe a timbre, or tone colour, of a complex sound by showing lines which represent the energy (loudness) and the frequency (pitch) of different simple components within that timbre.

The idea of a timbre being composed of many different frequency components is not at all confined to the world of electronic music. You have probably talked of "overtones" or "partials" or "harmonics" when referring to the timbre or colour of a sound yourself, but it was a scientist, Joseph Fourier, who in the early nineteenth century gave a mathematical basis to this idea, one that musicians had been talking about for centuries and indeed still are. This is explained more fully in Chapter 2 but simply means that certain complex sounds, such as that from a violin or trumpet, could be thought of as the sum of a collection of much simpler tones, at frequencies which are whole number (or integer) multiples of the fundamental frequency. A spectrum then, as in Fig.1.2, shows clearly the actual frequencies at which partials exist for a given sound, and the energy of each. The "X"-ample below shows how to make the sound from which the spectrum in Fig.1.2. was obtained. (This particular "X"-ample is not an exercise in itself, but it represents the form that the exercises ("X"-amples) will take throughout the rest of the text. However, please note the settings of the function controls as these occur in many of the coming exercises).

**"X"-ample 1.1**

<table>
<thead>
<tr>
<th>FREQUENCY</th>
<th>OUTPUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>99</td>
</tr>
<tr>
<td>3.00</td>
<td>87</td>
</tr>
<tr>
<td>5.00</td>
<td>79</td>
</tr>
<tr>
<td>7.00</td>
<td>75</td>
</tr>
<tr>
<td>9.00</td>
<td>72</td>
</tr>
<tr>
<td>11.00</td>
<td>71</td>
</tr>
</tbody>
</table>

**INSTRUCTIONS:** Starting from the [VOICE INIT?] position....

Select algorithm 32, and set the operator values as shown above.
Voice Init does not affect the function controls. In later examples, we will be using the MOD. WHEEL to control index (don't worry about the word "index" — that will be explained in Chapter 4), therefore set the function controls on your instrument as follows; they will serve for further "X"-amples.
Poly/Mono — POLY
Portamento time – 0
Mod. Wheel Range — 99
Pitch — OFF
Amplitude — OFF
**EG. BIAS — ON**
Foot Control, Breath Control, After Touch — all OFF for 0.
How do we interpret this "visual" representation of a sound? Well, in the example shown (Fig. 1.2), the line on the left represents the fundamental or the first harmonic. To dispel confusion each of the components, including the fundamental, is most generally called a "partial". Collectively they form a spectrum. In the special case where the partials fall in the "natural" harmonic series - in other words they are all whole number (integer) multiples of the fundamental - they are called "harmonics", with the fundamental being harmonic number one. Sometimes the word "overtones" is used, and here beware of confusion because the first overtone (bearing in mind the implication of the word) will actually be the second harmonic. From here on, when we are looking at harmonic spectra, we shall refer to harmonics, not overtones. Partials which do not fall in the harmonic series will be specially referred to as inharmonic components of the spectrum. But, as we can see from the diagram (Fig. 1.2), these partials are in the harmonic series. The line to the right of the fundamental shows that there is the presence of the third harmonic, the next shows the fifth harmonic. We can notice by the relative height of the lines, which represent energy (or loudness), that the harmonics become progressively weaker. The note played from which this plot was made was A=880 cycles per second or 880 Hertz (Hz). Although it is the relationship between the harmonics which is of prime importance, nonetheless from this spectrum we can tell exactly the frequencies at which these harmonics occur. The first is at 880 Hz (the fundamental in this case); the next is at 3 times 880 or 2640 Hz; the next is at 5 times 880 or 4400 Hz, and so on. This brings us to a quite important point, and that is that above about 15,000 Hz, ears stop working! That is about as high as we can hear, whether we are listening for a fundamental or a high harmonic component in a "bright" sound. Therefore the spectral plot has point this marked on it.

![Diagram of a line spectrum with energy on the y-axis and frequency on the x-axis.](attachment:line_spectrum.png)

Fig. 1.3
A line spectrum.
Having introduced the idea of the spectrum, it must be said that there is more than one way of presenting it. In this book, there are basically three types of spectral diagram. The first is as in Fig. 1.2, which is a print from a real-time spectrum analyser which shows a Fourier analysis of the output of a DX7. Sometimes, for precision and clarity, we will want to represent the spectrum in a second, more exact way — this is done with a line spectrum (see Fig. 1.3). Notice that there is some space left for negative frequency and negative energy (concepts hard to imagine, but useful in understanding FM, as you will see in chapter 4).

There is one other type of spectral image which can be useful in helping to visualise sounds and that is the three-dimensional plot (Fig. 1.4). Sometimes, we will want to see how a spectrum changes with time (and certainly most interesting sounds do) and for this purpose a “3D plot” can be used. It is a series of spectra (or traces) shown one behind the other, and slightly offset in order to see them better, each one showing the change in the spectrum as the sound develops. The sound change in this diagram, for instance, is brought about by moving the DX7 mod. wheel from min. to max. in about two seconds, and letting it control the output of operator two (that’s changing the modulation index, but more of that later!). If you would like to hear the sound represented in this plot, then follow the instructions in “X” -ample 1.2. The “how and why” all of these partials are produced with only two operators is, of course, the purpose of this book, and will be revealed in the following chapters.
**“X”-ample 1.2**

<table>
<thead>
<tr>
<th>FREQUENCY</th>
<th>OUTPUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>op 1</td>
<td>1.00</td>
</tr>
<tr>
<td>op 2</td>
<td>1.00</td>
</tr>
<tr>
<td>op 3</td>
<td>1.00</td>
</tr>
<tr>
<td>op 4</td>
<td>1.00</td>
</tr>
<tr>
<td>op 5</td>
<td>1.00</td>
</tr>
<tr>
<td>op 6</td>
<td>1.00</td>
</tr>
</tbody>
</table>

**INSTRUCTIONS:** Starting from the **VOICE INIT?** position.

Select algorithm 1 and set the parameters as shown, but this time control the output of operator 2, with the modulation wheel. Make the AMS for operator 2, equal 3, and the function settings as in “X”-ample 1.1. A movement of the mod. wheel from min. to max. in about 2 seconds produces the sound from which the spectral plot in Fig. 1.4 was produced.

