White Paper:
Digital Microphones and AES42

English-language version 2.1
May 11, 2010

Editors:

Claudio Becker-Foss, DirectOut
Stephan Flock, DirectOut
Helmut Jahne, StageTec
Christian Langen, SCHOEPS
Martin Werwein, LAWO
Helmut Wittek, SCHOEPS

Translator: David Satz

Contents

Introduction

Digital microphones for professional audio applications have been on the market for quite a few years, but their use in audio production has been rather limited thus far.

To facilitate the evaluation of this new technology, the authors of the present White Paper aim to offer a view of the subject that is as objective as possible. In the process, not only the advantages but also the possible shortcomings of working with digital microphones will be considered in an even-handed manner.

Most currently available digital microphones for the professional audio market (aside from special solutions, *e.g.* for mobile telephones) are based on the AES42 standard. That standard is an extension of the AES3 standard, and will be a major topic of discussion in this document.

1. The history of digital studio microphones

1.1. The first "digital microphone"

The likely first digital microphone was introduced to the press at the end of the 1980s. It was relatively small and unimpressive-looking, but held out the high promise (for that time) of "CD quality." The prototype, by Ariel, used a 16-bit A/D converter along with a conventional microphone capsule, in keeping with the then-current level of technology. Unfortunately no further details about this microphone are available today.

Figure 1: Ariel Digital Microphone

It was designed to match the newly-released NeXT Computer (1988) which had been brought to market by the newly-founded company of Apple cofounder Steve Jobs. The cube-shaped NeXT workstation had computing power far ahead of its time [1]. In addition to 32-bit graphics, it featured a 25 MHz Motorola DSP56001 chip which could perform real-time digital audio signal processing. The Ariel "digital microphone" was designed for connection to the NeXT's proprietary audio input port.

The microphone's inventor, Jon Paul, considered the idea of building an A/D converter directly into the housing of a microphone to be innovative enough that he filed immediately for a patent on it (U.S. Patent 5,051,799 "Digital Output Transducer" (1989) [2]). During its lifetime this patent made life difficult for other microphone manufacturers who wanted to sell digital microphones in the USA, since its very broad claims would be infringed upon by nearly any attempt to construct a digital microphone with an analog microphone capsule followed by a converter.

Unfortunately, this idea – like that of the NeXT cube itself – was so revolutionary that neither product became a commercial success. In 1993, as the victory train of standard PC hardware and Windows software became unstoppable, the microphone disappeared from the market along with the NeXT hardware.

1.2. The first digital studio microphone

Near the end of 1995 Kai Konrath, who was then studying audio and visual engineering at the University of Applied Science in Düsseldorf, began his thesis project in the development of a digital studio microphone at Beyerdynamic. The cardioid MC 834 was initially made digital by the inclusion of a 20-bit A/D converter (Crystal CS5390). A 7-pin XLR connector carried the bipolar 5-Volt power supply and word clock sync along with a balanced AES3 audio data stream. Since the A/D converter had a narrower dynamic range than the capsule, remote-controllable internal preamplification was provided for the capsule's signal.

At the beginning of 1996, however, soon after the first prototypes were completed, Beyerdynamic came into a close working relationship with StageTec. The StageTec "TrueMatch" gain-ranging system was licensed for use in the MCD 100, thus allowing its dynamic range to increase to 115 dB (A-weighted, RMS).

Through the use of a 6–12V phantom powering arrangement (150 mA), and by giving up the word clock synchronization and switchable preamplification features, it became possible to use XLR-3 connectors and available microphone cables.

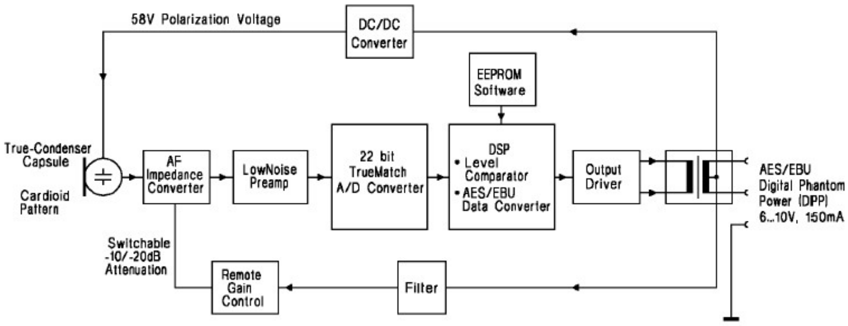

Figure 2: Block diagram of the Beyerdynamic MCD 100 digital microphone with internal analog preamplifier

Figure shows the block diagram of the MCD 100. The “TrueMatch” system was implemented in DSP—the first use of such technology in a microphone. Since no standardized interface was available for connecting this microphone to other

Figure 3: Beyerdynamic MCD 100 with power supplies MCD 50, MCD 100 and MCD 200

equipment, Beyerdynamic developed various powering units with AES3 outputs and remote controllable attenuation or amplification.

The polarization voltage for the condenser microphone capsule was derived from the 6–10 V / 150 mA phantom powering. An analog preamplifier raised the capsule’s inherent noise enough to cover (and thus mask) the noise level of the A/D-converter. As a consequence, very high sound pressure levels could overload the A/D converter. To reduce the likelihood of this, a remotely-switchable pre-attenuation was provided. It was set via control signals that modulated the phantom powering.

Though the goal of digitizing the full dynamic range of a condenser microphone capsule could not be met completely, the MCD 100 still represents the first usable digital microphone. It awakened the interest in an international standard interface for digital microphones. Despite its modest commercial success, the engagement of its manufacturer caused many ideas that had been realized in the MCD 100 (and its omnidirectional counterpart, the MCD 101) to flow into the later AES42-2001 standard. The presentation of the MCD 100 at the 20th Tonmeister Convention (1998) unleashed the interest of all the German microphone manufacturers in working out a standard for a digital interface under the aegis of the German Electrotechnical Commission (DKE).

At the 1999 AES Convention, Milab introduced the DM 1001 digital microphone. In contrast to the MCD 100 and MCD 101, its directional characteristic was adjustable through the use of two pressure-gradient transducers; signals from the dual-membrane capsule were digitized by a stereo A/D converter chip without the use of gain ranging. This limited the converter to a dynamic range of only 114 dB (A, RMS). The microphone's directional pattern, internal preamplification and filter functions could be remotely controlled via Windows-based software running on a PC.

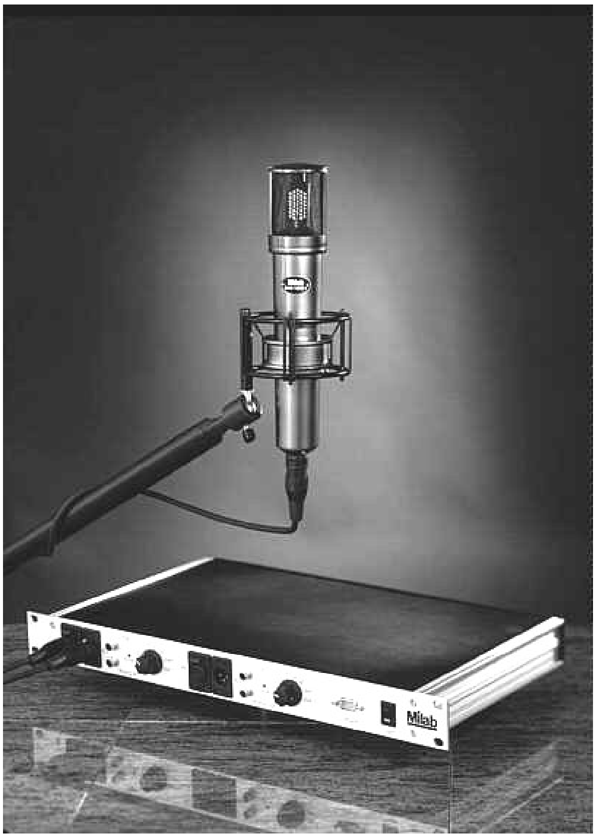

Figure 4: Milab DM 1001 B Digital Microphone

1.3. The first commercial success of digital microphone technology

The first larger application of this digital microphone technology occurred in public address applications rather than in the studio realm. In 1999 the Reichstag building in Berlin was refitted as a meeting place for the German Parliament and was equipped with new digital audio technology. Beyerdynamic's MCD 800 series (Figure 5) was specially developed for use as the delegates' microphone, following a much simpler design concept than that of the MCD 100. A simpler A/D converter (Crystal CS5360) with a dynamic range slightly greater than 100 dB (A, RMS) was used. The lack of gain ranging necessitated a high degree of internal analog preamplification.

Figure 5: Beyerdynamic MCD 800 Series

For the speaker's lectern it was planned to use the "cardioid plane microphone" (KEM) [3], which had been developed at the Institute for Radio Technology (IRT) and was being built by Microtech Gefell under license. Since exclusively digital microphones were to be used in the Reichstag, two KEM had to be "digitized," *i.e.* provided with TrueMatch A/D converters by StageTec.

Apart from the 220 delegates' microphones in the Reichstag (variants of the MCD 803) and the two KEM 970, presumably very few of these models are otherwise in use, since the completely digital setup of the Reichstag has not been imitated elsewhere. Nor did this prominent example help digital microphone technology to achieve a hoped-for market breakthrough, even though the setup performs its daily service in a manner that leaves no room for complaint.

1.4. The path toward a standard digital microphone interface

At the European AES Convention in March, 1997

Steven Harris, who was then with Crystal Semiconductors, visited all the microphone manufacturers to encourage them to work together on an AES standard for an interface for digital microphones. In the same year he introduced the first results of these efforts, together with Xue-Mei Gong and David Josephson [4].

