# INHARMONICITY, NONLINEARITY, AND MUSIC


## N.H. FLETCHER

***Research School of Physical Sciences and Engineering***
***Australian National University,***
***Canberra 0200***



*Neville Fletcher* was educated at New England University College (then part of Sydney University) and at Harvard University. After 20 years as Professor of Physics at the University of New England and five years as a CSIRO Institute Director, he is now a Visiting Fellow at both ANU and ADFA. His research interests were originally in solid-state physics and atmospheric physics, but are now mostly concentrated in acoustics.



*Western music is based upon the sounds of instruments with repeating waveforms and harmonic frequency spectra. While strings and pipes come close to this ideal, they achieve it only because of the influence of great levels of nonlinearity in the sound generating mechanism. Bells and gongs, on the other hand, generally have very inharmonic spectra, and when significant nonlinearity is present it leads to phenomena such as pitch glide, frequency-multiplication cascades, and transition to chaos.*


The music with which we are all familiar is based upon simple principles. Musical instruments produce sounds in which the overtones are exact integer multiples (harmonics) of the fundamental, and two notes sound pleasantly together if their fundamental frequencies are in the ratio of two small integers — 2:1 for an octave, 3:2 for a perfect fifth, and so on. And the reason why this works, we have all been told, is that the vibrational modes of taut strings, and of the air in cylindrical or conical tubes, have frequencies that form a simple integer progression, albeit a progression of odd integers in the case of a cylindrical tube stopped at one end.

If you have examined the subject a little further, you will know that even this simple prescription has its difficulties, for it is not possible to step out a sequence of fifths and get back to an octave of the note you started from, simply from the prime number theorem, which says there are no integers $n$ and $m$ such that $2^n = 3^m$. We return to discuss scales, harmony and music in the final section of this piece, but for the moment we concentrate upon physical phenomena.

Let us return to strings and tubes for a moment. Are the overtones of a real string really exact harmonics of the fundamental? The answer is "no", and the reason is that a real string is not infinitely flexible. For an ideal string, indeed, the frequency $f_n$ of the $n$th mode is $nf_1$, while for a simple stiff rod $f_n = (n + \varepsilon)^2 f_0$ where $\varepsilon$ depends upon the way in which the ends of the rod are fixed and has a value in the range –0.5 to +0.5. Putting these effects together in a fourth-order differential equation, we find that the modes of a real string have frequencies like

$$f_n = nf_1(1 + \beta n^2)^{1/2} \qquad (1)$$

where $\beta$ depends on the elastic modulus, radius and tension of the string. A skilled piano tuner adjusts the string tensions so that there is no beat between the string being tuned and the second mode of the string an octave below it, with the result that all the octaves are slightly stretched above a 2:1 ratio, the total stretch being nearly half a semitone over the compass of the piano.

Much the same thing occurs in the resonance of air columns in tubes, though for a different reason. The acoustic length of a tube, as is well known, is greater than the geometric length by a small additive quantity called the end-correction. For an open end, this correction is about 0.6 times the tube radius. The complication is that the end-correction depends on frequency and decreases towards zero as the frequency is raised, effectively vanishing when the sound wavelength is equal to half the circumference of the tube. The upper modes of an open pipe are therefore higher in frequency than true harmonics of the fundamental, the behaviour being very like that given by equation (1) above. This effect is even more exaggerated for a partly-open pipe termination such as the mouth of an organ pipe or the blowing end of a flute.

## Sustained-Tone Instruments

All this would cause us no concern except for the experimental fact that the sounds made by sustained-tone instruments such as violins, flutes and clarinets have exactly repeating waveforms. This can be demonstrated only over half a minute of so by an actual player, but for an indefinitely long period using a belt-bowing machine or a compressed air source and, as we all know, a repetitive waveform consists of precise harmonics of the fundamental. How can these be generated from modes that are quite significantly inharmonic? The surprising answer is "by nonlinearity!"

Figure 1 shows a system diagram for a sustained-tone instrument. A power source (the player) provides a steady flow of mechanical or pneumatic energy to a sound-generating mechanism (the bow in contact with the string, or the reed-valve on a clarinet) that is closely coupled to a slightly inharmonic resonator. The player controls this resonator by moving fingers along a fingerboard or by opening holes or valves in the resonator, and can also control the bowing speed and force or the air pressure delivered from the power source. There is feedback from each part of this system to its neighbours and to the player. Rather surprisingly, it turns out that the sound power actually generated by the instrument is typically less than 1 percent of the input power, most of which is wasted in viscous and thermal losses along the way.

