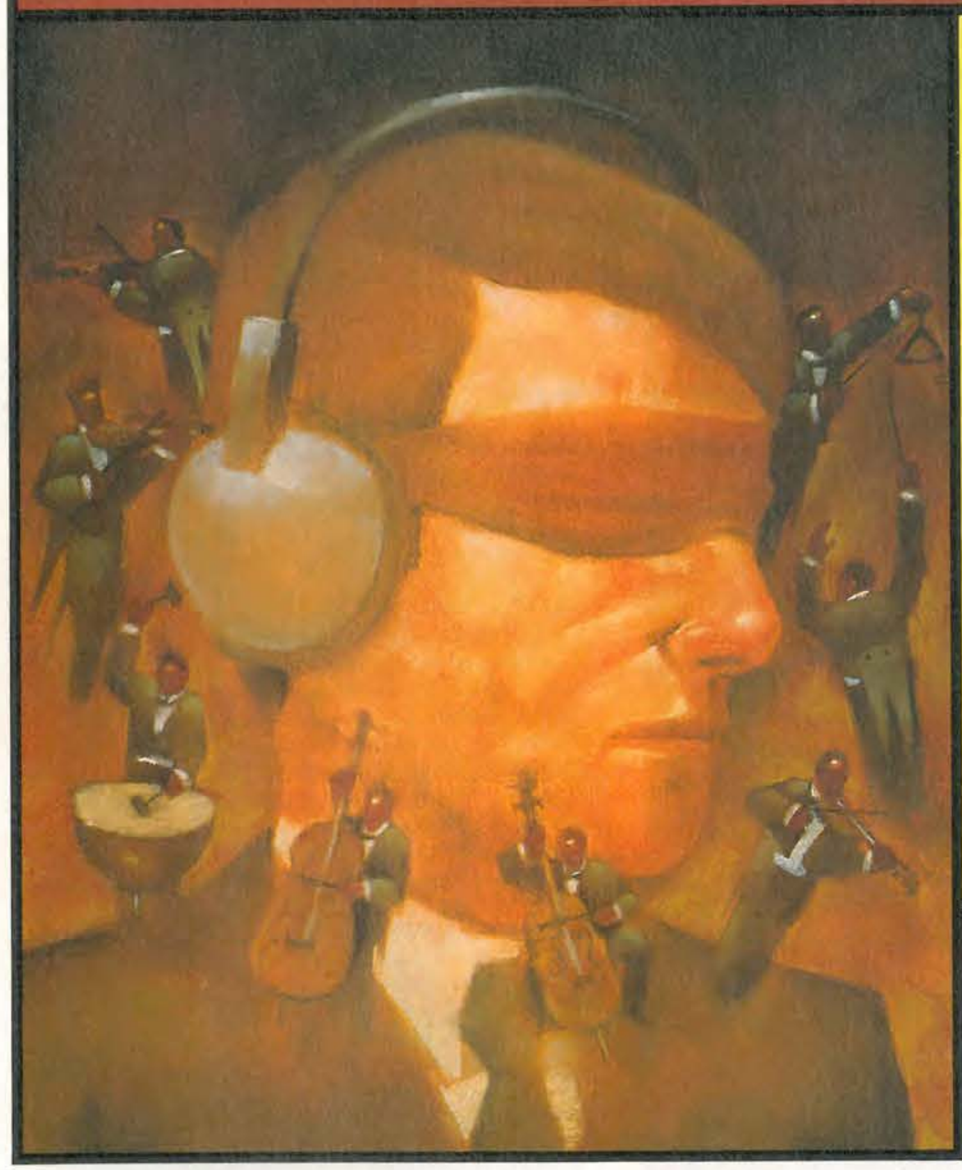
Improve your image



Now that stereo audio designers have all but eliminated crosstalk, Richard Brice explains why they shouldn't have. Richard also builds on Blumlein's work, presenting an improved stereo microphone technique, in this first article based on his new book.

Electronics plays two roles in the art-form we call music. Traditionally, its role is regarded as the conduit and its aspiration, the unimpeachable conveyer of music.

Many articles have appeared in these pages which discuss this role; the quest for the ever-more 'perfect' audio amplifier for example. But I'd like to take a different approach which takes a look at the creative role for electronics. Scrutinised from this perspective, what is in my view a rather moribund debate takes on a new vitality.

A perfect example of this more open-minded approach, concerns improvements to conventional stereophony. This article concentrates on just two techniques which aim to improve conventional stereo recording techniques;

- an unusual stereo microphone arrangement
- a very simple idea for improving two-loudspeaker stereophonic image quality.

A better image

Postponing the microphone technique until later, let's first look at a system which uses an apparently 'distorting' crosstalk signal to improve stereo quality. The irony here illustrates well the enlightened approach mentioned above.

For years, the reduction of left-right crosstalk has been a primary aim of recording system designers and one of the 'triumphs' of digital audio mooted as its elimination. But we now know that there are beneficial effects of controlled left-right channel crosstalk. So, the effects of this 'distortion' have been misunderstood and the deleterious results from its elimination in digital systems similarly mistaken.

It is possible that this has nurtured much of the debate surrounding the subjective differences between analogue and digital recording systems. It also explains many of the apparent 'shortcomings' of digital recording. In order to see why, we have to take a quick look at our ability to localise – i.e. determine the direction of – sounds in space.

How we locate sounds

Consider the situation in **Fig. 1**, where an experimental subject is presented with a source of steady sound located at some distance from the side of the head. The two most important cues the brain uses to determine the direction of a sound are due to the physical nature of sound and its propagation through the atmosphere and around solid objects.

