ENGINEERING IN RECORDING+

Jeremy Wells*

Audio Lab, Department of Electronics, University of York

Abstract: One does not have to look very far to find controversy in the use of the name 'engineer' in the field of audio recording. It is a ubiquitous term but there are those who firmly believe that the act of 'recording' is not 'engineering'. This paper briefly surveys definitions of engineering which exist in the literature and then applies these to specific, documented examples of recording processes. These processes are described in terms of the knowledge, training and technology they require for their execution. The purpose of these case studies is not to prove that recording is or isn't engineering; rather it is to highlight how activities undertaken by those who make sound recordings can overlap with generally accepted notions of engineering. The primary motivation for this work is pedagogical: the presented activities can be used as examples in general engineering education and to illustrate the nature of engineering within degrees in sound recording and music technology. Links to materials for supporting teaching are also provided.

Keywords; Audio, Sound Recording, Multidisciplinarity, Professional engineer.

**Correspondence to: J. Wells, Department of Electronics, University of York, YO10 5DD, UK. E-mail: jez.wells@york.ac.uk*

1. INTRODUCTION

1.1 Is Recording Engineering?

Whilst the popularity of sound recording and music technology courses continues (Boehm, 2007**)** it is often asserted that "engineering suffers from an image problem" (Robson, 2012). *Is Recording Engineering?* is a public engagement project, supported by the Royal Academy of Engineering, which is intended to explore the relationship between these two fields in both industry and education. Its inspiration is the idea that engineering as a discipline can be more widely understood by the broader public if the way in which it permeates a popular multidisciplinary subject area such as music technology and/or sound recording is explored with them. As the title of the project acknowledges, a connection between engineering and what happens in a music recording studio is not universally accepted. However this paper argues that there are specific aspects of sound recording, and approaches to its execution, which can act as examples of engineering in action. The project and this paper specifically focus on the term 'recording engineer' as opposed to 'audio engineer'; the latter is often used more generally and can be applied to less controversially defined endeavours such as the design of audio equipment. Before presenting these examples of 'recording engineering in action' it is first necessary to

+ This work is supported by an *Ingenious* public engagement fellowship with the Royal Academy of Engineering.

survey some definitions of engineering. With an overview of these definitions in place they are then applied to three varied recording situations/tasks which can be explored in engineering or music technology teaching. For the engineering student there is the opportunity to reflect on the nature of their discipline via an accessible and recognisable scenario. For the music, or music technology student, there is the opportunity to recognise facets of engineering in their subject area, and also to import good engineering practice in their approaches to that subject; in short, to become more ‘engineer-like’. For these latter students there is the employability benefit to being exposed to a more STEM orientated approach to such tasks, improving numeracy and science skills as well as gaining an understanding of engineering. The attractiveness to employers of such skills and understanding is well documented (e.g. CBI, 2008).

1.1 What is a recording engineer?

A common, broad definition of engineering can be summarised as the application of “scientific and technical knowledge to the design, creation and use of structures and functional artefacts” (Open University, 2011). This is a useful starting point: the use of organised and verifiable knowledge about the physical world to create and use things within it. However this simple statement does not specify how to move from the knowledge and understanding to the artifact and how success in doing so might be measured. Despite attempts to rouse public support for legislation to better define the identity and role of the engineer (e.g. Robson, 2012) it remains the case that in the UK anyone may refer to themselves as an engineer, in contrast to so-called protected titles such as ‘doctor’ and ‘architect’. In contrast, there is no such clamour to protect via legislation the term ‘musician’. Whilst the latter point may seem facile, the medical doctor and the musician might be considered to be at opposite ends of a continuum, and engineering must lie somewhere along this continuum. It seems reasonable to assume that there is widespread agreement that the protection of the title of doctor is beneficial but that a similar protection for musician is impractical and unnecessary. That said, recognised qualifications are often required for some roles, such as instrumental teaching (for example a ‘licence’ to teach, such as the licentiate of Trinity College, London). Using both music and medicine for comparison David Blockley (2008) explains one facet of engineering that pushes it towards one end of the continuum:

