AUDIO TOPICS

The development of the music scale

The musical scale in universal use today did not just happen, nor was it invented suddenly. As this article will show, it was developed over a period of centuries, from a few tones to the present complex structure of keys and modes.

By L. C. Debnam*

Since the advent of radio and electronic sound reproduction, music has become a part of nearly every person's experience. Although a great many people enjoy music, and many of these play musical instruments, its physical basis is not generally understood. Musicians are primarily interested in music as an art or a living and other people are interested in it as a form of relaxation or enjoyment.

Physicists in general are not interested in music as a science because a detailed study yields insufficient rewards. One of the main reasons for this is that music is not generally subject to fixed laws and equations, much of the enjoyment being due to physiological and psychological factors. "Correct" music may be played by anyone on a pianola, and computers may be programmed to play almost any instrument—but these "playings" pale to insignificance when compared with the same tune played by a Winifred Atwell or an Arthur Rubinstein.

Science associated with music appears to have started with Pythagoras of Samos, who discovered, about 525 B.C., that two or more strings when plucked at the same time made a pleasant note if the length of the strings were in the ratios of small whole numbers. The lengths of the strings are related to the frequencies of the notes emitted and it is more usual to discuss musical notes in terms of frequencies.

* The author is Instructor in Mathematics and Physics at the Darling Downs Institute of Technology.

The simplest ratio of frequencies involving whole numbers is the ratio 2:1, and if two notes are sounded with this ratio, i.e., one being twice the frequency of the other, the resultant sound is not unpleasant to the ear. Pythagoras admitted only three ratios which sounded pleasant. These were notes with the frequency ratios 1:2, 2:3 and 3:4 and from these ratios the tetrachords (four-stringed instruments) were derived. With this system as a basis, Greek music developed as melody in which any of the four fundamental notes may be followed by any other without the music seeming harsh. Although more than one note could be sounded together without discord, apparently only the frequency ratio 1:2 was used in this manner.

The sound of ancient Greek music was not consistent, as they used at least seven different scales which, although they bore the Pythagorean relationships, had different frequencies for their bases. These scales, or modes, were later adopted by the early Christian Church with some alterations and are known as the "Ecclesiastical Modes."

Figure 1. The tempered scale frequencies are shown in relation to a section of a piano keyboard, the names of the notes and their position on the musical stave.

Harmony (the simultaneous sounding of more than one note) was not generally recognised as legitimate in music until the fourteenth century A.D. Before then, nearly all music was melodic (consecutive) but during the period A.D. 1300-1600 harmony of the Pythagorean and other ratios was gradually accepted, along with melody of the same ratios. Since about A.D. 1600 music has developed more as harmony than melody.

Before the development of harmony the four-note scale of the Greeks had been improved by the admission of more ratios into the system of modes so that the scale took the form of the ratios

4, 5, 6, 8.

This scale may be seen to contain the Pythagorean ratios 1:2 (as 4:8), 2:3 (as 4:6) and 3:4 (as 6:8) with the addition of a "triad" of 4:5:6. This triad was found to be pleasant and to contain the Pythagorean ratios but the gaps between notes were large so a second triad of notes was fitted between.

The second triad has the ratios 5-1/3 : 6-2/3 : 8 (the same as 4:5:6) and the two triads interlock to form the scale:

4, 5, 5-1/3, 6, 6-2/3, 8.

About the year A.D. 1000 a third triad of 6 : 7-1/2 : 9 (again in the ratio 4:5:6) was introduced into the scale to fill the large gaps between 4 : 5 and 6-2/3 : 8. "9" is outside the range of 1:2 (4:8) so it was halved and inserted between 4 and 5 and the scale took on the ratio form

4, 4-1/2, 5, 5-1/3, 6, 6-2/3, 7-1/2, 8.

As there are eight notes in this scale it is known as an octave, from the Latin word for eight. The last note is exactly twice the frequency of the first and a new octave may be constructed using the same ratios but with "8" as a basis. Similarly a lower scale may be constructed which gives a total range of ratios for three octaves of the form

2	$2\frac{1}{4}$	$2\frac{1}{2}$	$2\frac{2}{3}$	3	$3\frac{1}{3}$	$3\frac{3}{4}$	4
4	$4\frac{1}{2}$	5	$5\frac{1}{3}$	6	$6\frac{2}{3}$	$7\frac{1}{2}$	8
8	9	10	$10\frac{2}{3}$	12	$13\frac{1}{3}$	15	16

and similar octaves may be constructed above and below these.

