ELECTRONIC MUSIC PITCH STANDARDS

Including the construction of a voltage-controlled oscillator with memory

BY DON LANCASTER

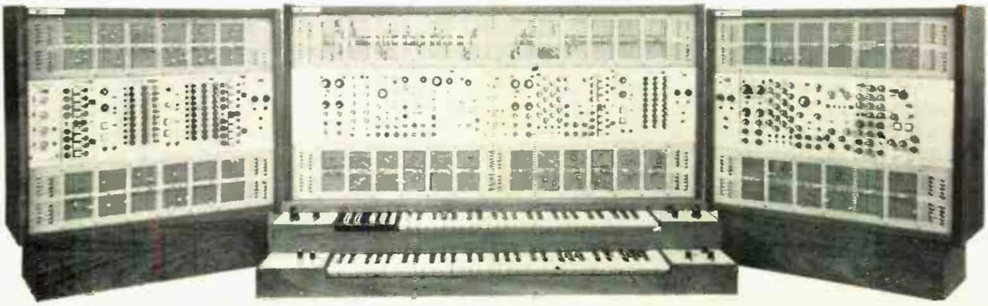
ANY electronic music system has to generate either a number of different tones or one or more tone producers have to be shifted around to get a desired tone sequence. Before we can look at how we produce tones, we have to ask some pretty basic questions: What is pitch? How stable must the tones be? What are their frequencies? And so on.

Pitch is the psychological perception of the *frequency* of a tone, just as *loudness* is the psychological perception of the *amplitude* of a tone. Since pitch is an "in the head" type of thing, it's affected by loudness, the presence of other notes, room acoustics, your particular mood, and a dozen other things. So, we really can't generate a pitch directly. All we can do is generate a note of a given frequency, amplitude, and stability, and hope that the final effect in the listener's head is the one we are after or at least close to it. Thus, our pitch generation problem is really one of making one or a group of related tone frequencies available.

There are two types of perceived frequencies. One is called *absolute pitch,* or the actual value as perceived in so many hertz or cycles per second. The other is the *relative pitch* or the frequency relationship between a certain note and a note played either with it or immediately before. Fortunately for music in general, very few people care about or can tell an absolute pitch, and it is only the *relative* relationship between notes on a frequency ratio basis that really matters. Thus, if *all* the notes produced change their absolute frequency slowly with time, very few listeners will notice. This also lets us move up and down a musical scale to get a different overall effect or to add variety or depth to a composition.

A pitch generator is a circuit or system that makes available several notes of different frequencies. It does this in such a way

The ARP Model 2500 electronic music synthesizer.



that the frequency ratio relationships can be combined to get an overall effect that at least the composer, and hopefully his listeners, will consider "music."

Theoretically, we could use any group of tones of any stability we like. This might be fine for special effects, music soundtracks, and TV commercials, but there are certain things that get in the way of using a more or less random selection of tone frequencies. The three most important are: the physical characteristics of the ear and hearing, the traditions and music that went before in the mind of the listener, and, finally, whether the electronic musical instrument you are building is going to be used either with other instruments or in an attempt to synthesize their sounds. While you as a composer can select any group of tones you like, you may not have very many listeners if your selection violates basic physical properties of the ear or goes completely against the cultural grain of your intended listeners.

What the Ear Likes To Hear. There are several characteristics of the ear that set up the basic ground rules. These are pretty much independent of the listener's cultural or musical heritage.

The first of these is that the ear is *not* a linear device. It works on a logarithmic basis both for the perception of pitch and for loudness. For instance, the sound energy of a whisper added to another whisper gives an apparent doubling of loudness. The same amount of energy added to a shout gives very little if any added loudness perception. What counts is the *change* in the amount of energy received, not on a "how much is it?" linear basis, but on a "how much is it in relation to what we have already?" log basis.

This, of course, gives us the familiar decibel or log relationship where a change in loudness of one decibel (around 10% amplitude or 20% power) is about the smallest change we can normally detect.

The same type of log relationship applies to pitch as well as loudness. It's very easy to spot a 10-hertz frequency difference as you go from 40 hertz to 50 hertz, since the *percentage* or relative change in frequency is so great. If you go from 4000 to 4010 hertz, the frequency difference between the two is still 10 hertz, but the *percentage* change is so small that a trained ear could just barely detect it.