> Many years ago I was listening to a well know jazz critic interviewing Louis Armstrong on the radio. After a long erudite discussion of the contrapuntal complexities of his trumpet playing the critic asked ‘Louis how do you do it?’ The reply delivered in typical rasping style was, ‘Man I just blows.’ Why is this relevant to engineering? Because, like Louis, much of what practitioners actually do is as a result of experience - they learn from doing the job. However if Louis had blasted a wrong note then his career would have suffered - but little else. The consequences of engineers or indeed many professional practitioners such as medics, making a wrong decision can be loss of life.

Although Blockley does not appear to be claiming that engineering must involve critical activities whose outcomes determine life or death, the reliance of modern civilisation on engineering is often cited when defining the importance of this discipline (e.g. Willis et al, 2009). Just as Armstrong blasting a wrong note is unlikely to be matter of life and death, neither can a poor recording place anyone in mortal danger. However, the transferability of skills in engineering from life-critical to relatively trivial applications (e.g. between image processing systems for medicine and those for domestic television) suggests that engineering can exist within the apparently trivial. Therefore the application is not considered here to be a determinant of whether an activity is engineering.

If we cannot use application as an indicator of what engineering is, then how else can it be separated from science, craft and art? Lewin (1983) identified ‘Two Cultures’ of Arts and Sciences which dominate universities and proposes Engineering as a third. He observes:

> Though today engineering artifacts draw extensively on scientific knowledge, it has not always been so, and up to the end of the nineteenth century the methods of manufacture and natural science were quite distinct. Indeed the development of the steam engine, often quoted as a ‘product of science’, was the work of men like James Watt and George Stephenson who were predominantly skilled craftsmen and certainly not knowledgeable in scientific matters. Thus engineering can be said to predate science, which is still relatively young, and man’s development has been, and still is, determined essentially by his capacity to make artifacts and improve upon his environment rather than the systematic accumulation of knowledge.

The distinction that Lewin makes between the two is that science often (although not always) deals with systems which are ‘closed’, isolated from the real world, whereas engineering is concerned with open systems which must serve and operate within real-world situations. From this we might consider that engineering, in a recording situation, delivers ‘what is possible’ given what is known about laws of nature and the limits of technology, along with ‘what is desired’ according to aesthetic considerations. However there is still Lewin’s delineation of the ‘designer craftsman’ and ‘designing engineer’ to consider. He elevates the status of engineering above that of a skilled craft or trade in stating that “the design of artefacts such as a communication system, computer software system ... etc. require[s] a high level of intellectual skills rather than simple craft skills”. Rogers (1983) states that “craft is the power to produce a preconceived result by consciously controlled action: the craftsman always knows what he wants to make in advance” whilst the engineer will innovate and have to choose from a number of solutions to a problem.

Engineer is not the only title (rightly or wrongly) attached to the process of making sound recordings. Borwick (1973) describes the ‘Tonmeister’ concept, which he adopted as the head of the first degree in music and sound recording to be offered at a UK higher education institution, thus:

> A literal translation from the German produces ‘sound-master’ ... implying that a tonmeister is someone skilled in the arts of sound recording, transmission and reproduction. The complication, and the special delight, is that it calls for a strange mixture of artistic flair and technical knowledge. How you combine the seemingly incompatible aptitudes of art and technology in a single individual and in what proportions, has always been a matter for discussion – and even of heated argument.

