Toward Reality Equivalence in Spatial Sound Diffusion
Author(s): David G. Malham
Source: *Computer Music Journal*, Winter, 2001, Vol. 25, No. 4 (Winter, 2001), pp. 31-38
Published by: The MIT Press
Stable URL: https://www.jstor.org/stable/3681954

David G. Malham
Department of Music
University of York
Heslington, York YO10 5DD UK
dgm2@york.ac.uk

Toward Reality Equivalence in Spatial Sound Diffusion

The diffusion of sound in space has always been a major topic interest to many composers working with music composed, produced, or performed using electronic means. Despite the rich possibilities inherent in current sound diffusion and spatialization systems, none of the current systems can fully mimic the spatial characteristics of natural sound systems. By examining the nature of our spatial hearing and the history of sound spatialization systems, this article attempts to map possible routes, taking us closer to full "reality equivalent" sound diffusion systems.

Since the remote transmission of sound became possible in the late 19th century, ways have been sought to enable sound systems to deal not only with the frequency and amplitude elements that make up a sound event, but also with the spatiotemporal ones. This is, naturally, of the greatest importance when the sounds being transmitted are musical in nature. Prior to the current era, all musical events were fixed in and wholly part of the acoustic location in which they were being performed. With the breaking of that age-old bond came enormous challenges for the audio engineer in finding ways of restoring the bond correctly after such space/time transmissions. Perhaps more importantly, enormous opportunities opened for composers explore and exploit the relatively little-known territory of space in music.

Edgar Varèse may be regarded as the first composer to grasp the importance of these new opportunities. When he began composing *Intégrales* (1925), he intended to explore something he first became aware of when listening to the scherzo of Beethoven's Seventh symphony:

> Probably because the hall [Salle Pleyel, Paris] happened to be over resonant . . . I became conscious of an entirely new effect produced by this familiar music. I seemed to feel the music detaching itself and projecting itself in space. I became aware of a third dimension in the music. . . . [I]t gives a sense of . . . a journey into space. (Varèse 1936)

The phrase "spatial music" appears to have evolved first in relation to *Intégrales* (Ouellette 1973) and the work itself can be regarded as ushering in the modern era of space in music. Varèse himself has said of the piece:

> *Intégrales* was conceived for a spatial projection. I constructed the work to employ certain acoustical means which did not yet exist, but which I knew could be realized and would be used sooner or later. . . . (Varèse 1959)

Since then, of course, many composers, particularly those involved in electronics-based genres, have ventured into this area, and engineers have been pushed to meet the demands for newer, better, and more interesting sound spatialization and diffusion technologies. As others have remarked in relation to the simulation of "real" or "natural" instrumental timbres, it is perhaps only by learning how to achieve a full simulation of the real and natural that we can appreciate the full extent of what is possible in the synthetic. It is therefore instructive to consider how far we have come today and where we might go in the future, in terms of building acoustic images in space that mimic reality.

Although much progress has been made, for instance with the development of HRTF-based binaural systems, cinema-style systems such as various 5.1 systems, ambisonics (Gerzon 1972,1975; Fellgett 1975), holophonics (Nicol and Emerit 1998), wave-field synthesis (Boone, Verheijen, and Jansen 1996), and hyper-dense transducer array technology (Malham 1999a), we are still a long way from having achieved performance that comes close to that of a natural soundscape. This article examines some of the issues involved as well as the history of our efforts to move toward this, and it uses that background to suggest possible future developments.

Computer Music Journal, 25:4, pp. 31–38, Winter 2001

"Full Reality" Sound

For those people who do not to suffer from significant impairment of either visual or aural senses, the visual sense tends to be regarded as more important than the sense of hearing in the perception of what we call "reality," at least when we are not discussing music. When we think of a beautiful scene, it is the shape of the hills the purple of the heather, and the white dots of sheep to which we pay the most attention. The sound of the wind moving through that heather and the calling of the sheep are in the background. Yet, imagine that scene without the sounds. In that situation, it ceases to be a living, breathing part of nature, losing much of its depth and becoming just another picture.