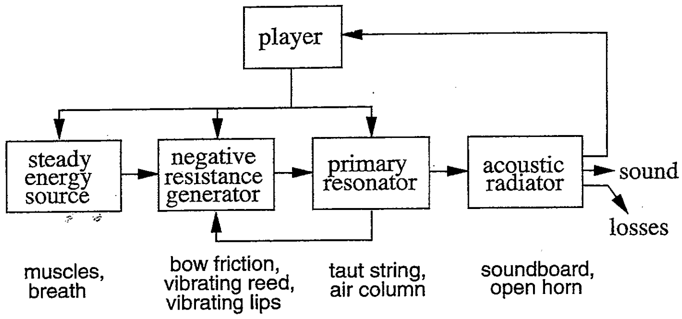


*Figure 1. A musical instrument as a system. A feedback loop between the passive resonator and the negative-resistance generator provides the sounding mechanism, and all is under the control of the player. Physically, the sound is a minor by-product!*

With one exception that we come to later, the resonator is driven at a level sufficiently low that its response is linear, but this is not true of the sound-generating mechanism. Indeed it is the nonlinear behaviour of the sound generator that makes the whole instrument function properly - we might term it "essentially nonlinear."

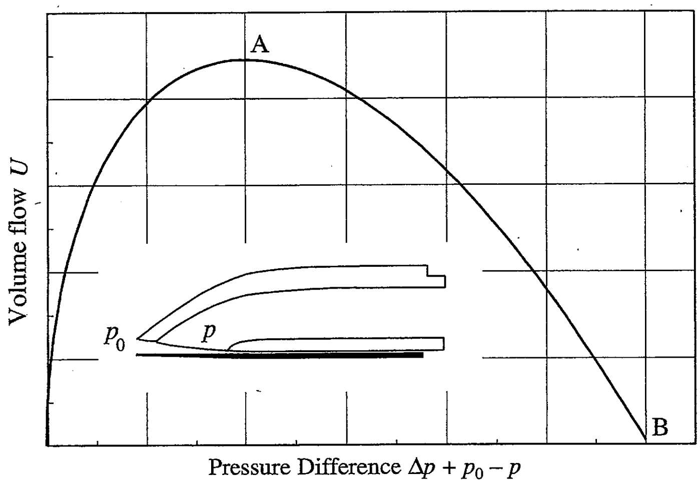


*Figure 2. Quasi-static flow characteristic of a reed generator such as that of the clarinet, shown in inset. The difference between the blowing pressure $p_0$ and the pressure p in the instrument mouthpiece drives the flow, but also tends to close the reed opening. Between the points A and B the acoustic conductance of the reed generator is negative.*

## Wind Instruments

Consider first the reed generator in a clarinet, as shown in Figure 2. The air pressure in the player's mouth causes a flow of air through the reed opening that is proportional to the square root of the pressure drop and directly proportional to the area of the opening between the reed and the mouthpiece. At the same time, this applied pressure tends to close the reed opening, so that the actual volume flow $U$ has the form

$$U = A\Delta p^{1/2}(1 - B\Delta p) \quad (2)$$

where $A$ and $B$ are constants and $\Delta p$ is the difference between the blowing pressure and the pressure in the instrument mouthpiece. The form of this curve is shown in the figure. In acoustic terms, the pressure can be considered the analog of electric potential and the volume flow the analog of current, so the slope of the curve gives the acoustic conductance of the reed valve. The first important thing is that this conductance is negative when the blowing pressure exceeds one-third of the pressure required to completely close the reed, so that for pressures above this threshold and below the closing value the reed valve can act as an acoustic sound generator. The second thing of note is that the slope of the curve varies with pressure, so that this negative driving resistance has a nonlinear behaviour.

Much the same thing happens for the lip-valve of a trumpet player, except that here the mouth pressure tends to force the lips apart instead of closed. It turns out that what is required to make the lips act as a generator is a change in the phase of their motion relative to the acoustic pressure, and this is brought about by having them driven a little above their resonance frequency, while the reed of a clarinet is driven well below its resonance. The details need not concern us here, except to note that, once again, the flow is a nonlinear function of pressure.

Flute-like instruments are rather different and more complex in behaviour, and depend upon wave propagation on an air jet. The important thing from our present viewpoint is that, when this air jet reaches the sharp edge of the mouth-hole in the instrument, it can do not more than blow completely into or completely outside of the lip. Thus, while flow into the instrument is a linear function of acoustic disturbance at small amplitudes, it saturates for large jet displacements in either direction.