So far, the very familiar word “waveform” has not appeared; yet, if you have had any previous experience working with synthesizers, it has probably been a basic part of “sound-talk” — rich waveforms, sawtooth waves, sine waves, square waves and so on. The problem with waveforms is that while they are very descriptive of the way air pressure or voltages change with time, they are not very descriptive of the actual timbre of a sound. Look at Figs. 1.5 and 1.6. If you are familiar with analogue synthesis you can probably imagine the sound which gives the first waveform (Fig.1.5), but can you imagine the second (Fig.1.6)?

![Fig. 1.5](image)

![Fig. 1.6](image)
Clearly the waveforms are both very different, but as a visual representation they really do not give us much immediate information about the nature of the sound which produced them. Actually, both of these waveforms are properly represented by the spectrum shown in Fig. 1.2, in other words, taken from the same sound on the DX7. One waveform you recognise (as a square wave), but the other you do not. And that is why we do not normally use the waveform when describing a sound. The spectrum on the other hand, which is unique to any one sound, gives audible information about that sound, and is our most important visual aid or descriptive tool. An appreciation of what a waveform is, however, is important when beginning to read about the basic ideas covered in Chapter 2.

There remains just one other type of visual picture which is used in the text, and that is the “graph”. Probably you are quite familiar with how a graph works, but in case you are not, read on .... A graph is merely a way of showing how two different things or variables change in relation to each other. Look at the diagram below, Fig. 1.7.

A graph showing the relationship between two variables $x$ and $y$. A value of 4 for $y$ can be seen to be equivalent to a value of 8.5 for $x$.

By choosing a certain value on the vertical line, then following across horizontally until we meet the curve of the graph, we can then read down vertically to a corresponding value on the horizontal line, or vice versa. The horizontal line is commonly called the $x$-axis, and the vertical line, the $y$-axis. So by using a graph, for any given value of one variable we can determine a corresponding value of the other — for example, modulation index is a “function” of modulator output level on a synthesizer. That means given a graph of this function, we could look up output level on one axis and read modulation index off of the other. When studying FM, this will be a very useful function and you will find this graph in detail on page 54.
While the mathematics used in this book are actually very simple (we require only that you can add, subtract, multiply and divide), there are one or two terms which you may have forgotten from your schooldays . . . .

A "power" or "exponent" is the amount of times a number is multiplied by itself — that is to say,

\[ 2^3 \]

or two to the power three, means \( 2 \times 2 \times 2 \).

The words numerator or denominator refer to the parts of a fraction . . . .

\[ \frac{12}{7} \]

\[ \text{numerator} \]

\[ \text{denominator} \]

Sometimes, subscripts are used when one particular item from a class of items is referred to; this is rather like house numbers and street names. So the term \( J_2 \) means that in a class of objects which are collectively labelled by \( J \), we want to consider object number, or the value of object number two. Subscripts are a convenient way of keeping and referring to lists, so that if, for example, we want to consider all the items in a class, or all the houses in a street, from zero right through to a number that we haven't decided yet (let's call that undecided number "n"), we can express this rather cumbersome sentence in the simple mathematical expression . . . .

\[ J_0, 1, 2, 3, 4, \ldots, n \]

There is one non-mathematical convention that we should establish here in the introduction, and that is "what C is middle C?". Well, in frequency terms, there is no problem — it is at 262 cycles per second. But notes are also given numbers, and conventionally on a grand piano middle C is C4 (that is the fourth C counting up from the bottom of the keyboard), whereas on our MIDI system "X" Series synthesizer C3 is found to be middle C (as this is more convenient with a shorter scale keyboard). As we are dealing with synthesizers in this book, we shall stick to the C3-middle C standard, but it is as well to be aware of the classical convention.

Occasionally in the text you will find "rules" in clearly marked boxes. These highlight some of the most useful and direct aspects of FM theory and should prove helpful guides to programming. Where these guides are perhaps less precise but nonetheless useful, they are called "Hints". Here is the first . . . .
CHAPTER 2

SOME BASIC IDEAS
The new digital technology allows an efficient implementation of a relatively new synthesis technique, Frequency Modulation Synthesis (FM). One of the advantages of FM synthesis is that with a few number of elemental units (oscillators or operators) an extraordinarily large number of different sounds can be produced. One might say that the timbral space is large while the computational space is small (all digital synthesizers can be considered as highly optimized computers for the computation of sound waveforms). But ... some knowledge in addition to musical knowledge is helpful in the effective use of FM. Some of this knowledge, such as acoustics and psychoacoustics, is interesting no matter what synthesis technique, or indeed instrument, is being used, and some is quite specific to FM. In this section we will present some basic acoustical ideas that are of interest in themselves and that will serve us well in gaining some insight into the workings of FM.

Acoustics is a discipline in itself and has a large body of knowledge which we cannot hope to explore here. It is concerned with a variety of auditory systems and receptors, vibrating mechanisms, reverberant spaces and transmission media. We will limit our discussion to the human ear and a loudspeaker in a room filled with air.

