International Telecommunication Union

ITU-R

Radiocommunication Sector of ITU

Recommendation ITU-R BS.2032
(01/2013)

Synchronization of digital audio sample clock to video references

BS Series
Broadcasting service (sound)

International
Telecommunication
Union

Foreword

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services, and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

Policy on Intellectual Property Right (IPR)

ITU-R policy on IPR is described in the Common Patent Policy for ITU-T/ITU-R/ISO/IEC referenced in Annex 1 of Resolution ITU-R 1. Forms to be used for the submission of patent statements and licensing declarations by patent holders are available from http://www.itu.int/ITU-R/go/patents/en where the Guidelines for Implementation of the Common Patent Policy for ITU-T/ITU-R/ISO/IEC and the ITU-R patent information database can also be found.

<div style="border:1px solid">

Series of ITU-R Recommendations

(Also available online at http://www.itu.int/publ/R-REC/en)

Series	Title
BO	Satellite delivery
BR	Recording for production, archival and play-out; film for television
BS	**Broadcasting service (sound)**
BT	Broadcasting service (television)
F	Fixed service
M	Mobile, radiodetermination, amateur and related satellite services
P	Radiowave propagation
RA	Radio astronomy
RS	Remote sensing systems
S	Fixed-satellite service
SA	Space applications and meteorology
SF	Frequency sharing and coordination between fixed-satellite and fixed service systems
SM	Spectrum management
SNG	Satellite news gathering
TF	Time signals and frequency standards emissions
V	Vocabulary and related subjects

</div>

Note: *This ITU-R Recommendation was approved in English under the procedure detailed in Resolution ITU-R 1.*

Electronic Publication
Geneva, 2013

© ITU 2013

RECOMMENDATION ITU-R BS.2032

Synchronization of digital audio sample clock to video references

(2013)

Scope

This Recommendation provides methods for synchronizing interconnected digital audio equipment, and to address synchronization of the audio sample clocks to video reference signals.

Synchronization of digital audio signals is a necessary function for the exchange of signals between equipment. The objective of synchronization is primarily to time align sample clocks within digital audio signal sources and align them with video frames/fields.

The provisions make use of the two-channel digital audio interface standard for professional use, Recommendation ITU-R BS.647.

The ITU Radiocommunication Assembly,

considering

a)   that digital audio implementations have been widely deployed;

b)   that there would be advantages if all digital audio equipment use the same synchronizing signal;

c)   that Recommendation ITU-R BS.647 requires the use of a video synchronizing signal;

d)   that a common synchronizing signal should be used for digital video and audio equipment;

e)   that for non-TV-video related applications, a reference signal is still required for multiple source applications,

recommends

1   that synchronization of digital audio sample clock to video references defined in Annex 1 should be used;

2   that compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g. interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements.

Annex 1

The primary area of application is the digital interconnection and synchronization of digital audio equipment contained within the production and acquisition environment. A further important objective is to define a means to synchronize digital audio sample clocks to a video reference signal.

1   Modes of operation

1.1   General

Equipment should provide the ability to lock an internal sample-clock generator to a Digital Audio Reference Signal (DARS).

1.2   Synchronizing methods

Equipment should be synchronized by one of three methods.

NOTE – In some circumstances, "word clock" may be used in a manner similar to the DARS signal. Word clock is not part of this Recommendation, but is described in Appendix 2.

1.2.1   DARS referenced

Equipment is synchronized to a DARS, which ensures that all input-output equipment sample clocks are locked to the same reference frequency and within the phase tolerance specified in this Recommendation.

In situations where some 96 kHz signals are carried in the mode described in Recommendation ITU-R BS.647 as "single channel double sampling frequency mode", it is necessary that the synchronizing reference signal has a frequency component at 48 kHz or lower, in order that two channels comprising a stereo pair shall be correctly related. Appendix 1 illustrates preferred phase relationships.

1.2.2   Video referenced

A master video reference is used to derive a DARS, locking video and audio signals at the sample-rate level and to video frame boundaries.

