Speaking Aloud

Like microphones, loudspeakers have resisted the leaps in development afforded by the digital revolution. Genelec's **Ilpo Martikainen** traces their development

IN SIMPLE TERMS, a loudspeaker is an electro acoustical transducer with the purpose of converting electrical signals into sound. But as there are no easy practical ways to move air molecules directly with electricity, the conversion is made in two phases. First the electric signal is converted to mechanical movement and then the mechanical movement is converted into an acoustic signal. The most striking feature of this transformation is its poor total efficiency, as only few per cent of the electrical energy is converted into sound and the rest is lost as heat.

The transducer principles having practical meaning are electrodynamic, electrostatic and piezoelectric. The electrodynamic loudspeaker, which is by far the most common, was introduced in its current form in 1925 by CW Rice and EW Kellog1[1]. In their experiments they went through many different ideas and found the voice coil and paper-cone combination to be the most promising solution.

The first transformation is done with a linear electric motor (a voice coil in a magnetic field) and the second with a diaphragm glued to the voice coil. The efficiency in the first step is good as we know from electric motors (it does not matter whether the motor is rotating or linear), but the second is poor due to the great mismatch between high impedance diaphragm and low impedance air. Variations of the electrodynamic theme are numerous; for example constructions such as ribbon, where the diaphragm is a flat, thin ribbon immersed in a magnetic field, and planar where several turns of wire covers the surface of a diaphragm, again immersed in magnetic field.

The electrodynamic transducer principle itself is simple and it looks fairly linear. In practice it has several inherent and well known nonlinearity mechanisms which have been subject to gradual improvements over the decades. At its best it can deliver a very good replica of the original electrical signal.

[1] Rice, CW; Kellog, EW, 'Notes on the Development of a New Type of Hornless Loudspeaker', *Journal of the American Institute of Electrical Engineers*, Vol 44, pp 982-991 (Sept 1925), reprinted in JAES, Vol 30, Nos 7/8 (1982)

In an electrostatic loudspeaker the first transformation is done using signal-driven electrostatic attraction and repulsion forces affecting a film diaphragm. The second transformation is made with the same diaphragm and in this case, because the film is very thin and light, it is well matched into air yielding high mechano-acoustical efficiency. This principle was refined by Peter Walker, who in 1955 introduced a new design[2]. This three-way system had a constantly charged diaphragm between acoustically transparent, rigid electrodes operating in a push-pull arrangement. As the force affecting the charged diaphragm is independent of the diaphragm's position, this speaker is inherently linear, and the acoustic distortion in the mid band can be around 0.05%. On the electrical side, the electrostatic loudspeaker appears as a capacitor, which means that mainly reactive current is flowing into it and heat is generated in the amplifier output stage (instead of voice coil in the case of dynamic loudspeakers). The total efficiency is about the same as for electrodynamic loudspeakers. Most electrostatic systems are dipoles - the radiation pattern is a figure-of-eight - which has the benefit that the system does not directly excite vertical or lateral room modes; the drawback is limited LF response which can be compensated by making the system larger, which again causes drawbacks in terms of directivity; the film has to be divided into smaller sections or to be made curved. The directivity dilemma was elegantly solved - again by PJ Walker - who introduced a new system (the Quad ESL-63) that approaches a curved radiator using concentric ring electrodes fed through a delay line. The diaphragm is thus excited first from centre, and as the outer parts are delayed, a curved wave front results. Regardless of the obvious merits of the electrostatic principle, the maximum output is regarded to be low for studio use. With a few exceptions, electrostatic speakers seem to be limited to high-end consumer applications, at least for the time being.

The piezoelectric loudspeaker uses ceramic materials where an applied electric field causes mechanical movement. The electro mechanic motor can be in a form of a thin wafer on which the radiating diaphragm is glued, but also piezoelectric films have been developed which change either their length or thickness. Most piezoelectric drivers suffer from displacement nonlinearity and hence they have been used for treble only. On the electrical side they appear as capacitors.

The electrodynamic transducer is the workhorse of the audio industry, and the following discussion refers mainly to the development of loudspeaker systems based on them. Also, while the emphasis here is supposed to be in professional applications, most of the research work done in loudspeakers is done as such, without restrictions to professional or consumer applications.

