Technical Essays For Electronic Music

Peter Elsea

Fall 2002

Music 80c Technical Writings 2000

These essays cover basic information that every electronic musician should know. The toys are getting easier and easier to use these days, so practically anyone can get some sounds going and call themselves an electronic musician, but if you want to go beyond the factory presets and software demos, you need to know what's going on under the hood.

To Quote Albert Einstein:
"Everything should be as simple as possible, but no simpler."

You can find slightly revised versions of these essays and some new ones on the electronic music web site

http://arts.ucsc.edu/ems/music

The Propagation Of Sound

The notion of sound is rather remarkable. Something happens there, and we know it here, even if we are looking the other way, not paying attention, or even asleep. The fact that some sounds can produce complex physical and emotional effects beyond the mere signaling of events is just short of astounding. These notes will perhaps remove some of the mystery associated with sound and hearing, but probably none of the wonder.

Sound, in the usual sense, is a disturbance of the atmosphere that human beings can hear. Such disturbances are produced by practically everything that moves, especially if it moves quickly or in a rapid and repetitive manner. For the purposes of discussion, I will limit possible sound generators to two: one that moves once to produce a sound IMPULSE (such as a cap gun), and one that moves the air repetitively to produce a TONE (such as a penny whistle).

You should be aware that the air is made up of molecules just like everything else. Most of the characteristics we expect of air are a result of the fact that these particular molecules are very light and are in extremely rapid but disorganized motion. This motion spreads the molecules out evenly, so that any part of an enclosed space has just as many molecules as any other. If a little extra volume were to be suddenly added to the enclosed space (such as by moving a piston into a box), the molecules nearest the new volume would move into the recently created void, and all the others would move a little farther apart to keep the distribution even.

Because the motion of the molecules is so disorganized, this filling of the void takes more time than you might think, and the redistribution of the rest of the air molecules in the room takes even longer. If the room were ten feet across, the whole process might take 1/100 of a second or so. If the piston were to move out suddenly, the volume of the room would be reduced and the reverse of the process would take place, again taking a hundredth of a second til everything was settled down. The interesting thing you would discover if you repeated this experiment is that no matter how far or how quickly the piston is moved, it always takes the same time for the molecules to even out.

In a larger room (or outdoors for that matter), times measured until the piston's motion affects a particular molecule vary directly as the distance from that molecule to the piston, so you could say the disturbance caused by the piston moves at a constant rate through the air. If you could make the disturbance visible somehow, you would see it spreading spherically from the piston, like an expanding balloon. Because the process is so similar to what happens when you drop an apple into a bucket, I call the disturbance line the WAVEFRONT.

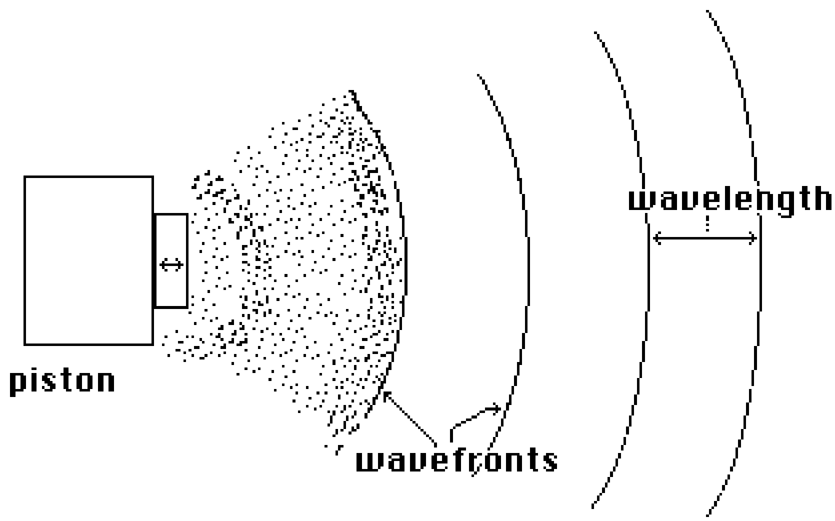


Fig.1 Propagation of pressure disturbance in air

Under normal conditions a wavefront moves through air at 1130 feet/second. This is the SPEED OF SOUND.

I mentioned that air gets its properties from the weight and speed of its molecules. Molecules of different weight (other gases or any material for that matter) or different velocity (that translates into temperature) will exhibit different speeds of sound. Motion of the sound source, the listener, or the air itself has no effect on the speed of sound, nor does the pressure of the air.

A question comes to mind at this point: how far can a wavefront go? You can think of the piston as having imparted a certain-sized chunk of energy to the air. This energy is manifest as a difference in pressure between the affected and not-yet-affected areas of air, which is to say across the wavefront. The front will continue moving until the energy is used up. As the front expands spherically, the energy is spread over a larger and larger area, in a way suggested by the relationship between the radius and area of a sphere, $A=4\pi R^2$. The total energy stays the same, the area expands, so energy per area unit decreases proportionally to the square of the distance from the source until there isn't enough to measure. The stronger the impulse, the farther the wavefront goes (the sound of the explosion of Krakatoa is said to have gone around the world three or four times).

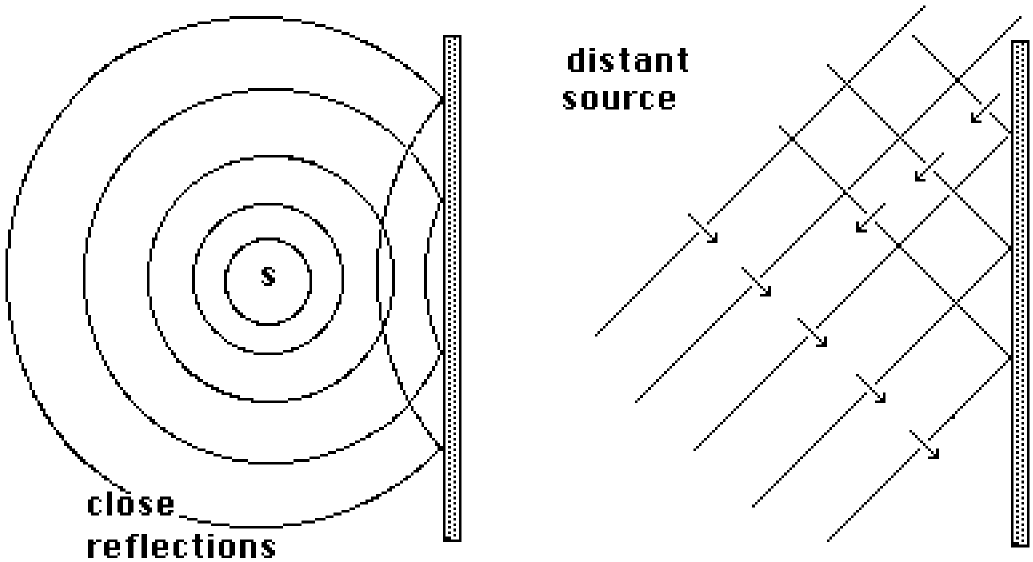


Fig. 2 Reflection of sound

REFLECTION

What happens if the wavefront hits something before it dies out? If the something is soft the energy of the wavefront is dissipated in moving the object (or its surface) around. If the object is rigid (like a wall) two things happen. Part of the energy of the wavefront will set up a wavefront within the wall; just how big a part is transmitted this way depends on the material the wall is made of. The rest of the energy is reflected off the surface according to the same rules that apply to light on a mirror. (The most important of these is that "the angle of reflection equals the angle of incidence".) Figure 2 shows this happening in two kinds of situations, one in which the source of disturbance is relatively close to the wall, and one in which the source is far enough away that you don't notice the wavefront is curved. Several repeated wavefronts are shown in each example.

At the risk of getting ahead of myself, I will point out two important effects of reflections. One is that a listener might hear something twice if the distances involved are great enough, the other is that the energy reflected back into a room is going to change our notions about how long the disturbance endures.

DIFFRACTION.

Here's another question; suppose the wavefront hits a wall with a hole in it? A small portion of the wave energy leaks through the hole and begins propagating as if the hole were the source, that is, spherically around the hole. The amount of energy available depends on the size of the hole. This process is known as DIFFRACTION.

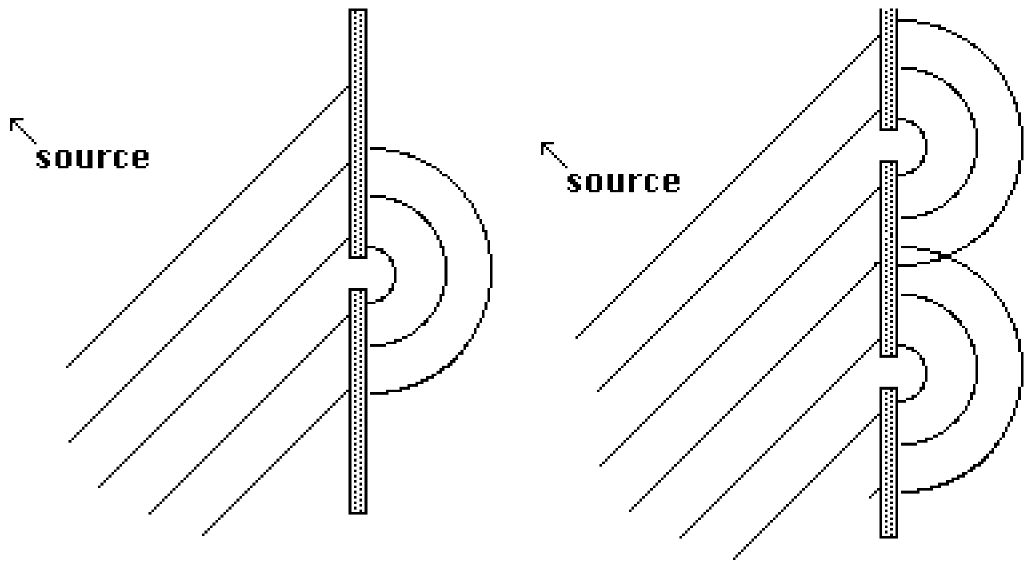


Fig. 3 Diffraction

An interesting effect is created when there are two holes in the wall. Each hole produces a wavefront, and the two separate wavefronts coincide at some point. If you happened to measure the strength of the wave at that point, the energies of both fronts would be combined, and the measurement would be quite different from one obtained only a few inches away. This effect is called INTERFERENCE, which I'll cover in a little more detail later.

Now let's backtrack and see what happens if we produce several disturbances or wavefronts at a steady rate. Initially, the separate fronts do not interact because each moves away from the source at the same speed. If we could freeze everything and take some measurements, we would see that the wavefronts were evenly spaced at a distance expected by considering the rate of repetition and the speed of sound.* We call this distance the WAVELENGTH. While things were still frozen, we could make a graph of the exact pressure changes in the space between the wavefronts. We call that graph the WAVEFORM. (Since the repeating

*Wavelength = speed of sound divided by frequency.

waves are created by something moving back and forth, the wave form will represent a pressure pattern that goes both above and below the "normal" or undisturbed value.)

I should take this opportunity to point out that the time between the creation of wavefronts is called the PERIOD, and that the number of waves produced per second is called the FREQUENCY. (Events per second is such a common concept in acoustics it needs an abbreviation. For a long time cps was used, but that can be confused with some other units so the name was changed to Hertz, abbreviated Hz.)

Now, letting our waves move again, we find the behavior of single wavefronts is the same as for those taken one at a time, but complications arise when we keep track of several.

For instance, we saw above that the movement of the wave through the air is not affected by any motion of the source or listener. Such motions will affect measurements of wavelength or frequency of repeating waves, however. If the source were to move between production of wavefronts, it would seem to be chasing the first front, and so the second front would be a little closer to the first than you would expect, at least if you measured in the direction the source was moving. If you measured the other direction, behind the source, the wavefronts would be too far apart. If you measure at right angles to the motion, you would get the "right" answers. As a corollary to this, any measurement of frequency would be high in the direction of movement, and low behind the source. This is called the DOPPLER EFFECT, and works the same way if the measuring apparatus is moving and the source standing still.

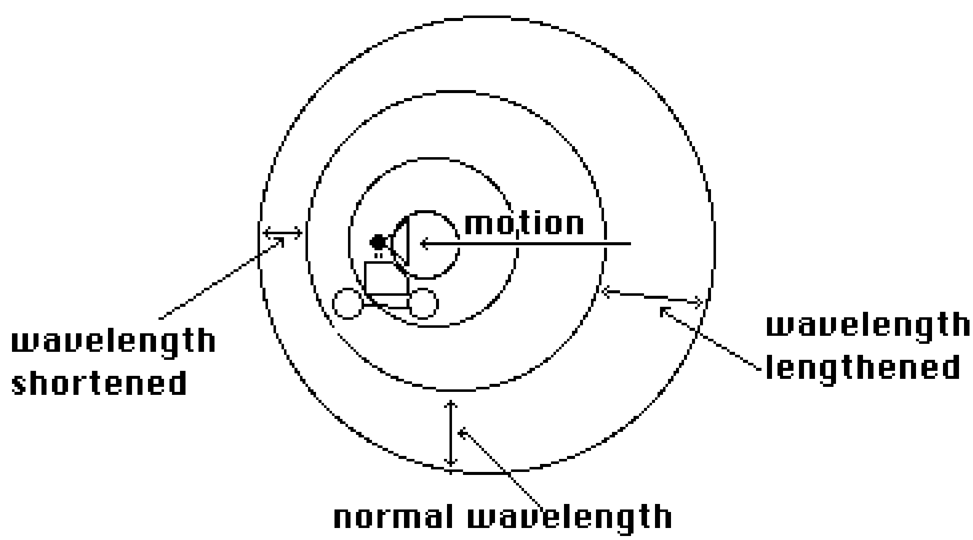


Fig. 4 Doppler Effect

Interference

When you analyze reflection and diffraction of repeating waves, you have the complication of waves running into each other. Now when two waves intersect, they have absolutely no effect on each other, but if you happen to measure the pressure at that particular point and time, you find the pressures of the two waves added together. If the two low spots coincide, you get a super low reading, and if the high of one wave coincides with the low of the other, you might easily get the normal value for the room and conclude nothing was going on at all.

You might think these are only momentary aberrations, but in fact it is quite easy for reflection or diffraction to establish locations where there will always be coinciding waves which are matched up high for high and low for low. (This is known as being IN PHASE.) These locations are any spot where two waves will pass, and for which one wave must travel exactly one wavelength (or multiple thereof) farther than the other one to reach. At this spot the two waves reinforce each other.

It also is common to find spots where two intersecting waves are matched up high for low and low for high (OUT OF PHASE). Such spots are anywhere that two waves must pass where one of the waves will travel one half (or 1+1/2, or 2+1/2 etc.) wavelength farther than the other. At this spot, the two waves cancel each other out.

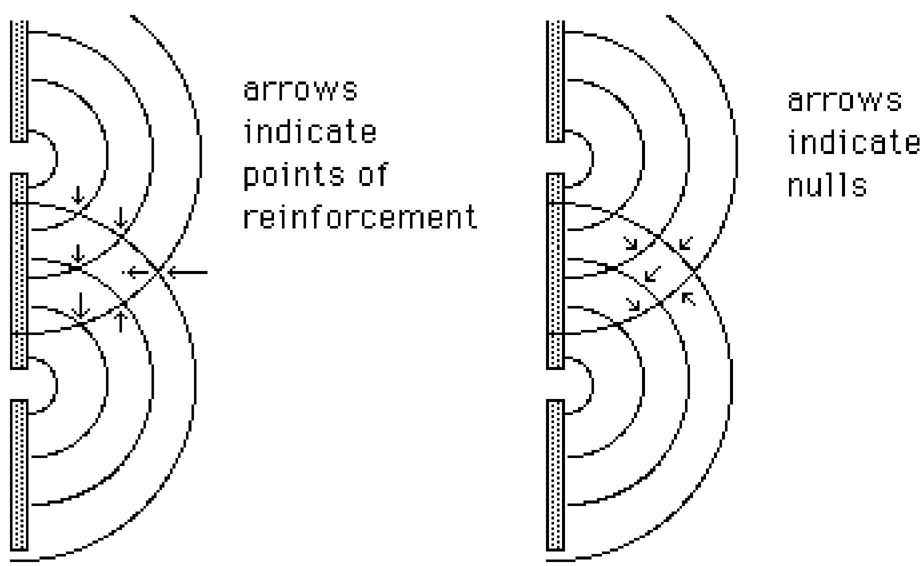


Fig. 5 Interference

If you study the interference drawing, you can find many places where the two sets of waves produced by the holes are in phase and, not far away, spots where the waves are out of phase. You should also realize that if the wavelength were to be changed, the in-phase and out-of-phase spots would move. Notice that as you move away from the holes, you find fewer such spots, so that at a distance of four or five wavelengths you are safe from this phenomenon. This whole effect, incidentally, is produced by any obstacle with a size equal to or larger than the wavelength. (This is called "shadowing", although we don't get reduced sound in the shadow, merely these interference effects.)

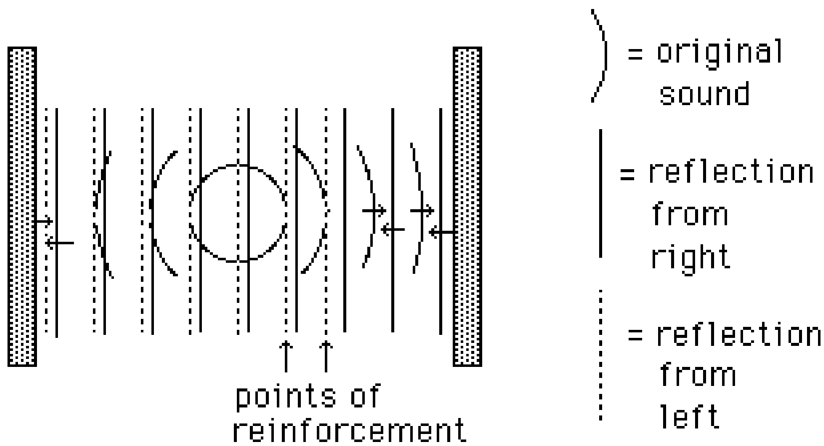


Fig. 6 Standing Waves

Standing Waves

This drawing shows how an extreme case of reinforcement can be established. Here, the source is placed between two parallel walls that happen to be exactly nine wavelengths apart. A wave will bounce off the right wall, travel to the left wall and bounce again, arriving at the source in phase with the new wavefronts. If the walls are fairly reflective, a particular front will pass many times, so the pressure differences measured at the points of reinforcement will be the sum of several waves. There are also null points, where no sound is detected. This process is very dependent on the wavelength of the sound. Change it only slightly and the system will settle down, double it or cut it in half and things go wild again. Standing waves are a pain in the neck in rooms, but can be very useful in the construction of musical instruments.

Acoustics For Music

Most of our music making is carried out indoors. In such a situation, the listener's experience is formed almost as much by the room itself as by the instruments. For a successful performance (or recording), the concert space (or studio, or living room with recorded sounds) must fulfill the following:

The audience must clearly hear all of the music with the proper balance between instruments, and the proper tonal balance for each instrument.

The performer must clearly hear himself and the other performers.

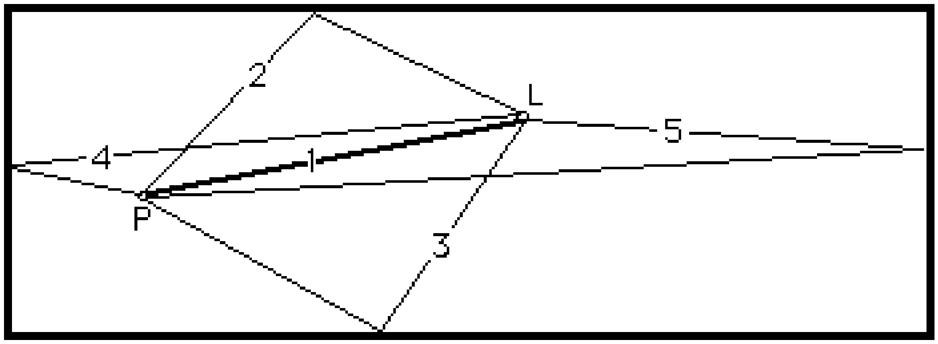
Reverberation should be appropriate to the style of the music.

Extraneous sounds must be inaudible in the concert space.

(The sound of the concert should be inaudible outside of the concert space.)

These goals are more or less in order of importance. The last requirement will not affect the concert itself, but may affect the possibility of holding future concerts. With these criteria in mind, we will examine the important structural factors of the the room which control them.

Sound In A Room



Paths of sound from performer P to listener L

Fig.1 Direct sound and early reflections

Figure 1 shows the paths taken by the sound as it travels from the performer to the listener. (The wavefronts of the sound are not shown, they would be perpendicular to the lines drawn.) The heavy line, number 1, shows the shortest path, the direct one. The other paths all involve one reflection, so must be longer than the direct path, although their relative lengths will change as the performer and listener move about the room. Since sound travels at a steady 1 foot per millisecond, the sound of a single event is going to arrive at the listener's ears several times as determined by the different path lengths. We can chart the arrival times on a graph:

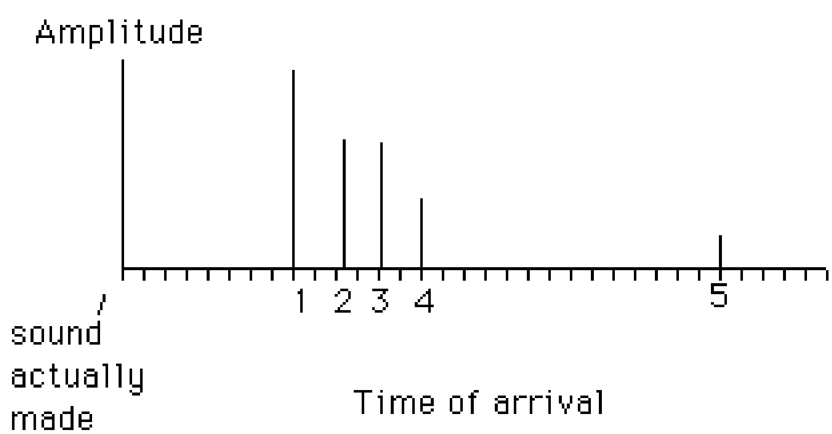
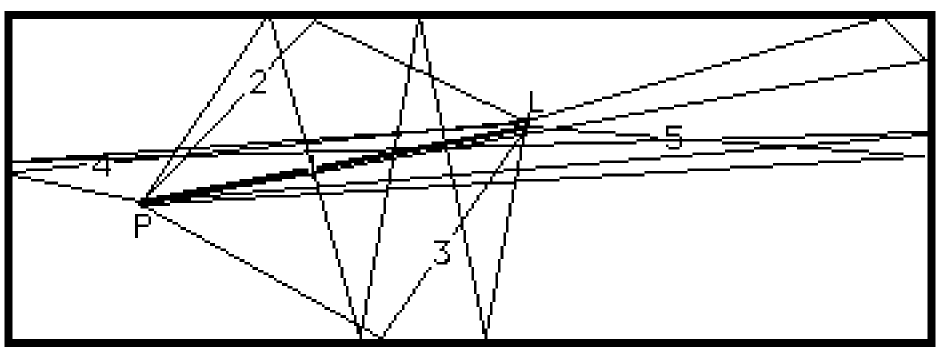


Fig. 2 Arrival times of a single sound

The amplitude of a particular reflection is determined by the path length and the efficiency of the wall in reflecting sound. That efficiency is described as the **coefficient of absorption** (any sound not reflected is absorbed). The coefficient of absorption is a number between 0 and 1, with 1 representing total absorption (an open window) and 0 representing total reflection.

We are very used to hearing sounds indoors, so we have learned not to be confused by the multiplicity of sounds arriving from various directions. We almost always realize the sound comes from the direction of the first arrival. (The whole issue of localization is too involved to get into here. It depends a lot on the number and shape of our ears.) Any reflections that arrive within 20 milliseconds of the first add to the impression of loudness of the sound. Any reflections that arrive more than 40 milliseconds after the first may be heard as a distinct echo, but are usually accepted as reverberation. Reflections that arrive between 20 and 40 milliseconds[1] after the direct sound can be confusing and interfere with understanding if the sounds are speech.

Sound does not stop at the listener's ears of course, it continues and is reflected again by the other walls of the room. If the coefficient of absorption is low, a sound may bounce several dozen times before it fades away.



Paths of sound from performer P to listener L

Fig. 3 More reflection paths

This drawing would be solid black if all of the possible reflections were shown.

[1]This time actually varies greatly from person to person. The numbers given are an average.

The arrival time graph is more informative:

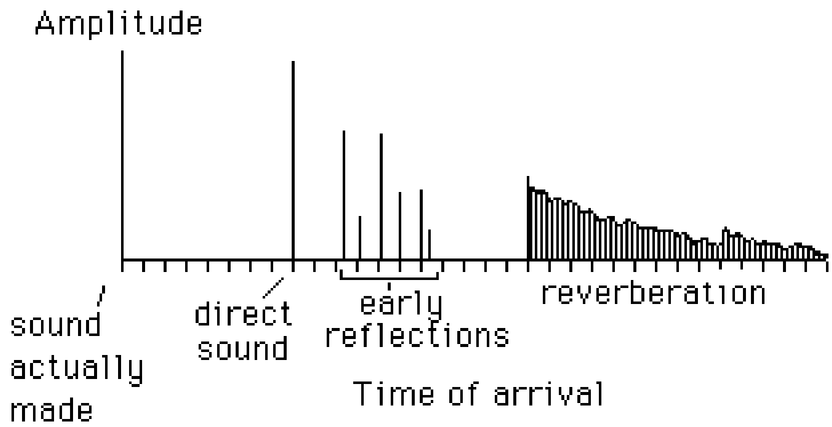


Fig. 4 Time and amplitude of sounds at listener's ear

This shows the complete picture of what is heard if a single, short sound is produced in a room. Most of the sound energy that is reflected twice or more is heard as reverberation, a sort of stretching of the sound event. The actual amplitude of reverberation is not very important (unless it is strong enough to obscure following sounds) but the time that it persists is. Short reverb times (a half to a full second) are comfortable for speech, whereas moderate times (1 to 3 seconds) work well with various kinds of music. Some music was written for very reverberant environments such as large stone churches, and should be heard that way.

Reverberation time is the most often quoted description of a performing space, but it is not really the most important. The frequency response of the reverb sound be reasonably flat, or slightly low pass, which is sometimes described as "warm reverb". That means that low partials of sounds will persist a little longer that high components, matching the decay characteristics of most instruments. The opposite effect, where high pitched sounds linger, can be very annoying. This is the situation in many indoor swimming pools.

The envelope of the reverberation should match that in figure 4, a fairly even decay, with no "lumps" of sound. A rectangular room with flat walls will not provide such an envelope; the reverberation will occur in bursts, often with distinct echoes ("slap-back). To provide even reverberation, the shape of the walls should be complex, but not very regular. A regular structure, such as a staircase, will often produce a series of echoes called flutter echo.

Isolation

Control of reflections and reverberation can satisfy the first three goals on our list. Isolation is a matter of the materials and techniques used to build the room. The walls must be heavy and solid, and for really excellent isolation, all walls, doors, floor and ceiling must be doubled; literally one room within another. Attention must be paid to such details as air ducts and holes for electrical cables, for sound can leak through any opening. Once an adequately isolated structure is finished, noise generating devices must be kept out. Light fixtures, (especially fluorescent), heaters, and backstage equipment can all create noise and must be chosen for quiet operation.

Adequate isolation is almost impossible to achieve after construction if it was not built in in the first place, but since it is an issue that is very important to low budget recording and electronic music, here are a few things that can be tried.

> First, find the leaks that sound follows between the studio and the outside world. Edges of doors, vent ducts, electrical outlets are all suspect. They can be treated with the materials sold for heat insulation, if the heavy, expensive versions are used.
>
> Direct attachment of sound sources to walls, floors or ceiling should be avoided. Swing speakers from ropes or mount them on stands. Put three layers of carpet on the floor, or set things on the canvas part of camp stools.
>
> Hang absorptive materials. Heavy curtains or rugs from floor to ceiling work well, as does four inch thick fiberglas insulation. (Thinner fiberglas has poor frequency response) There are plastic foams designed for this purpose, but they are expensive and a fire hazard. Egg carton material has a nice shape for diffusion, but is not particularly absorptive.
>
> If the above procedure makes the room too dead, hang some light hard panels in front of but not touching the absorption.

Building For Good Acoustics

Most concert halls have a lot of attention paid to acoustics. (I should hope anyway!) Here are some of the obvious features you will see:

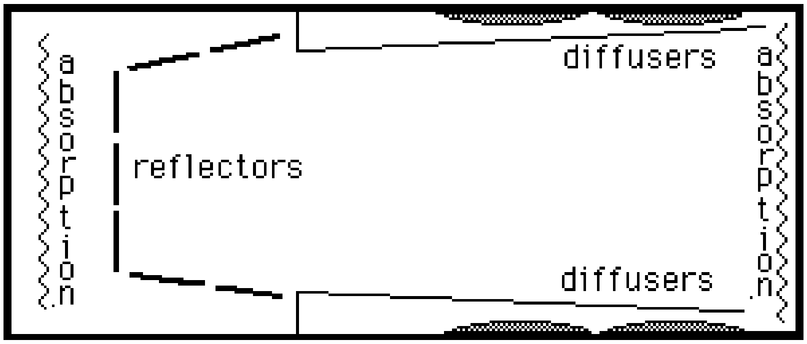


Fig. 5 Some structures to control reflections and reverberation

The diffusers smooth out the reverberation and make the sound reasonably uniform at different seats. The absorptive curtains allow the reverberation time of the room to be adjusted to control the loudness of ensembles of various sizes. The shell behind the performers serves to group the early reflections into the "sooner than 20 ms" range and also (probably more important) enables the performers hear each other better.

It’s impossible to build a single hall that is satisfactory for all kinds of music. Choirs prefer a long reverberation (up to 2 seconds), whereas Jazz sounds better in a much drier environment (say a stuffy night club with a few beers on the table.) Sometimes acoustic elements are added to fit the particular group—these can be reflectors to focus the sound for small groups in large stages, or absorptive “gobos” to tame a big drum set.

The issue of architectural acoustics is very complex, and often not handled well. It seems that most concert halls are constantly being tinkered with and occasionally rebuilt at fantastic costs; perhaps our expectations are unrealistic, now that we are used to hearing every note and nuance in our living room

The Numbers (and Initials) of Acoustics

Architectural acoustics is about half engineering, half art. The art comes from experience in choosing, placing and evaluating various materials and structures. The engineering comes from measuring the effects of the materials and structures and relating what can be measured to what can be heard. There are standards for making test measurements so you can compare specifications from various manufacturers when choosing materials and prebuilt items such as doors. The standards are also used in specifying and evaluating performance of finished construction such as floor, wall, and window isolation.

Graphs

Many acoustical numbers are presented as curves on graphs. In order to accommodate the nature of hearing, peculiar graph paper is often used. The horizontal divisions represent frequency, but they are spaced in such a way that an octave is the same width anywhere across the graph. Notice that each mark up to 100 represents 10 hz, each from 100 to 1000 represents 100hz, and from 1000 to 10,000, each is 1000 hz. The spacing is therefore logarithmic. The vertical divisions are equally spaced, but since they are marked in decibels, the graph is logarithmic in this direction also. This makes the perceived effect of any deviations in the curves the same anywhere on the graph.

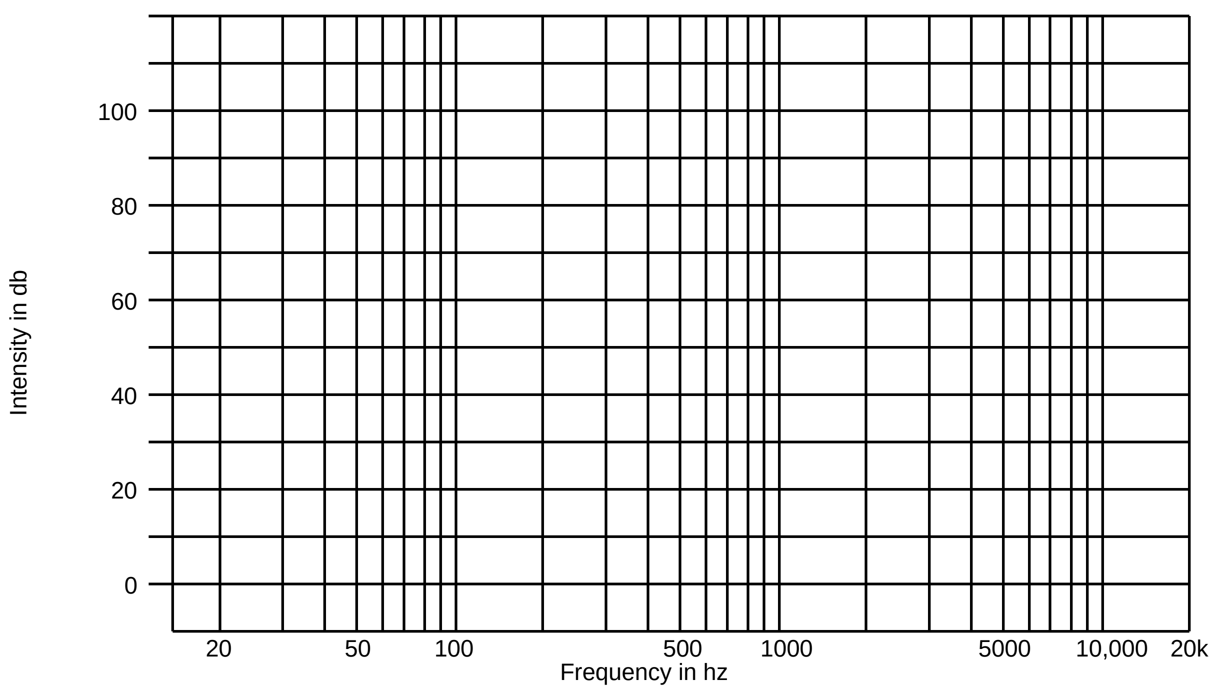


Fig 1 Logarithmic graph paper

Equal Loudness Curves

You will see lots of references to equal loudness curves or equal loudness contours- these are based on the work of Fletcher and Munson at Bell labs in the 30s, or perhaps refinements made more recently by Robinson and Dadson. These were made by asking people to judge when pure tones of two different frequencies were the same loudness. This is a very difficult judgement to make, and the curves are the average results from many subjects, so they should be considered general indicators rather than a prescription as to what a single individual might hear.

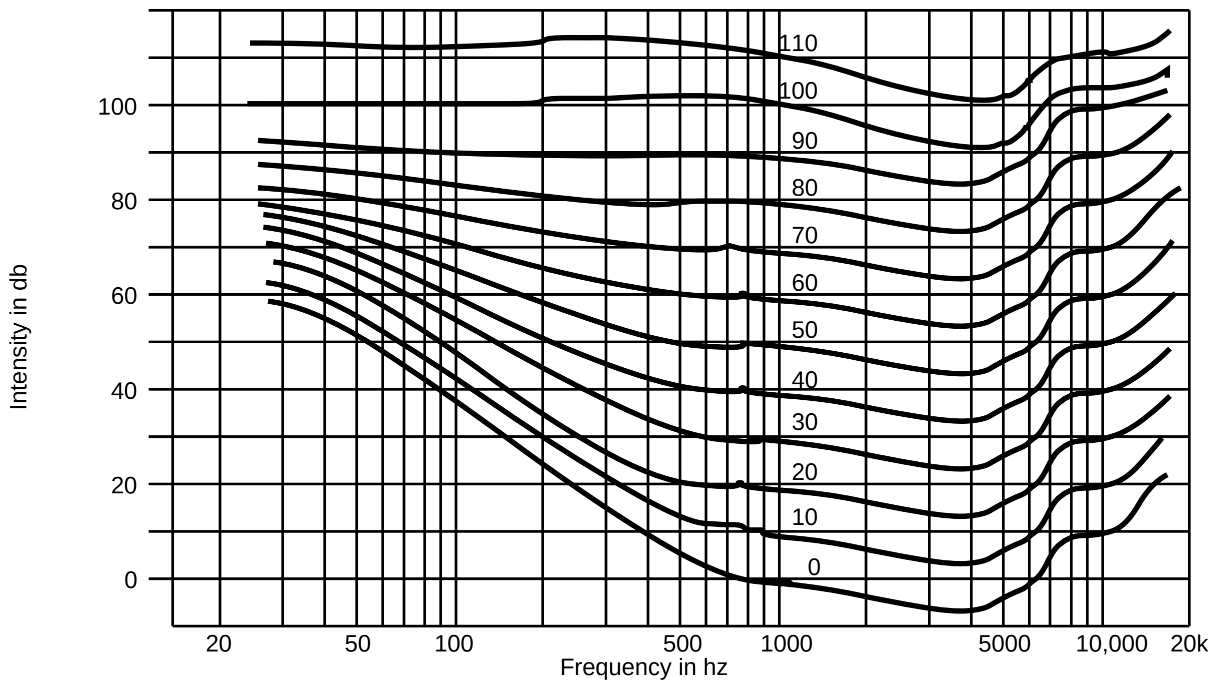


Fig 2. Equal loudness contours or Fletcher-Munson curves

The numbers on each curve identify it in terms of phons, a unit of loudness that compensates for frequency effects. To find the phon value of an intensity measurement, find the db reading and frequency on the graph, then see which curve it lands on.

The interesting aspects of these curves are that it is difficult to hear low frequency of soft sounds, and that the ear is extra sensitive between 1 and 6 kilohertz.

SPL

Sound Pressure Level is a single measurement of sound pressure in decibels relative to the threshold of hearing. That varies from person to person of course, but for the purposes of SPL measurements is defined as 2 X 10^{-5} Newtons per Meter^2 or 20 microPascals. We generally live in the mid 60s db_{SPL}, think music is loud when it gets above 90 db_{SPL}, and complain of pain at 120 db_{SPL} .

When SPL measurements are made, some adjustment for the ear's response to low frequency is usually included. This is done by using filters that follow the Fletcher-Munson curves - the A curve follows F&M at low levels and the B curve follows intermediate levels. The C curve is nearly flat.

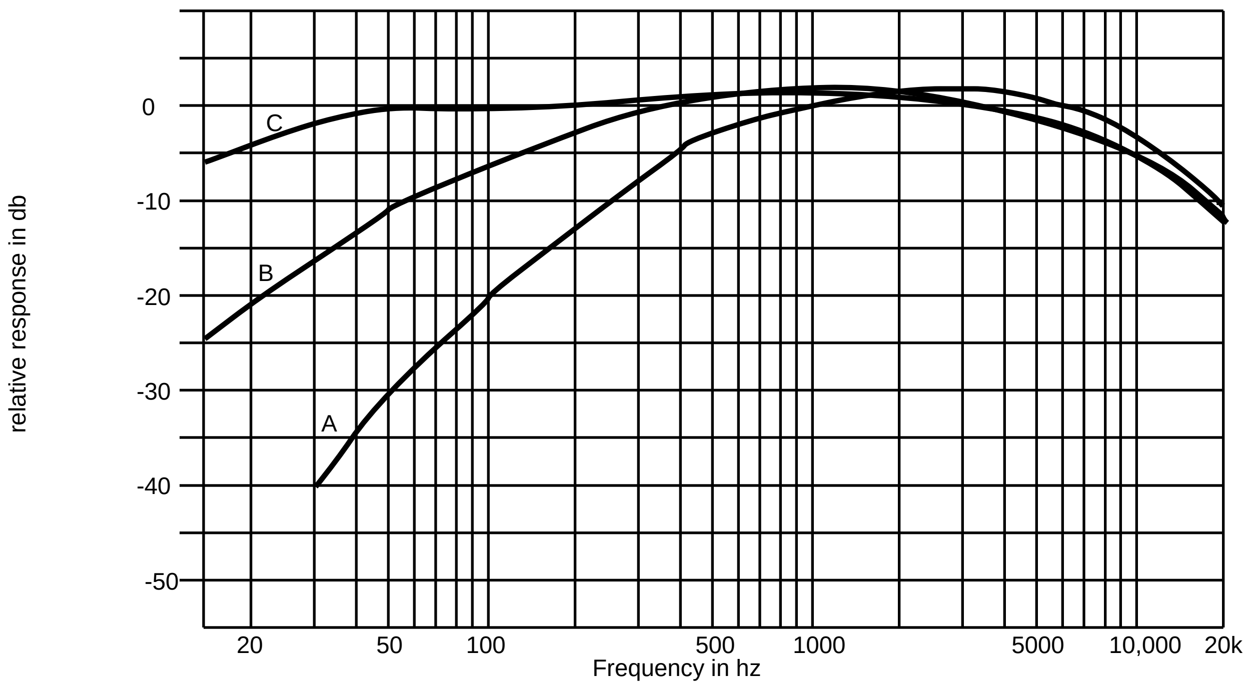


FIG.3 SPL Weighting curves,

If it seems odd that these curves turn down when the equal loudness curves turn up, remember that this is a frequency response and the loudness curve indicates sensitivity. The idea is that if you find noise at a low level and low frequency, it doesn't count for much since it is hard to hear.

SPL is usually measured with special meters that have the weighting curves built in. Use the A curve for soft measurements and the C curve for loud sounds, and the flat setting for comparative measurements like transmission loss.

NC

Noise Criteria levels or NC ratings are a common way of specifying the background noise in rooms. It's not quite the same as SPL- there is a special filter used which discounts low frequency sound even more. It also only includes about half of the audible spectrum. Generally, A weighted SPL readings run about 10db higher than NC readings.

