VIDEO SWEETENING BASICS FOR AUDIO ENGINEERS

© Eric Wenocur Rev.5 9/2/92

With the advent of stereo and high fidelity television equipment, the desire for high-quality sound to accompany video productions has grown dramatically. The main purpose of **video sweetening** (also called "audio sweetening", or "audio post-production") is to enhance the audio content of a video production. This enhancement might include sound effects, music, sync sound, or narration. It is often impractical for complex audio to be handled during video editing, so it is then necessary for the program to be "sweetened" on a multitrack audio recorder, or disk-based editing system (so-called "audio workstations").

In addition, development of reliable methods to link audio and video transports has allowed automation of such tasks as dialogue replacement for films ("looping"). Manufacturers have been quick to provide equipment for handling these tasks, but utilizing this equipment correctly requires knowledge of some possibly unfamiliar concepts.

The main purpose of video sweetening (also called "audio sweetening", or "audio post--production") is to enhance the audio content of a video production. This augmentation might include sound effects, music, sync sound, or narration. It is often impractical for complex audio to be handled during video editing, so it is necessary for the program to be "sweetened" on a multitrack audio recorder, or disk-based editing system (so-called "audio workstations"). In addition, development of reliable methods to link audio and video transports has allowed automation of such tasks as dialogue replacement for films ("looping"). The following description outlines the sweetening process using a simple, synchronized audio and video system.

The Sweetening Process

The typical video sweetening job takes place after the video of the program has been edited into its final, or nearly final form. This is done at a video editing facility, and the end product is an **edit master**--usually on 3/4", 1" or digital (D1 or D2) videotape. The audio channels on the edit master may contain nothing, but they will usually contain some reference audio, such as narration, or sync sound that was layed in during editing. This is usually live sound from the shoot, particularly of people speaking on camera ("lip-sync"). In the sweetening process, the preservation of the sound-to-picture sync relationship will be of continual concern.

From the edit master, two dubs are made. One is a 3/4" videotape with timecode on an audio channel (or address track, discussed later), and **burn-in** timecode numbers in the picture. This tape will be the video reference during the sweetening process. If one channel of the 3/4" tape has timecode, the other should contain a mix of the two audio tracks on the edit master.

The other tape is a 2 or 4-track audio dub of the edit master, with timecode on one channel or on centertrack. For this tape, the two audio tracks from the edit master must be dubbed one-to-one, as they will become the base tracks on the multitrack tape or hard-disk. Noise reduction is advisable, if it is available at both facilities. (Audio facilities capable of playing the edit master format may transfer tracks directly.)

Most importantly, the timecode on these dubs should be the exact same code as the edit master. This greatly simplifies sweetening, because the timecode for a given audio event will be identical to that for the corresponding video event on the 3/4" reference--thus eliminating the need for slave offsets.

Once these tapes are at the studio, the audio from the 2-track tape is dubbed, or **layed-up** onto the multitrack (or into the digital editing system). The multitrack now contains the base tracks upon which additional audio will be built. It also has the same timecode as the video master, which has been dubbed from one of the other tapes (usually the 2-track).

Now, the reference 3/4" tape is used as the **MASTER**, and the multitrack as a **SLAVE** (a hard-disk editing system would also typically run as a slave). The synchronizer runs the multitrack in exact sync with the videotape, and additional sounds are layed-up onto the multitrack in the exact location desired to fit with the picture. This usually amounts to a number of effects and/or music tracks. If a 3-machine synchronizer (or audio edit controller) is used, the source tapes may also be run in sync as the audio is layed-up. This makes precise placement of the new sounds easier.

When all the sounds have been layed-up, the multitrack tape is mixed back down to 2-track, while watching the video for reference ("mix-to-picture"). Again, this 2-track tape will have the same timecode as the other tapes (usually copied from the multitrack during mixing).

Finally, the 2-track mix is brought back to the video house where it is **layed-back** onto the edit master in place of the original tracks. As long as the audio and timecode transfers have been made correctly, the audio tracks will sync back up to the edited video.

A common variation of this process is when the video edit master is a format which the audio facility can use (many larger houses now have 1" and D2 VTRs). In this case the multitrack mixdown can be made directly onto this tape, eliminating the layback step. Of course, this requires extra care, since the recording is being made onto the edit master. It is also limited by the location of the timecode on the master and whether the mix is stereo (becoming more common in video projects). A 3/4" burn-in dub of the master will still be used for reference during sweetening, to avoid running the master.