This loss of "reality" that happens when the sound is lost, or is otherwise inadequately presented, is even more pronounced when the scene is an artificial rendition of reality. It is only necessary to consider what happens when the sound is lost on a film or a television set to see the truth of this assertion. The corollary to this is that we must ensure that the degree of reality equivalence that we can achieve is as high as possible. When dealing with musical events, of course, the importance of the visual element is usually significantly diminished if not absent altogether. It would be easy to say that music would therefore necessarily be more demanding of audio systems in terms of needing "reality-mimicking" performance, but this is not necessarily so. In some musical forms, the important elements may be ones that do not rely on specific acoustical locations or timbres to achieve their compositional effect (for instance, a simple melodic line). Those cases notwithstanding, the more capable our sound diffusion systems are of mimicking reality, the more options composers will have for exploring the use of spatial elements within their music.

For a system to produce reality-mimicking sound, it is necessary for all departures from reality equivalence to be below the relevant thresholds of perception. We need therefore to consider the mechanisms our hearing uses and to consider its capabilities.

Hearing is the only one of our five senses that is truly capable of providing us with fully three-dimensional information about remote (i.e., non-contact) events. We are able to perceive where acoustical sources are located in the space around us, including above and below, as well as whether they are moving or stationary. We can even estimate the distance of the sound-producing object and get some idea of its nature, size, and orientation. For creatures blessed with directional hearing, the three-dimensional sonic environment is one from that they cannot easily escape, because, unlike sight, hearing cannot be "cut off" by anything as simple as night or a blindfold.

Given the importance of spatial hearing to our perception of the world, it is self-evident that we must correctly render the spatial aspects of a sonic image. But to do that, we must first understand what is meant here by "correctly render." In the case of a piece of music, this will vary with the musical intention of the composer. Assuming that the composer requires the image to be a "natural" one, the major variable is the number of people to which the image must be presented simultaneously. For a single individual, rendering the sound scene correctly simply means presenting them with the same set of sonic precepts that they would have gained had the situation been real (i.e., "one-to-one" mapping).

Presenting full reality sound to a multiplicity of people simultaneously, which is the most likely case for most compositions, poses the question of what we mean by "reality." Do we treat each person as a separate observer of the scene, with their own location and independent perceptions (i.e., "many-to-many" mapping), or is it more appropriate to provide them all with the same perceptions simultaneously (i.e., "one-to-many" mapping)? Either is a legitimate option, and the choice of which to use should be made on aesthetic grounds (although economics, and to a lesser extent technological considerations, may well play a part).

Spatial Hearing

To assess what must be done to achieve full reality sound, it is worth reviewing briefly the known mechanisms that humans use for the spatial perception of sounds. This list should not be regarded as exhaustive, because it is by no means clear that our current level of knowledge can be regarded as complete.

Time of Arrival of the Wavefront

One important aspect of spatial hearing is the time of arrival at the ears of the wavefront emitted by a sound object, or, more specifically, the difference in arrival times between our two ears. A sound source anywhere on a plane from front, through above, to back (the median plane) will have its wavefront arrive at the two ears simultaneously. Moving the source away from this line and toward one ear will cause the ears to receive the wavefront one after another. This is known as the *interaural time delay* (ITD). The minimum difference in arrival times between the two ears that can be perceived is dependent on the nature of the sound, varying between 5 μsec and 1.5 msec (Begault 1994).

Level Differences

Another aspect of spatial hearing is the sound level difference between the two ears. Sound from a source to the left of the head, for example, will arrive directly at the left ear, but will be diffracted around the head to get to the right ear. Its amplitude will be less at the right ear than the left, both as a result of the obstructing effect of the head and, to a lesser extent, the extra distance travelled. This is referred to as the *interaural level difference* (ILD) or *interaural intensity difference* (IID).

Frequency Response Differences

Frequency response differences between the two ears also play an important role. The shape of the head and the external part of the ears imparts a frequency-dependent response that varies with sound position and which is, in general, different for each ear. Although this is often referred to as the *Head-Related Transfer Function* (HRTF), strictly speaking HRTFs also include ILDs and ITDs. For this reason, it will be referred to as the *Head-Related Frequency Response* (HRFR). For positions where ILDs or ITDs give ambiguous or nonexistent differences between ear signals (such as median plane signals) or where the listener has little or no hearing in one ear, this is the main positional sensing mechanism where head movement is not involved. It is also one of the two main mechanisms for distinguishing frontal sound sources from rear ones.

