Chapter 5
Authenticity Assessment

In certain cases, there may be a question about the *authenticity* of an audio forensic recording. Like any physical evidence, audio forensic recordings are subject to potential questions about authenticity: is the recording complete, unaltered, and consistent with the stated circumstances of its creation? For example, an individual may claim that a recorded conversation has been edited so that certain critical utterances are inserted or edited out. Other cases may involve suspicion that the asserted time, place, and circumstances are not what was claimed (Audio Engineering Society 2000). What is authenticity? Can it be guaranteed?

Recordings are always susceptible to accidental alterations or deliberate tampering, and detecting these changes may or may not be possible. The court must be convinced of the authenticity and integrity of the audio evidence. Audio forensic examiners must follow chain-of-custody procedures and avoid any possibility of unintended changes to the original evidence and must be diligent about potential signs of alteration. The court must also understand that the fact that an examiner does not find specific evidence of tampering does NOT necessarily mean that the audio recording is authentic: a particularly skilled adversary could conceivably create a tampered recording that defies detection.

5.1 Historic Context: Authenticity of Analog Magnetic Tape Recordings

Until the first decade of the twenty-first century, the primary medium for audio forensic evidence was analog magnetic tape. With the exception of a few mechanical recording systems, such as the dictabelt system used in the Dallas Police Department at the time of the John F. Kennedy assassination, magnetic tape was essentially ubiquitous as the means to capture live audio.

© Springer Nature Switzerland AG 2018

R. C. Maher, *Principles of Forensic Audio Analysis*, Modern Acoustics and Signal Processing, https://doi.org/10.1007/978-3-319-99453-6_5

Magnetic tape consists of a thin, flexible plastic ribbon that serves as a substrate for a thin layer of magnetic powder material impregnated in a binder substance and spread uniformly onto one side of the tape. A magnetic field deliberately brought into the vicinity of the tape can magnetize the surface material, leaving a telltale magnetic polarization. Later on, a magnetic detector circuit can measure the amount of magnetization of that particular segment of the tape.

The magnetic tape is stored rolled onto a spool, known as a reel. The magnetic tape recorder draws the tape off the supply reel at a fixed rate using a motor-driven capstan spindle and pinch roller. The tape path slides over three electromagnetic coils: the *erase head*, the *record head*, and the *playback head* (some less expensive tape recorders used just two heads: the erase head and a combined record/reproduce head). When recording, the moving tape first passes over the erase head to randomize the magnetic domains on the tape and then continues and passes over the small coil of wire in the record head that acts as a variable electromagnet. The electrical current through the record head's electromagnet is modulated by the analog audio signal, causing fluctuating magnetization of the tape to represent the audio information. The tape transport then collects the recorded tape and spools it onto a separate take-up reel.

To playback the recorded information, the tape is first rewound from the take-up reel back onto the supply reel. Next, the tape is passed again from the supply reel to the take-up reel, but the erase head and record head are not activated, while the playback head detects the magnetic field on the tape and regenerates the analog audio signal. Once the tape has been recorded, it can be played back repeatedly at will, with only gradual losses due to mechanical wear as the tape is moved through the player.

The relationship between the record head current and the magnetization of the tape is nonlinear, which causes distortion. To minimize this inherent distortion, a strong, inaudibly high-frequency (e.g., 40 kHz) AC *bias* signal is added (mixed) with the audio signal. The bias signal causes the tape to be strongly magnetized at the bias frequency, with a small residual magnetization being the audio signal component. The playback system reproduces only the audio frequency range, so the ultrasonic AC bias linearizes the overall behavior without interfering with the audible information.

Audio tape recorders can have multiple parallel magnetization heads to create multiple longitudinal *tracks* on a single tape. Multi-track recordings for consumer products typically have two tracks, corresponding to the left and right stereo channels. Common consumer devices, such as compact cassette recorders, also allow interleaved tracks: one stereo pair is recorded on two tracks of the tape, then the tape can be flipped over, and a second pair of left and right tracks recorded with the tape moving in the opposite direction. A few common tape track configurations are shown in Fig. 5.1.

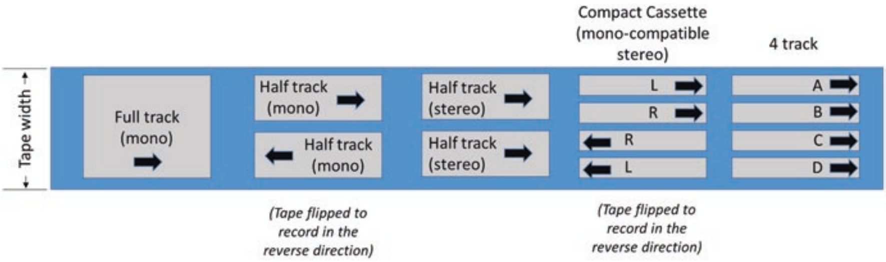

Fig. 5.1 Track format examples for analog audio magnetic tape

5.1.1 Physical Inspection

Audio forensic examination to assess the authenticity of an analog magnetic tape recording requires physical handling and examination of the tape itself (Audio Engineering Society 2000). Analog tape alterations made by physically cutting and then reattaching the tape, known as *splice* edits, involve adhesive tape used to hold the ends of the cut tape together.

The examiner will visually inspect the cassette housing, the reels, the entire length of the tape, and any related material, looking for spliced tape, broken housing, or other indications that the tape has been altered physically. The examiner will record any manufacturing serial numbers and tape batch designations, determining if the age of the tape is at least as old as the date the recording was reportedly made.

If the recorder used to produce the recording is available, the examiner inspects and tests the device. A qualified tape recorder technician can examine the track configuration, head alignment, azimuth setting, bias level, and so forth. If the recorder was out of calibration, it may be necessary to set up the playback head so that it matches the tape's alignment.

5.1.2 Magnetic Development

Authenticity evaluation of analog magnetic tape typically requires *magnetic development* to view the latent magnetic domains recorded on the tape. Magnetic development uses a ferromagnetic fluid (ferrofluid), which contains microscopic magnetic particles suspended in a solvent mixed with a surfactant to help keep the particles dispersed and suspended. The examiner spreads the ferrofluid uniformly, but sparingly, on the magnetic tape, which allows the suspended ferro particles to align with the invisible magnetic domains recorded on the tape. After allowing the solvent to evaporate, the examiner uses a microscope to observe the pattern of the magnetic particles adhering to the tape, known as *Bitter Patterns*, after Francis Bitter (1902–1967), a researcher at Westinghouse Electric Company and later MIT, who proposed the powder pattern method in 1931.

The erasing and recording process of analog tape creates a distinctive magnetic pattern when the recorder is started and stopped. The magnetic heads are energized as the capstan and reel motors start transporting the tape through the recorder, and the transient start-up magnetic fields leave a corresponding trace in the magnetic recording tape. Similarly, when the recording is stopped, the tape comes to a halt as the erase and record heads are de-energized. An example magnetization pattern is shown in Fig. 5.2. The image shows two regions of recorded material on a piece of analog cassette tape, caused by a stop/start recording sequence. The portion on the right is magnetization from the recording process up until the recorder was stopped, leaving the unmagnetized (dark) gap. Then the recorder was started again, which caused a slight offset of the magnetic pattern as the tape started moving again. The vertical striations in the magnetic patterns are due to the high-frequency AC bias mentioned previously (Koenig 1990).