Of course there are many other ways in which audio-for-video is used. For example, the construction of a soundtrack before the video is edited, with the soundtrack acting as the base for the entire show. In most cases, though, the same elements of audio, video and timecode are involved and must be handled correctly.

Video Basics

Although the television picture appears as a continuous image, it is actually a series of individual "snapshots" flashing by. These are not seen as snapshots because our persistence of vision causes them to appear continuous. Unlike analog audio, which is a continuous unbroken signal, these individual snapshots are electronically generated and recorded as discrete "packets" of information, or **fields**. Two consecutive fields together contain all 525 lines of signal information needed to the fill the television screen one time, thus producing a complete **frame** of video.

Between each field there is a vertical blanking period (or vertical interval) containing a **vertical sync pulse**. The vertical blanking period can be seen as the "black bar" that rolls through the picture when the vertical hold is misadjusted. Because of their construction, video signals can only be edited electronically, so video edits are made during the vertical interval where they will be unseen.

In American color television (NTSC), there are **29.97 frames** of video per second (59.94 fields). This does not mean that 29 "and a fraction" of actual video frames occur each second, that is impossible. It simply means that the video frame rate is slightly slower than 30 fps, so that at the end of one "realtime" (clock-on-the-wall) second, only enough time has elapsed for 29.97 frames to have occured. The reasons for this odd number have to do with the frequencies used in the color television signal--in the days of black & white the frame rate was an even 30 fps. Fortunately it is generally acceptable to think in terms of "30" frames per second for most applications, particularly video sweetening. The actual frame rate is only important under certain circumstances, which will be mentioned later.

On a VTR of any current analog format, the video is recorded by magnetic heads attached to a rotating head drum; a system known as **helical scan**. The tape is threaded around the drum at an angle, so that the tracks of the signal end up as slanted lines along the tape. Audio, timecode and control track signals are recorded longitudinally beside the video tracks by conventional audio heads. In addition there is a "full bar" erase head which will wipe all signals on the tape, and "flying" erase heads, on the video head drum, which erase video only. Digital VTRs also use helical scan recording, but the video and audio are recorded together as interleaved tracks of digital data. Timecode and audio cue channels are still recorded longitudinally.

On 1" videotapes, timecode is usually recorded on audio track 3, which is dedicated for this purpose. Betacam and digital VTRs also have a fully functional dedicated TC track. On 3/4" tapes, however, timecode can be on an audio track, or on a special **address track**. The address track is standard on professional 3/4" decks, but is usually an option on industrial (U-matic) models. Address track is tricky because it is a longitudinal track recorded in the same physical location as the vertical interval on the tape. This means that address track code cannot be poststriped; it can only be recorded simultaneously with video! Since it may be optional, studios must know if their VTR has address track capability before accepting these tapes. If not, matching timecode can be poststriped, at a video facility, onto an audio track (but, of course, that track is no longer useable for audio). Any other method of recording timecode on a 3/4" VTR is not "industry standard" and should be examined for compatibility before accepting tapes.

There are three types of video recording: hard record, assemble editing, and insert editing. When a VTR is put into **hard record**, it erases and records all tracks simultaneously. In addition to the video and audio, the VTR records a control track signal (NOT to be confused with address track). The **control track** is a series of pulses used to govern the tape speed on playback. This is necessary because off-speed playback on most VTRs will cause visible frame lines in the picture. In addition, a lack of control track will likely cause erratic tape speed, and make any audio or timecode unusable as well.

The hitch here is that a VTR will only record control track if it is receiving a video signal. If hard record or assemble edits are attempted with no video input to the VTR, control track will not be recorded! Any audio house that cannot provide a correct video (or composite sync) signal cannot make these types of recordings. This does not mean that no video sweetening can be done. It simply means that these functions of the VTR should be disabled, or the buttons taped over. This is the first step in avoiding potentially disastrous trouble. (Remember, a facility that cannot record video also cannot "stripe" timecode on a blank videotape.)

On VTRs with electronic counters, if the tape has no control track the counter will not advance. As long as the counter advances, the control track is present and audio can be accurately recordedeven if the video has been erased. Since it is recorded automatically, control track is often taken for granted, but a videotape without control track is virtually useless.