Two reliable observations can be made:

- at all frequencies, there is a delay between the sound reaching the near ear and the further ear,
- at high frequencies, the relative loudness of a sound at the two ears is different since the nearer ear receives a louder signal compared with the remote ear.

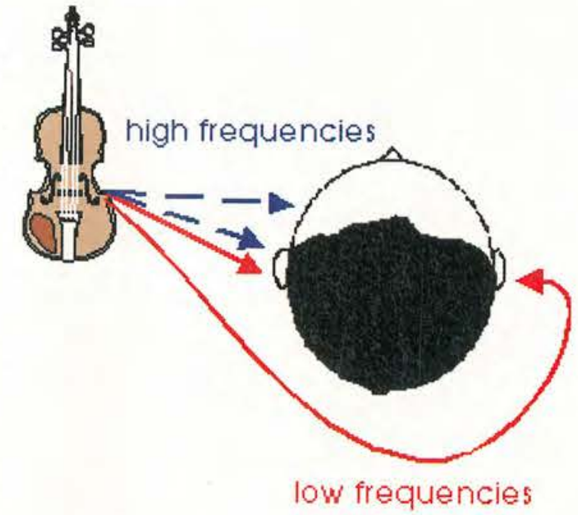


Fig. 1. Two key cues that the brain uses to locate sound sources are the difference in time taken to reach each ear and, at high frequencies, the intensity of sound reaching each ear.

It can be demonstrated that both effects aid the nervous system in its judgement as to the location of a sound source. At high frequencies, the head casts an effective acoustic 'shadow.' This shadow acts like a low-pass filter and attenuates high frequencies arriving at the far ear. In this way, it enables the nervous system to make use of intensity differences to determine direction.

At low frequencies, sound diffracts and bends around the head to reach the far ear virtually unimpeded. So, in the absence of intensity-type directional cues, the nervous system compares the relative delay of the signals at each ear. This effect is termed interaural delay difference.

In the case of steady-state sounds or pure tones, the delay manifests itself as a phase difference between the signals arriving at either ear. But, of course, this phase difference is only useful at low frequencies. Above about 500Hz, the distance between the ears is more than one wavelength.

The idea that sound localisation is based upon interaural time differences at low frequencies and interaural intensity differences at high frequencies has been called 'duplex theory' and it originates with Lord Rayleigh at the turn of the century.

Two-loudspeaker stereophony

Consider the 'classic' stereo arrangement, **Fig. 2**. A moment's thought will probably lead you to some fairly obvious conclusions: if all the sound comes out of the left loudspeaker, the listener will clearly experience the sound from the left. Similarly with the right.

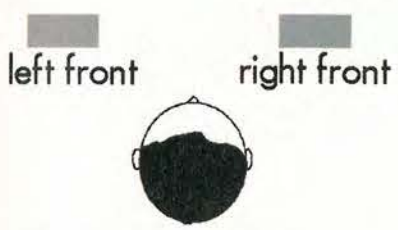


Fig. 2. It is easy to understand how the classic stereo arrangement works with high frequencies, but what about the lower frequencies, where the sound can diffract around the head and reach the other ear?

If both loudspeakers reproduce identical sounds at identical intensity, it is reasonable to assume that the listener's brain will conclude the existence of a 'phantom' sound, coming from directly in front. This is because, in nature, that situation will result in the sound at both ears being identical. And indeed it does; as experiments have confirmed.[1]

More surprisingly, proportionally varied interchannel signal intensities result in a continuum of perceived 'phantom' image positions between the loudspeakers. And the system works as well for high frequency sounds as for low frequency ones. But how?

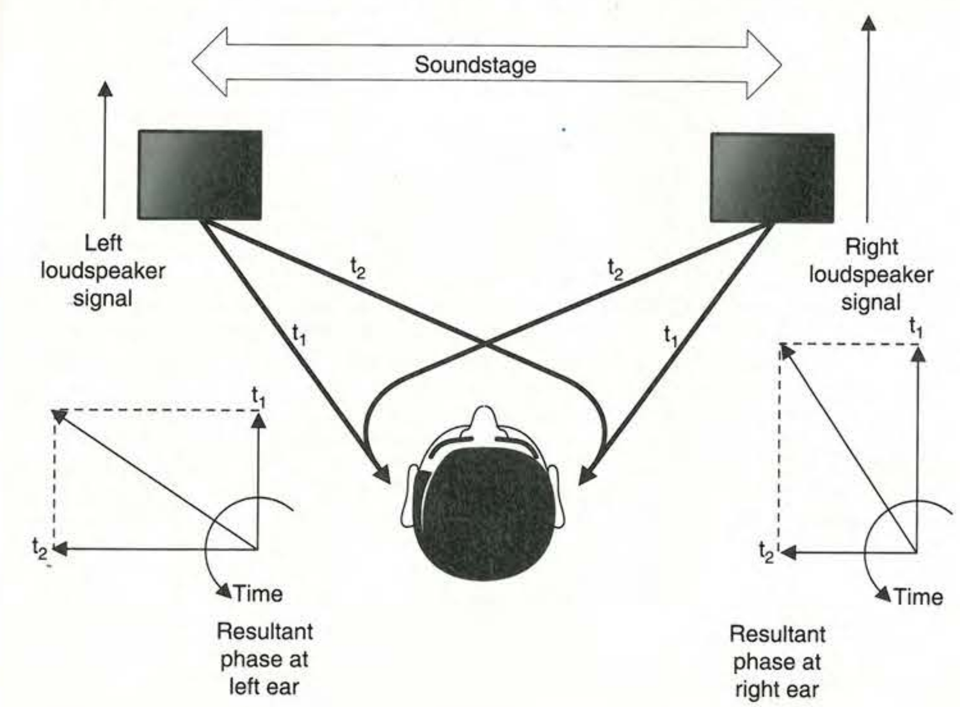
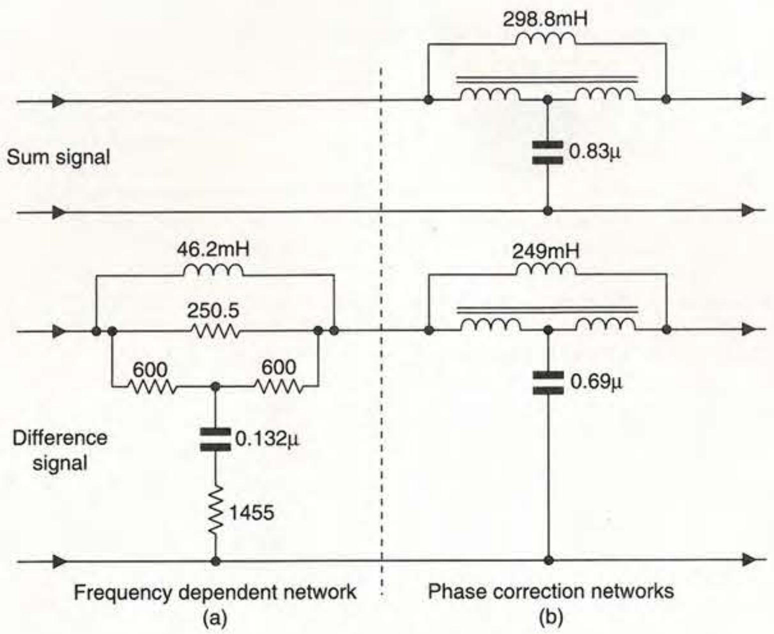