The proposed microphone interface ([4], p. 10) showed strong commonality with the initial designs of Beyerdynamic, although it was kept far more general. All operating voltages would be obtained through the cable. A/D conversion and optional signal processing could be remote-controlled. Synchronization capability was allowed as an option, although this was not specified in any detail. Even the powering showed parallels with the MCD 100: A phantom supply voltage would be coupled into the balanced line carrying the AES3 signal and extracted again within the microphone, with center-tapped transformers on both ends. Control signals would modulate this supply voltage.

After the Tonmeister convention of 1998, at the initiative of the German microphone manufacturers, Working Group AK 742.6.1 of the UK 742.6 Subcommittee (microphones and headphones) of the German Electrotechnical Commission (DKE) was founded. Early in 1999, three sessions of this working group worked out the essential elements of AES42-2001. This was then presented as a proposal to the AES in time for its European convention in May, 1999.

Nonetheless almost two more years passed before the first version of AES42-2001 was finalized [5]. During this time the DKE working group's proposal was augmented with a capability for external synchronization. It was intended that the commonly employed XLR-3 connector be used so as to allow further use of existing installations. But the this widely-used connector became a matter of discussion which greatly delayed finalization of the AES42 standard. The fear was expressed, particularly by those from the USA, that unschooled personnel could destroy the output circuitry of analog equipment by connecting it to powered AES42 inputs. For this reason a new, mechanically coded "XLD" (=XLR for digital) connector was advocated to prevent the very compatibility with XLR-3 connectors which had been provided for in the original draft standard.

1.5. More recent developments

Since 1999, StageTec has been the only manufacturer of mixing desks to offer an interface card for digital microphones in its product line. More recent developments by other manufacturers concern PC sound cards as well as portable recorders (see section 4.1).

Directly following the publication of the AES-2001 standard, Neumann introduced its digital "Solution-D" microphone system in 2001. Initially this consisted of the D-01 microphone and an interface unit that supported the newly-defined standard. By connecting the interface to a Windows-based PC, the microphone parameters defined in AES42 could be remote-controlled via a USB/RS422 port. The Solution-D microphone system is Mode 2-capable (see section 3.2.4).

Small-diaphragm microphones from Neumann and Schoeps arrived in the succeeding years. Neumann extended their Solution-D system with the KM-D microphones. With the CMD 2 amplifier module (Mode 1-capable; see section 3.2.4), Schoeps provided their modular "Colette" microphone series with the ability to operate digitally.

The current revision of the standard dates from 2006 (AES42-2006) [12], and forms the basis for the present document.

2. What is a digital microphone?

2.1. “Digital” or “digitalized”?

Traditional microphones use a membrane and convert its excursion or velocity into an electrical signal.

Typical examples include the dynamic microphone, in which a coil of wire is firmly coupled to the membrane while being able to move within a magnetic field; sound energy thus results in an analog voltage being emitted by the coil. A condenser microphone functions quite similarly, but in this case the membrane forms one plate of a capacitor which varies its capacitance with motion of the membrane; sound energy thus causes variations in capacitance which are transformed into an analog signal.

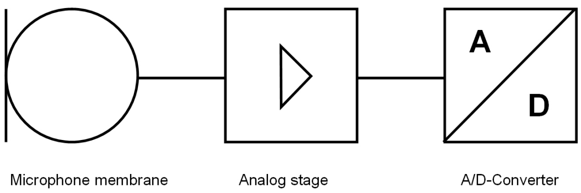

Figure 6: Schema of a contemporary “digitized“ microphone: Membrane, analog gain stage, A/D converter

Now if one gets the idea to build digital microphones the first thought is always to leave the analog world behind completely, and to produce a digital signal as early in the signal path as possible. The ideal microphone would keep only the retaining ring of the membrane, and would count the number of air molecules that pass through the ring in one direction versus those that pass through in the other direction. The difference would be the ideal output signal for a digital microphone. Unfortunately no one has yet succeeded in making a studio-quality microphone of this kind. If they could, then the sound pickup itself would be truly digital.

There have been countless attempts to build digital microphones with the aid of a membrane. Membrane excursion still offers a good measure of the difference between the molecules that strike the membrane from the one side versus those that strike it from the other side. At first glance it would seem not to be too hard, then, to measure the membrane’s excursion digitally. But the attempts at doing this with studio quality have reached the limits of technology sooner than might be wished. For example, it is certainly possible to make this measurement with a laser beam, but then the results are spoiled by noise because the laser obtains its values at only one point on the membrane. This reminds us that the good old membrane in a microphone integrates over all the molecules, and that our result values must take the entire membrane surface into account. And in order to reach the quality level of analog microphones, a digital microphone would have to register membrane displacements about the size of one oxygen or nitrogen molecule!

These considerations, plus the fact that analog microphones obviously work quite well, have led all manufacturers to retain a well-proven, economically affordable technology and to continue using membranes even in digital microphones. The excursions of these membranes also continue to be conveyed in analog fashion. Thus in digital microphones, customary kinds of audio A/D converters convert the analog electrical signals into digital electrical signals.

Ultimately the result is what counts. Sound enters the microphone at one end, and on the other end a digital signal can be obtained with all its pros and cons. What happens in between remains obscure to most users anyway.

A microphone manufacturer long ago might have called such a microphone “digitalized” rather than “digital”—which would have been completely correct in principle. What is generally called a “digital microphone” today is, strictly speaking, an analog microphone that has an additional digital part built in.

2.2. Dynamic range

Dynamic range is the ratio, measured in decibels, between the largest signal and the smallest measurable signal (*i.e.* the noise floor). But anyone who expects dynamic range to apply to an A/D converter in a digital microphone the same way as it applies to an analog microphone is making a huge mistake.

We must take a closer look at the way in which dynamic range is indicated for analog microphones. Determining the “largest signal” is not so simple; rather, it is defined as the largest signal that can be obtained within a given distortion limit—usually 0.5% or 1%. This means that still larger signals can definitely be conveyed, but the distortion will then exceed the chosen limit. Thus the overload limit of an A/D converter in a microphone must be distinctly higher than that of the microphone’s analog components (because overloading an A/D converter generally produces „hard clipping“ as opposed to the „soft clipping“ of the microphone’s analog electronics).

Similar considerations apply to the noise floor of a microphone. For one thing, the noise floor is usually measured with a weighting curve (*e.g.* A-weighting) that more or less closely approximates the sensitivity of human hearing at different frequencies. Thus in the area of the ear’s highest sensitivity, an even wider dynamic range is needed than would be the case with an unweighted noise measurement.

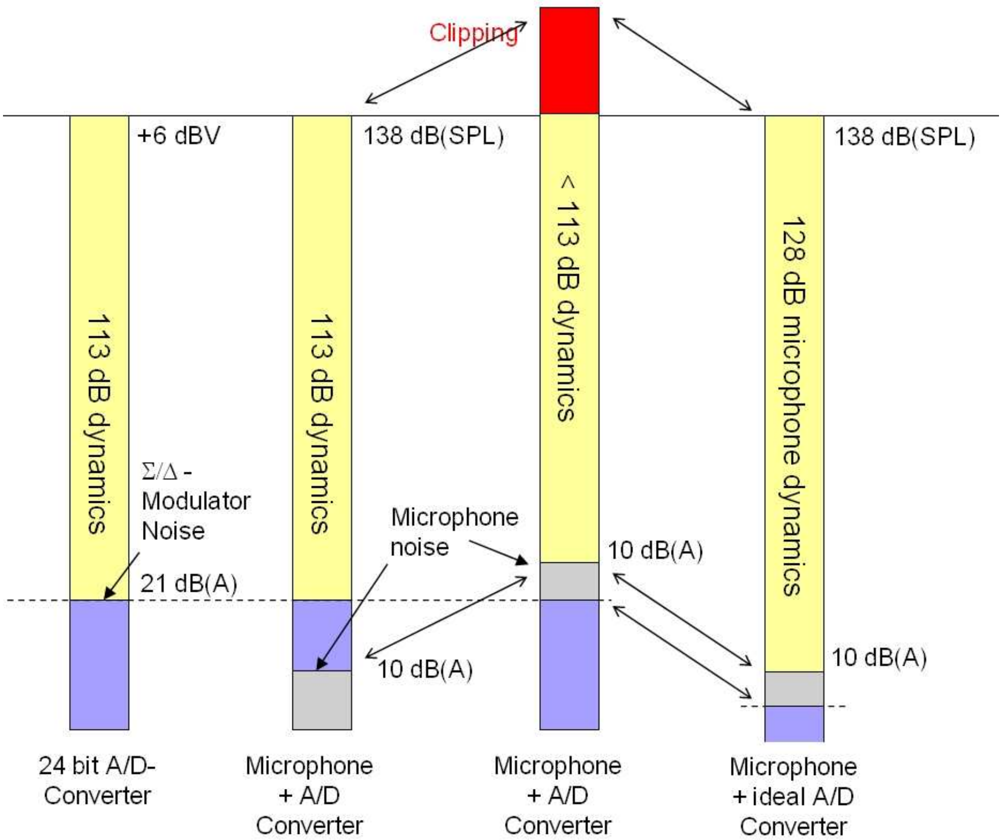

Figure 7: Dynamic range considerations for microphone signals connected to an A/D converter, after [6]

For another thing, no one would be well served if the quantization noise of an A/D converter were at exactly the same level as the inherent noise of a microphone, since the overall noise of the system would then be 3 dB higher. It should also be remembered that the quantization noise of an A/D converter is far less acceptable to the human ear than the thermal noise which (at very high levels of amplification) would ordinarily be audible in a microphone.

