ESSAY

Facing the music

At the heart of any scientific explanation of music is an understanding of how and why it affects us. In the first of a nine-part essay series, **Philip Ball** explores just how far we can hope to achieve a full scientific theory of music.

How strange that music is deemed a phenomenon in need of scientific explanation. We don't, in general, construct objective theories of how great paintings 'work', or great literature, dance or sculpture. We are interested in what is happening at a perceptual level when we experience these arts, but there is always a space in which we leave them to speak for themselves, beyond the reach of cold facts. Yet with music, scientific studies seem to be on the trail of an absolute, all-encompassing explanation that connects neurology with creativity, auditory physiology with acoustic physics. There seems to be a conviction that the composer Arnold Schoenberg was right when he cautioned: "One day the children's children of our psychologists will have deciphered the language of music."

This 'scientification' of music is part of a very old tradition. In antiquity and the Middle Ages music was not an art in the modern sense; it was one of the four sciences of the syllabus called the liberal arts, alongside geometry, arithmetic and astronomy. Scholars studied music to learn about the natural harmony of the world, and performed music was often dismissed as frippery. The early sixth-century Roman philosopher Boethius ranked it as the least of his three classes of 'music', and agreed with Pythagoras that music should ideally be studied while "setting aside the judgement of the ears".

The practice of music does have something of the mathematical about it. Some of the experiments in compositional symmetry, such as the palindromes and mirror reflections of Wolfgang Amadeus Mozart and Joseph Hadyn, are little more than the parlour tricks of an age that delighted in such amusements. But many other musical forms and theories have deeper, more formal organization, from the interwoven fugues of Johann Sebastian Bach to the quasi-mathematical laws of composition developed by Paul Hindemith.

In the final throes of Schoenberg's twelve-note serialism in the 1960s, composers such as Pierre Boulez insisted on a mathematical rigidity that almost sucks their music dry of expression and makes onerous demands of the listener's ability to perceive ordered forms. And in some types of non-Western music, pattern and structure rather than emotion or tone-painting provide the foundations of composition. This is the case in polyrhythmic African drumming, for instance, and the shimmering soundscapes of Javanese gamelan.

Even musicians are uncertain of what kind of art it is they are engaged in, and what, if anything, can be said about it. 'Is there meaning in music?' asked US composer Aaron Copland. He felt there was, but admitted to being unable to articulate what that meaning is.