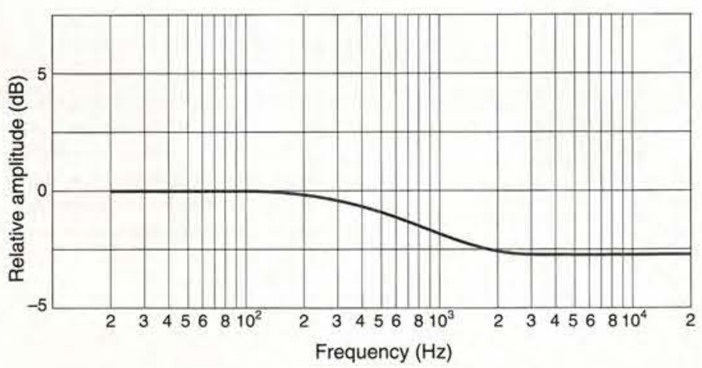
Fig. 3. While it is immediately obvious that left/right channel intensity differences result in high-frequency inter-aural intensity differences when listening to a stereo loudspeaker system, it is far from obvious that these same level differences do, in fact, translate into low-frequency interaural phase differences as well.

While it is fairly obvious that inter-channel intensity differences will reliably result in the appropriate inter-aural intensity differences at high-frequencies, what about at low-frequencies? Here, the sound can diffract around the head and reach the other ear. Where does the necessary low-frequency time-delay component come from?

Well, when two spaced loudspeakers produce identically phased low-frequency sounds at different intensities, the sound-waves from both loudspeakers travel the different distances to both ears and arrive at either ear at different times. **Figure 3** illustrates the principle. The louder signal travels the shorter distance to the right ear and the longer distance to the left ear. But the quieter signal travels the shorter distance to the left ear and the longer distance to the right ear.

The result is that the sounds add vectorially to the same intensity but with a different phase at each ear. Our brain interprets this phase information in terms of interaural delay. This means that stereophonic reproduction from loudspeakers requires only that stereo information be carried by interchannel intensity difference.





Confused literature...

Despite a huge body of confused literature to the contrary, there is no requirement to encode interchannel delay difference. If this were not the case, the pan control, which the sound engineer uses to 'steer' instruments into position in the stereo sound-stage would not the simple potentiometer control shown in **Fig. 4**.

You might think that it is incredibly fortunate that a given interchannel intensity ratio, leading to a particular interaural intensity ratio at high frequencies, causes an exactly appropriate interaural phase difference at low frequencies. It would be incredibly fortunate – if it were true.

For a given interchannel intensity difference, the direction of the perceived auditory event is further from a central point between the loudspeakers when a high-frequency signal is reproduced than when a low frequency is reproduced. Since music is itself a wideband signal, when two-loudspeakers reproduce a stereo image from an interchannel intensity derived stereo music signal, the high frequency components of each instrument or voice will subtend a greater angle at the listening position, than will the low-frequency components.

In fact the stereo image will be 'smeared.' There exists an analogy with chromatic aberration in a lens. This problem was appreciated even in the very early days of research on interchannel intensity related stereophony and, through the years, a number of different solutions have been proposed.

The shuffler

In his original stereo patent application, Blumlein mentioned that it was possible to control the width of a stereo image. He explained that this could be done by matrixing the left and right signal channels into a sum and difference signal pair and controlling the gain of the difference channel prior to re-matrixing back to the normal left and right signals.

Blumlein further suggested that, to alter the stereo image width in a frequency dependent fashion, all that was needed was a filter with the appropriate characteristics to be inserted in this difference channel. After his untimely death, the post-war team working at EMI on a practical stereo system and attempting to cure this frequency dependent 'smearing' of the stereo picture implemented just such an arrangement and introduced a low-pass filter into the difference channel.[2]

Figure 4 is an illustration of their practical Shuffler circuit – as they termed it – and its implementation in the difference channel. Unfortunately this circuit was found to introduce distortion and tonal colouring and was eventually abandoned.