The audio forensic examiner looks for the distinctive erase and record head magnetic signature patterns on the tape, as well as the magnetic tracks containing the audio information. If the recording is authentic, the examiner expects that there is a single start-up transient at the beginning of the recording, and then no other head transients until the recording was stopped. Any observation of additional start-stop sequences or erasures could indicate that the tape has been altered, deliberately or inadvertently, and the investigators would need to seek an explanation for why the recording appeared to be edited or truncated.

As was noted previously in the Watergate tape study, the investigators identified several overlapping erasures performed with a specific model of tape recorder that differed from the device that produced the original recording based on the characteristic start/stop magnetic signatures present on the June 20, 1972 tape (see Fig. 5.3).

In recent years, the use of a multi-track tape recorder to read information from the original track format on the tape has been used (Begault et al. 2005). Also,

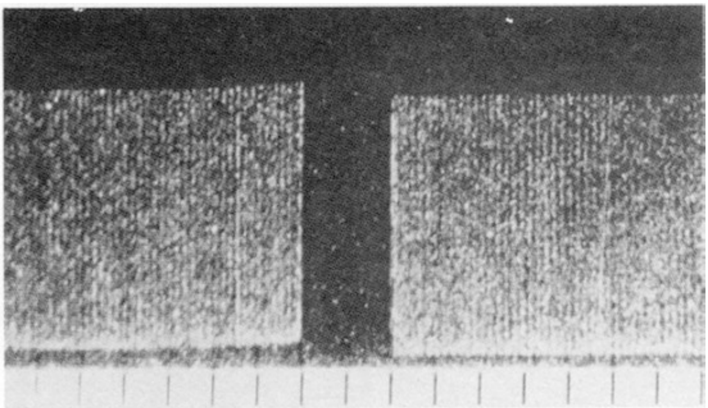

Fig. 5.2 Example magnetic domain "Bitter Patterns" magnified from an analog audio tape recording. The dark gap in the middle of the photograph shows where a recording had been stopped and then re-started (from Koenig 1990, reprinted by permission)

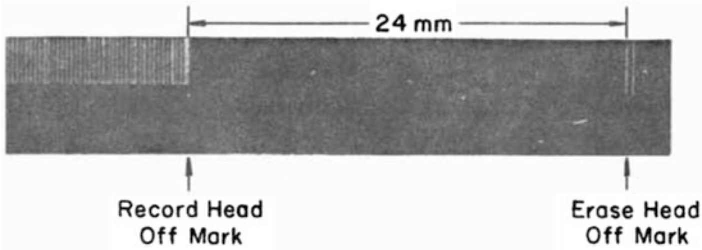

Fig. 5.3 Example of magnetic development as used in the Watergate tape investigation report (from Advisory Panel on White House Tapes 1974)

several high-resolution direct imaging methods have been developed using specialized equipment to reveal the recorded magnetic pattern without the use of the ferrofluid (Marr and Pappas 2008).

Some authenticity issues with analog tape recordings may be due to an edited tape that is subsequently copied and then presented as if it were an original recording. The copy can even be made with a single start and stop record sequence, so it may appear to be a continuous, authentic recording. In such a case, there may still be other evidence of tampering detectable by signal irregularities or gaps, as described below.

5.2 Current Context: Authenticity of Digital Audio Recordings

A digital audio recording presents many challenges for authenticity assessment (Brixen 2007). Digital audio recordings are essentially sequential lists of binary numbers stored in a digital computer file, and digital files can be copied, transmitted, and stored on a variety of media with perfect fidelity. What's more, it is often difficult to exclude the possibility that a digital file was adjusted and edited surreptitiously and then stored as a seemingly intact and pristine file. If all that is available is the digital audio file itself, the examiner must use other means to assess the integrity of the recording.

5.2.1 Identifying Edits: Splicing and Mixing

An audio forgery could consist of one or more edits made to an original recording by deleting certain time segments, by inserting audio material, or by additively mixing in the forged material. An unsophisticated forger could attempt to make such edits in the digital audio file with an abrupt insertion or deletion, often referred to as a *butt splice*. If the butt splice occurs at a point in the audio recording that is nearly silent, the butt splice edit may be essentially inaudible, but if the splice occurs during a louder passage of the recording, there may be telltale audible effects and discontinuities. Nevertheless, there may be detectable signal alterations due to the splice that can be observed in the waveform and/or the spectrogram even if there is minimal audible effect.

For example, consider the example recording of Fig. 5.4. The upper panel shows the time waveform, and the lower panel shows the spectrogram. The recording includes some speech utterances and background sounds.

If a forger wanted to remove a particular section of this recording with a butt splice edit, such as the portion of the recording indicated with dashed lines in Fig. 5.5, the result is depicted in Fig. 5.6.

Note that because the edit left an abrupt discontinuity in the waveform, the resulting waveform and spectrogram show evidence of a "click" in the signal. At this time scale, the effect is most easily seen as a vertical line in the spectrogram: the abrupt splice discontinuity in the waveform exposes spectral energy at all frequencies for a brief instant.

Zooming in on the butt splice, it is possible to see the abrupt change in the signal waveform caused by the deletion, as seen in Fig. 5.7. The abrupt change in the time domain waveform leads to the spread of high-frequency energy at the corresponding point in the spectrogram.

While this example makes it appear easy to identify a possible butt splice edit, a more skilled forger would conceal the edit by choosing the edit point in order to minimize the signal discontinuity and by using a short *cross-fade* instead of the butt splice. The cross-fade means overlapping a few samples from before the edit and a few samples after the edit and tapering the amplitude to blend the samples, thereby reducing the likelihood of a noticeable discontinuity at the splice point.