In order for the noise of a digital microphone to remain as low as that of an analog microphone, the inherent noise of the digital microphone's A/D converter must be so low that total noise is increased by no more than a few tenths of a dB. If the noise of an A/D converter is at least 12 dB lower than that of the analog microphone, then the noise level of the combination (*i.e.* the digital microphone) will be only about 0.3 dB greater than that of the analog microphone.

These two facts lead to the conclusion that to maintain the full dynamic range of a digital microphone, the dynamic range of its A/D converter must exceed that of its analog components by at least 20 dB. And that is also why the trend toward digitization of microphones has taken so long to get started. The best analog microphones today have a dynamic range of about 130—135 dB. Suitable A/D converters would thus have to offer a dynamic range of 150-155 dB so as not to decrease this range. Such A/D converters have never been built into any microphone so far.

Since even the AES42 (and any standard AES3) interface itself cannot convey a dynamic range like that (ca. 145 dB is the limit for a 24-bit interface), with digital microphones one is forced to make compromises somewhere (see Figure 7).

In digital microphones, most of the circuit stages required by the corresponding analog variants are present as before. These include the impedance converter (for audio-frequency condenser microphones) or the demodulator (for radio-frequency condenser microphones). But the digital versions can dispense with certain parts of analog microphones: the output stage, which is necessary for driving long cables, can be omitted and the A/D converter can be matched directly to the stage that precedes it.

Let us consider the case in which an analog microphone is connected to a preamplifier that boosts its signal by some 10 to 50 dB (depending on the microphone and application) so that an A/D converter can be driven to full scale—perhaps +15 dBu to +24 dBu, depending on the converter and its settings. The actual converter circuit within the A/D unit requires only some lower signal level to reach full scale—perhaps 0 dBu to +10 dBu. Elevating the signal level (which makes sense for reasons such as vulnerability to interference in the signal lines) and then reducing it again will not harm the signal quality in the ideal case, but it can never improve the signal. In a digital microphone it would not occur at all.

Conventional analog microphones, depending on their type and their method of construction, have various output voltages when excited by the same sound pressure levels. Their placement at varying distances from all kinds of sound sources can lead to great differences in output level. In order to drive the A/D converter as well as possible (*i.e.* avoiding the bad signal-to-noise ratio that would be caused by leaving too much headroom) but also for practical reasons (to mix in a sensible way, the faders on the mixing desk must be in a certain part of their operating range), at some point it must be possible to adjust the levels. Most often this occurs in the microphone preamplifier; its gain can be set in more or less coarse increments, or may be continuously adjustable.

For a digital microphone it is naturally possible to perform this gain adjustment digitally at the input of the mixing desk. This solves at least the practical problem of the faders' working range. The effects on headroom and signal-to-noise ratio require further consideration. If the A/D converter in the microphone has distinctly lower levels of noise and distortion than the inherent noise of the microphone, then it can be assumed that digital level adjustments will not cause any decrease in quality. In other cases it is helpful to provide at least one and perhaps several coarse level adjustment stages before the A/D converter.

2.3. A/D Conversion

2.3.1. Introduction

All A/D conversion today follows the delta/sigma principle [7]. This guarantees very good linearity and extremely low distortion. The noise level of some modern A/D converter chips is also very low, but unlike distortion, the noise is not always at the level of the microphone capsule that precedes it. Here, too, solutions already exist for representing the entire dynamic range of an analog microphone digitally (see section 2.3.2 following).

The delta/sigma approach requires a decimation filter to render the data in the usual digital distribution formats. The way in which this filter is realized determines further parameters of conversion such as aliasing distortion [7] and latency; these will remain small if the microphone is operated at a high enough sampling rate. Various possibilities exist for digital signal output; these range from simple digital microphones with USB ports to professional microphones that follow the AES42 standard (see Section 3).

2.3.2. Multi-stage converter schemes

In order to achieve a wider dynamic range than is possible with single A/D converter chips, multi-stage converter designs are available. The classic approach is that of gain-ranging (see Figure 8). The early systems were burdened with artifacts [9], but perfect results can be obtained by suitable means today [10][11][12].

In all existing gain-ranging systems, an input signal is digitized by conventional A/D converter chips at different levels of gain. Small signals, which otherwise would be strongly affected by the limited resolution of conventional A/D converter chips, are first raised in level, then digitized, then reduced in level again. In this way good digital resolution can be attained even for small signals. At the same time, the low amplification factor in the other branch of the amplification circuit provides for overload-free digital signals even at high levels.

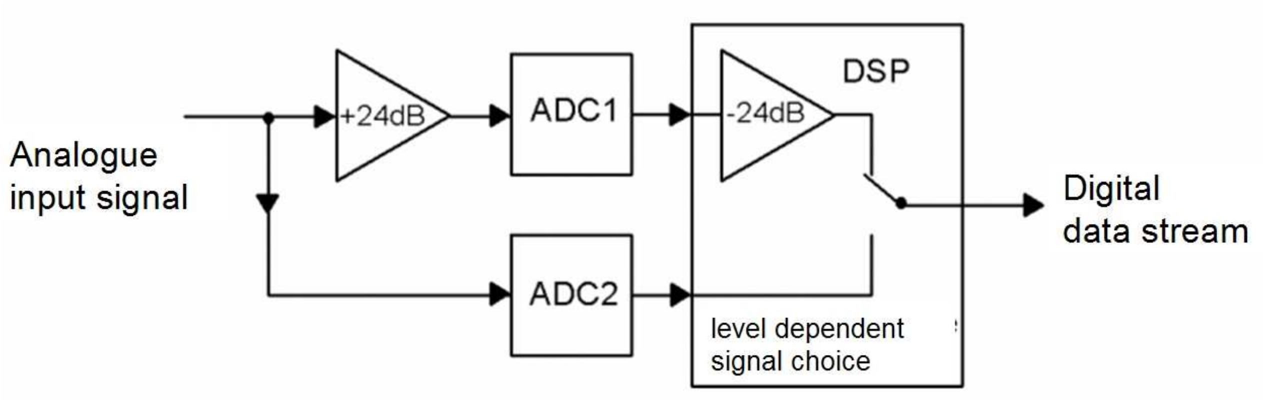

Figure 8: Basic principle of gain ranging

The real art in the design of gain-ranging systems consists in making the existence of different conversion paths unnoticeable to the user, so that the result is perceptually equivalent to that of a single, seemingly higher-resolution A/D converter. Thus any signal alteration that occurs in one of the conversion paths must be precisely reversible.

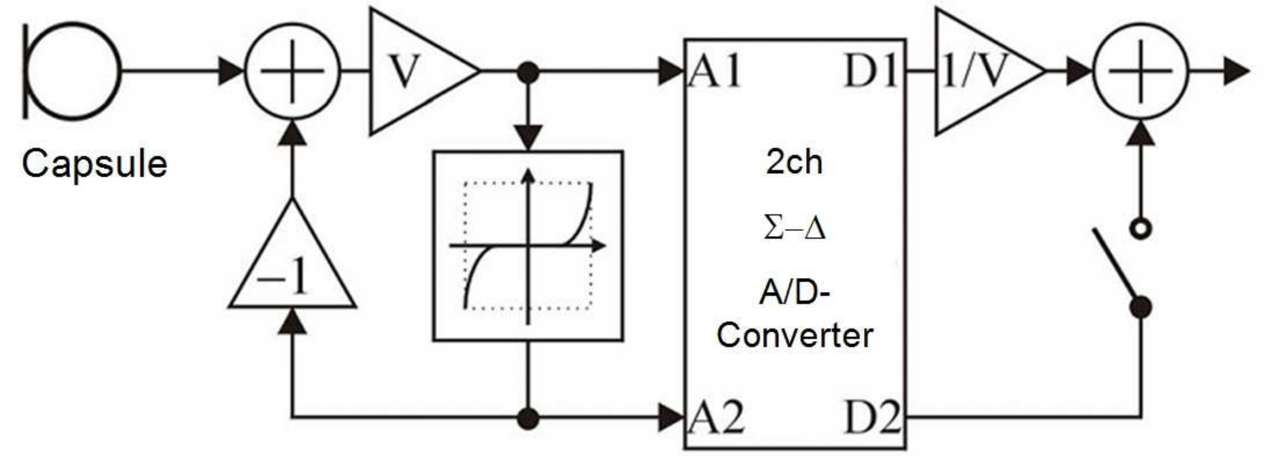

Figure 9: “Neumann system”: Multi-stage converter design according to [8]

There are various approaches to the realization of the principle of gain-ranging, from simple switchers and level matching via potentiometer up to complex signal analyses that correct for component aging, tolerances, and the phase errors of individual conversion pathways.

There also are gain-ranging A/D converters that operate with a non-linear gain characteristic rather than using linear amplification (see Figure 9 and [8]). Different opinions exist concerning the advantages and disadvantages of such systems; the actual reason for such disagreements might be issues of patent rights, given that the problem of compensating exactly for the analog gain in one path—the main problem in any gain-ranging converter—is unaffected by the use of non-linear amplification. This system has been used successfully by one manufacturer of digital microphones for the past several years.

3. The AES42 Standard

3.1. Origin and Motivation of the Standard

The AES42 standard was sketched in its earliest form at the AES conventions in 1997–99. The idea was to create a standard which would permit as much diversity among digital microphones as there already existed among analog microphones. Remote-controlled microphones as well as non-remote-controlled microphones (which of course represent more than 90% of all analog microphones) were to be allowed, as well as microphone power supplies both with and without remote control facilities.

It was decided that the minimum requirements should not be too complex; if an adequate supply voltage was available, that would be enough to ask. As a consequence, it was also necessary to allow for microphones that transmit no control information back to the power supply unit.

Stereo microphones were also to be permitted, in order to use the corresponding capabilities of the AES3 interface. Last but not least, the highest priority in the standard was for microphones to function reliably when connected, before control settings of any kind are made. That would have to be independent of the sampling rate of the microphone and receiver, since there might be no ability to communicate control information. Taken to its logical conclusion, sampling rate conversion (SRC, see Section 4.3) would be necessary.