## Nonlinear Mode-Locking

Let us briefly examine the behaviour of a linear resonator driven by a nonlinear negative-resistance generator. Suppose the angular frequencies of the resonator modes are $\omega_n$ and that they are not in harmonic relationship. Then the behaviour of the nth mode is governed by an equation of the form

$$x_n'' + kx_n' + \omega_n^2 x_n = g(x_m, x_m') \quad (3)$$

where $k$ is a damping coefficient, a prime implies differentiation with respect to time, and $g$ is a nonlinear function of all the mode amplitudes and velocities. There are potentially contributions to $g$ from all modes $m$, and there is an equation of this form for each of the modes. Notice that the right-hand side of (3) depends solely upon the values of the mode amplitudes $x_m$ and their time derivatives, and that if $g$ is positive it contributes a negative damping that can balance out the term $kx_n'$ on the left-hand side. If this happens, the system breaks into self-excited oscillation.

While this looks complicated, it turns out that it is quite easy to find a solution by writing $x_n = a_n \sin(\omega_n t + \phi_n)$ and assuming that both $a_n$ and $\phi_n$ are slowly varying functions of time. We can then find explicit expressions for both $da_n/dt$ and $d\phi_n/dt$ as time-integrals of $g$ multiplied by $\cos(\omega_n t + \phi_n)$ or $\sin(\omega_n t + \phi_n)$ respectively, and from these we deduce a shift in the actual frequency of mode n to a new value $\omega_n + d\phi_n/dt$. The extent of this frequency shift is proportional to the magnitude of the terms in $g$ that are in-phase with $x_n$ and that have frequency close to $\omega_n$. Similarly, the rate of change of amplitude is proportional to the terms in $g$ that are in-phase with $x_n'$. In systems with nearly harmonic modes, the simplest and most important of these terms arise from quadratic combinations of the form $x_i x_j$ with $\omega_i + \omega_j \approx \omega_n$. It turns out that, provided the nonlinearity is large enough, the system will settle down after an initial transient to a state in which the modified frequencies are in simple integral relationship, giving a repetitive waveform and a harmonic frequency spectrum.

Since the pipes of an organ are included among the instruments just discussed, we conclude that their sound spectrum is precisely harmonic and that the octaves on an organ should be in exactly 2:1 ratio. This is confirmed by examination of the tuning of these instruments. An organ does not go well with a piano!

All this may not sound like a big deal for a simple pipe with nearly harmonic modes, but for most notes on a woodwind instrument there are several finger holes open along the length of the pipe, and measurements show that there are then several modes that are in nothing like harmonic frequency relationship. If the nonlinearity is not large enough, then two or more of these modes may be excited independently, and the nonlinearity will then produce multiple sum and difference frequencies. The result is a non-harmonic chord-like sound termed a multiphonic. These sounds are much loved by those modern composers who are unable to write melody or harmony, but otherwise have only a very limited place in music.

Brief mention should also be made of a completely different nonlinear effect that occurs in instruments such as trumpets and trombones when played very loudly. Measurements and calculations show that in this case the internal sound pressure level can be as high as 175dB, which is equivalent to about 10kPa or one-tenth of normal atmospheric pressure! Under these conditions, sound waves propagating along the narrow cylindrical bore of the instrument develop into shock waves with steepened wavefronts. This effect transfers energy from low to high frequencies and gives the incisive tone quality that we associate with these instruments.

## Bowed-String Instruments

Consider now the bowed string of a violin. The frictional force between the bow and the string is a function of relative velocity, as well as of the normal force which we take to be constant. Static friction is higher than dynamic friction, and there is a smooth decrease in the coefficient of friction with increasing relative velocity. This situation is illustrated in Figure 3. Since the slope of the friction vs velocity curve is uniformly positive, more energy is supplied to the string when moving in the same direction as the bow than is lost when its motion is reversed, so that the frictional contact acts as a negative mechanical resistance. The slope of the curve is, however, not constant, and indeed the force has a catastrophic reversal of sign when the string velocity equals the bow velocity. Mild nonlinearity on the gently sloping part of the curve becomes pathological when the string catches up with the bow!