**Pressure Waves and Periodicity**

In as much as our "tools" include an "X" Series synthesizer and an amplification system consisting of an amplifier and a loudspeaker, we can make some connections rather easily between some elementary concepts of acoustics, some trigonometry — a simple form of mathematics dating from antiquity — and some basic sounds from the synthesizer.

When sounding any tone of any of the presets of the synthesizer we know that the loudspeaker cone is set in motion and that the nature of this motion has to do with the pitch, loudness and the quality of the sound which we hear. In fact, we can imagine slowing time to the point where we can trace this motion of the cone of the loudspeaker as it moves in and out around its rest point. We can think of this pattern of motion such that when it moves in an outward direction we can say it is positive and when in an inward direction it is negative. If we were able also to look at the pattern of motion of the ear drum in the presence of the loudspeaker, we would notice that it would be very similar except that where the cone moves "outward" the ear may move "inward" and the units of displacement would be very much smaller.

Knowing that the coupling of the cone and the ear is by means of air as the medium, we can infer with some certainty that the air in some way must transmit the same pattern, which of course it does. When the cone moves outward it causes a compression of air particles, and when it moves inward it causes a rarefaction, or a decompression, of particles. We can think of these particles as being elastic in that they tend to be spread equally throughout the room — always equidistant from one another. Therefore in a region that is compressed, the particles will move away from the point of greatest compression, whereas in a rarefied region they will move toward the point of greatest rarefaction. But since air particles have mass and
The pattern of motion of a loudspeaker cone as it is displaced in time in proportion to a changing voltage applied to it. Let us say that the cone moves a maximum distance of 0.5 mm in both a positive and negative direction. The motion is not perfectly smooth, however, as we can see in a, where the direction of movement changes as the cone approaches the maximum and minimum. The motion of the cone causes compression and rarefaction of air particles (b), which causes a pressure wave which travels in the direction of the ear, whose ear drum is set in motion in response to the pressure wave (c).
therefore momentum, they tend to “bounce” rather like rubber balls — their motion does not stop at the instant they are equidistant from one another but rather continues on, creating other regions of compression and rarefaction in a general direction away from the speaker-cone, but always losing a bit of their energy until once again they are at rest. Because this wave action is a wave of instantaneous changes in pressure, it is called a pressure wave.

An oscilloscope attached at some point between the loudspeaker and the amplifier speaker terminals would show us the changing voltage which, when applied to the coil attached to the cone of the loudspeaker, induces a varying-strength magnetic field which causes the cone to be attracted and repelled, causing its motion relative to the speaker frame. The voltage, then, must be proportional to the motion. The pattern of change in air pressure is in turn proportional to the motion of the loudspeaker cone, and so must be the motion of the ear drum as it responds to this continuous change in pressure. These patterns of generation, transmission, and reception are shown in Fig. 2.1.

Were we to look at a large number of such acoustic pressure waves we would notice that there are, in general, two types which can be fairly easily discerned. One type consists of patterns of motion or voltage which are largely repetitive as in Fig. 2.1.
and the other type consists of patterns where there are no apparent repetitions. The first is called periodic and the second aperiodic. Further examples of periodic and aperiodic pressure waves are shown in Fig. 2.2 and Fig. 2.3. Remember that these examples of waveforms are simply graphical representations — as you trace along the line of the graph you can see how pressure (in this case the vertical dimension) increases and decreases as time progresses. Time is represented in the horizontal axis.

There are two things we will note about these examples. First, those of Fig. 2.2 all produce a strong sense of pitch, while those of Fig. 2.3 are weak or ambiguous. Second, they are all plots of waves which were produced by an “X” Series synthesizer in which there is stored only one wave form, a sine wave. In fact, Fig. 2.2a is four periods or repetitions of a sine wave. The other five waveforms in Fig. 2.2 and Fig. 2.3 were all constructed from “mixes” of sine waves at differing periods and amounts.

How is it that such different shapes can be formed from a sine wave and what is it about a sine wave that makes it a fundamental unit of acoustics? In order to answer these questions we must recall some observations about right-angled triangles first made by Pythagoras in the 6th century BC.

Fig. 2.3
Three aperiodic pressure waves where there is no discernible pattern
Again, vertical axis is amplitude and horizontal is time.
Sine Wave

First we will see how a sine wave relates to a right-angled triangle, and then examine the properties of a sine wave. This may seem like an unnecessary excursion into mathematics, but in fact it is very simple and relates directly to our understanding of terms like "phase", and shows why a sine wave is such a powerful tool in both the analysis of complex sound and its synthesis.

Recall from trigonometry that a right-angled triangle has one angle which is 90° and that the side of the triangle opposite the right angle is, by convention, called the hypotenuse. The sine of either of the other two is defined to be the ratio of the length of the side opposite the angle and the length of the hypotenuse. As shown in Fig. 2.4 the sine of the angle theta (θ) is equal to side a, (the opposite) divided by side c (the hypotenuse) or:

$$\sin \theta = \frac{a}{c}$$

and the cosine of an angle is the ratio of the adjacent side and the hypotenuse or, again looking at Fig. 2.4, sides b and c or:-

$$\cos \theta = \frac{b}{c}$$

![Right-angled triangle diagram](image)

Fig. 2.4
A right-angled triangle.
These relationships between the lengths of the sides of a right-angled triangle are, of course, constant regardless of the actual size of the triangle. In other words, if by some means we had calculated that the value for $\sin 30^\circ$ is 0.5, then we could solve the following non-musical problem!