Other derivatives of the Shuffler have appeared using operational amplifier techniques. But the act of matrixing, filtering and re-matrixing is fraught with problems. This is because it is necessary to introduce compensating delays in the sum channel. These must exactly match the frequency dependent delay caused by the filters in the difference channel if comb filter colouration effects are to be avoided.

Moreover, the signal manipulation performed by the Shuffler must be very carefully defined and the EMI team did not have the benefit of more modern psychological

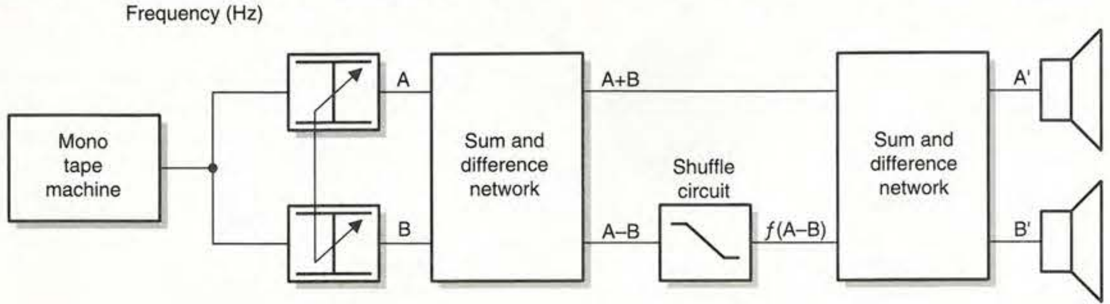


Fig. 4. Practical shuffler circuit and its implementation in the difference channel.

not, as they invariably suppose, due to digital taking something mysteriously away. It is due to the analogue equipment adding beneficial high-frequency crosstalk distortion.

This hf crosstalk technique is exploited in the *Francinstien* range of stereophonic image enhancement systems.

In blind listening tests conducted with 'expert' audiences of musicians and recording engineers using material specially recorded for the experiment, preference for *Francinstien* enhanced signals was overwhelmingly significant.[5]

These tests were scored using classic questionnaire type Likert scales, but room was also given for comments and typical observations included, "more air" and "more space" around each instrument. This is typical of my experience of listeners' reactions to the process.

Commercial units are illustrated in **Fig. 6**. The schematic is given in **Fig. 7**. The *Francinstien* technique is especially useful because of its wide application. It can be used to improve all pre-existing stereo recordings and may thereby be left in-circuit all the time.

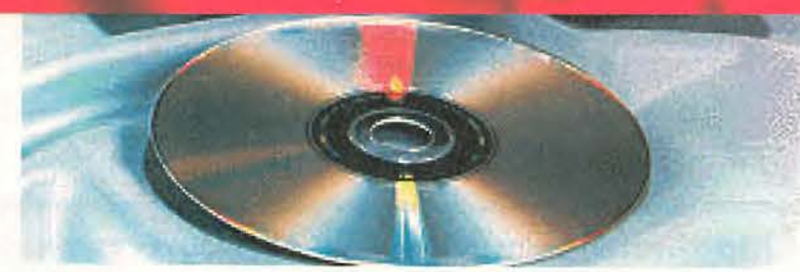


Fig. 6. Stereophonic image enhancement products incorporating beneficial hf crosstalk.

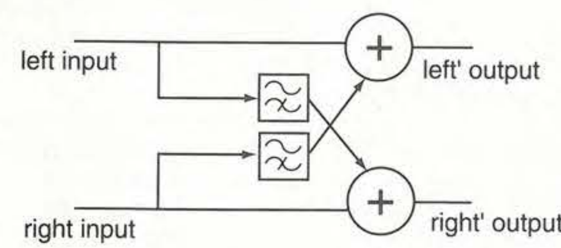


Fig. 7. Outline diagram of the stereo image enhancement circuit used in the products of Fig. 6.

Improved microphone technique

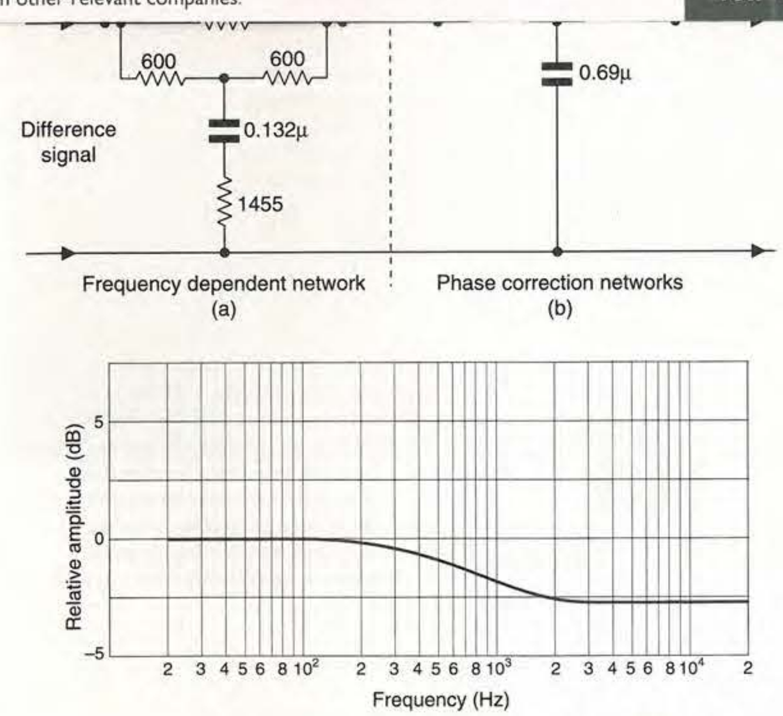
In commercial classical music recordings, it's theoretically possible to 'mike-up' every instrument in an orchestra and then – with electronic panning – create a stereo picture of the orchestra. But this is usually not done for several reasons.