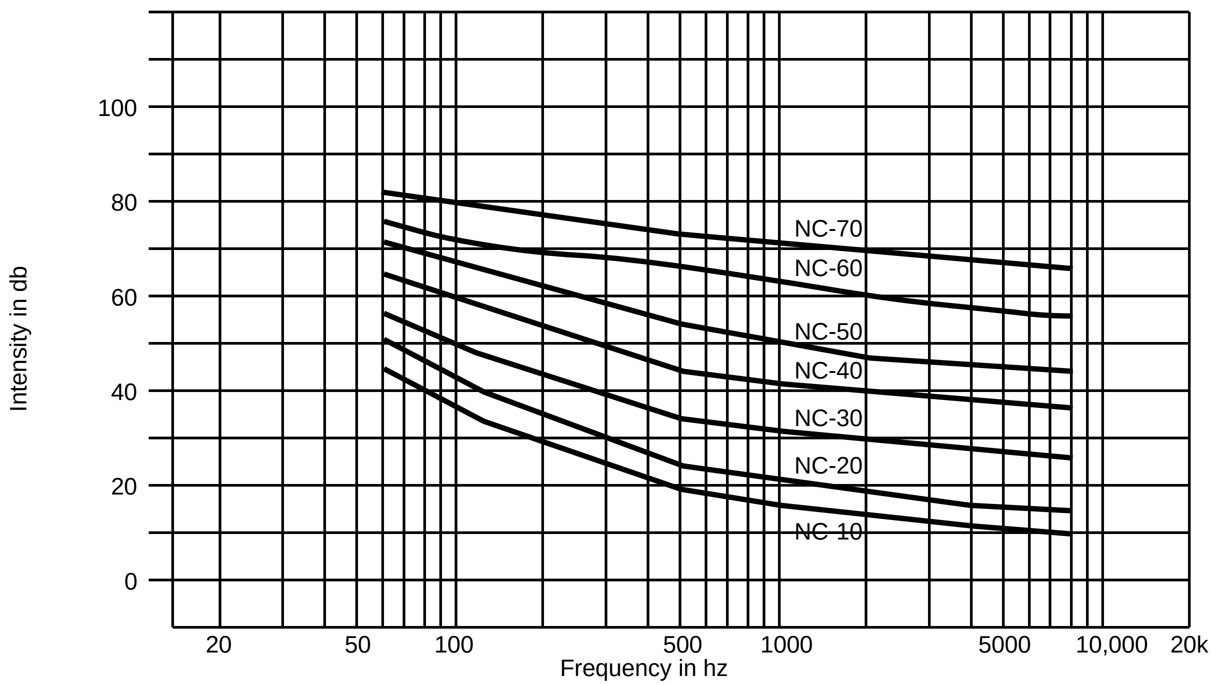


Fig. 4 Noise Criteria curves

To get an NC value, find the curve that is just below all of your measurements. You can see from these curves that an environment that measures NC-20 could have a 50 db level of 60hz hum. Although the Fletcher-Munson effect would make this tolerable for a classroom, it's not appropriate for a recording studio, because soft sounds are often amplified to the point where our hearing response is flat.

TL

Transmission Loss is simply the reduction of SPL as sound travels through a structure. It will vary with frequency, and should be presented as a curve or series of numbers at selected frequencies. Very often you see a chart like this:

125hz	250hz	500hz	1000hz	2000hz	4000hz	STC
15	19	21	28	33	37	27

Fig 5. Transmission loss

This gives the transmission loss in db at the frequencies listed. Notice the last entry, instead of a frequency, is an overall rating called STC.

STC

Sound Transmission Class is another set of curves, again relaxed in stringency at low frequency. To convert a TL curve to STC, you find an STC curve that fits the measured curve within 8 db, then specify the value from the STC curve at 500hz. This is useful for comparing products in a catalog, but the true transmission loss curve is necessary to predict what will happen with music.

Absorption Coefficient

The absorptive efficiency of a material is given by its Absorption Coefficient, which is the ratio of the sound energy that is reflected back to the arriving sound energy. A totally reflective material has an absorption coefficient of 0, and an open window has an AC of 1. As with most things acoustical, the value varies with frequency, although you will often see a single number specified. Here's how various thicknesses of Fiberglas stack up:

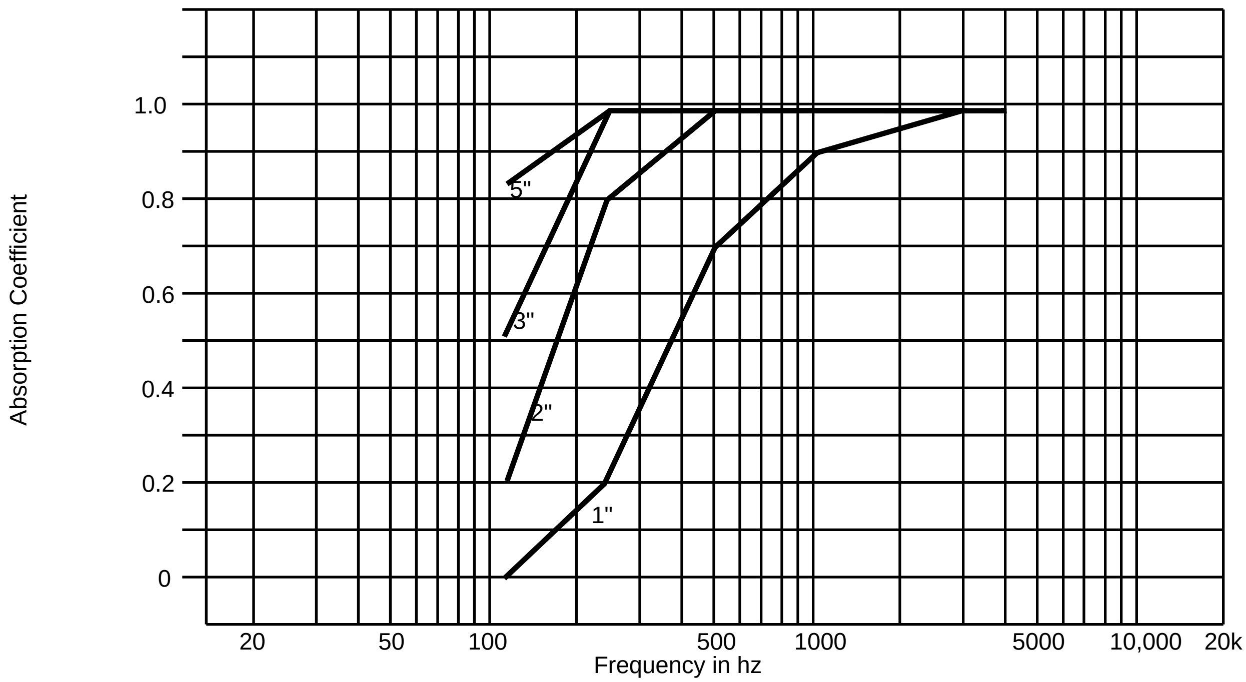


Figure 6. Absorption Coefficient of Fiberglas.

NRC

The Noise Reduction Coefficient of a material is the average absorption measured at 250, 500, 1000, and 2000 hz. It's useful for designing offices, but not sound studios. Even the complete specs given above don't tell what happens in the lowest octaves, although you can guess from the trend of the curves.

Speed of Sound

The most important number to remember is that the speed of sound under typical conditions is 1130 ft/second. This varies with temperature (slows down when it's cold) but not with any other conditions you will encounter in recording studios. For many purposes, we can use the rule of thumb that sound travels a little faster than one foot in a millisecond.

Frequency

This how often something happens. For a steady tone, the frequency is the number of wavefronts that pass your ear in one second. It is measured in hertz or hz, which can be though of as meaning 1/sec or "per second". We usually consider the audible range of frequency to be from 20hz, to 20,000hz. Few people can hear all of this range, but some can hear beyond.

Wavelength

This is the distance between wavefronts of a steady tone. It is often represented in formulas by the greek letter lambda, which looks like an upside down y.

The formula.

Remember the relationship between frequency, wavelength and speed of sound is:

Wavelength = Speed_of_Sound / Frequency

Hearing And Perception

The operation of the ear has two facets: the behavior of the mechanical apparatus and the neurological processing of the information acquired. The mechanics of hearing are straightforward and well understood, but the action of the brain in interpreting sounds is still a matter of dispute among researchers.

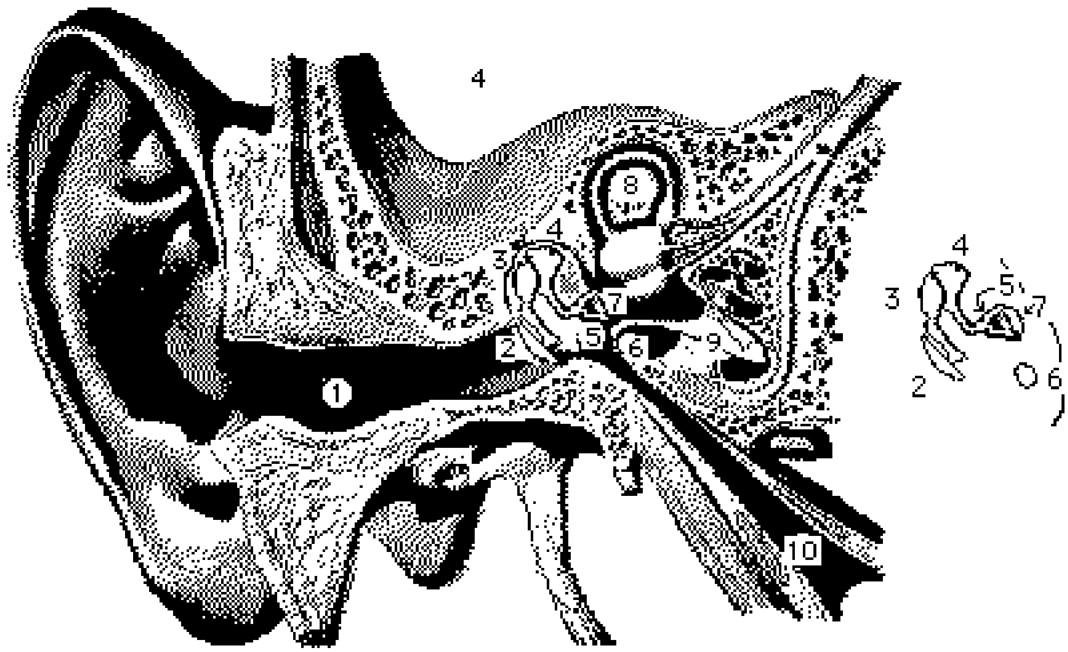


Fig. 1 Parts of the ear

1. Auditory canal
2. Ear drum
3. Hammer
4. Anvil
5. Stirrup
6. Round window
7. Oval window
8. Semicircular canals
9. Cochlea
10. Eustachian tube

THE EAR MECHANISM

The ear contains three sections, the outer, middle, and inner ears. The outer ear consists of the lobe and ear canal, structures which serve to protect the more delicate parts inside.

The outer boundry of the middle ear is the eardrum, a thin membrane which vibrates in sympathy with any entering sound. The motion of the eardrum is transferred across the middle ear via three small bones named the hammer, anvil, and stirrup. These bones are supported by muscles which normally allow free motion but can tighten up and inhibit the bones' action when the sound gets too loud. The leverages of these bones are such that rather small motions of the ear drum are very efficiently transmitted.

The boundry of the inner ear is the oval window, another thin membrane which is almost totally covered by the end of the stirrup. The inner ear is not a chamber like the middle ear, but consists of several tubes which wind in various ways within the skull. Most of these tubes, the ones called the semicircular canals, are part of our orientation apparatus. (They contain fine particles of dust-the location of the dust tells us which way is up.) The tube involved in the hearing process is wound tightly like a snail shell and is called the cochlea.

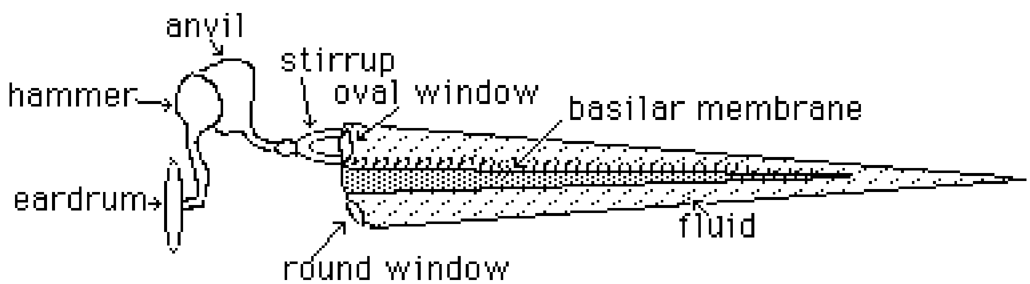


Fig 2. Schematic of the ear
This is a diagram of the ear with the cochlea unwound. The cochlea is filled with fluid and is divided in two the long way by the basilar membrane. The basilar membrane is supported by the sides of the cochlea but is not tightly stretched. Sound introduced into the cochlea via the oval window flexes the basilar membrane and sets up traveling waves along its length. The taper of the membrane is such that these traveling waves are not of even amplitude the entire distance, but grow in amplitude to a certain point and then quickly fade out. The point of maximum amplitude depends on the frequency of the sound wave.

The basilar membrane is covered with tiny hairs, and each hair follicle is connected to a bundle of nerves. Motion of the basilar membrane bends the hairs which in turn excite the associated nerve fibers. These fibers carry the sound information to the brain. This information has two components. First, even though a single nerve cell cannot react fast enough to follow audio frequencies, enough cells are involved that the aggregate of all the firing patterns is a fair replica of the waveform. Second, and probably most importantly, the location of the hair cells associated with the firing nerves is highly correlated with the frequency of the sound. A complex sound will produce a series of active loci along the basilar membrane that accurately matches the spectral plot of the sound.

The amplitude of a sound determines how many nerves associated with the appropriate location fire, and to a slight extent the rate of firing. The main effect is that a loud sound excites nerves along a fairly wide region of the basilar membrane, whereas a soft one excites only a few nerves at each locus.

Perception

The mechanical process described so far is only the beginning of our perception of sounds. The mechanisms of sound interpretation are poorly understood, in fact is not yet clear whether all people interpret sounds in the same way. Until recently, there has been no way to trace the wiring of the brain, no way to apply simple stimuli and see which parts of the nervous system respond, at least not in any detail. The only research method available was to have people listen to sounds and describe what they heard. The variability of listening skills and the imprecision of the language combined to make psycho-acoustics a rather frustrating field of study. Some of the newest research tools show promise of improving the situation, so research that is happening now will likely clear up several of the mysteries.

The current best guess as to the neural operation of hearing goes like this:

We have seen that sound of a particular waveform and frequency sets up a characteristic pattern of active locations on the basilar membranes. (We might assume that the brain deals with these patterns in the same way it deals with visual patterns on the retina.) If a pattern is repeated enough we learn to recognize that pattern as belonging to a certain sound, much as we learn a particular visual pattern belongs to a certain face. (This learning is accomplished most easily during the early years of life.) The absolute position of the pattern is not very important, it is the pattern itself that is learned. We do possess an ability to interpret the location of the pattern to some degree, but that ability is quite variable from one person to the next. (It is not clear whether that ability is innate or learned.) What use the brain makes of the fact that the aggregate firing of the nerves more or less approximates the waveform of the sound is not known. The processing of impulse sounds (which do not last long enough to set up basilar patterns) is also not well explored.

INTERPRETATION OF SOUNDS

Most studies in psycho-acoustics deal with the sensitivity and accuracy of hearing. This data was intended for use in medicine and telecommunications, so it reflects the abilities of the average untrained listener. It seems to be traditional to weed out musicians from such studies, so the capabilities of trained ears are not documented. I suspect such capabilities are much better than that suggested by the classic studies.

LOUDNESS

The ear can respond to a remarkable range of sound amplitude. (Amplitude corresponds to the quality known as loudness.) The ratio between the threshold of pain and the threshold of sensation is on the order of 130 dB, or ten trillion to one. The judgment of relative sounds is more or less logarithmic, such that a tenfold increase in sound power is described as "twice as loud". The just noticeable difference in loudness varies from 3 dB at the threshold of hearing to an impressive 0.5 dB for loud sounds.

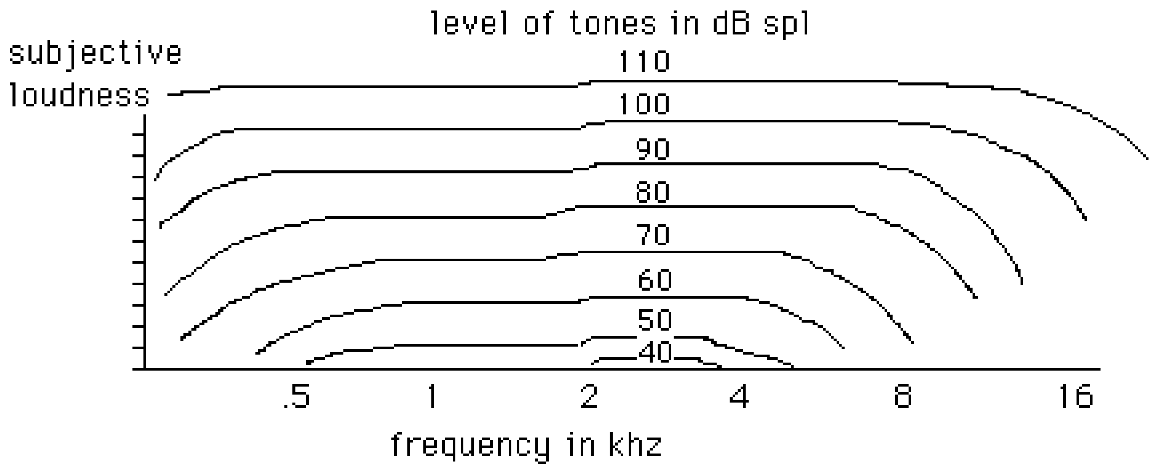


Fig. 3 Perceived loudness of sounds

The sensation of loudness is affected by the frequency of the sound. A series of tests using sine waves produces the curves shown. At the low end of the frequency range of hearing, the ear becomes less sensitive to soft sounds, although the pain threshold as well as judgments of relatively loud sounds are not affected much. Sounds of intermediate softness show some but not all of the sensitivity loss indicated for the threshold of hearing. At high frequencies the change in the sensitivity is more abrupt, with sensation ceasing entirely around 20 khz. The threshold of pain increases in the top octave also.

The ability to make loudness judgments is compromised for sounds of less than 200ms duration. Below that limit, the loudness is affected by the length of the sound; shorter is softer. Durations longer than 200ms do not affect loudness judgment, beyond the fact that we tend to stop paying attention to long unchanging tones.

MASKING

The threshold of hearing for a particular tone can be raised by the presence of another noise or another tone. White noise reduces the loudness of all tones, regardless of absolute level. If the bandwidth of the masking noise is reduced, the effect of masking loud tones is reduced, but the threshold of hearing for those tones remains high. If the masking sound is narrow band noise or a tone, masking depends on the frequency relationship of the masked and masking tones. At low loudness levels, a band of noise will mask tones of higher frequency than the noise more than those of lower frequency. At high levels, a band of noise will also mask tones of lower frequency than itself.

PITCH

People's ability to judge pitch is quite variable. (Pitch is the quality of sound associated with frequency.) Most subjects studied could match pitches very well, usually getting the frequencies of two sine waves within 3%. (Musicians can match frequencies to 1%, or should be able to.) Better results are obtained if the stimuli are similar complex tones, which makes sense since there are more active points along the basilar membrane to give clues. Dissimilar complex tones are apparently fairly difficult to match for pitch (judging from experience with ear training students; I haven't seen any studies on the matter to compare them with sine tone results).

Judgment of relative pitch intervals is extremely variable. The notion of the two to one frequency ratio for the octave is probably learned, although it is easily learned given access to a musical instrument. An untrained subject, asked to set the frequency of a tone to twice that of a reference, is quite likely to set them a twelfth or two octaves apart or find some arbitrary and inconsistent ratio. The tendency to land on "proper" intervals increases if complex tones are used instead of sine tones. Trained musicians often produce octaves slightly wider than two to one, although the practical aspects of their instrument strongly influence their sense of interval. (As a bassoonist who has played the same instrument for twenty years, I have a very strong tendency to place G below middle C a bit high.)

Identification of intervals is even more variable, even among musicians. It does appear to be trainable, suggesting it is a learned ability. Identification of exact pitches is so rare that it has not been properly studied, but there is some anecdotal evidence (such as its relatively more common occurrence among people blind from birth) suggesting it is somehow learned also.

The amplitude of sound does not have a strong effect on the perception of pitch. Such effects seem to hold only for sine tones. At low loudness levels pitch recognition of pure tones becomes difficult, and at high levels increasing loudness seems to shift low and middle register pitches down and high register pitches up.

The assignment of the quality of possessing pitch in the first place depends on the duration and spectral content of the sound. If a sound is shorter than 200ms or so, pitch assignment becomes difficult with decreasing length until a sound of 50ms or less can only be described as a pop. Sounds with waveforms fitting the harmonic pattern are clearly heard as pitched, even if the frequencies are offset by some additive factor. As the spectral plot deviates from the harmonic model, the sense of pitch is reduced, although even noise retains some sense of being high or low.

TIMBRE

Recognition of sounds that are similar in aspects other than pitch and loudness is not well studied, but it is an ability that everyone seems to share. We do know that timbre identification depends strongly on two things, waveform of the steady part of the tone, and the way the spectrum changes with time, particularly at the onset or attack. This ability is probably built on pattern matching, a process that is well documented with vision. Once we have learned to identify a particular timbre, recognition is possible even if the pitch is changed or if parts of the spectrum are filtered out. (We are good enough at this that we can tell the pitch of low sounds when played through a sound system that does not reproduce the fundamentals.)

LOCALIZATION

We are also able to perceive the direction of a sound source with some accuracy. Left and right location is determined by perception of the difference of arrival time or difference in phase of sounds at each ear. If there are more than two arrivals, as in a reverberant environment, we choose the direction of the first sound to arrive, even if later ones are louder. Localization is most accurate with high frequency sounds with sharp attacks.

Height information is provided by the shape of our ears. If a sound of fairly high frequency arrives from the front, a small amount of energy is reflected from the back edge of the ear lobe. This reflection is out of phase for one specific frequency, so a notch is produced in the spectrum. The elongated shape of the lobe causes the notch frequency to vary with the vertical angle of incidence, and we can interpret that effect as height. Height detection is not good for sounds originating to the side or back, or lacking high frequency content.

Taking The Waveform Apart

The WAVEFORM of a sound is the graph of the way the pressure changes between the wavefronts. This is usually a very convoluted pattern and the actual sound of the wave is not apparent from looking at the waveform. As a matter of fact, the waveform does not usually repeat exactly from one cycle to another.

At this point I am going to have to digress into some mathematics. (See essay 11 if this does not seem simple to you.)

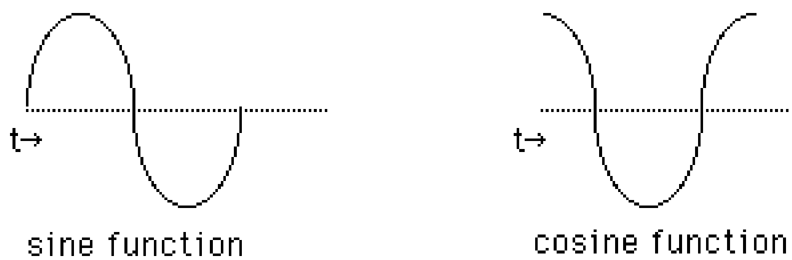


Fig 1. Sine and cosine functions

The waveform produced by simple harmonic motion is the SINE WAVE. We graph a sine wave by plotting the function:

$$f(t)=A \sin 2\pi ft$$

To do this we divide up our graph paper horizontally into equal chunks to represent a time scale, and for each time t we want to plot, we multiply t by $2\pi f$ (f=frequency) and look up the sine of the result. That sine value is what gets used for the vertical part of the graph.

There is also a function called a cosine wave. The expression is

$$f(t)=A\cos 2\pi ft$$

and it looks just like the sine wave. The difference is that the cosine of an angle is equal to the sine of an angle 90 degrees bigger. When we have two waveforms which have the same shape and frequency but are offset in time, we say they are out of phase by the amount of angle you have to add to the $2\pi ft$ term of the first to move them together. In other words the wave defined by $\sin(2\pi ft)$ is out of phase with the wave defined as $\sin(2\pi ft+p)$ by the angle p.

The second simplest waveform is probably the combination of two sine waves. Any combination of waves is interpreted by the ear as a single waveform, and that waveform is merely the sum of all of the waves passing that spot. Here are a few rules about the addition of two sine waves:

- If both have the same frequency and phase, the result is a sine wave of amplitude equal to the sum of the two amplitudes.

- If both have the same frequency and amplitude but are 180 degrees out of phase, the result is zero. Any other combinations of amplitude produce a result of amplitude equal to the difference in the two original amplitudes.

- If both are the same frequency and amplitude but are out of phase a value other 180 degrees, you get a sine wave of amplitude less than the sum of the two and of intermediate phase.

- If the two sine waves are not the same frequency, the result is complex. In fact, the waveform will not be the same for each cycle unless the frequency of one sine wave is an exact multiple of the frequency of the other.

If you explore combinations of more than two sine waves you find that the waveforms become very complex indeed, and depend on the amplitude, frequency and phase of each component. Every stable waveform you discover will be made up of sine waves with frequencies that are some whole number multiple of the frequency of the composite wave.

The reverse process has been shown mathematically to be true: Any waveform can be analyzed as a combination of sine waves of various amplitude, frequency and phase. The method of analysis was developed by Fourier in 1807 and is called FOURIER ANALYSIS.

The actual procedure for Fourier analysis is too complex to get into here, but the result (with stable waveforms) is an expression of the form:

$$A\sin\ t + B\cos\ t + C\sin 2\ t + D\cos 2\ t + E\sin 3\ t \ldots$$

and so forth. (The symbol   represents the frequency in radians per second, also known as angular frequency.) The inclusion of cosine waves as well as sine waves takes care of phase, and the letters represent the amplitude of each component. This result is easily translated into a bar graph with one bar per component. Since the ear is apparently not sensitive to phase, we often simplify the graph into a sine waves only form. Such a graph is called a SPECTRAL PLOT.

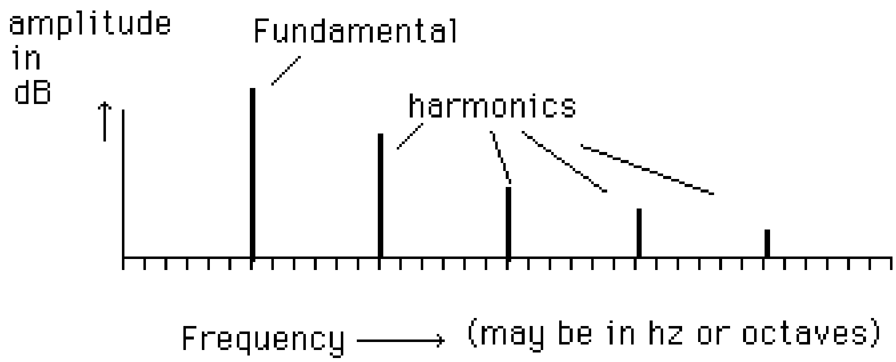


Fig. 2 A typical spectral plot

The lowest component of the waveform is known as the FUNDAMENTAL, and the others are HARMONICS, with a number corresponding to the multiple of the fundamental frequency. The second harmonic is twice the fundamental frequency, the third harmonic is three times the fundamental frequency, and so forth. It is important to recognize that the harmonic number is not the same as the equivalent musical interval name, although the early harmonics do approximate some of the intervals. The most important relationship is that the harmonics numbered by powers of two are various octaves.

Non-repeating waveforms may be disassembled by Fourier means also, but the result is a complex integral that is not useful as a visual aid. However, if we disregard phase, these waveforms may also be represented on a spectral plot as long as we remember that the components are not necessarily whole number multiples of the fundamental frequency and therefore do not qualify as harmonics. We should not say that a non-harmonic waveform is not pitched, but it is true that the less the spectral plot fits the harmonic model the more difficult it is to perceive pitch in a sound.

There are sounds whose waveforms are so complex that the Fourier process gives a statistical answer. (These waveforms are the sounds commonly called noise.) You can express the likelihood of finding a particular frequency as a component over a large enough time but you cannot assign any component a constant amplitude. To describe such sounds on a spectral plot, we plot the probability curve. A very narrow band of noise will sound like a pitched tone, but as the curve widens, we lose the impression of pitch, aware only of a vague highness or lowness of the sound.

Noise that spreads across the entire range of hearing is called WHITE NOISE if it has equal probability of all frequencies being represented. Such noise sounds high pitched because of the logarithmic response of the ear to frequency. (Our ears consider the octave 100 hz to 200 hz to be equal to the octave 1000 hz to 2000 hz, even though the higher one has a much wider frequency spread, and therefore more power.) Noise with emphasis added to the low end to compensate for this is called PINK NOISE.

A sound event is only partially described by its spectral plot. For a complete description, we need to graph the way the sound changes over time. There are two ways in which such graphs are presented. In the Sonogram, the horizontal axis is time, the vertical axis is frequency, and the amplitude is represented by the darkness of the mark. There is a machine that produces this kind of chart by mechanical means.

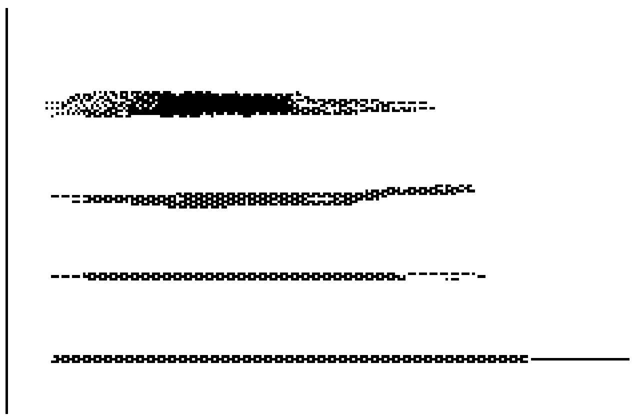


Fig 3. A Sonogram

Computer analysis usually produces a three dimensional graph on plotter or printer:

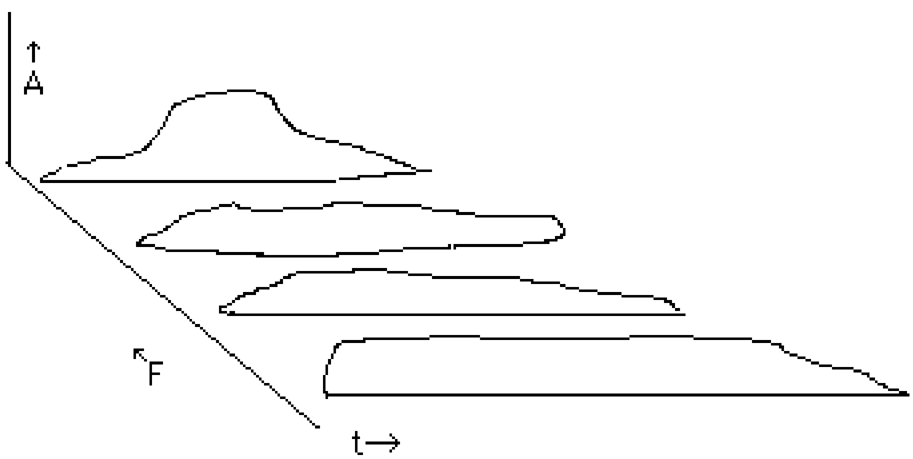


Fig 4. Spectragraph

The three dimensional graph gives a clearer sense of how the amplitudes of various components of a sound change.

Some Basic Electronics

It is adequate for a music class to take a rather simplistic view of the precise nature of electricity, so the following discussion is not complete or rigorous.

The phenomenon known as electricity involves the exchange of very small particles (electrons) between atoms. Now a particular atom generally possesses a constant number of electrons, and there are some rather potent forces within the atom working to keep the correct number of electrons, so if an atom looses an electron because of some mechanical or chemical process, it will very quickly pull in a replacement. There are seldom spare electrons around, so the replacement likely comes from an adjacent atom, which, shy one electron, steals from its neighbor, and so forth. The exchange of electrons propagates in much the same way as the pressure disturbance I discussed in the section on sound, and exhibits all of the associated wave effects. The speed of propagation is so high however, (almost the speed of light) that we may ignore the wave characteristics of electricity for all but the briefest events.

The electron exchange has two important side effects: it produces a tiny amount of heat or light, and supports a weak magnetic field. These two effects represent an energy loss, and must be compensated for by the original electron moving mechanism. They can also be thought of as energy transference from the point of generation to any place where these effects are apparent. Most of the science of electrical engineering deals with getting these effects to happen at the right place.

Since the heat and magnetism effects are small, in order to produce any useful work we must move a lot of electrons and ensure that it is a continuous process. To get a steady flow of electricity, we must have something that absorbs electrons (called a sink) something that supplies electrons (a source), and some path in between for the electrons to follow. If we want the process to be continuous forever, we must realize the sink will eventually fill and the source will be depleted, so the two should be combined and the path from one to another be made circular. The source / sink combination can be realized chemically with a battery or mechanically with a generator, or in a variety of ways I'll lump together and call the power supply.

VOLTAGE, CURRENT, RESISTANCE

If we set up a power supply and a wire connecting its source and sink terminals, we can make some observations about the flow of electricity in this simple circle, or circuit.

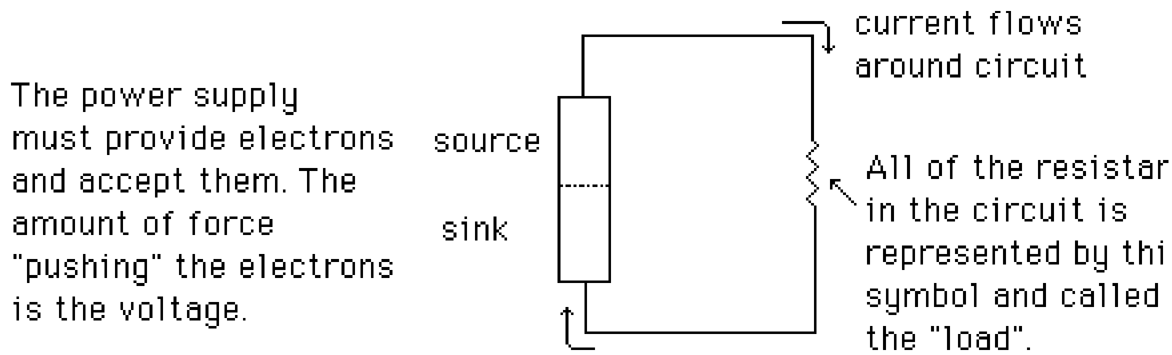


Fig. 1 A simple circuit.

The amount of heat or light produced is determined by the number of electrons moving around. The concept of "number of electrons" is called CURRENT and is measured in units called AMPERES. The symbol used to stand in for current in mathematical formulas is **I**.

The amount of heat produced for a given current depends on the material used for the wire. The heat producing mechanism opposes the flow of electrons, so it is termed the RESISTANCE. It is measured in units called OHMS, and the symbol is R. Many materials have such a high resistance that they will not carry measurable current. Such materials are termed insulators, whereas materials that will carry current are conductors.

If you want to move a current against a resistance, you must supply some ELECTROMOTIVE FORCE. This force is measured in VOLTS, is symbolized by V or E, and is a property of the power supply or battery. Electromotive force is a bit of a mouthfull, so it is also called VOLTAGE. Since this force acts in moving current from one spot to another (from source to sink) a voltage measurement must be made between two places. If the two places are on the same piece of wire, the voltage will be unmeasurable. Usually, we designate a convenient spot as being "0 volts" and use that as a common point in all measurements.

These three properties are everything we need to know about a constant current flow. They are related in the following way:

> If the voltage is held constant, and the resistance reduced, the current will increase.
>
> If the resistance is not changed, but the voltage is reduced the current will decrease.
>
> If we know two of the quantities, we can calculate the third, because the units are defined that way.

The mathematical statement of these relationships is:

$$I=E/R$$

Where (to review),
I is current in Amperes,
Eis electromotive force in Volts, and
R is resistance in Ohms.

This relationship is the cornerstone of electronics, and is known as **OHM'S LAW**.

There are many uses for current flow, of course. In fact, our modern civilization is pretty dependent on it. The two most important are the transmission of power and the transmission of information. In electronic music, we use electricity mostly to transmit and modify information; the flow of current represents the pressure of air as it changes to produce sound. To do this, we use devices that generate current from the energy of the sound waves, and ultimately, devices that generate sound from changes in current.

IMPEDANCE

Ohm's law is true for any steady state condition but if the current is changed, Ohm's law is not true while the change is taking place. This is because there are effects that oppose any change in the amount of current flow. Opposition to current change is called **REACTANCE**, and is measured in ohms because it must be combined with resistance in order to describe the current to voltage relationship. (For the time being, think of reactance as the electrical equivalent of momentum.)

This combination is not a simple addition because it must take into account the rate of change. The rate correction is usually computed on the basis of frequency of a sine wave. Some structures have a reactance which increases with frequency, and some devices have a reactance that decreases as frequency rises. There are complex devices that favor certain frequency regions but react against frequency content above or below the magic value. The combination of resistance and reactance is called **IMPEDANCE** (symbolized by Z).

Ohm's law for impedance is stated:

I=E/Z

Where I= the current amplitude of the signal*
E= the voltage amplitude of the signal*
Z= the impedance of the circuit at the frequency of the signal.

The moral of all this is that for a given voltage, low impedance circuits require large currents, and high impedance circuits require small currents.

POWER

There is little point in moving these electrons around unless the current can make something happen. How much can happen is expressed by the POWER of the circuit. In a resistive circuit (i.e. no reactance in any of the components) the power is the product of the voltage and the current, and is expressed in watts.

P=EI

For instance, a light bulb that passes a half ampere of current if one hundred volts is applied to it will produce 50 watts worth of light and heat. If the bulb were redesigned to pass a whole ampere, it would provide 100 watts. If 200 volts were applied to the original bulb, it would also provide 100 watts, but probably not for long.

Power calculations in reactive or partially reactive circuits are complex. The ratings on light bulbs and appliances assume you are going to use the resistive power formula to determine whether a fuse will blow, even if the

*The magnitude of a complex number which involves the phase difference between the voltage and current waveforms. Luckily, our meters give these values directly, so we do not have to do any calculations with imaginaries.

device is reactive. (In the latter case the label should read VA for volt–amperes instead of W for watts.)

INFORMATION

When electricity is used to transmit information, we lose interest in voltage and current values* and pay attention to how much information can be represented in a second and how accurate that representation is. We use two basic systems to represent information. In **analog** systems, the amount of current varies directly with the represented quantity, such as air pressure. In **digital** systems, the information is coded into binary numbers, and the value of each bit of the number is represented by the presence or absence of current.

For analog systems, the accuracy of representation is measured as **percent of distortion**; which is the difference between whatever went into the system and the actual output compared to the "perfect" output. The amount of information that may be represented is termed the **bandwidth**. For instance, a high fidelity sound system that can reproduce sound from 20hz to 20,000hz has a bandwidth of 19980hz.

In digital systems, the amount of information is determined by the number of bits that may be transmitted per second, also known as the **baud rate**. The accuracy is the percentage of bits that are transmitted wrong or the **error rate**. Translation of baud rate to bandwidth and error rate to distortion is a function of the encoding scheme used; some are fast and sloppy, others are accurate and slow. Low distortion, high bandwidth data transmission is expensive no matter what system is used.

In both analog and digital audio systems, the representation of sound is usually called the **signal**. We will often trace the **signal path** through various devices, which, because they modify the information, are called **signal processors**. Any output the system produces which is not the signal is considered **noise**.

The following essays in this reader cover all of these topics in more detail.

---

*As long as there is enough of both for the circuit to work.

Decibels And Dynamic Range

The decibel (abbreviated dB) must be the most misunderstood measurement since the cubit. Although the term decibel always means the same thing, decibels may be calculated in several ways, and there are many confusing explanations of what they are.

The decibel is not a unit in the sense that a foot or a dyne is. Dynes and feet are defined quantities of force and distance. (You can go to the National Bureau of Standards and look at a foot or a dyne if you want to.) A decibel is a RELATIONSHIP between two values of POWER.

Decibels are designed for talking about numbers of greatly different magnitude, such as 23 vs. 4,700,000,000,000. With such vast differences between the numbers, the most difficult problem is getting the number of zeros right. We could use scientific notation, but a comparison between 2.3 X 10 and 4.7 X 10^{12} is still awkward. For convenience, we find the RATIO between the two numbers and convert that into a logarithm. This gives a number like 11.3. As long as we are going for simplicity, we might as well get rid of the decimal, so we multiply the number times ten. If we measured one value as 23 hp and another as 4.7 trillion hp, we say that one is 113dB greater than the other.

$$\text{Power difference in dB} = 10 \log \frac{\text{power A}}{\text{power B}}$$

The usefulness of all this becomes apparent when we think about how the ear perceives loudness. First of all, the ear is very sensitive. The softest audible sound has a power of about 0.000000000001 watt/sq. meter and the threshold of pain is around 1 watt/sq. meter, giving a total range of 120dB. In the second place, our judgment of relative levels of loudness is somewhat logarithmic. If a sound has 10 times the power of a reference (10dB) we hear it as twice as loud. If we merely double the power (3dB), the difference will be just noticeable.

[The calculations for the dB relationships I just gave go like this; for a 10 to one relationship, the log of 10 is 1, and ten times 1 is 10. For the 2 to one relationship, the log of 2 is 0.3, and 10 times that is 3. Incidentally, if the ratio goes the other way, with the measured value less than the reference, we get a negative dB value, because the log of 1/10 is -1.]

Converting voltage ratios to decibels

Remember that the dB is used to describe relationships of POWER. Power is not often conveniently measured, especially in electronic devices. Most often we measure voltage and use the formula $P=E^2/R$ to get power. Squaring a value doubles its logarithm, so our dB formula becomes:

$$\text{Power difference in dB} = 20 \log \frac{\text{Voltage A}}{\text{Voltage B}}$$

Power of sound varies as the square of pressure, so this formula is also appropriate for SPL (sound pressure level) calculations.

Reference Levels

The final confusion comes from the concept of RELATIVE power. The question "relative to what?" has no single answer. The standard level (0dB) is chosen to be some convenient value for the application. For acoustics, 0dB often means the threshold of hearing, 0.0002µbar (Microbars: a bar is the "normal" pressure of air). Acousticians deal with positive values and call their measurements dB SPL. Electrical engineers use several meanings for 0dB. They sometimes remember to add a letter to the dB symbol to indicate which is intended.

0 dB	= 6 millivolts at 500 ohms
0 dBj	= 1 millivolt
0 dBk	= 1 killowatt
0 dBm	= 1 milliwatt at 600 ohms
0 dBv	= 1 volt

There are many more. The power calculations must also take spectrum into account: it is not valid to compare a noise signal to a sine wave without some correction factor. The simple rule is to always compare similar signals.

VU

The most common reference used to be 0 dBVU. dBVU is calculated just like dB with some extra restrictions on bandwidth and ballistics of the meter used. The VU (or Volume Unit) system is a hangover from early radio usage when 0 VU meant 100% of the legal modulation for the particular radio station. The level meters were all marked with percentage numbers as well as dBVU, and the numbers above 0 were in red. When tape recorders were invented, the same meters were used, and 0 dBVU came to mean the recommended operating point for the tape in use. The tape manufacturer supplied calibration tapes, and the machines were adjusted to give a 0 dBVU reading on the meter when those tapes were played.