So, our notes have to be spread out on an exponential or log basis, in order that the frequency difference between the low notes is low and the frequency difference between high notes is higher—perhaps maintaining the same percentage change between successive notes over the scale. Thus, any scheme to make all the notes equally spaced in frequency is doomed to failure because the ear just doesn't work that way. The notes would end up much too far apart at the low end and too cramped together at the high end to allow very much meaningful to be done with them. Any reasonable pitch generator has to spread out the high notes and cramp together the low ones. The type of scale that best fits the ear's characteristics turns out to be one where each successive tone is an equal percentage or constant ratio higher in frequency. This is called an *equally tempered* scale, and produces tones that appear to be just as far apart in the low register as in the high.

TABLE 1—STANDARD FREQUENCIES FOR TWELVE-NOTE EQUALLY TEMPERED MUSICAL SCALE

Octave Number	Note (in hertz)											
	C	C#	D	D#	E	F	F#	G	G#	A	A#	B
0*	16.351	17.324	18.354	19.445	20.601	21.827	23.124	24.499	25.956	27.500	29.135	30.867
1	32.703	34.648	36.708	38.891	41.203	43.654	46.249	48.999	51.913	55.000	58.270	61.735
2	65.406	69.296	73.416	77.782	82.407	87.307	92.499	97.999	103.83	110.00	116.54	123.47
3	130.81	138.59	146.83	155.56	164.81	174.61	184.99	195.99	207.65	220.00	233.08	246.94
4	261.62	277.18	293.67	311.13	329.63	349.23	369.99	391.99	415.31	440.00	466.16	493.88
5	523.25	554.36	587.33	622.25	659.26	698.46	739.99	783.99	830.61	880.00	932.32	987.76
6	1046.5	1108.7	1174.7	1244.5	1318.5	1396.9	1479.9	1567.9	1661.2	1760.0	1864.7	1975.5
7	2093.0	2217.5	2349.3	2489.0	2637.0	2793.8	2959.9	3135.9	3322.4	3520.0	3729.3	3951.1
8	4186.0	4434.9	4698.6	4978.0	5274.0	5587.7	5919.9	6271.9	6644.9	7040.0	7458.6	7902.1

*Octave zero is very seldom used.
This scale applies to most musical instruments except the piano.

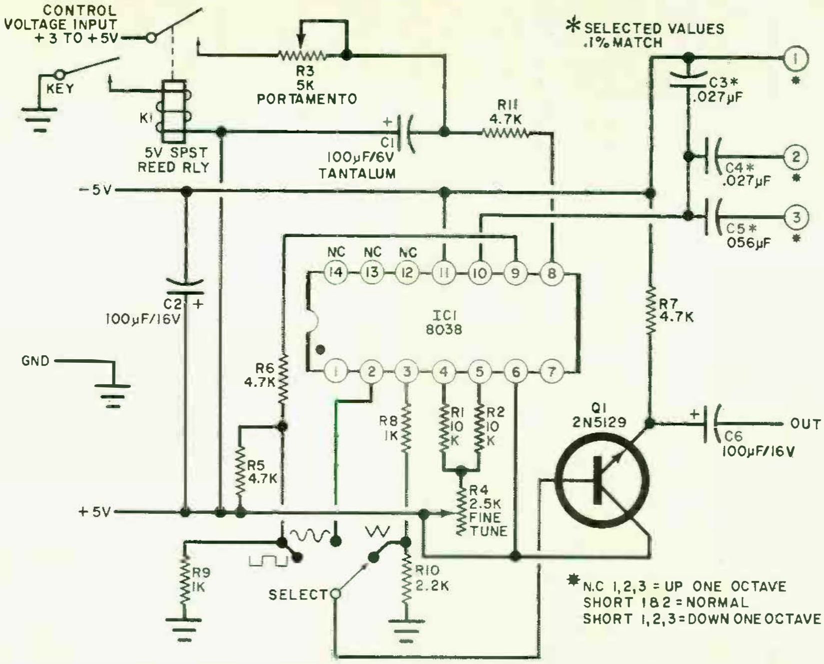


Fig. 1. Voltage controlled oscillator has memory to hold notes after key release.

A VCO WITH MEMORY

The vco or voltage controlled oscillator circuit described in the October issue (p. 37) gave us a very low-cost, extremely stable way of generating simultaneous sine, square and triangle waves. By the way, the temperature stability is such that around 20 degrees F are needed for a 1-cent change in frequency. Supply sensitivity is about 10 cents per volt, so a reasonably regulated supply is recommended.

Figure 1 shows how we add memory to the basic circuit and provide an emitter follower output to handle the normally high output impedance of the sine wave generator section.