Assemble editing is just like hard record, except that the VTR rolls into the edit, and starts recording new control track in a perfect continuation of that already on tape. Assemble editing is used when additional new recording is being done on a partially blank tape, and it is desired to "pickup" at the end of each segment and continue on. To make assemble edits, the first video on the tape must have been created by hard recording. In addition, once an assemble edit is made a "hole" is left at the end, so assemble edits must be made from then on. Again, the VTR must receive video to make assemble edits!

In **insert editing**, only video or audio tracks are recorded. On VTRs with this capability, there are buttons to select which track will be put into recordlike record safeties on an audio deck. There is also a separate "record" button for editing (vs. hard record). No control track is recorded during an insert edit, even if video is being recorded. For this reason, insert edits are made only on tapes which already contain control track. Keep in mind that only fresh videotapes are actually blank. Tapes that have been "blacked" for editing contain at least control track and video, so insert editing can be done. Audio insert editing is the only type likely to be used at an audio facility doing sweetening.

Timecode Basics

Timecode is a system that was designed to allow accurate identification of every video frame on a tape by assigning each frame a unique number. There is only one valid timecode number for each video frame. Each frame of timecode is made up of an 80-bit stream of digital data containing bits for hours, minutes, seconds and frames, as well as other information. There is a **sync word** at the beginning of each 80-bit group which identifies the start of a new frame. The timecode signal is recorded longitudinally on videotape (except for Vertical Interval Timecode, mentioned later), and is a continuous, frequency--modulated signal that sounds like a dirty "warble tone", much like old-fashioned drum machine sync tone.

Timecode comes in a variety of different formats: 30 fps, 30 fps drop-frame, 29.97 fps, 29.97 fps drop-frame (all called "SMPTE", for the American standard), 24 and 25 fps (EBU, the European standard). Each type is used for a certain purpose, and they are not interchangeable. 30 fps code is always used in conjunction with 24 fps film shoots (unless otherwise specified), and is generally acceptable for audio post-production work, as long as audio leaving the studio for layback is locked to the same type of code that it came in with. In virtually all cases, this will be 29.97 fps code, which is used in video houses because the timecode rate must match the video frame rate. (Bear in mind that video houses rarely need to consider what rate code they are using because the VTRs and timecode generators are all locked to a common sync generator. Don't be surprised if your contact at a video house acts confused when the subject of timecode rates comes up.)

Then again, while standard 29.97 fps code is most common, in some cases it causes a problem with determining program length. Since the code runs at 29.97 frames per second, but counts to an even 30, the result is a loss of 1.8 frame counts per actual minute of time. If standard 29.97 fps code is used to determine program length, after one realtime hour the timecode will read only 00:59:56:12 (59 minutes, 56 seconds, 12 frames). When the timecode finally reaches the 01:00:00:00 count, the program is 3 seconds and 18 frames too long. In broadcast this is thoroughly unacceptable. Since the video frame rate really is 29.97 fps, it is necessary for the timecode numbers to reflect the actual number of frames that have gone by in order to time programs accurately.

This is done with the aid of drop-frame timecode. **Drop-frame** is a variation whereby certain frame numbers are "dropped" from the count, so that the final count is 01:00:00:00 at the end of each realtime hour. The drop-frame system eliminates the first two frames in every minute, except the tens minutes (0, 10, 20, etc), thus reducing the count by the necessary 108 frames. Remember, **drop-frame runs at the same rate as non-drop-frame code (either 30 or 29.97 fps)**, only the counting system is different.

From a video sweetening perspective, the use of drop-frame or non-drop code is relatively unimportant, and will usually be determined during the video editing. However, any tapes that will be synchronized with a drop-frame master should also contain drop-frame code. This will eliminate alot of confusion. Another complication to remember is that with drop-frame certain frame numbers "do not exist". Thus if we ask a synchronizer to park the tape at 1:01:00:00, it will be unable to do so because this frame number is not on the tape! The count actually changes from 01:00:59:29 to 01:01:00:02. The lack of certain frames also complicates offset calculations and such. Most timecode generators and synchronizers can deal with drop-frame code without problems, but, given a choice, it may be preferable to use nondrop to simplify manual timecode calculations.

The other two timecode formats, 24 and 25 fps, are used for film work and European television (PAL, in which video runs at 25 fps), respectively. My own recommendation for audio studios is to stick with 29.97 fps code for all in-house work. This will avoid any possible problems that arise from mixing frame rates with the video houses. One such problem is that the tempo and absolute pitch of audio material locked to one rate of code will change if played back synchronized at the other rate. The difference is only .1% (1/10 of 1%), but it might cause trouble in certain situations (for example, synchronizing live MIDI instruments with the system, which will play the pitches stored in their sequencer, regardless of speed changes of the master tape).