0 VU is not the maximum allowable signal on analog tape recorders. Most tape decks will cope with +6 or even +15 for brief times (such levels might damage the VU meters if sustained) and other devices will go up to +25.

Digital VU

The VU readings on digital recorders are referenced to the full value of the digital word. That means 0VU is as hot as the signal can get! Anything above that will be clipped mercilessly. Usually, digital 0 VU is equivalent to +14 or +20 VU on analog machines[2].

The difference between the strongest signal a system can handle and the typical signal (the old 0VU) is called the HEADROOM.

The minimum useful signal is limited by the level of the ever present system noise. This is the NOISE FLOOR, and may be as high as -40 VU on a cassette deck or as low as -100 VU on a digital recorder.

[2] So, when setting levels on digital machines, aim for –14 or so.

DYNAMIC RANGE

The difference between the clipping level and the noise floor of any piece of equipment is the DYNAMIC RANGE, the useful range of signal levels. (In sales literature this range is called the signal to noise ratio.) The dynamic range concept is useful when you consider a signal that goes through several devices:

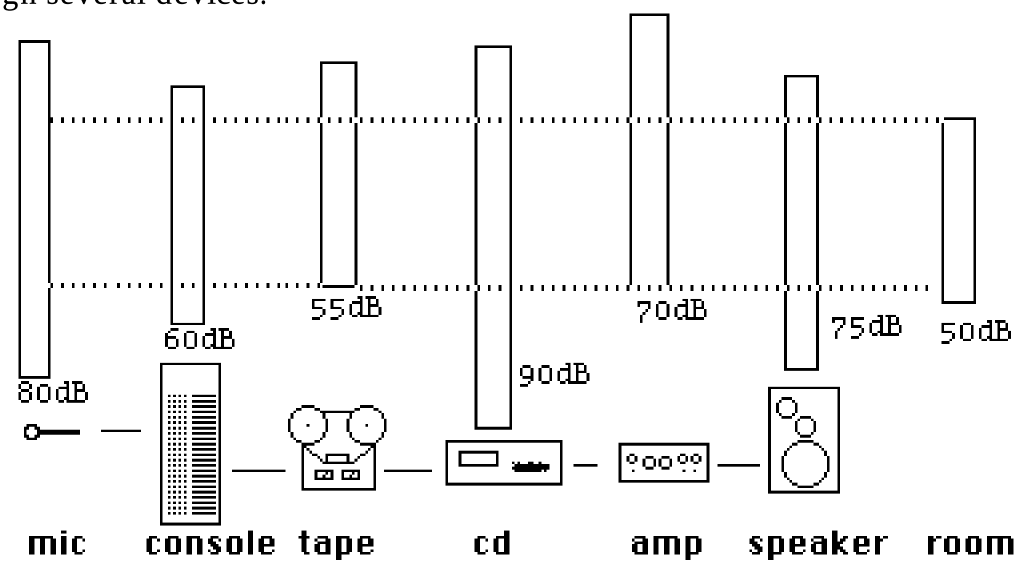


Fig. 1 Dynamics of the recording chain

These dynamic ranges are typical of high quality equipment. Although there is room for improvement almost everywhere, the dynamic limits of the living room (neighbor's complaints minus the noise of the refrigerator) are the most constraining.

The proper use of any device involves getting a match between its own dynamic range and that of the previous device. On the diagram you will notice a mismatch between the console and tape deck. The console will overload before the deck, stealing 5 dB of dynamics. This could probably be avoided by an adjustment of the tape deck input control. Note that most of the capabilities of the CD player are wasted in this system.

Analog Sound Processors

Messing With The Sounds

The devices collectively known as audio processors were originally added to the recording studio to allow compensation for frequency response or dynamic range problems in the equipment. When carefully used, they can add to the fidelity of the sound, and occasionally improve on reality. If improperly used, they can seriously degrade the sound or even produce laughable results. It is possible to use audio processors to manipulate sounds into unrecognizablilty, so naturally they are popular with composers of electronic music.

EQUALIZATION (E.Q.)

An equalizer is a device that can alter the spectral content of a signal. This can be done with any circuit that has an adjustable frequency response, the most familiar being the tone controls on a home stereo set. These tone controls typically affect the amplitude in two frequency regions, the treble and bass. This is sufficient for the minor changes the end user may wish to make in the program, but the recording engineer needs more flexibility and coverage of the entire audio spectrum. The complete studio will have complex e.q. systems that might include 30 or more regions of control. Many of these machines have sliders to adjust the amplitude of each band, and those sliders are laid out in such a way that their positions visually indicate the frequency response. This feature gives rise to the name GRAPHIC EQUALIZER.

A 30 band equalizer is going to be an expensive device, simply because of the sheer duplication of circuitry. Much of this circuitry is wasted most of the time because typical use only involves adjustment to two or three bands while the others are left alone. An equalizer with bands of adjustable frequency is a more economical approach to the problem, because only three to five circuits are necessary. Such an equalizer is called a PARAMETRIC EQUALIZER. Parametric equalizers typically have three controls for each band, allowing adjustment of frequency, amplitude, and width of the band.

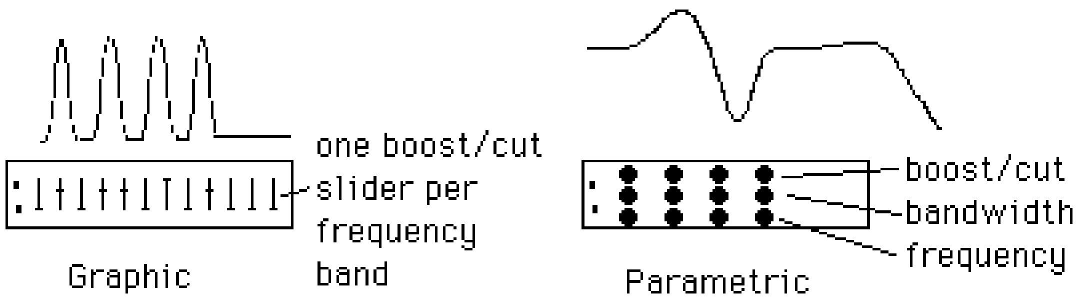


Fig. 1 Equalizers

FILTERS

Equalizers are deliberately designed to create fairly minor changes in the signal. For more drastic effects, such as removing some region of the signal entirely, a FILTER is required. A filter is a circuit that sharply reduces the amplitude of signals of frequency outside of specified limits. The unaffected region is called the PASSBAND, and the filter type is named after the passband as low-pass, high-pass, or band-pass. The point where the signal attenuation just becomes noticeable (a reduction of 3 dB) is termed the CUTOFF FREQUENCY. A low-pass filter with a cutoff of 500 hz will attenuate any signal of frequency above 500hz.

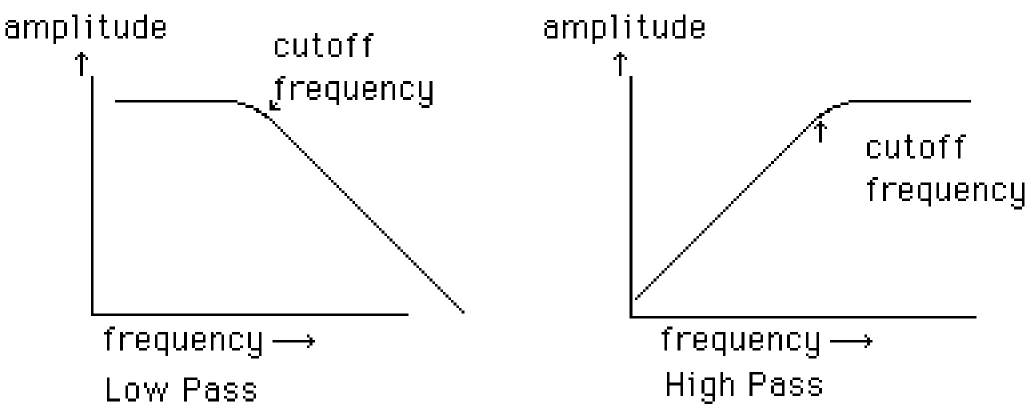


Fig. 2 Filter response curves

The attenuation provided by a filter is never absolute. A graph of the frequency response of the filter shows that the amplitude of a signal decreases as the frequency moves beyond the cutoff. The actual rate of this decrease is a parameter of filter design called the SLOPE. For arcane reasons the slope is always some multiple of 6 dB of attenuation per octave of signal frequency change. (A 6dB per octave filter is really not much more than an equalizer, and the drastic effects used in electronic music generally require 24 dB per octave circuits.) The slope is a general description of response at some remove from the cutoff. Various circuits differ in the shape of the curve near cutoff, and the possible shapes have names such as Bessel and Chebychev, after the originators of the math formulas involved.

Another design parameter which affects the shape of the filter curve is known as "Q". The derivation of Q is too complex for this discussion, but it is useful to know that filters with a high value of Q have an amplitude bump near the cutoff frequency and have a tendency to oscillate at that frequency.

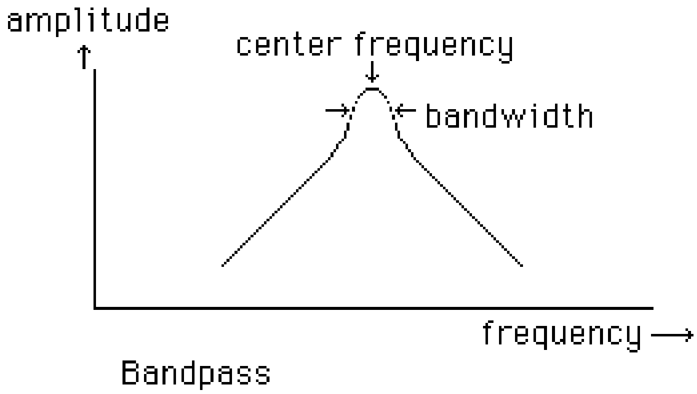


Fig. 3 Bandpass response curve

Bandpass filters have two cutoff frequencies. The difference between these frequencies is the BANDWIDTH, and the mean of the two is the CENTER FREQUENCY. Bandpass filters are sometimes encountered in large groups of fixed frequency circuits similar to graphic equalizers. These filter banks are often called 1/2 octave or 1/3 octave filters after the spacing of the filter bands. Such devices used to be the mainstay of many tape music studios before stable synthesizer filters and parametric equalizers were available. The characteristic pitches associated with those machines is almost a trademark of early 60's tape music.

The 1/3 octave filter was originally designed as a research tool for use in the spectral analysis of sounds. The amplitude of the output of each filter band can be separately measured and those measurements plotted on a graph to show the spectral content of the signal. New versions of the spectral analyzer generate data suitable for feeding directly to a computer for measuring complex time/spectrum relationships.

The most complex filter system I have heard of is the vocoder. This device contains a filter bank set up for spectral analysis of a signal and a similar filter bank set up to process a second signal. The measurements derived from each filter in the analysis bank are used to control the amplitude of the signal from the corresponding filter in the processing bank. This system will impose the spectral shape of the analyzed signal onto the processed signal. The effect produced is quite striking, especially if voice is used as the sound for analysis.

LIMITERS

It is a very common requirement of electronic devices that the amplitude of the input signal must not exceed some value. The typical result of failure to obey that rule is extreme distortion and (if the device in question is a radio transmitter) a stiff note from the FCC. If the operator of an audio system cannot be trusted to keep the signals within bounds, it is necessary to include a LIMITER, a device which turns the gain down whenever the signal exceeds the limit.

The action of reducing the amplitude of a signal can easily become a form of distortion, so the design of limiters is a delicate art. The usual approach is to set up the circuit so the transition from free to limited operation is somewhat gradual, and so that the transition back (called the RELEASE) is even more gradual. Appropriate release times vary with the nature of the program material, so this parameter is often adjustable. The limiting point (called the THRESHOLD) is of course also adjustable.

COMPRESSION

It is also possible for the soft spots of a program to become too low in amplitude, especially in a noisy medium like AM radio. A program with a wide dynamic range may have to have that dynamic range reduced in order to fit a particular audio system. (Surprisingly, the switch to digital audio is made this problem more common as stations try to fit the wide range of compact discs through the restricted dynamic window of FM transmitters.) A device which reduces the overall dynamic range of a signal is called a COMPRESSOR.

A compressor works by measuring the average amplitude of the input signal and using the information from that measurement to control the gain of an amplifier. If the measurement comes out above an adjustable threshold, the gain is reduced, if the measured input amplitude is below the threshold, the gain is increased. The amount of effect is usually adjustable and is expressed as a ratio of input dynamic range to output dynamic range, such as 2 to 1 or 3 to 1.

The release time is adjustable just as with the limiter, but the adjustment is much more critical, because the compressor is constantly changing gain whereas the limiter only cuts in occasionally. It is impossible to find a release time that is perfect for all situations. A very short release time will cause the circuit to attempt to trace the waveform of low frequency components of the signal, resulting in intermodulation distortion of high frequency components; whereas a long release time will leave the amplifier in the wrong mode if the signal changes quickly. That second error results in brief bursts of expansion, a very disconcerting effect. (Sounds like heavy breathing.)

EXPANSION

Expansion is the opposite of compression. It is accomplished with the same circuit as compression, but the rules are changed so that any signal with amplitude above the threshold will be amplified and any signal with amplitude lower than the threshold will be attenuated. Expansion is used primarily for special effects, such as super loud taco chips.

NOISE GATING

Sometimes a program is encountered which has proper dynamic range but which has objectionable background noise. Very little can be done to reduce the noise without affecting the signal, but the spaces where there is no signal can often be cleaned up with a NOISE GATE. This circuit is an expander that only works on the low amplitude parts of the signal. Any section that is below the threshold will be attenuated rather drastically, so the space becomes nice and quiet. You can hear this device in action on old movies on TV. Release time adjustment is very critical in this application also, because if the circuit is too slow, the turning down of the noise will be painfully obvious, but if the circuit works too fast, the reverb at the ends of the sounds will be lost.

ARTIFACTS OF COMPRESSION SYSTEMS

Once you have learned to hear such things, you will be able to spot compression and expansion most of the time they are used. That’s because no dynamics processor can work right with all music. It can’t react instantly, otherwise it would be fooled by low frequency signals, and produce an awkward sound known as PUMPING. On the other hand, if it reacts too slowly, the attacks of high frequency sounds would be let through unaffected—this adds PUNCH to a sound, making flutes and such sound brittle, and can produce distortion with the louder instruments. (On the other hand, this is sometimes desirable, and is commonly done with bass and drums).

At the end of a sound, the compressor will tend to chop off the ends of notes or exaggerate the noise between notes. When you can hear the noise between notes change level, the effect is called BREATHING.

These artifacts are controlled by adjustments on the compressor that must be tweaked for the type of music coming through. There are rather expensive multiband compressors that minimize these effects. When compression is done tracks recorded in a computer , these effects can be avoided altogether.

NOISE REDUCTION

If you happen to know ahead of time that you are going to pipe the signal through some noisy medium (such as analog tape), it is possible to reduce the effect of the noise by compressing the signal before recording and expanding afterward. The DBX NOISE REDUCTION SYSTEM does just that with a 2 to 1 compression and an exactly complementary expansion. The DOLBY noise reduction system goes one step better by using filters to divide the audio spectrum into regions and applying a compression/expansion cycle to each band. A recorded signal that has been treated with the first half of one of these systems is said to be ENCODED and sounds rather strange if it is played back without DECODING.

Distortors

Perfect fidelity recording is the ultimate goal of recording engineers and audiophiles, but many musicians have found charm in deliberately distorted sounds. For instance, if you clip or rectify the sound of a guitar the nasty buzz that results will easily be heard against a very loud accompaniment. In fact, slightly distorted sound can give the impression of being loud when it actually isn’t. (If you listen to renaissance music on things like krumhorns and rebecs, you will realize this is not a new idea.) Guitarists and others routinely carry an array of effects devices, many of which came straight our of electronic music studios. Since they are often built into foot pedals, they are often called “stomp boxes”.

A more subtle distortion is commonly applied to pop vocalists. Called “exciters” these devices add a slight edge or third formant to a singers tone, which helps make lyrics understandable.

Another common type of distortion has become popular with the advent of the digital age. Surprisingly, that is the sound we were trying to escape with the conversion to transistors and digital, the often bizarre distortions introduced by vacuum tubes! You can now buy very expensive versions of everything from microphone preamplifiers to power amps that feature the “warmth” of tubes. Most have actual tubes in them, but some simulate the tube sound via digital algorithms. Admittedly, when you are not using a whole series of them to provide gain, and when the circuit is well designed, the subtle effect of a single tube can produce a charming sound.

REVERBERATORS

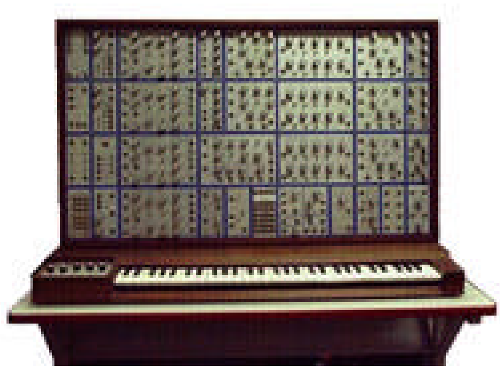
Recording is usually done in tiny soundproof studios (or garages, not soundproof, but equally tine), which produces tracks with dry, non reverberant sound. A device that would add a convincing concert hall like reverberation to such recordings was for a long time the holy grail to audio device manufacturers. Down the years many wonderful gadgets were invented, from Hammond’s slack spring to giant steel plates. The problem has been more or less solved with digital reverberators, but it is interesting to note that the sound of the classic devices is still sought after for its own sake—in fact, the settings on digital units almost always include “Plate” and “Spring”.

Spring and Plate reverbs worked pretty much as you might imagine, with the sound being injected into the device at one and picked up by a transducer at the other. If you put your ear to one, you can hear the signal being processed, and it is possible to get feedback through them from monitor speakers. Also the sound you get when you kick one is awesome!

See essay 9 for more about reverbs and other digital effects processors.

The Analog Synthesizer

(PATCH: A particular connection of modules on a synthesizer.)



An analog synthesizer is a collection of electronic circuits that are useful for generating and manipulating signals that will ultimately be heard as sound. In the classic MODULAR SYNTHESIZER each circuit is separate and interconnections are left up to the user. In the PERFORMANCE SYNTHESIZER the circuits are already interconnected in a pattern the manufacturer thinks is most likely to sell. This discussion is going to deal with modular synthesizers.

Each module of a synthesizer is designed to fill one of three functions. These are: SIGNAL GENERATION, SIGNAL PROCESSING, or CONTROL of the other two kinds of module. (You may notice a resemblance to the parts of acoustical instruments: driver, resonator, and pitch control mechanism.) The manner in which the modules perform their functions can be controlled in two ways: by adjustment of panel knobs or by application of a voltage to an appropriate input jack.

SOME MODULES

The principal signal generator of any synthesizer is the OSCILLATOR. This device produces a steady signal at a frequency determined by the combination of the knob settings and the input voltage. There is usually a choice of waveforms at various output jacks; most often available are sine, sawtooth, triangle and pulse. The pulse often has an adjustable duty cycle and sometimes a blend of the other waveforms is provided.

The most important processors on the machine are the AMPLIFIERS and the FILTERS. These modify the amplitude and spectrum respectively of any signal applied to their signal inputs. The gain of the amplifier and the cutoff frequency of the filter are the parameters controlled by the knobs and voltage input jacks.

The basic controller is the ENVELOPE GENERATOR. In its simplest form, this module produces a voltage that changes relatively slowly, rising from 0 volts to some preset high value, stops at that value for a time and then falls back to zero. The rate of rise and fall are adjustable and often voltage controlled. A slightly more complex version that is very widespread has a drop to an intermediate voltage level just after the initial rise. This device is fundamentally different from an oscillator in that it only performs its function once upon command. This command is usually in the form of a brief voltage pulse called a TRIGGER. Most envelope generators also require duration information. This is provided in the form of a voltage that stays high as long as the generator is supposed to stay at its upper or intermediate level. This abruptly changing voltage is called a GATE. (Some brands of synthesizer do not use triggers, since almost the same information can be derived from the leading edge of the gate.)

THE FIRST PATCH

This group of four modules is sufficient to produce a variety of useful sounds. The traditional way of connecting them is to patch the output of the oscillator to the signal input of the amplifier, the output of the amplifier to the signal input of the filter, and the output of the filter to the playback or monitoring system. The output of the envelope generator is applied to the control input of the amplifier. When the envelope generator is triggered, the result will be some sort of beep with pitch and timbral aspects determined by the adjustments of the modules. The one required setting of this patch is that the filter cutoff frequency be set so that the oscillator frequency falls within the filter passband. If silence between notes is desired, the amplifier gain control must be set so that the signal is shut off when the output of the envelope generator is zero volts. A very different sound can be produced by connecting the envelope generator to the filter. The sweeping of the peaks in the filter response create a characteristic sound often described as "vocal". If a low pass filter is used it is possible to dispense with the amplifier in this variation.

HOOKING IN THE MUSICIAN

The described patch will generate enough different sounds to produce a fairly complex composition (Stockhausen produced Kontakte with a similar setup), but preparing each note using only the adjustment knobs is a bit tedious. This problem is overcome with a set of control modules that are designed to provide a comfortable INTERFACE with the musician.

The KEYBOARD is the most familiar module in the system. In a way this is a disadvantage, because the traditional keyboard implies capabilities and a manner of performance that are not necessarily appropriate for synthesizers. The keyboard module has two outputs; a steady voltage corresponding to the last key pressed, and a gate voltage that is at some high value whenever any key is pressed. The obvious connection to the above patch is to apply the voltage output to the control input of the oscillator and the gate signal to the envelope generator. (If the particular system has envelope generators that require triggers the keyboard will provide those also.) The output of the keyboard and the input of the oscillator are very carefully calibrated so that playing the keys will produce the "right" pitches; i.e. 12 equally tempered notes per octave. This system will play melodies, but any attempt at playing chords will give unpredictable results.

The keyboard should not be irrevocably associated with pitch. Since its output is a control voltage, the keyboard can be used to control anything that requires arbitrary stepwise modification. This can be volume, filter settings, tempo, or conceivably even temperature. (The "Bay Area School" of synthesizer builders try to reduce traditional associations of the keyboard to pitch by providing keyboards with touch plates instead of the usual white and black keys.)

There are many alternate interfaces available, some of them quite weird and wacky. There are touch ribbons, joysticks, wheels, theremin antennae, springy bits of metal, and foot pedals, not to mention imitations of traditional instruments such as clarinets, trumpets, and guitars. I have yet to see a faucet handle used as an interface, but that doesn't mean it hasn't been done.

MORE MODULES

The system described so far is the least you would expect to find in any synthesizer. The range of timbres available is wide but not inexhaustible, and these tones would not be described as subtle. The book-perfect harmonic structure of the geometric waveforms the oscillator produces is a very unrealistic and ultimately boring sound.

There is a need for a palette of non-harmonic timbres in a synthesizer, but it turns out that a simple module to produce them directly is not available, at least not short of expensive digital techniques. It is possible however, to combine basic modules to generate such sounds. Here are a few procedures used in the quest for anharmonicity:

There is a module that produces WHITE NOISE. This is often mixed into the sound in small amounts or short bursts to imitate some percussive timbres.

Two or more oscillators may be used in parallel for ADDITIVE SYNTHESIS. This is most effective if each oscillator has its own amplifier and envelope generator, but simple combinations are useful.

If an oscillator is connected to the control input of an amplifier that is processing another oscillator, two interesting possibilities arise. If the control oscillator is running at a sub-audio frequency, a vibrato of amplitude is produced, similar to that performed on wind instruments. If the control oscillator is running at audio frequency, the result is a mix of the two frequencies, plus extra components at frequencies equal to the sum and difference of the two. This technique is known as AMPLITUDE MODULATION.

If an oscillator is connected to the control input of another oscillator, again there are two possibilities. If the control oscillator is sub-audio, the sound will have a frequency vibrato, similar to the performance practice of the violin. If the control oscillator is running at audio frequency, the result is FREQUENCY MODULATION. The output is a very complex spectrum, consisting of the original frequency plus a set of SIDEBANDS, which are components above and below the original spaced at intervals equal to the frequency of the control oscillator. The number and amplitude of the sidebands is determined by the amplitude of the control, according to a mathematical distribution known as a Bessel function.
There is a special module called a RING MODULATOR which produces the sidebands associated with amplitude modulation but suppresses the original frequency content. The output is a very messy spectrum, especially if one of the input signals is harmonically complex to start with. A further refinement of the same device suppresses sidebands above or below the input signal, producing FREQUENCY SHIFTING. (This must not be confused with pitch shifting as discussed in the essay on sampled processors. The upper sidebands are produced by adding some constant value to the frequencies of all components of one input, pushing the whole spectrum off of the harmonic series. The pitch shifter

effectively MULTIPLIES the frequencies of the components by a constant, maintaining the harmonic relationships.)

You noticed that I have been throwing the word MODULATION around rather freely. This use of the word is derived from radio broadcasting and is directly related to the usual definition of modulate as "to change". In electronic music, modulation is used to describe any audio frequency manipulation of the signal.

OTHER FORMS OF PATCH

The patch described models the loosely coupled form of musical instrument (See essay 13). This is the most common use of synthesizers, simply because this is the built in patch that comes with most non-modular machines. It is perfectly reasonable to model the other forms of instrument as well. Please note that modeling an instrument does not mean a patch will sound like the instrument. What we are copying is the relationship of sound producing and modifying mechanisms. The timbre of any instrument depends on the construction of such structures.

A synthesizer patch may model the tightly coupled variety of instrument if some FEEDBACK to the oscillator from some point further along the signal path is provided. Since the input to the oscillator is for control, the result is not the "howl-round" you might expect, but a change in waveform and frequency. If the feedback is controlled and processed the results can be subtle and delicate.

To model the impulse resonant instruments, we do something unusual for audio electronics; we find a poorly performing circuit and use it in the "wrong" manner. The engineers who design filters go to a lot of trouble to produce circuits that have a smooth and uncolored response in the passband and do not ring when a square wave is applied. For a percussive sound, we need a filter that does ring, and a funky passband lends character. The engineers have compromised their principles far enough to allow the user to "de-adjust" part of the filter circuit by a panel control labeled RESONANCE. As this knob is turned, the filter starts emphasizing signals near the cutoff frequency and shows a tendency to ring at the cutoff frequency when impulsive inputs are applied. If a single pulse such as the trigger from the keyboard is connected to the input of such a filter, the sound produced is quite percussive, with a sound similar to a wood-block. This basic sound may then be processed into a variety of timbres.

MORE CONTROL

The use of performer interfaces to control the synthesizer produces a strong implication: the synthesizer is intended to emulate or extend existing instruments. In fact, this is a very limited view of the synthesizer. To the sophisticated synthesizer user, the machine is a powerful composing tool which can produce gestures, segments, or even whole compositions. To fulfill this promise, the synthesizer must be fairly large, and must include some modules designed for PROGRAMMED CONTROL of the system.

The best controller for a synthesizer is a computer of course, but it is worth considering a few obsolete techniques because many computer systems emulate them.

> A LOW FREQUENCY OSCILLATOR produces a cyclical voltage change that may produce long swells and fades if connected to an amplifier and slow timbral changes if connected to a filter. Many pieces have been composed that use several of these devices for the main formal control. A slight variation is the RANDOM VOLTAGE GENERATOR which avoids the periodicity of the LFO.
>
> The SAMPLE AND HOLD, borrowed from digitizing systems and run at a very low sampling rate, can produce quite complex patterns of voltages as the audio frequency signals it samples are aliased. If noise is sampled, the output is very random. The subtlety in this technique comes from well thought out control of the sampling times.
>
> The SEQUENCER allows the composer to set a pattern of voltages by tuning a row of knobs. An internal switching system applies one setting at a time to the output. Usually when the last control has been used the machine starts over with the first, a feature that has generated many mindless ostinati. Again the subtlety comes from careful control of rhythm.

You can find even more modules by visiting
http://www.obsolete.com/120_years/

Sampled Sound Processors

There are times when we want to store a sound briefly without resorting to tape or some other permanent medium. An example is the simulation of reverberation, where there are various delays of the sound as it reflects off distant walls. There is some propagation delay in electronic circuits, but that delay is of the order of five microseconds per mile, and is not very useful. The response delay of some circuits is longer, but is by nature frequency dependent.

Analog delay

The contemporary solution to this problem is to store samples of the signal rather than the whole thing. It is fairly easy to construct a circuit that samples a changing voltage and keeps the value it obtains as a constant output. (That circuit is called, not surprisingly, a SAMPLE AND HOLD, and will turn up again in the discussion of analog synthesizers. The circuit that times the sample is called a CLOCK). If we had a lot of those circuits, it would be possible to store a series of samples and then read them back, reconstructing the signal much in the same way that a motion picture uses a series of still photographs to reconstruct a moving image.

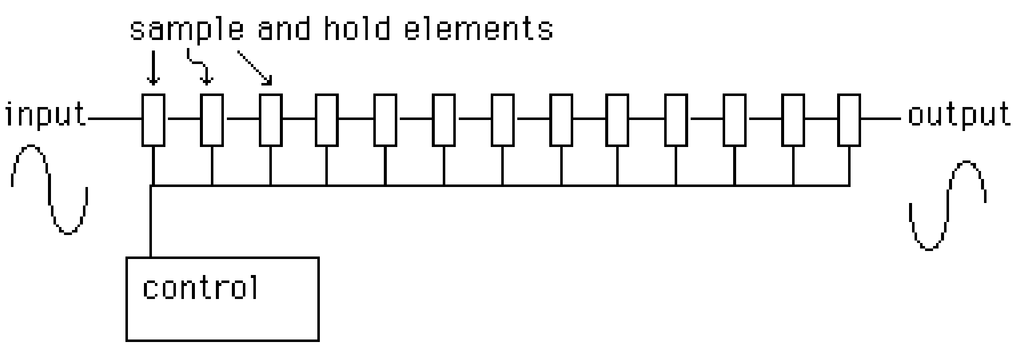


Fig. 1 analog delay

The simplest way to hook up all of these sample and hold circuits is in series with a common control circuit, so that on each clock pulse the last S&H (number N) samples the output of the preceding one, then number N–1 samples number N–2, and so forth, until the first S&H samples the signal. The signal value then marches through the circuit one stage per clock pulse. The output of the final sample and hold is read as the delayed signal. (This series of sample and hold circuits is called a BUCKET BRIGADE.) Building large numbers of identical circuits is a strong point of integrated circuit technology, and we can now buy bucket brigade chips with more than 4,000 elements.

The amount of delay we can get with a given length brigade depends on how often we sample the signal. If the signal were sampled once a second, a 4000 unit delay line would give a total delay of slightly more than an hour. Our one per second sampling rate does not give us sound however, since sound contains information changing at the rate of 20-20,000 hz. You should remember from the essay on digital audio that the 20,000hz upper limit of hearing implies a Nyquist frequency of 40,000hz. At that rate, our 4,000 unit bucket brigade yields 1/10 of a second delay.

Bucket brigades longer than 4,000 units are not very practical because each transfer has a tiny error, and those tiny errors accumulate into serious noise. For this reason bucket brigade circuits are limited to short delays or low fidelity systems such as guitar effect boxes.

Digital Delay

The sample and hold circuits are really analog memories. If we replace them with digital memory circuits, we can eliminate the transfer errors and build very long delays. Of course we are now subject to all of the errors inherent in A-to-D and D-to-A conversion and have to make the same trade-off between noise and word size as discussed in the essay on digital recording.

Digital delay systems are constructed from chips originally designed for computer use. These chips contain parallel access circuitry, so the bucket brigade configuration is not used. The memory locations are numbered, and a counting circuit, called a POINTER, controls which location is connected to the A-to-D or the D-to-A. The pointer is designed so it starts over with zero after it hits the highest location, so the whole process is best envisioned as circular:

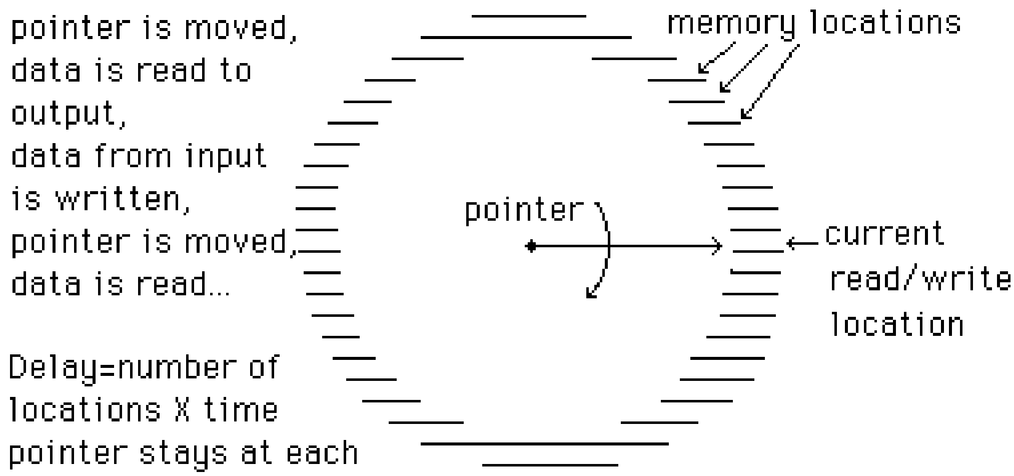


Fig.2 How a digital delay is organized

Operation of the system is controlled by a complex clock circuit. At each sample time, the location indicated by the pointer is connected to the D–to–A, and the value at that location is read out. Then the same location is connected to the A-to-D and a new value is written in. After that the pointer is moved up one value and the system waits for the next sample time. Since a value will wait in its location until the entire memory has been filled, the delay time is determined by the number of memory locations available. As in the analog delay, minor changes in the clock rate will fine tune the delay time.

Large changes in the delay time can be made by using separate pointers for the read and write functions. When that is done, the delay time is determined by the difference between the two pointers.

Pitch Changing

There are some interesting tricks that can be done with a digital delay system. For instance, suppose we stop the data writing process but continue to read. The output will be the same series of values over and over, and part of the signal will repeat endlessly. The result is the same as a short tape loop with a splice at the point where the writing stopped. (The quality of the splice is a matter of luck, and generally produces a recurring pop or GLITCH.) Now if you adjust the clock rate, the pitch of the output signal will change, just as it would if you changed the speed of a tape recorder.

The pitch will change any time you adjust the clock in a sampled system. This change is similar to doppler shift, and is caused by the difference between the writing rate and the reading rate. Usually the change is brief, lasting only until the memory is filled with samples taken at the new rate.

We can make the pitch change work continuously if we start the writing circuit again at the original sampling rate, and continue to read at the changed rate. The delay time will change as the read and write pointers pull away from each other, and there will be a glitch whenever the two pointers cross (most commercial pitch changers perform some simple signal modification to avoid this glitch as much as possible).

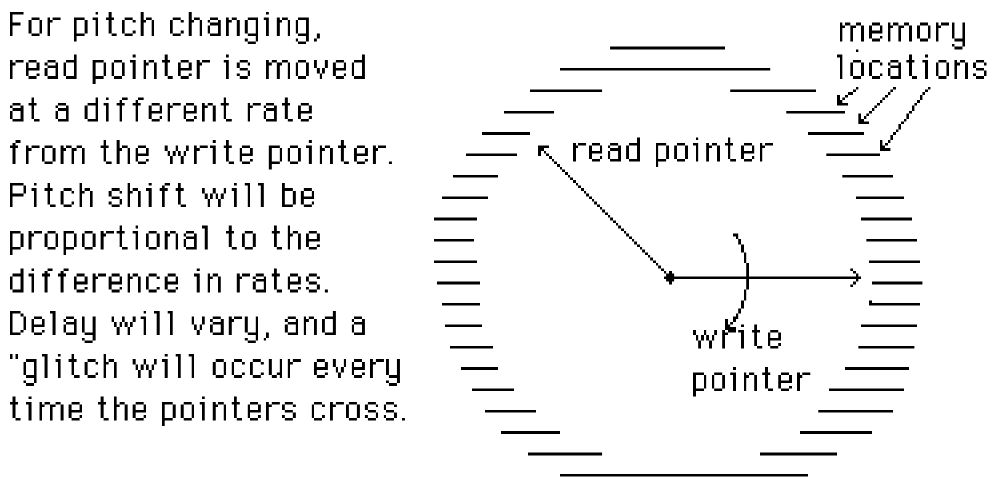


Fig 3. Changing pitch

Reverberation

To use a delay as a simple reverberator, we apply a little of the output signal to the input along with the new material. (Any time the output of a device is connected to its own input we have FEEDBACK. Feedback must be carefully controlled- you probably can name examples of uncontrolled feedback.) Then if a short signal is applied to the system, the sound will cycle through the system a few extra times before dying out. If you do this with only one delay the result will be like the reverberation of a rain barrel or shower stall, but it will be reverberation nonetheless. More realistic reverberation can be generated using several delays of different lengths. Delays constructed with reverberation in mind feature multiple outputs and are called TAPPED DELAYS.

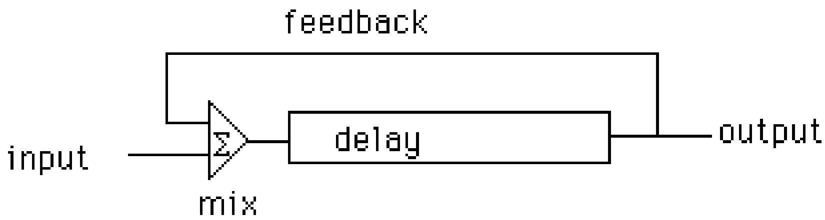


Fig 4. A simple reverb.

Digital Filters

Delays can be used to modify the frequency content of a signal. To do this some of the undelayed signal is added to the output of the delay. (This is called FEEDFORWARD.) In the essay on spectral analysis it was noted that if sine waves of equal frequency but different phase are added together, the result is a sine wave of amplitude determined by the phase difference, that amplitude being zero at a difference of 180 degrees. If you mix a complex signal with a delayed version of the same signal, some components of the signal will have the equivalent of a 180 degree phase shift caused by the delay, and will be suppressed. This will happen for any component with a half period equal to an even fraction of the delay time. This means that for any delay time d, there is a frequency F (such that F=1/2d) that will be filtered out of the output. In fact, the entire series F, 2F, 3f... will be missing. If the input is inverted before the combination occurs, F will equal 1/d. This process is called **COMB FILTERING**, because a graph of the frequency response looks like a comb.

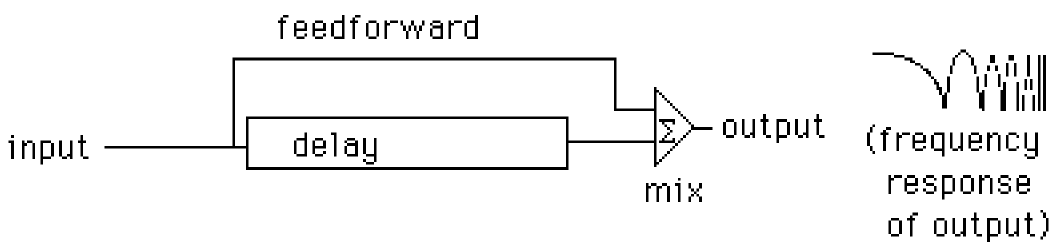


Fig. 5 Comb filter.

If you build a delay and add both feedforward and feedback, you can construct other kinds of filters. In fact, any kind of filter can be specified by adding multiple delay units and carefully controlling the amounts of feedforward and feedback. The basic unit of a digital filter is called a **pole**: this example has two of them.

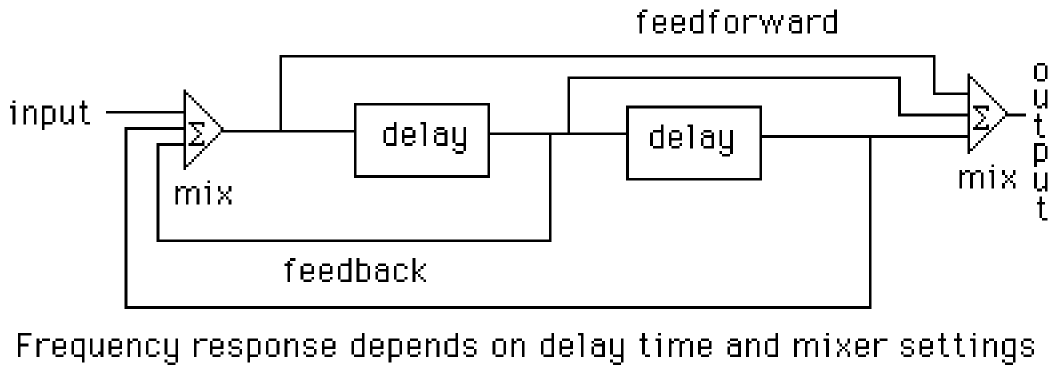


Fig. 6 Digital filter (simplified)

Digital filters are actually programmed in computers rather than built in hardware. The techniques used are a bit advanced for this discussion.

An Overview of Computer Music

Any computer music system has three major components: the **host** computer, the **synthesis hardware**, and the **programs** that make the two work together. Systems may be classed as research, proprietary, or MIDI based.

Reseach Systems

Almost all **mainframe based** music systems were found at research institutions, primarily large U.S. universities. These were set up during the 60's and 70's to explore computer applications in music and acoustics. As a rule, they were used for development of new tools for composition and more efficient techniques of sound generation, and were financed by corporate and government grants. Hardware usually consisted of a very powerful central computer with a custom made synthesis box, such as the IRCAM[3] 4x, which had the ability to generate 1024 independent sine waves in real time with 24 bit resolution. Users worked at terminals that might be several miles from the mainframe. Most operations required some compilation time before they can be heard, ranging from minutes to hours.

Mainframes have pretty much disappeared in favor of networked workstations, but much remains the same:

The operating system is usually UNIX, and composers as well as programmers work in extendable high level languages such as C and LISP. It is still not unusual for a researcher to spend two or three years writing code for a two or three minute piece, but the mature institutions also have their share of productive composers. Often people who are strong programmers team up with those who are principally musicians for collaborative works. Each lab has its own software library, but there is enough communication and compatibility between installations that musicians trained at one can work at another with a little training. Some of the widely used research packages are Common Music developed at Stanford, and Csound from MIT.

The difficulty of learning these systems and the awkward operating conditions are justified by true state of the art performance. With the ability to generate four channels of 24 bit sound at 100khz sampling rates, the sound produced is exquisite, and the procedures available to the

[3]Institut de Recherche et Coordination Acoustique / Music, Paris.

composer are breathtaking. Imagine being able to record a trombone solo and have it gradually evolve into soprano voice, or to explore the reverberation of a room with constantly changing dimensions!