This specialist term is very useful, not least because it began life (half a century ago) as a discussion about the skills and sensibilities that someone making recordings should possess. To the composer Schoenberg, the tonmeister “should be trained in music, acoustics, physics, mechanics and related fields to a degree enabling them to control and improve the sonority of recordings, radio broadcasts and sound films” (Maconie, 1984). In fact, music aside, this is a list of subjects that might be expected in many engineering courses. The following question then remains: what is the person who makes recordings of music required to do and does what they are required to do encompass the innovation and informed decision making that characterises engineering? At one extreme there is the existence of performance venues with microphones permanently rigged in one place (often positioned with reducing visual obstruction in mind, rather than optimising sound quality), connected to recording systems with a set of step-by-step instructions to enable recording by a minimally trained person. This example demonstrates that certainly not all recordings are engineered: in this situation no one is exercising choice or innovating, there is just the “preconceived result” which Rogers associates with craft. However it could be argued that the overall system, as embodied in the microphone positioning, the choice of system components may have been engineered to best match a specification: a system capable of making the best recordings given that the microphones must not visually obscure performances, their position must be permanently fixed, the operator of the system will have little or no training etc. But an individual who has knowledge of acoustics, is aware of the capability of the equipment at his/her disposal, is able to adapt to different recording situations, can weight the outcomes of different decisions against each other whilst understanding the musical intentions of performer(s) and producer is designing a recording, making choices and is often innovating. As the microphone designer Jörg Wuttke of Scheops has observed (1999):

> A person who understands something about the way microphones work will be well able to master the most diverse assortment of recording tasks. “Cookbook” approaches are certainly easier, and they can be in step with the latest fashion trends, but they are not much help in the constantly changing conditions of live recording. Recording is not easy; being a recording engineer is truly a profession

2. RECORDING ENGINEERING SCENARIOS

2.1 Time-domain processing of audio

The digital delay line forms the basis of many widely used audio effects. Chorus, flanging and automatic double tracking (ADT) are all effects based mixing a signal with a single, delayed version of itself. Many of these effects arose in the pre-digital era as a result of physical interaction with analogue magnetic tape or running two tape machines containing the same signal simultaneously at varying speeds. Phasing is sometimes also included in this group (e.g. Elen, 1994), but elsewhere it is considered a process which requires short, *frequency dependent* delays and therefore a network of all-pass filters rather than a simple frequency independent delay line (e.g. Cousins, 2007).

Here it is proposed that there are three levels which understanding and use of these effects can be grouped into. Firstly there is the ‘operator’ level, where there is an understanding of how to summon a preset flange or chorus and to access and adjust the parameters offered. At this level there is no understanding of how the effect is produced and, therefore, no understanding of how the adjustment of parameters affects the processing – there is only audition, memory and adjusting ‘to taste’. In fact this analogy to cooking can be continued, since these types of short-delay effects are often described as ‘thickening’ the sounds that they are applied to. At the second level there is the understanding that these ‘thickening’ effects are produced by the addition of a delayed version of the sound. It is understood that the delay is short enough that it is not perceived as a separate echo and that the movement within the sound (e.g. the ‘swooshing’ of the flanger heard at the first level) is created by dynamic variation, or modulation, of this short delay time. Armed with this understanding the recordist can follow ‘cookbook’ recipes to adapt a general purpose digital delay line to become a flanger or chorus unit; i.e. they do not need processors with preset configurations. Also, with the connection made between ‘movement’ and modulation a connection can be made between the parameters ‘modulation depth’ and ‘modulation speed’. At the third level is the engineer who is able to apply science (acoustics and psychoacoustics), along with the numeracy required for understanding linear superposition and the relationship between the time and frequency domains, to the problem of developing a particular desired texture for the sound that they are working with. This is someone who understands, via superposition, that a single feed-forward delay will produce a comb notch filter and that if there is feedback within the delay path then a peak comb filter will be produced. They realise, as the feedback level increases and the peaks in the frequency response of the system become more localised, that the increasingly 'metallic' quality which is attributed to the sound output is related to the fact that metal resonators tend to have strong, widely spaced modes due to the relatively low damping and faster speed with which sound travels in this medium.