For example, let's suppose that we are map makers and we need to know the precise distance between two points which we cannot measure directly (as shown in Fig. 2.5). We want to know the distance between points A and B but cannot measure directly because of an intervening body of water. We can, however, locate a point C from which we can see points A and B and that we can determine to be at an angle of $90^\circ$ to B in relation to A. We can only measure directly the distance C to B, but we can see the other points. With only a protractor to estimate angles, a fairly reliable gait to step off the distance between C and B, and a table of values for the sine of angles from $0^\circ$ to $90^\circ$, we could find an approximation of the distance AB.

![Diagram showing right-angled triangle with labels for hypotenuse, opposite side, adjacent side, angle, right angle, and points A, B, C.]

If the distance BC has been paced off to be 100 metres and angle $\theta$ has been gauged to be $30^\circ$, then we can determine the distance AB if we know that $\sin 30^\circ = 0.5$. Since

$$\sin \theta = \frac{a}{c} \text{ (see fig. 2.4) then } c = \frac{a}{\sin \theta} \text{ or, substituting our values, } c = \frac{100}{0.5} = 200 \text{ metres.}$$

But back to the original assumption — How DO we know that $\sin 30^\circ$ is equal to 0.5? Let us figure out how to calculate a table of sine values. While the means we will use
are a little unorthodox, it is not difficult and will lead us directly to the sine wave. And that is important, because we use the term frequently in our discussions of music and synthesis, perhaps without fully understanding exactly what it actually means. The tools we need are a protractor to measure angles and a ruler to measure the lengths of lines.

With our protractor we first draw a circle and divide it into four quadrants I, II, III, IV, the co-ordinate system as shown in Fig. 2.6. The radius of the circle is equal to 1 and we divide the y axis into units 0.1, positive above the origin where the x axis and y axis cross and negative below the origin. Now with the protractor, mark off the circle at 10° increments. We can now approximate the sine of these angles in the following way.

Knowing that:-

\[
\sin \theta = \frac{\text{opposite}}{\text{hypotenuse}} \quad \text{and in our unit circle}
\]

\[
\text{hypotenuse} = 1
\]

\[
\sin \theta = \frac{\text{opposite}}{1} = a
\]

We simply draw a perpendicular line from the angle in question to the x axis, which becomes our opposite side, a. For example, in Fig. 2.6 if we draw a line from 40° to the x axis we can measure the length of that line to approximately 0.64, therefore:-

\[
\sin 40° = \frac{\text{opposite}}{\text{hypotenuse}} = \frac{0.64}{1} = 0.64
\]

Because the hypotenuse (the radius of the circle) is equal to 1 for all of the right-angled triangles formed from the angles, we can approximate all of the sines by simply measuring the successive perpendiculrars.

\[
\begin{align*}
\sin 10° & = 0.17 \\
\sin 20° & = 0.34 \\
\sin 30° & = 0.50 \\
\sin 40° & = 0.64 \\
\sin 50° & = 0.77 \\
\sin 60° & = 0.87 \\
\sin 70° & = 0.94 \\
\sin 80° & = 0.98 \\
\sin 90° & = 1.00*
\end{align*}
\]

*Here we must accept some mathematical abstraction for, as we approach 90°, the perpendicular approaches the length of the hypotenuse (the radius). If we were calculating very small increments of angle, a right-angled triangle would exist even if the angle \( \theta = 89.999° \). But at \( \theta = 90° \) our right triangle collapses to a line and we have two rather abstract angles both of which equal 90°.
A circle whose radius = 1 can be divided into a number of sections (in this case every 10°) from which a succession of right-angled triangles can be formed. For every triangle the radius of the circle is considered its hypotenuse thus allowing a straightforward visualization of the function \( \sin \theta \) which equals the height of the side \( a \)

\[
\sin \theta = \frac{\text{opposite}}{\text{hypotenuse}} = \frac{a}{1}
\]
With simple tools we have been able to approximate the sine function and therefore solve our map-making problem. We were able to do this through the use of one quadrant of a circle whose radius is equal to one. Such a circle is sometimes called a "unit circle" and we see in Fig. 2.6 that there are 270° which remain (90° in each of quadrants II, III, and IV). We can move on from our 90° angle then to 100°, 110°, 120°, etc., always incrementing the angle by 10°. To maintain our right-angled triangle, we let the adjacent side of the angle become negative x and draw the perpendicular as we did before. Obviously the height of line a measured on the y axis will decrease in the second quadrant with the same values of the first quadrant except in reverse order. Thus:

\[
\begin{align*}
\sin 100° &= 0.98 \\
\sin 110° &= 0.94 \\
\sin 120° &= 0.87 \\
\sin 130° &= 0.77 \\
\sin 140° &= 0.64 \\
\sin 150° &= 0.50 \\
\sin 160° &= 0.34 \\
\sin 170° &= 0.17 \\
\sin 180° &= 0.0*
\end{align*}
\]

Continuing beyond 180° we notice that the values for \(\sin \theta\) are the same as they were in quadrant I, except for the fact that the side a opposite the angle \(\theta\) is now negative because below the origin, y is negative. For example (see Fig. 2.6),

\[\sin 240° = -0.87\]

The continuation into quadrant IV will be the negative of quadrant II.

Now we will look at all of these values in the form of a graph, Fig. 2.7, where on the horizontal axis we mark the degrees for successive values of angle \(\theta\) and on the vertical axis we plot \(\sin \theta\), the values of which we have just determined from our unit circle. Thus we have sine of the angle theta as a function of theta.

![Diagram of a sine wave with labeled degrees](image)

Fig. 2.7

A plot of \(\sin \theta\) as a function of \(\theta\) reveals the familiar shape of a sine wave.