Luckily, we do not have to move to Palo Alto to take advantage of the most promising of the developments of the big labs; we need only wait, and the technology will come to us in a conveniently packaged, affordable form. The magazines are bursting with ads for systems that equal the best imaginable ten years ago, and the current research projects will come to the marketplace even faster.

Proprietary Systems

These are complete packages produced by a single manufacturer, with no attempt at compatability with any other product. In the BM (before MIDI) days, there were many of these, and the studios of the world are littered with their carcasses. Two such products survived the revolution, and now that the realization is beginning to dawn that MIDI is not the ultimate answer, a few new systems are beginning to appear.

The Synclavier was originaly developed at Dartmouth, and was the first commercially available FM machine. The first product was rather simple, but it got some well timed support from a few Hollywood performers that financed the development of a rather flexible and nice sounding machine. It is modular in design, so you could start with the basic device and add extra voices, sampling, MIDI, and so forth as your budget allowed. The full blown system included something called the "tapeless studio", in which tracks are recorded on hard disks and can be edited and processed very quickly and accurately. Synclaviers are no nonger produced, but most are still running.

The Fairlight was the first commercial sampler. It was developed in Australia, and like the Synclavier, first caught on with people scoring TV movies. Fairlight is also an expandable system, and although it is still a sampler at heart, it features some powerful sample generating software and extremely high quality sound. The most recent versions are MIDI controlable to some extent. Both the Synclavier and Fairlight are rather expensive[4], but many recording studios have them installed in a small suite. Composers develop pieces on economy instruments and then rent studio time to produce the finished product with high class sounds.[5]

---

[4]$15-50,000.

[5]Unfortunately, the makers of both Fairlight and Synclavier are no longer in business. The instruments are supported by user's groups.

Kyma is a realatively new addition to the propriatory crowd. It is very reasonably priced, and interfaces well with the MIDI world. It's really just a single box (called the Capybara) and a smalltalk based programs that can run on any platform. The heart of the Kyma system is a series of dedicated audio processing cards.

Most systems currently for sale are designed as recording systems and are known as Digital Audio Workstations or DAWs. They consist of high speed analog to digital and digital to analog converters designed to interface with a fast computer with lots of hard disk space[6]. The software that comes with these systems is for digital editing of recordings stored on disk. Because there is no rewind time, editing in this manner is very efficient, and unlike analog editing, it is very easy to undo splices that don't work out. One interesting feature of this kind of system is they all embrace software plugins, so if there's some function of Cuebase you would like to use in your Digidesign system, it can probably be arranged. There are now a lot of software synthesis programs offered, most of them emulating old analog systems.

MIDI

MIDI based systems are the current workhorses of studio and stage. Because the standard is well defined and supported by most manufacturers, individual composers can assemble a setup that fits their own needs and budget. New functions are being added every day, and the price keeps going down!

[6]About 200 kbytes of storage per second of sound. The big systems accomodate four or five hours.

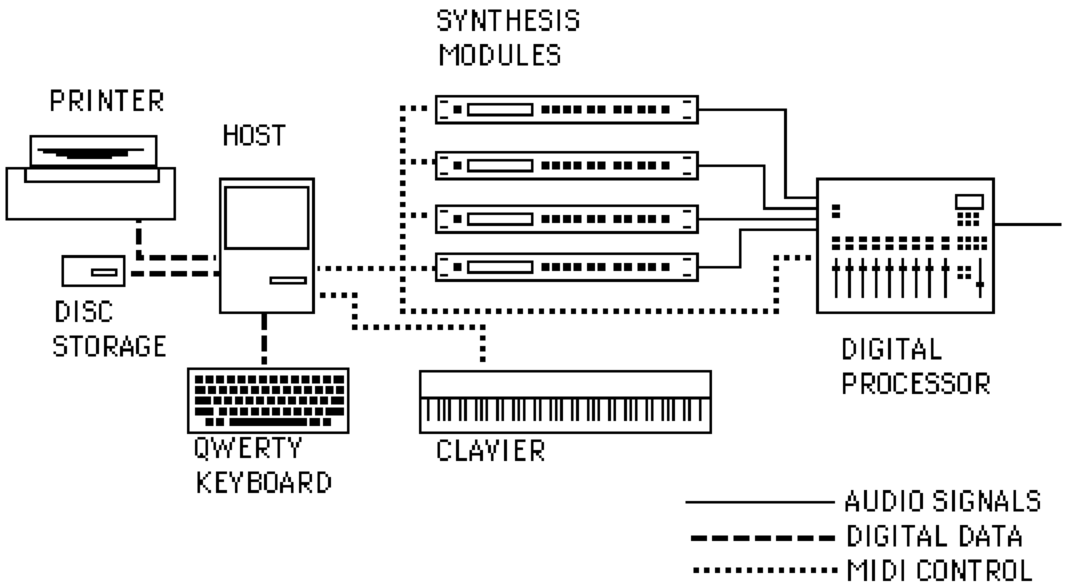


Fig. 1 A typical computer music system (MIDI based)

Figure 1 shows the relationships of the various pieces of hardware you might find in a large system. The host computer controls the operation of the other devices, sending commands and data that shape the music as it is happening. The synthesis modules (which are powerful computers in their own right) generate the audio signals, which may then be processed under further control of the host. The composer communicates with the host through a "QWERTY" (typewriter) keyboard, and a variety of interface devices, of which the most common is the clavier (organ keyboard). The host also performs the usual computer functions of data storage and output on printer or screen. An actual setup may not seem to have this many pieces, as various devices such as clavier and synthesis modules are often combined in a single package. In fact, it is possible to buy the entire system, including permanent software, in a single box called a "MIDI workstation".

- See essay 18 for more detail about the MIDI protocol.

The computers used for hosts in MIDI systems are usually Macintosh or Windows, but you can get MIDI working on LINUX.

MIDI SYNTHESIZERS

Twenty years ago the portable keyboard synthesizer was little more than a toy, useful for density and special effects in pop performances, but not much else. True power for interesting synthesis was found only on modular systems, which while expensive, were fundamentally economical-- if you had eight oscillators, you could configure them to produce eight simple voices, four modulated voices, or one voice with detailed control over eight partials. You could do any of that with keyboard systems, but each approach required a different synthesizer. As the price of portable synthesizers came down, it became possible for a musician to own a stack of several different brands of machine, and it became possible for the keyboard instrument to provide the sound quality and some of the flexibility of the modular systems. In fact, the concept eventually returned to modularity, with separate tone generators, keyboards (or other controllers), processors, and mixers. The difference is that a new generation MIDI module is much more powerful than any analog synthesizer ever was.

There are several different approaches to synthesis being manufactured today. **ANALOG** synthesis is still popular, although modular systems are rare. There is a growing trade in refurbished "classic" machines such as the MiniMoog and arp Odessy.

Most systems utilize a digital emulation of the classic analog synthesizer, using oscillators that work by the wavetable[7] method to create a number stream that is fed through the digital equivalent of a VCA (a multiplier) converted to an analog signal, and then filtered. Analog filters are used because digital filtering is pretty expensive, and an analog filter is always required at the output of a digital system anyway. This system, like the analog system, produces fixed, perfectly harmonic waveforms.

There are a lot of variants on this approach, as each manufacturer strives to appear to offer unique instrumens. One common feature is the use of wavetables based on sampled sounds. The sounds are in read only memory chips, which may be changable, but are not recordable, so these systems are not true samplers. Some very successful drum machines are designed this way, but it is a rather limited approach for a keyboard synthesizer.

A few instruments have been introduced that use additive synthesis, but they have not been widely successful, probably because they are hard to program. Several wavetable machines use a Fourier program to design the waveform, but they do not offer dynamic control of each partial.

The Yamaha DX and TX line of synthesizers implement FM synthesis. In Yamaha's system, digital oscillator, amplifier, and envelope generator are grouped together as an **operator.** The oscillator produces a sine wave, which is modified by the amplifier (digitally speaking, a multiplier)

---

[7]See essay 17.

according to the envelope. The power of this system is that we don't simply listen to the output of the operators. Each voice may have as many as eight operators, and there are 32 available configurations of those operators, most of which involve modulation of one by another. These machines have a very large palette of sounds, with a decided preference for the non-harmonic. DX and TX machines are history, and the patent has run out on FM systhesis, but yamaha still releases an FM machine from time to time.

Many companies are making systems that feature **SOUND SAMPLING.** The first sampling machine was a device called the Mellotron (made in the sixties), which contained thirty or so tape loops with the desired sound recorded at different pitches. The contemporary machines work much the same way, but digital recording technology gives a very high quality playback. The best allow editing and other modifications of sounds you record yourself. A few systems have the possibility of inventing sounds on a computer, and then transferring those sounds to the sampler for easy performance. The most popular are play only samplers, chock full of imitations of traditional instruments.

A new wrinkle called **Vector Synthesis** or AWM is really an all of the above type. Short samples are combined with FM waveforms under control of complex envelopes. This gives a pretty rich and familar sound at low cost.

The newest thing in digital synthesizers is **modeling**. This requires a lot of computing power, so the machines are either expensive or limited in polyphony. A modeling synth uses a description of the physical characteristics of an instrument to calculate sounds. For instance, with a flute algorithm, you specify the size of the tube, which holes are open, the breath pressure and similar things. The synthesizer evaluates the equations that describe the sound producing process and generates a tone. Surprisingly, the most popular modeling synthesizers model the circuit behavior of old analog synthesizers!

MIDI PROCESSORS

We also have MIDI controlled digital sound processors. The advantage of a digital processor over an analog device is that the dsp is reprogramable: if you buy an analog equalizer and an analog compressor, you come up short if you need two e.q.s, but if you have two d.s.p.s they can both be e.q.s, compressors, reverbs, or almost anything else[8]. The MIDI part is that

[8]This is the ideal- the reality as of 1989 is that dsps offer at most three or four functions. However, there is a lot of variety in the

the adjustments of the device can be stored in the host computer and quickly changed in the middle of a piece.

MIDI connectors are now found on mixing consoles. Not only is it possible to capture and recall all of the e.q. and track assignments of a complex mix setup, and control the levels of each channel (On the most elaborate machines the faders move.), but operation of the controls by the engineer generates MIDI data which can be recorded in a sequencer program, edited, and played back. Since the host computer can follow timing signals from a tape recorder, it is possible to perfect a multitrack mixdown (one program allows you to draw in the fader moves) before hitting the record button on the master machine.

MIDI CONTROLLERS

The traditional performer's interface for a MIDI system has been the clavier (keyboard), because that is the easiest to build and fits the manufacturer's preconception that synthesizers are really fancy organs. However, many musicians with perfectly good money are not keyboard players and prefer something that uses their hard earned playing skills on guitar or whatever. The manufacturers have responded and we can now use guitar, violin, xylophone, or drum-like controllers to make our real time music. There are many non-traditional controllers around too. The Buchla Lightning is a box that generates MIDI data when special wands are waved in the air. The most exotic system I have seen is the Hands, built by Michael Waisvisz of The Netherlands. These contraptions are strapped to the performer's hands and translate various finger and wrist motions into MIDI data. Waisvisz' performances show what a mistake it is to think of MIDI in traditional musical terms.

PROGRAMS

Any computer music system will have several types of program to perform different musical tasks. All programs may be available at once through concurrent processing, or you may have to reload software and convert data files to move from one job to another; that is mostly a function of how much you spend on your computer. Here are some of the programs you would expect to find.

---

functions availablem on different models, and you can probably find the combination that best suits your needs. Incidentally, not all MIDI controlled multi-processors are digital, some actually contain three or four separate analog circuits that are switched in when needed. One advantage of these is that all functions may be available at once!

The least apparent and the most influential is the **system software**. This is the package of programs that handle routine chores such as disk operations and the starting and stopping of the other programs. It is the system that gives each computer its personality and ease of use (or lack thereof).

The first composition programs (on mainframe computers) were **event list** editors. These work as the name implies; the composer types a list of things that are to happen using some kind of shorthand notation. Plain event list programs are no longer very popular, but many sequencer programs use some sort of event list as an editing option.

A more traditional approach to composition uses **CMN** or conventional music notation. These programs allow you to compose at the QWERTY keyboard, while displaying your music with fancy graphics and give flawless performances of your score. The most advanced allow you to record a performance using a sequencer program, edit the music on the crt screen, and then play back or print out the perfected piece.

The most common MIDI program is the **tape recorder simulator**, usually known as a **sequencer**. A program of this sort records a performance at the clavier rather than the sound output. Thus it can be played back with different voicing or other changes, but with first generation fidelity. Most systems offer multitrack-like operation with different MIDI channels controlled by each track. There are dozens of these for sale for various computers. One early feature of these programs was quantization[9]- if a performance was not exactly rhythmical, the program would move notes to the nearest "correct" time. When it became apparant that this produced rather mechanical sounding music, new versions appeared that allowed precise recording, slight quantizing, or even random changes ("humanizing") in timing.

Advanced versions of sequencers offer some level of recording and playback of audio, so it is possible to combine MIDi and acoustic instruments in the same program. Coming the other way, most DAW programs now include some accomodation of MIDI data.

An **algorithmic composition** program adds a different dimension to the composition process. Instead of entering each note, the composer specifies a set of rules and the computer calculates the notes. This can be done with ordinary programming languages such as BASIC or C, or with special

[9]Actually this was a bug, proving the adage "If you can't fix it, feature it!"

purpose programs that are somewhat easier to learn. The rules might be a set of jazz changes, a probability distribution for a variety of parameters, or a "grammar" describing the style of a specific composer.

One group of programs might be called **intelligent performers**. These programs have a score of some kind, and have the ability to follow a performance and play their own part in tempo, even improvising in some cases. The simplest are triggered by a direct input from a baton or drum stick, but the best actually analyze the sounds of the other instruments and respond appropriately.

For most synthesizers and processors, there are **voice editor/librarian** programs. These allow manipulation and / or display of the instrument settings, and most importantly, storage of those settings for recall at a later time. These are a great help in dealing with many instruments because the manufacturers seem to be in a competition to see how few operating controls they can get away with.

Max is a program that really defies definition. It consists of hundreds of tiny modules that each preforms some simple task. The composer can link these modules together into “patches” that perform any function imaginable. The interface is graphic, so the composer only has to draw a diagram of what he wants. The modules are always active, to the results are audible as you work. The end result can be converted into an application that will run on any macintosh without the parent program. The basic set of modules manipulate MIDI data, but you can add MSP to do digital audio or NATO+55 for video processing. If that’s not enough power for you, it is possible to add modules of your own written in C.

The Mathematics Of Electronic Music

One of the difficult aspects of the study of electronic music is the accurate description of the sounds used. With traditional music, there is a general understanding of what the instruments sound like, so a simple notation of 'violin', or 'steel guitar' will convey enough of an aural image for study or performance. In electronic music, the sounds are usually unfamiliar, and a composition may involve some very delicate variations in those sounds. In order to discuss and study such sounds with the required accuracy, we must use the tools of mathematics. There will be no proofs or rigorous developments, but many concepts will be illustrated with graphs and a few simple functions. Here is a review of the techniques you will encounter:

Hertz

In dealing with sound, we are constantly concered with frequency, the number of times some event occurs within a second. In old literature, you will find this parameter measured in c.p.s., standing for cycles per second. In modern usage, the unit of frequency is the Hertz, (abbr. hz) which is officially defined as the reciprocal of one second. This makes sense if you remember that the period of a cyclical process, which is a time measured in seconds, is equal to one over the frequency. (P=1/f) Since we often discuss frequencies in the thousands of Hertz, the unit kiloHertz (1000Hz=1kHz) is very useful.

Exponential functions

Many concepts in electronic music involve logarithmic or exponential relationships. A relationship between two parameters is **linear** if a constant ratio exists between the two, in other words, if one is increased, the other is increased a proportianal amount, or in math expression:

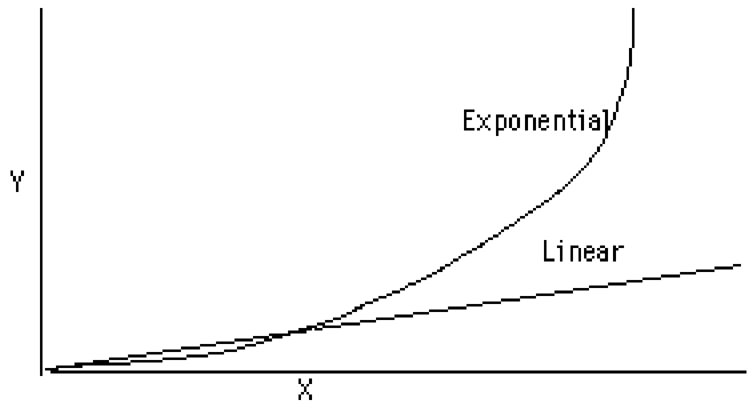
$$Y=kX$$

where k is a number that does not change (a constant).

A relationship between two parameters is exponential if the expression has this form:

$$\mathbf{Y=k^x}$$

In this situation, a small change in X will cause a small change in Y, but a moderate change in X will cause a large change in Y. The two kinds of relationship can be shown graphically like this:



One fact to keep in mind whenever you are confronted with exponential functions: $X^0=1$ no matter what X is.

Logarithms

A logarithm is a method of representing large numbers originally developed for use with mechanical calculators. It is the inverse of an exponential relationship. If $\mathbf{Y=10^X}$, X is the logarithm (base 10)[10] of Y. This system has several advantages; it keeps numbers compact (the log of1,000,000 is 6), and there are a variety of mathematical tricks that can be performed with logarithms. For instance, the sum of the logarithms of two numbers is the logarithm of the product of the two numbers-if you know your logs (or have a list of them handy), you can multiply large numbers with a mechanical adder. (This is what a slide rule does.) Two times the logarithm of a number is the log of the square of that number, and so forth.

We find logarithmic and exponential relationships many places in music. For instance the octave relationship may be expressed as Freq= $Fx2^n$ where F is the frequency of the original pitch and n is the number of octaves you want to raise the pitch. We discuss the logarithmic nature of loudness as length in essays 3 and 6.

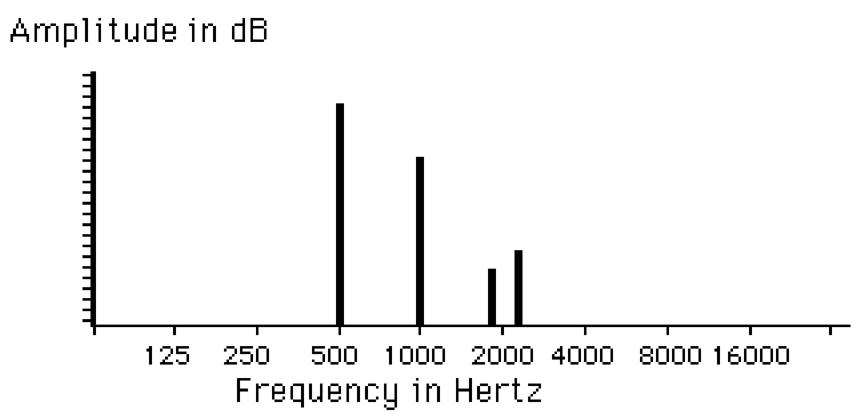
The number you raise to the power of a logrithm to get the number represented is called the base of the logarithim. For instance, the logarithm of 2 (base 10) is 0.301. There are two numbers commonly used as bases for logarithms: 10, in which case the notation is log x, and e (2.71828...), where the notation is ln x.

Decibels

The strength of sounds, and related electronic measurements are often expressed in decibels (abbr. dB). The dB is not an absolute measurement; it is based upon the relative strengths of two sounds. Furthermore, it is a logarithmic concept, so that very large ratios can be expressed with small numbers. The formula for computing the decibel relationship between two sounds of powers A and B is 10 log(A/B). Please see essay 6 for more complete information.

The Spectral Plot

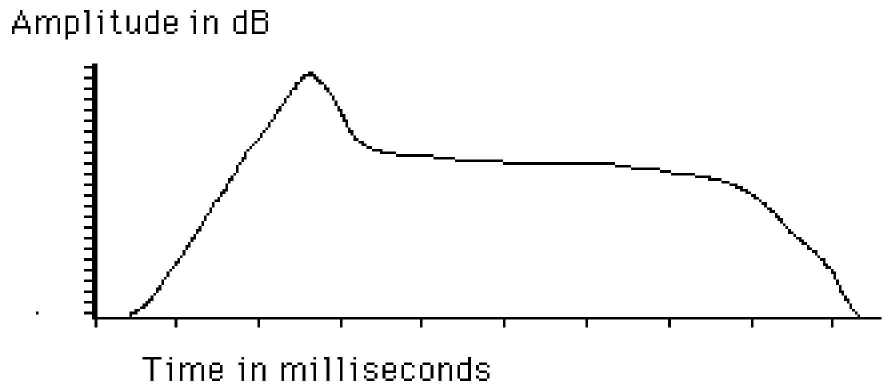
A spectral plot is a map of the energy of a sound. It shows the frequency and strength of each component.



Each component of a complex sound is represented by a bar on the graph. The frequency of a component is indicated by its position to the right or left, and its amplitude is represented by the height of the bar. The frequencies are marked out in a manner that gives equal space to each octave of the audible spectrum. The amplitude scale is not usually marked, since we are usually only concerned with the relative strengths of each component. It is important to realize that whenever a spectral plot is presented, we are talking about the contents of sound. In the example, the sound has four noticable components, at 500 hz, 1000, just below 2000 hz, and just above 2000 hz.

Envelopes

Envelopes are a very familiar type of graph, showing how some parameter changes with time.



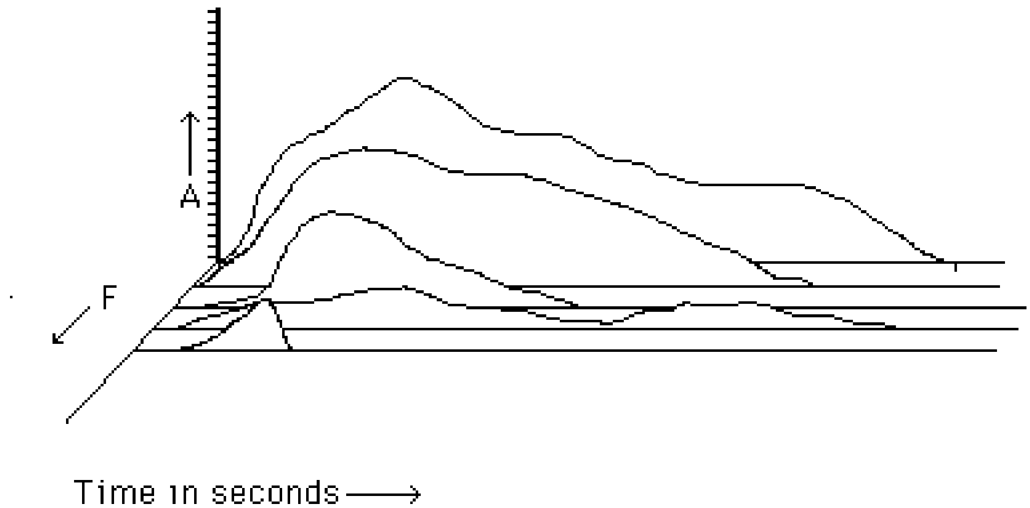
This example shows how a sound starts from nothing, builds quickly to a peak, falls to an intermediate value and stays near that value a while, then falls back to zero. When we use these graphs, we are usually more concerned with the rate of the changes that take place than with any actual values.

A variation of this type of graph has the origin in the middle:

Even when the numbers are left off, we understand that values above the line are positive and values below the line are negative. The origin does not represent 'zero frequency', it represents no change from the expected frequency.

Spectral envelopes

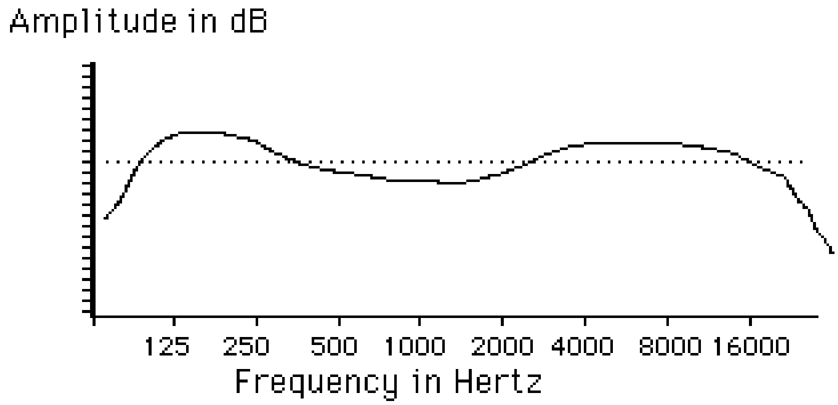
The most complex graph you will see combines spectral plots and envelopes in a sort of three dimensional display:



This graph shows how the amplitudes of all of the components of a sound change with time. The 'F' stands for frequency, which is displayed in this instance with the lower frequency components in the back. That perspective was chosen because the lowest partials of this sound have relaltively high amplitudes. A different sound may be best displayed with the low components in front.

Frequency Response

When we are discussing the effects of various devices on sounds, we often are concerned with the way such effects vary with frequency. The most common frequency dependent effect is a simple change of amplitude; in fact all electronic devices show some variation of output level with frequency. We call this overall change frequency response, and usually show it on a simple graph:

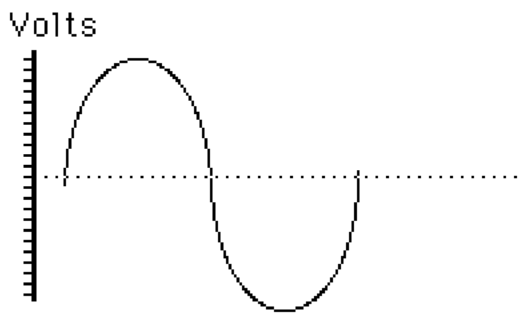


The dotted line represents 0 dB, which is defined as the 'flat' output, which would occur if the device responded the same way to all frequencies of input. This is not a spectral plot; rather, it shows how the spectrum of a sound would be changed by the device. In the example, if a sound with components of 1 kHz, 3kHz, and 8kHz were applied, at the device output the 1kHz partial would be reduced by 2dB, the 8kHz partial would be increased by 3dB, and the 3kHz partial would be unaffected. There would be nothing happening at 200Hz since there was no such component in the input signal.

When we analyze frequency response curves, we will often be interested in the rate of change, or slope of the curve. This is expressed in number of dB change per octave. In the example, the output above 16kHz seems to be dropping at about 6 dB/oct.

Waveforms

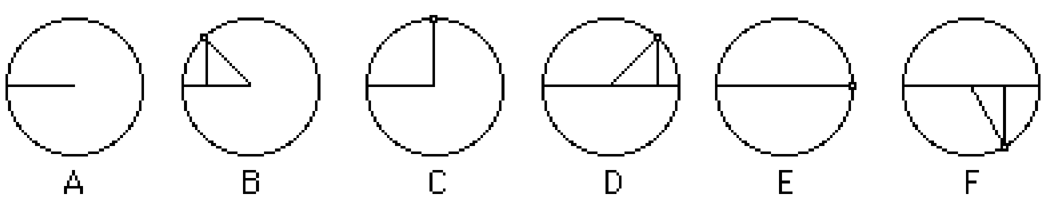
Once in a while, we will look at the details of the change in pressure (or the electrical equivalent, voltage) over a single cycle of the sound. A graph of the changing voltage is the waveform, as:



Time is along the horizontal axis, but we usually do not indicate any units, as the waveform of a sound is more or less independent of its frequency. The graph is always one complete period. The dotted line is the average value of the signal. This value may be zero volts, or it may not. The amplitude of the signal is the maximum departure from this average.

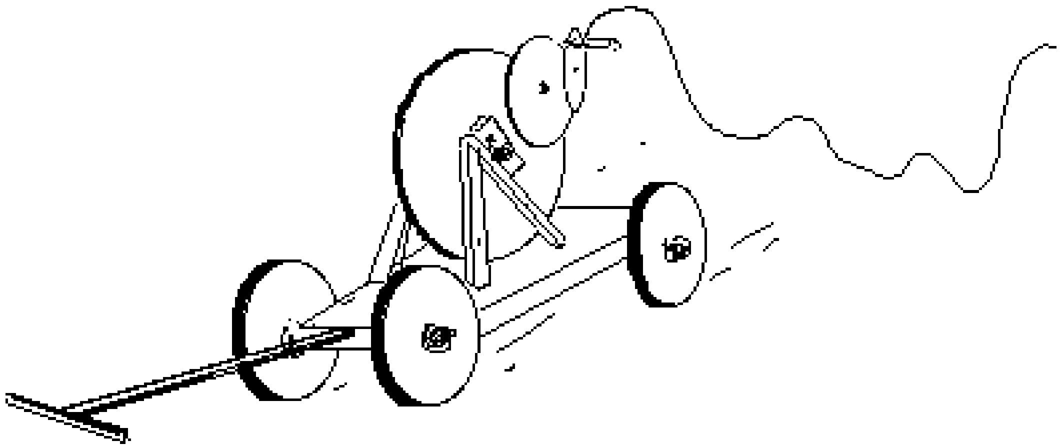
Sine Waves

The most common waveform we will see is the sine wave, a graph of the function v=AsinT. Understanding of some of the applications of sine functions in electronic music may come more easily if we review how sine values are derived.



You can mechanically construct sine values by moving a point around a circle as illustrated. Start at the left side of the circle and draw a radius. Move the point up the circle some distance, and draw another radius. The height of the point above the original radius is the sine of the angle formed by both radii. The sine is expressed as a fraction of the radius, and so must fall between 1 and -1.

Imagine that the circle is spinning at a constant rate. A graph of the height of the point vs. time would be a sine wave. Now imagine that there is a new circle drawn about the point that is also spinning. A point on this new circle would describe a very complex path, which would have an equally complex graph. It is this notion of circles upon circles upon circles which is the basis for the concept of breaking waveforms into collections of sine waves. (See the essay Taking Waves Apart for more information.)



The Fourier Engine

The Harmonic Series

A mathematical series is a list of numbers in which each new member is derived by performing some computation with previous members of the list. A famous one is the Fibonacci series, where each new number is the sum of the two previous numbers (1,1,2,3,5,8 etc.) In music, we often encounter the harmonic series, constructed by multiplying a base number by each integer in turn. The harmonic series built on 5 would be 5,10,15,20,25,30 and so forth. The number used as the base is called the fundamental, and is the first number in the series. Other members are named after their order in the series, so you would say that 15 is the third harmonic of 5. The series was called harmonic because early mathematicians considered it the foundation of musical harmony. (They were right, but it is only part of the story.)

Temperament

One of the aspects of music that is based on tradition is which frequencies of sound may be used for 'correct' notes. The concept of the octave, where one note is twice the frequency of another is almost universal, but the number of other notes that may be found between is highly variable from one culture to another, as is the tuning of those notes. In the western European tradition, there are twelve scale degrees, which are generally used in one or two assortments of seven. For the past hundred and fifty years or so, the tunings of these notes have been standardized as dividing the octave into twelve equal steps. The western equal tempered scale can then be defined as a series built by multiplying the last member by the twelfth root of two (1.05946). The distance between two notes is known by the musical term interval. (Frequency specifications are not very useful when we are talking about notes.) The smallest interval is the half step, which can be further broken down into one hundred units called cents.

Equal temperament has a variety of advantages over the alternatives, the most notable one being the ability of simple keyboard instruments to play in any key. The major disadvantage of the system is that none of the intervals beside the octave is in tune. To justify that last statement we have to define "in tune". When two musicians who have control of their instruments attempt to play the same pitch, they will adjust their pitch so the resulting sound is beat free. (Beating occurs when two tones of almost the same frequency are combined. The beat rate is the difference between the frequencies.) If the two attempt to play an interval expected to be consonant, they will also try for a beat free effect. This will occur when the frequencies of the notes fall at some simple whole number ratio, such as 3:2 or 5:4. If the instruments are restricted to equal tempered steps, that 5:4 ratio is unobtainable. The actual interval (supposed to be a third) is almost an eighth of a step too large.

It is possible to build scales in which all common intervals are simple ratios of frequency. It was such scales that were replaced by equaltemperament. We say scales–plural, because a different scale is required for each key; if you build a pure scale on C and one on D, you find that some notes which are supposed to occur in both scales come out with different frequencies. String instruments, and to some extent winds can deal with this, but keyboard instruments cannot. If you combine a musical style that requires modulation from key to key with the popularity keyboards have had for the last two centuries you have a situation where equal temperament is going to be the rule.

I wouldn't even bring this topic up if it weren't for two factors. One is that the different temperaments have a strong effect on the timbres achieved when harmony is part of a composition. The other is that the techniques of electronic music offer the best of both systems. It is possible to have the nice intonation of pure scales and the flexability for modulation offered by equal temperament. Composers are starting to explore the possibilities, and some commercial instrument makers are including multi-temperament capability on their products, so the near future may hold some interesting developments in the area.

Simple Harmonic Motion

You have noticed many objects have a tendency to return to their original location after they have been moved slightly. You have probably also noticed most objects not only return to their original position but continue too far, so that they may swing back and forth several times before they come to rest. A playground swing is a convenient example for a discussion of this kind of motion, because we have all ridden one, and more importantly, pushed one.

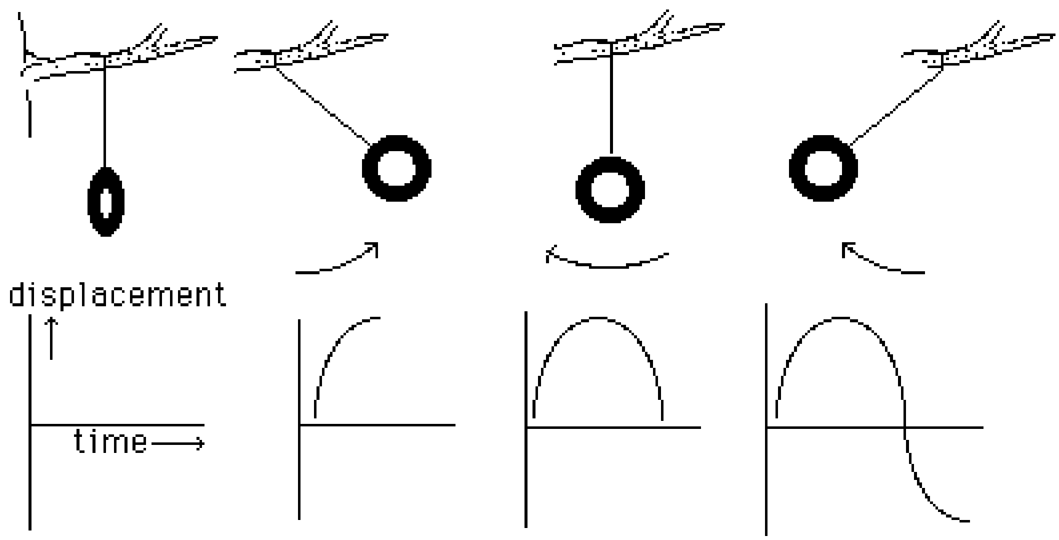


Fig. 1 Motion of a swing

If given one strong push, a swing will move back and forth because the earth's gravity is pulling down while the rope is making the swing move in a partial circle. When the swing has returned to the starting point, it can't go down any more, but still has a lot of momentum to use up, so it moves beyond the center and up on the other side until gravity slows it down and pulls it back to the center, but it still has too much momentum.. and so forth. This whole operation is periodic: each complete swing takes the same amount of time, regardless of how far the swing moves. When the arc of swing is large, the swing moves quickly, when the arc is small, the swing moves slowly. The amount of time it takes the swing to complete its cycle depends on the length of the rope and nothing else. If you make a graph of the distance the swing travels crossways with time, the result is a sine wave.

The only thing slowing the swing down is friction with the air and within the rope. If you were to replace the energy lost to this friction, you could keep the swing going forever. If you try this by giving the swing extra pushes, you very quickly discover you must push at the right time and in the right direction, or you waste your push or perhaps slow the swing down. Specificaly, the time between pushes must correspond with the period of the swing, and you must push in the same direction the swing is already going.

Now there are many objects that behave like swings in their motion (In classical physics this behavior is termed SIMPLE HARMONIC MOTION). In fact, any object that is even slightly movable, has mass, and has some tendency to return home when disturbed will do this. Sound waves can contribute to the simple harmonic motion of an object if the period of the wave and period of the motion of the object are about the same. This phenomenon is called RESONANCE. In the next essay you will see that resonance is an important factor in the operation of acoustic musical instruments.

The Helmholtz Resonator

One harmonic motion machine that is particularly interesting is an ordinary pop bottle.

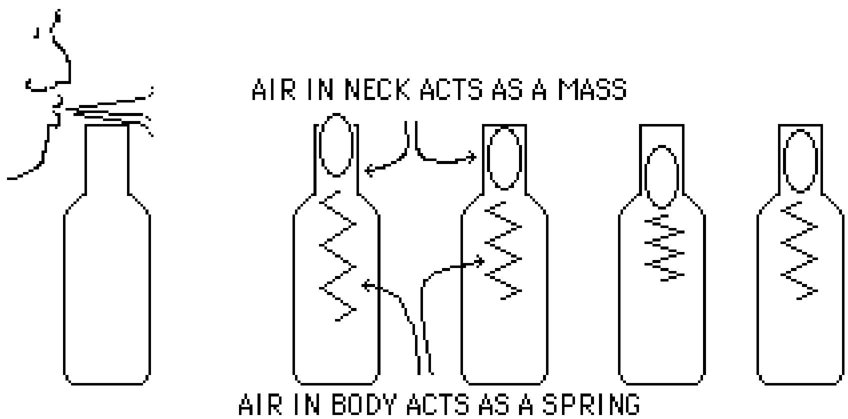


Fig. 2 Harmonic motion in a bottle.

When you blow across the bottle opening, the air in the neck behaves as a mass on the end of a spring. (The air in the body of the bottle is the spring.) The flow of air across the top pulls the neck mass a little bit out of the bottle, which reduces the pressure within the body. That low pressure pulls the neck mass back in, and the result is an oscillation of the air in the neck. The rate of oscillation is determined by the mass of the air in the neck and the size of the body of the bottle. This system is resonant enough that it will respond to any sort of air disturbance with some oscillations at its characteristic frequency. Helmholtz used bottles in this way when he proved that complex sounds contained a mix of partials of various frequencies.

Impulse Response

If you apply a burst of energy to most objects (by hitting them with a stick, for instance) they will generally "ring" for a time. What is happening is that each part of the object is going through its own version of harmonic motion. (Imagine a swing hanging from a swing.) These motions interact in complex ways, as they pass energy back and forth in and out of phase with the various motion components. Ultimately everything dies down, but you may be able to hear different pitches at different times, some even growing stronger for a while. A recording of this complex sound is the IMPULSE RESPONSE of the object, which can be used in a variety of interesting ways beyond just plying it back. In particular, if you get the impulse response of a reverberant space, you can apply this response to a recording by a computer process called convolution. This gives the effect of the recording having been made in the space.

A Look At Un-Electronic Musical Instruments

A little later in the course we will be looking at the problem of how to construct an electrical model, or analog, of an acoustical musical instrument. To prepare for this, we need to consider how such instruments work.

Resonance

All acoustical instruments are built around some kind of RESONATOR. A structure is resonant if it responds to an energy impulse by vibrating for a noticeable length of time. The frequency of vibration is determined by the size and material of the resonator, and the pattern of vibration may be simple harmonic motion or some more complex action. If the vibration dies away quickly, the resonator is said to be DAMPED. A repeating series of impulses will sustain the vibrations if the frequency of the pulses matches to some degree the natural frequencies of the resonator. If the resonator responds to a wide range of input frequency, it is BROADLY TUNED. If the input frequency has to match the frequency of the resonator pretty closely before resonance occurs, the resonator is NARROWLY TUNED.

Acoustic instruments also require some sort of DRIVER, a mechanism that applies energy to the resonator in the appropriate form. The driver may be as simple as a stick (or bare hand), or it may be an elaborate resonant structure itself. If the driver supplies the energy all at once, it is an IMPULSE driver; if the energy is applied as a repeated stream of pushes, the driver is often called a SOUND GENERATOR.

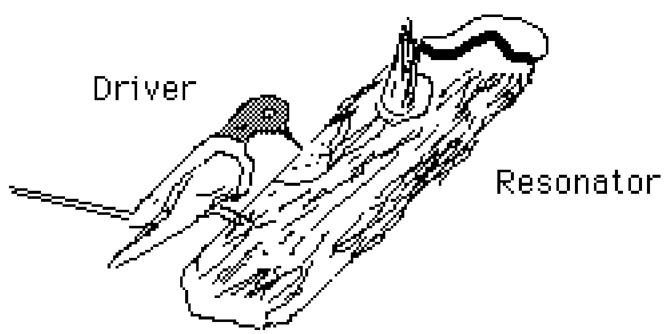


Fig.1 A simple instrument

Most instruments also possess some kind of pitch control mechanism. Pitch is controlled at two levels, tuning and performance. The tuning of an instrument determines the pitch possibilities that the artist may exploit during the performance. An instrument's tuning is largely in the manufacturing process, although there is often some provision made for adjustments. Pitch controllers may modify the operation of the resonator, the driver, or both. Some instruments provide pitch selection by duplication of tuned structures, trading flexibility of intonation for the possibility of polyphonic performance.

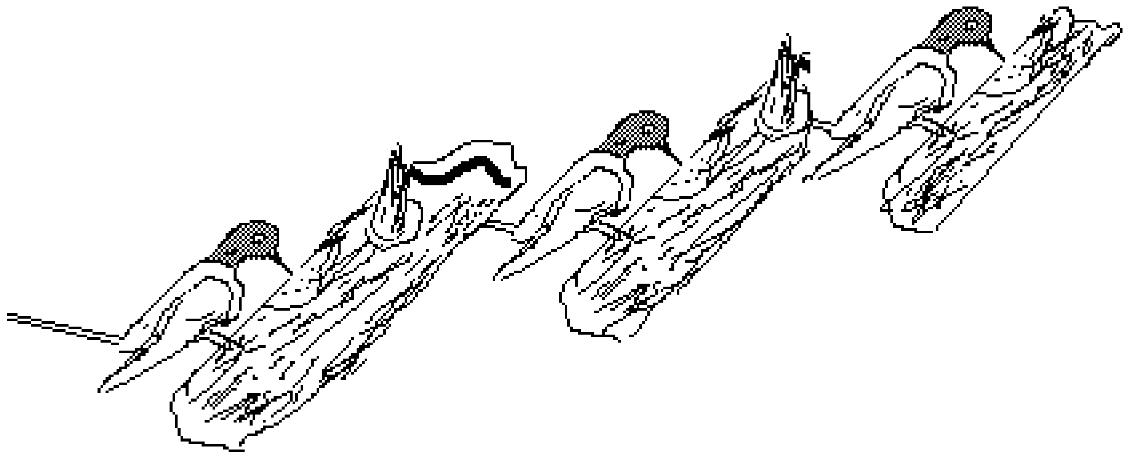


Fig. 2 A polyphonic instrument

It is difficult to organize a general discussion of operating principles of instruments because there are so many varieties, but for engineering purposes we can divide instruments into three classes based on the style of driver: the familiar strings, winds, and percussion instruments.