In the early part of the eleventh century, a Benedictine monk named

Guido d'Arezzo invented the four-line staff (which in modern music has been translated to the five-line staff) and placed letters on certain lines to indicate their pitch. These letters have gradually turned into the Clef signs known today, the G or treble clef; the F or bass clef; and the C clef. It is also traditional that he composed a hymn to John the Baptist which is as follows:

UT queant laxis REsonare fibris,
MIra gestorum FAmuli tuorum,
SOLve polluti LAbii reatum
Sanite Ioannes.

The capitalised syllables of this hymn correspond to ratios in the octave as follows:

4	4-1/2	5	5-1/3	6	6-2/3	7-1/2	8
Ut	re	mi	fa	so	la	—	—

which was the forerunner of the "do-re-mi" scale learnt today.

The fundamental frequency of the scale of modern music often causes some confusion. The ratios were in the eleventh century given letters to distinguish them but the fundamental note was quite variable. There are two scales in use in the world, the Physical Scale, based on the note "C" of 256Hz and the Musical Scale based on the note "A" of 440Hz. These may be compared with the ratios as follows:

Ratio	Name	Physical frequency	Musical frequency
4	C	256	261.6
4-1/2	D	288	293.7
5	E	320	329.6
5-1/3	F	341-1/3	349.2
6	G	384	392.0
6-2/3	A	426-2/3	440.0
7-1/2	B	480	493.9
8	C	512	523

The musical scale does not follow exactly the ratios expressed in the first line but the differences are generally less than 1 per cent when compared with C and are barely discernible to the ear.

The reason for the changes of frequency in the musical scale is the introduction of "altered notes," the sharps and flats represented by the black keys of a piano.

The altered notes (accidentals in musical terminology) were introduced in a confused process during the thirteenth to seventeenth centuries. The strict modal melodies were somewhat monotonous and harmony was not generally in fashion so that musicians occasionally **modulated** from one key (fundamental frequency) to another which imposed changes in the scale separations. If the "true" musical scale (corresponding to the true ratios with A=440Hz as fundamental) is used the ratios between consecutive notes are as follows:

Notes	True Ratio
D/C	1.125
E/D	1.111
F/E	1.067
G/F	1.125
A/G	1.111
B/A	1.125
C/B	1.067

When musicians changed from one note to another as "fundamental," the scale did not sound the same, because the ratios were different. This may be tried on a piano. If the scale is played in the sequence CDEFGABC' the ratios sound different from that played in the sequence DEFGABC'D'. To overcome this difficulty and at the same time to keep the notes at the same frequency, five new notes were introduced between the eight notes of the octave. These notes were fitted between C and D, D and E, F and G, G and A, and A and B. These "half-notes" enabled the scale to be filled so that approximately the same intervals could occur in an octave regardless of which fundamental note was used. It is interesting to note that the use of "altered" notes was an infringement of canonical law for ecclesiastical singing and our present knowledge of fourteenth and fifteenth century musical practice is incomplete because accidentals were rarely marked on musical scores of this period.

The accidentals were given names which correspond to the nearest "whole" notes, and termed "sharp" (indicated by the symbol #) and "flat" (indicated by the symbol ♭) to indicate whether the frequency was above or below the whole note to which it referred. For example the black note on a piano between C and D is known both as "C sharp" and "D flat." When written on a musical score the sharp and flat symbols are placed before the dot signifying the note.

With the insertion of the accidentals, the scale over one complete octave now contains 13 notes (including the first note of the next octave) which is shown as follows:

B and the introduction of A flat into scales such as the key of E flat requires other variations in its frequency. As G sharp and A flat are the same note this presents difficulties on a keyboard instrument such as a piano or when string instruments such as the violin (the notes of which are not fixed but continuously variable) are to be played in unison. The same trouble occurs in all of the accidentals and to overcome this a TEMPERED SCALE was introduced.