While we can employ the same mathematical analysis set out above to this case, it is actually easy to see how the extreme

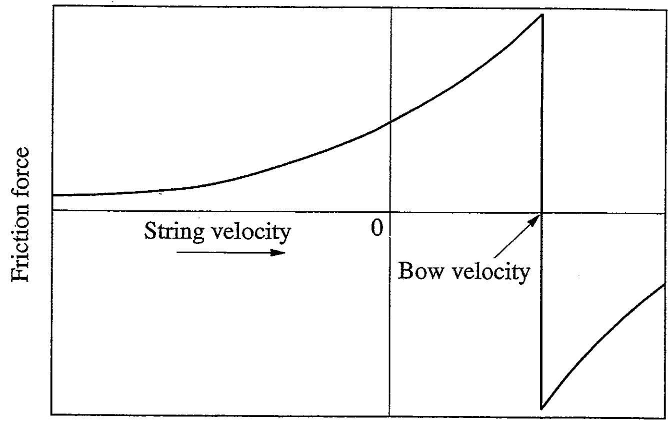


*Figure 3. Behaviour of the frictional force between a moving bow and a string as a function of string velocity. When the string velocity equals the bow velocity, the frictional force is undetermined. When the string is moving more slowly than the bow, the mechanical resistance of the contact is negative and promotes string oscillation. The normal motion is of "stick-slip" type.*

nonlinearity leads to mode-locking and a repetitive waveform. The string simply sticks to the bow for a large part of each oscillation cycle and then slips rapidly to the further extreme of its motion, giving a repetitive waveform.

Both Helmholtz and Raman are among those who investigated bowed strings in detail, and there has been much recent progress using computer analysis of the motion. There is a well-defined range of possible bowing speeds, for a given bowing position along the string and a given normal force, within which the stick-slip mechanism works well. Outside this range, despite the extreme nonlinearity, mode-locking fails to occur, and we experience the excruciating sounds made by some beginning violinists!

From this discussion and that in the previous section, we conclude that the sustained-tone instruments upon which Western music is based have repetitive waveforms and harmonic frequency spectra. In a later section we return to consider the influence that this has had upon the development of music.

## Impulsively Excited Strings

Musical instruments such as the guitar, the harp and the piano are based upon the impulsive excitation of taut strings by a plucking or hammering action. The energy supply is thus disconnected from the string after the initial excitation, and the only nonlinearity that remains is that inherent in the oscillating string itself. If the string tension is low and its elastic modulus high, then the tension can be significantly increased by the stretching involved in an oscillation of large amplitude. This extra tension raises the vibration frequency of all modes and causes the string to "twang" unpleasantly as its frequency drops during the decay of the oscillation. For this reason, among others, guitars usually use nylon strings with rather low Young's modulus, while the steel strings of pianos are tensioned to almost their breaking point.

Apart from this minor nonlinear effect, impulsively excited strings oscillate with a combination of modes that have frequencies given by equation (2) above and amplitudes that are determined by details of the pluck or hammer impact. The nylon strings of a guitar are not appreciably inharmonic, but the steel strings of a piano have stretched octave modes so that the scale of the whole instrument must be slightly stretched, as discussed above. The scale of a harpsichord is very little stretched because the strings have very small diameter.

## Bells, Gongs and Cymbals

Percussion instruments of the drum family need not concern us here, interesting though they are, but we concentrate on those instruments made from metal. Bells are perhaps the most familiar, and produce a more-or-less harmonious sound, depending upon their design. Tradition plays a large part here, and the shapes of Western bells have evolved over many centuries. The vibration amplitude of the bell is so small, and the stiffness of its thick cast metal so great, that nonlinearity is insignificant and the bell sounds its characteristic mode frequencies. As many as six modes frequencies are adjusted during manufacture by turning metal off the inside of the bell on a vertical lathe, and the result is a sequence like 1/2, 1, 6/5, 3/2, 2,..., the nominal pitch being 1 and the octave below that being the "hum" note. The presence of a minor third interval, 6/5, is what gives to a bell its characteristic sound.

The thick-walled metal gongs of the Indonesian gamelan are similarly uninfluenced by nonlinearity, but their mode frequencies are very far from being in harmonic ratio. As we discuss later, this leads to a characteristic gamelan scale that is different from the familiar Western scale but, since the decay time of these gongs is quite short, harmonies are not very much in evidence in their music.

The two small gongs of the Chinese Opera have quite a different sound. They consist of a nearly flat central portion, surrounded by a shallow conical section and terminated by a turned-down rim. They are struck centrally with a padded stick to excite just the fundamental mode. The larger of the pair has an exactly flat vibrating section, and large oscillation causes appreciable radial tension which raises the vibration frequency just as in a string. When the gong is struck vigorously, the frequency therefore starts high and falls back towards its small-amplitude value as the sound decays, the pitch glide being as much as a major third (5/4). The smaller gong, on the other hand, has a central portion that is very slightly domed - only about 1mm over its 100mm diameter - and this is just enough to reverse the behaviour, so that the pitch glides up instead of down. Analysis of the radial stress shows why this occurs.