* The last of course is the reverse of the abstraction at 90°; only this time it is 0.
Starting at $0^\circ$ again we could go through this entire process again and figure the values of \textbf{cosine} of theta rather than sine. As we said previously:

$$\cos \theta = \frac{\text{adjacent}}{\text{hypotenuse}} = \frac{b}{c}$$  \hspace{1cm} (see Fig. 2.4)

We can see in our unit circle, Fig. 2.6, that while sine of theta is 0, cosine of theta will be 1; that is, when theta is 0, then both the hypotenuse and the adjacent side \(c\) are equal to 1, as must be the ratio. (Remember we are looking at line \(b\) now).

![Fig. 2.8](image)

A plot $\cos \theta$ as a function of $\theta$.

These two functions, sine and cosine, are related in another way. In Fig. 2.9 we see several periods of a sine function from which is extracted a shape that is exactly that of a cosine. That is, by plotting a sine function beginning at $90^\circ$ and passing through $360^\circ$ and ending at $90^\circ$, we will have plotted the same as the cosine of theta. Therefore:

$$\cos \theta = \sin (\theta + 90^\circ)$$

which leads us to the important concept of \textbf{Phase}.

![Fig. 2.9](image)

Several periods of a sine function, where one period beginning at a phase of $90^\circ$ can be seen to be the same as a cosine function.
We can also make reference to the phase of a function derived from uniform circular motion, as we now know a sine function to be, where phase is a point to which the rotation (angle) has advanced relative to some reference point. Therefore, we can say that a cosine function is the same as a sine function beginning at a phase angle of 90°, where the reference point is assumed to be 0°.

What is it about these functions that are so special in their relation to acoustical theory? It has long been known that many natural sounds, including the steady state or sustain part of most musical instrument sounds, are more or less periodic in that the pattern of pressure variation is largely constant (see Fig. 2.2). Near the beginning of the 19th century a French physicist/mathematician, Joseph Fourier, realised that any periodic function could be resolved into mixtures of sine and/or cosine functions which may differ in amplitude and period but whose periods are related by whole number (integer) multiples. This applies to periodic sound waves as well.

For example, the square wave, which is a periodic function having a characteristic timbre familiar to all those who have acquaintance with analogue synthesizers, can be resolved into an infinite number of sine functions (this is sometimes called a Fourier transformation or analysis), whose frequencies are related by all being multiples of the odd integers (1, 3, 5, 7, etc.), and whose amplitudes are related by the reciprocal of those integers (1/1, 1/3, 1/5, 1/7, etc). With the "X" Series synthesizer we can add six sine functions together and approximate a square wave. Set up "X"-ample 2.1 (this is the "X"-ample we used in the introduction, when explaining the exercise format and the spectral diagram).

```
**"X"-ample 2.1**

<table>
<thead>
<tr>
<th>FREQUENCY</th>
<th>OUTPUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>op 1</td>
<td>1.00</td>
</tr>
<tr>
<td>op 2</td>
<td>3.00</td>
</tr>
<tr>
<td>op 3</td>
<td>5.00</td>
</tr>
<tr>
<td>op 4</td>
<td>7.00</td>
</tr>
<tr>
<td>op 5</td>
<td>9.00</td>
</tr>
<tr>
<td>op 6</td>
<td>11.00</td>
</tr>
</tbody>
</table>

**INSTRUCTIONS:** Starting from the [VOICE INIT?] position....

Select algorithm 32. Set the values of the operators to those shown above, then play any note to hear the familiar "clarinet" type sound characteristic of the square wave.
```
Of course, the wave that is produced by these six operators will not be exactly a square wave — for that we need an infinite number, not just six. But in this book we will be dealing with FM synthesis, not additive synthesis which is what we are experimenting with now at this learning stage. We used some new terms however, frequency and amplitude, which we need to define. Looking at Fig. 2.11 on the next page, we see six patterns, all of which seem to be sine waves but which differ in the relative number of periods on the \( x \) axis and in their heights on the \( y \) axis. When we measure the height of the wave at the point where it is at its greatest distance from 0, we are measuring its *amplitude* as in Fig. 2.11a.

![Spectrum relating to "X"-ample 2.1. For note A440Hz (that's A just above middle C).](image)

Reset the values in "X"-ample 2.1, turn off all of the operators except op. 1, and press the key A (=440). What you hear is a wave whose pressure varies in a manner as shown in Fig. 2.11a at a frequency of 440 periods, or cycles, per second. That is "cycles" because it is 440 cycles around our unit circle. Now look at the spectrum shown in Fig. 2.10, turn off op. 1 and turn on op. 2. You now are hearing the second component of the square wave at frequency three times that of the first — at 1320 cycles/sec (commonly called Hertz and abbreviated Hz) — and at an amplitude of \( \frac{1}{3} \) that of the first, as seen in Fig. 2.11b. Because the frequency is greater, the length of each period must be shorter. Frequency, then, is inversely related to period. Therefore:

\[
\text{period} = \frac{1}{\text{frequency}}
\]

i.e. a wave whose frequency is 440Hz (cycles/sec) will have a period of \( \frac{1}{440} \) sec.

