A THEORETICAL FRAMEWORK FOR AUDIO PRESERVATION[7]

Marcos Sueiro Bal, New York Public Radio, New York, USA

"You've got to help me remember"
—The Replacements, "Never Mind"

Abstract: Drawing parallels from the world of telecommunications, we suggest that only two factors affect the information extracted from an audio carrier: the intrinsic signal to noise ratio and the resources available for extraction. We then explore ways in which these factors vary over time, and suggest directions for future research.

Audio archivists have two basic concerns: (a) retrieval of past audio information and (b) transmission of audio information into the future. In this respect one could say we have strong parallels with the telecommunications business, whose engineers "must develop a signal from a source and deliver it to a sink to the satisfaction of a customer."[8] In our parallel world one might say we play the roles of both listener and broadcaster: listeners are concerned with retrieving a distant signal, while radio stations are interested in broadcasting a signal towards points a certain distance away. In other words, telecommunications tries to ford a spatial distance, while audio archiving attempts to do the same across a temporal distance.[9] In order to ford these distances, the information is encoded (modulated) in a carrier and decoded (demodulated) at the other end, and that decoded information is presented to an end user.

The present article suggests a theoretical framework for the process of extracting audio from a carrier in a preservation context, i.e., role (a) above, as presented in Fig.1. It posits that the quality of a retrieved signal depends on only two factors: (1) its inherent quality and (2) resources available for signal extraction. Both factors vary across (3) time. Let us examine how these three factors may interact.

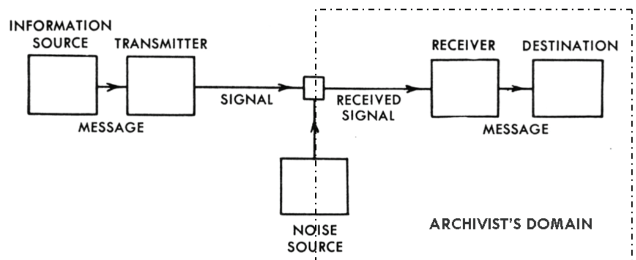

Fig. 1. The Archivist's Domain (after Shannon)

7 This article developed from a presentation at the 2012 Association for Recorded Sound Collections Conference, and I thank Mike Casey of Indiana University for inviting me to speak there. The entire article is pretty much a refinement of Fig. 1 in Richard Hess' "Tape Degradation Factors and Predicting Tape Life", Convention Paper 6970 of the Audio Engineering Society's 121st Convention (2006).
I would like to thank Andy Lanset, Director of the Archives at New York Public Radio, for generously allowing me to spend time to write this article. The staff at the New York Public Library and Dan Biddle at the Library of New York University's IFA Conservation Center were especially helpful.

8 Freeman, Roger L. ***Telecommunications Transmission Handbook***. 2nd ed. New York: John Wiley & Sons, 1981, p.1.

9 Gleick, James. ***The Information: A History, A Theory, A Flood***. New York: Pantheon, 2011, pp. 8-9.

1. Signal quality vs. time

All archivists, not just audio archivists, are faced with information loss over time. This information loss is ultimately due to the change in some physical or chemical aspect of the carrier: dyes fade, paper becomes brittle, magnetism is rearranged. Moreover, over time the chance of sporadic damaging events increases.[10]

The combined effect of these irreversible changes is to reduce the Signal to Noise Ratio (SNR), which in a band-limited "channel" (such as an audio carrier) strongly correlates to the maximum retrievable information.[11] (In this context "noise" can encompass not just what we think of as traditional audio noise, but any type of distortion—in essence the reduction in likelihood of correct retrieval.[12]) Once information is encoded in a carrier, SNR decreases with time. Eventually, when SNR=0, the amount of available information is zero. Assuming a linear decrease, then

$SNR(t)=SNR_0-\alpha t \quad / \quad 0 \text{ if } H(t)>=I_0$ [1]

Where

$SNR(t)$=Signal to noise ratio at time $t>t_0$
SNR_0=Signal to noise ratio at the time of encoding
α=attenuation coefficient
t=time

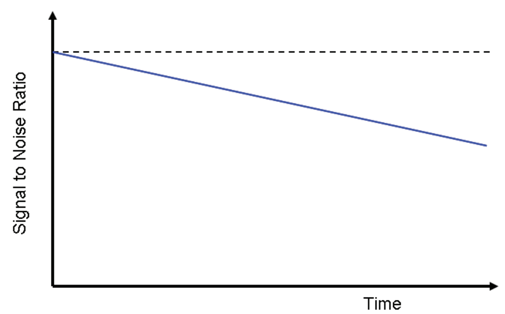

Fig. 2

Calculating the attenuation coefficient α would of course be useful. Telecommunications engineers routinely calculate similar coefficients (which are usually frequency- and temperature-dependent) per unit of length by comparing a signal at the start of a medium with the signal at the other end of the line.[13] But in this respect telecommunications engineers have two significant advantages over archivists. Firstly, they can easily measure the original signal, while archivists seldom have such a luxury: the most we can usually hope for is having reference signals of a known age, which can take such forms as test tones or error levels. Secondly, telecommunications engineers can choose well-known, homogeneous media within controlled