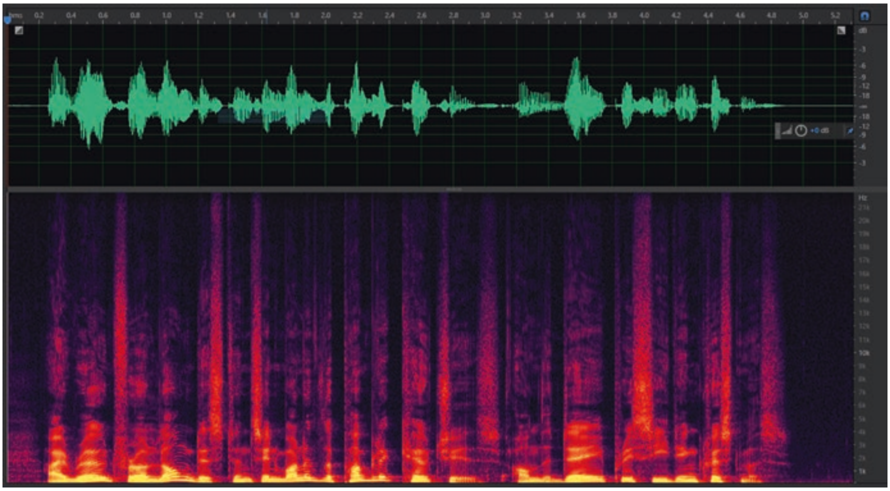

Fig. 5.4 Example audio recording of speech. Upper panel: signal waveform. Lower panel: signal spectrogram (overall duration 5.3 s, frequency range 0–22 kHz, linear scale)

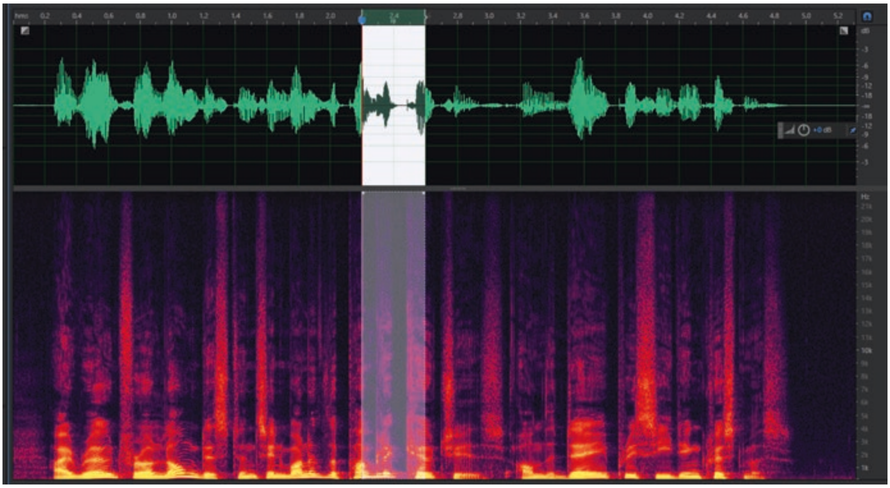

Fig. 5.5 Signal portion (0.4 s) to be removed by butt splice deletion (overall duration 5.3 s, frequency range 0–22 kHz, linear scale)

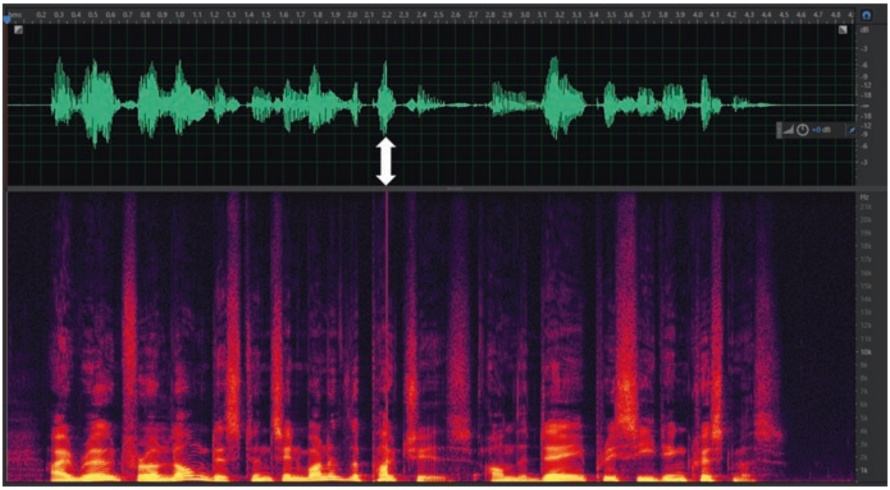

Fig. 5.6 Signal after highlighted portion of Fig. 5.5 is removed (overall duration 4.9 s, frequency range 0–22 kHz, linear scale)

Performing the same deletion edit depicted in Fig. 5.5, but with a 2-ms cross-fade instead of the butt splice, conceals the waveform effect in this example, as shown in Fig. 5.8 and enlarged in Fig. 5.9.

A forger may also attempt to introduce new material into an existing recording, either by opening up a gap in the file for the insertion or by additively mixing the contrived material into the recording. As with the deletion, the boundaries of the insertion could be a butt splice or a concealed cross-fade. If the forger uses skill and care, the edit points may be virtually undetectable in the waveform itself.

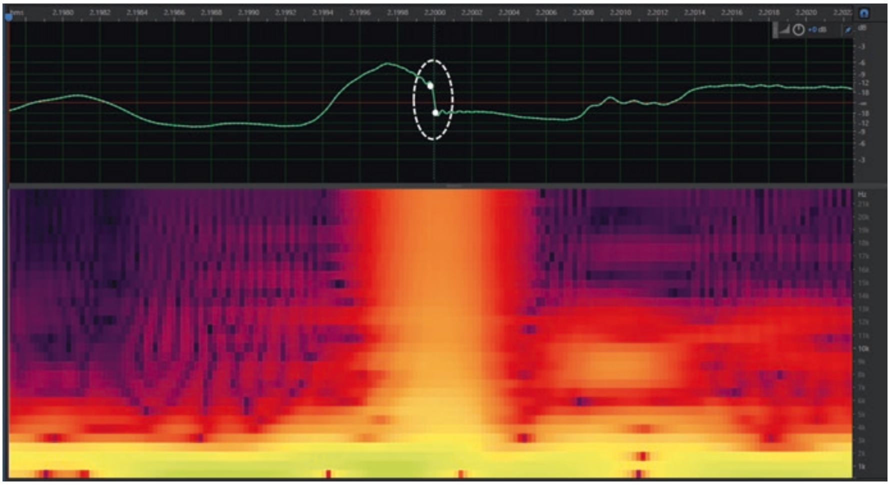

Fig. 5.7 Enlargement of time interval containing the butt splice discontinuity (overall duration 4.8 ms, frequency range 0–22 kHz, linear scale)

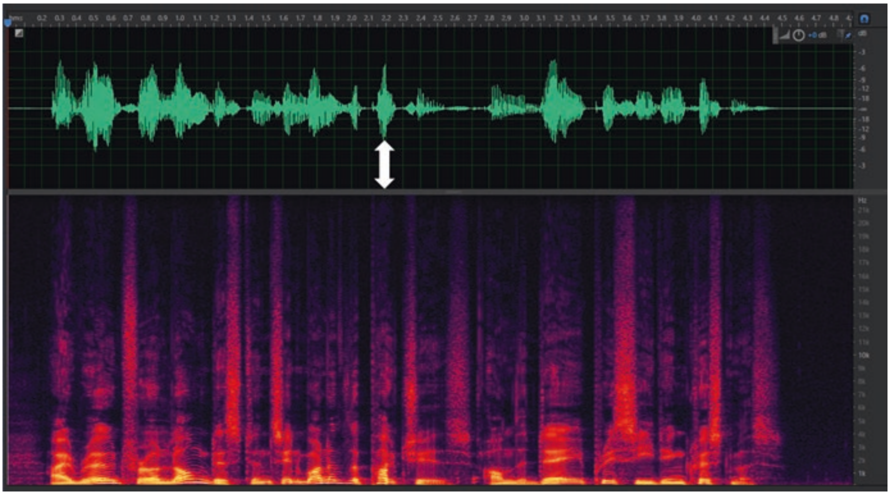

Fig. 5.8 Edit region of Fig. 5.5, but with 2 ms cross-fade instead of the simple butt splice (overall duration 4.9 s, frequency range 0–22 kHz, linear scale)

5.2.2 Other Authenticity Observations

While a smooth edit may reduce the likelihood of first-order detection of an alteration, there may still be signal observations that could raise questions about authenticity. These include background sounds, reverberation, and other acoustic information present in the recording.