Handling Timecode

Timecode is a relatively volatile signal in two ways. It has alot of high frequency content, and seems to "leak" into everything, and it is easily degraded due to dubbing. Timecode should never be dubbed from any tape source without first being regenerated or restored. If this is not done, each generation of code will become less and less readable.

Regenerating is a process whereby the code leaving a tape is used to feed a timecode reader which simultaneously drives a generator that makes new code. This new code is fresh and can be recorded again. Since regenerating utilizes a generator, the code that comes out is subject to the sync frame rate that the generator is locked to. In the case of most smaller studio synchronizers, the generator will simply lock to the incoming code, and thus the regenerated code will be exactly the same rate as the original.

At least, that is one option. Many units also allow the generator (or entire synchronizer) to lock to other sources of sync. This is where the utmost care must be taken! If the generator is inadvertantly locked to the wrong source of sync, the new code will no longer be locked to the video or audio it was originally associated with (it will emerge at the rate that the generator is locked to). For this reason, the generator must always be locked to the reader when dubbing code (unless all equipment is locked to a common sync generator).

When striping or post-striping timecode on an audio tape, the generator should be internally referenced to its own crystal (non-resolved mode, if the generator is part of a synchronizer), and the ATR should be recording at its internally determined speed.

Conversely, when timecode is recorded on a videotape, the timecode generator and the VTR must be “locked” together, so that each timecode number actually corresponds to a video frame. This can be accomplished by referencing the generator to the VTR’s video output. If they are not locked together, the timecode will “float” in relation to the video, causing massive synchronizing problems.

Regenerating is often referred to as **jam-syncing** the timecode, because the reader is "jamming" new numbers into the generator. The thing to be watchful of is whether or not the code is being duplicated frame-for-frame. In a "momentary" jam-sync the generator looks at the reader at the start of the transfer, but then continues to count upward on its own, regardless of what comes out of the reader. This is used to replace timecode that has been partially erased or has unwanted breaks in it. A "continuous" jam-sync (called "transfer" by at least one synchronizer manufacturer) causes the generator to duplicate the reader exactly, even if the incoming code jumps or stops. This second method is generally used for dubbing of code during video sweetening.

A simple alternative to regenerating is **restoring** the code. This is electronic "reshaping" of the timecode signal, so that the data bits are clearly defined. It does not involve a generator, and does not have the same inherent dangers. Some timecode devices provide a Restored Code output, which can be used for dubbing code with little worry. Also the timecode output of most 1”, Betacam and digital VTRs is already regenerated internally and suitable for re-recording.

Timecode should not be recorded at "0" VU on audio tracks of an ATR or VTR as it tends to crosstalk. Recommended levels are between -10 and -5 VU. Also be sure that any noise reduction is bypassed on the track being used for timecode. Keep in mind that "timecode present" indicators on many VTRs and ATRs actually detect the presence of any audio signal. If the light is on, but still no timecode reading, patch the code output into a speaker and see if there is really timecode there!

Lastly, a word about **Vertical Interval Timecode**. This is a signal that is recorded in the vertical interval on a videotape by the video heads; it is not a longitudinal signal. The usefulness of this method, besides saving an audio track, is that the timecode can be read at very slow speeds--even when the tape is not moving at all--because the video heads are always moving (longitudinal code is innaccurate at very slow speeds). Hence it is excellent for locating and cueing up to exact points on the videotape (such as a sound effect "hit"). VITC can be recorded on any videotape, but requires a special generator and reader (offered as an option by some synchronizer manufacturers).

Using Timecode

There are a few commonly accepted practices in the video industry that should be adopted by audio studios when dealing with timecode. All synchronizing devices require a few seconds to bring their transports up to speed and lock before the actual audio can be used. For this reason there must be sufficient timecode before the audio material to allow for this **preroll** time. The upshot of this is that the standard audio practice of tight leadering between cuts is a no-no! There MUST be at least 10 seconds of magnetic tape, with timecode, before each selection for adequate preroll--as much as 30 seconds might sometimes prove useful. And please, do your video colleagues a favor by applying this practice to any audio tapes going to video houses!