In listening to first op. 1 and then 2, we notice that as frequency is greater, the sensation of pitch is higher, and that as amplitude is lower, loudness is less. The physical properties of frequency and amplitude are indeed linked to the perceptual properties of pitch and loudness, sometimes in rather complicated ways. It is not always correct to assume that those terms which we use to describe the physical world of sound can be applied directly to the perceptual world which happens inside our heads and which we call music.
The output of each of the six operators with the frequency and output values indicated. Added together they constitute the first six components of a square wave. On the right one can see that with the addition of each component the wave becomes more square-like.
While looking at Fig. 2.11 a-f and always pressing the key for A (440Hz), listen to each of the operators in turn, 1 through 6. With each operator the frequency is greater and the amplitude less. The frequency coarse parameter, then, specifies an integer which multiplies whatever frequency is associated with the key being pressed. (See appendices for a table of note frequencies.) The operator, then, produces a sine wave at that frequency. In this case op. 2, having a freq. coarse of three, produces a sine tone at $3 \times 440 = 1320$Hz; op. 3 at $5 \times 440 = 2200$Hz, etc. (If we now press the key B, a whole step higher, the frequencies produced by each of the operators are integer multiples of 494Hz).

Return to A440 and listen to the sound as we add each operator, first op. 1, then op. 2 (you must repress the key every time to activate the additional operators), then op. 3, etc. The brightness of the tone increases as we add the higher frequency sine tones.

Frequencies which are related in this way by integers (that is to say they are products of whole numbers and not fractions) fall into what is called the harmonic series, and each frequency is called a harmonic. A harmonic is a sinusoidal vibration whose frequency is an integral (integer) multiple of a fundamental frequency. In the case above, for a fundamental of A (440Hz), we listened to the 1st, 3rd, 5th, 7th, 9th, and 11th harmonics. Whenever the term harmonic is used, then, it refers to a sinusoidal vibration or sine wave that is related to other sinusoidal vibrations by integral (whole number) multiples.

Before going on to our next subject, we can remind ourselves of the power of FM synthesis by trying another "X"-ample which produces the same sound as "X"-ample 2.1 but with the use of only three operators. Understanding how this is possible will become clear over the next chapters when we begin to study FM.

```
```

<table>
<thead>
<tr>
<th>FREQUENCY</th>
<th>OUTPUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>op 1</td>
<td>1.00</td>
</tr>
<tr>
<td>op 2</td>
<td>2.00</td>
</tr>
<tr>
<td>op 3</td>
<td>4.00</td>
</tr>
<tr>
<td>op 4</td>
<td>1.00</td>
</tr>
<tr>
<td>op 5</td>
<td>1.00</td>
</tr>
<tr>
<td>op 6</td>
<td>1.00</td>
</tr>
</tbody>
</table>

**INSTRUCTIONS:** Starting from the [VOICE INIT?] position.....

Select algorithm 3. Set the values of the operators to those shown above, then play any note to hear the familiar "clarinet" type sound characteristic of the square wave, this time produced by FM. Of course any algorithm which has a stack of three operators will serve for this example, as long as they have the values shown above. Store this sound and make a comparison with the sound made additively in "X"-ample 2.1.

31
Set up the following "X"-ample:

<table>
<thead>
<tr>
<th>FREQUENCY</th>
<th>OUTPUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>op 1</td>
<td>1.00</td>
</tr>
<tr>
<td>op 2</td>
<td>2.00</td>
</tr>
<tr>
<td>op 3</td>
<td>3.00</td>
</tr>
<tr>
<td>op 4</td>
<td>4.00</td>
</tr>
<tr>
<td>op 5</td>
<td>5.00</td>
</tr>
<tr>
<td>op 6</td>
<td>6.00</td>
</tr>
</tbody>
</table>

**INSTRUCTIONS:** Starting from the VOICE INIT position.....

Select algorithm 3.2. Set the values of the operators to those shown above, then turn off all operators and listen to each one successively.

Remember that the frequency coarse parameter multiplies the frequency associated with a key by an integer. Therefore what we have in this exercise, assuming that you play A (440Hz) just above middle C, is as follows:

<table>
<thead>
<tr>
<th>Freq. coarse parameter</th>
<th>freq. of key pressed</th>
<th>freq. of harmonic</th>
</tr>
</thead>
<tbody>
<tr>
<td>op 1 1 x 440</td>
<td>440 (harmonic No. 1, called the fundamental)</td>
<td></td>
</tr>
<tr>
<td>op 2 2 x 440</td>
<td>880 (harmonic No. 2)</td>
<td></td>
</tr>
<tr>
<td>op 3 3 x 440</td>
<td>1320 (harmonic No. 3)</td>
<td></td>
</tr>
<tr>
<td>op 4 4 x 440</td>
<td>1760 (harmonic No. 4)</td>
<td></td>
</tr>
<tr>
<td>op 5 5 x 440</td>
<td>2200 (harmonic No. 5)</td>
<td></td>
</tr>
<tr>
<td>op 6 6 x 440</td>
<td>2640 (harmonic No. 6)</td>
<td></td>
</tr>
</tbody>
</table>

Perhaps you have noticed that the change in pitch between harmonic No. 1 and harmonic No. 2 is an octave, whereas the change in pitch between harmonic No. 2 and No. 3 is only a fifth, yet the difference in both cases is 440Hz. In fact, you can notice that the difference between each successive harmonic is a smaller musical interval than the preceding one although the arithmetic interval (the number of cycles per second) is the same (440Hz in this case). What we have noticed is of fundamental importance in the understanding of musical perception, i.e. the perception
of constant musical intervals or pitch distance is not based upon constant differences in frequency. But what then?

Now we will make a modification to the previous "X"-ample and demonstrate to ourselves something important regarding the perception of pitch. We want to have an understanding of pitch and frequency, not only because it is interesting, but because we will learn about forms of graphic representation which will be helpful in our understanding of FM beginning in the next section. So, using "X"-ample 2.3, change only the freq. coarse values from 1, 2, 3, 4, 5, 6 to 1, 2, 4, 8, 16, respectively. Never mind 32 for the time being — this experiment can be restricted to 5 operators.