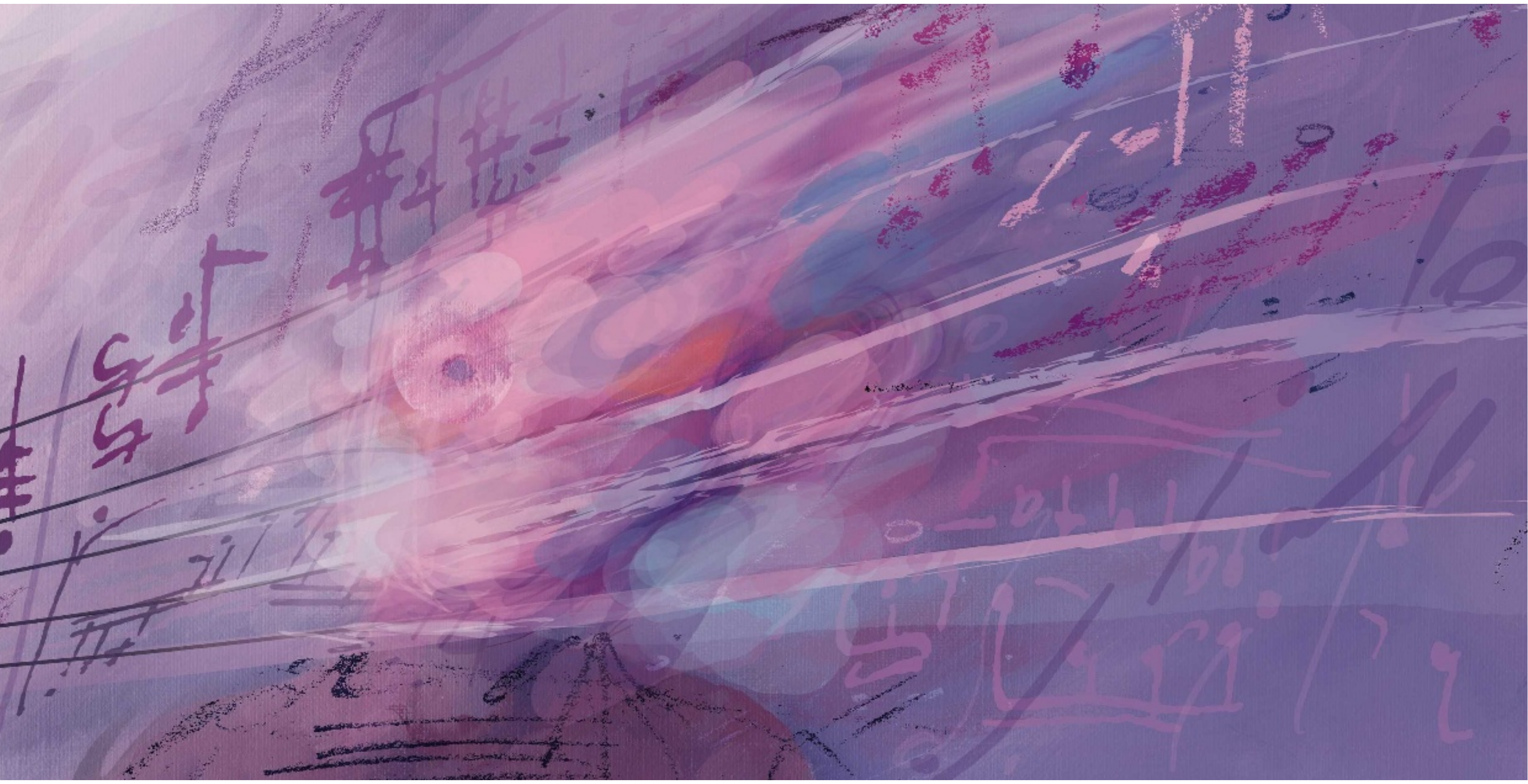
Almost the only thing we can say about music as a cultural phenomenon is that it seems to be universal. Music serves very diverse ends, sometimes with more apparent emphasis on the ritualistic than the hedonistic. Even when it is taken very seriously — in some Native American cultures a ceremony has to be started again if a single note is out of place — anthropologists have often struggled to understand how or to what extent cultures apply intellectual and aesthetic judgements. Sometimes music is a commodity for sale and exchange; elsewhere it is inseparable from dance.

Given this range of what music is and what functions it serves, how can we make sense of it as an acoustic, cognitive, cultural and aesthetic phenomenon? That need not be deemed an entirely hopeless task, but it is not one that science will accomplish alone.

What is to 'understand'?

Trying to understand music is a little like trying to understand biology. The problem is so hard that you have to be reductionist, breaking it down into the building blocks and how they function. Then you find that the original problem has evaporated: in this atomistic view, 'life' or 'music' ceases to be visible at all.

Nonetheless, it makes sense to start with the nucleotides of music: single notes, idealized perhaps to pure tones with a single acoustic frequency. Here, like Pythagoras with his



D. PARKINS

vibrating strings, we ponder the apparent primacy of simple ratios in proportion and frequency, getting an octave from 2:1, a perfect fifth interval from 3:2 and a perfect fourth from 4:3.

Already there are complications. For one thing, musicians do not generally create pure sine tones. Instruments, like any resonating natural object, produce blends of a fundamental tone and harmonics with frequencies that are integer multiples of the fundamental. To us, these complex tones sound like a single note and not a chord.

The brain seems to have evolved simple empirical rules of interpretation that furnish a good guess about the nature of the sounds we are hearing. As nearly all natural sounds are harmonically complex, the brain attributes related tones to a single sound source, combining them into a single auditory object. This is one of the 'Gestalt principles' the brain uses to decode sounds. It seems to be a basic aspect of sound processing — part of the natural auditory conditioning through which music is filtered.

These non-musical acoustic processing principles influence how music is composed. The rules of counterpoint developed in the early Baroque period, for example, use Gestalt grouping mechanisms to prevent separate melody lines from merging[1]. Bach and others also exploited the capacity to generate auditory illusions in the technique called virtual polyphony: splitting a single melodic line into two by means of alternating large pitch jumps.