In the tempered scale (the ratios of which are shown in Table 1) the octave is divided into 12 equal ratios (or semitones) of value 1.059463. Thus the note one higher than C, i.e., C sharp, has a frequency of 1.059463 times the frequency of C and the next note, D (which is one higher than C sharp) has a frequency of 1.059463 times that of C sharp or (1.059463) x 1.059463) =

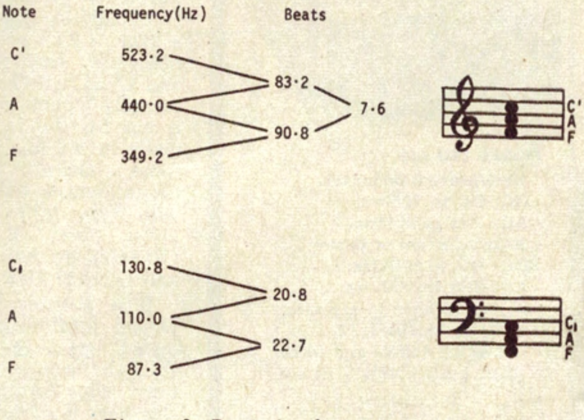

Figure 2. Beats in the major triad.

$$C,\ \left\{\begin{matrix}C\# \\ D\flat\end{matrix}\right\},\ D,\ \left\{\begin{matrix}D\# \\ E\flat\end{matrix}\right\},\ E,\ F,\ \left\{\begin{matrix}F\# \\ G\flat\end{matrix}\right\},\ G,\ \left\{\begin{matrix}G\# \\ A\flat\end{matrix}\right\},\ A,\ \left\{\begin{matrix}A\# \\ B\flat\end{matrix}\right\},\ B,\ C\ .$$

By using the accidentals it is possible to play an eight-note scale over the octave with almost the same frequency ratios regardless of which note is used as a starting note. For example, comparison may be made between the two scales in the "key of C" and the "key of D," as follows:

Key of C		Key of D	
	True ratios		True ratios
C to D	1.125	D to E	1.111
D to E	1.111	E to F#	1.130
E to F	1.067	F# to G	1.061
F to G	1.125	G to A	1.111
G to A	1.111	A to B	1.125
A to B	1.125	B to C#	1.12
B to C	1.071	C# to D	1.071

As may be seen from the above illustration the ratios of the scales in the key of C and the key of D are very nearly the same if accidentals are introduced. This almost solves the problem of the scales but a few minor anomalies still occur with this system. One of these may be seen by considering the ratio required to obtain the note G sharp which occurs in the key of A. In determining the value of G sharp it is found to require a frequency of 1.5625 times the frequency of C, but the frequency of G sharp in the key of E requires a frequency of 1.600 times the frequency of C. Similarly the introduction of G sharp into the key of

1.12246 times the frequency of C. The tempered scale has the added advantage that when a scale is played in **any** key the ratios are exactly the same—unlike the earlier example using the "true scale" when they are only approximately the same.

Few musicians are able to detect the difference between the true scale and the tempered scale, which differ by a maximum of 13 parts per thousand (on the note B).

The frequencies of the tempered scale and the musical notation for the "white" keys for three octaves are shown compared with part of a piano keyboard in figure 1.

The development of the musical scale, although the most fundamental part of music, is far from being the only part. The scale was developed on the principles of consonance and dissonance and the physical meanings of these words have changed with time. "Consonance" occurs when notes sound pleasant whether played harmonically or melodically and "dissonance" occurs when notes sound unpleasant. These words have real meaning only to the observer, and consonances such as the major triad F-A-C' which sound pleasant today would have been treated as dissonances by the ancient Greeks.

Harmonic consonance (the sounding of notes at the same time to produce a pleasant sound) is allied with the production of overtones, or harmonics, by the instrument being used. Generally dissonance occurs between two notes sounded together if a beat frequency between 10 and 50 per second is formed by the notes, either by their fundamentals or any of their overtones. For example, the major triad F-A-C', when sounded near the centre of a piano keyboard, sounds pleasant and is considered the most perfect consonance, but when sounded two octaves lower the result is not pleasant. It can be seen from figure 2 that beats of 90.8, 83.2 and 7.6 per second occur with the fundamentals near the centre of the piano keyboard and these are not in the range of 10-50 per second, but two octaves lower, beats of 22.7 and 20.8 per second occur. The reverse effect can also occur, as in the case of B and C', near the middle of the piano keyboard (B = 493.9Hz and C' = 523.2Hz) the notes are dissonant with a beat of 29.3 per second. These same notes two octaves lower (B=123.5Hz and C'=130.8Hz with a beat of 7.3 per second) would not sound harsh were it not for the fact that a piano, especially in the low notes, is rich in overtones. On the stopped pipes of an organ, which have nearly pure notes with very little overtones, these two lower notes (B=123.5Hz and C'=130.8Hz), do not; whereas B=493.9Hz and C'=523.2Hz do sound dissonant.