Now turn all the operators off except op. 1 and listen while pressing the key an octave below A440; that’s A220Hz just below middle C. Turn off op. 1 and turn on 2 and listen again. Always sounding the same key, listen to each of the remaining operators one by one. You may be surprised to hear a constant pitch distance of one octave as you progress through the operators. What, then, is the relationship of the frequencies that produced this constant pitch distance? The following paragraphs will help unravel the mystery which surrounds the idea of a "logarithmic" representation. Let’s start by looking at the table of frequencies in this case.

<table>
<thead>
<tr>
<th>Freq coarse parameter</th>
<th>freq of key pressed</th>
<th>freq of harmonic</th>
</tr>
</thead>
<tbody>
<tr>
<td>op 1 1</td>
<td>220</td>
<td>220 (harmonic No. 1)</td>
</tr>
<tr>
<td>op 2 2</td>
<td>220</td>
<td>440 (harmonic No. 2)</td>
</tr>
<tr>
<td>op 3 4</td>
<td>220</td>
<td>880 (harmonic No. 4)</td>
</tr>
<tr>
<td>op 4 8</td>
<td>220</td>
<td>1760 (harmonic No. 8)</td>
</tr>
<tr>
<td>op 5 16</td>
<td>220</td>
<td>3520 (harmonic No. 16)</td>
</tr>
</tbody>
</table>

We can see that, while we hear a constant pitch distance of an octave from operator to operator, the frequency distance always doubles. Or we can say that to change the pitch by a constant perceptual distance, the frequency must change by a constant factor (in this case of an octave, by a factor of 2). In Fig. 2.12 we see a line representing frequency in Hertz. The line is divided into equal units of frequency on which are marked the points which we have heard to be successive octaves.

Fig. 2.12

With frequency plotted linearly along the horizontal axis, we can see clearly that equal octaves are not represented by equal distances, or frequency intervals.
Because this representation of frequency does not preserve visually the equally perceived pitch distance, we should be aware of this when looking at our linear representations of FM spectra in the following chapters. When we think of frequency as it is perceived, it would be helpful to have a representation that is closer to how the ear works, and that is just what a logarithmic scale for frequency does. Fig. 2.13 shows two identical spectra, one with a linear frequency scale and one with a log frequency scale — offering a visual representation which is more similar to our perception of frequencies.

Fig. 2.13

Above are two representations of one typical DX7 sound. The first (a) represents the components close to a manner in which they are perceived; the second (b) maintains a constant frequency interval between the components and is perhaps more clear from a mathematical point of view.
Imagine stretching and squashing the graph in Fig. 2.12 as if it were printed on a rubber page so that the octaves were always represented by equal distances, and you have an idea of what a logarithmic representation is. But what exactly happens to the frequency scale if we do this, and how do we plot it? Well, it is not absolutely necessary to understand how to move from a linear representation to a logarithmic one, and indeed in this text for reasons of mathematical clarity, we shall be using linear frequency scales. However, it is an interesting and simple process which also reveals an understanding of frequencies in a musical scale, so a full explanation is given in appendix 1. There are two different ways, then, to plot frequency.

1) Logarithmically — which shows equal pitch intervals. 2) Linearly — which shows equal frequency intervals.

We have in Fig. 2.10 introduced another aspect of representation, one which has to do with loudness. We know intuitively that any frequency we hear has another quality besides pitch, and that is loudness; i.e., a single pitch can be heard at many different loudnesses, and conversely, a single loudness can be heard at many different pitches. In Fig. 2.10 we have shown on the vertical axis the output level of each of the operators of algorithm 32, such that we represent the six sinusoids’ frequencies on one axis and their levels on the other. When we plot frequency (either log or linear) on the x axis, the representation is called the frequency domain, whereas when we plot time on the x axis, as in Fig. 2.11, the representation is in the time domain. What we are not able to see in the frequency domain representation is the phase relationships of the sinusoids (for example, all of the components in Fig. 2.11 have an initial phase relationship of 0°, but in many cases, phase relationships between harmonics are only minimally perceptible, if at all).

The second and third periodic pressure waves of Fig. 2.2 are in fact made up of identical components in both frequency and output level. They differ in the initial phase relationships of the components. The waveforms are very different, yet they sound virtually identical. In this case the time domain representation does not necessarily reveal the odd harmonic or “square wave-lish” sound, while in the frequency domain, Fig. 2.10, we can see the relationship of odd harmonics independent of phase. There are other instances where the time domain is more revealing — amplitude envelopes, for example — and others where neither is an adequate representation. We are for the moment restricting the discussion to tones which are artificially simple, that is, tones which do not change in time. We want to understand these simple ideas before we consider the interesting but more complicated cases where the tones evolve in time. Frequency domain representations are often referred to as spectra. As with a prism we are able to see light broken up into a spectrum (singular of spectra) of light frequencies (colours); with sound we can see a tone broken up into a spectrum of audio frequencies (sinusoids). When we make reference to spectra, then, we mean frequency domain representations of sound, where frequencies are plotted on the horizontal axis and their levels (amplitude or intensity) on the vertical axis.
Loudness Perception and Intensity