Some have suggested that, through the series of overtones, conventional Western scales are inherent in a single, harmonically complex note. This idea was formalized by the eighteenth-century French composer Jean-Philippe Rameau, who showed that the major triad (tonic–third–fifth) can be derived from the harmonic series. Other music theorists have claimed that all the notes in the major scale (the white piano notes starting on C, say) originate this way, as higher harmonics of the fundamental tone.

"The practice of music does have something of the mathematical about it."

Similar proposals have been made to explain why certain combinations of notes are consonant, or comfortable to the ear: simple arithmetical relationships between their frequencies are said to confer a favoured status in auditory processing. Such pythagorean reasoning fails to unravel convention and conditioning from physiology. For one thing, several 'consonant' intervals in the conventional Western 'equal-tempered' major scale (the major third in particular) have frequency ratios far from their 'ideal' values. Equal temperament, which divides the octave into twelve equally spaced semitones, is needed because transposing between keys does not preserve the ideal ratios. An E in the key of C does not have the same pitch as an E in the key of A, say. So a scale defined by pythagorean proportions in C will be increasingly out of tune the farther the key moves away from C. Our ears don't seem to object to the adjustments that equal temperament demands.

There is no apparent reason why a scale based on the harmonic series should sound more 'natural' than any other. This is borne out when we look at non-Western music. Some Javanese scales, for instance, pay no heed to intervals with small-integer ratios. Yet scales are not arbitrary. Most have between four and seven notes arranged asymmetrically within the octave, with some pitch steps bigger than others. They also tend to avoid steps much smaller than the Western semitone. The idea that some systems, such as those of Indian classical music, use finer visions, or 'microtonal' scales, has given way to a recognition that these microtones are basically embellishments of a simpler basic scale.

So how do different cultures decide on their musical scales? Cognitive studies on infants and primates offer some evidence that the brain recognizes the octave, and possibly the fifth as 'special'. Indeed, these intervals feature in nearly all musical cultures that use scales. The other notes in a scale seem to be constrained in other ways, too. If there are too many notes per octave, it is hard to tell them apart, and instruments are difficult to tune. There is probably a good reason why most scales have unequal steps, as in the way the Western diatonic scales switch between whole notes and semitones. This asymmetry offers clues about a melody's tonal centre, letting a listener quickly figure out 'where the tune is' in relation to the tonic note.

It is also not obvious how much of the relative consonance and dissonance of different intervals, if any, is a 'natural' phenomenon. Certainly, notions of consonance in Western music have been fluid, defined largely by convention. But there does seem to be a genuine sensory dissonance in some combinations of tones, caused by the unpleasant sensation of beating between two tones that differ only slightly in frequency. Hermann von Helmholtz first did the maths in the nineteenth century and showed that sensory dissonance dips at the intervals corresponding to the Western scale, suggesting that physics does play a part in determining this scale. Yet there is considerable flexibility in the range of tunings that our ears will tolerate. It may even be that acclimatization to a convention can completely override these acoustic facts.

"How interval, melody and harmony act on the emotions is central to our understanding of music."

Why does music move us?

How interval, melody and harmony act on the emotions is central to our understanding of music. But we still have only hints of the mechanisms — or even of the basic phenomena. Take the tonic–minor third interval, for example: this ubiquitous musical element does not feature in the harmonic series at all. Some theorists have tried to argue that 'pulling down' the major third by a semitone to create the minor third creates an intrinsically sad effect. A passing acquaintance with Irish, Spanish or Hungarian folk music demolishes any idea that this association is anything more than a convention. Medieval church music largely ignored major keys, but not because it was trying to be 'sad'; the secular music of medieval troubadours used the major third, but not to be 'happy'. And Aristotle insisted that the Phrygian mode, a kind of scale that includes a minor third, "inspires enthusiasm".

Unlike literature, music cannot convey complex semantic ideas in any universal way. Musical phrases did have particular, conventional 'meanings' in the Western classical era of Mozart and Haydn, and some have tried to argue that instrumental music of other times has specific meanings that can be objectively decoded. Generally such efforts end up as dogmatic assertion, evoking French composer Albert Roussel's poignant image of the musician "alone in the world with his more or less unintelligible language".