In assessing a continuous recording, the audio forensic examiner can observe the acoustic reverberation and background sound level and detect whether there are any

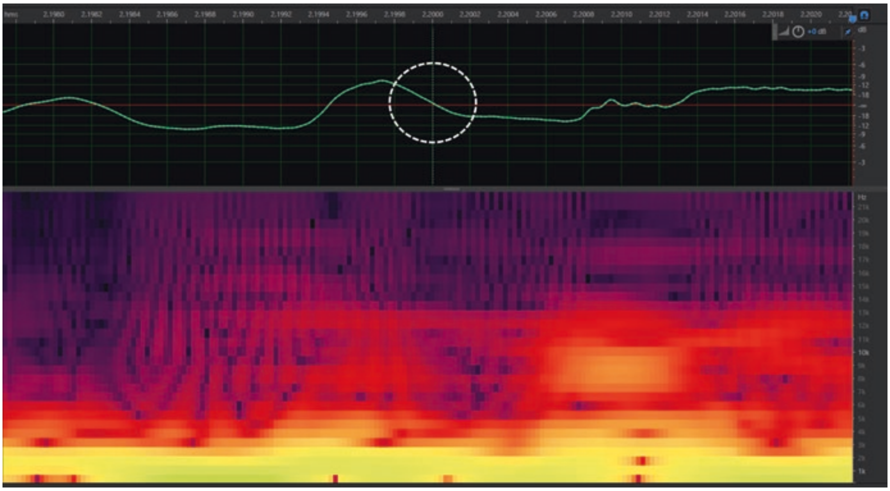

Fig. 5.9 Enlargement of time interval of Fig. 5.8 containing the cross-fade edit (overall duration 4.8 ms, frequency range 0–22 kHz, linear scale)

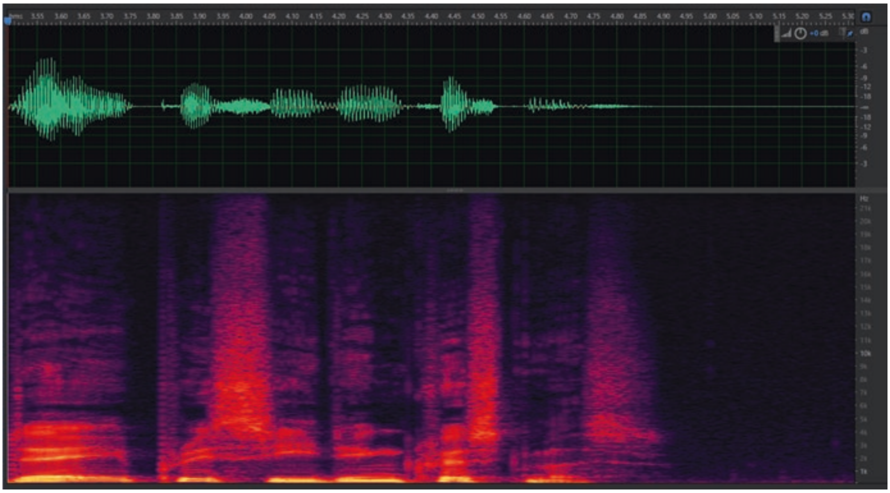

Fig. 5.10 Speech recording with little reverberation (overall duration 1.8 s, frequency range 0–22 kHz, linear scale)

unexplained changes in the background characteristics that could indicate a deletion or an insertion. As noted previously, a recording microphone picks up the direct sound of a source, such as a person talking, but also picks up the acoustic reflections of that sound source from the floor, walls, and other nearby surfaces. The microphone will also pick up any other sounds in the recording environment, such as wind, doors closing, mechanical sounds and alarms, etc.

For example, Fig. 5.10 shows a segment of speech recorded in a room with little reverberation, often referred to as a "dry" recording environment.

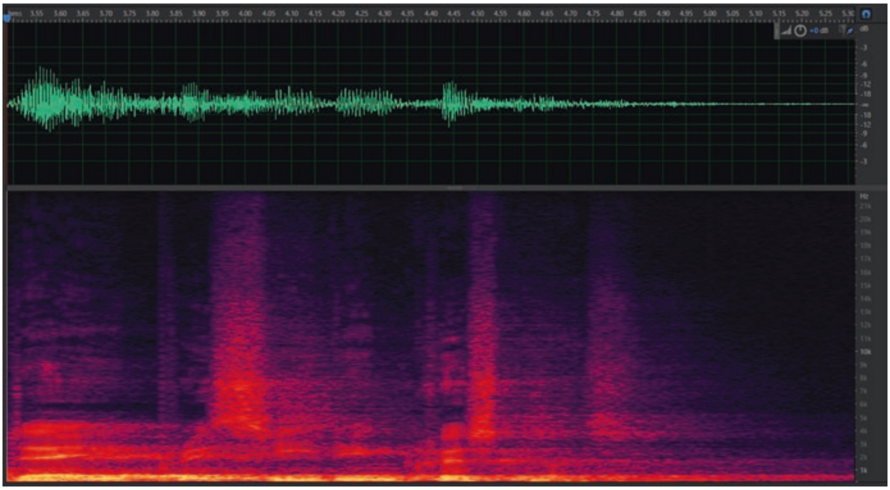

Fig. 5.11 Speech recording with strong reverberation present (overall duration 1.8 s, frequency range 0–22 kHz, linear scale)