Another standard practice is that of starting actual program at the 01:00:00:00 (one hour) count. Besides being neat and tidy, this avoids the possibility of crossing the 24 hour count during preroll, thus sending the transports screaming backward to find a number less than 00. In addition, it is wise to start the timecode at about 00:58:00:00 (58 minutes), to allow two minutes for tone, slates, and silence before the show.

In a related rule, timecode on a given tape should always be ascending. If there is a point where the timecode jumps backward (say at an edit), there will be some duplicate numbers on the tape which will confuse the synchronizer no end.

And speaking of edits, I do not recommend making razor blade edits anywhere on a timecoded tape during program material. While it is possible to make a frame-accurate splice, if any numbers are spliced out of the code it may cause a speed glitch when synchronized or, worse yet, loss of lipsync after the edit.

Lastly is myown little pet peave that people in audio seem to refer to SMPTE Timecode as just "SMPTE". For those who don't know, SMPTE is a professional organization that would probably like to collect dues from everyone who uses its name, instead of the more accurate term "timecode" (which is what video people call it). For that matter, SMPTE Timecode should not be confused with MIDI Timecode, MIDI SPP, FSK, or other similar sounding signals.

Resolving and Synchronizing

Now that we have discussed timecode, and its relation to video, the final issue is relating the timecode to the audio. Before sweetening, the audio from the video edit master is dubbed off and transfered to audio tapes. The premise here is that the finished audio will ultimately be layed-back onto the videotape, so there must be a means of ensuring that the audio will remain in sync with the picture.

The link that keeps the audio/video relationship intact is a **time reference**. Recording audio with a known time reference is like marking it with the ticking of a clock to guide it to the correct speed whenever it is played. Different time references have particular rates of "ticks" per second. These include video control track (59.94 per second), timecode frames (29.97 or 30), pilot-tone cycles (60) or film perforations (24 per second). The key to successfully using a time reference to retain audio sync is that **once the time reference and audio material are put in an established relationship, that relationship must never be changed!** How this is accomplished requires some further background...

The method used to make tapes play at a speed determined by the time reference is known as resolving the tape speed. **Resolving** is the process of regulating a tape's playback speed so that the time reference from the tape matches the rate of an external time reference. The resolving device continually compares the time reference signal off the tape with the external reference, and controls the tape playback speed (via the capstan) so that the two rates match exactly. This comparing/compensating feedback loop happens continuously, so the resolver always keeps the references locked together.

Resolving playback speed has the inherent property of compensating for speed deviations that may have occured during recording. However, the main purpose of resolving is not to keep the absolute pitch of the audio correct, but to ensure that the audio will remain in sync with video, or with audio on another machine.

Unlike VTRs (with control track), audio machines traditionally have no means of resolving their own speed; the ATR runs at a speed determined by an internal crystal (or the AC line frequency), regardless of the exact speed the tape was recorded at. When set for "15ips", the crystal provides a speed close enough to 15ips for most audio work. However, when picture sync is involved, the longterm speed drift of two non-synchronized machines is not acceptable. Consequently, audio tapes must use timecode as a time reference. For this reason timecode must be present on an audio tape for it to be run in sync with other tapes. The timecode allows the ATR's playback speed to be resolved so that the audio remains in sync with the picture. It also allows the tape to be synchronized with other transports, and cued to an exact location repeatedly.

For resolving to take place, the capstan speed must be controlled by an external device (the synchronizer), rather than the internal crystal. ATRs with this capability generally have some type of interal/external selector for capstan control. A simple synchronizer will run a Slave transport in sync with a Master transport by using the timecode present on both tapes. The synchronizer compares the timecode from both machines, and controls the slave transport's speed so that the slave timecode is playing at the same rate as the master timecode. Thus the slave tape is being resolved to the master timecode. The actual speed of the Master is not even an issue--within reason--since the Slave will be locked precisely to it. Likewise, any variations in the Master tape's speed will be followed by the Slave.

During sweetening, tachometer (tach) pulses from the transports may be used by the synchronizer to keep rough track of tape position when timecode cannot be read (such as fastwinding). This is known as tach-pulse updating. But only the timecode can be used to resolve the ATR's speed! This is because only the timecode has the precise relationship to the audio necessary to retain sync with the picture.

Procedures for Audio and Timecode

Since the sweetening process involves a number of audio and timecode transfers between tapes, care must be taken that the audio and timecode never become "unlocked" from one another. These types of mistakes are the easiest to make and, naturally, are the hardest to repair!