As we have seen with pitch and frequency, our ears do not line right up with straightforward physical measures. Frequency components at a constant distance in Hertz are not perceived to have constant intervals in pitch, as we have demonstrated listening to the harmonic series. And so too with loudness. At the beginning of this section we noted that air particles, when excited by some vibrating source, alternate between states of compression and rarefaction. If the source continues its vibration at the same rate (frequency) but over a greater distance (for example, if the loudspeaker cone's "in-out" excursion increases), then the maximum compression and rarefaction of the pressure wave will also increase. The measure of the pressure at the instant when it is greatest, or when the air particles are most compressed, is called the peak amplitude. As was the case with frequency, changes in amplitude or pressure must be transformed into another scale in order to approach the way that the ear senses loudness. A common and useful measure is sound pressure level (SPL), or intensity, the units of which are decibels (dB). A decibel is 20 times the log of the ratio of two pressures (two pressures since level is relational).

$$SPL = 20 \times \log_{10} \left( \frac{P_1}{P_2} \right)$$

The output level of the "X" Series synthesizer is in a log scale, although not in decibels. The reason that it is not in a dB scale has to do with the way the arithmetic is most efficiently done in the hardware and the fact that of the three basic musical parameters, loudness, time and pitch, loudness is the least precisely perceived (doubling a note's time value or its frequency is much more perceptually apparent than doubling a note's loudness). The available computational power of the instrument was thought to be better used in other ways — the accurate determination of pitch, for example. In the explanation of FM which follows, we will especially rely upon the amplitude (which is a linear measurement) of frequency components rather than intensity (log).

Aliasing

We now want to demonstrate a rather surprising phenomenon which is utterly counter-intuitive, especially to those whose experience has been in the analogue domain. What will at first appear to be a deficiency of digital audio will finally be seen to be a useful attribute, especially in FM synthesis.

What "X"-ample 2.4 will demonstrate is a fundamental attribute of digital audio, whether synthesis, recording or sampling. Known as aliasing (described by the Nyquist Theorem), it states that no frequency can be reproduced that is greater than one half the sampling rate. The sampling rate is the number of samples (numbers or readings) per second used to form the sound wave. (The compact disc has a sampling rate of 44.1 kHz while the DX7's rate is just below 60kHz.) Furthermore, the theorem states that frequencies greater than one half the sampling rate will reflect about the half sampling rate. That is, if the sampling rate is 60kHz, the half sampling
rate is 30kHz. So a frequency introduced at 31kHz cannot be produced, but will reflect and thus be heard at 29kHz. A frequency of 35.5kHz will reflect at 24.5kHz, and a frequency of 59.9kHz at 100Hz (0.1kHz) – the half sampling rate minus the amount by which the frequency exceeds it; 30,000Hz – (59,900Hz – 30,000Hz).

**“X”-ample 2.4**

<table>
<thead>
<tr>
<th>FREQUENCY</th>
<th>OUTPUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>op 1</td>
<td>1.00</td>
</tr>
<tr>
<td>op 2</td>
<td>5.00</td>
</tr>
<tr>
<td>op 3</td>
<td></td>
</tr>
<tr>
<td>op 4</td>
<td></td>
</tr>
<tr>
<td>op 5</td>
<td></td>
</tr>
<tr>
<td>op 6</td>
<td></td>
</tr>
</tbody>
</table>

**INSTRUCTIONS:** Starting from the VOICE INIT? position.....

Select algorithm 1 and set up the values shown above. Now, note by note, play an ascending chromatic scale starting from the lowest C on the keyboard. What do you hear? The tone itself is rather rich in harmonics, and as you would expect, it rises in pitch as you ascend the scale. Then at about B3, that is, B above middle C, there appears to be a slight change to the sound, a roughness that was not previously present. Continuing to ascend the scale, at E3 or E4 there is a very noticeable change, where not only is there a change in tone quality, but also there seems to be an additional pitch present. All of the remaining notes above this point have similar noticeable perturbations of the original sound and, what is more, in a seemingly random manner. Why should this be?

In digital recording, the signal to be recorded must first be (low-pass) filtered to ensure that there are no frequencies higher than the half sampling rate, which would reflect and distort the image. Remember, we are not only considering fundamental pitches, but also those high frequency components in very bright sounds. In the “X”-ample above, the highest harmonics began to reflect around the half sampling rate at B3 and as the pitch ascended, so more and more harmonics were reflected. (Appendix 7 shows a calculation which allows us to estimate roughly the sampling rate of the DX7 from this simple observation, but you will need to know something about FM before it is clear — so read on!) Where we require stability in the tone, this effect is undesirable and can be eliminated by progressively reducing the bandwidth as the pitch is raised by means of key scaling, as we shall see in Chapter 5. Of course we could also make use of aliasing for special effects, where this sort of randomness is desirable.
In digital audio, no frequencies can be produced beyond the half sampling rate, $\frac{SR}{2}$. Any frequency above the limit will reflect, and be produced below it by the same amount, as shown in the two cases above. A spectrum is shown with some components above the half sampling rate — these are not produced, but reflections of these frequencies, indicated by the arrows, are produced instead.

**Oscillators and Operators**

There is one final topic that we should touch upon briefly before actually beginning some experiments with FM synthesis in the following chapter. For our purposes, it is sufficient to know that an Operator in an “X” Series digital synthesizer is equivalent to an oscillator in an analogue synthesizer. In the latter case, the oscillator produces a changing voltage according to some selected pattern such as sine, sawtooth, pulse, etc. An operator performs essentially the same tasks, but instead of a changing voltage it produces a series of changing numbers (samples) whose pattern is always that of a sine wave.

Knowing that most interesting sounds have many components, how is it that with only six operators we can generate and control considerably more than six harmonic or inharmonic components, when an operator alone can only produce a single pure tone? That is what FM synthesis is about, and we are about to discover how!
In the digital case, the sine wave is stored as a series of numbers which are then changed to a form which we can eventually perceive — by the DAC (Digital to Analogue Converter).