At the first level there is an ability to operate a specific machine to produce a very narrow range of outcomes and to express a judgement or preference, at the second there is the ability to use a general purpose tool to arrive at a more adaptable set of outcomes and at the third there is the total understanding of the process from both causative and perceptual viewpoints. This offers ultimate adaptivity, in fact the engineer can move beyond set recipes and create, from scratch and as necessary, novel short-delay-time processes which fit the exact aesthetic requirements of the situation. These three levels might be labelled: operator, craftsperson and engineer levels of artifact creation. Of course all of these require audition as the final arbiter, but it should be the engineer who is able to present options of the highest quality and arrived at with the greatest expediency. To quote the famous words of Rayleigh (1878):

> Directly or indirectly, all questions connected to sound must come for decision to the ear, and from it can be no appeal. But we are not therefore to infer that all acoustical investigations are conducted with the unassisted ear. When once we have discovered the physical phenomena which constitute the foundation of sound, our explorations are in great measure transferred to another field lying within the dominion of the principles of Mechanics. Important laws are in this way arrived at, to which the sensations of the ear cannot but conform.

A Steinberg Virtual Studio Technology (VST) 2.3 plugin for Windows PC which implements a delay with sufficient flexibility for these types of effect to be explored is available from this author (Wells, 2012). This can be combined with a compatible application (such as *Audacity*) to process audio, and the output can be analysed within scientific computing applications such as *Matlab* or *Scilab* to demonstrate both time and frequency domain aspects of the processing. Notes on how this tool might be used are also included in this resource.

2.2 *Dynamic range control and the 'loudness war'*

This scenario concerns the mastering stage of recording post-production. It is typically the case (with the notable exception of legacy analogue material being released on CD) that the resolution of the production and post production media will be higher than that of the target distribution format. Also, the useful dynamic range of the domestic listening environment is likely to be much lower than that of acoustically isolated and otherwise optimised professional monitoring facilities. Therefore a common task at the mastering stage is to reduce the dynamic range and to adjust its overall level. The practice of mastering is discussed in detail in Katz (2002). In order to maximise the available resolution of the destination format the gain of the audio signal is adjusted such that its maximum level is approaching 0 dB_{FS} (i.e. 0 dB below full digital scale). In fact there is some complexity to this issue, which the recording engineer should be able to understand and negotiate in order to arrive at the best decision for the mastering situation and material to hand. On the one hand, since digital reconstruction filters within digital to analogue converters or oversamplers are integrators whose output might overshoot the level of two consecutive input samples, a maximum level of 0 dBFS might lead to clipping distortion later on in the signal path and, where oversampled level meters are not available, -3 dB_{FS} is regarded as 'safe' (Nielsen and Lund, 2003). On the other, it is possible that masking may obscure clipping distortion to the extent that a number of consecutive clipped samples may be inaudible and so setting the level such that it is clipped may not lead to audible distortion as a result of that clipping. A detailed summary of auditory masking can be found in Moore (2012). Here the engineer is well equipped to obtain the best outcome since an understanding of what is quite nuanced digital signal processing and psychoacoustics is required. The craftsperson will understand that there is some form of tradeoff between distortion and noise level and have developed a strategy through trial and error for dealing with this. The operator will either require some form of automated process with fixed parameters (such as level normalisation) or will simply adjust the level and audition the result; the former is unlikely to lead to the optimal solution, the latter will be time-consuming. Again, whilst the preference of the auditory system is the ultimate criterion, the optimum result for a given scenario, arrived at within a reasonable period of time, is more likely with the fuller understanding of the engineer. In fact, immediate audition may not be the best determinant of quality. It is proposed in some parts, e.g. Lund (2006), that there may also be an effect of long-term 'listening fatigue' to consider. However, it is important to note at this stage there has been very little research conducted in this area and it should not be confused with the established concept of 'auditory fatigue' (see Moore (2012) for a summary).