Basically there are two methods that can be used to record a tape which will preserve the original relationship between the timecode and audio:

Record the audio and timecode simultaneously, with the recorder's capstan speed internally locked. This applies when the audio and timecode already have a fixed relationship (say, from the edit master). Since they are recorded onto the tape at the same time, they are subject to the same speed variations and their relationship remains constant. Even if the recorder were running off-speed, when the tape is later resolved to the master's timecode, the playback speed will be adjusted so that the audio is correct.

Record the timecode first and resolve the ATR while the audio is being recorded. In this case the tape always runs at the correct speed while recording the audio, because it is already being resolved to its timecode. If this method is used when transferring audio and timecode tracks, there must be an external reference for the recorder to resolve to while the audio is being transferred, and this reference must already be related to the audio. The external reference usually comes from the source tape that is providing the audio being transferred (or from a sync generator, in a video house).

Either method results in a tape that can be resolved so that the audio plays at the speed which will keep it in sync with the video it relates to. The method chosen depends on the circumstances at the time. Most disasters result from careless transfers, where the audio and timecode somehow become unlocked (for example, being dubbed individually without resolving the recorder).

Finally, the exception to these rules is a circumstance where the audio and timecode do not have a fixed relationship to begin with. For example, when adding timecode to a tape of background music which will be layed-up simply by where it sounds appropriate. In this case, the timecode is recorded "wild", with the recorder capstan internally locked. From this point on, however, the tape must always be resolved on playback, if the newly established relationship is to be of any use

Tricks and Miscellaneous

Bear in mind that by striping timecode on an audio tape we are artificially "creating" frames, for the purpose of resolving the tape speed and synchronizing. These timecode frames have no inherent relationship to the audio because audio does not come in discrete "packets". For this reason the timecode will not represent every audio event discretely, since it occurs only 30 times per second (33 milliseconds, or approximately 1/2" of tape at 15ips). This is not generally a problem, but it can lead to occasional questions of accuracy when checking sync between sound and picture. Fortunately, most people cannot visually perceive lip-sync differences of less than about two frames, and the most any audio event can be off is 1/2 frame of timecode (if it occurs exactly between two timecode numbers). But a buildup of these kind of offsets can become noticeable.

Actually, the tolerances that audio engineers are accustomed to are a good deal tighter than video (audio comb-filtering becomes evident at around 5ms delay). Synchronizers often provide the ability to offest Master and Slave timecode numbers by fractions of a frame. This capability can be used to correct one-time sync errors (such as a jump in the timecode causing lip-sync to shift suddenly), but cannot be used to correct for unlocked audio and timecode--at least not without numerous punch-ins to "pull up" the sync.

When making audio and timecode transfers during sweetening, it is a good idea to check the sync at each step to be sure everything is ok. One method to do this is by listening to the reference audio on the 3/4" videotape, while also listening to the newly transferred audio running in sync with the videotape. If the tranfer was succesful, the two audio tracks should overlay perfectly. If there is any phasing or echo, something may be wrong! This method is more critical than simply watching the mouth of the person on camera, and will disclose sync problems which worsen as the program runs.

Another method of checking sync is to view the timecode from the Master and Slave on an oscilloscope. Using dual-trace, observe both timecode signals while triggering the scope on the sync word of the VTR's timecode (or the video output of the VTR). The sweep rate will need to be very slow. Over a period of several seconds (up to a minute) the two timecode signals should not drift in relation to each other. If they do, something has become unlocked, or a machine is not being resolved (check the capstan reference switch). This method can also be used to check that the timecode on the videotape is locked to the video.

If it can be determined that the difference in timecode rates is around .1%, that may indicate that the synchronizer/timecode generator was locked to AC mains at some point, since .1% is the difference between 59.94 and 60 Hz. Or it may indicate a problem with mixed timecode rates, or inadvertant cross--resolving during some part of the process (cross-resolving does have some uses, but they are beyond the scope of this article).

A final point is to remember that a video sweetening project involves interaction with a video facility--where some things are done differently! Make sure all your tapes are adequately labeled with the usual audio information, and also location and type of timecode. Do not assume that the equipment used by a video house to layoff or layback audio is similar to a synchronizer used at a studio, or that it behaves the same way (for example, slow-relocking after a splice). Most importantly, do not hesitate to get further information if something about a project is not understood. There are too many variables involved to allow questions to go unanswered!