Overall the intention was to produce as simple a protocol as possible—one that would not even require a CPU and that would encompass only the functions that were absolutely necessary in the microphone.

From the beginning, AES42 was very strongly influenced by four companies: Beyerdynamic, Neumann, Schoeps and StageTec. A few other companies were also involved during part of the time.

The current version of the standard has existed since 2006 (AES42-2006, [12]). A discussion is currently taking place in AES Working Group SC 0404-D concerning a revision of the standard, which is planned for publication in 2011. The questions being debated include the storage of parameters in the microphone as well as the definition of compressor/limiter parameters.

3.2. Operating data and protocols

3.2.1. Powering

One requirement for the operation of digital microphones is an energy supply. Of course this is also quite common with analog microphones, but not every transducer type requires it; most dynamic microphones do not require an energy supply, but condenser microphones do as a rule. Usually this is realized by means of phantom powering via the analog signal cable, which allows up to ca. 170 mW to be provided per channel. The energy requirements of digital microphones are considerably greater, mainly because of the A/D converter; for that reason the same phantom powering arrangement cannot be used. To avoid having to run a second cable to every microphone, any kind of input for a digital microphone must be able to support the electronics of that microphone. USB microphones use the voltage offered by standard USB port.

AES42-compliant microphones are supplied phantom powering (DPP = Digital Phantom Power: 10V +/- 0,5V) that is imposed on a shielded, two-conductor cable. This voltage is also modulated between 10 and 12 Volts as a way of sending control signals to the microphone. The standard allows current consumption of up to 250 mA, enabling the same power transfer as a USB port.

10 Volts was chosen because the resistance of an AES3 cable amounts to 7 Ohms per 100 meters per conductor on average. Thus for the maximum cable length allowed by AES3 this resistance would be 10.5 Ohms, and with the maximum allowable supply current, the voltage drop would be 2.625 V. The components with the highest power consumption in digital microphones operated according to the AES42 standard are digital components driven by (at most) a 5-Volt supply via a voltage regulator. Thus even with the maximum allowable current consumption and cable length, the supply voltage of 10 Volts guarantees that more than 7 Volts will always reach the microphone.

Despite the high maximum current (250 mA) of digital phantom powering, balanced analog microphones that are connected accidentally will not be harmed since they either do not allow current to flow to ground (dynamic microphones), or else are built for similar or significantly higher voltages (condenser microphones).

Matched pairs of supply resistors could not be used because of the high current; instead, the microphone and power supply each contain an AES3-compatible transformer. The transformer in the supply is fed via a center tap, while the powering is extracted within the microphone by way of an identical center tap.

To ensure reliable remote control via pulses imposed on the supply voltage, the input capacitance of the microphone should not exceed 120 nF. These remote control pulses are overlaid on the powering as common-mode modulation with a deviation of +2V (±0.2V).

3.2.2. Audio format

Considering now the “opposite direction,” a digital microphone is expected to deliver an audio signal that complies with the AES3 standard (www.aes.org). Possible sampling rates are 44.1 kHz / 48 kHz and their multiples, with 24-bit resolution. Microphones with support for up to 192 kHz are currently available on the market.

The audio format is thus completely identical to that of ordinary AES3 signals. However, the “user bits” (= bits left free in AES3 for the user’s own purposes) convey information about the microphone and its settings. Unfortunately there is no channel status bit thus far in the AES3 data stream which could identify it as an AES42 data stream and thereby indicate the meaning of the user bits. An AES42 controller must take precautions here to avoid (for example) misinterpreting the user bits if a conventional AES3 signal source is connected to it.

3.2.3. Compatibility

If a microphone fulfills the two basic requirements—digital phantom powering (DPP) and the AES3 output format—then it is AES42-compatible. The same holds true for an interface input.

3.2.4. Synchronization

The AES42 standard distinguishes between two types of operation: *Mode 1* and *Mode 2*.

Mode 1

Mode 1 means that the microphone does not support external synchronization. It runs on an internal clock with a fixed sampling rate. If the input of a controller or mixer (for example) cannot be driven at that rate, the input must provide sampling rate conversion. This will occur particularly when multiple free-running microphones are connected to the inputs of a single system. But *Mode 1* microphones are of course much simpler to design and implement.

Mode 2

As an alternative, *Mode 2* operation allows the simultaneous use of multiple microphones without sampling rate converters. For this purpose, each microphone's sampling rate is set to match that of the master clock in the receiving unit, and then the clock within each microphone is led to agree with the master clock in frequency and phase by means of a split phase-lock-loop (PLL). In this topology one or more voltage-controlled crystal oscillators (VCXOs) replace the VCO of the PLL circuit. The output of the phase comparator and proportional integral/differential (PID) controller are converted from analog to digital in the power supply unit and sent as a control signal to the microphone. This form of "follow-up" synchronization is not unconditionally reliable. The user must ascertain in every case that the microphone is set to the same sampling rate as the master clock, so that the capture range of the VCXO will be wide enough for synchronization to occur.

Does the AES42 standard require Mode 1?

The standard prescribes that every microphone must also support *Mode 1* if synchronized operation is not possible. Actually such a rule is redundant since in the absence of synchronization, each microphone's oscillator will run freely anyway, thus meeting the definition of *Mode 1*.

Equally clear rules for the receiver side would be important here. A multi-channel microphone receiver which lacks sampling rate conversion, and thus does not support *Mode 1* operation, can currently be considered "AES42 compliant" nonetheless. Multiple microphones operating in *Mode 1* might be connected to it, but the arrangement would be unusable because the receiver could not synchronize their signals. Thus compatibility is not assured even when all components follow the standard.

If the multi-channel receiver were to support *Mode 1*, it would not only gain support for *Mode 1* microphones, but its overall reliability would also improve significantly since all synchronization problems would be avoided. If synchronization were lost, sampling rate conversion would ensure undisturbed operation of the microphones. It is simply a fact that to use *Mode 2*, it is mandatory for the user to ensure the correct settings of all connected components.

The manner of synchronization used by *Mode 2* has both advantages and disadvantages. The advantages are that the use of VCXOs allows the quality of the clock signal to be kept very high with regard to short-term fluctuation (jitter). Thus the quality of the A/D conversion is not restricted in advance by poor clocking. Furthermore, the very slow transmission of frequency-regulating information from the receiver to the microphone means that the clock frequency of the microphone can change only very slowly. This, too, has an inherent positive effect on the A/D conversion.

The disadvantage of the system is that the capture range of the VCXO is very narrow. Typically a digital microphone with a sample rate of 44.1 kHz or 48 kHz can detune its frequency by only 6 to 8 Hz up or down at most. Thus it is not possible to use the kinds of pull-up and pull-down accomodations that are sometimes necessary when working between PAL and NTSC video standards (0.1%, corresponding to ca. 44–48 Hz). And only with the aid of sampling rate converters can synchronized *Mode 2* digital microphones be used for varispeed applications, which are everyday occurrences in recording studios.

This raises the question of whether such a restriction in the standard can still be considered up-to-date, or whether other solutions might already be thinkable given the current state of the technology. Since the characteristics of the VCXO are “hard coded” into the *Mode 2* phase regulation scheme of the standard, however, it surely would not be altogether trivial to bring this about.

3.2.5. Remote control

The Appendix to the AES42 standard specifies a series of commands which can be used for controlling extended functionality in a microphone. These control codes are sent to the microphone via pulse-width modulation (+2 V, ±0.2 V) of the DPP. Their transmission is relatively slow (689 or 750 bits/s at 44.1 or 48 kHz respectively), though for most commands this poses no problem. To allow as well for the time-critical synchronization of the microphone through this connection, the standard provides that at least every fourth command code must be a sync word. A higher data rate for issuing firmware updates (8.8 – 9.6 kbit/s) is currently being considered for a future revision of the standard.

A command set is defined for remote control of microphone parameters and *Mode 2* synchronization of the sampling rate. Unfortunately this command set is not arranged orthogonally, *i.e.* in the most straightforward way. The user will not experience this directly, but it makes implementation of the interface by further manufacturers entering the market more difficult. Thus there is a simple and an extended set of commands, but the default settings of the two command sets are not entirely the same.

The transmission format uses a 2-byte protocol. The *address byte* identifies the particular command; then the *data byte* contains the value that is to be set. This is followed by a *transmission pause* with a length equal to at least one byte. This structure enables a rudimentary implementation of the protocol (simple command set) in a microphone with very simple logic components.

Simple command set

The simple command set supports microphone parameters which can usually be set in an analog large-diaphragm microphone. These are:

- Signal “pre-attenuation” (none, 6 dB, 12 dB, 18 dB). The standard doesn't say where in the circuit this "pre-attenuation" should occur.
- Directional pattern in 15 steps (omnidirectional through figure-8)
- Low-cut filter (none, 40 Hz, 80 Hz, 120 Hz)
- Mute function
- Signal limiter disabled/enabled
- Signal gain (unity through 63 dB in 1 dB steps)
- *Mode 2* synchronization

The internal amplification is an interesting point. Even if the 24-bit data format makes it unnecessary to set levels for the microphone signals, this amplification capability was provided in order to support the tape-based 16-bit recording formats which were still common in the late 1990s (DAT, ADAT, TASCAM). Their data format offers a dynamic range 8 bits (about 50 dB) narrower than that of the AES3 interface.

The sole purpose of the peak limiter is to act like an airbag—to contain the damage caused by overload if the user turns the gain up too high.

Extended command set

An extended command set is intended to allow operation without a mixing desk, by implementing all the usual features in the microphone itself.