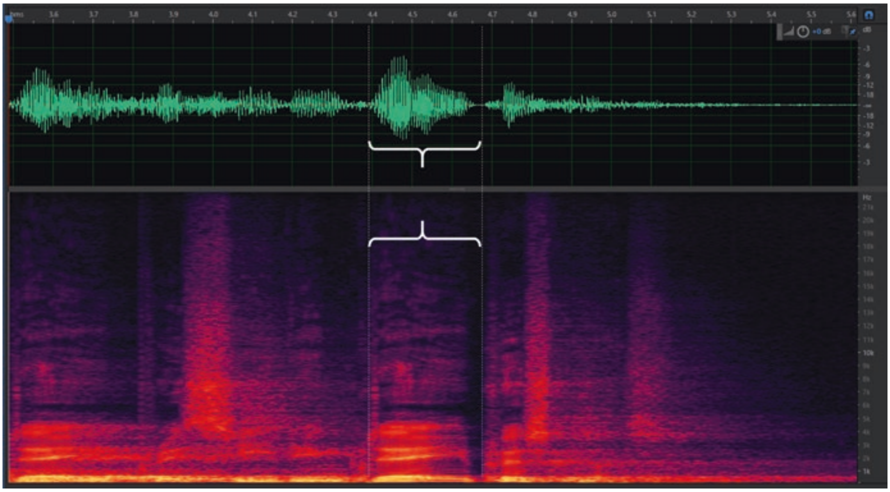

Fig. 5.12 Reverberant recording of Fig. 5.11 with a "dry" insertion (overall duration 2.1 s, frequency range 0–22 kHz, linear scale)

The gaps (dark areas) between the uttered words are particularly visible in the spectrogram, and the lack of background noise and reverberation is apparent. A recording of speech with reverberation present is shown in Fig. 5.11. The gaps between words visible in Fig. 5.10 are now filled with the lingering echoes and reverberation of the preceding sounds.

If there were an attempt to insert newly created material into a reverberant recording such as Fig. 5.11, the forgery could show a change in the reverberation pattern in the spectrogram, as well as in critical listening. Figure 5.12 shows an example in which a short utterance of dry speech is inserted into the recording of

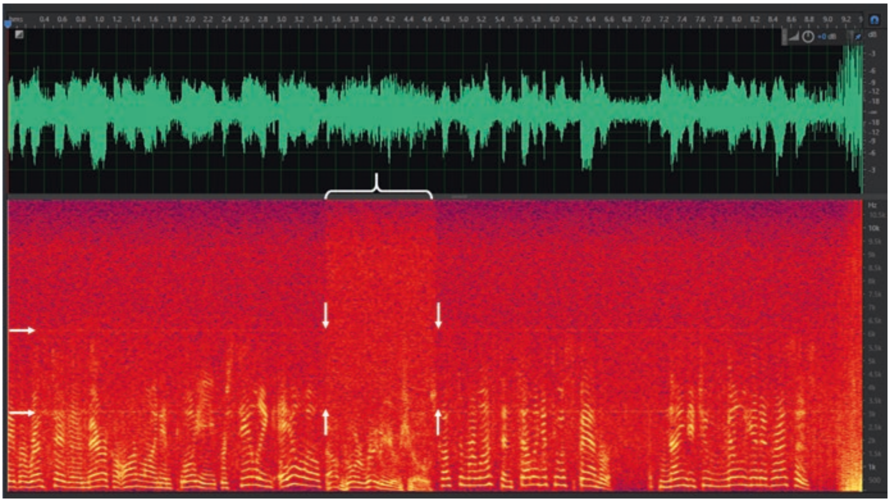

Fig. 5.13 Noisy speech recording with an apparent insertion. Note gap in continuous tones (horizontal lines) denoted by arrows and change in spectral "texture" (overall duration 9.3 s, frequency range 0–11 kHz, linear scale)

Fig. 5.11. The inserted speech lacks the reverb tail apparent after the other recorded words, indicating a likely edit point.

Another example is shown in Fig. 5.13. As seen in the spectrogram, the recording has significant background noise and two continuous discrete tones (horizontal lines indicated by the arrows on the left). In this example, there is a brief section indicated with a subtle difference in noise texture and the absence of the tones. These observations indicate a likely edit insertion into the recording.

5.2.3 Electrical Network Frequency (ENF) Analysis

An interesting potential technique for audio forensic investigations involves a particular background sound that may be present in a recording: the residual "hum" of the electrical power network. This "hum" is usually considered undesirable interference, but there are potentially some possibilities for using this background sound to assess authenticity.

The electrical network frequency (ENF) in the United States and some other countries is nominally 60 Hz, and 50 Hz ENF is common in Europe and many other parts of the world. The operation of the contemporary electrical power system requires that all of the AC electrical generators interconnected cooperatively through the electrical power network, or "the grid," operate synchronously: all of the 60 Hz power waveforms anywhere in the electrical grid are kept at exactly the same frequency and in-phase with each other. The United States power network is comprised of three large grids: eastern grid, western grid, and Texas. Within each grid, the power frequency is the same at every generator and outlet.

The electrical grid operating organization has to control the power system so that the amount of electricity being generated exactly matches the amount of electricity needed at any point in time, which keeps the ENF at the 60 Hz nominal value. However, if electrical use declines at a given time, the rotating electrical generators have less load and tend to turn a bit faster, increasing the ENF. On the other hand, if the demand for electricity increases, the electrical generators have a greater load and tend to slow down, decreasing the ENF. The grid operating organization must keep the variation to within about ±0.5 Hz by generating more or less power as needed, and the precise ENF frequency fluctuates gradually and unpredictably within the allowable range.

Because all of the generators attached to the grid operate synchronously, the instantaneous ENF will be the same everywhere on the entire electrical grid. If an audio recording includes hum from the electrical power system, the frequency of the hum is the electrical network frequency, and therefore it should be possible—at least in principle—to compare the recorded ENF fluctuations with a database of known power grid ENF measurements to identify the date and time of the recording.

Audio recording systems are generally designed to minimize the effects of AC (alternating current) power line interference, but low levels of residual power signals may appear in the audio circuitry and become part of the audio recording. This is most likely to occur when a line transformer powers the recording device, but some residual line frequency pickup is possible even with battery-powered equipment if the recording device is susceptible to the magnetic fields emanating from nearby wiring (Brixen 2007, 2008; Grigoras 2005, 2007).

In addition to needing a reference power grid frequency database, ENF analysis requires several important assumptions and measurements.

First, the recording must contain a detectable hum signal of sufficient strength that its precise frequency can be determined several times per second. The extraction process can be difficult because the 60 Hz ENF (and its harmonics) is within the regular audio bandwidth, so there may also be acoustic signals in the same frequency range as the ENF.

Second, the length of the recording and the corresponding duration of the ENF record need to be sufficiently long that the extracted ENF pattern is reliably distinguishable from any other span of time.

Third, the extracted ENF depends upon the actual sampling rate (or analog recording speed) of the audio signal, and any discrepancy in the recording process will introduce a systematic frequency shift.

An example procedure for extracting the ENF from audio recordings is shown in Fig. 5.14 (Cooper 2008). An example comparison of ENF data obtained from an audio recording and the reference ENF data from the electrical power system are shown in Fig. 5.15.