STRING INSTRUMENTS

The driver or sound generation device of the string instrument is a tightly stretched string. When the string is excited, which may be done by a hammer blow, a pluck, or a continuous scrape, it is set into motion at a rate determined by its length, mass, and tension. The motion is complex and contains energy at many (almost) harmonically related frequencies. This motion is transmitted to the resonator via the bridge, a light piece of wood supporting one end of the string.

The resonator of a string instrument is commonly an oddly shaped box or a wide thin board. The resonator is not sharply tuned; it responds to broad bands of frequencies and radiates sound at those frequencies from its entire surface area. The response of the body or soundboard is not flat within these bands, however, so some frequencies are transmitted more efficiently than others. These response peaks are called FORMANTS, and play a very important part in establishing the timbral identity of an instrument.

Since the tuning of the resonator is very broad, the string frequency is the controlling factor in the pitch of the instrument. (The string itself is a narrowly tuned resonator.) String frequency is controlled by adjusting tension for tuning and by manipulating length for performance. The formant frequencies do not change, so the waveform produced varies somewhat from one pitch to another.

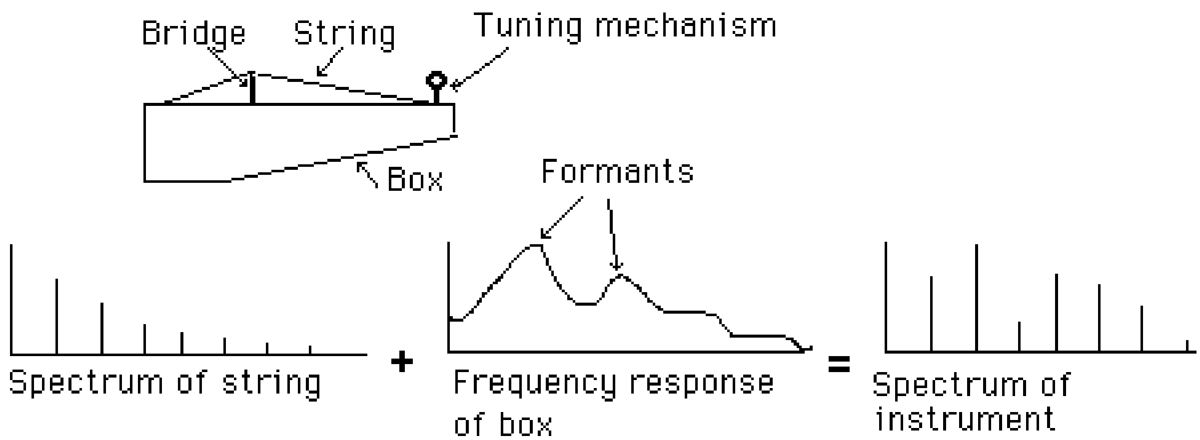


Fig. 3 A string instrument

WIND INSTRUMENTS

With wind instruments, the resonator is usually in the shape of a pipe and the energy is supplied as a stream of air into the pipe. The driving mechanism is some kind of valve that periodically interrupts or modulates the air flow. The reed of some woodwinds and the lips of the brass player are examples of modulating valves. These respond to back pressure from the resonator, so the resonator has almost total control of the frequency of the instrument. The resonant frequency of a pipe is determined by its length but the system will respond at harmonics of that frequency with a little encouragement from the driver. (The actual mechanism of resonance is a standing wave.) Notewise pitch control in the winds is usually done by adjusting the length of the resonator. Slight changes to the driver cause slight changes in pitch, whereas major changes in the driver will cause the pipe to shift modes of vibration and produce a large jump in pitch.

The spectral content of pipe resonators follows the harmonic series closely but the upper components usually deviate somewhat from the predicted frequencies. The amplitudes of the various partials are determined by the shape of the pipe, particularly by the configuration at the ends.

In the WOODWINDS the pipe length is changed by opening or closing holes along the side of the instrument. The part of the instrument that extends beyond the open holes acts as a second resonator, modifying the sound produced by the primary resonator in a manner that changes somewhat from note to note. Woodwinds typically only use the three lowest vibratory modes of the pipe, so enough holes have to be provided to fill in notes for an octave or more.

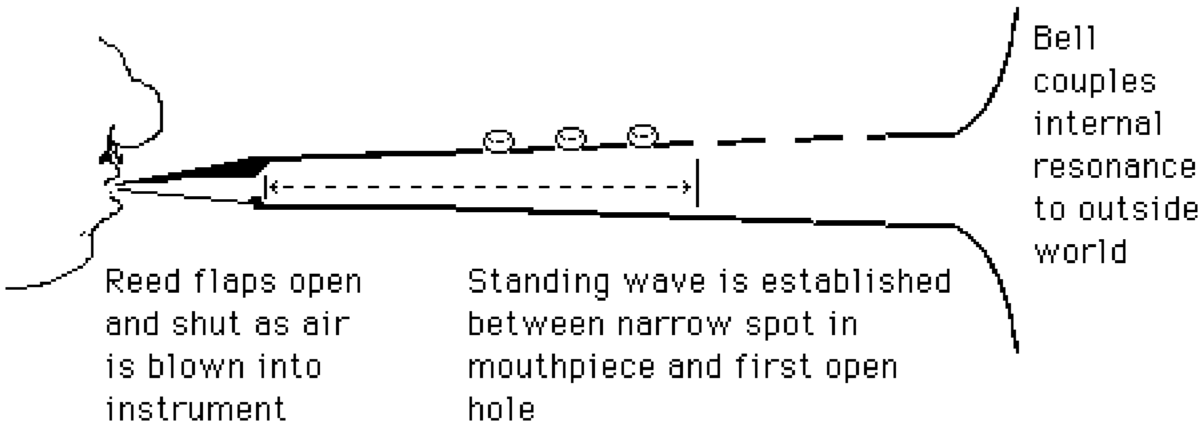


Fig. 4 A wind instrument.

In the BRASS instruments the pipe length is manipulated directly, adding sections by the use of valves or pulling slides in or out. Since the air modulating valve is part of the player's body (the lips) it is very responsive, allowing use of many high pipe modes. In fact, the fundamental mode of vibration is not used at all (except for special effects), so only enough valves or slide positions are required to fill in the space between the second and third modes, an interval of a fifth. Much of the brass timbre is attributable to the bell, which is frequency selective in the way it transmits sound power into the open air. The sound is drastically changed if the shape of the bell flare is modified by the addition of a mute.

There are non-pipe wind instruments:

The ocarina is a Helmholtz resonator that is tuned by opening holes in no particular order. The more holes open, the lower the pitch because the holes add to the vibrating mass.

The harmonica and accordion have reeds that sound into a rudimentary resonator. The resonator provides a weak formant, but no pitch control.

In the voice, the resonant structures are an assortment of body cavities, including the mouth. The volume of these cavities can be changed, producing tunable formants. The major driving mechanism of the voice is the larynx, containing two loosely stretched flaps of muscle that can modulate the air flow from the lungs. The frequency produced is controlled by muscular tension, with no effective feedback from the resonators. The result is an instrument with independently controlled pitch and timbre. The timbral range is extended by an alternate driving mechanism, the tongue, which can provide a variety of noise and impulse inputs to the system.

PERCUSSION INSTRUMENTS

Loosely speaking, a percussion instrument is anything you can hit. If we must make a generalization, we might say that percussion instruments usually lack a complex driving mechanism that could be separated from the resonator. The unifying principle is that impulse energy is applied directly to the resonator, which responds with vibrations for a short period of time. You can see that almost any instrument can be played in a percussive mode.

The resonator may be an air chamber of the Helmholtz or pipe variety, or may simply be a particularly resonant chunk of metal or wood. The air resonators have spectra that fit the harmonic model to some degree, giving a fairly definite pitch, but the solid body resonators vibrate in extremely complex ways, with spectra that are non-harmonic clusters of components or even broad band noise. Pitch on these instruments is usually not very discernable beyond a general highness or lowness.

SUMMARY

You can see from this discussion that there are (at least) three common relationships between drivers and resonators. We might call these driver controlled, feedback controlled, and resonator controlled. In the strings and non-pipe winds the frequencies of the resonator do not strongly affect the frequency of the driver or the pitch of the instrument; pitch control is a function of the driver. In the pipe winds, the resonator and the driver affect each other, producing a pitch suitable to both. In the percussion instruments pitch is entirely up to the resonator, since the driving energy is applied as an impulse.

Acoustic Treatment For Home Studios

Part 1: Soundproofing

Silence is golden, or at least pretty expensive. Commercial recording studios cost hundreds of thousands of dollars to build because they must allow absolutely no sound to enter from a usually noisy urban environment. Double and triple walls, isolated concrete slabs, custom steel doors are all standard but high priced items used in their construction. A studio's sound is its number one asset and most owners will go to any lengths to get it right.

Luckily, electronic music does not normally require the extreme isolation needed for recording live ensembles. The use of microphones is infrequent enough that it can be scheduled for predictably quiet times, and close mic techniques, (which are usually appropriate for sampling or vocal lines) don't pick up much noise. Given a reasonably quiet, solidly built house to start with, a decent home studio can be created with modest expense and effort.

SOME THEORY

Sound can travel through any medium-- in fact it passes through solids better than through air. Sound intensity is reduced in the transition from one material to another, as from the air to a wall and back. The amount of reduction (called the transmission loss) is related to the density of the wall-- as long as it doesn't move in response to the sound. Unfortunately, all walls are somewhat flexible. Any motion caused by sound striking one side of the wall will result in sound radiated by the other side, an effect called coupling. If the sound hits a resonant frequency, the wall will boom like a drum. Most isolation techniques are really ways to reduce coupling and prevent resonances.

NEW CONSTRUCTION

The most effective soundproofing must be designed into a house when it is first built. A typical residential wall is made of a frame of 2x4 wood studs covered with 5/8" thick gypsum board. Properly built (no holes!) this will provide about 35 dB of isolation. Fiberglas filler, R-7 or better, will increase this by 5 to 8 dB and decrease wall resonance. Doubling the thickness of gypsum gives another 3 to 6 dB of overall isolation, but its most important effect is lowering the resonant frequency, hopefully below the audio range.

There are two common strategies for reducing coupling between the two sides of the wall. One is to make the gypsum to stud connection springy, either by using metal studs or by hanging the gypsum board on resilient metal bars. The most effective trick is to use separate studs for each face of the wall so there is no direct connection. This eats up a lot of space, but can give a transmission loss of over 60 dB. This is actually better performance than simple cinder block or poured concrete construction!

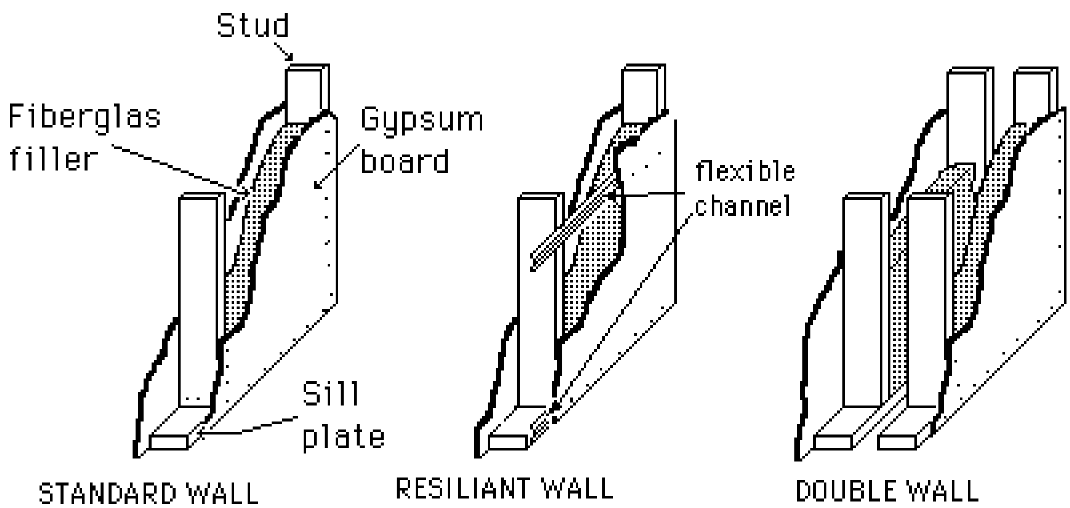


fig. 1

These same principles can be applied to floors and ceilings. A heavy false ceiling hung on springs can match the performance of a double wall-- If there is a room below the studio, it should get a double ceiling too.

INTERIOR WINDOWS

The window between control room and studio used to be a traditional feature of a recording facility. The home studio doesn't really need one, because you can get a decent video camera and a large monitor for less than what a good window costs to build. If you want a window, figure 2 shows what has to be done:

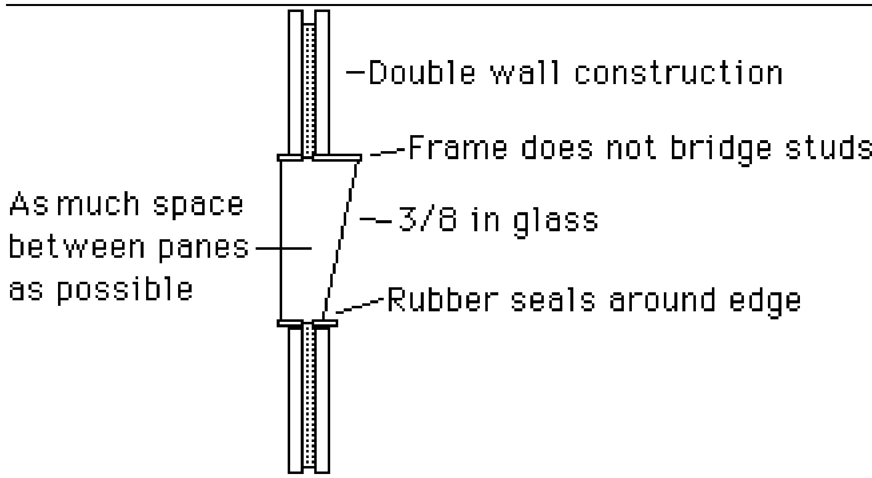


figure 2. The traditional window.

The effectiveness of these kinds of construction depends a great deal on the craftsmanship of the builder. There must be no loose studs, and the sill plates must really hug the floor. The board must be well fitted and all potential cracks must be caulked. (Caulk is soft and will not crack when the building settles.) Do not put holes in sound walls for outlets or pipes--use surface mount electrical fittings and caulk around any wires that pierce the wall.

STRUCTURE BORN NOISE

The peskiest isolation hassle is dealing with sounds transmitted through the frame of the building. The problem is caused by machinery such as air conditioners and refrigerators which are mounted on floors or walls and can actually shake the structure. Footsteps can cause similar effects to a somewhat lesser extent. This is usually not severe with wood frame construction but can be a nightmare in a concrete and steel building.

This kind of noise must be treated at the source. Walking areas should be carpeted, and heavy appliances have to be mounted with shock absorbers or placed on thick rubber pads. In a wooden house sound tends to be transmitted along the floor joists, so some problems can be solved by simply moving the offending machines. With concrete and steel buildings, you usually wind up completely "floating" the studio floor, a very complex and expensive operation.

Water pipes are distressingly efficient at carrying sound. If any pass under the floor of the studio (pipes in the walls are a definite no-no) make sure they are on flexible hangers. If your pipes are prone to "water hammers" consult a plumber about possible cures.

Retrofitting

We seldom have the luxury of building our home studio from scratch. More often we are trying to fix up an existing room, and budget or landlords limit the techniques we can use. The best approach to adding soundproofing is to try simple techniques first and to move to the high caliber options only when needed.

STEP ONE: THE OBVIOUS

You can reduce the amount of isolation you need if you give some careful thought to the choice of rooms you are putting your studio in. Pick a room that does not adjoin a kitchen or bathroom, or the place where your housemate does taxes or watches TV. In other words, put some space between the studio and the noisemakers or sound sensitive activities. The fewer walls in common with the rest of the building, the better. Clearly, a house is a better location than an apartment because you don't have to worry as much about sound traveling through the ceiling or floor. An outside corner room away from the street would be a good choice, a basement would be even better.

Some people consider a garage the ideal location for a studio. This may be true, but you will encounter special problems with the big door and with getting heat and ventilation. The way most garages are built, you are really working outdoors.

STEP TWO: TIGHTEN UP

Most builders are more concerned with how walls look rather than how solid they really are. This is unfortunate, because any air path from one room to another will limit the wall's effectiveness. You can make an amazing contribution to keeping the sound in your studio by filling all cracks and holes, no matter how small or indirect.

The worst sound leaks will be around doors. Your neighborhood hardware store has the fittings and gaskets to fill these up, sold for weather stripping but effective for sound too. Some common styles of gasket are illustrated in fig.3.

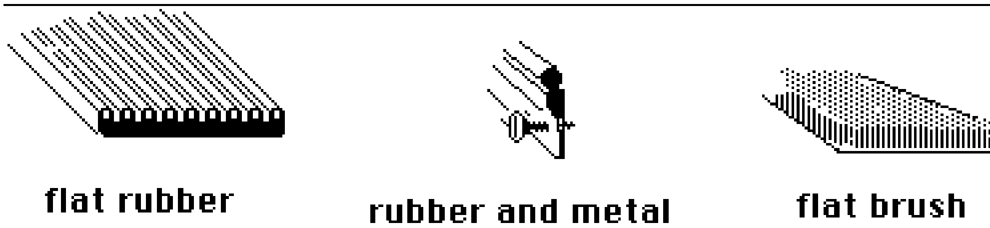


Fig. 3

The flat rubber type is used in a door that doesn't fit well, rubber and metal gaskets work on doors that are pretty tight already. The brush material is for sliding surfaces. Don't forget the bottom of the door-- the best gaskets are spring loaded and drop down when the door is closed.

Once the door is sealed there still may be leaks around the door frame. Carefully remove the trim and fill any gap between the frame and gypsum board with caulk or spray polystyrene foam. As long as you are pulling off trim, check for gaps behind the baseboards and around any window frames.

Incidentally, many interior doors are hollow and light and don't really stop sound well even when tightly gasketed. Such a door should be replaced with a solid one. Manufacturers will supply data on the amount of transmission loss a door can provide. Alternatively, the door can be reinforced with a layer of thick plywood, or you may want to hang a second door that opens the other way in the frame. If none of this is practical, a really heavy curtain over the door will help some.

External windows are a real problem, since a single layer of ordinary glass is only slightly better sound insulation than nothing at all. Storm windows are a big help, especially if you fill the space between panes with Fiberglas. (You don't really need to see outside, do you?) Thermal glass is actually worse than a single pane window because the narrow air space tends to resonate. Seal the movable part of the window with good gaskets, then cover the whole opening with heavy drapes. Make sure the drapes fit snugly against the wall all the way around. An inexpensive alternative to drapes is a solid piece of 3/4" plywood, gasketed just like a door. This can be hinged to the wall as shutters or in a sliding track, just as long as it fits tightly.

Electrical fittings are another source of leakage. Take the plates off light switches and receptacles, fill the gaps between the box and the wallboard, and add a sealing gasket when you put the plate back on. If switches or receptacles are found back to back in both sides of the wall, the gasket will not be enough to stop sound. Replace the electrical box with a surface mount type, and patch over the original hole. If you aren't up for rewiring, cover the offending outlets with a weatherproof hinged cover.

Air ducts present a special problem. You don't want to cover them up (even keyboard players have to breathe), but they are a veritable freeway for sound. To soundproof air vents build a baffle as illustrated in figure 4. Start with a rectangle of 3/8 in plywood as large as you can fit into the space. Cut a series of slots for the air to pass through, and cover the back with Fiberglas, leaving the slots clear. Hang this at an angle in front of the duct and fit triangular pieces over the ends.

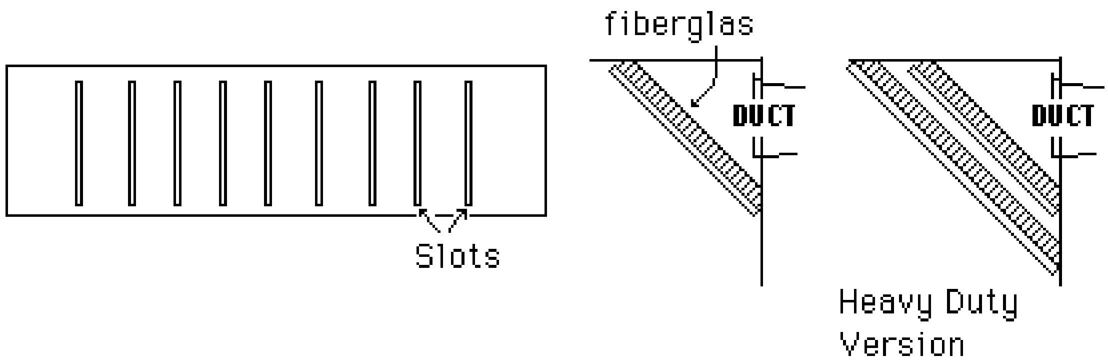


fig. 4 vent baffles

Put one of these on each vent connected to the studio ducts, even those in other rooms. If this isn't enough, you can add a second unit in front of the first.

STEP 3: BEEF IT UP

As I mentioned earlier, low frequency sound can actually resonate a gypsum board wall just as it will rattle a drum head. When this happens, the bass might seem louder in the next room than it is in the studio! The only cure short of tearing the wall down and rebuilding is to add weight, usually another layer of wallboard. This is most effective if the new wallboard is thicker than the original and if it is glued in place rather than nailed to the studs.

There is no point in doing this halfway-- you must cover all internal walls, preferably on both sides. Additional gypsum gives diminishing returns, but an intermediate layer of soft fiber board can be helpful.

Even if there is no direct air route for sound to follow, there can be flanking paths around heavy walls through thin ceilings or floors. The sound will then pass through the attic or crawl space into adjoining areas.

You can add wallboard to a ceiling either directly on top of the existing material or suspended a few inches below. In some cases it would be simpler to extend the side walls all the way up to the roof. A properly built hardwood floor should not leak much sound, but sometimes contractors cut corners when a house has wall to wall carpet-- you should lift a corner of the carpet and see what is really below.

THE ULTIMATE SOLUTION: DOUBLE UP

Truly isolated spaces are created by building a separate room within the room. Both the external room and the internal room have to be tight and heavy and there must be no solid connection between the two, not even the floor. You can buy prefabricated isolation rooms (at a hefty cost), or you can build one using construction techniques similar to that of the house. Something like this should really be designed by an architect to fit your situation, but here is a typical plan to give you the idea.

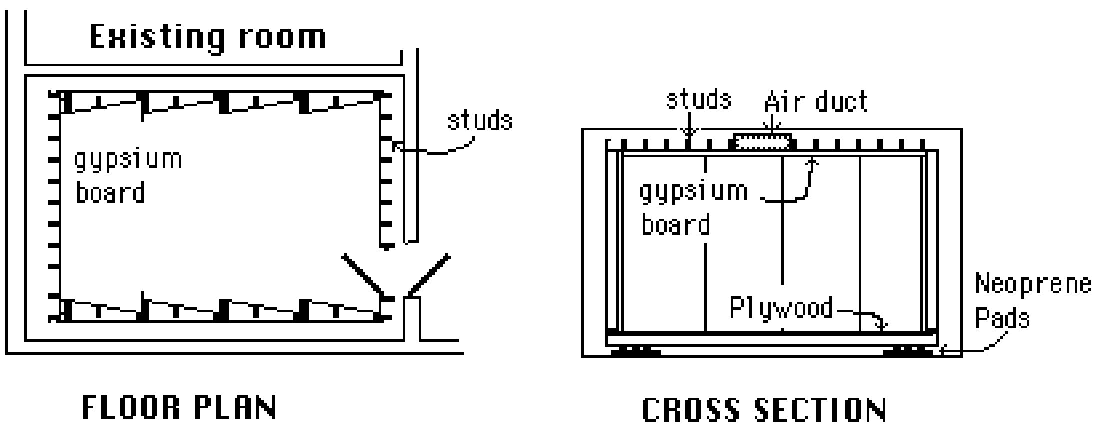


fig 5

The inner room is built on a platform of 2X4s covered with two layers of 3/4 inch plywood. The platform is supported by neoprene pads that line up with the floor joists. There must be no other connection between the room and the house. The walls and ceiling are built on the platform using 2X4 studs and double wallboard on the inside only. The space between the walls should be at least one inch (wider if practical) and lined with Fiberglas. The air duct should be very long and lined with sound absorbent material. Get the heaviest solid door and frame you can find, and add gaskets as described above.

These steps can result in a very quiet space, but they get progressively more expensive-- the real question is when is it quiet enough? The easy test is to make a recording of the space. No sound, just a tape of the mic levels at their usual setting with nothing going on. Now turn up the gain and play it back. If you can't hear any difference between the unrecorded and recorded portions of the tape you have reached your goal.

The only way to get an objective measurement of sound levels is to use an SPL meter. (There are some inexpensive models by Gold Line/loft or Radio Shack.) As measured by the "C" scale on these meters you will find the following numbers appropriate for these uses.

Good restaurant	35-45
Quiet office	30-40
Hospital room	25-35
Church	20-30
Concert hall	15-25
Recording studio	10-20

A decent home studio should measure in the 20s. Assuming all noise sources are outside the room, you can calculate the amount of transmission loss the walls have to provide by measuring the sound level with the door open. Close the door and you can figure what you already have. If the level does not change when you close the door, you know where to start!

Part 2: Room Treatment

Has this ever happened to you? You are playing your latest masterpiece at a party at a friend's place, and when the best song comes on you want to hide under the couch-- the bass is boomy, the highs screech, and along with the backup vocals you can definitely hear Gilligan's Island. If you find this experience familiar, you are probably the victim of BAD ACOUSTICS.

You won't be surprised to hear that the shape and furnishings of a room can affect the way things sound-- we have all experienced extreme cases such as large echoey bathrooms and overstuffed restaurants. These effects can easily happen in a subtle way in your studio, causing inaccuracies in the sound from the monitors. When you record or mix you adjust the music till it is right in your control room, but when you play the tape in a neutral environment the sound is overcompensated and strange.

There are expensive instruments available to measure the quality of sound in a space, but the best ones are on the sides of your head. You can compare rooms by listening to familiar recordings. (It doesn't have to be on CD-- you can tell a lot from the quality of hiss on a tape.) In a good room, the bass is balanced and clear, cymbals "shine" without being harsh, you can understand words without effort. A mono signal appears to come from a spot exactly between the speakers, and that spot does not jump around with changes of pitch. Now listen to the quiet-- can you hear a refrigerator, a TV, traffic on the street? Clap your hands--you should hear a slight broadening of the sound, but little reverberation and certainly no pitches or echoes.

These simple tests should tell you about any severe problems the room may have. Subtle ones will show up in the music produced in the room, as described above. You may be surprised to find that the control of the sound of a room is not really very complicated and can usually be accomplished with inexpensive materials.

SOME MORE THEORY

The goal is very simple-- we want to get the sound from the speakers to your ears without messing it up. This is really just a matter of what becomes of the sound after it passes your ears.

There are three things that can happen when sound hits a wall. It can be reflected, absorbed, or diffused. If the wall is flat and hard, the sound will be reflected. A single strong reflection can sometimes be heard as an echo, but in most rooms a lot of reflections (including reflections of reflections) combine into the reverberation. The aspect of reverberation you hear about the most is reverberation time. This is the amount of time it takes a loud short sound to die away. "Dying away" can be defined more scientifically as a drop in loudness of 60 dB, so acousticians call reverberation time RT60.

The amount of reverberation desired in a room depends on the activity going on. Musicians like fairly long reverberation times; between one and two seconds. This allows them to hear themselves play and enhances the harmonic effects of the music. (In larger rooms even more reverb is desirable because it helps fill the hall with sound.) For listening to speech or music played through loudspeakers this amount of reverb is too much--values around a second are more comfortable, and for critical listening to speakers the RT60 should be close to a half second.

Reverberation time is determined by the volume of the room. It can be reduced by replacing some of the hard, reflective parts of the the walls with soft, absorptive sections. Every material has some absorptive qualities. This is described by its coefficient of absorption, a number between 0 and 1, with 0 being totally reflective and 1 being an open window. For instance the COE of brick is 0.04, whereas that for heavy drapes is around 0.6. The effective absorption of a surface is simply the COE times the area of the surface in square feet. These numbers can be used to compare materials and to predict the results of treatment. The absorption ability of most materials is frequency dependent, which can cause problems as described later.

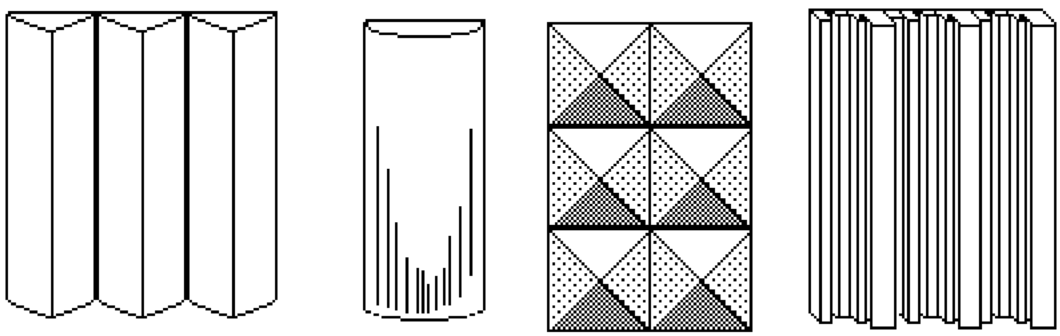
Reflections off flat walls can sometimes combine to produce undesirable effects. The worst of these is the standing wave.

STANDING WAVES

Standing waves are created when you have two parallel facing walls. There will be a particular set of frequencies that are reinforced by the distance between the walls (the sound makes exactly one round trip on each cycle of the speaker and the pressure fronts pile up). This is what happens in bathrooms- you probably know one where the deep tones of your voice are tremendously supported (doesn't everybody sing in the shower?). Most rooms have three pairs of parallel surfaces, and the dimensions are usually just right to affect music. An eight foot ceiling, for instance, reinforces 70 hz. (This is called a room mode.)

This phenomenon can be prevented by designing the room with nonparallel walls. It can be cured in existing rooms by making one of the walls absorptive or by breaking up the flat surfaces. When sound is reflected off a rounded or complex surface, it is diffused. Diffusion spreads the reverberant sound evenly throughout a room, which not only prevents standing waves but also eliminates "dead spots"-- places where components of the sound are missing.

We can break up flat surfaces by hanging large objects called diffusers. The shapes chosen for diffusers are really a matter of taste and cost. Avoid concave curves, which focus sound instead of dispersing it, but otherwise pyramids, lattices, or computer designed random surfaces all work well. The depth of a diffuser determines the lowest frequency that will be affected. A diffuser one foot deep will scatter sound down to 160 hz.



Some popular shapes for diffusers.

Reflections can cause a further problem when the principal activity in a room is listening to loudspeakers.

INTERFERENCE

You may be familiar with phase interference from recording work with multiple microphones. If a sound arrives at a single point via two paths at slightly different times, certain frequencies will be reinforced and others will be weakened. You can easily hear this by putting your ear close to a wall: the quality of sound will change because the reflections off the wall interfere with the direct sound. The effect is at its worst when the distance the reflected sound travels is only slightly longer than the direct distance.

Phase interference is attacked by careful consideration of the placement of speakers and the listener. In general avoid locating either so that there are short reflective paths off of walls, ceiling, or equipment. The worst problems occur when a speaker winds up in a corner. If this is unavoidable, figure out where the reflections occur, and make that part of the wall or ceiling absorptive.

COLORATION

What I've said so far might seem to imply you can take care of all acoustic problems by making every surface absorptive, completely deadening the room. Actually, such a room is rather unpleasant to work in, but even if it weren't, any attempt to create it would probably be a disaster. The problem is that all absorptive materials are frequency selective. As a general rule, high frequency sound is absorbed more readily than low, so as absorption is added to a room, the reverberation becomes more and more bassy in tone. Some of this coloration is ok, even preferable, but eventually the room develops a tubby response. If we need a very dead room and bass buildup occurs there are devices called bass traps and Helmholtz resonators that absorb a restricted range of very low frequencies. The specifics for designing these are beyond the scope of this article, but the general principle is the larger they are, the lower the frequency. The moral is that absorption should be used only in moderation, and only materials that soak up the full range of sound should be used.

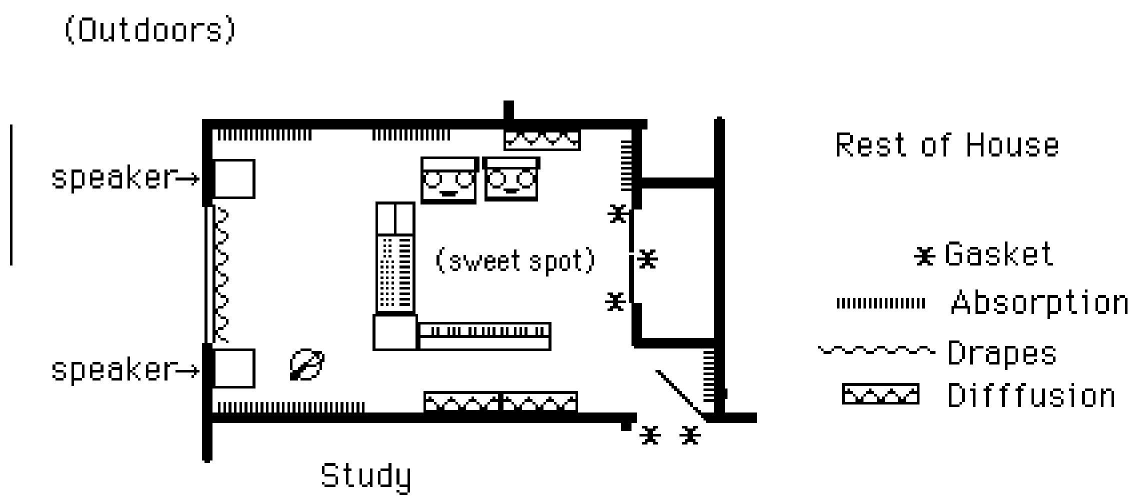
Such materials need not be expensive. In fact, ordinary R-19 fiberglass insulation (about 6 inches thick) is as good a general purpose absorber as you can find and costs about 30 cents a square foot. You can tack it right to the wall, paper side down. Of course this is ugly as sin and breathing fiberglass is not good for you, so you want to cover it up with some lightweight cloth. More attractive absorbers can be made from Insul-shield (a solid wall insulation material) or various foam products sold through audio supply houses. (Again see sidebar) These all work down to 100 hz or so. Carpet on a thick pad is a decent absorber down to about 250hz. It is the simplest way to control floor to ceiling standing waves, and if hung in deep pleats works well as a wall treatment also.

We can get away with materials that poop out below 100 hz because normal wall and floor construction is absorptive in the low end but very reflective above 200 hz. This means that the reverb in an empty room is almost always bass shy.

These facts suggest a fairly simple recipe for tuning a room: Add absorption until you reach the point where the new material balances the original curve of the room, yielding a reverberation with a nice flat frequency response. Place the first panels near the speakers where they will eliminate interference paths, then spread the rest through the room to cut out any standing waves. If you are left with parallel hard surfaces, put diffusers on them. This method does not allow direct control of the reverberation time, but for any room smaller than 2000 cubic feet the RT60 should fall into the usable range.

A SAMPLE DESIGN

As an example of how to apply these principles, let us look at an ordinary room in a typical house. This room is rectangular, about 11' by 13' with an eight foot ceiling. There is a large closet at the back of the room and a window at the front looking onto a suburban street. The closet helps isolation because it provides something of a double wall between the studio and the living room.



INTERIOR TREATMENT OF A TYPICAL ROOM

There was a plush carpet over a thick pad on the floor, but no other absorptive material in the room to start with. The clap test in the empty room suggested a moderately long, primarily high frequency reverberation and produced the characteristic "chirp" of a severe standing wave problem.

SOME ISOLATION

After adding gaskets to the doors, isolation from the rest of the house is adequate as long as recording is limited to quiet times. (We checked this out before we moved in!) Noise from the street is an occasional problem which was helped a little by drapes on the window. An additional drape across the doorway made only a slight improvement in isolation and was really in the way, so we gave it up.

POSITIONING THE EQUIPMENT

After some experimentation, we decided to locate the speakers each side of the window. Since speakers tend to move gypsum as well as air, outside walls are always your first choice if you are concerned with sound control. Incidentally, these are obviously not near field speakers. Near field monitors should not be against a wall, but most large systems depend on a wall backing for extended bass response. The speakers were hung about 6 ft from the floor. This is a bit on the high side, but was necessary to allow the placement of a writing table underneath them.

The speakers wound up eight feet apart. This placed the "sweet spot" eight feet from the wall along the center line of the room. This in turn dictated the location of the mixing board and other equipment. Once the equipment was set in place, we checked for reflective phase interference from the console or cabinet tops. This can be tested with a mirror and a flashlight. Set the mirror on the console and hold the flashlight by your ear aimed at the mirror. If the light beam falls on or near the speakers there is a potential reflection problem. This can usually be fixed by propping up the back of the board.

WALL TREATMENT

At this point we were down to two problems: the rising frequency response of the reverberation and the standing wave. We attacked both problems at the same time with some carefully placed absorptive panels. These were made of R-19 fiberglass and measured 2 ft by 6 ft. (They do not need to extend down to the floor because the furniture scatters sound at that level.) Most of this absorption wound up on the walls near the speakers-- this cleaned up the last of the short delay reflections and resulted in a very clear sound image between the speakers. The absorption was brought along the side walls to soak up the standing wave. We wanted to keep the room symmetrical, so we spaced out the absorptive panels, winding up with a pattern where bare wall on one side was opposed by absorption on the other. A large section of absorptive wall near the left speaker created a dead corner for recording vocals.

The curtain over the window Is too light to be a really broadband absorber, but it combines with the low frequency absorption of the glass to give a reasonably flat overall effect. The carpet and wooden floor interact in much the same way.

We found the sound to be balanced in frequency when the walls were about one third covered with fiberglass. This left the side walls near the back of the room untreated so we added diffusion. This is provided by some homemade diffuser panels on one side and some very cluttered bookshelves on the other.

EVALUATION

This particular project cost about $30, and I must say I was quite pleased with the results. Gilligan stays in the bedroom down the hall, and quiet activity in adjoining sections of the house causes no problems with close mic recordings. A mono signal fed to both speakers appears to be centered precisely between them, and any imbalance in a stereo signal is immediately obvious. The room has a soft, comfortable ambience and the music produced there sounds just fine out in the real world.

Further Study

There is a lot more to acoustics than the principles explained here, but these techniques are the most likely to be effective in an existing small room with a limited budget. If you are planning a new building or just want to study more you can find some excellent articles in Handbook for Sound Engineers; edited by Glen Ballou, and Sound studio construction on a Budget by F. Alton Everest.

Making Connections

It is not necessary to have an engineering degree to be a successful electronic musician. In fact, most of us will never take the cover off of an electronic device, let alone design one. We do have to connect devices together however, and since the instructions for doing that are usually written by electronics engineers, we have to understand a few fundamental concepts.

First, remember that electricity flows in circles. That means that connecting two pieces of equipment requires two wires (called "hot" and "return"), that run from the output terminals of the device originating the signal to the input terminals of the receiving device. Do not confuse this with stereo; stereo connections require four wires, two for each channel.

Impedance Revisited

The toughest concept to grasp is impedance. You learned in essay five that impedance is related to the amount of current that will flow in a circuit with a particular voltage. (Low impedance implies high current.) You also learned that impedance can vary with the frequency of a signal.

When we discuss the impedance of a device, we are talking about either the input impedance or the output impedance. (Most devices have an input and an output.) The input impedance of a device is an indication of how much current is required to make the device function. It is a real quantity that can be measured with a little ingenuity. The output impedance of a device is a fiction that suggests how much current is available. It really indicates the expected input impedance of the next device along the line. In other words, when a manufacturer says "output impedance: 150 ohms"; he really means: "This thing should work properly if it is connected to a device of 150 ohms or greater input impedance."

There was a time when output impedance (also called source impedance[11]) was carefully matched to the input impedance of a connected device, because that is the most efficient way to transmit power. However, since a connection will work if the input impedance is higher than the source impedance but will not if the input impedance is

[11]If you are speaking of a particular device, the source impedance is the output impedance of whatever is connected to the input, and the load impedance is the input impedance of whatever is connected to the output.

too low (low impedance requires more current, remember.) engineers tried to design circuits with somewhat higher input impedance than actually necessary, and lower output impedance. (A low output impedance implies that lots of current is available.) There is a limit to how high input impedances can go; if it is too high, unwanted low current sources like radio stations will begin to affect the circuit. The usual ratio of input impedance to source impedance is 10 to 1[12] .

Signal Level

Another important concept is signal level. This is a measurement of the voltage expected at the output of a device. If the device has a VU meter this is the output at 0VU; otherwise, it is the strongest undistorted output available. Again, there are elements of fiction in the published specs. Engineers like to add extra capacity (called headroom) in the expectation that the device will be operated incorrectly. The measurement is specified in dB, which as you have learned from the essay on dBs and dynamics, requires a reference value. The reference will either be 1 milliwatt (about .775 volts if the load is 600 ohms) or 1 volt (with the circuit unconnected). The first reference is noted dBm, the second dBv; the difference between them is about 2 dB.

Both the impedance and signal level have to be compatible if devices are to work properly together. Since manufacturers want their products to be useful, they tend to match existing standards for similar devices. Here is a rundown of what you will probably find:

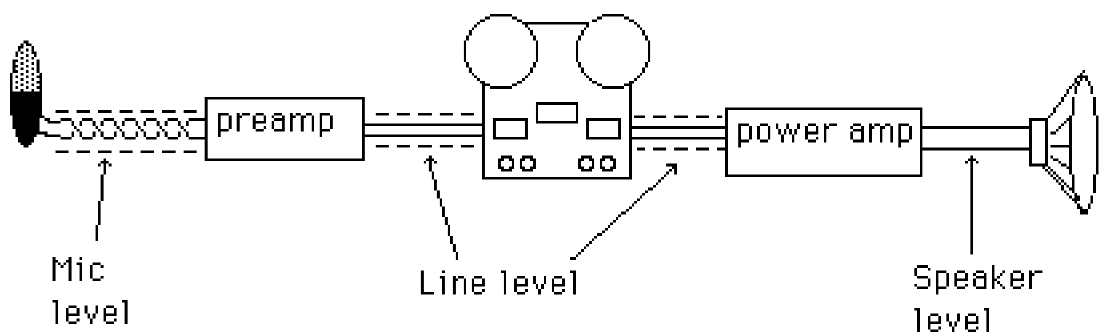


Fig. 1 typical connections

---

[12]You still see some equipment that requires proper loading to provide a flat frequency response. Such a device will be marked "terminate with 600 ohms". If it is not used with a 600 ohm load, a 600 ohm resistor should be connected across the output pins.