Firstly, the technique is costly and complicated. Secondly often it is simply not practicable. The fact that a multi-miked stereo technique would not work for a recording of the dawn chorus goes without saying. Finally, this 'multi-miked' technique has rarely found favour when it has been tried. Critics, musicians and audiophiles all agree that it fails to provide as faithful a representation of the real orchestral experience.

For these reasons, recordings of real sound fields depend almost exclusively on the application of simple, or 'purist' microphone techniques where the majority of the signal that goes on to the master tape at a classical recording session is derived from just two microphones.

Surprisingly, there are no fixed rules as to how these main microphones should be arranged, although a number of popular deployments have evolved over the years.

The way that the microphones are arranged achieves a certain character of sound. Often it betrays a 'house style'. For instance Deutsche Grammophon currently use two pressure-zone microphones taped to huge sheets of Perspex. Essentially, this arrangement is essentially the same as wide spaced omni-directional microphones much beloved by American recording institutions.



matrixing back to the normal left and right signals.

Blumlein further suggested that, to alter the stereo image width in a frequency dependent fashion, all that was needed was a filter with the appropriate characteristics to be inserted in this difference channel. After his untimely death, the post-war team working at EMI on a practical stereo system and attempting to cure this frequency dependent 'smearing' of the stereo picture implemented just such an arrangement and introduced a low-pass filter into the difference channel.[2]

Figure 4 is an illustration of their practical Shuffler circuit – as they termed it – and its implementation in the difference channel. Unfortunately this circuit was found to introduce distortion and tonal colouring and was eventually abandoned.

Other derivatives of the Shuffler have appeared using operational amplifier techniques. But the act of matrixing, filtering and re-matrixing is fraught with problems. This is because it is necessary to introduce compensating delays in the sum channel. These must exactly match the frequency dependent delay caused by the filters in the difference channel if comb filter colouration effects are to be avoided.

Moreover, the signal manipulation performed by the Shuffler must be very carefully defined and the EMI team did not have the benefit of more modern psychological

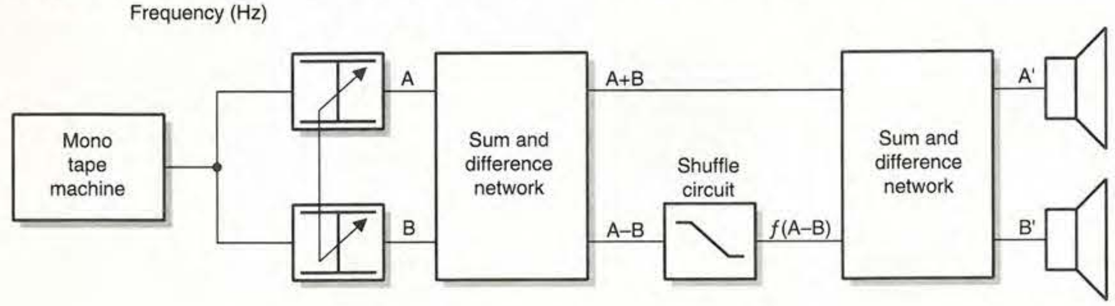


Fig. 4. Practical shuffler circuit and its implementation in the difference channel.

research into the signal manipulation required to bring about stereo-image improvement.

Edeko
Remember, it's a fundamental characteristic of the blurring problem that the brain perceives the high-frequency intensity derived image as generally wider than the low-frequency, delay-derived image. With this in mind Dr Edeko[3] conceived of a way of solving the problem acoustically and therefore of side-stepping the problems which beset electronic solutions.

Edeko suggested a specially designed loudspeaker arrangement, **Fig. 5**. In this arrangement, the angle between the high-frequency loudspeaker drive-units subtended a smaller angle at the listening position than the mid-range drive-units. In turn, the mid-range drivers subtended a smaller angle than the low frequency units.

This device, coupled with precise designs of electrical crossover network, enabled the image width to be manipulated with respect to frequency.

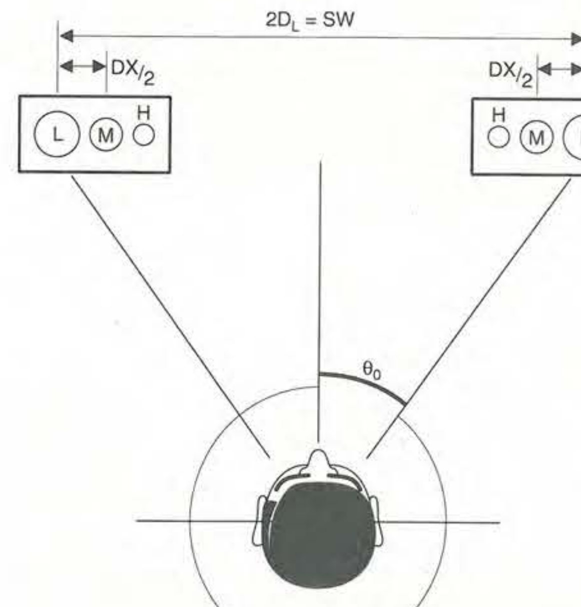


Fig. 5. Edeko (1988) suggested a specially designed loudspeaker arrangement.