5.2.4 Metadata Consistency

Contemporary digital audio recordings are stored as computer files in a number of standard or proprietary formats. The audio file format includes the bytes containing the digital audio data, along with additional useful information *about* the recording, known

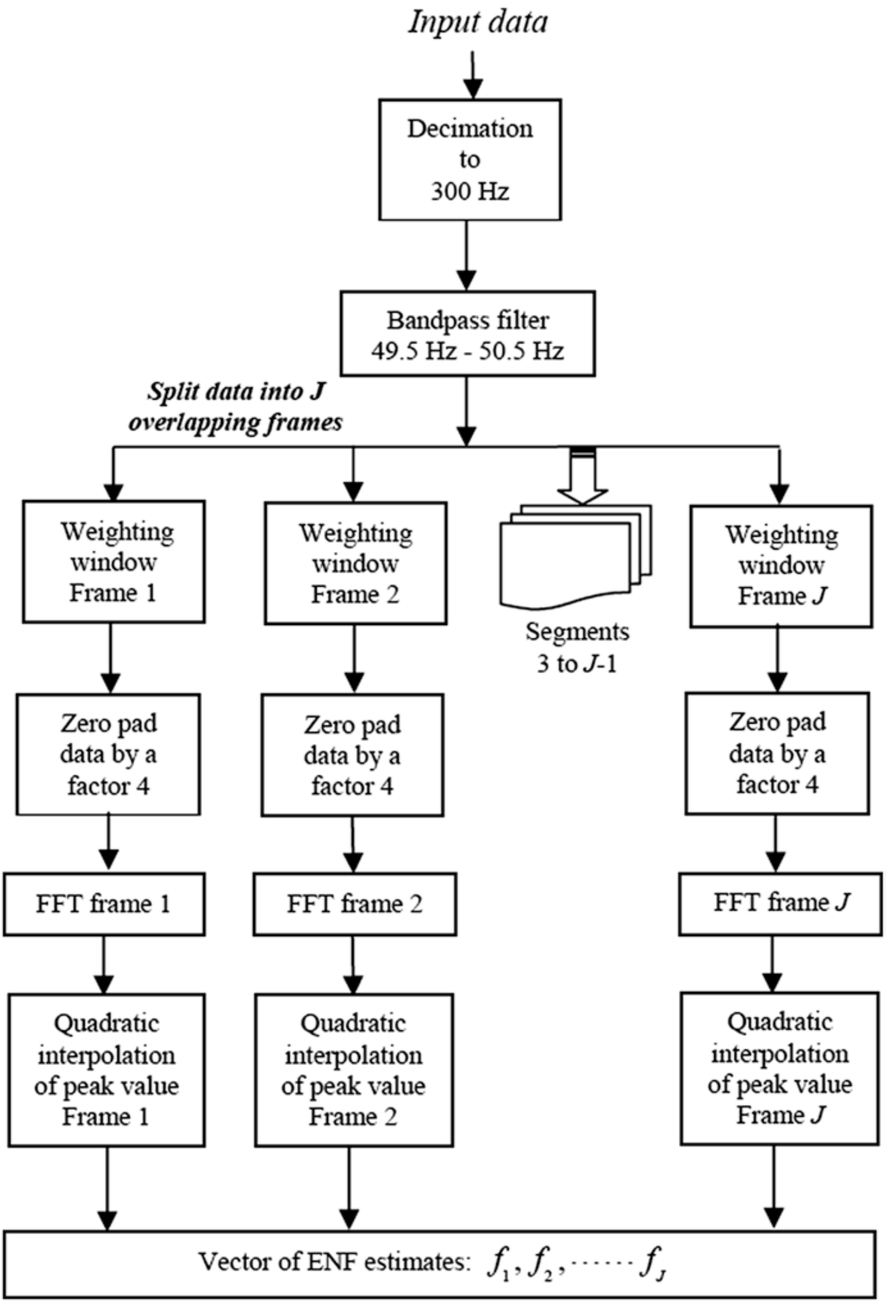

Fig. 5.14 Proposed ENF processing procedure (from Cooper 2008, reprinted by permission)

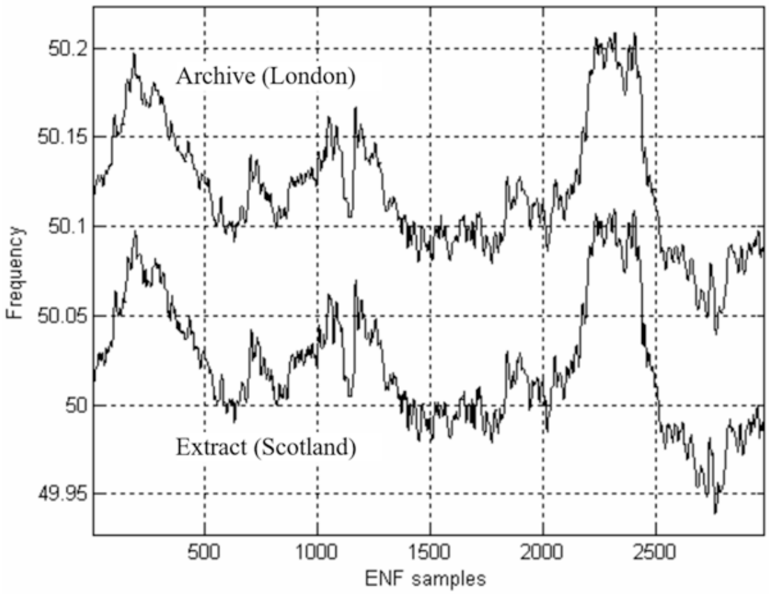

Fig. 5.15 Automated match between extracted ENF and database. The total time represented is approximately 70 min (1.4 s per ENF sample). The extracted waveform has been offset by 0.1 Hz to aid visual comparison (from Cooper 2008, reprinted by permission)

as *metadata*. Metadata included in an audio file format might include the sampling rate, the number of audio channels, the brand and model number of the recording device, the date of the recording, and so forth (Koenig and Lacey 2014). A forensic examiner can nondestructively observe the metadata using an editor program capable of displaying the binary information as readable characters (typically hexadecimal notation).

An example listing of the file metadata for an example MP3 audio file is shown in Fig. 5.16. The information is from a binary display program and shows the file contents as text characters in the right column, and the binary data values are given as hexadecimal (base-16) values in the body of the figure. Note that in the right column, there are recognizable strings of characters such as “ID3,” identifying the beginning of the file to be containing standard “tags” from id3.org. The file was recorded by an Olympus brand model 702 memo recorder on July 11, 2018. Note the strings OLY (Olympus), mp3 (recording mode), 702 (recorder model), and 20,180,711 (date code). Other bytes in the metadata header may also be meaningful if the manufacturer provides technical information regarding the recording device. Often this information is proprietary, and its meaning can only be determined empirically.

When a recording device opens a file, performs the recording operation, and then saves and closes the file, the device also updates the metadata. An authentic recording will have metadata that is consistent with the stated circumstances of the recording and the file contents. As in the example above, the metadata must match the type and model of recorder, the recording date, and the recording duration.