Due to the age of the electrodynamic driver, the general principles of loudspeaker design are by no means new either. In 1940 Harry F Olson published Elements of Acoustical Engineering, a comprehensive book, which - with the updates on its later editions, and the second book Acoustical Engineering, published in 1957 - contains the bulk of the necessary information for good speaker system design. A third invaluable book, Leo L Beranek's Acoustics, was published in 1954.

Understanding and analyzing the system performance has been subject to wide research. Low-frequency reproduction has been simplest to analyze and hence the work has started there. A modern closed-box speaker design was introduced in 1954 by Edgar M Villchur (Acoustic

[2] Walker, PJ, 'Wide-range electrostatic loudspeakers', *Wireless World*, Vol 61, pp 208, 265, 381 (1955)

Research AR-1). In 1958 James K Novak published comprehensive analysis of both closed and vented enclosures. A few years later, in 1961, Arthur N Thiele applied filter transfer functions to bass reflex speaker design, and he was followed in 1972 by Richard H Small who continued his work with both small and large signal analysis of closed, vented and (in 1974) passive radiator boxes, with associated synthesis methods. A third Australian, JE Benson, published several detailed articles in 1971-73 about the same subject. After this work the loudspeaker LF design has been pretty straightforward, and several CAD programs have been created to make the job even easier. Still the LF performance of actual products is far from perfect which only indicates the number of compromises to be done in the actual design.

Crossover design was a topic of many papers in the seventies and eighties, and of those Linkwitz-Riley and D'Appolito are often referred. Also KEF should be mentioned again in this context, as they voiced the now obvious fact that the electrical filter response has to be the difference between the desired acoustical response and the raw response of the driver.

While the calculating power of computers has increased, the interest has moved to modeling of transducers. Diaphragm vibrational modes and response can now be calculated with good accuracy, and the same is true for enclosures and their diffraction. We are approaching a state where the speaker system can be designed and modeled on-screen (and maybe even auditioned) before the actual system has been built.

WHILE THE PAPER cone performs well in LF applications, mid-band coloration prompted new developments when HD Harwood at the BBC started to look for other materials in the sixties. Plastic cones, first made from mixture of polystyrene and neoprene and later from polypropylene, offered much improved performance. Various forms of polypropylene diaphragms are now widely used as either LF-MF or MF units. Most current two-way monitors use some form of polypropylene cone in the LF-MF driver. The number of variations of the 'two-way 160 mm or 200 mm woofer/25 mm tweeter' theme is astonishing. The influence of the BBC work has been significant also beyond UK designed monitors.

While paper, polypropylene and similar materials are inherently soft, there has been another approach to find extremely rigid materials. In 1962 DA Barlow introduced sandwich cones, made from two skins of aluminium glued on expanded polystyrene. The idea was to use the driver only in its piston range and move to a similar but smaller driver at HF. (KEF made roughly similar but flat face woofers, also Philips had one woofer with an expanded polystyrene cone but no skins on the surface). Subsequently, the Japanese have made significant improvements in this respect; Technics introduced their flat honeycomb drivers in the seventies, Pioneer and Yamaha used beryllium. Yamaha's NS-1000M with beryllium MF and HF domes was favored in seventies and eighties and it offered very good performance. The continuing Japanese interest in exotic materials has resulted in interesting solutions such as carbon fiber and boron-titanium composites. In Europe woven Kevlar fibers, either as such or as skin material in a sandwich cone are also used. Although a rigid diaphragm offers a lot of benefits, if the internal damping of the material is too small, very pronounced break-up modes occur, and, although they are outside the intended pass band they usually are audible. When correctly constructed the performance of rigid diaphragms can be clearly better than with the softer counterparts, especially at high levels. The most common application is to use a titanium or an aluminium dome for tweeters.

Increasing the power handling has been a particular American interest. 102 mm voice-coils were used by JBL, Altec and Electro-Voice long before European manufacturers. Heat resistant coil formers, glues and wire insulators were developed for this purpose; JBL developed ideas for ventilation of the coil. All these were necessary for sound-reinforcement applications.