The sample-hold circuit on the input uses a reed relay and a tantalum capacitor. It remembers what key was pressed after it is released so that the envelope generating circuit still has a note to work on during the decay portion of the note. Pins 1, 2 and 3 add capacitance or remove it so that the vco has to work only over a two-octave range, greatly easing the control voltage and stability problems. Note that a reasonably regulated 10-, 12-, or 15-volt supply is needed. Figure 2 shows the control voltages to apply for the various notes. Potentiometer *R3* controls the *portamento*.

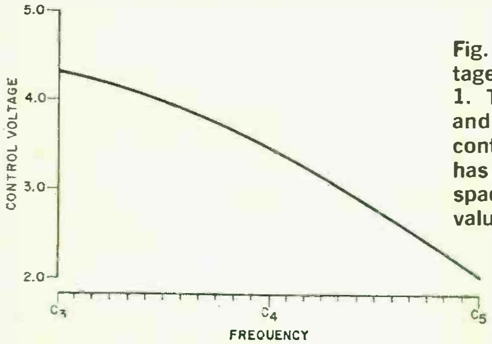


Fig. 2. Typical variation of control voltage vs frequency for the circuit in Fig. 1. The vco by itself is extremely linear and is related to the difference between control input and positive supply. Curve has to be suitably "bent" to fit the log-spaced, equally tempered notes. The exact values can be calculated by using Table II.

Beats. Since the ear has to be nonlinear to accommodate such a wide range of tone frequencies and amplitudes, it reacts to two tones sounded at once by producing the sum and difference frequencies. The difference between two tones is called a *beat*. Beats of a very few hertz or fractions of a hertz are normally considered a pleasant sensation or a *consonant* sound. Beats higher in frequency may often be perceived as a "wrong" sound, called a *dissonance*.

Beats can be perceived even with two low-level sine waves. When tones have a more elaborate *timbre* (harmonic and overtone structure), the ear will also beat the harmonics of the tones against each other, generating a wide range of different beat frequencies. The overall result depends on how strong these beats are and where they lie in frequency.

Intervals. The characteristic of the ear to mix and multiply tones together has led to some fundamental and important musical ratios called *intervals*. Just about anyone can tell when two notes are of identical pitch, particularly if they have an identical overtone or harmonic structure and if the fundamental is strong.

Two notes of identical frequency played together are said to be in *unison*. If they are absolutely identical in frequency and phase and coming from the same source, you won't be able to tell there are two tones at all. They sound like one. If the tones differ slightly in frequency or phase, you get a *chorus* effect which adds *warmth* to the tone. This is why a group of violins in an orchestra sounds much richer than does a single violin. Of course, if the two tones get too far apart in frequency, the notes become dissonant or *out of tune*.

The 1:1 frequency ratio is very easy to spot. The next easiest ratio is a 2:1 frequency spread between two notes. If the notes have a complex harmonic structure, the beats between harmonics will minimize as you hit exactly a 2:1 ratio. Even with pure sine waves at low amplitudes, there is something "special" about this 2:1 relationship to the ear. This interval is called an *octave* because in one system of music there are eight notes before the frequency is doubled.

It turns out that the ear "likes" to hear or finds musically pleasant, any tone ratio that consists of *small whole numbers*. Another important interval is called the *fifth*, and its frequency ratio is 3:2. As the number ratios get larger, the tones tend toward sadness and then to dissonance.

A *scale* can be based entirely on fifths. For instance, we can start with a note, arbitrarily called C. We can find a new note at a 3:2 ratio above it, and call this G. We can find a new note at a 3:2 ratio above G. This will be in the next octave, but we can divide it down by two to put it between C and G. It may be called D. D will be (1/2) (3/2) (3/2) C or a ratio of (9/8). A fifth above D will be (3/2) (9/8) = (27/16) above C. We call this A, and so on. When we have generated all the notes, we have a seven-note scale in the *key* of C. The notes in order turn out to be C,D,E,F,G,A,B, and finally back to C one octave higher. This is called a *chromatic* scale or *Pythagorean* tuning.

The space between E and F and the space between B and C are only about half that of the other intervals. The other intervals are called *whole tone* steps, while the narrow ones are called *semitone* intervals. It became logical to add new semitone intervals between the other notes, so five new keys are added between the others, and called C♯ (or C sharp), D♯, F♯, G♯, and A♯. The scale then consists of twelve notes of one semitone each. This allows the playing of *minor* or generally sadder tonal sequences as well as the *major* or brighter, happier tonal sequences.