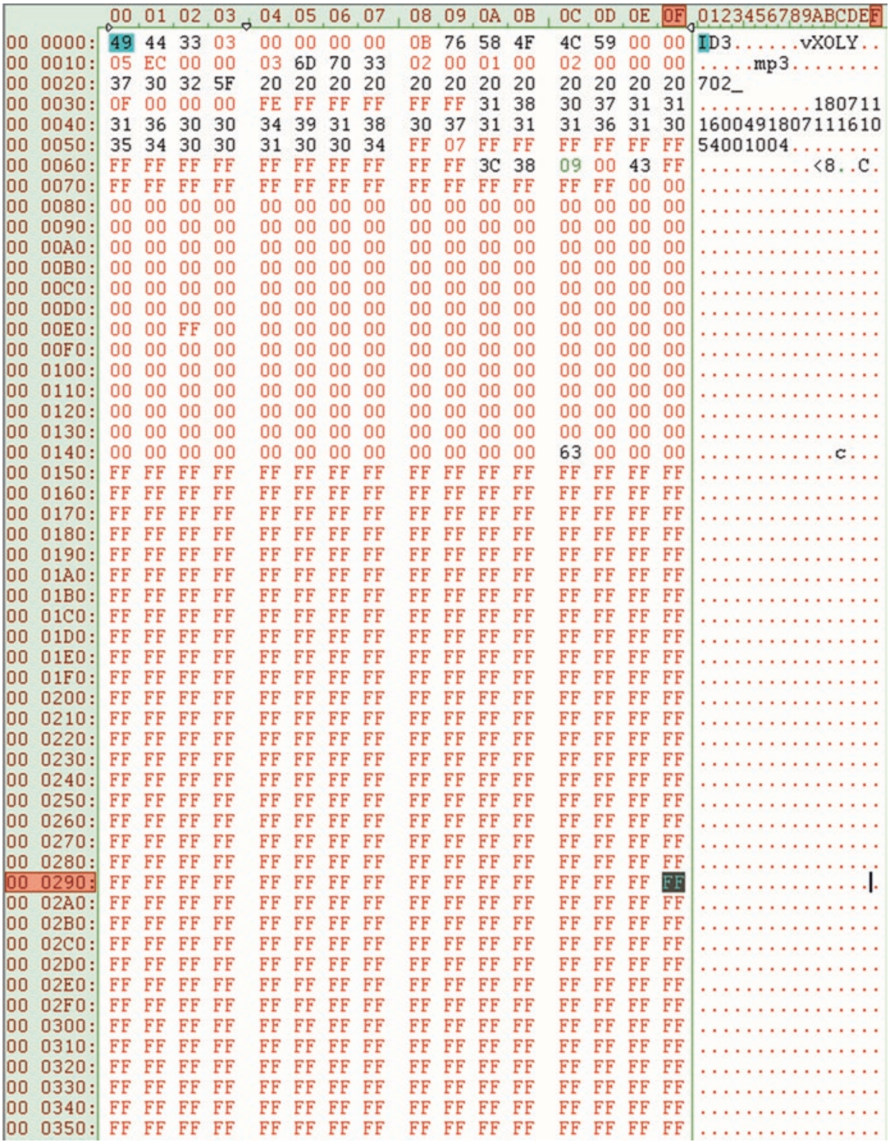

Fig. 5.16 Example of metadata in a digital audio file, displayed as hexadecimal and text. This file is from an Olympus brand model 702 memo recorder, and the file was created on July 11, 2018. Note the strings “ID3,” “OLY,” “mp3,” “702,” and the date code “20,180,711”