Sharper images using crosstalk
There is a much simpler technique which may be used to narrow a stereo image at high frequencies and that is by the application of periodic interchannel crosstalk.[4]

Distortion mechanisms in reproduction from vinyl and other analogue media are predominantly hf crosstalk caused by electrical or mechanical negative reactances. Interestingly, investigations reveal that these distortions may be similar to those required to bring about an improvement in the realism of the reproduced stereo image.

This suggests that there may be something in the hi-fi cognoscenti's preference for vinyl over cd and for many recording musicians' preference for analogue recording over the, apparently better, digital alternative. But the reason is not, as they invariably suppose, due to digital taking something mysteriously away. It is due to the analogue equipment adding beneficial high-frequency crosstalk distortion.

This hf crosstalk technique is exploited in the *Francinstien* range of stereophonic image enhancement systems.

In blind listening tests conducted with 'expert' audiences of musicians and recording engineers using material specially recorded for the experiment, preference for *Francinstien* enhanced signals was overwhelmingly significant.[5]

These tests were scored using classic questionnaire type Likert scales, but room was also given for comments and typical observations included, "more air" and "more space" around each instrument. This is typical of my experience of listeners' reactions to the process.

Commercial units are illustrated in **Fig. 6**. The schematic is given in **Fig. 7**. The *Francinstien* technique is especially useful because of its wide application. It can be used to improve all pre-existing stereo recordings and may thereby be left in-circuit all the time.

Fig. 6. Stereophonic image enhancement products incorporating beneficial hf crosstalk.

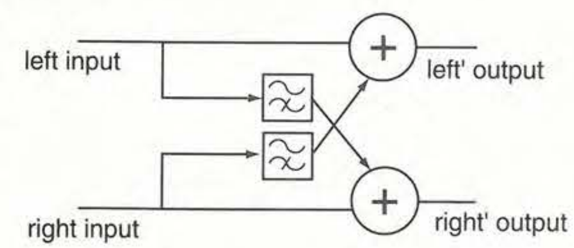


Fig. 7. Outline diagram of the stereo image enhancement circuit used in the products of Fig. 6.

Improved microphone technique
In commercial classical music recordings, it's theoretically possible to 'mike-up' every instrument in an orchestra and then – with electronic panning – create a stereo picture of the orchestra. But this is usually not done for several reasons.

Firstly, the technique is costly and complicated. Secondly often it is simply not practicable. The fact that a multi-miked stereo technique would not work for a recording of the dawn chorus goes without saying. Finally, this 'multi-miked' technique has rarely found favour when it has been tried. Critics, musicians and audiophiles all agree that it fails to provide as faithful a representation of the real orchestral experience.

For these reasons, recordings of real sound fields depend almost exclusively on the application of simple, or 'purist' microphone techniques where the majority of the signal that goes on to the master tape at a classical recording session is derived from just two microphones.

Surprisingly, there are no fixed rules as to how these main microphones should be arranged, although a number of popular deployments have evolved over the years.

The way that the microphones are arranged achieves a certain character of sound. Often it betrays a 'house style'. For instance Deutsche Grammophon currently use two pressure-zone microphones taped to huge sheets of Perspex. Essentially, this arrangement is essentially the same as wide spaced omni-directional microphones much beloved by American recording institutions.

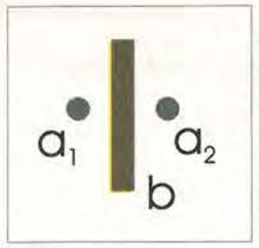


Fig. 8. In Blumlein's patent, two pressure microphones a_1 and a_2 are mounted either side of block of wood b. This wooden baffle provides the high-frequency intensity differences at the microphones in the same way as the human head affects the ears.

British record companies have developed their own arrangements too, while the BBC has stuck almost exclusively to coincident crossed pairs until relatively recently.

This part of the article describes a microphone technique which re-creates the original work by Blumlein with modern signal processing. In controlled recording and listening tests with musician-engineers, the system was preferred to other microphone techniques.[5]

It is possible this is as a result of the system recording extra directional cues over and above those coded in conventional amplitude derived stereophony.[6]

Blumlein's stereo

What makes Blumlein's 1933 patent relating to stereo[7] so important is his originality in realising the principle, as explained above, that interchannel intensity differences alone produce both high-frequency interaural intensity differences and low-frequency interaural phase differences when listening with loudspeakers.

Intriguingly, Blumlein regarded the principle of pan-potted stereo as trivial. It seems that even in 1933, the principle of positioning a pre-recorded single mono sound-signal by means of intensity control was well known.

The technological problem Blumlein set out to solve was how to 'capture' the sound field; so that directional information was encoded solely as intensity difference.

Blumlein noted that a crossed pair of velocity microphones mounted at 45° to the centre of the stereo image has the technological advantage that a pure intensity-derived stereo signal may be obtained from such a configuration without the use of electrical matrixing.

His instinct proved right because this has become one of the standard arrangements for the acquisition of intensity coded stereophony. This is so to such an extent that the configuration has become associated exclusively with his name. It is often referred to as the 'Blumlein-pair', an eponymous, and somewhat incorrect label.

In fact, the greater part of Blumlein's patent is concerned with a primitive 'dummy-head', or quasi-binaural, stereophonic microphone arrangement in which,

"...two pressure microphones a_1 and a_2 [are] mounted on opposite sides of a block of wood or baffle b which serves to provide the high frequency intensity differences at the microphones in the same way as the human head operates upon the ears..." (**Fig. 8**).