In all but the quietest listening environments the resolution of typical digital distribution media should be sufficient to deliver audio without audible quantisation noise even with levels set at -3 dB_{FS}. However it is often the case that the dynamic range of program material is too wide to be accommodated by that of the typical listening environment and so it must be compressed. In conjunction with level adjustment so that the signal peak is just below clipping level, compression has the effect of raising the RMS signal level. Perceptually, the loudness of a signal is closely related to its short-time (local) RMS, although this relationship is complex and is not consistent between individuals (again, see Moore for a summary). Loudness has been seen as a way of gaining listener attention and enhancing gratification; as a result of this some perceive that a 'loudness war' is underway (e.g. Vickers, 2010) and this is an area of some considerable debate at all levels (e.g. Deruty, 2011). Whatever the merits are, maintaining signal quality whilst reducing its dynamic range is quite a challenge requiring expert control of non-linear devices (compressors and limiters): there are a number of possible solutions and each will require insightful analysis due to the interactions of complex program material and non-linear processing. Our operator will only be able to access pre-configured compression patches within digital processors. They are unable to negotiate between the different ways in which compression might be applied (e.g. via additional processing of the variable compressor gain): the signal can only get louder in one or two different ways. The craftsperson will have experience of dealing with the problem in a particular way which, via trial and error, they have found to work well on a particular kind of source material. The engineer will understand and be able to quantify the tradeoffs, such as shorter attack/release times for better following of the signal versus amplitude envelope distortion, modulation and (possibly) aliasing effects. Because they are able to understand how limiting, compression and clipping processes will affect the signal they focus audition on to these areas to check whether unwanted artifacts are becoming audible. More plainly: engineers can know what to listen out for and where to focus their attention.

2.3 Microphone array design

This final scenario concerns the positioning of microphones for recording an ensemble at-once (i.e. not overdubbed instrument by instrument) in an acoustically ‘live’ environment (e.g. concert hall as opposed to recording booth). Here the design requirement is very often that the microphones used should be selected and positioned such that the character of the space and the internal balance of the ensemble is retained in the reproduced sound. Beyond this the requirements are more aesthetic-orientated: precise translation of instrument position to the stereo image produced between the loudspeakers versus a sense of envelopment and extended low frequency response and so on. Beginning with the expression of some preference in this regard (such as from the producer or performing musicians, or from their own ear) the task is then to design a microphone array, such as a main pair, or tree, with spot microphones, that meets these requirements (which quite possibly conflict to an extent). There is a rich array of high quality literature on this subject, from Blumlein’s patent of coincident dipoles for accurate positioning of sources between loudspeakers through to Jecklin’s baffled omnidirectional microphones, much of it collected in Eargle (1986). More recent work, such as that by Martin (2005) in developing arrays for surround (e.g. 5.1) recording is also highly relevant to the understanding of arrays for two-channel stereo. There is controversy in this area see, for example, Lipshitz (1986). There are ‘off the shelf’ stereo microphones which the operator can use. One example is the Rode NT4 which has a fixed configuration of two coincident cardioids at 90 degrees to each other. This spacing is conducive to a good front/back rejection ratio but there will be considerable correlation between both microphones for many signals, leading to a narrow image. Therefore it is a safe configuration – not much can go wrong with it (e.g. no out-of-phase regions) but it will only represent the optimum in a few situations. The craftsperson will be more comfortable selecting individual microphones and using these in established recipes for stereo capture such as the Blumlein pair or Decca Tree. Of course, engineers should adopt already proven solutions where they are optimal and it can certainly be argued that there is no better stereo technique than Blumlein’s pair for imaging accuracy. However, an understanding of how that technique actually functions enables adaptation to a more optimised system for the task at hand (for example dipole microphones tend to have a poor at-distance low frequency response, and such an array which is equally sensitive in all directions may well not be suited to an overly reverberant space or one with an audience in close proximity to the performers). The engineer can adapt existing solutions to fit the specifics of the scenario at hand, whereas the craftsperson will perhaps select between a more discrete range of pre-conceived options and tend to select those with which they are most familiar. The operator can merely apply ready-made solutions with modification possible only by trial and error. One area of interest is the differences in spatial perception for a mono signal panned by either level or time differences. The vast majority of panning systems offer only level difference positioning. A VST panner which offers both time and level difference simultaneously is available from the author (Wells, 2012).