10 MacDonald, Lindsay, and Ralph Jacobson. "Assessing Image Quality". In: ***Digital Heritage: Applying Digital Imaging to Cultural Heritage***. Lindsay McDonald, ed. Amsterdam: Elsevier, 200, p. 353.
11 Shannon, Claude E. "A Mathematical Theory of Communication". ***Bell System Technical Journal*** 27 (3): 379–423. (July–October 1948)
12 Cover, Thomas F. ***Elements of Information Theory***. 2nd ed. new York: Wiley, 2006, p. 13
13 Freeman, p. 14

environments to test, whereas the particular chemical composition and the history of a carrier or a collection are often unknown (although this uncertainty can be reduced by testing many carriers). Signals (like test tones) whose quality was registered at the time of encoding can help archivists can get a sense of signal loss by allowing them to compare the present signal with the (presumed) original; accelerated aging tests fall into this type of test. It is important to note that such coefficients do not take into account sporadic events which may further reduce the SNR. Expression [1] is thus a best-case scenario.

2. Signal quality vs. cost

Unlike books and still images, audio and video materials need an extra layer of decoding before they can be understood by a human user. This decoding essentially takes the form of a signal measurement using a specific device, and this adds an additional layer of complexity; in fact, the extraction point is often the weakest point in the information chain from creation to presentation.

In general, measurement tries to maximize accuracy (lack of bias) and resolution, and the latter is closely related to the signal to noise ratio.[14] Accuracy and resolution are affected by the quality of the measuring instrument, the skill of the operator, and the time and care spent on making the measurement. More experienced operators, better instruments, and more time spent when making a measurement will translate into higher accuracy and resolution, but also higher costs. Audio extraction operates along similar principles.

If detected, simple bias can usually be corrected exactly, but (assuming linear encoding or the entire communication channel(s)) the upper limit on SNR at time of retrieval is the SNR of the stored signal at that time: "the quality of the signal heard by the listener depends very largely on the nature of the signal present at the receiver."[15] We can infer that additional cost (resources) will result in improved SNR (perhaps even arbitrarily small), but the SNR will not transcend this limit. For analog sources, the upper limit is approached logarithmically, while for digital sources the upper limit can be reached.

$$SNR(c)=SNR(t)-[1/Rc] \qquad [2]$$

Where

c= cost
SNR(t) = signal to noise ratio at time ***t***
R=resource factor, expressed in dB per currency (e.g. $)

There are some potentially interesting consequences from expression [2]:

1. Since noise is additive, you should try to minimise *1/Rc*
2. It may be useful to try to estimate the resource function *R*. (We again present this kind of factor as a constant for simplicity.) Given a certain machine and an operator, how many dB of SNR do you gain by spending two hours instead of one to retrieve a signal? How about if you buy a machine twice as expensive? Or an operator twice as pricey? The answers to these questions will likely vary with each specific situation.
3. There is a minimum cost of extraction below which the SNR is zero and thus no extraction can be accomplished. This seems to correlate with real-life experience.

14 Lathi, B.P. ***Modern Digital and Analog Communications Systems***. 2nd ed. Philadelphia, Pa.: Holt, Rinehart and Winston, 1989, pp.558-586.
15 Hood, John Linsley. "Tuners and Radio Receivers". ***Audio & Hi-Fi Book***. 3rd ed. Ian R. Sinclair, ed. Oxford, U.K.: Newnes, 1998.

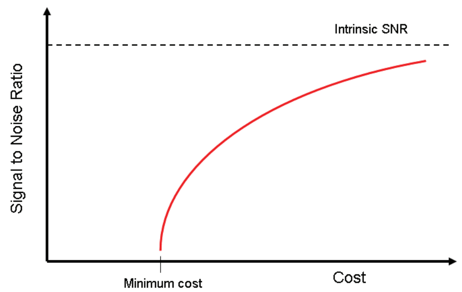

Fig. 3

3. Cost vs. time

Resources available for playback over time will vary from situation to situation, but we can try estimating an average cost over time. Audio formats tend to have short periods of market dominance.[16] If we let t=0 be the point at which a format is introduced and t_p the point in its history of peak popularity, the resources necessary to retrieve a signal with a certain resolution may generally follow a curve such as the one below, which approximates the inverse of sales of most formats. Sales of a specific format may generally mimic the general availability of resources (such as expertise or maintenance) necessary to play back that particular format. Primary markets well past their peak can become quite erratic, since monopolies and oligopolies have different incentives than competitive markets. In our case, we have estimated that after the format's peak popularity costs eventually grow almost linearly with time.

$$R(t) = R\,[(t^2 + t_p)/t] \qquad [3]$$

Where

R(t)=Cost factor at time t
R=Cost factor
t = time

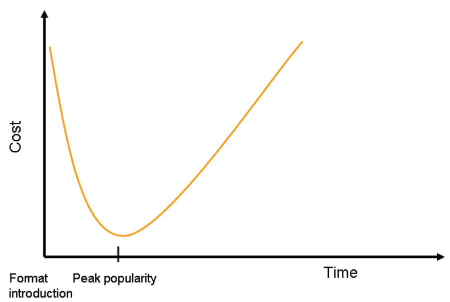

Fig. 4

16 Franklin Templeton Investments. "Innovation Will Surprise Us… Again", in "The Case for Equities". https://www.franklintempleton.com/funds/2020vision/innovation#.Uz7p1qLfQrc. Retrieved April 2013.