It is possible to arrange two notes played together so that the beat produced acts as a third note which will fully harmonise with the two initial notes, thus forming a triad. Although this beat is of lower intensity than the two initial notes it is readily apparent on the diapason stop of an organ and can form a more "mellow" triad than the playing of the three notes together. An example of this is the sounding of C (261.6Hz) and F (174.6Hz) together which produces a difference tone (beat) of 87.0Hz, one octave lower than the F initially played. Similarly playing the major triad C-E-G (261.6Hz, 329.6Hz and 392Hz) produces beats very close to $C_{''}$(65.6Hz) two octaves lower and $C_{'}$ (130.8Hz) one octave lower than the original. If the true scale, rather than the tempered scale, is used the beats in these examples are at the exact frequencies of the lower notes.

The "dissonance" figures of 10-50

NOTE	TRUE SCALE	TEMPERED SCALE	PERCENTAGE DIFFERENCE
C	1·00000	1·00000	0
C#,D♭	....	1·05946	—
D	1·12500	1·12246	-0·3
D#,E♭	1·20000	1·18921	-0·9
E	1·25000	1·25992	+0·8
F	1·33333	1·33484	+0·1
F#,G♭	....	1·41421	—
G	1·50000	1·49831	-0·1
G#,A♭	1·60000	1·58740	-0·8
A	1·66667	1·68179	+0·9
A#,B♭	....	1·78180	—
B	1·87500	1·88775	+1·3
C'	2·00000	2·00000	0

INHARMONIC

Instrument	Overtones as multiples of fundamental frequency		
Drum	1·594	2·136	2·296
Xylophone	2·76	5·43	
Thin Bells	2·928	5·423	8·771

HARMONIC

Instrument		Harmonics in order of intensity
Flute		1, 2, (4,3) .
Violin	E String	1, 3, (2,4)
	A String	1, 5, 2, 4, 3, (6,7,8)
	D String	1, 2, 3, 5, 4, 6, (7,8,9,11,12)
	G String	4, 5, 3, 2, 1, 6, (7,8,9,11,12,13,14)
Clarinet, Saxaphone		8, 9, 10, 3, 1, 7, 11, 12, (5,4)
Oboe		5, 4, 6, (10,37,11,8,2,1)
Piano	Low notes:	Weak fundamental. All harmonics up to 42^{nd}
	Mid notes:	1-10 . All about equal.
	High notes:	2, 1 .
Organ Pipes		1, (2,3)

Tables of true and tempered intervals (above) and overtones produced by musical instruments (right).

per second are only approximate and often depend on the individual. For example the chord of the "augmented eleventh" (as used by George Gershwin in "Rhapsody in Blue") comprising C, E, G, B, D' and F', contains second order beats of 11.2, 31.4 and 46.0 per second but this is still recognised as consonant, and is common in jazz.

Because of the different harmonic content of different types of instruments, music which sounds good on one type of instrument may sound dissonant when played on another, and special "arrangements" of the musical score have to be made to translate from one type of instrument to another.

Harmonic content is also an important consideration when two or more types of instruments are to be played together. The harmonics of one type of instrument can beat with the harmonics of another to produce dissonance even though the tune may be consonant if played individually by each instrument. For this reason the string quartet was favoured by classical composers and the addition of the piano to the string quartet to form a quintet is not nearly as good as the clarinet quintet in which the piano is replaced by the clarinet. The harmonics of the clarinet are sufficiently like those of the strings that the tone combination is homogenous when required (which is not true of the piano quintet), yet is sufficiently different to be readily discernible when required, as may be seen in Brahms' Clarinet Quintet, opus 115.

Some instruments produce "partial tones" which are not harmonics of the fundamental and care must be exercised in the use of these with other instruments. Such instruments are referred to as "inharmonic" and the partial tones produced by some of these (in terms of the fundamental frequency) are listed in Table II. Also listed in Table II are the harmonics produced by some other common instruments.

In addition to the harmonic content, many different instruments play over different frequency ranges. As an example, the piano notes cover the range 27.5Hz to 4.1856KHz; the pipe organ may extend down to 16.35Hz (C_0—a 30ft pipe) and up to 16.5KHz; the piccolo extends from 523.3Hz to 4.699Hz. Most other instruments are limited to cover only part of the range of a piano. The use of high fidelity amplifying equipment with responses up to 40KHz is not necessary to reproduce the fundamental frequencies but to reproduce the harmonics of the instruments, and, although many of these harmonics are outside of the range of detection by the human ear, beats produced by them are detectable and the absence of these beats can often be detected even though the beats, when present, are faint. ■