More interesting from our present perspective is the large Chinese gong or tam-tam, which makes occasional appearance in Western symphony orchestras. It is typically about a metre in diameter and nearly flat, though closer examination shows a raised central hump, a turned-down edge, and several circles of hammered bumps. When struck in the centre with a large padded stick, it gives out an impressive sound that starts as a low-pitched simple tone and develops over several seconds into a high-pitched shimmering gloss. Examination of the development of the spectrum over time, as shown in Figure 4, shows that energy is indeed transferred from a low-pitched centro-

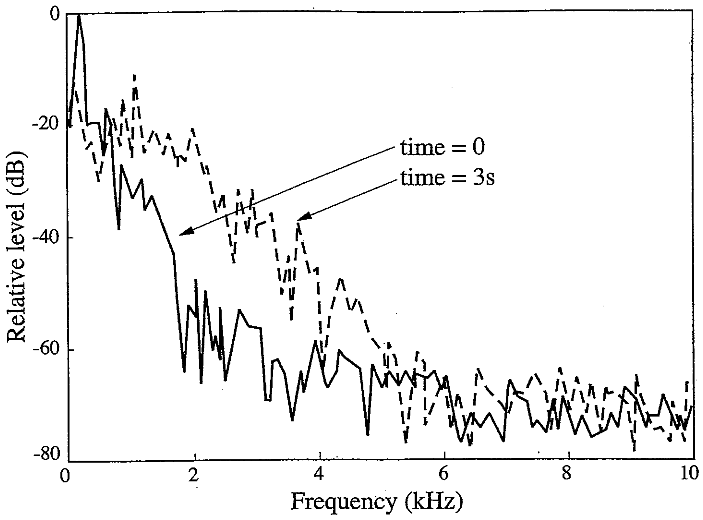


*Figure 4. Sound spectrum of a Chinese tam-tam immediately after being struck, and after the lapse of 3 seconds. Note that the initial low-frequency peak has disappeared, and energy has been transferred from it to high-frequency modes.*

symmetric fundamental mode into a multitude of high-frequency modes distributed towards the edge of the gong.

Again we see nonlinearity at work. Excitation begins with a simple central mode at frequency f. Because the gong is made from thin metal and vigorously excited, it develops significant radial tension forces that oscillate at twice the frequency of the mode concerned, and thus at $2f$. When this tension stress encounters a sharp change in slope, as at the rings of hammered bumps, some of its energy is transferred to the next part of the surface as a transverse vibration at this doubled frequency. In addition, it can interact in the same place with the original mode, giving an excitation at $3f$, and can similarly interact with all other modes present. The result is a cascade of energy to progressively higher frequencies, just as our ears inform us.

The behaviour of cymbals is rather similar, except that they are normally supported in the centre and struck near the edge with a hard stick, thus exciting the higher modes immediately. Perhaps surprisingly, both the tam-tam and the cymbal appear to exhibit chaotic behaviour in their final vibration. Interesting examination of this can be made by exciting the gong in its centre using a sinusoidal shaker with variable amplitude and frequency. Experiments of this kind show regions of period-doubling, behaviour, period multiplication by other factors such as 3 or 5, and ultimately a transition to chaos.

## Implications for Music

Since this is a discussion of physics rather than of music, there is space for only a brief note on the implications of what we have seen. The aim of music is to produce pleasant sounds, and what is pleasant depends on human auditory perception. For two simple sinusoidal sounds played together, frequency differences of a few hertz are not unpleasant, but simply produce a rhythmic fluctuation in loudness. When the frequency difference approaches 20Hz, however, our ears cannot follow the time variation and we hear an unpleasant rough sound. For still greater frequency differences, we hear the two tones individually, and it makes little difference what their relative pitches are. Perhaps surprisingly, musical intervals such as octaves of fifths do not show any greater degree of concord.