Blumlein noted that, when listened to with headphones, the direct output from the microphones produced an excellent stereo effect. But, when replayed through loudspeakers, the stereo effect was very disappointing.

The transformation Blumlein required was the translation of low-frequency, inter-microphone phase differences into interchannel intensity differences. He proposed the following technique:

"The outputs from the two microphones are taken to suitably arranged network circuits which convert the two primary channels into two secondary channels which may be called the summation and difference channels arranged so that the current flowing in the summation channel will represent the mean of the currents flowing in the two original channels, while the current flowing into the difference channel will represent half the difference of the currents in the original channels... Assuming the original currents differ in phase only, the current in the difference channel will be π/2 different in phase from the current in the summation channel. This difference current is passed through two resistances in series between which is a condenser which forms a shunt arm. The voltage across this condenser will be in phase with that in the summation channel. By passing the current in the summation channel through a plain resistive attenuation network comprised of resistances a voltage is obtained which remains in phase with the voltage across the condenser in the difference channel. The voltages are then combined and re-separated by [another] sum and difference process... so as to produce two final channels. The voltage in the first final channel will be the sum of these voltages and the second final channel will be the difference between these voltages. Since these voltages were in phase the two final channels will be in phase but will differ in magnitude."

Blumlein's comments on the perpendicularity of the sum and difference vectors are far from obvious. But consider **Fig. 9**.

A modern practical implementation

The circuit described below is designed so that maximum stereo obliquity is achieved when the inter-microphone delay is 500µs. Other calibrations are possible *mutatis mutandis*. **Table 1** tabulates the phase-angle which 500µs represents at various frequencies.

Consider the 30Hz case. The circuit operates by first deriving the sum and difference of the phasor (vector) quantities derived from the primary left and right channels. ie, let,

$V_1=(0,1)$

and,

$V_2=(\sin 5.4°,\cos 5.4°)=(0.1,0.996)$

$V_{sum}=V_1+V_2=(0.1,1.996)$,

which has a magnitude of 2.

$V_{diff}=V_1-V_2=(-0.1,0.004)$

has a magnitude of 0.1. So, at 30Hz, the difference channel is 20 times, or 26dB, smaller than sum channel's signal.

Now consider the situation at 300Hz, where

$V_2=(\sin 54°,\cos 54°)=(0.81,0.59)$

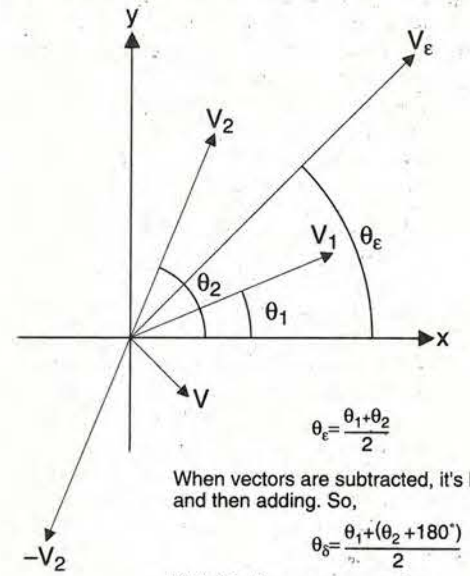
$V_{sum}=(0.81,1.59)$, magnitude=1.78

$V_{diff}=(-0.81,-0.41)$, magnitude=0.9.

At 300Hz the signal is approximately 2 times smaller, i.e. 6dB, compared with the signal in the sum channel.

Now 300Hz is nearly three octaves away from 30Hz and the gain difference is 20dB, demonstrating that the signal in the difference channel rises by 6dB/octave. This confirms Blumlein's statement that, "for a given obliquity of sound the phase difference is approximately proportional to frequency, representing a fixed time delay between sound arriving at the two ears."

Looking now at the circuit diagram for the binaural to summation stereophony transcoder illustrated in **Fig. 10**, consider the role of the integrator circuit implemented around IC_{3a}.



$$\theta_\varepsilon=\frac{\theta_1+\theta_2}{2}$$

When vectors are subtracted, it's like adding 180° and then adding. So,

$$\theta_\delta=\frac{\theta_1+(\theta_2+180°)}{2}$$

Which is the same as saying,

$$\theta_\delta=\frac{\theta_1+\theta_2}{2}+90°$$

Fig. 9. Provided the magnitude of two vectors remains identical, the sum vector and difference are always perpendicular, as shown.