Microphone Signals

Microphone levels are of necessity very weak signals, generally around –60dBm. (The specification is the power produced by a sound pressure of 10 µBar) The output impedance will depend on whether the mic has a transformer[13] balanced output (see essay 20 for details). If it does not, the microphone will be labeled "high impedance" or "hi Z" and must be connected to an appropriate input. The cable used must be kept short, less than 10 feet or so, to avoid noise problems.

If a microphone has a transformer, it will be labeled "low impedance", and will work best with a balanced input mic preamp. The cable can be several hundred feet long with no problem. Balanced output, low impedance microphones are expensive, and generally found in professional applications. Balanced outputs must have three pin connectors ("Cannon plugs"), but not all mics with those plugs are really balanced. Microphones with standard or miniature phone plugs are high impedance. A balanced mic can be used with a high impedance input with a suitable adapter.

Any microphone cable must be shielded to reduce pickup of radio signals and other unwanted electromagnetic effects. The shield is a tube of braided thin wire that surrounds the current carrying wires and is connected to the metal case of the microphone preamp. Shielding works by the "Faraday effect", which states that a charge induced in a metal body accumulates at the outer surface of that body. If the microphone is unbalanced, there may only be one wire within the shield; return current is brought back to the mic along the shield itself.

Speaker Lines

Speaker connections are also specified in terms of impedance. The tradition is to match amplifier impedance to speaker impedance. Again, it does not cause a problem to connect a low output impedance to a somewhat higher input impedance, but the opposite situation can result in a blown power amp. Most speaker systems are specified as 8 ohms, although a few are 4 ohms. Amps will clearly state the impedance expected. (This spec is another fiction. Speaker impedance varies wildly with frequency; the rated number is the lowest value encountered.)

---

[13]A transformer is a device that changes a low voltage, high current signal into a high voltage, low current signal or vice versa (the product of current and voltage remains constant). The impedance of the high current side will be low, and the impedance of the low current side will be high.

Amplifiers and speakers are rated for power output or power handing capacity. You should match these figures also; otherwise, you run the risk of damaged speakers or not enough sound.

Transformer coupled speaker lines are used in situations where one amp must drive a lot of speakers (as in, for instance, a hotel PA system.) These are called high impedance or "70 volt" systems, because the transformers are put in backwards compared to mic and line connections. Equipment designed for this approach is not appropriate for hi-fi music systems.

Speaker connections should not be shielded, because the power transmitted in speaker lines is much greater than any induced signal, and there are side effects of shielding that can change the frequency response.

Line Level Signals

Most of the interconnections between audio devices pass line level signals. There are two standards in use in this country: consumer products usually operate at –10dBv, while professional devices produce outputs at +4dBm. In professional installations, all connections are brought to a central patch bay for convenience in changing configurations of equipment. This may entail rather long cables, so transformer balanced inputs and outputs are often used. The distinction between "consumer" and "pro" audio gear is rather vague; it really has more to do with reliability than fidelity. There is even a class of equipment called "semi–pro", which has most of the features of pro gear but operates at the –10dBv level.

Interconnecting the two kinds of equipment can be a headache, but it is a common problem, particularly in electronic music. The usual solution in a studio is to convert everything to match the level of the majority of equipment. The decision to convert all low level gear to +4dBm balanced line is an expensive one, as that requires buffer amplifiers or transformers for all the affected inputs and outputs, but a pro studio is an expensive item anyway. Converting a few +4dBm devices to consumer levels is fairly simple. The output level can be unbalanced and lowered with a 15 cent part, and most pro gear is designed with enough input gain to make up the difference with no modification at all. Unbalanced patch bays are satisfactory if total wire length is kept under thirty feet or so.

Line level signals must be shielded.

A new signal level is becoming established for **synthesizer outputs**. Some digital instruments seem to be designed so that their maximum output is

–10dBv. The typical output of an interesting voice will be maximum only briefly; a VU type measurement of the output would be around –20dBv. These instruments must have their outputs amplified if they are to be used in +4DBm studios, and can even benefit from a buffer when used with low level gear.

Grounding

An electronic device responds to changes in current flowing through the input circuitry, which implies that a changing voltage is applied across the input terminals. In a balanced system, the voltage at one terminal mirrors the other; for example, if the voltage on the hot lead rises from 0 to + 4 volts, the voltage on the return lead would fall from 0 to -4 volts. Current flows equally in the hot and return leads and relative differences in reference voltage from one device to another are unimportant.

In an unbalanced system, the voltage on one lead changes while the voltage on the other remains steady; the hot lead would rise from 0 to +8 volts and the return would stay at 0 volts to produce the same result as the previous example. In this system, it is important that 0 volts be properly defined at both ends of the wire. If the two devices disagree as to what zero means, current will flow along the shield, often at 60 hz AC. This current will induce hum into the hot lead. The techniques used to establish a reliable zero volt level are lumped together as "grounding" and often seem to involve as much witchcraft as engineering.

WIRES

Hi-fi enthusiasts talk a lot about the effects of the wire that is used to interconnect audio equipment. As with many things audiophiles are concerned with, this is a case of valid basic principles followed to extremes. Herewith; a rational point of view:

The prime concerns for microphone cables are flexibility and durability. The main difference between a cheap cable and an expensive one is the effectiveness of the shield. The connectors are expensive enough that no one has thought to gold plate them.

Balanced line cables are identical to mic cables. High impedance unbalanced connections are somewhat sensitive to the kind of wire used. With shielded wire, there is some capacitive coupling between the center conductor and the shield. Capacitive coupling acts as a low impedance path for signals of high frequency. Since current follows the path of least impedance, a capacitive cable connected to a high impedance input will show a low pass characteristic; in other words, some high frequency energy will be shorted out. The actual frequency at which this is apparent depends on the material the cable is made of, the length of cable, and the actual input impedance of the device. In my experience, with typical equipment and ordinary cable, the problem develops with runs longer than twenty feet. The only exception is with phono inputs, which with their extremely high impedance are sensitive to cables longer than three feet. There is no point in gold plating a phono plug unless it is connected to a gold plated jack[14]. (The gold is supposed to prevent the buildup of resistive corrosion. This can happen with nickel plated connectors if they are in a warm humid place, but the corrosion is removed each time the plug is inserted. Anyway, it is the jack that corrodes, not the plug.)

In speaker wires the important issue is resistance of the wire itself. With circuit elements connected in series, the voltage across each is proportional to the impedance of each. If a speaker impedance is eight ohms, and the wire impedance is two ohms, twenty percent of the power of the amplifier will be used up in the wire. The resistance of a piece of wire depends on its length and thickness, or **gauge**[15]. A six foot chunk of #22 "speaker wire" sold by Radio Shack measures 0.3 ohms, so you probably should not use it for very long runs (I never use the stuff). Ordinary #18 lamp cord ("zip cord") is less than a third as resistive, and I routinely use it for runs of up to twenty feet or so. For very long runs, I use #14 electrical cable and count on some power loss. You can buy so-called "monster cable", which has very low loss over fairly long runs, but the cost of the cable is more than the cost of the wattage you are wasting. You should never use shielded cable for speaker leads.

[14]A gold connector is a good idea on a power amp, or any device that is exposed to high temperature or humidity.
[15]At audio frequencies, the impedance of a wire is almost all resistance.

Connectors

There are two parts required for a connection, a jack
Which is usually on the equipment, and a plug, which is usually on the cable. A plug with a prong on it is called a “male” connector, and a plug with receptacles that fit the prongs is called “female”.

As this picture shows, there are an amazing variety of connectors used in audio. The most common kinds are called phono, Phone, and XLR (also called Cannon), which sort of refers to who invented them.

Phono

The phono (or RCA) connector is designed to be cheap. It only has one central conductor inside the shield, so it is always used for unbalanced wiring. It is usually found with –10dBv audio signals, although it is seen a lot in consumer video connections and the SPDIF variety of digital audio.

They are sometimes gold plated, but they are still cheap. The jacks wear out fast if you constantly connect and disconnect them. On the other hand, they need to be disconnected and reconnected at least once a year, because the nickel plating corrodes. (The act of inserting cleans the corrosion off.)

Phone

The phone plug comes in both single conductor and two conductor versions. Single conductor is used for unbalanced –10 dBv audio signals. They are favored for musical instrument applications because they can be replugged many more times than RCAs, but don't take up as much room or cost as much as XLRs. They are also occasionally used for speaker connection on portable systems.

Dual conductor phone plugs are used for balanced +4dBm and for stereo headphones.

Phone plugs can have either plastic or steel cases- the steel ones don't break when you step on them.

Phone plugs also come in a miniature size-- either 1/8 inch or 3.5mm. These are not the same size! A 1/8 plug in a 3.5 mm jack will be loose. Fortunately, as most of the world is metric, 1/8 inch plugs are pretty rare. Miniature phone plugs are generally fond on $10 walkmans and $2000 computers.

There is yet another variant, called the MIL P-642. This is the original phone company plug (circa 1920), as codified by the military. These are used with high end patch bays. There is also a miniature version of this. (Patch bays are panels with dozens of jacks on them—they are found in elaborate audio installations such as recording studios. All of the connections for all of the equipment is brought to the patch bay, where short patch cords are used to make connections. Patch bays usually have built in connections, so the cords are only needed for special hookups..)

XLR

XLR connectors used to be called cannon connectors, but the Switchcraft version (model XLR) was more popular. XLR plugs come in both male and female versions, as do the jacks. Male XLR plugs are usually associated with outputs. They are found in balanced microphone and balanced +4dBm situations. They are expensive but rugged. I have some I've been using for 40 years.

XLR type connectors can have up to 7 pins in them.

DIN

DIN connectors follow a European standard for –10 dBv audio connections. The standard allows for two channels each way, so the most common plug has 5 pins (one for ground). In the US they are only used for MIDI connections.

Digital Audio Connections

There are several standards for interconnecting digital audio devices, and more seem to appear every day. There are committees of organizations such as the Audio Engineering Society (AES) and Institute of Electrical and Electronic Engineers (IEEE) developing standards, but the manufacturers seem to prefer a process of putting out their own versions and letting the market decide which survives.

All of the systems are based on high speed serial data transmission, as this is a well understood and reliable system. The systems are all pretty much the same electrically- the major difference are in formatting of the data, and the type of connector. They all handle sample rates up to 96khz and word sizes up to 24 bits. These are the most common types:

AES

Developed by the Audio Engineering Society early on in an attempt to forestall what happened, this is a simple system that uses the balanced (XLR connector) wiring already in use in pro audio. It can be connected to standard network wiring (CAT5) with a simple transformer, which makes it simple to add to older buildings. All that is transmitted is the audio signal, two channels in a cable. There is a multi channel variant called MADI that has not really caught on.

S/PDIF

This was developed by consumer gear manufacturers to interconnect home stereo equipment. It's supposed use ordinary unbalanced (phono) wiring, but it was soon realized that it would only work properly with video grade cable. Even with good cable, things like adapters and patchbays can keep the signal from getting through. Alternatively, a SPDIF signal can be sent through optical cables, an expensive but reliable solution. SPDIF transmits more than just the audio. There is provision for "subcode" which can include things like titles and copyrights, as well as track and index marks.

After a series of lawsuits, SPDIF was amended to include the serial copyright management system (SCMS). This allows the maker of a CD or any digital recording to include codes that restrict the number of copies that can be made. (You can either copy freely, make a copy that can not be copied, or cannot copy).

Alesis Lightpipe

The Alesis ADAT recording system had a profound effect on the audio industry, bringing high quality multitrack capability to budget studios. One feature of the device was a digital interconnection that would carry 8 channels of audio over an optical cable. This made it easy to copy ADAT tapes, and is also used on other Alesis products like synthesizers and effects boxes. Other companies are just starting to bring out products that feature the interconnection, such as multitrack computer audio interfaces.

TDIF

When Tascam brought out their competing multitrack digital system, they also introduced a multi track interconnection system. This is basically four SPDIF lines in a single cable with a computer style (DB25) connector. Only Tascam and devices (like digital mixers) designed to interface with Tascam products are using it.

Wordclock

If two digital devices are going to work properly together, they have to have exactly the same sample rate. If one is 48000 Hz and the other is 48001 Hz, eventually the two will be on different words in the audio stream. (If you are copying from one to another, this will result in a pop in the sound. If you are mixing similar material from two sources, there may be flanging.) Since it is impossible to build clocking circuits more accurate than a few parts per million, digital audio devices need a common source to establish the sample rate.

All of the interconnection systems described are self clocking, meaning the receiving device has to lock its internal circuitry to the data rate of the incoming signal as its own sample rate. This can lead to interesting surprises, such as a gross pitch change when a tape is played back. If carefully used[16], this system works well enough, but if there are three or more digital devices it gets harder to make them talk flexibly to each other. Large digital studios define one device as the "wordclock" and connect separate synchronizing cables to the rest of the digital gear. To work with this system, a device must have a special wordclock input and a switch that sets the synchronization source. Wordclock is carried on professional grade video cable with BNC connectors.

---

[16] Just connect a DAT running at 44.1 khz to one that is set at 48. Then disconnect them and play back the tape you just made.

Wordclock is a simple signal running at the sample rate. Some systems are touting "superclock" that is 2 or 4 times the sample rate and presumed to be more accurate.

Converting Sound Into Numbers

In a digital recording system, sound is stored and manipulated as a stream of discrete numbers, each number representing the air pressure at a particular time. The numbers are generated by a microphone connected to a circuit called an ANALOG TO DIGITAL CONVERTER, or ADC. Each number is called a SAMPLE, and the number of samples taken per second is the SAMPLE RATE. Ultimately, the numbers will be converted back into sound by a DIGITAL TO ANALOG CONVERTER or DAC, connected to a loudspeaker.

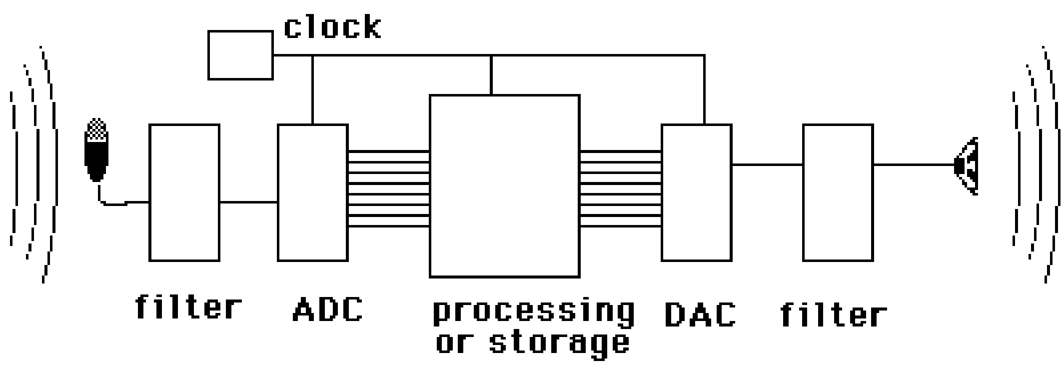


Fig. 1 The digital signal chain

Binary Numbers

Figure 1 shows the components of a digital system. Notice that the output of the ADC and the input of the DAC consists of a bundle of wires. These wires carry the numbers that are the result of the analog to digital conversion. The numbers are in the binary number system in which only two characters are used, 1 and 0. (The circuitry is actually built around switches which are either on or off.) The value of a character depends on its place in the number, just as in the familiar decimal system. Here are a few equivalents:

BINARY DECIMAL
0=0
1=1
10=2
11=3
100=4
1111=15
1111111111111111=65535

Each digit in a number is called a BIT, so that last number is sixteen bits long in its binary form. If we wrote the second number as 0000000000000001, it would be sixteen bits long and have a value of 1.

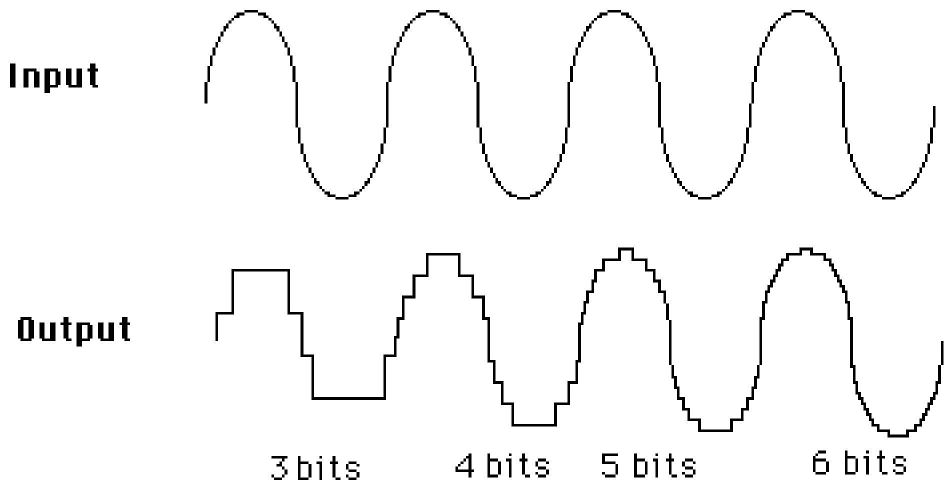


Fig. 2 Effect of word size

Word Size

The number of bits in the number has a direct bearing on the fidelity of the signal. Figure 2 illustrates how this works. The number of possible voltage levels at the output is simply the number of values that may be represented by the largest possible number (no "in between" values are allowed). If there were only one bit in the number, the ultimate output would be a pulse wave with a fixed amplitude and more or less the frequency of the input signal. If there are more bits in the number the waveform is more accurately traced, because each added bit doubles the number of possible values. The distortion is roughly the percentage that the least significant bit represents out of the average value. Distortion in digital systems increases as signal levels decrease, which is the opposite of the behavior of analog systems.

The number of bits in the number also determines the dynamic range. Moving a binary number one space to the left multiplies the value by two (just as moving a decimal number one space to the left multiplies the value by ten), so each bit doubles the voltage that may be represented. You remember from the essay on dB that doubling the voltage increases the power available by 6 dB, so we can say the dynamic range available is the number of bits times 6 dB.

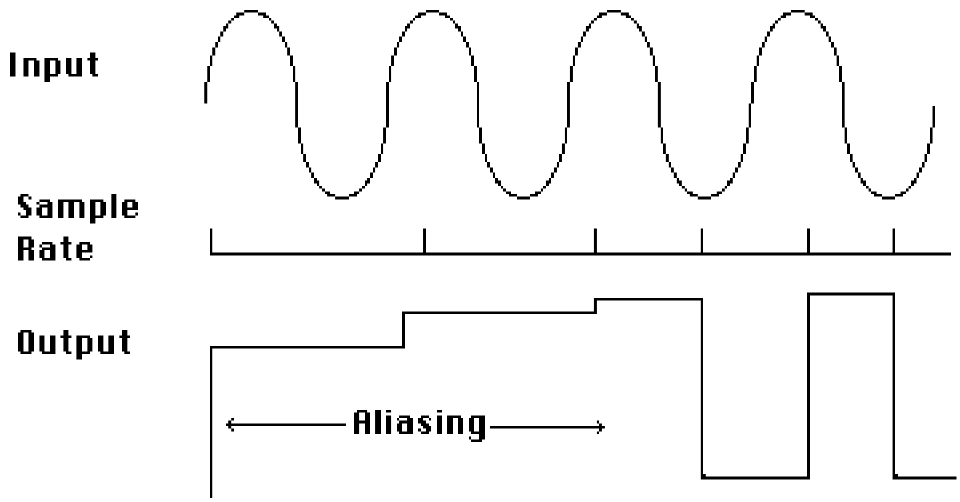


Fig. 3 Effects of sample rate changes

Sample Rate

The rate at which the numbers are generated is even more important than the number of bits used. Figure 3. illustrates this. If the sampling rate is lower than the frequency we are trying to capture, entire cycles will be missed, and the decoded result would be too low in frequency and might not resemble the proper waveform at all. This kind of mistake is called ALIASING. If the sampling rate is exactly the frequency of the input the result would be a straight line, because the same spot on the waveform would be measured each time. This can happen even if the sampling rate is twice the frequency of the input if the input is a sine or similar waveform. The sampling rate must be greater than twice the frequency measured for accurate results. (The mathematical statement of this is the NYQUIST THEOREM.) This implies that if we are dealing with sound, we should sample at least 40,000 times per second.

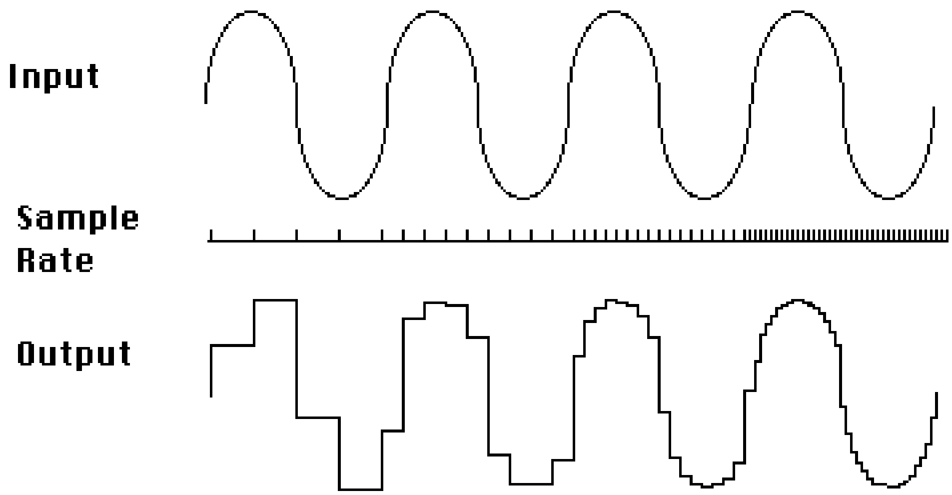


Fig. 4 Effect of increasing sample rate

The Nyquist rate is the lowest allowable sampling rate. For best results, sampling rates twice or four times this should be used. Figure 4 shows how the waveform improves as the sampling rate is increased. A Fourier analysis of the result would show that everything belonging in the signal would be there along with a healthy dose of the sampling rate and its harmonics. The extra junk must be removed with a low pass filter set just higher than the highest desired frequency. (An identical filter should be placed before the ADC to prevent aliasing of any unsuspected ultrasonic content, such as radio frequency interference.) If the sampling rate is only twice the frequency of interest, the filter must have a very steep characteristic to allow proper frequency response and satisfactorily reject the sampling clock. Such filters are difficult and expensive to build.[17]

Digital audio recording

Once the waveform is faithfully transformed into bits, it is fairly easy to record. The major problem is finding a scheme that will record the bits fast enough. If we sample at 44,100 hz, with a sixteen bit word size, in stereo, we have to accommodate 1,411,200 bits per second. This seems like a lot, but it is within the capabilities of techniques developed for video recording (in fact, some digital audio systems are built around VCRs). To record on tape, a very high speed is required to keep the wavelength of a bit at manageable dimensions. This is often accomplished by moving the head as well as the tape, resulting in a series of short tracks across the tape

[17] Most of the complaints about the "digital sound" of the first CD units were caused by filter problems.

at a diagonal. On a Compact Disc, the bits are microscopic pits burned into the plastic by a laser.

Even with these techniques, the bits are going to be physically very small, and it must be assumed that some will be lost in the process. A single bit can be very important (suppose it represents the sign of a large number!), so there has to be a way of recovering lost data. This issue is known as ERROR CORRECTION, and has two problems; how to detect an error, and what to do about it.

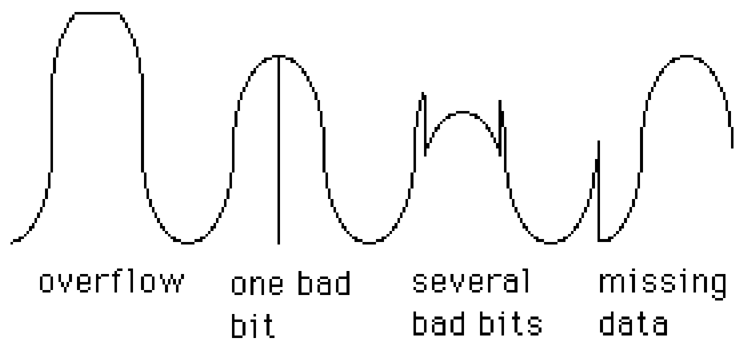


Fig. 5 Effects of data errors

Error Correction

The most common error detection method is PARITY computation. A extra bit is added to each number which indicates whether the number is even or odd. When the data is read off the tape, if the parity bit is inappropriate, something has gone wrong. This works well enough for telephone conversations and the like, but does not detect serious errors very well.

In digital recording, large chunks of data are often wiped out by a tape dropout or a scratch on the disk. Catching these problems with parity would be a matter of luck. To help deal with large scale data loss, some mathematical computation is run on the numbers, and the result is merged with the data from time to time. This is known as a CYCLICAL REDUNDANCY CHECK or CRC. If a mistake turns up in this number, an error has occurred since the last correct CRC was received.

Once an error is detected, the system must deal gracefully with the problem. To make this possible, the data is recorded in an odd order. Instead of word two following word one, as you might expect, the data is INTERLEAVED, following a pattern like:

words 1,5,9,13,17,21,25,29,2,6,10,14,18,22,26,30,3,7,15,19,27 etc.

With this scheme, you could lose 8 words, but they would represent several isolated parts of the data stream, rather than a large continuous chunk of waveform. When the CRC flags a bad block, the system can use the previous and following words to reconstruct a passable imitation of the missing one. One of the factors that makes up the price difference in various digital systems is the sophistication available to reconstruct missing data.

The Advantages of Digital

You may be wondering about the point of all of this, if it turns out that a digital system is more complex than the equivalent analog circuit. It is possible to build analog circuits that match digital performance levels, but they are very expensive and require constant maintenance. Digital circuits are complex, but very few of the components must be precise; most of the circuitry merely responds to the presence or absence of current. Improving performance is usually only a matter of increasing the word size or the sample rate, which is achieved by duplicating elements of the circuit. The bottom line is that good digital systems are cheaper than good analog systems.

Digital devices usually require less maintenance than analog equipment. The electrical characteristics of most circuit elements change with time and temperature, and minor changes slowly degrade the performance of analog circuits. Digital components either work or don't, and it is much easier to find a chip that has failed entirely than one that is merely 10% off spec. Many analog systems are mechanical in nature, and simple wear can soon cause problems. Digital systems have few moving parts, and such parts are usually designed so that a little vibration or speed variation is not important.

In addition, digitally encoded information is more durable than analog information, again because circuits are responding only to the presence or absence of something rather than to the precise characteristics of anything. As you have seen, it is possible to design digital systems so that they can actually reconstruct missing or incorrect data. You can hear every little imperfection on an LP, but minor damage is not audible with a CD.

The aspect of digital sound that is most exciting to the electronic musician is that any numbers can be converted into sound, whether they originated at a microphone or not. This opens up the possibility of creating sounds that have never existed before, and of controlling those sounds with a precision that is simply not possible with any other technique.

Another Way Of Looking At Audio

Throughout this reader I have presented audio information in two ways, as a waveform, or as a spectral plot or graph. The digitization method just described digitizes the waveform, and is simple, cheap, and the most common way of doing it. But there are sometimes advantages to taking the spectral approach to digital audio.

Luckilly, we can have it both ways. With computer technology, there is a straight forward (I won't say simple) way of transforming data from a waveform to a spectrum and back. The algorithm used is called the Fast Fourier Transform or FFT. It leaves us with a series of numbers giving spectral representation of the sound instead of a piece of waveform. This is called a frequency domain representation. (As opposed to a time domain representation, which refers to the waveform.) You can get the original waveform back with the Inverse Fast Fourier Transform.

Of course a spectral plot is a snapshot of sound at one moment, but if we take a series of such snapshots we can accurately represent a changing sound much the same way a series of pictures make television. In order to do this with any fidelity, several factors must be considered.

Windows

The fft cannot work on a single sample—there must be a series of samples, and the more samples the more accurate the transform will be. What you really need is a complete cycle of the lowest component of the waveform. The number of samples defines a duration that is called the window. A typical window has 1024 samples or about 23 ms, enough to get down to about 40 hz. If you cut sound into 40 chunks a second and try to play it back, a 40 hz buzz will be part of the result, so we do a trick—we overlap windows, and fade from one to another on both analysis and playback. The way the windows fade in and out is important[18], as is the number of windows overlappping at once (8 is considered good). With a proper choice of window and sample rate, the transformation to and from the frequency domain is essentially inaudible. (Of course as with all audio choices, if you cut corners, the results can be pretty ugly.)

The application the Fourier transform to a window of a waveform gives a block of data called a “frame’. The frame is divided into “bins”; one bin per frequency channel in the analysis. A bin actually consists of two

[18] You can get your name on a window shape, so you hear about “gaussian” windows and “Hamming” windows.

numbers, one giving amplitude and one giving phase for the frequency represented. These are usually output as two separate lists, and it is the amplitude list that is displayed[19]. This display is typically a row of vertical lines. The frequency spacing between the lines is the sample rate divided by the number of samples in the window.

There is an important distinction between this display and what you see in acoustics texts. The spectrum usually given for (for instance) a flute is derived knowing what pitch the flute is playing, so the window size can be adjusted to fit the period of the waveform. In the real world we don't have that knowledge, so we choose arbitrary windows based on the nature of the music played and the computing power available. In sophisticated systems, the window size is adjusted automatically, either attempting to track the pitch of the signal or by measuring the difference between the original and resynthesized signals.

There are other ways of converting waveforms into frequency domain representations, such as emulations of banks of band pass filters. These may require more computing power than the direct methods (the FFT always turns up somewhere) but give more musically useful results, such as spectra with an equal number of bands per octave.

Frequency Domain Processing

Once we have frequency domain data, what can we do with it? Well, this is a really hot area of development right now, and new applications appear every week. Here are some of the interesting ones:

Vocoding

Taking the bins of a signal and applying them as filter parameters to another signal results in something that sounds somewhat like each. This can give us singing dogs or steam radiators that "morph" into sopranos.

Time Stretching

If the signal is represented as a series of snapshots of spectra, you can play the snapshots back at a different rate than they we taken. (The number of samples between the start successive frames is called the "hop size". If the playback hop size is different from taht used in the analysis, you will

[19] Because of the math underlying this process, the two lists are sometimes called the "real" and "Imaginary" parts of the spectrum.

get a time change.) This will change the duration and tempo of tones without changing the pitch. There are plug ins for most digital editors that allow this.

Pitch Changing

Given the ability to do the above, it's no trick to change the pitch of a recording by conventional means (resampling) and then correct the time with an appropriate time stretch. The result: new pitch, same tempo.

Formant Correction

The munchkinization of sounds you get on samplers results from the transposition of formants in an instrument's tone when the pitch is transposed. If you analyze the spectrum of the untransposed sound and apply that to the same sound after transposition, you can correct the formants and get a more realistic result. The high end sampler manufacturers are all into this.

Resysnthesis

If you apply the bin amplitudes to the amplitudes of sinewave oscillators[20] of the same frequencies, the result is pretty much the same as the original signal. However, changing the waveform or pitch of the oscillators can produce very interesting sounds. There's a Macintosh program called MetaSynth that lets you do this, and also lets you redraw the spectrum (or create your own for that matter) in a graphic display.

Data Compression

If you want to shove sound through the internet, you have to use a very low sample rate or take a long time to transmit a file, at least if the waveform representation is used. Spectral representation of sound requires even more data than the time domain, at least at the start. However, in most music, many of the bins will be zero, and others can be discarded because they are psycho-acoustically masked. If you only transmit the significant bins, you can do reasonable audio at internet friendly data rates. That's what's going on with MP3 and Minidisc™ music.

Further reading

[20] Digital oscillators, naturally—you need a thousand or so.

Ken Pohlman; Principles of Digital Audio give the basic principles and details of consumer audio formats.
Curtis Roads: The Computer Music Tutorial gives the math.

Making Waves From Numbers

Wavetables

Nearly all digital music systems use some form of **wavetable synthesis** to generate signals. The wavetable is a section of memory that contains a list of values corresponding to the desired waveform. The computer reads the numbers from the list at a steady rate (the sampling rate), repeating the table when the end is reached. If the table contains a single cycle of the waveform, the frequency produced would simply be the sample rate divided by the number of values in the table:

$$F=SR/n$$

The output is a very high fidelity copy of the waveform:

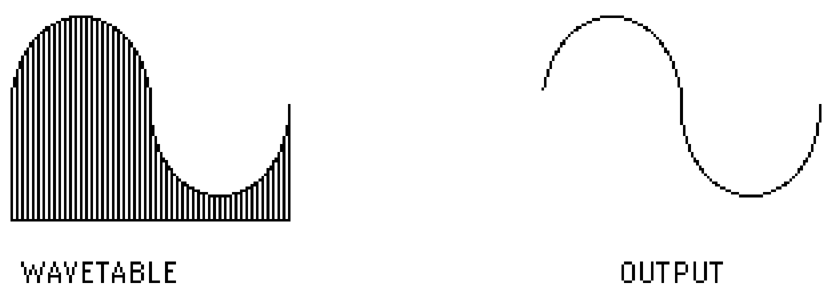


Fig. 1 Using all values in the wavetable gives an exact copy of the stored waveform.

To produce higher pitches, the system skips some values each time. The number of values skipped is the **sampling increment**. A sampling increment of 4 (reading every fourth value) gives an output two octaves higher than the original.

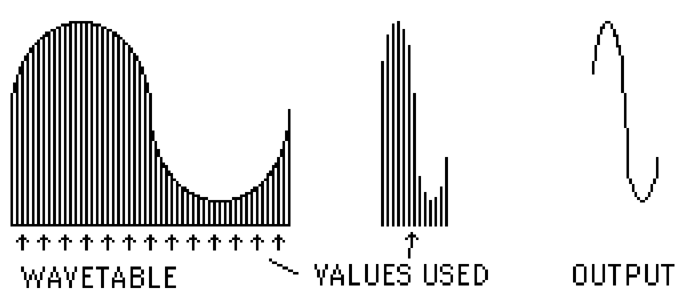


Fig. 2 Effect of increases sampling increment.

The frequency produced is the original multiplied by the sampling increment.

F=SIxSR / n

It is possible to have fractional increments; the computer interpolates between listed values, or simply reads a number twice[21]. (If all numbers are read twice, the pitch is one octave down.) This distorts the waveform somewhat.

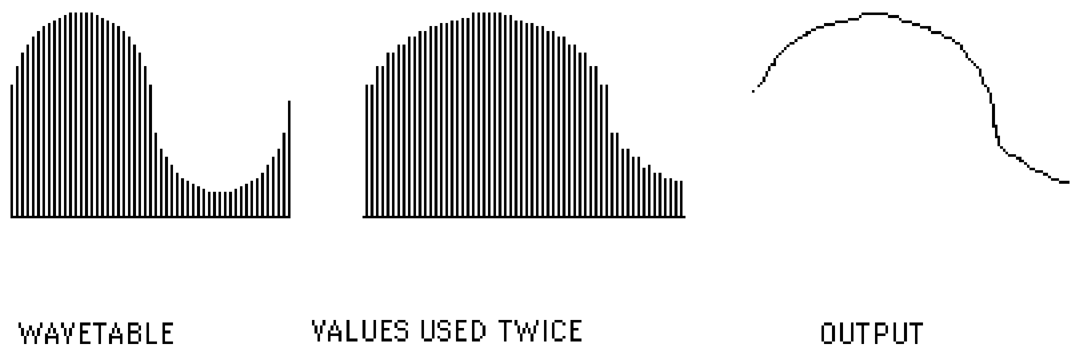


Fig. 3 A sampling increment of .5.

Amplitude control can be added with a variety of techniques. The straightforward way is to simply multiply the sample value by a number derived in a similar manner from an envelope table. A more efficient technique is available if the waveform is a sine. During each sample period two values are taken from the table: one found the usual way, and another at a location offset from the first according to the envelope. The two values are then added before moving to the output. The sum of two sine waves that are out of phase is a sine of amplitude determined by the phase difference. If the offset equals half the table size, the output will be zero.

Frequency Modulation

Frequency modulation is a very powerful algorithm for creating sounds. The heart of the technique is the way extra tones (sidebands) are created

---

[21]In point of fact, the number of the most recent value chosen is kept in a register known as the phase accumulator, which has more precison than neccessary to handle the size table used. The high part of the p.a. points into the table, and the low bits contain the fraction. The sampling increment is added to the phase accumulator during each iteration, which will produce the appropriate stepping action.

when one oscillator is used to modulate the frequency of another[22]. These sidebands are symmetrically spaced about the frequency of the carrier[23], and the size of the spaces is equal to the frequency of the modulator. Increasing modulation increases the number of sidebands, but the amplitude of the sidebands varies in a rather complex way as the modulation changes.

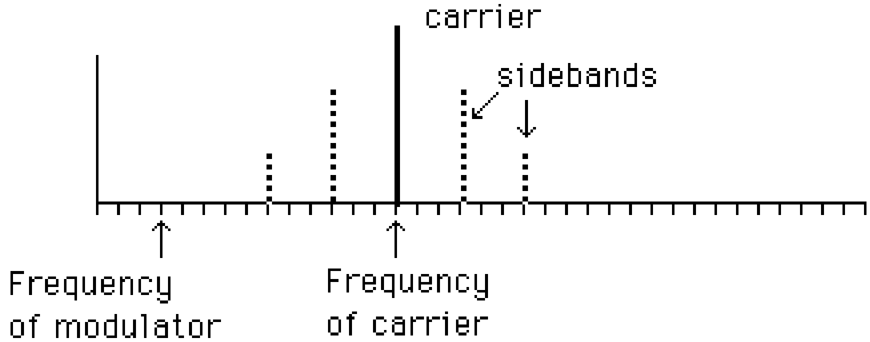


Fig. 4 Spectrum of simple frequency modulation

There are three kinds of relationship between the frequencies of the carrier and modulator, and each produces a different family of sounds.

If the modulator and carrier are the same frequency, all of the sidebands will be harmonics of that frequency, and the sound will be strongly pitched. You may wonder how that can be if there are supposed to be sidebands at frequencies lower than the carrier. If the spacing of the sidebands is the same as the carrier frequency (as it will be if modulator equals carrier), the sideband just below the carrier will be zero in frequency. The sideband just below that will be the carrier frequency, but negative. When that concept is applied in reality, the result is the carrier frequency, but 180° out of phase. That sideband therefore weakens or strengthens the fundamental, depending on the modulation index. Further low sidebands interact with upper sidebands in the same way. The regularity of the sidebands produces the strongly harmonic sound usually associated with synthesizers, but if the modulation index is changed during the note (dynamic modulation) the intensity of the sidebands will change in some very voicelike effects.

---

[22]Digitally speaking, the value obtained from one wavetable lookup is added to the sampling increment for another wavetable lookup.
[23]The <u>carrier</u> is the oscillator we listen to; the <u>modulator</u> is an oscillator that changes the frequency of the carrier at an audio rate.

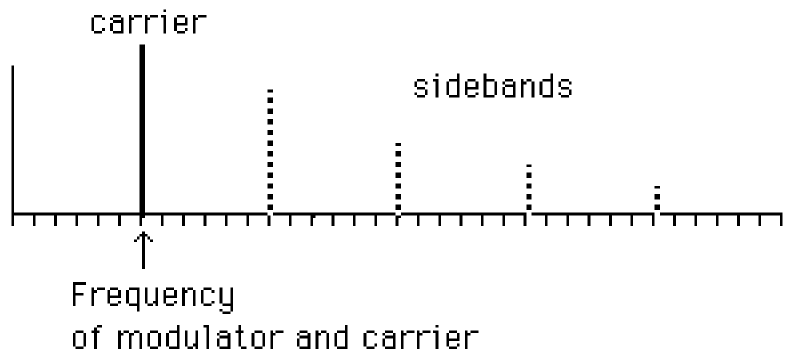


Fig. 5 Harmonic spectrum generated with FM

If the frequencies of the carrier and modulator are different but rationally related, the result will again be strongly harmonic, and the pitch will be the root of the implied series. (For instance, frequencies of 400hz and 500hz imply a root of 100hz.) If the carrier is the higher frequency, the resultant sound will be quite bright, sounding like a high pass effect at low modulation and becoming very brash as the modulation increases. The frequency of the carrier is always prominent. If the carrier is the lower frequency, the sound will have "missing" harmonics, and those that are present will appear in pairs (see figure 6). At low modulation index, you will hear two distinct pitches in the tone; as the index is increased, the timbre of the upper pitch seems to become brighter.

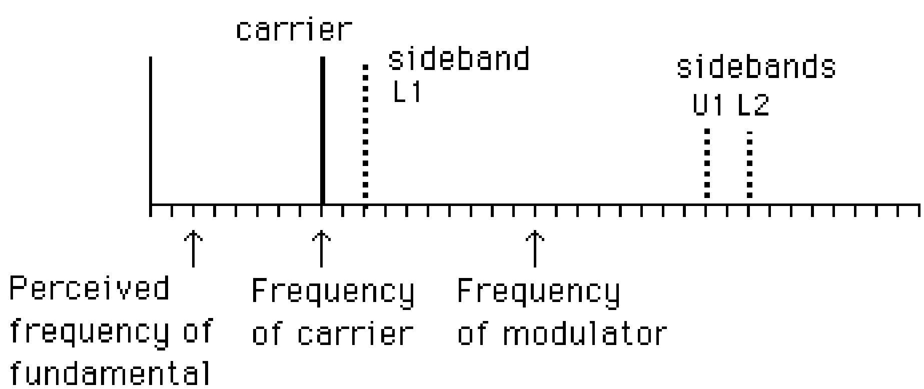


Fig. 6 FM with modulator frequency higher than carrier

If the frequencies of the carrier and modulator are not rationally related, the tone will have a less definite pitch, and will have a rich sound. Very often the effect is of two tones, a weak pure tone at the carrier frequency, plus a rough sound with a vague pitch. With careful adjustment of the operator level of the modulator, the carrier tone can be nearly eliminated. If the frequencies of the carrier and modulator are close to, but not quite harmonic, **timbral beating** will occur at a rate that equals the difference.