It extends the protocol with 31 further commands, which are:

- Control of two lights, for communicating with the performers in the recording room
- Generation of four different (but not standardized) test signals. With an unfamiliar microphone this could always be good for a surprise.
- Calibration of the A/D converter in the microphone. Converters in the late 1990s could still be a little touchy when the sampling rate setting was changed (they might emit noise at 0 dBFS!). That could be avoided with a calibration procedure, but this is no longer needed today.
- Resetting (=rebooting) the DSP software (this would ideally be unnecessary in an embedded system)
- Choice among four pages for microphone identification and status flag indicators
- Dither and noise shaping of the quantization error with 16(!) possible settings. (One would be enough, if its noise shaping characteristic matched the spectral curve at the threshold of hearing.)
- Sampling rate (eight possible settings: 44.1 kHz, 48 kHz, 88.2 kHz, 96 kHz, 176.4 kHz, 192 kHz, 352.8 kHz, 384 kHz)
- M/S vs. X/Y selection for stereo microphones (an M/S matrix may be implemented in the microphone interface)
- L/R or M/S balance coefficients (128 possible settings)
- Equalizer (255 manufacturer-specific settings or "no equalizer")
- Polarity ("phase") of the output signal (0°, 180°)
- Stereo/mono selection. In mono mode, both channels carry the same signal. The standard does not specify which signal is chosen.
- Compressor/limiter attack time (8 settings from 0–100 ms, plus 8 "auto attack" modes)
- Compressor/limiter release time (7 settings from 50 ms–5 s, plus 8 "auto release" modes)
- Compressor/limiter ratio, 8 standardized settings (1.2:1 through ∞:1)
- Frequency response of the compressor/limiter side chain, with four settings (flat, 1 kHz high-pass, 2 kHz high-pass, 4 kHz high-pass)
- Threshold of the compressor/limiter (64 settings: 0 dBFS through -63 dBFS in 1-dB steps)
- Brightness settings for the two lights, 16 steps each

The following items are currently being discussed:

1. Ground-lift switch
2. Data storage for parameter settings
3. Reverting to stored parameter settings
4. Reverting to factory parameter settings
5. Reverting to default parameter settings
6. Threshold for the peak limiter, 16 steps from 0 dBFS through -15 dBFS)

The protocol allows proprietary, manufacturer-specific data (*e.g.* firmware updates) to be uploaded to the microphone by using a special command in the extended command set.

Microphone Identification and status flag information

Indications of a microphone’s identity and the settings that may be implemented within it, as well as the current status of those settings, shall be sent from a microphone to its powering unit. One user bit per data word is employed for this purpose. The maximum block length of this data stream is 192 bits. Even though only one bit is transmitted per audio sample, at a 48 kHz sampling rate the transmission of a complete data block requires only 4 ms. To increase the information density of the transmission, four different storage pages or data blocks are defined. By cycling among them, a complete set of information can be transmitted every 16 ms.

Each of the four storage pages delivers information about muting, overload, activity of the peak limiter, and the identity of the current data page.

The additional content of **storage page 0** consists of the low-cut filter, directional pattern, pre-attenuation and gain settings. The availability of various remote control features is also reported. Since the AES42 interface is also intended for use with receivers for wireless microphones, information on the state of the transmitter’s battery charge and field strength, as well as error corrections, are reported. The status of two “call buttons“ for conference setups is also reported.

Storage page 1 has as its additional content the identification of the microphone manufacturer and microphone model.

Storage page 2 has as its additional content the microphone’s serial number, hardware and software revisions, as well as its delay time measured in samples.

3.2.6. DSP

The possibility of using the AES42 interface for remote control of a microphone also makes new microphone concepts thinkable. In principle, within the specifications given in the standard (current consumption, supply voltage, etc.), it remains each manufacturer’s choice as to how much DSP power will be realized in a given microphone. The protocol allows for manufacturer-specific commands which could be assigned to the control of any arbitrary DSP functions (*e.g.* special equalizers or other audio processing filters). To the extent that such commands are kept unpublished, of course, their functions would be available only to the users of that same manufacturer’s AES42 controllers. This can be a trap, since it is impossible to tell by looking at a microphone what settings may be stored in it. With equalizers in a mixing desk, for example, one can normally see whether or not they are activated and if so, what their settings are. But if someone stores processor settings in a microphone and the next user has a power supply unit with no remote control capabilities, that microphone could behave in an unexpected manner without the cause being immediately evident.

Some microphones actually do not “fall back” to a default but reload a previously stored setting. This can be a nice feature, but could also cause errors.

There is a further risk with digital microphones: Whenever digital signal processing is being performed, the transit time can depend on the DSP program that is running. In an extreme case two microphones of the same type could differ in their signal latency (see also Section 4.2.1). The user would need to take appropriate precautions to avoid this situation.

4. Practical issues in the use of digital microphones

4.1. What AES42 controllers or interfaces exist?

Some solutions for the use of AES42 digital microphones are currently on the market. On the one hand are the modules and interfaces developed and sold by the microphone manufacturers themselves. On the other hand are implementations of the interface in existing mixing desks, and naturally--just as in the analog realm--interfaces from third-party manufacturers which offer the controller purely as a format converter with outputs in various other audio signal formats. The actual offerings are easy to summarize. Unfortunately, some products exist only as announcements or declarations of intent from the various manufacturers.

The available range of functionality is very wide. Some simple modules are available which can be inserted directly into the cable path from the microphone. They produce digital phantom powering either from AC (mains) or batteries, and offer an AES3 signal at the other end. In this case the microphone runs in *Mode 1*. Specifically for mobile applications (*e.g.* film sound recording), there also are AES42 inputs in film sound mixers and portable recorders. Most such equipment allows *Mode 1* operation with sampling rate conversion.

Beyond that, AES42 inputs are also offered in large mixing desk systems. In this case only the most basic, necessary remote control functions are usually available, since large consoles offer their own DSP functions (equalizers, dynamics processing), and a complete implementation of the standard would be difficult to integrate into the console’s existing control surface.

Controllers that implement the full range of the AES42 protocol exist in various sizes starting with two-channel solutions, but eight-channel units are becoming increasingly common. Depending on the intended application, conversion to other digital formats (MADI, Ethersound etc.) may also be offered so that digital microphones can also be connected to larger audio networks.

In addition to external controllers and those built in to mixing desks, AES42 ports are also being planned as audio cards for the PC or Mac. Up to four digital microphones could be connected directly to a PCI or PCIe card, for example, so that their signals would be immediately available in the DAW software. The microphones could then be controlled via a driver dialog box.

4.2. What is the difference between using digital and analog microphones?

4.2.1. Latency

When using digital microphones it should be kept in mind that their delay times may differ. On the one hand this depends on the A/D converter chips being used; there is considerable difference in latency among different types. If signal processing is being performed within the microphone, its latency will also come into play. Surely the largest contribution (0.5–2 ms, see Section 4.3) will be that of a sampling rate converter, although it is located in the receiving device rather than in the microphone. Differing delay times will occur above all with simultaneous use of *Mode 1* and *Mode 2* microphones as well as “mixed operation” with analog microphones and external A/D converters. In such cases the actual acoustical and electronic delays should be determined and appropriately compensated. This can be done with suitable acoustic measurement systems or more simply by recording the sound of a “click frog.“ The AES42 standard also provides for determining and conveying the actual latency time of a microphone.

4.2.2. Power consumption

Digital microphones can require more power than analog microphones (max. ca. 2.5 W for digital microphones vs. <200 mW for P48 analog microphones). That could make them less suitable for portable (battery-powered) applications. However, the power consumption certainly does not need to be so high. The AES3 output stage will require ca. 100 mW to deliver the prescribed signal voltages into the prescribed resistance. It is conceivable that in the future, much less power will be required for the rest of a microphone's circuitry than is the case with currently available digital microphones.

4.2.3. Cables

An analog microphone "always works somehow" even if it is connected via less-than-ideal cable. But a digital microphone operating at 192 kHz might not function at all if it is connected to its receiver through a long XLR cable of unsuitable type; the output signal might not arrive correctly at the receiver. At first glance that is a constraint, but on second thought one can also see an advantage in being forced to use suitable cable. Unsuitable cables can also cause a reduction in audio quality with analog microphones, but sometimes this might fail to be noticed because "it works anyway."

One of the claims made for AES42 is that existing microphone cables can be used with digital microphones, but that is true only to a limited degree in practice; this claim must be regarded with caution.

Digital microphones have the advantage of continuing to use 3-conductor XLR cables, although they can be carrying two audio channels as well as the control signals for the microphone. For analog microphones this would be possible only with special cables or add-on units (*e.g.* the capability of remote pattern control by varying the P48 supply voltage).

4.2.4. Are digital microphones more complicated to use than analog microphones?

We have become well accustomed to working with analog microphones and 48-Volt phantom powering. The introduction of digital microphones calls for special effort which in some cases will approach that of a laboratory. Until the powering and controlling of digital microphones has come into such common use as (for example) the P48 standard, digital microphones will not be as simple to use as analog microphones. A closer parallel would then exist with analog microphones that require special power supplies or control units (*e.g.* tube microphones with special power supplies and cables, or microphones with remote-controlled directional patterns).

4.3. Sampling rate conversion

4.3.1. Necessity

Here different opinions and many prejudices exist. A sampling rate converter (SRC) is needed whenever the sampling rate of a digital audio signal cannot simply be accepted by a system, and must instead be adapted to suit it. This can occur even with rates that are nominally the same, if they have been driven by different clocks.

That is not fundamentally a bad thing, but of course the question of audibility comes up immediately.

For a digital microphone that follows AES42, in principle one could synchronize the receiving interface to the sampling rate of the microphone.