Head Movements

Our ability to change the position of our head in such a way that we minimize the ITD, ILD, and the difference between the HRFRs at the two ears is another important aspect of spatial hearing. This minima are, or should be, the points at which we are directly facing toward (or away from) sound sources. This is also the other (and possibly main) mechanism for front-back discrimination, which is accomplished by observing whether interaural differences are increasing or decreasing for a particular direction of head movement.

Distance Perception

The preceding mechanisms largely deal with the angular positions of sound objects. For determining the distance of a sound source, the main cues are the ratio of direct to reverberant sound, early reflections, loudness, high-frequency losses, and air-related distortion.

Ratio of Direct to Reverberant Sound

In a reverberant environment, the energy in the reverberant field stays more or less constant for all combinations of listener-source positioning. This means that, for a given source level, the loudness of the reverberation remains the same, whereas the source loudness decreases with increasing distance. (It is this factor in particular that makes it difficult to place a "sound object" closer than the nearest loudspeaker in a diffusion system.)

Early Reflections

The pattern of directions and delays for the first reflected sounds off surfaces in the environment also contributes to distance perception. These early reflections change in a manner that is dependent on both source and listener positions.

Loudness

The loudness of a sound decreases with further distance owing to the spreading of the wavefront. Note that close sources show larger variations in loudness with listener movement or head turning, because the distance change is a larger fraction of the overall distance to the source than for more distant sources.

High-Frequency Losses

Higher frequencies are progressively attenuated with distance, largely due to absorption by water molecules in the atmosphere.

Air-Related Distortion

For high sound pressure levels, the increase in distortion with distance from the source, which results from the differing speeds of propagation of the positive and negative peaks of the pressure wave (Czerwinski et al. 2000), is a possible extra cue to source distance.

The interpretation of these last four cues is heavily dependent upon acquired knowledge of both the spectra and loudness of the sound source—something that should be considered when using heavily manipulated or wholly artificial sound objects. Loudness as a distance cue is, in particular, known to be of very dubious value, because experiments in anechoic chambers (Neilson 1993) have shown errors of more than 2:1 when subjects were asked to estimate the distance of an unseen sound source.

Other Mechanisms

These are not the only ways that the body perceives sound, and indeed there are other perceptual mechanisms which also provide directional cues. Unfortunately, because of the difficulty of working experimentally on, for instance, chest cavity pickup or bone conduction mechanisms, little work has been published on these means of perception and their directional discrimination capabilities. Instead, because of the relative ease with which headphone-based measurements can be made, almost all the major studies of directional hearing have concentrated on headphone-presented information. Informal experimentation by the author has, however, shown that such non-aural sound perception mechanisms should be investigated more thoroughly and should probably be taken seriously. In particular, there is reason to believe that the chest cavity may play a role in low frequency directional discrimination and that the commonly held belief that we cannot determine the direction of low-frequency sources, where the phase difference between the ears becomes very low, may be true only for headphone presentation. If proven, this would have serious implications for diffusion systems where the bass is presented over a limited number of subwoofers or where replay is over headphones. Additionally, it is worth remembering that the mechanisms of directional hearing described above are components only of the holistic, integrated, directional perception facility that we possess.

History and Assessment of Existing Systems

Reality-mimicking systems can be roughly divided into those which use headphone presentation and those which use loudspeakers. There are, of course, hybrids based on the transaural system developed by Cooper and Bauck (1989). These use *interaural crosstalk cancellation* to allow material intended for headphone presentation to be presented over loudspeakers. Over speakers, sounds that should go only to the left ear also reach the right ear, and vice versa. Unless a cancelling signal for the crosstalk reaching the each ear from the opposite loudspeaker is emitted from its own speaker, the image will be severely disrupted. Even this system is less than perfect, but nevertheless, it is now widely used. There are other hybrid forms, such as the Personal Sound Environment (see http://www.sonics.com/products/pse.html) and combined headphone–loudspeaker systems that are used for some IMAX films. We will return to the latter system later in this article.

Headphone-Based Systems

Headphone-based systems, more commonly known as binaural systems, are those in which the sound is introduced directly into the ear canal via transducers mounted close to or inside the ear, without directly interacting with any of the spatial hearing mechanisms mentioned above. As such, in one sense, the very first excursion into artificial spatial sound—Clément Ader's use of multiple sets of telephone transmitters and receivers in the first multichannel audio transmission system (Askew 1981)—was essentially binaural. Although it is unclear whether the original intention was, in fact, to give some sense of sound source position via the use of spatially separated transmitters, the effect was certainly noted and remarked upon by many visitors to his exhibit at the 1881 Paris Exhibition of Electricity. However, there was no attempt in this system to implement anything that would mimic the cues generated by a human head, for instance by mounting the microphones in the "ears" of a dummy head, so this was not binaural in the fullest sense.