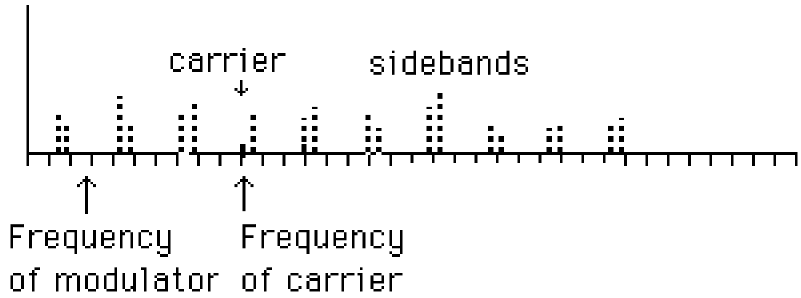


Fig. 7 Nonharmonic FM spectrum

A particularly powerful aspect of frequency modulation as a music generating technique is that the timbres can be dynamically varied. By applying an envelope function to the amount of modulation or the frequencies of carrier and modulator, sounds can be produced that have a life and excitement far beyond that available with the older synthesis methods.

Granular Synthesis

In granular synthesis, very short chunks of sound (40 samples or so) are combined into a continuous complex result. To reduce buzzing related to the size of the grains, they are actually overlapped and faded in and out in a manner similar to the fft windowing scheme. An interesting application of this process is to chop up an input signal, change the spacing of the grains slightly, and play it out. This can give real time pitch shifting. Granular synthesis is particularly useful for creating rough or noise-like sounds.

MIDI

History

Before 1983, all schemes for connecting synthesizers to computers were homemade or sold in small quantities by tiny companies. This led to a variety of devices, mutually incompatible and so idiosyncratic that only their inventors could write software for them. The usual approach was to connect extra circuitry to the innards of the computer that either generated sounds directly or provided several channels of voltage control for modular synthesizers. In 1983, several synthesizer manufacturers agreed on a communications protocol that would allow keyboard synthesizers to control each other (MIDI). This was very quickly picked up for computer applications, and today we have a mix and match situation, where any of several computers can be connected to one or more synthesizers, provided you have the proper software. MIDI is not perfect (the keyboard orientation and the rather slow data rate cause hassles), but it has provided an impetus for the development of software, has lowered the costs of computer assisted music, and has attracted many new musicians into the field.

Hookup

The **M**usical **I**nstrument **D**ata **I**nterface specification defines both the organization of the information transmitted and the circuitry used to connect systems together. The wiring is similar to that used for microphone cables, two wires within a shield. (The MIDI connector has five pins on it, but two of those are not connected. This is done for economy: the five pin DIN plugs, widely used overseas for connecting consumer stereo gear, costs about a third of what the three pin model does.) Exactly one input may be attached to each output; no multiples are allowed, but most devices have a "MIDI-THRU" output that simply passes the input data to the next device down the line. The basic configuration of equipment is a "daisy-chain", with one master device controlling a series of "slave" synthesizers.

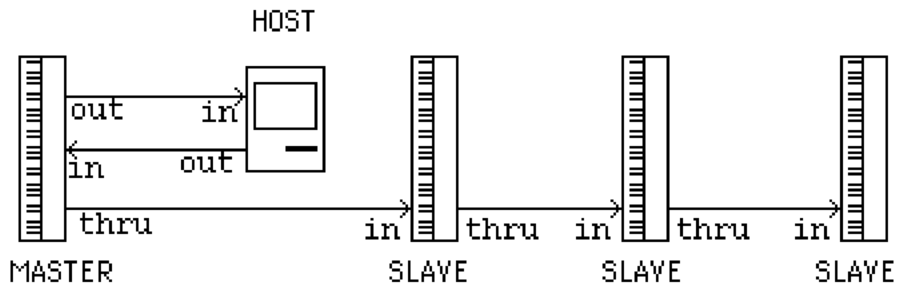


Fig. 2 Daisy chain MIDI setup

An alternative arrangement is sometimes used where the data from the controller goes to a splitter box that feeds the data to several outputs, each connected to one synthesizer in what is called the "star" configuration.

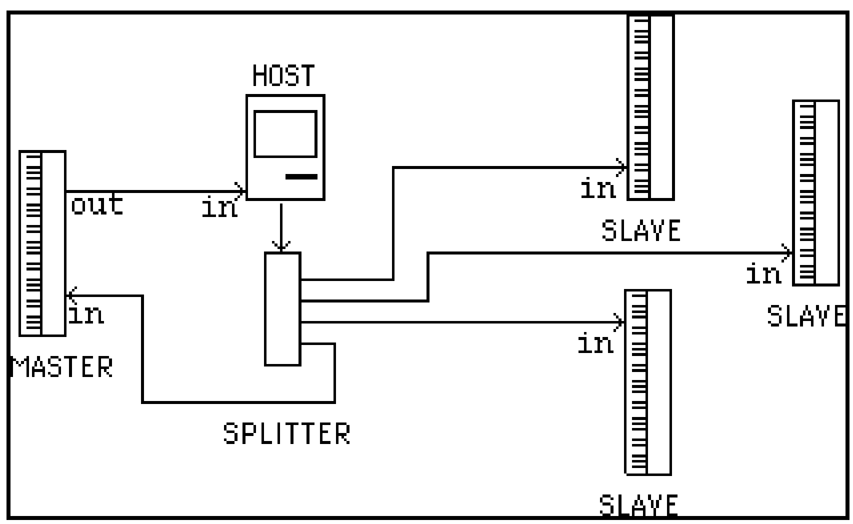


Fig. 3 Star MIDI setup

MIDI is a serial system. That means data is fed down a single wire pair one bit at a time. The bits are generated at the rate of 31,250 per second, but it takes ten bits to make a character and up to three characters to make a message, so it takes most of a millisecond to get anything said. As a rule, each action taken on the keyboard (such as releasing a key) generates a message. The typical message contains a channel number, a code for the key or other control affected, and descriptive data, such as key velocity. The channel number indicates which instruments are to respond to the data. There are sixteen channel numbers, and MIDI equipped synthesizers can be set to any of them.

Two streams of MIDI data cannot be mixed together in the simple manner two analog signals can. The group of bits that makes up a message must be kept intact or the meaning will be garbled. A device that combines

MIDI signals, called a "**merger**", has a microprocessor in it that can recognize messages, assign a priority to to them, knows how long the message should be, and prevents potential collisions by storing low priority messages until the output line is available. (This process is similar to switching freight trains onto a common track.)

It is surprisingly easy to generate a lot of MIDI data. For instance, many keyboards have "aftertouch"; a feature that measures how hard you press on a key as you hold it down and feeds that information into the data stream. If you hit a chord and wiggle your wrists, you might generate several thousand bytes of data. This data may be vital, or it may be useless, depending on exactly how other instruments in the MIDI chain are voiced. When the data stream gets too full, bizarre things begin to happen. Instruments slow down, or useful messages like "note-off" can get lost. For this reason, many instruments and programs have a "**filter**" feature which removes selected types of data. You can even buy a special purpose box to do this.

MIDI Processors

There are some other special tricks available in boxes. For instance there is a "**MIDI delay**" which simply stores data awhile before sending it along. If you connect an instrument's MIDI out to its own MIDI in through one of these, you get some complex echo effects. Another type of box is a "**mapper**" which can change data to compensate for differences in synthesizers. For instance, instruments often vary in the number of presets the can store. If you are using a fancy machine to control several simple ones, the fancy machine may implement all 128 preset locations, and the cheapies may only have 32. When you select preset 33 on the main synthesizer, it will send program change 33, which may have a bizarre result on the slave. The mapper can be set to change that program 33 to anything you desire. [These features are also available a part of better computer programs.]

A type of box that is very popular is the **MIDI patcher**. This device has a lot of inputs and outputs, say eight of each. Controls on the box electrically switch inputs to various outputs, so you don't have to fish around for the MIDI cables to change your system configuration. A particularly intriguing feature is that a configuration can be assigned a program number, so that the patch can be controlled over the MIDI line. Patchers usually cost about $300.00. MIDI plugs and jacks for making patch bays cost about $2.00 per connection. Which do you suppose you will find in the UCSC studio?

The most useful MIDI box is the **synchronizer.** With this device you can record a special tone (called a **sync stripe**) on one track of a tape deck. Once the stripe is in place, the box will then convert the data encoded in the tone into MIDI messages that can be followed by a sequencer or computer program. This allows a variety of procedures, such as building highly layered synthesizer compositions or MIDI controlled mixdowns.

Nuts And Bolts Of Midi

A MIDI message can consist of from one to several thousand bytes of data. The receiving instrument knows how many bytes to expect from the value of the first byte. This byte is known as the status byte, the others are data bytes. Status bytes always have the most significant bit equal to one, data bytes must have an msb of zero. If a status byte is received where a data byte is expected, the system assumes a transmission error has occurred. Because the msb is used to flag status bytes, data bytes are limited to numbers less than 128. This limits many things in the MIDI universe, such as the number of presets available. Bigger numbers may be transmitted of course, by using several bytes.

Status bytes inform the receiver as to what to do with incoming data. Many of the commands include the channel number (0-15) as the four least significant bits of the status byte.

The most common status is **note on.** [The actual bit values are: 1001nnnn, where nnnn gives the channel number.] Note on is followed by two data bytes, the first is the note number, the second is key velocity. If a keyboard is not equipped to sense velocity, it is supposed to send the value 64. Not too surprisingly, there is a status called **note off**, with the same data format. Note off is actually not used very much. Instead, MIDI allows for a shorthand, known as **running status**. Once a note on is received, an instrument interprets each pair of data bytes as instructions about a new note. If the velocity data is zero, the instrument performs a note off with velocity of 64.

This manner of thinking, requiring separate actions to start and stop a note, greatly simplifies the design of receiving instruments (the synthesizer does not have to keep time), but creates the potential for hung notes when a note off gets lost. The MIDI designers provided some features to compensate for this problem. There is a panic command, **all notes off**, which is generated by some computers and even some special effects boxes. There is also a special command called **active sensing**, which warns an instrument if there is a serious malfunction. Once the active sensing command has been received, the instrument expects something on the MIDI line at least every 300 milliseconds (If the controller has nothing to say, it sends more active sensing bytes.). If nothing is received, for instance because a cable has pulled out, the instrument shuts all notes off.

Commands are defined for about everything you would expect a synthesizer to do, to wit:

program change (also known as preset change)
aftertouch (for the entire keyboard, set by the heaviest push)
polyphonic aftertouch (values for each key down)
pitch bend (this can be a 14 bit value)
channel mode (1 to 4)

Channel modes take some explaining: mode 1 is OMNI ON, POLY. In this mode, the instrument accepts all note commands from all channels and plays them with the same voice to the best of its ability. Mode 2 is OMNI ON, MONO. In this mode, the synthesizer assigns one voice to each channel, as far as possible. A voice responds to each note event by changing pitch, with no overlap of sounds, much as an analog synthesizer would. Mode 3 is OMNI OFF, POLY; notes on one channel are played polyphonically. Mode 4 is OMNI OFF, MONO; only one note is played at a time.

There is a group of commands called **control changes**, that relate to actions of things like foot pedals, modulation wheels, and sliders. These are not very rigidly defined, so many systems allow assignment of controllers as part of preset definition. There is a sort of gentlemen's agreement that controller 7 will adjust volume.

Most of the preceding messages are **channel voice messages** which apply only to instruments set to the specified channel. Another group of messages are called **system messages**, and affect everything that is equipped to respond. You have already seen active sensing and all notes off. Some others are:

local on/off (disconnects the keyboard !)
reset
tune

These are in the spec, but seldom actually turn up on synthesizers. More useful are the synchronization commands:

song select

song pointer
start
stop
clock

With these commands, several sequencers or computers can be cued to a preset point in a composition and run together. The clock command is a single byte that is "broadcast" by a master sequencer at the rate of 24 per quarter note. Sequencers can follow this clock and stay in tempo. This clock can be recorded on tape and played back with a suitable adapter. If this recording happens to be on a multi-track tape deck, complex sequences can be built up using many passes with a single synthesizer. A song pointer goes one better, indicating the current measure and beat, allowing sequencers to cue up to the middle of pieces. Once cued, start and stop messages will control the operation of the sequencers.

MIDI time code is an even more sophisticated synchronization system. In this system, SMPTE time code, the standard syncronization signal for video and audio recorders is converted into a form that is compatable with MIDI. Sequencers and CMI programs can follow the time code to play scores with the pictures.

The final group of commands are the **system exclusive commands**. These are commands that the manufacturer may define as they like. (Each manufacturer is assigned an ID code to prevent confusion.) The data stream may be arbitrarily long, terminating with a command known as **EOX**. These messages are used for passing preset information, sequences, and even sound samples from one machine to another, and provide the foundation for the editor/librarian computer programs. Message are not limited to program data however, on the Yamaha instruments, system exclusive commands can be used to control everything but the power switch.

Problems

The MIDI protocol is often badmouthed because the original intentions of the designers are misunderstood. The system was created to allow a simple, cheap, and universal interconnection scheme for instrument controllers and synthesizers. The specification was developed by a committee made up of representatives from several companies, and contains many compromises between the needs and opinions of various people. The specification was inadvertently modified in translation to japanese, but since the company that made the mistake (Yamaha) has sold more synthesizers than all other companies combined, their implementation became the standard. The MIDI committee is still active, and adds features to the specification from time to time.

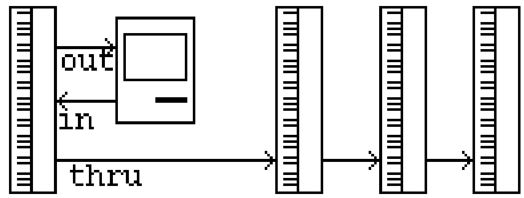
The complaint heard most often about MIDI is that it is too slow. It takes one millisecond (1/1000 sec) to send the command that starts a note. This is musically imperceptible (in normal notation, MM=60,000) in simple pieces, but the delay across a twenty note chord can be noticed by a keen ear. The actual effect of this problem on the music is arguable (very few bands are together within twenty milliseconds). Probably the worst case for a performer is when the delay is unpredictably varied. Elaborate computer controlled performances are the activities that generate the most frustration. The series connection MIDI system can clog up quickly when detailed control of a lot of instruments is attempted. The cure for this is to use a parallel connection scheme where the computer itself has several MIDI outputs.

The other common complaint is that MIDI sends the wrong information. It is clear that the standard was written with keyboard controllers in mind, and that is sensible, since the organ type keyboard is the most common controller for polyphonic single performer instruments. It is quite difficult but not impossible to design controllers with a continuous effect, such as a wind or bowed string instrument has, but the speed problem becomes extreme in such cases.

The channelization scheme chosen causes a lot of confusion, but is not a problem. The channel numbers are really a tag on each command, and instruments have the option of ignoring commands that are not tagged a certain way. Difficulties arise when sending devices and receiving devices are not set to the same channel.

Overcoming these problems is a challenge, but one similar to challenges musicians are already familiar with. Here are a few guidelines to maintain sanity.

> **Use a simple configuration, and stay with it.** The MIDI system is designed to have one master controller running a bunch of slaves. Mergers allow the use of two or more controllers, and switchers allow quick reconfiguration of the system, but there is usually little to be gained. Any good computer program will allow you to play along with running tracks.



Don't overload the system. Always filter out unnecessary information. Aftertouch, for instance should never be sent unless some device is responding to it. If you are playing with a sequenced track, the pedals are probably of interest only to the synthesizer you are playing.

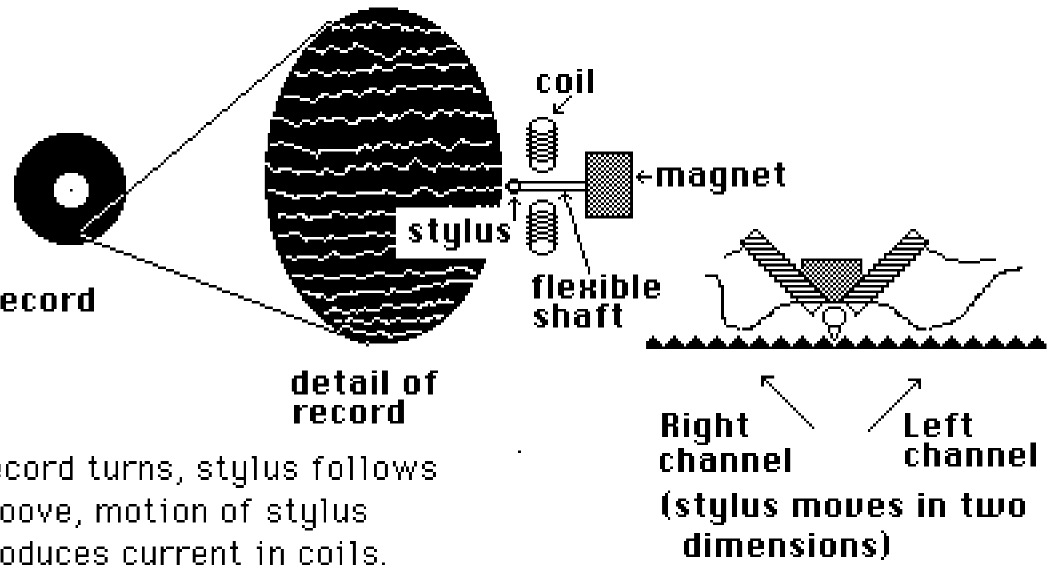
Know the difference between OUT and THRU. OUT is information generated by the instrument. THRU is a copy of the input data. Neither one is a mix of the input and the locally generated data.

Take care of your cables. The MIDI connector is not noted for ruggedness and reliability. It is possible for a plug to look like it is in, but be loose enough to stop the data (this is common when short cables are stretched between devices).

Read the manual. Read the Manual. READ THE MANUAL. Especially the part in the back that shows which MIDI features actually work. Pay particular attention to how to set the channel number and how to turn OMNI mode off.

Analog Recording Of Sound

Records



Record turns, stylus follows groove, motion of stylus produces current in coils.

The basic principle of disk recording is very simple. Displacement of the microphone diaphragm is transformed into a wiggley groove on a moving piece of vinyl. A stylus tracing the wiggles exactly reproduces the motion of the diaphragm at the time the recording was made. Electricity is really incidental to the process, used as a convenient way to connect the microphone to the cutter and the pickup to the speaker.

Most of the development in record technology has been devoted to putting a lot of music on a single record. The obvious approach, slow speed and a narrow groove, reached a practical limit in the middle of the century with the 33 1/3 rpm microgroove record. At that speed, (9 inches per second in the inner part of the groove) a 20 khz signal has a wavelength of .0004 inch. It is very difficult to manufacture a stylus that would handle wavelengths smaller than that.

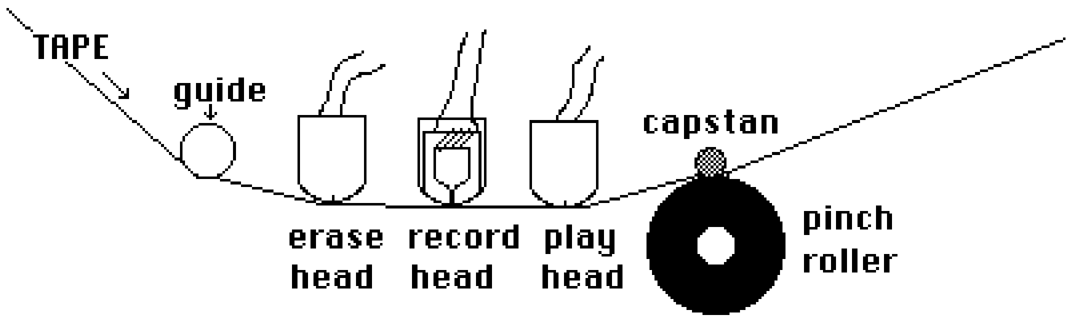
The major consumer of real estate on the record is low frequency content. This seems strange until you remember that the amplitude of the electrical signal produced is proportional to the side to side velocity of the stylus. Given equal velocities, a low frequency wiggle will swing wider than one of high frequency because at low frequencies the cutter will not turn around as often as it does at high frequency. To counteract this effect, the low frequency content of the record is deliberately reduced, and this low end rolloff has to be compensated by a bass boost in the playback system.

The high frequency content is given a treatment opposite to that of the lows. High frequency information is emphasized during recording, and reduced during playback. This is an attempt to reduce the noise generated by the roughness of the vinyl. That noise is white noise, and as we saw in an earlier essay, sounds like a high frequency phenomenon. When the playback system reduces the high frequency content to its proper level, the noise in that range is reduced by the same amount.

The combination of bass roll-off and treble boost is called the recording characteristic, and the complementary response of the playback system is called RIAA equalization after the manufacturer's association which standardized this feature in 1956.

The LP is an endangered species with the advent of Compact Disk technology, but it will not disappear overnight. Even if no new records are produced, there are hundreds of millions in existence, including many unique performances and compositions.

Analog Tape

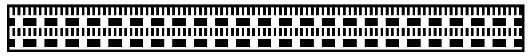


Tape Deck Mechanism

The principle of tape recording is just as simple as that of disk recording. The tape is a strip of plastic which has been coated with a material that is easily magnetized. (The most commonly used material is highly refined rust, or iron oxide.) The capstan is a spinning post. The tape is held tightly against the capstan by the pinch roller and dragged across the three heads at a steady rate.

All three heads are essentially the same in construction: a C-shaped piece of metal with the very narrow gap of the "C" near the tape. A coil of wire around the metal can serve to either detect or produce magnetic fields at the gap. If a strong current is passed through the coil, a field is produced which creates a magnetic spot on the nearby tape. The amount of magnetism will be proportional to the amount of current. If the tape is moved and the current varied in a periodic way, a "track" of magnetic areas will be imprinted on the tape. All of this happens at the record head. When the tape subsequently passes the play head, the varying magnetic field on the tape produces a varying current in the play head coil, which can be detected by some sensitive electronic circuitry. (The erase head works just like the record head, but at a super high frequency which will not be recorded but which will obliterate any existing information.)

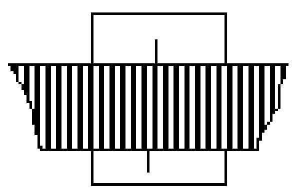
The signal applied to the record heads is equalized in a manner similar to LPs, for essentially the same reasons. The equalization is varied according to speed and type of tape. That is automatic on most decks, but must be set manually on some cassette decks.



Magnetic patterns on the tape
(top channel is high frequency)

This drawing shows how the magnetic fields are oriented on a stereo tape. You can see four tracks, two of which are played when the tape is going one direction, the other two when the tape is reversed. (Multi-track tape decks use all four tracks or more at one pass. There are also formats which use only one or two tracks, recorded on the entire width of the tape.) The width of a single track is an important factor in the strength of signal that can be recorded, which ultimately limits the noise of the system. The width of a track depends on the width of the tape as well as the number of tracks; various sizes ranging from 1/8 inch to 2 inches are used, recording up to 24 tracks.

The distance between spots in a single track is the wavelength of the signal, which depends on the frequency of the signal and the speed of the tape. The higher the frequency of the signal, the shorter the wavelength (a familiar formula). There is a limit to how small the shortest spot of magnetism can be; namely the width of the gap in the recording head. The practical result of this limitation is that the highest frequency that can be recorded is limited by the speed of the tape. Again, various standards are in use, from $1\frac{7}{8}$ to 30 inches per second.

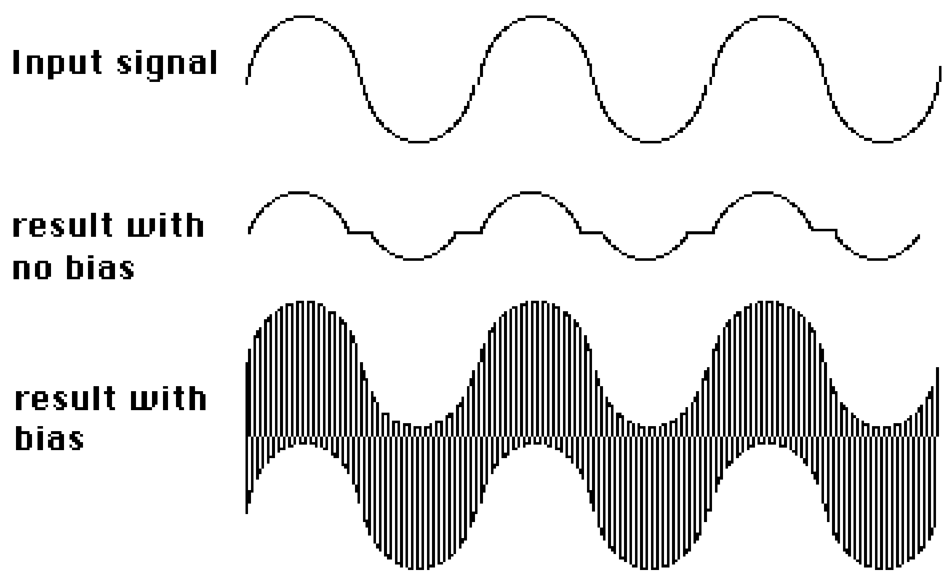


azimuth error

This drawing shows another feature of the magnetic spots. At high frequencies, the magnetic domain is a rather slender, tall shape, almost a line. If that line is not exactly the same angle as the gap in the play head (which is supposed to be perpendicular to the tape), the energy represented by the magnetism will not be accurately detected, resulting in a reduced output signal. Low frequency signals are not affected by this factor (known as azimuth adjustment), so the net result is a loss of highs. In a cassette system, this alignment is a function of the plastic shell, and is generally rather sloppy in all but the most expensive tapes.

AC Bias

Tape recording would be a very low fidelity business without AC bias. The process of magnetization is linear only when applied to fields of medium strength. There is a limit to the strength of field that the tape can accept; once the tape is completely magnetized, no amount of extra current in the head will increase the resultant field. That condition is called saturation. As the strength approaches saturation, there is a gradual falloff in the effectiveness of the magnetization process. This results in a phenomenon called "soft clipping", which is definitely distortion, but as distortions go is reasonably unobjectionable. Incidentally, since the record equalization increases the high frequency content of the signal, this clipping will happen to the highs first: that is why a cassette seems to lose its top end response when it is recorded "hot".



The effect of Bias

Clipping is something we live with in all electronic systems, and is easy to avoid; simply keep the gain down. Another region of nonlinearity is more difficult to deal with. The magnetization process produces field regions that alternate in polarity: one north, one south, north again, and so forth. In between there are regions where the field strength is zero. When the oxide is not magnetized, a fair amount of current is required to produce any magnetization at all; this leaves a flat spot in the middle of the waveform, as illustrated in the second waveform in the diagram above.

We avoid the effects of this nonlinear region by adding a very high frequency (over 100 khz) bias* signal to the signal we are trying to record. The result is the third waveform. The center of the bias frequency is distorted, but the original signal, which is the shape of the overall waveform, is clean. The playback head cannot respond to the bias signal, and simply returns the original. The amplitude of the bias signal has to be carefully adjusted to provide a distortion free recording. Many tape decks (especially cassettes) offer a switch to make a coarse change in bias for different tape types, but a finer calibration is really required for optimum results.

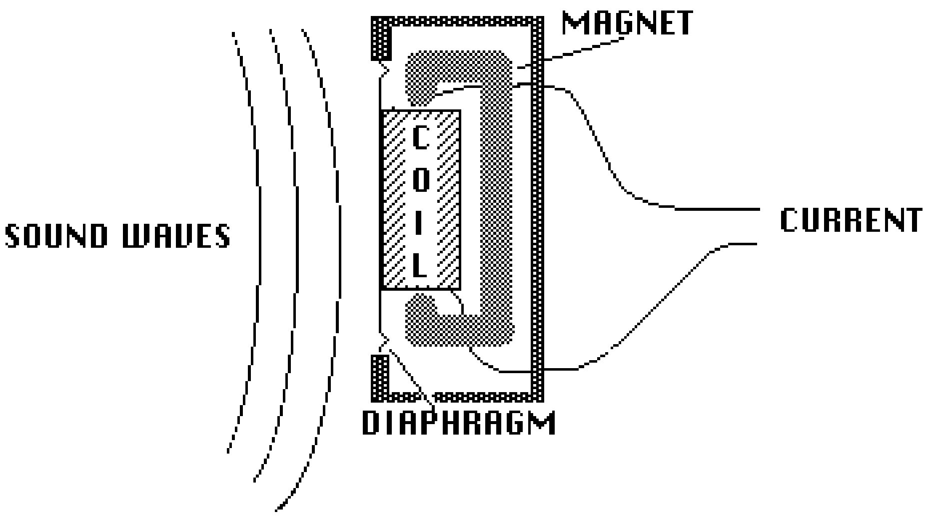
*In electronics, the word bias is used to indicate the circuitry or process that changes the midpoint of waveforms.

A Primer On Microphones

I. How they work.

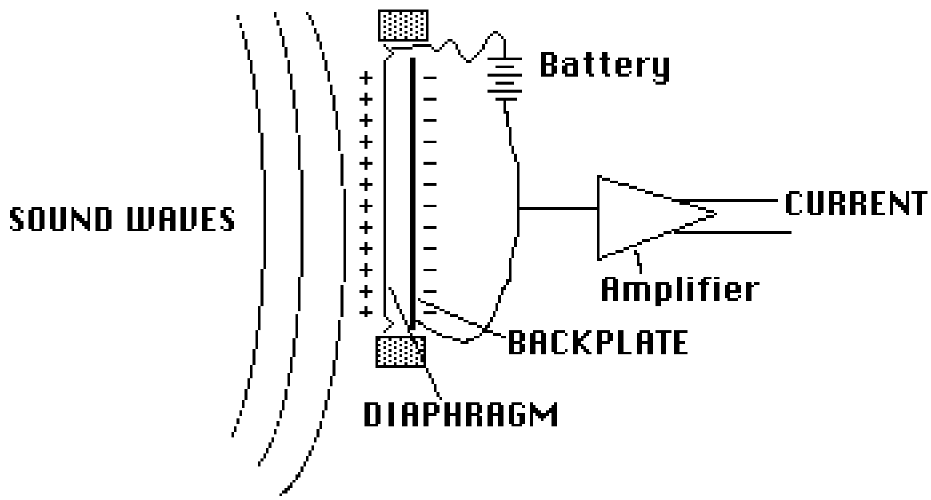
A microphone is an example of a transducer, a device that changes information from one form to another. Sound information exists as patterns of air pressure; the microphone changes this information into patterns of electric current. The recording engineer is interested in the accuracy of this transformation, a concept he thinks of as fidelity.

A variety of mechanical techniques can be used in building microphones. The two most commonly encountered in recording studios are the magneto–dynamic and the variable condenser designs.



The Dynamic Microphone.

In the magneto–dynamic, commonly called dynamic, microphone, sound waves cause movement of a thin metallic diaphragm and an attached coil of wire. A magnet produces a magnetic field which surrounds the coil, and motion of the coil within this field causes current to flow. The principles are the same as those that produce electricity at the utility company, realized in a pocket-sized scale. It is important to remember that current is produced by the <u>motion</u> of the diaphragm, and that the amount of current is determined by the speed of that motion. This kind of microphone is known as **velocity sensitive.**



The Condenser Microphone.

In a condenser microphone, the diaphragm is mounted close to, but not touching, a rigid backplate. (The plate may or may not have holes in it.) A battery is connected to both pieces of metal, which produces an electrical potential, or charge, between them.* The amount of charge is determined by the voltage of the battery, the area of the diaphragm and backplate, and the distance between the two. This distance changes as the diaphragm moves in response to sound. When the distance changes, current flows in the wire as the battery maintains the correct charge. The amount of current is essentially proportional to the **displacement** of the diaphragm, and is so small that it must be electrically amplified before it leaves the microphone.

A common variant of this design uses a material with a permanently imprinted charge for the diaphragm. Such a material is called an **electret** and is usually a kind of plastic. (You often get a piece of plastic with a permanent charge on it when you unwrap a record. Most plastics conduct electricity when they are hot but are insulators when they cool.) Plastic is a pretty good material for making diaphragms since it can be dependably produced to fairly exact specifications. (Some popular dynamic microphones use plastic diaphragms.) The major

---

*This is called a potential because electricity could flow between the two if they were shorted together, even after the battery were removed. This charge is really an excess of electrons on the backplate and an equivalent lack of electrons on the diaphragm.

disadvantage of electrets is that they lose their charge after a few years and cease to work.

II. How Well They Work.

There is no inherent advantage in fidelity of one type of microphone over another. Condenser types require batteries or power from the mixing console to operate, which is occasionally a hassle, and dynamics require shielding from stray magnetic fields, which makes them a bit heavy sometimes, but very fine microphones are available of both styles. The most important factor in choosing a microphone is how it sounds in the required application. The following issues must be considered:

Sensitivity. This is a measure of how much electrical output is produced by a given sound. This is a vital specification if you are trying to record very tiny sounds, such as a turtle snapping its jaw, but should be considered in any situation. If you put an insensitive mic on a quiet instrument, such as an acoustic guitar, you will have to increase the gain of the mixing console, adding noise to the mix. On the other hand, a very sensitive mic on vocals might overload the input electronics of the mixer or tape deck, producing distortion.

Overload characteristics. Any microphone will produce distortion when it is overdriven by loud sounds. This is caused by various factors. With a dynamic, the coil may be pulled out of the magnetic field; in a condenser, the internal amplifier might clip. Sustained overdriving or extremely loud sounds can permanently distort the diaphragm, degrading performance at ordinary sound levels. Loud sounds are encountered more often than you might think, especially if you place the mic very close to instruments. (Would you put your ear in the bell of a trumpet?) You usually get a choice between high sensitivity and high overload points, although occasionally there is a switch on the microphone for different situations.

Linearity, or distortion. This is the feature that runs up the price of microphones. The distortion characteristics of a mic are determined mostly by the care with which the diaphragm is made and mounted. High volume production methods can turn out an adequate microphone, but the distortion performance will be a matter of luck. Many manufacturers have several model numbers for what is essentially the same device. They build a batch, and then test the mics and charge a premium price for the good ones. The really big names throw away mic capsules that don't meet their standards. (If you buy one Neumann mic, you are paying for five!)

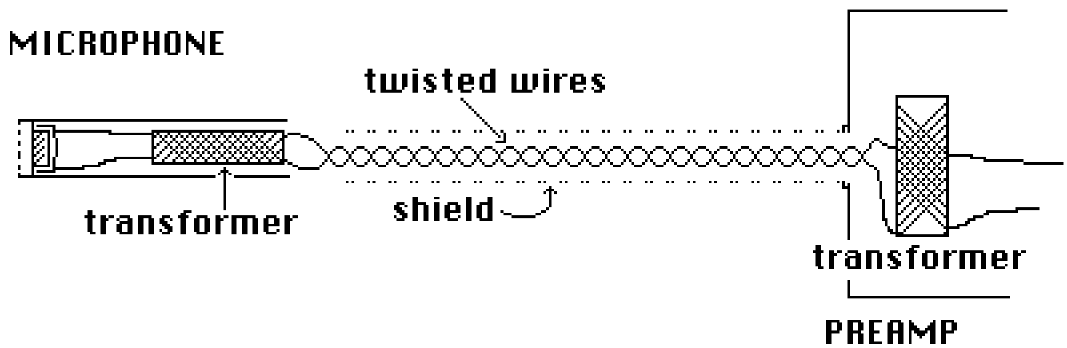
No mic is perfectly linear; the best you can do is find one with distortion that complements the sound you are trying to record. This is one of the factors of the microphone mystique discussed later.

Frequency response. A flat frequency response has been the main goal of microphone companies for the last three or four decades. In the fifties, mics were so bad that console manufacturers began adding equalizers to each input to compensate. This effort has now paid off to the point were most professional microphones are respectably flat, at least for sounds originating in front. The major exceptions are mics with deliberate emphasis at certain frequencies that are useful for some applications. This is another part of the microphone mystique. Problems in frequency response are mostly encountered with sounds originating behind the mic, as discussed in the next section.

Noise. Any microphone produces a very small amount of current, which makes sense when you consider just how light the moving parts must be to accurately follow sound waves. To be useful for recording or other electronic processes, the signal must be amplified by a factor of over a thousand. Any electrical noise produced by the microphone will also be amplified, so even slight amounts are intolerable. Dynamic microphones are essentially noise free, but the electronic circuit built into condenser types is a potential source of trouble, and must be carefully designed and constructed of premium parts.

Noise also includes unwanted pickup of mechanical vibration through the body of the microphone. Very sensitive designs require elastic shock mountings, and mics intended to be held in the hand need to have such mountings built inside the shell.

The most common source of noise associated with microphones is the wire connecting the mic to the console or tape deck. A mic preamp is very similar to a radio receiver, so the cable must be prevented from becoming an antenna. The basic technique is to surround the wires that carry the current to and from the mic with a flexible metallic shield, which deflects most radio energy. A second technique, which is more effective for the low frequency hum induced by the power company into our environment, is to balance the line:



Balanced Microphone Connection

Current produced by the microphone will flow down one wire of the twisted pair, and back along the other one. Any current induced in the cable from an outside source would tend to flow the same way in both wires, and such currents cancel each other in the transformers. This system is expensive.

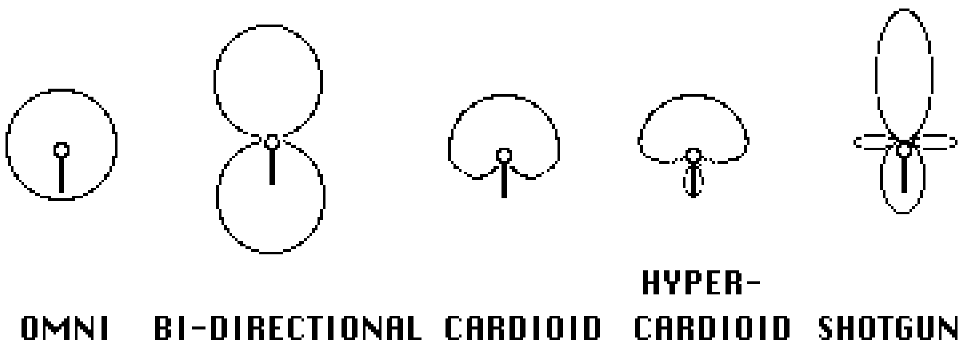
Microphone levels are of necessity very weak signals, generally around –60dBm. (The specification is the power produced by a sound pressure of 10 μBar) The output impedance will depend on whether the mic has a transformer balanced output . If it does not, the microphone will be labeled "high impedance" or "hi Z" and must be connected to an appropriate input. The cable used must be kept short, less than 10 feet or so, to avoid noise problems.

If a microphone has a transformer, it will be labeled low impedance, and will work best with a balanced input mic preamp. The cable can be several hundred feet long with no problem. Balanced output, low impedance microphones are expensive, and generally found in professional applications. Balanced outputs must have three pin connectors ("Cannon plugs"), but not all mics with those plugs are really balanced. Microphones with standard or miniature phone plugs are high impedance. A balanced mic can be used with a high impedance input with a suitable adapter.

You can see from the diagram that there is a transformer at the input of the console preamp. This is the most significant difference between professional preamplifiers and the type usually found on home tape decks. You can buy transformers that are designed to add this feature to a consumer deck for about $20 each. (Make sure you are getting a transformer and not just an adapter for the connectors.) With these accessories you can use professional quality microphones, run cables over a hundred feet with no hum, and because the transformers boost the signal somewhat, make recordings with less noise. This will not work with a few inexpensive cassette recorders, because the strong signal causes distortion. Such a deck will have other problems, so there is little point trying to make a high fidelity recording with it anyway.

III. Which Way They Work

Many people have the misconception that microphones only pick up sound from sources they are pointed at, much as a camera only photographs what is in front of the lens. This would be a nice feature if we could get it, but the truth is we can only approximate that action, and at the expense of other desirable qualities.



Microphone Patterns

These are polar graphs of the output produced vs. the angle of the sound source. The output is represented by the radius of the curve at the incident angle.

Omni

The simplest mic design will pick up all sound, regardless of its point of origin, and is thus known as an omnidirectional microphone. They are very easy to use and generally have good to outstanding frequency response.

Bi-directional

It is not very difficult to produce a pickup pattern that accepts sound striking the front or rear of the diaphragm, but does not respond to sound from the sides. This is the way any diaphragm will behave if sound can strike the front and back equally. The rejection of undesired sound is the best achievable with any design, but the fact that the mic accepts sound from both ends makes it difficult to use in many situations. Most often it is placed above an instrument. Frequency response is just as good as an omni, at least for sounds that are not too close to the microphone.

Cardioid

This pattern is popular for sound reinforcement or recording concerts where audience noise is a possible problem. The concept is great, a mic that picks up sounds it is pointed at. The reality is different. The first problem is that sounds from the back are not completely rejected, but merely reduced about 10-30 dB. This can surprise careless users. The second problem, and a severe one, is that the actual shape of the pickup pattern varies with frequency. For low frequencies, this is an omnidirectional microphone. A mic that is directional in the range of bass instruments will be fairly large and expensive. Furthermore, the frequency response for signals arriving from the back and sides will be uneven; this adds an undesired coloration to instruments at the edge of a large ensemble, or to the reverberation of the concert hall.

A third effect, which may be a problem or may be a desired feature, is that the microphone will emphasize the low frequency components of any source that is very close to the diaphragm. This is known as the "proximity effect", and many singers and radio announcers rely on it to add "chest" to a basically light voice. Close, in this context, is related to the wavelength of the sound and the size of the microphone diaphragm, so the nice large mics with even back and side frequency response exhibit the strongest presence effect. Most cardioid mics have a built in lowcut filter switch to compensate for proximity. Miss-setting that switch can cause hilarious results. Bidirectional mics also exhibit this phenomenon.

Tighter Patterns

It is posible to exaggerate the directionality of cardioid type microphones, if you don't mind exaggerating some of the problems. The Hypercardioid pattern is very popular, as it gives a better overall rejection and flatter frequency response at the cost of a small back pickup lobe. This is often seen as a good compromise between the cardioid and bidirectional patterns. A "shotgun" mic carries these techniques to extremes by mounting the diaphragm in the middle of a pipe. The shotgun is extremely sensitive along the main axis, but possesses pronounced extra lobes which vary drastically with frequency. In fact, the frequency response of this mic is so bad it is usually electronically restricted to the voice range, where it is used to record dialogue for film and video.

Stereo microphones

You don't need a special microphone to record in stereo, you just need two (see below). A so called stereo microphone is really two microphones in the same case. There are two kinds: extremely expensive professional models with precision matched capsules, adjustable capsule angles, and remote switching of pickup patterns; and very cheap units (often with the capsules oriented at 180°)

that can be sold for high prices because they have the word stereo written on them.