Unlike visual art, there is no tangible referent in music that we can point to and contemplate at our leisure. Music works its invisible magic for a moment, then it is gone. It refers to nothing else in the world except by occasional, intentional mimicry. It is hard enough to understand why we can make the slightest sense of these eliding acoustic signals, let alone why they move us to tears and laughter, make us dance or rage. Acculturation alone does not seem sufficient to account for it: we can enjoy the music of other cultures intensely, yet miss all its important allusions and structures.

Someone coming to the psychology of music to understand how those passions are stirred will probably be disappointed. Current studies are often woefully inadequate, even simple-minded. When neurologists examine how and when people classify musical extracts as 'happy' or 'sad', the music lover may reasonably say it traduces the emotive qualities of music, as though we sit through Igor Stravinsky's *The Rite of Spring* either beaming or moping.

But we have to start somewhere. And even these reductive attitudes to musical affect can reveal useful things about how the brain processes musical input. Take the case of a patient with brain damage studied by Isabelle Peretz and her colleagues at the University of Montreal in Canada, for example. This patient had lost all ability to recognize either melody or rhythm but somehow retained the ability to make a 'correct' distinction between happy and sad. She claimed to still enjoy music but lacked what seemed to be the basic neural mechanisms for comprehending it[2]. This finding suggests that affective and cognitive processing of music might involve different neural pathways.

Crudely speaking, there are two camps in the analysis of emotion in music. One says that the emotional content is inherent in musical cues: the choice of mode (major/minor), tempo, timbre, melodic contour and so forth. The other says that it is all about how the music unfolds in time: how a combination of innate and learned responses set up expectations about what the music will do, and emotional tension and release flow from the way these are manipulated, violated and postponed[3]. Much of this is achieved during performance. An expressive performer uses subtle changes of timing, loudness, phrasing and improvized ornamentation to bring out emotional qualities that become barren under the hands of an inexpressive, mechanical player.

What's missing?

All this still falls pitifully short of telling us 'how music works'. It can provide endless narratives about musical events that tend to be somewhat arbitrary and untestable. Do several repeated notes create an expectation of further repetition, or of imminent change? When is a violation of expectation pleasing, and when does it jar, confuse or irritate? In complex polyphonic music, violations can take too many potential directions for us to develop meaningful expectations about them. Beyond a rough sense of shifting tonal centres, I feel no real expectations at all when confronted with the dense, multifaceted slabs of sonic matter that make up, say, Arthur Honegger's first symphony.

We need a better understanding of how the alchemy of music depends on texture. This may be less easily atomized than melody, rhythm and harmony, but it is a more 'musical' characteristic. Here, perhaps, music is working like visual art, just as Mark Rothko's paintings are not mere rectangles of maroon but complex, textured paint surfaces that massage and enliven the brain. Violations of expectation can't account for how that happens.

A lot of music is, after all, less about sequences of notes or beats than about sound sculptures, rich in timbre and composed of interlocking and overlapping layers that function as composite entities. This type of music, whether it is by Messiaen or Ministry, is far ahead of music psychologists still looking for tension peaks in a Mozart sonata.

Far from discouraging scientific studies as futile, these current lacunae should be a stimulus to it. We might start by accepting that it is fruitless to try to define 'music'. We will either leave something out, or include a lot of noise. We might accept too that we should not expect anything like a fully scientific theory of something so fluid. Perhaps there will always be some fundamental limitation in connecting how our brains work with what we do with them. ■

Philip Ball is a consultant editor for *Nature*. His new book *Universe of Stone: Chartres Cathedral and the Triumph of the Medieval Mind* is published this month by Bodley Head.

1. Huron, D. *Music Perception* **19,** 1–64 (2001).
2. Peretz, I., Gagnon, L. & Bouchard, B. *Cognition* **68,** 111–141 (1998).
3. Meyer, L. B. *Emotion and Meaning in Music* (Univ. Chicago Press, Chicago, 1956).

See Editorial, page 134, and http://tinyurl.com/559f2c for further reading.