3. SUMMARY

Three examples of where engineering as a discipline can be observed in the sound recording industry have been presented with reference to the relevant literature. For two of the scenarios there is demonstration software available, along with teaching notes. These examples have identified three ‘levels’ of creative and intellectual involvement which relate to popular theories and opinions about the nature of engineering. It is hoped that this paper will serve as a useful starting point for those who wish to explore how recording can embody engineering and how engineering can serve recording.

5. REFERENCES

Blockley, D., 2008. Integrating Hard and Soft Systems. *Proceedings of the Royal Academy of Engineering Workshop on Philosophy and Engineering* (*WPE2008*).

Boehm, C., 2007. The discipline that never was: current developments in music technology in higher education in Britain. *Journal of Music, Technology and Education*, 1/1, 7-21.

CBI (Confederation of British Industry), 2008. Taking Stock: CBI education and skills survey 2008. URL:http://www.investorsinpeople.co.uk/Documents Hidden/Research/CBI education and skills survey.pdf

Cousins, M., 2007. Chorus, flange and phase. *MusicTech Magazine*, September, 56-57.

Deruty, E., 2011. ‘Dynamic Range’ and the Loudness War. *Sound on Sound*, September

Eargle, J., (ed) 1986. *Stereophonic Techniques.* New York. The Audio Engineering Society

Elen, R., 1994. Sound processing. In: Borwick, J., ed. *Sound Recording Practice*. 4th ed. Oxford. OUP, 204-229.

Katz, B., 2002. *Mastering Audio: the art and the science*. Waltham, MA. Focal Press.

Lewin, D., 1983. Engineering Philosophy – The Third Culture. *Leonardo*, 16/2, 127-132.

Lipshitz, S., 1986. Stereo Microphone Techniques: Are the Purists Wrong? *Journal of the Audio Engineering Society*, 9, 716-744.

Lund, T., 2006. Stop Counting Samples. *The Proceedings of the 121st Audio Engineering Society International Conference*. Paper number 6972.

Maconie, R., 1984. Sound Recording in a Music Department. *International Journal of Music Education*, 4, 51-56.

Martin, G., 2005. A New Microphone Technique for Five Channel Recording. *The Proceedings of the 118th Audio Engineering Society International Conference*. Paper number 6427.

Moore, B., 2012. *An Introduction to the Psychology of Hearing*. 6th ed. Bingley. Emerald.

Nielsen, S. and Lund, T., 2003. Overload in Signal Conversion. *The Proceedings of the 23rd Audio Engineering Society International Conference*. Paper number 11.

Open University, 2007. *Contexts, Conquests and Consequences - A Contemplation of Engineering*. Milton Keynes. Open University Worldwide.

Robson, D., 2012. Make Engineer a Protected Title: HM Government e-petition. URL: http://epetitions.direct.gov.uk/petitions/6271

Rayleigh, Lord, 1878. *The Theory of Sound*. Reprinted New York 1945. Dover.

Rogers, G., 1983. *The Nature of Engineering*. London. The MacMillan Press, 1983.

Vickers, E., 2010. The Loudness War – Background, Speculation and Recommendations. *The Proceedings of the 129th Audio Engineering Society International Conference*. Paper number 8175.

Wells, J., 2012. URL: www.jezwells.org/engineering_recording

Wuttke, J., “The Microphone: Between Physics and Emotion”, 107th Convention of the Audio Engineering Society, September 1999, Preprint No. 4991.