The first real work of any significance in this field was conducted in the 1920s when a binaural, headphone-based system was developed by Dr. Harvey Fletcher and his team at Bell Labs (Sanal 1976). However, the work was not applied at the time, as interest shifted instead to loudspeaker-based systems because they could be used in cinemas (Fox 1982). However, binaural systems implemented either directly over headphones or indirectly over two speakers using crosstalk cancellation have become very popular in recent years. They have achieved widespread acceptance in virtual reality systems (mostly direct presentation), computer gaming (mostly indirect) and home theater applications (indirect). This is largely because they have relatively low resource demand in relation to their performance, at least when compared to systems using large numbers of loudspeakers. This is certainly true for simple recordings made using dummy head techniques, but was not so initially for synthesized soundscapes. When first used, the computing power necessary limited the application of the technique to areas where it use could be justified by the fact that it was obviously the "correct" way of approaching the problem of full three-dimensional spatialization of sound. Computing is now so inexpensive that the power required is no longer a consideration. The perception that binaural is the "correct" system was, and still is, supported by the concept that if we achieve exact duplication of what the ear would hear in a natural situation, we will produce the best reproduction.

Unfortunately, even ignoring the possibility that binaural systems do not act upon all significant sound perception mechanisms, there are major problems with binaural systems. No satisfactory methodology exists for recording real sound fields which allow the use of head movement cues, reducing reality equivalence. The wide variance in individuals makes the use of dummy heads and computational models of the head and body that do not closely match that of a particular individual render the system unreliable at best and completely unusable at worst. Finally, the need to wear a piece of technology, namely headphones, can provide a strong counter-cue, diminishing the degree of reality equivalence.

Crosstalk-cancelled, loudspeaker-based binaural systems may generally be rejected for either one-to-many or many-to-many mappings, because the very small "sweet spot" where it works correctly disrupts the sense of reality for most members of the audience. However, this does have applications in specific situations that are essentially one-to-one mapped, such as computer gaming and televisions.

Loudspeaker-Based Systems

In our search for a high level of reality equivalence, we can immediately dismiss two-channel stereo, satisfying though this might be when the recording is good, because, unless the image is highly distorted, it only presents a small part of the horizontal image with no vertical component, except inadvertently. We can also dismiss the existing cinema-style systems because, despite appearing to cover the whole of the horizontal image, this fails to meet homogeneity criteria and, in any case, no serious attempt is made to deal with the vertical dimension.

There are currently only three systems that even attempt any significant degree of reality equivalence. They are wave-field synthesis (Boone, Verheijen, and Jansen 1996), holophony (Nicol and Emerit 1998), and ambisonics (Gerzon 1972, 1975; Fellgett 1975). The three systems are all related, as Nicol and Emerit have shown (1999), in that they all attempt to recreate the wavefronts in a sound field, based on the Huygens Principle, but using multiple speakers as the "secondary" sources. The Huygens Principle envisages each wavefront as a consisting of point sources of secondary waves which only add up constructively in the direction of propagation of the wave (Richards et al. 1963). Of the three, only ambisonics can achieve an exact, reality-equivalent image, and that only at the central point. The others sacrifice this in favor of being able to provide an acceptable image over a larger area.

In essence, all three attempt to produce appropriate reality-equivalent wavefronts using large numbers of speakers (although ambisonics can successfully use as few as four for correct operation in the horizontal plane only) disposed around the listening area. Holophonics and wave-field synthesis are *volume solutions* that both use large numbers of microphones to sample the acoustic wavefronts over a significant area, whereas ambisonics uses a small array of microphones to sample the wavefronts crossing a single point in the recording space, thus making it a *point solution*. For the volume solutions mentioned above, spatial aliasing results in non-zero errors, which are similar at all points in the listening area and which increase with frequency, owing to the finite spacing of the transducers in both recording and reproduction arrays. As a point solution, ambisonics is immune to such *spatial aliasing* at the sampling point. It does, however, produce increasing frequency-dependent errors as the listener moves away from the central point.