IV. How To Use Them

Single Microphone Use

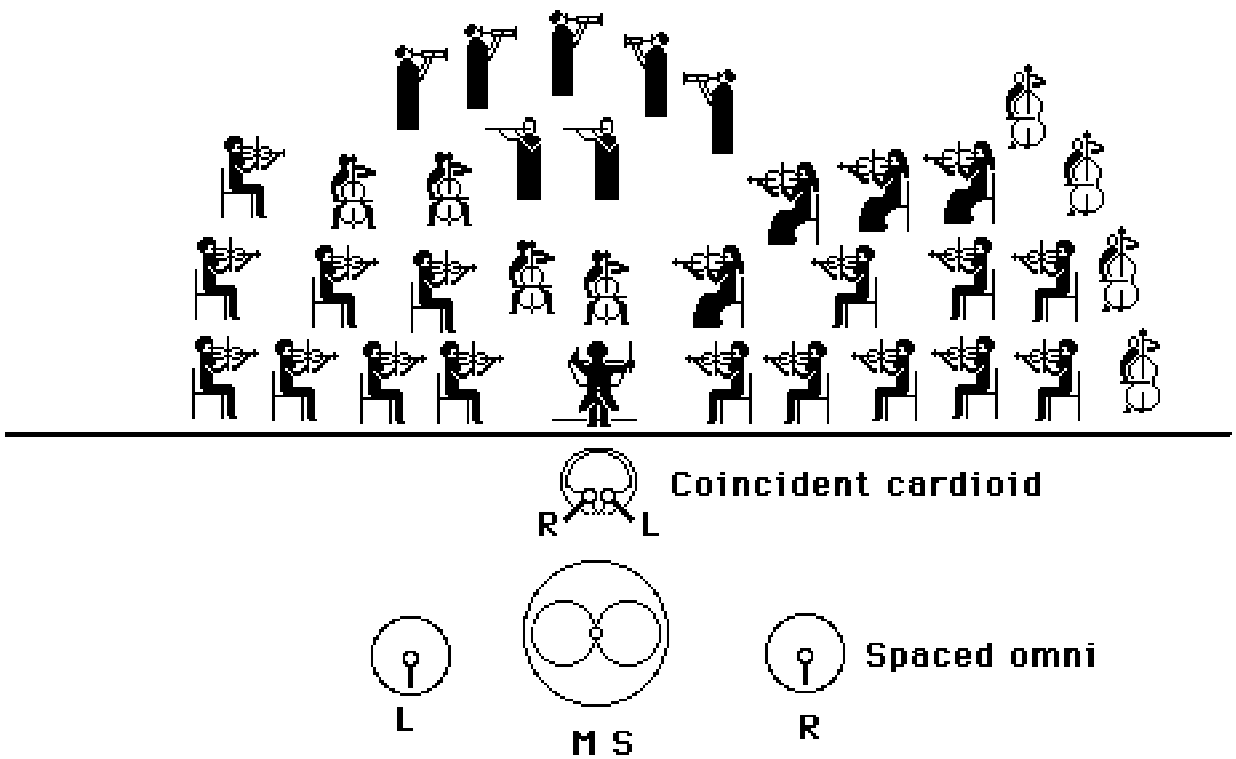
Use of a single microphone is pretty straightforward. Having chosen one with appropriate sensitivity and pattern, (and the best distortion, frequency response, and noise characteristics you can afford), you simply mount it where the sounds are. The practical range of distance between the instrument and the microphone is determined by the point where the sound overloads the microphone or console at the near end, and the point where ambient noise becomes objectionable at the far end. Between those extremes it is largely a matter of taste and experimentation.

If you place the microphone close to the instrument, and listen to the results, you will find the location of the mic affects the way the instrument sounds on the recording. The timbre may be odd, or some notes may be louder than others. That is because the various components of an instrument's sound often come from different parts of the instrument body (the highest note of a piano is nearly five feet from the lowest), and we are used to hearing an evenly blended tone. A close in microphone will respond to some locations on the instrument more than others because the difference in distance from each to the mic is proportionally large. A good rule of thumb is that the blend zone starts at a distance of about twice the length of the instrument. If you are recording several instruments, the distance between the players must be treated the same way.

If you place the microphone far away from the instrument, it will sound as if it is far away from the instrument. We judge sonic distance by the ratio of the strength of the direct sound from the instrument (which is always heard first) to the strength of the reverberation from the walls of the room. When we are physically present at a concert, we use many cues beside the sounds to keep our attention focused on the performance, and we are able to ignore any distractions there may be. When we listen to a recording, we don't have those visual clues to what is happening, and find anything extraneous that is very audible annoying. For this reason, the best seat in the house is not a good place to record a concert. On the other hand, we do need some reverberation to appreciate certain features of the music. (That is why some types of music sound best in a stone church) Close microphone placement prevents this. Some engineers prefer to use close miking techniques to keep noise down and add artificial reverberation to the recording, others solve the problem by mounting the mic very high, away from audience noise but where adequate reverberation can be found.

Stereo

Stereo sound is an illusion of spaciousness produced by playing a recording back through two speakers. The success of this illusion is referred to as the image. A good image is one in which each instrument is a natural size, has a distinct location within the sound space, and does not move around. The main factors that establish the image are the relative strength of an instrument's sound in each speaker, and the timing of arrival of the sounds at the listener's ear. In a studio recording, the stereo image is produced artificially. Each instrument has its own microphone, and the various signals are balanced in the console as the producer desires. In a concert recording, where the point is to document reality, and where individual microphones would be awkward at best, it is most common to use two mics, one for each speaker.



Microphone placement for stereo recording.

Spaced Microphones

The simplest approach is to assume that the speakers will be eight to ten feet apart, and place two microphones eight to ten feet apart to match. Either omnis or cardioids will work. When played back, the results will be satisfactory with most speaker arrangements. (I often laugh when I attend concerts and watch people using this setup fuss endlessly with the precise placement of the mics. This technique is so forgiving that none of their efforts will make any practical difference.) The big disavantage of this technique is that the mics must be rather far back from the ensemble- at least as far as the distance from the leftmost performer to the rightmost. Otherwise, those instruments closest to the microphones will be too prominent. There is usually not enough room between stage and audience to achieve this with a large ensemble, unless you can suspend the mics or have two very tall stands.

Coincident Cardioids
There is another disadvantage to the spaced technique that appears if the two channels are ever mixed together into a monophonic signal. (Or broadcast over the radio, for similar reasons.) Because there is a large distance between the mics, it is quite possible that sound from a particular instrument would reach each mic at slightly different times. (Sound takes 1 millisecond to travel a foot.) This effect creates phase differences between the two channels, which results in severe frequency response problems when the signals are combined. You seldom actually lose notes from this interference, but the result is an uneven, almost shimmery sound. The various coincident techniques avoid this problem by mounting both mics in almost the same spot.

This is most often done with two cardioid microphones, one pointing slightly left, one slightly right. The microphones are often pointing toward each other, as this places the diaphragms within a couple of inches of each other, totally eliminating phase problems. No matter how they are mounted, the microphone that points to the left provides the left channel. The stereo effect comes from the fact that the instruments on the right side are on–axis for the right channel microphone and somewhat off–axis (and therefore reduced in level) for the other one. The angle between the microphones is critical, depending on the actual pickup pattern of the microphone. If the mics are too parallel, there will be little stereo effect. If the angle is too wide, instruments in the middle of the stage will sound weak, producing a hole in the middle of the image. [Incidentally, to use this technique, you must know which way the capsule actually points. There are some very fine German cardioid microphones in which the diaphragm is mounted so that the pickup is from the side, even though the case is shaped just like many popular end addressed models. (The front of the mic in question is marked by the trademark medallion.) I have heard the results where an engineer mounted a pair of these as if the axis were at the end. You could hear one cello player and the tympani, but not much else.]

You may place the microphones fairly close to the instruments when you use this technique. The problem of balance between near and far instruments is solved by aiming the mics toward the back row of the ensemble; the front instruments are therefore off axis and record at a lower level.* You will notice that the height of the microphones becomes a critical adjustment.

M.S.

The most elegant approach to coincident miking is the M.S. or middle-side technique. This is usually done with a stereo microphone in which one element is omnidirectional, and the other bidirectional. The bidirectional element is oriented with the axis running parallel to the stage, rejecting sound from the center. The omni element, of course, picks up everything. To understand the next part, consider what happens as instrument is moved on the stage. If the instrument is on the left half of the stage, a sound would first move the diaphragm of the bidirectional mic to the right, causing a positive voltage at the output. If the instrument is moved to center stage, the microphone will not produce any signal at all. If the instrument is moved to the right side, the sound would first move the diaphragm to the left, producing a negative volage. You can then say that instruments on one side of the stage are 180° out of phase with those on the other side, and the closer they are to the center, the weaker the signal produced.

Now the signals from the two microphones are not merely kept in two channels and played back over individual speakers. The signals are combined in a circuit that has two outputs; for the left channel output, the bidirectional output is added to the omni signal. For the right channel output, the bidirectional output is subtracted from the omni signal. This gives stereo, because an instrument on the right produces a negative signal in the bidirectional mic, which when added to the omni signal, tends to remove that instrument, but when subtracted, increases the strength of the instrument. An instrument on the left suffers the opposite fate, but instruments in the center are not affected, because their sound does not turn up in the bidirectional signal at all.

*With an orchestra, you can zero in on the woodwinds, often achieving a better balance on the recording than exists in the hall, were brass and strings often dominate.

M.S. produces a very smooth and accurate image, and is entirely mono compatible. The only reason it is not used more extensively is the cost of the special microphone and decoding circuit, about $1,000.

Large Ensembles

The above techniques work well for concert recordings in good halls with small ensembles. When recording large groups in difficult places, you will often see a combination of spaced and coincident pairs. This does produce a kind of chorusing when the signals are mixed, but it is an attractive effect and not very different from the sound of string or choral ensembles any way. When balance between large sections and soloists cannot be achieved with the basic setup, extra microphones are added to highlight the weaker instruments. A very common problem with large halls is that the reverberation from the back seems late when compared to the direct sound taken at the edge of the stage. This can be helped by placing a mic at the rear of the audience area to get the ambient sound into the recording sooner.

Studio Techniques

A complete description of all of the procedures and tricks encountered in the recording studio would fill several books. These are just a few things you might see if you dropped in on the middle of a session.

INDIVIDUAL MICS ON EACH INSTRUMENT. This provides the engineer with the ability to adjust the balance of the instruments at the console, or, with a multitrack recorder, after the musicians have gone home. There may be eight or nine mics on the drum set alone.

CLOSE MIC PLACEMENT. The microphones will usually be placed rather close to the instruments. This is partially to avoid problems that occur when an instrument is picked up in two non-coincident mics, and partially to modify the sound of the instruments (to get a "honky–tonk" effect from a grand piano, for instance).

ACOUSTIC FENCES AROUND INSTRUMENTS, OR INSTRUMENTS IN SEPARATE ROOMS. The interference that occurs when an instrument is picked up by two mics that are mixed is a very serious problem. You will often see extreme measures, such as a bass drum stuffed with blankets to muffle the sound, and then electronically processed to make it sound like a drum again.

EVERYONE WEARING HEADPHONES. Studio musicians often play to "click tracks", which are not recorded metronomes, but some one tapping the beat with sticks and occasionally counting through tempo changes. This is done when the music must be synchronized to a film or video, but is often required because the performer cannot hear the other musicians because of the isolation measures described above.

20 OR 30 TAKES ON ONE SONG. Recordings require a level of perfection in intonation and rhythm that is much higher than that acceptable in concert. The finished product is usually a composite of several takes.

"POP FILTERS" IN FRONT OF MICS. Vocalists like to move around when they sing; in particular, they will lean into microphones. If the singer is very close to the mic, any motion will produce drastic changes in level and sound quality. (You have seen this with inexpert entertainers using hand held mics.) To keep the artist the proper distance, the engineer will often mount a small nylon screen in front of the microphone. The performer may move slightly in relation to the screen, but that is a small proportion of the distance to the microphone. (Singers are often not told exactly why this screen is used.)

V. The Microphone Mystique

There is an aura of mystery about microphones. To the general public, a recording engineer is something like a priest, privy to a secret arcana, and capable of supernatural feats. A few modern day engineers encourage this attitude, but it is mostly a holdover from the days when studio microphones were expensive and fragile, and most people never dealt with any electronics more complex than a table radio. There are no secrets to recording; the art is mostly a commonsense application of the principles already discussed in this paper. If there is an arcana, it is an accumulation of trivia achieved through experience with the following problems:

Matching the microphone to the instrument. There is no wrong microphone for any instrument. Every engineer has preferences, usually based on mics with which he is familiar. Each mic has a unique sound, but the differences between good examples of any one type are pretty minor. The artist has a conception of the sound of his instrument, (which may not be accurate) and wants to hear that sound through the speakers. Frequency response and placement of the microphone will affect that sound; sometimes you need to exaggerate the features of the sound the client is looking for.

Listening the proper way. It is easy to forget that the recording engineer is an illusionist- the result will never be confused with reality by the listener. Listeners are in fact very forgiving about some things. It is important that the engineer be able to focus his attention on the main issues and not waste time with interesting but minor technicalities. It is important that the engineer know what the main issues are. An example is the noise / distortion tradeoff. Most listeners are willing to ignore a small amount of distortion on loud passages (in fact, they expect it), but would be annoyed by the extra noise that would result if the engineer turned the recording level down to avoid it. One technique for encouraging this attention is to listen to recordings over a variety of sound systems, good and bad.

Learning for yourself. Many students come to me asking for a book or a course of study that will easily make them a member of this elite company. There are books, and some schools have courses in recording, but they do not supply the essential quality the professional recording engineer needs, which is experience.

A good engineer will have made hundreds of recordings using dozens of different microphones. Each session is an opportunity to make a new discovery. The engineer will make careful notes of the setup, and will listen to the results many times to build an association between the technique used and the sound achieved. Most of us do not have access to lots of professional microphones, but we could probably afford a pair of general purpose cardioids. With about $400 worth of mics, and a reliable, properly adjusted tape deck, it is possible to learn to make excellent recordings. The trick is to record everything that will sit still and make noise, and study the results: learn to hear when the mic is placed badly and what to do about it. When you know all you can about your mics, buy a different pair and learn those. Occasionally, you will get the opportunity to borrow mics. If possible, set them up right alongside yours and make two recordings at once. It will not be long before you will know how to make consistently excellent recordings under most conditions.

Notes on Fourier Transforms

The Fourier transform is something we all toss around like we understand it, but it is often discussed in an offhand way that leads to confusion for those just learning their way around DSP. I'm not ready to write a comprehensive manual on the thing (Smith devotes 14 chapters to it), but here are some assorted bits of trivia that may clear the air some.

Buzzwords, or fftspeak:

Time domain representation means a graph with time along the bottom. Waveforms are usually represented this way.

Frequency domain representation means a graph with frequency along the bottom. Spectral plots are like this.

Polar representation means graphing in terms of an angle and radius. Since sine waves are inherently angular this can be useful. To map the frequency domain onto a polar plot, the angle π radians represents 1/2 the sampling rate. The region on the bottom of the circle represents negative frequency plotted from 0 to $-\pi$. Points on a polar plot can also be indicated by their rectangular (Cartesian) coordinates.

The **unit** means one. The unit circle on a polar plot is a circle of radius 1. Setting things to equal 1 often makes math clearer.

The letter **j** is the imaginary square root of -1. Some mathematicians use i for this, but i means something else in electronics. We aren't interested in roots of negative numbers per se, but complex numbers are handy.

Complex numbers are the sum of a real and imaginary part such as (a + bj). The math works as if the imaginary part were at right angles to the real part, which is the way a lot of audio phenomena behave. Some authors indicate variables that represent complex numbers with a capital letter. Numbers whose imaginary parts are zero are called **real** numbers.

The letter ω (Greek lower case omega) is often used to refer to angles. You will see $\omega=2\pi f$ as a way to convert a frequency f to an angle. When derived this way, ω may be called angular frequency.

The letter **e** means Euler's constant, a number like π that is one of the fundamental features of the universe. It is used to calculate interest, and is the base of natural logarithms. The interesting item here is that you can represent a sine wave as $e^{j\omega}$. The j makes this a complex number, the ω makes it polar notation. $e^{j\omega} = \cos(\omega) + j\sin(\omega)$.

A **function** is the mathematical representation of any sort of curve. A sine wave is a function that could be written $f(t) = k\sin(\omega t)$. Functions are always functions of some thing, in this case t. Some texts make the distinction that f(x) is a continuous function (i.e. an unbroken curve) and f[x] is a discreet function, made up of a lot of points. (Like a sampled waveform.) f(x) is pronounced "f of x".

A **coefficient** is a value that adjusts the overall value of a term in a function. For $k\sin(\omega t)$, k is a coefficient. Most of the hard part of DSP design is finding the coefficients that make a function work the way you want it to.

The **delta function** is a 1 followed by as many 0s as you want. It's the mathematical equivalent of hitting something with a stick to see what it will do. When you apply a delta function to digital filter, you get its **impulse response**. The Fourier transform of the impulse response is the frequency response.

A **transform** is a method for converting a function of time into a function of frequency (or back). In audio, it converts a chunk of waveform into a spectral representation.

You **multiply** two **functions** by multiplying the value of each point on one curve by the value of the equivalent point of the other curve.

Correlation of functions is performed by multiplying each point on one curve by all of the other curve. This gives one complete curve per point. You then add all of these together.

Convolution of functions is performed by multiplying each point on one curve by the reverse of the other curve (that's the other curve backwards) and adding all the results. Convolution of two time domain functions is equivalent to multiplying the frequency domain versions, and vice versa.

Transforms

There are several transforms out there - Laplace, Z-transform, and Fourier being the big names.

The **Laplace** transform converts a waveform into a series of exponentially changing sinusoids. This results in a whole family of curves of amplitude vs. frequency (one curve for each possible exponent) These are represented by parallel curves in a three dimensional space called the s domain. This is very useful for designing analog filters, whose response is a combination of exponential shapes.

The **Z transform** converts waveforms into something similar to the s plane, but in a polar scheme known as the Z plane. The frequency is represented by the angle, and the exponent by the radius, so the amplitude vs. frequency curves are wrapped in circles. This is needed for designing digital filters.

The **Fourier transform** converts waveforms into a series of sinusoids. A sinusoid is a waveform shaped like a sine wave, but not necessarily starting at 0.0. The Fourier transform actually results in two curves, one that represents the amplitude of the sinusoids and another that shows the phases.

Types of Fourier transform

There are four types of signal encountered in audio. These signal types are:

- Non-periodic analog signal
- Periodic analog signal
- Non-periodic digitized (discrete) signal
- Periodic digitized (discrete) signal

Different variants of the Fourier transform are appropriate to each. Some of these are referred to with initials.

- In the case of the non-periodic analog signal, the sinusoids may have any frequency, and there may be an infinite number of them. The original Fourier transform deals with this.
- In the case of the periodic analog signal, the sinusoids take frequencies that are multiples of the fundamental established by the period (The **Fourier series**). There

still may be an infinite number. (You need an infinite number of sinusoids to represent a corner in a waveform, so square waves and triangles have infinite Fourier representations.)

- In the case of the non-periodic digitized signal, the sinusoids are discrete themselves, and are limited in frequency to the range of one half of the sampling rate. The **D**iscrete **T**ime **F**ourier **T**ransform is used here.
- In the case of the periodic digitized signal, the sinusoids are discrete themselves, and limited in frequency to one half of the sampling rate. This can be analyzed by the **D**iscrete **F**ourier **T**ransform, or the **F**ast **F**ourier **T**ransform.

All of these transforms have inverse transforms, which take us back to the original waveform.

The transforms for analog signals are the theoretical basis for all of the math, but you can't actually do any of them in a computer. Computers deal with discrete signals only, because all they know is lists of numbers. These are processed by the DFT. The waveform goes in as a list of numbers, and two lists come out. What numbers do we need in the lists?

Details of the output

A sinusoid has frequency, amplitude and phase. These three parameters must be included for each of the components of a waveform. When we are dealing with periodic waveforms, the frequency of a component can be implied by its position in the list. The first component is DC, the next the fundamental, then the second harmonic, and so on. So lists with values for each component will be sufficient. What do the two values mean?

Amplitude and phase are one possibility. In that case each pair of numbers is a **polar representation** of the component (phase is an angle, and a number and an angle define a point on a polar plot.) This is very nice, because the list of amplitudes can be charted as a graph of the frequency response. This is what a spectral plot actually is. We generally ignore phase when we are just looking, since phase has no audible effect.[24] For

[24] The inaudible effects of phase are significant, and forgetting about phase when you are processing audio and not just listening is a serious mistake.

mathematical convenience, the amplitude range is normalized from 0 to 1.0, and the phase is from $-\pi$ to π.[25]

Phase is hard to do math with. The fact that the phase wraps around from $-\pi$ to π, and that math leads to division by 0 make the code awkward. For computation, a **rectangular representation** of amplitude and phase is handier. This can be done by specifying each component as a sum of a cosine wave and a sine wave. This is the most common definition of the Fourier series:

$$1/2a_0 + (a_1\cos(x) + b_1\sin(x)) + (a_2\cos(2x) + b_2\sin(2x)) + \ldots$$

This is usually output as two lists, one with the a values and the other with the b values. To keep the lists the same length, a b_0 of value 0 is included in the b list. The lists are called the cosine and sine parts or sometimes the real and imaginary parts. (Even if there are no complex numbers around). If there are N points in the input sample, there will be N/2 + 1 points in each list. The frequency x is the sample rate divided by N. A 512 point DFT with a sample rate of 44.1 khz gives a fundamental frequency of 86.13 hz. The highest frequency represented is half the sample rate.

This is called the real DFT. The algorithm that does this is rather inefficient (there's a correlation with every potential sinusoid), so a streamlined version called the **F**ast **F**ourier **T**ransform is preferred. The FFT also outputs two lists, real and imaginary, which actually are complex numbers.

$$(a_0\cos(0) - b_0 j\sin(0)) + (a_1\cos(x) - b_1 j\sin(x)) +(a_2\cos(2x) - b_2 j\sin(2x))$$

These can be changed to amplitude and phase pairs with a simple Cartesian to polar conversion.

There are N components in each list from the FFT . They still represent multiples of x = SR/N. The values beyond half the sampling rate represent negative frequency components running from - (N/2 -1)x to -x.[26]

[25] We usually do angles in radians. It's nicer for the computers. And yes, negative phase is as likely as positive.

The **FFT** takes complex numbers as it input[27]. If the input is a not a complex signal, the imaginary parts of the input are zero and the output curves will be symmetrical about the N/2 point. This implies a lot of wasted computation. This can be skipped, giving the **real FFT**. The output looks just like the complex FFT.

At this point, many may be worrying about how waveforms that are harmonic series on other fundamentals can be accurately represented by components based on the arbitrary frequency SR/N. Consider an example: a 100 hz tone. This falls between the 86.13 hz and the 166.32 hz components of the transform, so would show as an increase in each. These, plus the phase information are enough for the iFFT to give a 100 hz tone back. Of course, the more points in the analysis, the more accurate the reconstructed signal.

Transforming continuous signals

The FFT works with a chunk of signal N points long. The algorithm requires that N be a power of two. If a recording is too short, we can just add zeros, but more likely, we are interested in recordings much longer than N samples. We are also interested in how the transform changes over time, not just the overall average frequency content.

This problem is overcome with a system of windowing, which is similar to the practice of showing moving pictures by projecting a series of still pictures. In essence, the incoming signal is broken into chunks of N samples and analyzed as a series of frames. To smooth out the errors caused by this arbitrary chopping, overlapping chunks are processed-- when the signal is reconstituted, the overlapped frames are mixed together. For further smoothing, the frames are faded in and out before the analysis. There are several schemes for doing this.

Putting the FFT to work.

So, what's the point? Well, there are several pretty neat tricks we can do with Fourier transforms of signals. In real time systems the fft is a pair of synchronized signals, one

[26] This makes the outputs symmetrical curves. For many purposes the negative frequencies can be ignored, but the iFFT uses them when converting back to waveforms, so the result would be a loss in amplitude.
[27] Most waveforms are real, not complex. Where would you find a complex waveform? As the output of an inverse FFT.

with the coefficients of the magnitude terms and the other with the coefficients of the phase terms. These can be routed and processed just like any other signal. (But you wouldn't want to listen to them!)

Viewing Spectra.

Real time spectral analysis programs display the amplitude coefficients on the screen. Usually the phase terms are ignored.

Filtering

When the fft signal has been converted into magnitude and phase format, it should be clear that you can change the amplitude of the reconstructed signal just by changing the magnitude part of the fft signal. If you isolate particular bands in the fft output, you could change just that part of the spectrum. The ifft will then give you the modified sound.

Signal morphing

The next step is to use an input signal for the modification source. Transforms are derived from two signals, then the amplitude coefficients from one are multiplied by those of the other. The result is something of both, like vocoding. A gradual transition from one to the other creates a smooth change of timbre.

Cross synthesis

You can also add spectra. This gives a type of mixing when you add magnitude coefficients, but adding the phase terms will result in new kinds of sound.

To learn more about the Fourier transforms, I suggest you brush up on your math, then study

Smith, Steven; The Scientist and Engineer's Guide to Digital Signal Processing, 1997. (download in pdf or order from www.DSPguide.com)

Roads, Curtis; The Computer Music Tutorial 1996. MIT Press

A Glossary Of Technical Terms Used In Electronic Music

AC abbr. **alternating current** 1: current which varies over time at an audio rate. 2: the electrical power supplied by the utility company

AC bias ultrasonic component added to signal during the recording process in order to improve fidelity of analog tape recordings

ADC abbr. **analog to digital converter**: circuit which derives binary numbers from voltage measurements

additive synthesis method of production of complex tones by combining the outputs of several oscillators

ADSR acr. attack, decay, sustain, release: module which produces a voltage envelope of the pattern: rise, partial fall, steady state, fall to zero

algorithm procedure for solving a problem

aliasing errors caused in sampling systems by sampling at a rate less than twice the frequency of the input signal

ampere unit of measurement for electrical current

amplification process whereby a low power signal is converted into a high power signal

amplifier device capable of increasing the power of a signal

amplitude 1 : on a graphic representation, the vertical component of a function 2: in electronics, the average strength of a signal measured as voltage, current, or power

amplitude modulation process of changing the amplitude of an existing signal, usually at an audio rate; produces sidebands at sum and difference of the frequencies involved

analog **n**: electrical or mechanical representation of a signal that is continuously equivalent to the measured aspect of the signal : **adj**: electronic circuit designed to utilize such representations

anharmonic not part of a harmonic series

aperiodic not repeating

artifaction the use of material primarily for its symbolic function

attenuate to reduce the amplitude of a signal

band-pass filter filter which attenuates signals of frequency above and below specified limits

band-reject filter filter which attenuates signals of frequency within specified limits

bandwidth difference between the highest and lowest frequency signals that will be properly processed by a given device

bin A single frequency complnent of a Fourier Analysis of a waveforem.

binary having two possible states, such as on and off or 0 and 1

binary number number represented as a group of binary digits

binary word a group of binary digits of length convenient for a given computer system

bit a binary digit, usually 1 or 0

black box a device with a single function and few if any controls

bug in digital systems, a software error that prevents proper operation of a program

byte a binary word 8 bits long

carrier the signal that is to be processed by a modulation technique

CD abbr **compact disc** a medium for the recording and mass distribution of digitally recorded music or information

channel number data encoded into a MIDI message that specifies which devices the message applies to

clock **n** 1:circuit which provides a constant frequency stream of pulses used for timing operations 2: the signal produced by such a circuit **v**: to control a stepwise process as "the computer may clock the sequencer"

CMN acr. **common music notation** used in reference to composition programs that allow specification of notes by traditional looking graphics

comb filter device which is capable of attenuating signals within several narrow frequency bands

compander device capable of compression or expansion of signals

complex tone continuous sound with a waveform that is not a sine wave

compression process of reducing the dynamic range of a signal

conductor material that will carry electrical current

control voltage signal that is used to modify the operation of a synthesizer module

CPU acr. **CENTRAL PROCESSING UNIT** : the part of a computer that manipulates data

current a measurement of the amount of electricity flowing in a circuit

cutoff-frequency frequency at which a filter will attenuate a signal by 3dB : the demarcation point between attenuated and non-attenuated signals

DAC acr. **DIGITAL TO ANALOG CONVERTER** : circuit which converts numbers into continuously changing voltage

DASH acr. **digital audio, stationary head** a professional level format for digital recording and playback of sound

DAT acr. **digital audio tape** a consumer level format for digital recording and playback of sound

decibel expression of the relative powers of two signals: equal to $10 \log P_1 / P_2$ or $20 \log V_1 / V_2$

difference tone a signal created by the interaction of two existing signals with a frequency equal to the difference between the two existing signals

digital recording information stored as sampled numeric representations of a measured quantity

digital system circuitry implemented using discrete state circuit elements

digital computer programmable data processing device built with discrete state circuit elements

dither random variation imposed upon the least significant bits of a data stream to remove accidental periodicity in the quantization noise

dsp acr. **digital signal processor** a device or circuit which modifies audio signals by digital procedures

duty cycle description of a pulse waveform which represents duration of one state as a percentage of the period

dynamic range measurement of the difference in dB between the loudest and softest parts of a recording

dynamic window expression of the maximum dynamic range that can be properly handled by a recording system

enharmonic two ways of writing the same pitch

envelope 1: (amplitude envelope) a description of the manner in which the amplitude of a sound changes with time 2: pattern of voltage change similar to the above

envelope follower module of an analog synthesizer which produces a voltage proportional to the short-term average amplitude of a signal

envelope generator module of an analog synthesizer which produces pattern of smoothly changing voltage upon command

equalizer device with a complexly adjustable frequency response

equalization process of modifying the spectral content of a signal

event list 1: a list of notes and procedures in a composition ordered by time of occurrence 2: used in reference to composition programs that require specification of notes by typographical symbols

expansion process of increasing the dynamic range of a signal

FFT acr. **fast fourier transform** : algorithm for waveform analysis; result is spectral plot of the analyzed signal

filter device capable of attenuating signals with frequencies outside of specified limits

firmware programs which are built into a computer system

FM acr. **frequency modulation** the process of changing the frequency of a generated signal at an audio rate, spectrum of result contains the carrier plus sidebands spaced at the frequency of the modulator

FM synthesis algorithm which uses frequency modulation to create complex waveforms in digital synthesizers

formant frequency region of an instrument's sound spectrum which is emphasized independently of the pitch played

Frame A spectrum plot, usually one of a series of snapshots of a continuing signal

frequency rate of occurrence of a repetitive event

frequency response description of the manner in which the output of a device varies with frequency of the input signal

Fourier synthesis algorithm which uses addition of values taken from a sine table to create complex waveforms in digital synthesizers

fundamental 1: the lowest frequency component of a harmonic waveform, usually the perceived pitch 2: the root of a harmonic series

gain expression of the amount of amplification provided by a particular device; the ratio of the amplitude of the output signal to the amplitude of the input signal

gate **n** 1:timing signal within a synthesizer which provides duration information for notes 2: in digital systems, a circuit which performs a simple logical operation **v**: to command a circuit to allow a signal to pass

glitch 1: a discontinuity in a normally smooth function 2: in digital systems, a hardware produced error in program operation

grammar the traditional principles or rules of a specific language or art

graphic equalizer device which allows the adjustment of signal amplitude within several fixed frequency bands

hardware the physical components of a computer system

harmonic series a set of numbers, each of which is a whole number multiple of a specified root value

harmonics the components of a complex tone which have frequencies that lie on a harmonic series

headroom dynamic region above the usual operating level and below the clipping level of a device

Hertz abbr **hz** unit of frequency, equal to 1/sec

high-pass filter a device which attenuates signals of frequency lower than some specified value

hybrid synthesizer a complex signal generating system which consists of a group of analog synthesis modules controlled by a digital computer

impedance total opposition to current flow in a circuit, including both resistance and reactance

insulator material that will not carry measurable electric current

integrated circuit complex electronic circuit realized on a single chip of silicon

juxtaposition compositional process of arranging sound elements in a series

kilo- prefix indicating one thousand units

level rough indication of the amplitude of a signal

loop n 1: a short section of a recording which is constructed to repeat continuously 2: the sound produced by a loop v 1:to create a loop 2: to repeat endlessly 3: (film) to synchronize dialog to existing footage

loop point particular sample in the memory of a sampling synthesizer which marks the beginning or end of a repeated section

loudness 1: the perceived strength of a sound, strongly related to the amplitude of the sound 2: a function found on some consumer stereo amplifiers, which adds a low frequency boost to the signal

low-pass filter a device which attenuates signals of frequency higher than a specified value

mega- prefix indicating one million units

merge v to combine two or more streams of digital data

message within the MIDI protocol, a single performance instruction

micro- prefix indicating one millionth of a unit

MIDI acr. **Musical Instrument Data Interface** a protocol for the exchange of music performance instructions between devices

milli- prefix indicating one thousandth of a unit

mix n : a particular balance of several musical elements **v**: to combine several simultaneous analog signals

mixdown the process of combining the elements of a multi-track recording into the final composition

mixer device which is capable of combining several signals, often includes other functions

montage compositional process of ordering sounds in simultaneous or overlapping combinations

modulation process of change : in synthesis, to change some parameter at an audio rate in order to create complex waveforms

MP3 Short for **M**otion **P**icture **E**xperts **G**roup level 3 audio encoding algorithm, a method of lossy data compression for audio files

multiplex v to combine several information streams into one channel

multi-timbral adj implies that an instrument is capable of producing several simultaneous notes which differ in timbre

multitrack recording system of synchronizing sounds by recording them side by side on the same piece of tape

noise 1: continuous signal or waveform lacking any periodicity 2: any undesired sound or signal

note-off a MIDI message that stops a note with a specified pitch and velocity

note-on a MIDI message that starts a note with a specified pitch and velocity

Nyquist theorem description of the frequency limits of an information sampling system; the sampling rate must be twice the frequency of the sampled signal

operating system the collection of software in a computer that deals with specific tasks such as disk access

oscillator electronic circuit which generates a continuous signal of a specified waveform

oscilloscope instrument which displays voltage and time relationships of a signal on a screen

overtones (obs.) those components of a complex tone which are not the fundamental

parameter any of a set of physical properties whose values determine the behavior or characteristics of something

parametric equalizer device which permits adjustment of the amplitude of signals within variable frequency bands

partials components of a complex tone

patch **n**: 1 a particular set of connections in an analog synthesis system 2: a list of the parameter settings of a digital synthesis system **v**: to make connections between devices in a system

PCM abbr. **pulse code modulation** a technique for recording digital information on tape or disc

period time between the corresponding points of two succeeding cycles of a repeating waveform

phase description of the time difference between two signals of equal frequency

phase modulation technique of imposing information on a carrier signal by manipulating its phase

phone adj phone plug or jack, the type of connector established as a standard by the telephone company.

phones headphones

phoneme the smallest sound component of speech

phono abbr phonograph **adj** phono connector**-** the type of connector associated with consumer audio gear

pico- prefix indicating one millionth of a millionth of a unit

pink noise sound of random spectral content with an equal distribution of power in each octave

pitch description or name of a tone based on perception of its frequency

polyphonic adj implication that a system is capable of producing several independent musical voices

portamento effect of gliding into the pitch of a note rather than changing abruptly from one note to another

potentiometer circuit element which is capable of reducing the amplitude of a signal by an adjustable amount

program list of instructions a computer will carry out to serve a particular function

"Q" quality of a filter circuit which describes, among other things, the tendency to accentuate signals close to the cut-off frequency

quantization noise noise in a digital system attributable to the fact that all values must be rounded to a whole number of bits

RAM acr. **random access memory** memory locations in computer system that may be altered by a running program

reactance opposition to a change in current flow in a circuit

reciprocal quantity obtained by dividing a specified value into one

resonance tendency of a physical structure to oscillate in sympathy with an applied sound

resynthesis any of a number of techniques for reconstructing a waveform out of the data from some analysis such as the Fourier transform

ring modulator analog synthesizer module which produces a variation of amplitude modulation in which the original signals are suppressed in favor of the sidebands

ROM acr. **read only memory** memory locations in a computer that may not be altered by a running program

sample-and-hold acr **S&H** module of an analog synthesizer which, upon command, captures and maintains the instantaneous value of a changing input voltage

sampled information system any system in which a continuous signal is converted into a series of discrete values for storage or processing

sampler a digital synthesis system based on the controlled playback of recordings of individual notes

sampling rate the frequency at which samples are taken or generated

sawtooth wave simple geometric waveform with a gradual rise in voltage and an abrupt fall

sequencer 1: module of an analog synthesizer which produces a timed series of control voltages 2: a computer program 3: a device which has the capability of recording and playing back MIDI note data

sidebands components added to the spectral plot of an existing waveform by various modulation procedures; usually symmetrically spaced about the fundamental of the original signal

signal the electrical manifestation of the information of interest; usually the electrical analog of sound

signal path the route of the signal from input to output of a device or group of devices

signal-to-noise ratio an expression in dB of the ratio between the most powerful signal a device can provide and the inherent noise that device produces in the absence of any signal

sine wave simple geometric waveform which follows the values found in a sine table

software 1:the collection of programs available for a computer system 2: the aspects of system behavior which may be modified by the operator

sound reversal effect generated by playing a sound recording backwards

spectral plot graphic representation of the amplitude and frequency of the components of a complex waveform

spectrum description of the frequency distribution of energy in a waveform

square wave simple geometric waveform with two voltage states, each present for one half the period

state-variable filter circuit or analog synthesizer module capable of providing low pass, band pass, and high pass filter functions simultaneously

subtractive synthesis process of deriving simple waveforms by filtering complex ones

supersonic obs. above the normal frequency range of the human ear

synthesizer electronic system capable of generating a variety of sounds with detailed control of the sound parameters

system exclusive a MIDI message which contains data defined by the manufacturer

temperament description of the relative frequencies of the pitches of notes in a scale

timbre 1: perceived sound quality 2: characteristic other than pitch or loudness that allows us to distinguish between sounds

tone steady sound of noticeable duration

track 1: the area on an analog tape where a single signal is recorded 2: an emulation of 1 in a digital recording system or sequencer program

trigger electronic impulse which commands certain modules to perform their functions

transducer device for converting energy from one form to another

transformer circuit component consisting of two adjacent coils of wire; useful for efficient transfer of power from one device to another

tremolo musical effect produced by the rapid alternation of two notes or rapid repetition of a single note

triangle wave simple geometric waveform with a gradual increase in voltage followed by a symmetrical gradual decrease in voltage

tuning frequency relationship of the notes produced by an instrument, either absolute or relative

ultrasonic above the range of human hearing (above 20kHz)

vaporware software which should be ready any day now

VCA acr. **VOLTAGE CONTROLLED AMPLIFIER** : module of an analog synthesizer which varies the amplitude of a signal according to an applied control voltage

VCF acr. **VOLTAGE CONTROLLED FILTER** : module of an analog synthesizer which attenuates signals of frequency outside limits that are set by an applied control voltage

VCO acr. **VOLTAGE CONTROLLED OSCILLATOR** : module of an analog synthesizer which generates various simple geometric waveforms at a frequency determined by an applied control voltage

vibrato musical effect produced by the rapid but sub-audio variation of pitch or amplitude of a tone

vocabulary in electronic music, the collection of sounds used as source material for a composition

vocoder device capable of imposing the spectral characteristics of one sound upon another

voice n 1: Circuitry or software of a synthesis system with the capability of producing a musical note; "a four voice instrument" implies four more or less independent notes may be produced at once. 2: A list of the parameters that specify a particular sound 3: Within certain composition programs, a part meant for performance by a particular instrument v: To specify the values of parameters that describe a sound on a synthesis system

volume rough description of loudness of sound produced by an electronic sound system

wavefront the locus of a single pressure disturbance propagating through air

wavelength measurement of the distance between equivalent points on two succeeding wavefronts

wavetable-lookup algorithm for producing complex tones in digital systems by taking sample values from a pre-calculated list

white noise sound of random spectral content with an equal distribution of power throughout the audio range

Some Good Books For Technical Background

Recording

Modern Recording Techniques 3rd Ed
Huber And Runstein
373 pp. $29.95
The usual text for recording courses.

Sound Recording Handbook
John Woram
586 pp. $49.95
The definitive reference for pro studios.

Handbook For Sound Engineers
Glen Ballou Ed.
1,247 pp. $100
This is where Woram looks stuff up.

Sound Reinforcement

Sound Reinforcement Handbook
Davis & Jones
417 pp $34.99
Published by Yamaha, Everything you need to know about live sound.

ACOUSTICS

Sound Studio Construction on Budget
F. Alton Everest
The only acoustics book that actually tells you what to do.

Mainframe Computer Music

The Computer Music Tutorial
Curtis Roads, et al
1234 pp. $50
Very thorough coverage of the topic.

The Music Machine
Curtis Roads ed.
725pp. $47.50
A good overview.

Foundations Of Computer Music
Roads And Strawn ed.
712 pp. $22.95

Current Directions In Computer Music Research
Mathews And Pierce ed
432 pp. $37.50

Pretty heavy going, but solid information about what's really in there. These are reprints from Computer Music Journal.

Musical Applications Of Microprossors
Hal Chamberlin
802 pp. $39.95
Mostly for builders, but has the best descriptions of digital algorithms and analog synthesis(!) circuits around.

Programming Midi

MIDI 1.0 Detailed Specification
the International MIDI Association
5316 W.57th st
Los Angeles CA 90056
63 pp. $45 - cost depends on whether you join or not.

Midi Programming For The Macintosh
Defuria & Scacciaferro
371 pp+ disk $37.95
Very basic, and no use at all unless you are willing to spend another $150 for MIDI pascal & a Pascal Compiler. Too old to discuss MIDI Manager.

C Programming For Midi
Jim Conger
219 PP. +DISC $37.95
Bad code but some good basic information.

Maximum MIDI
Paul Messick
460 pp + Disc $49.95
This will get you going in writing MIDI code for Windows, if you are already familiar with the Windows API

DIGITAL AUDIO

Principles Of Digital Audio
Ken Pohlmann
474 pp. $29.95
All about the inner workings of digital recording.

An Introduction To Digital Signal Processing
John Karl
341 pp. $39.95
If you really mean it (examples are in FORTRAN!) this book shows how digital fiters and such are coded.

C Language Algorithms for Digital Signal Processing

Embree & Kimble
456pp. About $75
This one is usefull.

The Scientist and Engineers Guide to Digital Signal Processing
Steven W. Smith
On-line at www.DSPguide.com

JOURNALS

These are in the library:

Computer Music Journal

Doings of the big time research institutions.

Electronic Musician
All about what's available at the local music store.

Also, Keyboard, EQ, Mix and others available at local news stands and music stores.