It is interesting to note that, given a certain level of available resources, extracting the signal may become impractical in the future, as we distance ourselves from the peak of popularity for that format. This is not because the signal has intrinsically diminished in quality (although it probably also has), but rather because it is more difficult to extract. At one point, it either becomes impractical to extract the signal or you need to allocate more resources to extract it. This is the problem commonly called "format obsolescence".

Combining all equations [1]-[3], we get the maximum SNR of a recorded signal at a specific time t with resources c:

$SNR=SNR_0-at-[t/(t^2+t_p)c]$ [4]

Where

SNR = signal to noise ratio (dB)
SNR_0= original SNR of encoded signal at time t_0 (dB)
a = attenuation factor (dB/year)
t = time since the introduction of the format (years)
t_p = time since peak popularity of the format (years)
c = cost (currency)

An interesting exercise is to combine all these factors in one graph to get a sense of the quality of the signal extracted, at a certain point in time, with a certain budget. This is represented graphically below.

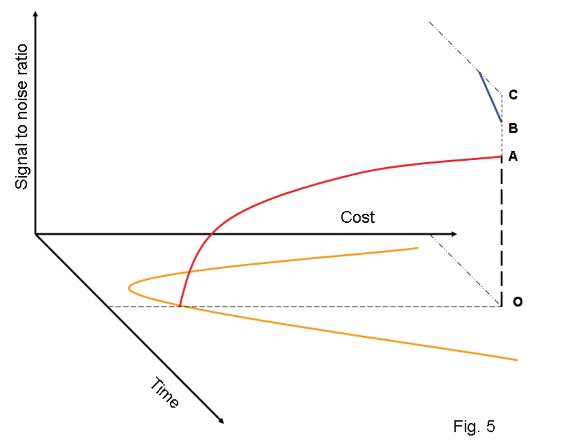

Fig. 5

Fig. 5 shows the SNR achieved at a certain cost from a hypothetical carrier at a certain point in time relatively close to the carrier's peak popularity. The carrier was recorded at the peak of popularity of the format. Segment OA is the achieved SNR. Segment AB is not achieved due to lack of resources, and segment BC is not achieved due to the intrinsic SNR of the carrier at that point.

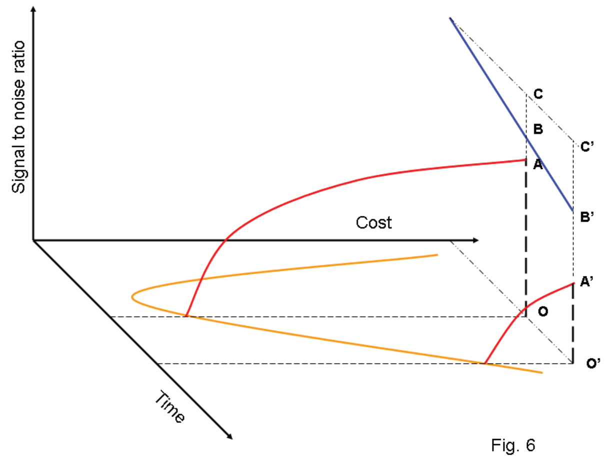

Fig. 6

Fig. 6 shows the SNR achieved at the same cost but long after the format has disappeared from the market. Two factors are conspiring against the later sound archivist: intrinsic signal quality has diminished (segment B'C' is the additional SNR loss), and the effort necessary to extract it with a certain resolution has increased (which results in the loss represented by segment A'B'). The signal extracted is O'A', and it is evident that, with the same resources, we can extract far less signal today than we could at the time of peak popularity for that particular format. In other words, you are getting far less bang for the buck. Since this is the trend for *any* format, it stands to reason that delaying signal extraction amounts to a less effective use of resources.

However, all of these predictions are not yet quantified. We need to establish a solid basis to estimate long-term signal extraction.[17] We stand at an interesting point in time, as there are several mass audio digitization projects around the world whose metadata could be used to establish a basis for our path moving forward. It is my hope that it will be so. Until then, our knowledge of content degradation will come only from anecdotal evidence.

17 van der Reyden, Dianne. "The Science-Based Fight Against 'Inherent Vice'". ***International Preservation News***, No. 50 (May 2010), p. 5.