When we come to complex sounds, the degree of concordance can be evaluated by summing the effects of all pairs of partial tones interacting with each other. If the initial sounds are harmonic, as for sustained-tone instruments, then the pair will sound well together when their fundamental frequencies are in the ratio of two small number, like 2:1 (octave), 3:2 (fifth), 5:4 (major third), and so on, for then there are no beats between their harmonics and so no roughnesses. Very slightly out-of-tune intervals are tolerated because slow beats do not worry us too much, but differences of only a few hertz in the fundamental frequency translate to differences of 15-30Hz in some of the higher partials and give unpleasant discords. For this reason, the Western musical scale is based on small-integer frequency ratios, and "playing in tune" is critical. As mentioned at the beginning, we run into trouble when we try to devise a fixed scale that will work in more than a single musical key - but that is another story.

If we begin with sounds that are not harmonic, but do have a regular structure, like the sounds of bells or of gamelan gongs, then the same principles of concordance can be used to construct scales in which notes played together on these instruments sound pleasant. Musicians in the societies concerned have evolved these scales over the centuries, but it is now possible to construct them as a scientific exercise for any arbitrary sound spectrum, simply by minimising the dissonances between upper partials..

## Further Reading

There are many places where more information about the matters discussed here can be found. In particular, a recent survey by the present author [1] covers the whole subject in some detail, and a book [2] gives extensive information about the physics of musical instruments. Both these sources give extensive references to the literature. Finally, a recent book by Sethares [3] gives an excellent account of harmonic and inharmonic musical scales, and even goes so far as to construct a scale based on the diffraction spectrum of morphine!

## References


1. N.H. Fletcher "The nonlinear physics of musical instruments," *Rep. Prog. Phys.* **62**, 723-764 (1999).
2. N.H. Fletcher and T.D. Rossing *The Physics of Musical Instruments*, (second edition), Springer-Verlag, New York (1998) [756 pp.]
3. W.A. Sethares *Tuning, Timbre, Spectrum, Scale,* Springer-Verlag, London (1997) [345 pp. plus CD]


## AUSTRALIAN SCIENTIST RECEIVES RECOGNITION FOR SURVEY MEASUREMENT ADVANCES

The International Association of Geodesy (IAG), which governs large-scale geographical measurement, has adopted a significant advance in the accuracy of surveying which was made by an Australian scientist.

Philip Ciddor, an honorary fellow at CSIRO's National Measurement Laboratory, developed new equations while on the Association's working party. His research was carried out in consultation with colleagues from several countries, including Dr RJ Hill of the National Oceanic and Atmospheric Administration in the United States and Professor JM Rueger, Associate Professor of Surveying at the University of NSW.

Modern surveying is carried out by measuring the time taken by light to travel over the distance being measured, a kind of 'optical radar'. While the speed of light is constant in a vacuum, it varies with the composition and conditions of the atmosphere.

The equations commonly used to correct for this variation were based on outmoded data. Errors of one part in ten million were occurring which caused problems in measurement of large distances, for example from the earth to man-made satellites.

The new equations are being applied to geodesy where distances are measured through the atmosphere between terrestrial stations or to satellites.

Last year the International Association of Geodesy (IAG) adopted the new equations by recommending them to the surveying community for use in the most accurate measurements, such as measuring distances to the moon and artificial satellites and in synchronising atomic clocks around the world.

Accuracy in distance measurement is vital for many aspects of geodesy and is being applied to satellite communication, meteorology and mining.

The work's importance lies in it being an international collaboration which can be applied to a range of applications, including the study of slight variations in the orbits of the earth's satellites.

The IAG working party is making further studies, including the effect of varying amounts of carbon dioxide in the atmosphere, the influence of molecular resonance lines of water vapour, and a revision of the analogous equations used for radio-frequency waves.

Dr John Luck, Director of SLR Network Management in AUSLIG (Australian Surveying and Land Information Group) last year instituted the working group of the International Laser Ranging Service to study formulae for atmospheric corrections applied to Satellite Laser Ranging.

Luck says Ciddor's work is of fundamental importance in all aspects of distance measurement using electro-magnetic radiation at radio, optical or laser wavelengths.

An incidental outcome of Ciddor's research was that the uncertainty of the measured speed of light in standard atmospheric conditions set a limit to the measurement of the diameter of silicon spheres used at CSIRO's National Measurement Laboratory in developing an 'atomic kilogram'. This resulted in the apparatus used to measure the spheres being set to operate at a pressure of a few percent of normal conditions, scaling the underlying uncertainty to an acceptable level.

Most recently Ciddor's research has involved a consideration of the validity of some of the basic equations that describe the optical properties of gases, including a classical experiment performed in the National Physical Laboratory in Britain in the 1930s.

*Philip.Ciddor@tip.csiro.au.*