That would work as long as there is only one microphone, but it is no longer a solution if there are two or more. This was understood long before the AES42 standard. The choices are to learn to live with divergent sampling rates or to adapt the microphone to an existing setup (*e.g.* mixing desk). The standard permits either approach.

4.3.2. Operating principles and characteristics

Basic principle

The basic task of any sampling rate converter is to calculate sample values at the target sampling rate which are equivalent to those of the source signal. Theoretically one could interpolate additional samples into the source signal until sample values for any other given sampling rate became available; then the needed values could simply be picked out. But this "infinite interpolation" would be equivalent to an analog conversion, and in this sense a D/A converter followed by an A/D converter running at the target rate would represent the simplest form of asynchronous sampling rate converter.

To avoid this detour via analog, the required sample values can be calculated with the aid of an interpolation filter (Figure 10). Conversion quality sufficient for a high-quality digital microphone can be achieved with an oversampling ratio of 2^{20} or higher. But the calculation of such an enormous quantity of sample values is economically infeasible, and the resulting sampling frequencies in the gigahertz range would no longer be practical, either. For that reason the interpolation process calculates only those sample values that are actually needed for the output signal (Figure 11). For this purpose, it must first be determined which points in time within the source sampling period correspond to the points in time of the output sample values.

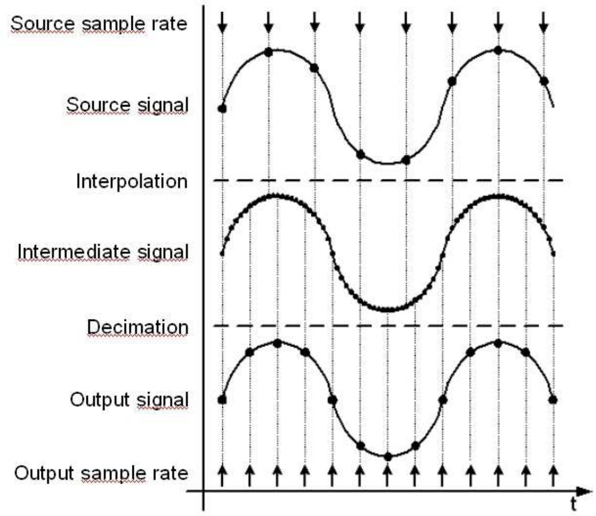

Figure 10:
Top: Source signal with sampling rate fq,
Middle: intermediate signal interpolated from the above, with a sampling rate f >> fq,
Bottom: Target signal with sampling rate fz ≠ fq

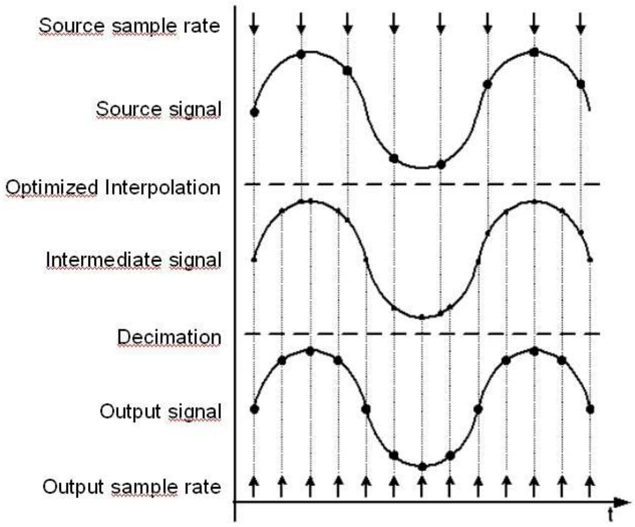

Figure 11:
Top: Source signal with sampling rate fq,
Middle: Intermediate signal which has been interpolated in an optimized manner,
Bottom: Target signal with sampling rate fz ≠ fq

Because of the mutually asynchronous clocks of the microphone and target system, the moment of interpolation varies within the source sampling period on a sample-by-sample basis, and must therefore be recalculated for each target sample value. This creates the possibility of jitter between the two clocks. To find the proper moment for interpolation, the relationship between the two sampling rates is determined continuously, and the needed samples are calculated on the basis of that information.

Intermediate storage

Since the relationship cannot be determined in less than one sampling period, an average of several periods is taken; this also allows any variance caused by jitter to be suppressed. In the interim, intermediate storage (buffer memory) is needed to store any extra samples resulting from sudden changes in the previously calculated sampling rate relationship, or if needed, to provide additional samples until a valid new value of the sampling rate ratio can be determined.

If the sampling rate ratio should change in one particular direction over a longer period of time, the buffer memory could over- or underflow. To avoid this, the determination of the sampling rate ratio is regulated so as to take the "fullness level" of the buffer memory into account. If it is too low, the conversion rate is briefly increased, or if it is too high, the conversion rate is decreased until the buffer reaches its nominal "fullness level."

The speed of this regulation process influences the quality of the sampling rate conversion. If it is set to be very slow, a large buffer will be necessary. This will effectively suppress errors caused by jitter, but will increase the transit time (latency). A very small buffer will suffice if the regulation is designed to be fast; lower latency will be achieved, but then one would have to reckon with some reduction in quality as the result of any clock jitter.

With a digital microphone as source, the sampling rate in relation to any given target system will be constant. Intermediate storage thus has the sole function of buffering any short-term variations in the sampling rate ratio that occur due to jitter. Since the microphone's internal clock will usually be quartz accurate and thus nearly jitter-free, this leaves only the target system's clock (depending on its origin) as a possible source of jitter. Here, too, in good implementations the jitter is so low that a memory buffer of just a few sample values is perfectly adequate. Of course the previously described sampling rate corrections within the buffer regulation are inaudible; not even a brief change in musical pitch will occur, for example. The input signal is correctly converted at every moment, but the exact moment chosen for the calculation and the calculation rate are adjusted dynamically to fit the situation within the sampling rate converter.

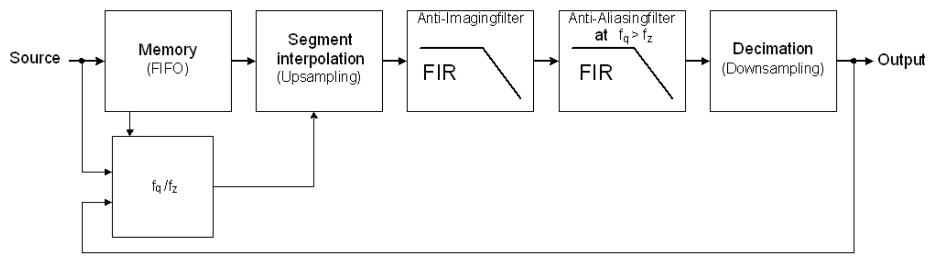

Figure 12: Block diagram of a sampling rate converter

Interpolation and decimation filters

If a microphone is operating at a higher sampling rate than the target rate, interpolation alone will not suffice; the source signal's greater bandwidth cannot be represented at the sampling rate of the target signal. To prevent aliasing distortion, an appropriate band-limiting process must precede the conversion. This requires a decimation filter.

Interpolation and decimation filters are usually implemented as linear-phase FIR filters [7]; the two functions can also be combined into a single filter. Such filters are responsible for the greatest part of the total latency, with the absolute transit time being directly related to the translation ratio. But it also depends on the sampling rate. If both the microphone and the target system operate at 48 kHz, for example, latencies of less than a millisecond are possible without diminishing the microphone's quality through noise or distortion from the sampling rate converter.

Half-band filters are often employed, which can lead to certain (usually negligible) aliasing errors around the Nyquist frequency [7].

Figure 12 shows the overall block diagram of an asynchronous sampling rate converter.

Latency

Typical signal delay times for SRCs are 0.5–2 ms, depending on the sampling rates at which they operate. Myths to the contrary notwithstanding, the latency time of an SRC is precisely defined, and is always given in the manufacturer's specifications. It depends on the sampling rate. The input values of asynchronous signals are also interpolated with complete phase accuracy, *i.e.* precisely the same sample values are generated as would have been generated with a synchronous signal at the input.

4.3.3. Modern sampling rate converters and audio quality

It is immediately clear that sampling rate conversion works better the higher the sampling rate of the intermediate signal is (oversampling) and the better the low-pass filter is at allowing only the original signal components to pass through. As with all things in life, there are possible ways to implement an SRC less expensively. It is very tempting to set a lower oversampling ratio, since this would avoid having very high frequencies in a semiconductor circuit while reducing the number of semiconductor components. The low-pass filter offers a further cost-saving potential: Insufficient bit depth, a lower-grade filter, and the previously mentioned half-band filters can be used. Each of these measures (or all of them combined) offers a savings in semiconductor components.

In the early days of digital audio, semiconductor technology itself had not yet reached a stage allowing truly high-quality SRCs to be built. Thus early on, SRCs acquired the sense of being somewhat disreputable—always audible in their effect, and better not used at all. Unfortunately, people easily overlook the fact that quality has its price. And it is damnably difficult to redeem a bad impression—nearly impossible, since now that people no longer hear artifacts from SRCs, they also tend not to ask whether an SRC is in the circuit.

Truly high-quality SRCs have a dynamic range greater than 140 dB, “distortion” less than 0.00001% and no modulation-dependent distortion of any kind. SRCs today genuinely achieve near-24-bit quality, and can be employed with no loss of quality worth mentioning. Thus the operation of a digital microphone in *Mode 1* can be seen as a true alternative to synchronized operation in *Mode 2*. Generally the only remaining argument against *Mode 1* operation is the transit time of the SRC.

SRCs have long ago become ubiquitous. Practically every CD has had its sampling rate converted downward, and nearly every DVD or Blu-ray Disc has had its sampling rate converted upward. Often the process goes unnoticed since the SRC is concealed within the audio preparation software.