This can be ameliorated by going to a higher order system, for instance a second-order system (Malham 1999b, 2001). Even with a second-order system, however, whereas the performance remains better in most respects than other systems, the errors outside the relatively small central area are sufficient to disqualify the ambisonics approach from being fully reality equivalent, at least when used on its own. Even higher-order systems (Daniel 2000) only ameliorate the problem but soon run into the problem of diminishing returns, because each additional order n requires $2n + 1$ extra channels and $2n + 2$ extra speakers. Each added order, however, provides an ever-diminishing degree of improvement. Nicol and Emerit have proposed a hybrid of ambisonics and holophonics that combines the best features of both, but this still seems some way from reality equivalence.

Another approach, proposed recently by the current author, departs from the idea of attempting to create wavefronts based on the Huygens Principle and uses instead real sound sources. In the hyperdense transducer array (Malham 1999a), a transducer is present for every individually distinguishable sound source position, which, depending on the nature of the sound source and whose research results one accepts, requires transducers placed between one and six degrees apart. A fairly simple calculation then shows that there would be from over a thousand to many tens of thousands of transducers placed on the surface of a notional sphere around the listener. Of course, in practice, they would probably be placed on flat surfaces with delays and amplitude shading used to place them effectively on the spherical shell. The driving force behind this idea was the comment made by many people that any system based on phantom imaging was bound to sound artificial, because the perceptual characteristics of phantom images and real ones diverge, at least in part because their phase velocities are different (Daniel, Rault, and Polack 1998).

With this system and good transducers, many-to-many mapped reality equivalence is potentially achievable for any source position outside the sphere, but for source positions inside the sphere, the system would have to revert to one of the Huygens-based wavefront reconstruction systems. As we have seen, this would result in a departure from reality equivalence except potentially in the central area. Here, the angular separation no longer translates to a linear separation large enough to cause spatial aliasing at the frequencies of interest. Unfortunately, this effectively rules out this interesting approach in terms of reality equivalence when used over a large area (e.g., a concert hall), but still leaves the system as a potentially interesting one. Although the concept of using many thousands of transducers rather than just two, five, or eight, seems somewhat daunting, the technological problems are well within the reach of current technologies. Concepts such as distributing low-frequency sounds out among a group of transducers to increase the air movement available, high-order spherical harmonic encoding of the sounds to reduce the number of required channels, distributed processing using low-cost processors driving only a small number of transducers, self-calibration of the array exploiting the reversible nature of transducers to use them as microphones for position determination and frequency-response correction, the low power needed out of each individual transducer, and other factors all combine to make the system only an expensive—not an impossible—one.

Conclusions

The conclusions to be drawn from the foregoing is that none of the approaches to spatialization mentioned can meet the criteria for true reality equivalence on their own. Although it is possible that new approaches may be developed that will be able to meet this criteria, the best option at present appears to be to develop the hybrid speaker-headphone approach that was mentioned earlier. In this system, sounds near the head are reproduced binaurally, preferably using unobtrusive head-tracked headphones and individualized HRTFs, and all other sounds are handled by one of, or a combination of, the Huygens-based volume solutions using loudspeakers. In this way, both the direct air–ear and the non-air–ear sound perception mechanisms would be properly dealt with for both one-to-one and many-to-many mappings. For one-to-many mappings, however, where the listeners share the same acoustical space, the hybrid approach would be difficult to use, because the speaker-based presentations would not be isolated. It should, however, be usable when the "many" are isolated, as in non-concert situations like broadcasts or recordings.

So where does all this leave the composer? The technologies available are in a considerable state of flux, with the commercially available ones lagging decades behind those in the experimental stage, and none of them really yet being capable of results that are indistinguishable from reality. For composers interested in the concept of articulating a particular performance space as part of their musical approach, there is little need for attempts to produce reality equivalence. For them, the need is for the training of more virtuosic diffusion artists together with the creation of more and better orchestras of loudspeakers. On the other hand, for composers interested in exploring a more abstract, less "performance space–centered" spatial music, the choices really come down to two. First, they can compose for a particular system, whether it be ambisonic or binaural or cinema-style or even, perhaps, an individually specified one designed just for the particular piece. They can then either treat the piece as evanescent, only existing as long as the technological framework within which it was composed exists, or they can hope that the later systems may be able to incorporate the earlier systems. The second choice is to carefully retain the source materials as separate entities, producing versions of the piece appropriate to the limitations of various current systems, but with clear descriptions of the spatial elements and musical structures that should be striven for in future realizations under new technologies. In the end, the choice of which approach to take must, as always, be based firmly on musical considerations, not technological ones.