This is a fine scale as long as everyone is willing to play in the key of C. But certain instruments have a range that favors other starting points. Every time you pick a different key or starting point, you generate 12 new notes. Some of these new notes are so close to the others it doesn't matter; others are so far apart and so bad that they are called *wolves*.

To make a long story short, a compromise approximation has to be made, one that ends up with a reasonable number of notes per octave, but that sounds almost as good as the chromatic scale, yet lets each instrument play in its own key as needed.

To an engineer, the solution is obvious—use twelve notes and space each one an equal percentage above the one below, exactly fitting what the ear likes to hear. It took musicians centuries to get around to the same consclusion. The result is called the *equally tempered 12-note scale*. To standardize usage around the western world, the center C was called *middle C*, and the A above middle C was eventually standardized to a frequency of 440 hertz. The notation was also standardized. The lowest note is called C_0 at 16.35 hertz. The first octave continues $C\#_0$, D_0 . . . . . B_0, and the next octave starts at C_1 and so on. Middle C turns out to be C_4 at 261.6 Hz and the standard pitch A above middle C is A_4 at 440 hertz. Table I shows all the notes of the equally tempered scale with 12 notes per octave.

Mathematically, if you are going to double something by multiplying it by itself twelve times, the basic interval must be the twelfth root of two, or 1.0595. This is roughly six percent. Since you might like to design your own way of generating tones in this sequence, Table II lists the twelfth-root-of-two ratios and some number series that *approximate* the ratios. Note that the twelfth root of two is an irrational number—in no way will you generate it exactly.

How Stable? The equally tempered scale then provides a reasonable approximation to all the individual chromatic scales and lets us play Western music with only twelve notes per octave. Our next question should be how accurate must our approximation to each note be? Also, how stable does our electronic instrument have to be?

We've already found out that an absolute drift by a small amount of all the notes together really doesn't matter too much unless you get too far off. What really matters is the relative change between notes.

The interval between any two successive notes is called a semitone, and corresponds to roughly a 6% change in frequency. We obviously have to be much better than this. It's convenient to break the semitone interval down into 100 parts, and call each 1/100th of a semitone a *cent*. A 1-cent change in frequency is a change of 0.0595 percent, or roughly 0.06% or 600 parts per million.

Under carefully controlled conditions, certain musicians can sometimes spot a cyclical variation of a few cents in pitch. So, a good criterion for any electronic musical instrument is to set and hold the pitch to within one cent of its design value; perhaps a note or two could be off by twice this without too much harm, so long as the overall average worked out to under one cent.

Must We Use Western Music Standards? The obvious answer to this is, "no, but. . . . ."

TABLE II—SOME PROPERTIES OF THE EQUALLY TEMPERED 12-NOTE SCALE

Note	Ratio	Series A	Series B
C	1.0000	232	478
C#	1.0595	219	451
D	1.1225	207	426
D#	1.1892	195	402
E	1.2599	184	379
F	1.3348	176	358
F#	1.4142	164	338
G	1.4983	155	319
G#	1.5874	146	301
A	1.6818	138	284
A#	1.7818	130	268
B	1.8877	123	253
C	2.0000	116	239

From the fundamental physical properties of the ear, we should use a scale that is related by octave intervals, and we should use a scale of equal tempering. These are pretty much independent of cultural musical heritage. This leaves several choices.

You can use the traditional 12-note equally tempered scale and the traditional rules of composition and chording and end up with musical sequences that are more or less in the traditional heritage of western music. You also will be able to play with and synthesize tones of the conventional musical instruments this way.

Or, you can use the 12-note equally tempered scale *without* the traditional rules of the game. You treat every note as an equal and do your own thing. This is called *well tempered* composition and it does go back a way and has had limited success.

Or, you can change the number of notes per octave but keep equal tempering and octave relationships. Up to 31 notes per scale have been used, and a 31-note scale easily accommodates eastern musical instruments (such as the Indian sitar) as well as conventional. There are some ground rules for these exotic scales. Today, the hardware to generate them really isn't too bad, but the keyboards get rather complex and hard to play.

Finally, you can say that everything everyone else has ever done in music is wrong and go out and do your own thing, using completely unstructured or random tonal sequences. While this can generate some interesting compositions and is quite useful for special effects, for a steady diet of this sort of thing, you might have to end up renting your listeners by the hour. Synthesizing or accompanying a traditional instrument will also be rather tricky. ◆