Terrible things can occur when sound is added to picture. Image processing has no equivalent to SRC; entire frames must be repeated or skipped when differing frame rates need to be matched up. If sound is already attached to the picture, then either 40 ms is repeated or 40 ms is cut out. Today, audio companies that once built their legendary reputation with dynamic range improvement offer equipment that tries to use cross-blending to conceal the transitions between these 40-ms segments. That, too, is a kind of "sampling rate conversion"!

4.4. Operational reliability

Operational reliability is a decisive criterion for the professional user when evaluating any new technology. The following section will illustrate various aspects of the use of digital microphones.

4.4.1. Increased likelihood of failure due to greater complexity of the electronics?

Moving mechanical parts have the highest rate of failure—connectors above all. If the number of cables and connectors is reduced, the failure rate goes down as well. In this respect the integration of the preamplifier and A/D converter into the microphone is helpful. Setting levels with a digital signal processor rather than a potentiometer eliminates another possible source of mishap.

Of course all electronic components must be operated within their specified temperature range to maintain their guaranteed service life. The developer must also keep in mind the maximum number of write operations that non-volatile memory allows, as well as its maximum data retention time (up to 100 years!).

In summary, high-quality microphones, be they analog or digital, must be dealt with as professional rather than consumer products.

4.4.2. Electromagnetic compatibility and radio-frequency interference

Every serious manufacturer can tell stories about the difficulties of getting analog microphones and amplifiers to be "RFI-proof". This task will become more and more demanding in coming years, as communications services move increasingly into the UHF frequencies set free by the recent changeover to digital television.

There are analog microphones which have exemplary immunity to interference from external signals. But due to the high signal levels of the AES3 data stream, digital microphones are immune to interference into the cable. The same is also true for interference from magnetic fields.

4.4.3. Phantom powering

The AES42 standard defines digital phantom powering (DPP) so that it will have enough current in all cases. Digital technology also offers the manufacturer possibilities for recognizing defects in powering, and above all, for notifying the user of the problem in some appropriate way—*e.g.* via the return data stream or a flashing LED. This capability doesn't exist with analog technology; insufficient powering may be recognized only when high signal levels occur, *i.e.* while recording.

4.4.4. Humidity

The absence of humidity problems is a side effect which the manufacturers discovered only at a rather late stage. The power consumption of half a Watt to one Watt dispels condensation in the microphone without noticeably raising the temperature of its housing.

4.4.5. Cable capacitance

With analog microphones, the use of long cables can cause an overall loss of level as well as a reduction of maximum output level at high frequencies. But even with cable lengths far beyond the specified limit (usually 300 m) a more or less acceptable signal generally still comes through—which speaks well for the operational reliability of analog microphones.

With digital microphones the situation is somewhat different. Cable capacitance has no effect at all on the audio quality of digital signals. If the "sound" of analog cables was ever a topic for discussion, this worry now belongs to the past.

Due to multi-bit sigma-delta converters, the problem of jitter has shifted increasingly to the background. For distances up to around 100 meters, AES3 signals can be conveyed by "normal" microphone cables. For lengths up to 300 m, impedance-matched 110 Ohm cables can be used. 75-Ohm coaxial cables can carry AES3-id signals[1] up to 1000 m without any loss of audio quality. Digital technology offers a suitable transmission medium for every application, but the use of unsuitable media or the overstepping of stated maximum cable lengths does not work.

4.4.6. Input impedance of the equipment to which the microphone is connected

The input impedance of an AES3 port is standardized. No manufacturer would dare to deviate from a 110 Ohm input impedance.

4.4.7. Synchronization

This set of problems does not exist for analog microphones, making life much simpler for them.

Synchronization to a master clock signal which is received by each piece of connected equipment has been used since the very beginning of digital studio technology. This approach obviates the need for sampling rate conversion. The concept of distributing a centralized clock is continued in AES42 *Mode 2*. The necessity to arrange for synchronization naturally reduces operational reliability, since the correct distribution and functioning of synchronization is left up to the user. No one loves this responsibility; thus for "quick and dirty" setups, analog equipment is often preferred.

AES42 *Mode 1* was conceived as an uncomplicated "plug-and-play" solution which doesn't even require the user to know the sampling rate for which a microphone has been set; the microphone can simply be connected to any *Mode 1* input. Unfortunately, *Mode 1* support on the receiving side is not prescribed in the AES42 standard as obligatory (*cf.* Section 3.2.4).

4.4.8. Embedded software

This is another problem unknown to the users of analog microphones. Where there is no software, no software failures can occur. A basic discussion of software reliability may be useful in this connection.

DSP software is fundamentally very reliable because it can never get into an "undefined state." This is a side effect of its deterministic behavior; the program runs in a sample-driven loop with known execution times for routines that never vary. "Dead" code branches and untested software states are both rare, so bugs seldom occur.

But this matters only to the extent that no other tasks are running asynchronously on the same processor, *e.g.* control programs or code that reads from or writes to storage media. Whenever several asynchronous tasks exist, an operating system must be used, and of course it, too, must be tested for reliability. The additional processor overhead increases the need for computational power and raises current consumption; the resulting increase in heat further reduces opera-

[1] Explanation: AES3 id (Information document) describes a way to transmit standard AES3 signals via a coaxial 75 Ohm cable.

tional reliability. User access to data storage via software is the most dangerous of all, since corruption can occur.

An excess of gimmicks and software tricks can be harmful to reliability; they must all be tested very conscientiously before the manufacturer can be sure of getting a market advantage from them. Certainly "less is more" in this area.

On the other hand, the software in a microphone can also monitor its own operational status. Thus, for example, one can imagine a microphone which notifies the user of any deviations from the manufacturer's specified tolerances, assuming of course that the corresponding sensors are built in to the microphone. Some present-day microphones, for example, already check for proper installation of a capsule and report its status to the receiver.

4.4.9. Summary

From a technical viewpoint, the use of digital microphones increases operational reliability. Signal integrity is improved, the likelihood of electronic failure is decreased because of the reduced number of components in the signal chain, electromagnetic compatibility and immunity to RFI and magnetic fields are improved, defective powering can be detected by the microphone, and finally, the increased power consumption makes the microphone less sensitive to humidity in the air. The audio signal is not influenced by the cable.

Operational reliability will surely not be notably influenced by the age of the products. It will be endangered far more by contradictory "interpretations" of the AES42 standard, by differing software conditions between the microphone and the receiver, by parameters stored in the microphone, by sampling rates that are too high for unsuitable cables and the like. All problems with AES42-conforming digital microphones thus far have been traceable to these issues.

A digital system is definitely susceptible to malfunction due to improper operation or defective components. The need for paying careful attention is thus greater.

The problems of synchronization and embedded software should be solved by manufacturers limiting themselves to reliable solutions and refraining from too many "gimmicks."

4.5. Areas of application for digital microphones

One can imagine uses for digital microphones in many fields. One is in fixed installations (*e.g.* theaters or opera houses) where each program might call for different settings. The directional pattern, for example, could be set comfortably from the control desk without requiring an adjustment to be made at the microphone itself (*e.g.* on a lighting bridge above the stage).

Digital microphones make even more sense in an environment of that kind, because the fact that A/D conversion takes place inside the microphone makes them exceptionally insensitive to interference from light dimmers, for example.

In connection with the converter principles mentioned above, digital microphones could also be of interest in critical recording environments. An appropriately constructed microphone would make level-setting for each microphone unnecessary. That would also mean that normally, no overload would ever occur, either. A system of this kind is helpful most of all if, for lack of personnel or time, the signal levels cannot be properly set in advance, *e.g.* for news reporting. Actuality recordings can be either over-recorded or (due to an excess of caution) under-recorded, often requiring that much time be spent in post-processing to reconstruct the signals so that they are suitable for broadcast. As with film sound recording, the previously mentioned immunity of digital microphones to interference is also a great advantage in this situation.

Naturally digital microphones are also very good for use in the studio sector, even though the special environment makes their advantages somewhat less noticeable. The dynamic range improvement due to the avoidance of possibly inferior microphone preamplifiers and A/D converters could have a positive effect especially in the semi-professional arena.

5. Future directions for digital microphones

5.1. Compatibility with existing systems

The specific way in which digital microphones must be powered prevents them from being considered directly compatible with existing mixer and preamplifier systems. But once this powering has been provided, the digital audio signal can be fed into nearly all digital production environments without difficulty. To use the microphone with an external controller requires a format converter, but some interfaces already offer a great variety of well-known output formats, either natively or via expansion options.

Integration into existing systems is even simpler if the manufacturer of the mixing board offers an AES42 input port. Then the user would experience no difference between analog or digital microphones in terms of integration with the mixer and the signal processing within it.

5.2. Investment costs (amplifiers, capsules, converters)

Upgrading to digital microphones is naturally an economic issue as well as a sonic one. Many studios already have a large investment in analog microphone technology and do not feel called upon at present to make this investment a second time. At least the microphone amplifiers (the internal circuitry of the microphones) would need to be exchanged for a digital variant. With some manufacturers that is the only requirement for refitting a microphone; all capsules (and potentially accessories) that were already on hand for the analog microphone can then continue to be used on the digital amplifier.

But in some cases, especially with microphones that have not been designed as the combination of an amplifier and one or more interchangeable capsules, an entire new microphone must be bought. Until about two years ago, the digital versions of microphones were still distinctly more expensive than their analog counterparts. But the manufacturers have adapted to the market in the meantime. For a complete new microphone (which might be necessary anyway), it hardly makes any difference in price whether a digital or an analog microphone is being bought.

Naturally, suitable controllers are needed for operating the microphones. Here, too, some things have happened in the last few years. The offerings have increased, and the manufacturers have recognized that the decision for or against a digital microphone should not be made solely on the basis of price—so they have seen to it that the receivers are now priced interestingly in comparison with analog microphone preamplifiers.