Further Information

1. Larry Blake, "Timecode and Synchronization", Recording Engineer/Producer, Aug/Sept/Oct. 1987
2. "SMPTE/EBU Longitudinal and Vertical Timecode", EECO, Inc.
3. "Using Timecode in the Reel World", Nagra/Kudelski Electronics
4. M. Hubatka, F. Hull, R. Sanders, "Audio Sweetening for Film and TV", Tab Books, Inc. 1985

General Rules for A/V Sync Work

1. The relationship between sync audio and its timecode must NEVER be altered once it is established. Audio transfers can only be made by methods which do not alter this relationship.
2. The timecode generator must be locked to its internal crystal (non-resolved mode) when generating new code for audio tapes. It must be locked to the source when dubbing timecode. It must be locked to the video when post-coding videotapes.
3. Audio transports must be in "internal" capstan reference mode when recording timecode wild, or simultaneously with audio.
4. The timecode on a tape being sent for layback should be the same rate as the code on the original audio source tape (this will happen automatically if the timecode and audio are always dubbed together).
5. Timecode should always be restored or regenerated when it is dubbed.
6. Never attempt a hard record or assemble edit unless video is being fed to the VTR.
7. Audio insert edits can only be made on videotapes with control track present.
8. There MUST be at least 10 seconds of timecode before the start of audio material to allow for synchronizer preroll.

Acceptable Methods

Dub both audio and timecode simultaneously to a "wild" (internall locked) recorder.
Dub audio to a timecoded tape, with the machines locked (resolved) together. The prestriped timecode should match the original.

Unacceptable Methods

Never dub audio first, then timecode (or viceversa) when copying a tape with an established audio/timecode relationship.
Never dub audio to a prestriped tape unless the two machines are locked together, or locked to a common sync generator.

Glossary of Terms

ADR — Automatic Dialogue Replacement, a method of dubbing (re-recording) dialogue for motion pictures using a VTR and multitrack ATR to record the actor's lines while he or she watches the scene on videotape.

Burn-in — A videotape in which a "window" displaying the timecode count on the tape is superimposed over part of the picture. Eliminates the need to watch a TC reader.

Control track — A series of pulses (at field rate) recorded automatically on a videotape to resolve the playback speed. Control track is essential, and only recorded if the VTR receives video (see Section 2).

Double system sound — Sound and picture on seperate transports (film and Nagra, film and mag dubber, videotape and audio tape).

EBU — European Broadcast Union (identifies 25 Hz timecode standard).

Edit Master — Video industry term for the tape which is the finished (edited) program.

Field rate — Frequency at which video fields occur (59.94 Hz in NTSC, 50 Hz in PAL).

Frame rate — Frequency at which video frames occur (29.97 Hz in NTSC, 25 Hz in PAL).

Jam-Sync — The process of driving a timecode generator with a reader to regenerate or reclock code from a tape (see Regenerating, Section 3).

Layback — Transferring of the finished audio mix back onto the video edit master.

Layoff — Transferring of audio and timecode from the video edit master to an audio tape.

Layup — Transferring of audio onto the multitrack or hard-disk.

Lip-sync — The sound and picture of someone speaking on camera, which should match.

NTSC — National Television Standards Committee, established U.S. color television system.

PAL — Television standard used in parts of Europe (25 Hz frame rate, 625 lines per frame).

Pilot-tone — A sinewave signal recorded by Nagra field ATRs at a known frequency, used to resolve the tape speed on playback to retain sync with film camera footage.

Resolving — The process of regulating tape speed by comparing a reference signal on the tape with an external reference and adjusting the speed so that they match (see Section 4).

Sync —
1. When two or more audio or video elements occur simultaneously.
2. The process of resolving two or more transports to a common reference.
3. A generic term for any signal that provides a synchronizing reference.

Sync sound — Any sound that has precisely corresponding picture (lip-sync, sound effects).

Video editing — The process of assembling a program from many tapes of original footage, and adding transitions (cuts, wipes, dissolves) and special effects. Video editing is done by re-recording the desired footage onto a new tape (edit master).

VITC — Vertical Interval Timecode, a timecode signal that is written in the vertical interval by the rotating video heads, thus allowing it to be read when tape is not moving. Requires special equipment to read and write.

Wild — An audio element (or transport) that is not running "in sync" with a reference.