AN IMPORTANT CHANGE of the measurement methodology took place when Richard C Heyser published Time Delay Spectrometry (TDS) in 1967. Not only did it give full transfer function of the speaker, it also made it possible to make measurements in non-anechoic sites, with frequency resolution limited by the size of the space. In TDS the impulse response is calculated from the actual sine wave frequency response through inverse Fourier transform. This allows also distortion measurements. Another option was used later by Laurie Fincham at KEF, who used impulse excitation and a gated measurement window. The impulse response was achieved directly, and frequency response via FFT. Due to the excitation signal distortion measurements are not possible. The impulse method is very fast but often needs averaging to improve SNR because the energy content of the excitation signal is small. This was overcome by using special periodic noise (Maximum Length Sequence) excitation where the crest factor is very low. MLS signals are used by, for example, Doug Rife's MLSSA and Audio Precision.

The current measurements can reveal audible coloration and distortion; also there is basic understanding of what kind of frequency response gives good fidelity ratings in listening. Still there are many unknown areas between the current measurement methods and human perception, and understanding the hearing mechanism is a necessary element for development of meaningful measurement methods. The measurements are made in mono; however, we listen in stereo and multichannel, and the human perception of spatial information is of utmost importance which is currently ignored. These issues have been discussed over the years with no real progress being made; the most recent voice in the desert is that of John Watkinson[3].

Two or even three drivers have been mounted on a single chassis. Jensen had its two-way and three-way (coaxial and tri-axial) speakers, James B Lansing made the Altec 604 in 1943 and it was revitalized in late seventies by UREI. The Altec tweeter had a separate magnet structure and a horn that caused some shadowing and irregularities in the MF response. These were avoided by the Tannoy Dual Concentric design from 1947. The benefit of the Tannoy construction was the use of the woofer cone as extension of the tweeter horn which was machined in the pole piece. Another clever solution was the use of two air gaps in the magnetic structure, one for the woofer and the second for tweeter. Later the use of ferrite magnets instead of Alnico made the single magnet structure obsolete. Although the woofer cone displacement causes a level and frequency dependent modulation to the HF signal, the benefits are also obvious: all frequencies are radiated from the same direction, the structure is symmetrical, the radiation pattern is well controlled and hence the directivity is smooth, all important features for good stereo localization. Tannoy has refined the art of coaxial drivers.

[3] 'Are you listening?', *Studio Sound*, Sept 1999.

With neodymium magnet materials the tweeter size has been reduced and it has become possible to mount a dome tweeter in the centre of the woofer cone; this was first done by KEF in their Uni-Q drivers.

Even with the improved measurement methods and better understanding of their correlation to the hearing experience, by far the most common performance information is the on-axis magnitude response of sound pressure vs frequency. This is insufficient to characterize the total performance because it is not what we hear. The phrase 'near field monitoring' is misleading because acoustically the listener is in far field even with the speakers on top of the console. What we hear is the direct sound plus reverberation of the room and the reverberant part is dictated by the off-axis radiation (power response) of the speaker and the room characteristics. This fact is well known and documented, but still not well understood. Excellent work was done in the mid eighties in Canada by Floyd Toole who documented in listening tests that the best correlation of subjective quality vs measured frequency response was achieved when the frequency response was flat not only on axis, but spatially averaged in a listening window of roughly ±30°(H) x ±15°(V), and when the power response is smooth. Rapid changes in directivity, as is common in two-way designs around the crossover frequency (tweeter is directive, woofer is not) are clearly audible. Floyd Toole is currently at JBL and the fruits of the past work are pleasantly visible and audible in the new LSR monitors.

Avoiding the changes of directivity can take several forms. With horn-loaded tweeters in two-way systems the problem is the narrow radiation pattern of the horn. This was first improved with lenses, multi-cellular and later with radial designs. A major improvement came in the early eighties from JBL who introduced Bi-Radial horns offering roughly 100°x 100° radiation. This matched well with the woofer at the crossover.

At low frequencies the drivers are omni directional because they are small compared to the wavelength. If this state of affairs can be maintained at all frequencies, there will be no change in directivity; the system is almost omni directional. This approach is favored for example by Bill Woodman from ATC who, in the seventies, designed a soft-dome mid range driver. A 75mm radiator is non directional up to about 3 kHz, and switching there to tweeter avoids a rapid change in directivity. However, as this kind of system radiates in all directions, the perceived sound is highly dependant on the room characteristics--what the reverberant field frequency spectrum will be due to room surface absorption as many early reflections affect the stereo imaging. Several companies, like Quested and PMC, have developed systems around the 75mm mid-range dome driver; Stan Kelly designed a still larger, 100mm dome for Boxer, there the crossover frequencies have to be lower.