However, if an authentic original file is transferred to another device or to a computer, edited on that device, and then saved to a new file, the editing device or software package typically updates various details in the metadata. For example, the file depicted in Fig. 5.16 was subsequently opened with the *Adobe Audition®* software package and then saved with a new file name, resulting in the metadata shown in Fig. 5.17. The software has clearly altered the metadata at the head of the file.

```
         00 01 02 03  04 05 06 07  08 09 0A 0B  0C 0D 0E 0F  0123456789ABCDEF
00 0000: 49 44 33 03  00 00 00 00  30 01 54 59  45 52 00 00  ID3.....0.TYER..
00 0010: 00 06 00 00  00 32 30 31  38 00 54 44  41 54 00 00  .....2018.TDAT..
00 0020: 00 06 00 00  00 31 32 30  37 00 54 49  4D 45 00 00  .....1207.TIME..
00 0030: 00 06 00 00  00 31 36 34  39 00 50 52  49 56 00 00  .....1649.PRIV..
00 0040: 0F C7 00 00  58 4D 50 00  3C 3F 78 70  61 63 6B 65  ....XMP.<?xpacke
00 0050: 74 20 62 65  67 69 6E 3D  22 EF BB BF  22 20 69 64  t begin="..." id
00 0060: 3D 22 57 35  4D 30 4D 70  43 65 68 69  48 7A 72 65  ="W5M0MpCehiHzre
00 0070: 53 7A 4E 54  63 7A 6B 63  39 64 22 3F  3E 0A 3C 78  SzNTczkc9d"?>.<x
00 0080: 3A 78 6D 70  6D 65 74 61  20 78 6D 6C  6E 73 3A 78  :xmpmeta xmlns:x
00 0090: 3D 22 61 64  6F 62 65 3A  6E 73 3A 6D  65 74 61 2F  ="adobe:ns:meta/
00 00A0: 22 20 78 3A  78 6D 70 74  6B 3D 22 41  64 6F 62 65  " x:xmptk="Adobe
00 00B0: 20 58 4D 50  20 43 6F 72  65 20 35 2E  36 2D 63 31   XMP Core 5.6-c1
00 00C0: 34 34 20 37  39 2E 31 36  32 30 34 35  2C 20 32 30  44 79.162045, 20
00 00D0: 31 38 2F 30  31 2F 32 33  2D 30 36 3A  30 35 3A 35  18/01/23-06:05:5
00 00E0: 32 20 20 20  20 20 20 20  20 22 3E 0A  20 3C 72 64  2        ">. <rd
00 00F0: 66 3A 52 44  46 20 78 6D  6C 6E 73 3A  72 64 66 3D  f:RDF xmlns:rdf=
00 0100: 22 68 74 74  70 3A 2F 2F  77 77 77 2E  77 33 2E 6F  "http://www.w3.o
00 0110: 72 67 2F 31  39 39 39 2F  30 32 2F 32  32 2D 72 64  rg/1999/02/22-rd
00 0120: 66 2D 73 79  6E 74 61 78  2D 6E 73 23  22 3E 0A 20  f-syntax-ns#">. 
00 0130: 20 3C 72 64  66 3A 44 65  73 63 72 69  70 74 69 6F   <rdf:Descriptio
00 0140: 6E 20 72 64  66 3A 61 62  6F 75 74 3D  22 22 0A 20  n rdf:about="". 
00 0150: 20 20 20 78  6D 6C 6E 73  3A 78 6D 70  44 4D 3D 22     xmlns:xmpDM="
00 0160: 68 74 74 70  3A 2F 2F 6E  73 2E 61 64  6F 62 65 2E  http://ns.adobe.
00 0170: 63 6F 6D 2F  78 6D 70 2F  31 2E 30 2F  44 79 6E 61  com/xmp/1.0/Dyna
00 0180: 6D 69 63 4D  65 64 69 61  2F 22 0A 20  20 20 20 78  micMedia/".    x
00 0190: 6D 6C 6E 73  3A 78 6D 70  3D 22 68 74  74 70 3A 2F  mlns:xmp="http:/
00 01A0: 2F 6E 73 2E  61 64 6F 62  65 2E 63 6F  6D 2F 78 61  /ns.adobe.com/xa
00 01B0: 70 2F 31 2E  30 2F 22 0A  20 20 20 20  78 6D 6C 6E  p/1.0/".    xmln
00 01C0: 73 3A 78 6D  70 4D 4D 3D  22 68 74 74  70 3A 2F 2F  s:xmpMM="http://
00 01D0: 6E 73 2E 61  64 6F 62 65  2E 63 6F 6D  2F 78 61 70  ns.adobe.com/xap
00 01E0: 2F 31 2E 30  2F 6D 6D 2F  22 0A 20 20  20 20 78 6D  /1.0/mm/".    xm
00 01F0: 6C 6E 73 3A  73 74 45 76  74 3D 22 68  74 74 70 3A  lns:stEvt="http:
00 0200: 2F 2F 6E 73  2E 61 64 6F  62 65 2E 63  6F 6D 2F 78  //ns.adobe.com/x
00 0210: 61 70 2F 31  2E 30 2F 73  54 79 70 65  2F 52 65 73  ap/1.0/sType/Res
00 0220: 6F 75 72 63  65 45 76 65  6E 74 23 22  0A 20 20 20  ourceEvent#".   
00 0230: 20 78 6D 6C  6E 73 3A 64  63 3D 22 68  74 74 70 3A   xmlns:dc="http:
00 0240: 2F 2F 70 75  72 6C 2E 6F  72 67 2F 64  63 2F 65 6C  //purl.org/dc/el
00 0250: 65 6D 65 6E  74 73 2F 31  2E 31 2F 22  0A 20 20 20  ements/1.1/".   
00 0260: 78 6D 70 3A  4D 65 74 61  64 61 74 61  44 61 74 65  xmp:MetadataDate
00 0270: 3D 22 32 30  31 38 2D 30  37 2D 31 32  54 31 36 3A  ="2018-07-12T16:
00 0280: 34 39 3A 33  33 2D 30 36  3A 30 30 22  0A 20 20 20  49:33-06:00".   
00 0290: 78 6D 70 3A  43 72 65 61  74 6F 72 54  6F 6F 6C 3D  xmp:CreatorTool=
00 02A0: 22 41 64 6F  62 65 20 41  75 64 69 74  69 6F 6E 20  "Adobe Audition 
00 02B0: 43 43 20 32  30 31 38 2E  31 20 28 57  69 6E 64 6F  CC 2018.1 (Windo
00 02C0: 77 73 29 22  0A 20 20 20  78 6D 70 3A  43 72 65 61  ws)".   xmp:Crea
00 02D0: 74 65 44 61  74 65 3D 22  32 30 31 38  2D 30 37 2D  teDate="2018-07-
00 02E0: 31 32 54 31  36 3A 34 39  3A 33 33 2D  30 36 3A 30  12T16:49:33-06:0
00 02F0: 30 22 0A 20  20 20 78 6D  70 3A 4D 6F  64 69 66 79  0".   xmp:Modify
00 0300: 44 61 74 65  3D 22 32 30  31 38 2D 30  37 2D 31 32  Date="2018-07-12
00 0310: 54 31 36 3A  34 39 3A 33  33 2D 30 36  3A 30 30 22  T16:49:33-06:00"
00 0320: 0A 20 20 20  78 6D 70 4D  4D 3A 49 6E  73 74 61 6E  .   xmpMM:Instan
00 0330: 63 65 49 44  3D 22 78 6D  70 2E 69 69  64 3A 35 61  ceID="xmp.iid:5a
00 0340: 31 62 39 30  32 36 2D 62  65 31 37 2D  32 39 34 66  1b9026-be17-294f
00 0350: 2D 39 31 61  36 2D 37 61  39 34 64 36  34 30 38 38  -91a6-7a94d64088
```

Fig. 5.17 Example of metadata for the file of Fig. 5.16 opened and saved with a different software package, thereby altering the metadata

If the examination reveals inconsistency between the metadata and the expected circumstances of the recording, this could indicate that the recording was edited or modified in some manner, possibly representing a forgery.

As explained previously, one of the difficulties associated with digital files is the inability to distinguish between an authentic file and a forgery prepared by a skillful adversary. This caveat applies to metadata as well, since a forgery could have metadata altered in such a manner as to appear consistent with an authentic recording. Thus, an examiner is generally only able to report upon inconsistencies that could indicate inauthenticity, not to guarantee authenticity if no inconsistencies are found.

References

Advisory Panel on White House Tapes. (1974). *The executive office building tape of June 20, 1972: Report on a technical investigation*. Washington, D.C.: United States District Court for the District of Columbia.

Audio Engineering Society. (2000). *AES43-2000: AES standard for forensic purposes – Criteria for the authentication of analog audio tape recordings*. New York: AES.

Brixen, E. B. (2007). Techniques for the authentication of digital audio recordings. In *Proceedings Audio Engineering Society 122nd Convention, Vienna, Austria, Convention paper 7014*.

Brixen, E. B. (2008). ENF–Quantification of the magnetic field. In *Proceedings Audio Engineering Society 33rd Conference, Audio Forensics—Theory and Practice, Denver, CO* (pp. 1–6).

Begault, D. R., Brustad, B. M., & Stanley, A. M. (2005) Tape analysis and authentication using multi-track recorders. In *Proceedings Audio Engineering Society 26th Conference, Audio Forensics in the Digital Age, Denver, CO* (pp. 115–121).

Cooper, A. J. (2008). The electric network frequency (ENF) as an aid to authenticating forensic digital audio recordings – An automated approach. In *Proceedings Audio Engineering Society 33rd Conference, Audio Forensics—Theory and Practice, Denver, CO* (pp. 1–10).

Grigoras, C. (2005). Digital audio recording analysis: The electric network frequency (ENF) criterion. *International Journal of Speech, Language and the Law, 12*(1), 63–76.

Grigoras, C. (2007). Application of ENF analysis method in authentication of digital audio and video recordings, In *Proceedings Audio Engineering Society 123rd Convention, New York, NY, Convention paper 1273*.

Koenig, B. E. (1990). Authentication of forensic audio recordings. *Journal of the Audio Engineering Society, 38*(1/2), 3–33.

Koenig, B. E., & Lacey, D. S. (2014). Forensic authenticity analyses of the metadata in re-encoded WAV files. In *Proceedings of the Audio Engineering Society 54th International Conference: Audio Forensics, London, U.K.*

Marr, K., & Pappas, D. P. (2008). Magneto-resistive field mapping of analog audio tapes for forensics imaging. In *Proceedings of the Audio Engineering Society 33rd Conference, Audio Forensics Theory and Practice, Denver, CO* (pp. 1–7).