But how future-proof is such an investment? With analog microphones that are well taken care of, it can be assumed that they can be resold even many years later with only a moderate loss in comparison to their cost when new. But in many areas the introduction of digital technology has caused developments to advance more rapidly, and product life cycles to become shorter. What will happen if a digital microphone that you buy today because of its technical capabilities (*e.g.* the DSP or logic chips that it uses) is no longer “up to date” in five years and can no longer be used, either due to shortcomings in its features or (far worse) changes to the interface? The AES42 standard is being asked to form a reliable basis upon which manufacturers and users alike can be certain that their investments are “future-proof.”

5.3. Differences among implementations of the AES42 standard

The AES42 standard allows a wide range for implementing a digital microphone. One basic, defining choice is whether to synchronize the microphone or not. In the standard the two possibilities are described as *Mode 1* and *Mode 2* operation (see Section 3.2.4).

Currently available digital microphone solutions are rather strongly manufacturer-specific. The controls offered by the powering units of manufacturer "A," for example, match the accessible parameters of the microphones from the same manufacturer very closely. Manufacturer "A" offers only microphones which support *Mode 2*, and for this reason the power supplies and controllers of this company lack the sampling rate converters that would be necessary for their use with microphones from manufacturer "B." Mixing desk manufacturer "C," on the other hand, supports only a few microphone parameters and offers no support for *Mode 2* whatsoever; thus few of the special characteristics and functions of the microphones from manufacturer "A" can be used.

Further differences occur in the extent of the microphone settings that can be controlled remotely. Far more possibilities are defined in the standard than are offered by any practical implementation at present.

In the simplest case a digital microphone might neither have any settable parameters (*e.g.* level control, filters, directional pattern) nor be synchronizable; it might fulfill the requirements of the standard only with respect to the interface: operating voltage, current consumption, protocol of the output signal, etc. A "full-fledged" digital microphone by contrast could be synchronizable via *Mode 2*, support two-channel operation, have adjustable directional patterns, and perhaps even offer special filters.

Since the standard is so broadly encompassing that any given microphone can never serve all of its aspects simultaneously (for example, specific parameters are defined for wireless microphone systems), the requirements for the receiver—the powering and control unit—can be very complex.

5.4. Which areas of application hold the future of digital microphones?

Digital technology has moved into many areas of audio that were dominated by analog technology as recently as a few years ago; in some areas its hold has become nearly exclusive. This applies as well to telecommunication, radio and television. Given that, one might dare to predict that microphone inputs and microphones in general will follow this trend. This doesn't necessarily mean that in the future all microphones will be built to the AES42 standard, but the number of digital microphones will certainly continue to increase.

As in the other fields just mentioned, what may now seem difficult and expensive to obtain will, in a few years, very likely become less expensive and more readily available as digital technology enters the field.

5.5. How can digital microphones be made (still) more appealing to work with?

To make the use of digital microphones more appealing, it would surely be helpful to place much more emphasis on compatibility and interchangeability. The current AES42-2006 standard does not yet ensure full compatibility between all AES42-compliant items of equipment.

In so doing it would seem advisable to revisit the fundamental question of which special features make sense for a digital microphone. Filters and dynamic processors in a microphone can be dispensed with if similar capabilities are available in the mixer or workstation that one is using.

Gain control inside the microphone, if it merely reduces the available dynamic range, seems not to make sense, either. But depending on the way in which the amplification and conversion in the microphone are carried out, retaining gain control might still make sense.

Limiting the number of settable parameters, each of which the user must know about and keep track of, would contribute greatly to the user‘s knowing what all is going on in a microphone.

5.6. Summary

Advantages of digital microphones:

- Good for use if the microphone cables will be subject to strong electromagnetic interference
- No need to set levels at the microphone preamplifier, with the 24-bit digital systems of today
- Various manufacturer-independent remote control options are available, so that for example the directional pattern can be set directly at the mixing desk
- Microphones can give status indications, *e.g.* to show which microphones are active at a given moment. Indications of microphone type, manufacturer and other information bring a high level of operating comfort
- Microphones can have “personalized” user settings

Advantages of analog microphones:

- A potentially greater dynamic range than digital microphones; the 24-bit interface of AES42 limits the dynamic range to ca. 144 dB; analog microphone amplifiers of the highest grade can surpass 155 dB
- Much simpler handling: Plug it in and it works. An experienced sound engineer will know what kind of sound to expect from a given microphone type
- Microphones can be exchanged at any time since there are no sound-altering settings which could be stored in the microphone
- No signal delay; microphones of all brands and types can be used together; the most suitable microphone can be used for each application, instrument and type of placement
- No incompatible software versions for the interface; the microphone never requires a software update
- The microphone interface is standardized with much lower electrical power; considerably less energy is consumed by using analog, phantom-powered microphones, low-current microphone preamps and energy-optimized A/D converters than digital microphones (if they conform fully to the interface standard); better suited for battery-operated systems
- Modern, RF-proof analog microphones and modern microphone preamp architectures also allow for mostly interference-free operation, even in the presence of strong electromagnetic fields.

Questions remain …

The attractive possibility of “going digital” directly at the microphone, and the equally attractive possibility of remotely controlling many aspects of the microphone’s performance, unfortunately tend rather to make operation more obscure, particularly with more extensive systems. New problems need to be considered, such as:

- What is the A/D converter latency in each microphone?
- How will synchronization occur?
- Are there latencies from sampling rate converters to be considered?
- The delay time for signals coming from digital microphones can no longer be determined by simply measuring the distance from the sound source; it will depend as well on the characteristics and operating modes of the microphones. If different microphones (or similar microphones with different settings) are used simultaneously, the delay times must be adjusted carefully to be equal.
- Where must additional delay be applied to prevent comb-filter effects?
- Which settings have been set in which microphones? Even the menu of possible settings for one microphone can fill many pages of text.
- Do any microphones have settings left over from a previous production, such as a limiter that is still active?

The user of digital microphones will be set free from certain simple tasks, but must also come to terms with tasks that may be far more complicated. Apart from very simple microphone setups (*e.g.* two microphones with no remote control), preparing for a production such as an orchestral recording with digital microphones could require considerably more time. The advantages that stem from the overwhelming number of new possibilities can quickly become disadvantages due to the difficulty of keeping track of all the details.

6. The Authors

This document has been created through the co-operation of several authors from different fields and companies. It is intended to provide information in a neutral and professional manner.

The following summary lists the authors who are primarily responsible for each section. Helmut Wittek and Claudio Becker-Foss have overall responsibility for the organization and moderation of this White Paper.

Comments and suggestions for improvement are gladly accepted at aes42@hauptmikrofon.de.

Section	*Primary Author(s)*
Introduction	Becker-Foss
1	Langen
2.1.	Jahne
2.2.	Jahne, Flock
2.3.	Jahne, Werwein, Langen
3.1.	Jahne, Becker-Foss
3.2	Werwein, Becker-Foss, Langen, Flock
4.1	Becker-Foss
4.2.	Flock
4.3.	Jahne, Werwein
4.4.	Becker-Foss, Langen, Jahne
4.5.	Becker-Foss
5.1.	Becker-Foss
5.2.	Becker-Foss
5.3.	Flock
5.4.	Flock
5.5.	Flock
5.6.	Jahne

7. Bibliography

[1] "NeXT Computer: when cool wasn't enough": http://www.vnunet.com/Features/1112630 , retrieved August 2009

[2] Paul, Jon et. al.: "Digital Output Transducer", United States Patent, Patent Number 5,051,799, Sept. 24, 1991

[3] Goossens, Sebastian: „Das KEM-microphone. Praktische Erfahrungen bei Musikproduktionen", 21. Tonmeistertagung, Hannover, November 2009, http://www.irt.de/IRT/FuE/ak/pdf/KEM-MIC.pdf

[4] Harris, Steven; Gong, Xue-Mei; Josephson, David: "Towards a Digitally Interfaced Microphone", Standard, presented at the 103rd AES Convention New York 1999, preprint 4518 (B-8)

[5] AES42-2001: AES standard for acoustics – Digital interface for microphones. Audio Engineering Society, Inc. 2001

[6] Langen, Christian: „Geschichte der "digitalen" Studiomicrofone", 22.Ton meistertagung 2002, Hannover, Germany

[7] Weinzierl Stefan: „Handbuch der Audiotechnik", Springer 2008

[8] Leschka, Stephan: „Digitale Microphontechnik", 34. Jahrestagung für Akustik, http://www.neumann.com/download.php?download=lect0052.PDF, retrieved August 2009

[9] Myers, John P.; Feinberg, Abe: High-Quality Professional Recording Using New Digital Techniques, Journal of the Audio Engineering Society (JAES) Volume 2, October 1972, No. 8, S. 622 – 628.

[10] Tani, Yasunori et. al.: 20 bit A-to-D & D-to-A Converting Systems for Digital Audio. IEEE Transactions on Consumer Electronics, Vol. 33, No. 3, August 1989

[11] Knoth, Matthias; Reußner, Thomas: Stage Tec Entwicklungsgesellschaft für professionelle Auditechnik mbH: Vorrichtung zur Umwandlung analoger Audiosignale in einen digitalen Datenstrom, Offenlegungsschrift P 44 20 713.1, DE 44 20 713 A1, 1996

[12] Jahne, Helmut: Stage Tec Entwicklungsgesellschaft für professionelle Audiotechnik mbH: Verfahren and Anordnung zur Analog-Digital-Wandlung von Signalen: Offenlegungsschrift 196 06 261.6, DE 196 06 261 A1, 1997

[13] AES42-2006: AES standard for acoustics – Digital interface for microphones. Audio Engineering Society, Inc. 2006

8. Table of Illustrations, Sources