Another approach was taken by Genelec in 1983 with its Directivity Control Waveguides used in an egg-shaped enclosure. There the aim was also to make the directivity uniform and smooth, but not as wide as possible, instead to limit the radiation angle so that the stray radiation is reduced and the direct to reverberant field ratio improved. Thus the system's performance is less dependent on room characteristics. This result in very accurate imaging which is further supported by the cabinet shape - with no sharp corners there is no diffraction[4]. However, the audience did not like their appearance, and more conventional enclosures had to be used. Fortunately, as the DCW™ prevents radiation at the plane of the

[4] In the sense of the radiation pattern it was not too different from the constant directivity horns; smoother though, and because it was a direct radiator three-way design, the distortion of compression drivers was avoided.

baffle, the diffraction is very small also in rectangular enclosures. In 1988 the DCW™ was scaled up to large flush-mounted systems, and in 1991 downscaled to two-way systems. In the nineties the idea has been adopted by several companies.

UNTIL RECENTLY, the most common way of constructing a speaker system was to use, passive crossover network feeding the drivers and a separate power amplifier. Another solution--later called an active system--is to use one power amplifier for each driver, and a low-level crossover network to feed the power amplifiers. This was tried in the thirties, but it took a long time to gain acceptance. At the end of the sixties Sennheiser marketed a 'no compromise' hi-fi system using this principle, in the early seventies several German companies like Heco and Klein+Hummel, supplying German broadcasters, did the same. Similar ideas emerged in Bulgaria where NIVOX used motional feedback, and Hungary where BEAG later developed cardioid LF and MF radiators for their monitors. Philips was the best-known advocate of motional feedback in the early eighties, and it seems that the active design has been more European than American or Asian specialty. The first American active monitor was Meyer's HD-1 (launched in the late eighties) which used a complicated analogue delay to correct the tweeter position in relation to the woofer. Also some Finnish people started developing active systems in 1976 that led to the founding of Genelec in 1978. It took until the mid nineties before this principle gained industry-wide acceptance, now favored by 38 manufacturers of monitoring speakers.

Today's speaker systems are far from perfect and there is lots of room for improvement. It is still easy to recognize whether the sound comes from a speaker or from a natural source; this of course depends also on microphones. There is a stupid dilemma concerning microphones and loudspeakers: we are all used to lowering prices of electronics, and the same trend is somehow expected from transducers as well. The problem lies in the fact that transducers are fairly mature technology, and their price-performance does not follow the Moore law of processors. The electronics of active speaker systems can and will be made at lower cost, but the price of high quality transducers themselves is likely to go up, and this trend will get stronger in the necessary search for still better transducers. At the same time people are ready to pay high sums for a piece of cable, in a hope that it would make the sound somehow better.

The digitalization of the signal path is soon to extend to speakers as well. This is not to say that attempts would not exist - there are examples of digital crossovers from 80s - but the audible benefits have to be heard, which simply means that this may not have been the most important aspect of improvement. The currently available DSP-based designs do not yet seem to offer any acoustical performance benefits compared to good analog designs. The homework for the acoustical domain should be done first before attempting to correct obvious flaws with clever digital algorithms. Naturally the time will come when the benefits are audible, and this opens doors for the other possibilities, which can be done, but again, this will inevitably increase the cost.

Concerning the multichannel work it seems that we are in similar position as in the very early days of two-channel stereo - enjoying the possibility of five or even more different directions. This seems to be the practice, whether the program is an action film or an opera recording. Adding the LFE channel is a source of numerous misunderstandings, and using a subwoofer often causes cancellation dips in the combined system response. Having five full range speakers capable of handling the LFE signal as well seems to be waste of potential; the other option of using a single subwoofer for all LF and LFE signals and five satellites may not be good either, as there may be psycho-acoustical needs for more than one low frequency source in the room. There is obviously some work to be done before we find the good ways of using all the channels.

In the centre of this level of practical questions we raise issues such as higher sampling rates and hearing in the ultrasonic domain. This is kind of bizarre because simultaneously we are ready to throw a major part of the information away in an attempt to compress the data to more convenient space. I am not against technical development, just the contrary, but the fact-orientated engineering mind asks why not make sure that the current technology works in practice as well as it theoretically should, before entering into new areas without proper knowledge?

Ilpo Martikainen
President, Genelec.