References

Askew, A. 1981. "The Amazing Clément Ader." *Studio Sound* 23:9–11.

Begault, D. R. 1994. *3-D Sound for Virtual Reality and Multimedia*. Boston, Massachussetts: AP Professional.

Boone, M. M., E. N. G. Verheijen, and G. Jansen. 1996. "Virtual Reality by Sound Reproduction Based on Wave Field Synthesis." Paper presented at the 100th Convention of the Audio Engineering Society, May 1996, Copenhagen, Denmark.

Cooper, D. H., and J. L. Bauck. 1989. "Prospects for Transaural Recording." *Journal of the Audio Engineering Society* 37(1/2):3–19.

Czerwinski, E., et al. 2000. "Propagation Distortion in Sound Systems: Can We Avoid It?" *Journal of the Audio Engineering Society* 48(1/2):30–48.

Daniel, J., J.-B. Rault, and J.-D. Polack. 1998. "Ambisonics Encoding of Other Audio Formats for Multiple Listening Conditions." Paper presented at the 105th Audio Engineering Society Convention, September 1998, San Francisco, California, USA.

Daniel, J. 2000. "Représentation de Champs Acoustiques, Application à la Transmission et à la Reproduction de Scènes Sonores Complexes dans un Contexte Multimédia." Ph.D. diss., Université Paris 6.

Fellgett, P. 1975. "Ambisonics. Part One: General System Description." *Studio Sound* 17(8):20–22, 40.

Fox, B. 1982. "Early Stereo Recording." *Studio Sound* 24(5):36–42.

Gerzon, M. A. 1972 . "Periphony: With-Height Sound Reproduction." *Journal of the Audio Engineering Society* 21(1):2–10.

Gerzon, M. A. 1975. "Ambisonics. Part Two: Studio Techniques." *Studio Sound* 17(8):24, 26, 28–30.

Malham, D. G. 1999a. "Homogeneous and Non-Homogeneous Surround Sound Systems." In *Proceedings of the AES UK Conference: Audio and The Second Century*. New York: Audio Engineering Society, pp. 25–34.

Malham, D. G. 1999b. "Higher-Order Ambisonic Systems for the Spatialisation of Sound." In *Proceedings of the 1999 International Computer Music Conference*. San Francisco, California: International Computer Music Association, pp. 484–487.

Malham, D. 2001. "Higher Order Ambisonics." Available online at http://www.york.ac.uk/inst/mustech/3d_audio/secondor.html.

Nicol, R., and M. Emerit. 1998. "Reproducing 3D-Sound for Video Conferencing: A Comparison Between Holophony and Ambisonic." In *Proceedings of the First COST-G6 Workshop on Digital Audio Effects (DAFX98)*. Barcelona, Spain: The Audio-Visual Institute, Pompeu Fabra University, pp. 17–20.

Nicol, R., and M. Emerit. 1999. "3D-Sound Reproduction over an Extensive Listening Area: A Hybrid Method Derived from Holophony and Ambisonic." Paper presented at the Audio Engineering Society 16th International Conference on Spatial Sound Reproduction, Helsinki, Finland, April 1999.

Nielsen, S. H. 1993. "Auditory Distance Perception in Different Rooms." *Journal of the Audio Engineering Society* 41(10):755–770.

Ouellette, F. 1973. *Edgar Varèse*. London: Calder and Boyars, p. 82.

Richards, J. A., et al. 1963. *Modern University Physics; Fields, Waves and Particles*. Reading, Massachusetts: Addison-Wesley.

Sanal, A. J. 1976. "Looking Backward." *Journal of the Audio Engineering Society* 24(10):832.

Varèse, E. 1936. "Varèse envisions 'Space' Symphonies." *New York Times*, 6 December, sec. 2, p. 7.

Varèse, E. 1959. *Poème électronique* Le Corbusier. Paris: Editions de Minuit, p. 193. Quoted in "Edgard Varèse" by Fernand Oullette, Calder and Boyers, London: 1973, p. 83.