1.2.3   GPS referenced

A GPS receiver is used to reference a DARS, providing frequency and phase (from one-second pulses), and time-of-day sample address code in bytes 18 to 21 of channel status to support a time of day reference in locked equipment.

1.3   DARS distribution

The DARS shall be distributed in compliance with Recommendation ITU-R BS.647.

1.4   External signals

1.4.1   General

When connecting external signals to an otherwise synchronous digital audio studio or production facility, § 1.4.2 shall apply.

1.4.2 Phase correction

Where the incoming signal is identical in sample frequency but is out of phase with the DARS, Recommendation ITU-R BS.647 frame alignment shall be applied.

1.5 Video referencing

1.5.1 General

In the case of a combined video and audio environment, the source of the DARS shall be locked to the video source so that the mathematical relationships given in Table 1 are obtained precisely.

TABLE 1

Audio-video synchronization

Sample rate (kHz)	Samples per frame					
	25 Hz	**30 Hz**	**29.97 Hz**	**50 Hz**	**60 Hz**	**59.94 Hz**
48	1920	1600	8008/5	960	800	4004/5
96	3840	3200	16016/5	1920	1600	8008/5

1.5.2 Integer ratios

For video systems with an integer number of Recommendation ITU-R BS.647 frames in one video frame, the Recommendation ITU-R BS.647 audio can be locked synchronously to the video.

1.5.3 Non-integer ratios – 5-frame lock

For video systems with fractions of Recommendation ITU-R BS.647 frames in one video frame, Recommendation ITU-R BS.647 audio sample clocks can be locked using the video reference signal defined in Appendix 3. In this case, there is a 5-frame locking cycle.

1.5.4 Non-integer ratios – Random lock

It is common practice that the audio sample clock is synchronized on a video frame at random. While the frequency of the sample clock will be locked, the absolute phasing will not be maintained. Systems that utilize this means of locking do require buffers to ensure phasing of all equipment.

NOTE – While every effort is taken to maintain a 5-frame sequence, it is not unusual for there to be less or greater number of samples within the 5-frame sequence. These discrepancies occur when material is edited or switched.

2 Recommended practice for equipment synchronization

2.1 DARS requirements

2.1.1 DARS format

The DARS shall have the format and electrical configuration of the two-channel digital audio interface and use the same connector as given in Recommendation ITU-R BS.647. However, the basic structure of the digital audio interface format, where only the preamble is active, shall be acceptable as a digital audio synchronizing signal.

2.1.2   DARS grade

A DARS may be categorized as either grade 1 or grade 2. See also § 2.2.

To be compliant with this Recommendation, only a grade 1 DARS signal is permitted.

2.1.2.1   Grade 1

A grade 1 DARS is a high-accuracy signal intended for synchronizing systematically a multiple-studio complex and may also be used for a stand-alone studio.

2.1.3   DARS identification

A DARS, which has the prime purpose of studio synchronization, shall be identified as to its intended use by byte 4, bits 0 and 1, of Recommendation ITU-R BS.647 channel status:

TABLE 2

Byte 4, bits 0 and 1: DARS

Bits	1 0	Digital audio reference signal
States	0 0	Not a reference signal (default)
	1 0	Grade 1 reference signal
	0 1	Grade 2 reference signal*
	1 1	Reserved and not to be used until further defined

*   Not used in this Recommendation.

2.1.4   DARS non-audio content

A DARS shall be identified in channel status as "not linear PCM" when it contains other data rendering it unusable as a normal audio signal. See Recommendation ITU-R BS.647, channel status.

NOTE – A DARS may carry an alignment tone in linear PCM form.

2.1.5   DARS date and time

Where a DARS is used to carry date and time information in the user channel, this shall be signalled in channel status using the bits specified in Recommendation ITU-R BS.647[1] for the carriage of metadata in the user channel.

2.1.6   DARS sampling frequency

Sampling frequencies distributed by a DARS shall be 48 kHz or 96 kHz (see § 2.2.2).