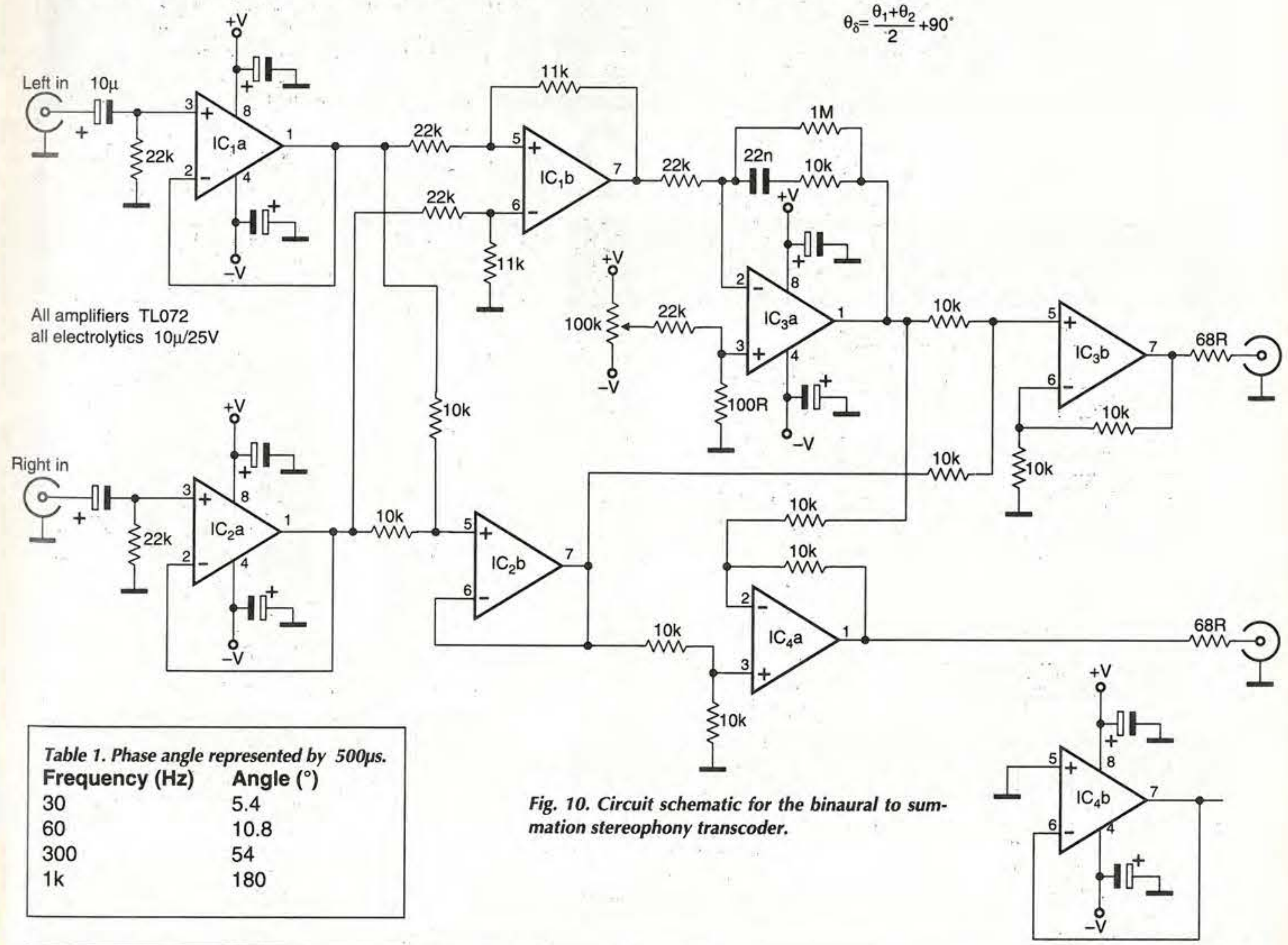


Table 1. Phase angle represented by 500µs.

Frequency (Hz)	Angle (°)
30	5.4
60	10.8
300	54
1k	180

Fig. 10. Circuit schematic for the binaural to summation stereophony transcoder.

The role of this circuit is twofold. It has to rotate the difference phasor by 90° – and thus align it with the axis of the phasor in the sum channel. It also has to provide the gain/frequency characteristic to compensate for the rising characteristic of the signal in the difference channel.

These requirements could be achieved with a simple integrator. But at frequencies above 1000Hz, it is necessary to return the circuit to a straightforward matrix arrangement. This is because the circuit needs to transmit the high frequency differences obtained due to the baffling effect of the block of wood directly into the stereo channels.

The straightforward matrix is implemented by returning the gain and phase characteristic of the integrator-amplifier to 0dB and 0° phase-shift at high frequencies. This is the function of the 10kΩ resistor in series with the 22nF integrator capacitor.

Note that the actual circuit returns to 180° phase shift at high frequencies – ie. not 0°; this is a detail which is compensated for in the following sum and difference arrangement.

Making modifications

Clearly all the above calculations could be made for other microphone spacings. For instance consider the situation in which two omnidirectional mics spaced 2m apart are used as a stereo pick-up arrangement.

With this geometry, 30Hz would produce nearly 22° of phase shift between the two microphones for a 30° obliquity. This would require a magnitude sum phasor with a value of 1.97 and a magnitude difference phasor of 0.39. In other words, an integrator with a gain of 5 is needed.

The gain at high frequency would once again need to fall to unity. At first this seems impossible because it requires the stand-off resistor in the feedback limb to remain 10kΩ as drawn in Fig. 10. A little more thought reveals that the transition region must begin at a commensurately lower frequencies for a widely space microphone system. This is because phase ambiguities of more than 180° will arise at lower frequencies. As a result, all that needs to be scaled is the capacitor, revealing that there is a continuum of possibilities of different microphone spacings and translation circuit values. ■

References

1. Blauert, J. (1983) Spatial Hearing MIT Press
2. Brittain, F.H. and Leakey, D.M. (1956) Two-channel stereophonic sound systems Wireless World Vol 62 No. 5, 206-210
3. Edeko, F.O. Improving stereophonic image sharpness Electronics and Wireless World Vol 94 No. 1623 (Jan 1988)
4. Brice, R. (1998) Music Engineering. Newnes
5. Private correspondence (with F. Rumsey)
6. Brice, R. (1997) Multimedia and Virtual Reality Engineering. Newnes
7. Blumlein, A. British Patent 394,325 June 14th 1933

The *Francinstien* range of stereophonic image enhancement systems mentioned in the article is developed by Perfect Pitch Music Ltd, Tel., +33 1 47 23 54 02.