2.2   Sample frequency tolerances in equipment

2.2.1   Long-term frequency accuracy

A grade 1 DARS signal shall maintain a long-term frequency accuracy within ±1 part per million (ppm) relative to its rated frequency. Equipment designed to provide grade 1 reference signals shall only be required to lock to other grade 1 reference signals.

---

[1]   Recommendation ITU-R BS.647-3 (2011), Part 3 § 3.3.

2.2.2   Capture range

The minimum capture range of equipment oscillators designed to lock to external inputs should be ±2 ppm for grade 1 equipment.

2.3   Equipment timing relationships

2.3.1   General

The timing-reference point is used to define the timing relationship between the DARS and digital audio input and output signals.

2.3.1.1   Output timing phase

The difference between the timing-reference points of the DARS and all output signals, at the equipment connector points, shall be less than ±5 % of the Recommendation ITU-R BS.647 frame period.

The timing reference point of an output signal with double-rate or higher sampling frequency shall be within the stated tolerance at its own Recommendation ITU-R BS.647 frame rate.

2.3.1.2   Device delay

Receivers shall be designed so that the number of samples of delay through a device remains constant and known while the difference between the timing-reference points of the DARS and all input signals is less than ±25 % of the Recommendation ITU-R BS.647 frame period.

2.3.2   Timing limits

Table 3 specifies tolerances values for the sample frequencies for use with this Recommendation.

TABLE 3

Synchronization of digital audio: limits

Professional sampling frequency (kHz)	Synchronization window µs		
	$1/f_s$	Permitted variation, Input (§ 2.3.1.2)	Permitted variation, Output (§ 2.3.1.1)
48	20.83	±5.2	±1.0
96	10.41	±2.6	±0.5

2.4   Video reference[2]

One of the objectives of this Recommendation is to define the start of the X or Z preamble of the DARS to a known point in the video reference signal.

2.4.1   25 or 50 Hz reference

It is common practice in 25 or 50 Hz related TV systems to utilize a colour black analogue PAL signal. The PAL signal waveform is defined in Recommendation ITU-R BT.1700.

---

[2]   For the purpose of this Recommendation, the interface line numbers are used for the establishment of the timing references.
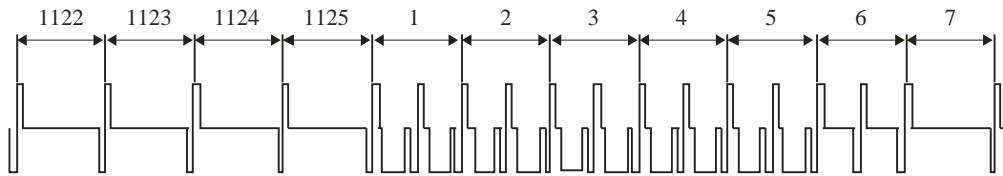
2.4.1.1 Video interface format timing

V and H sync phase relationship between 1125/50/I and 1125/25/P tri-level sync, and 625/50/I analogue sync is shown in Fig. 1.
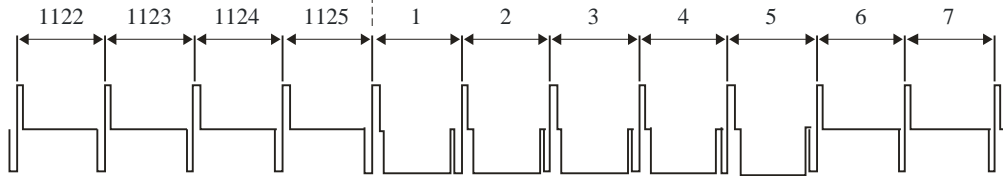
FIGURE 1

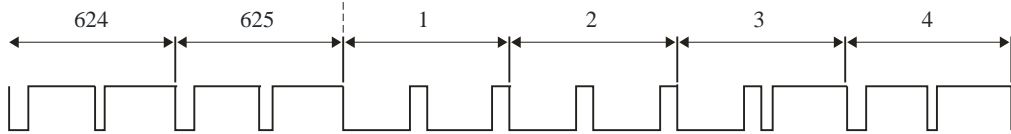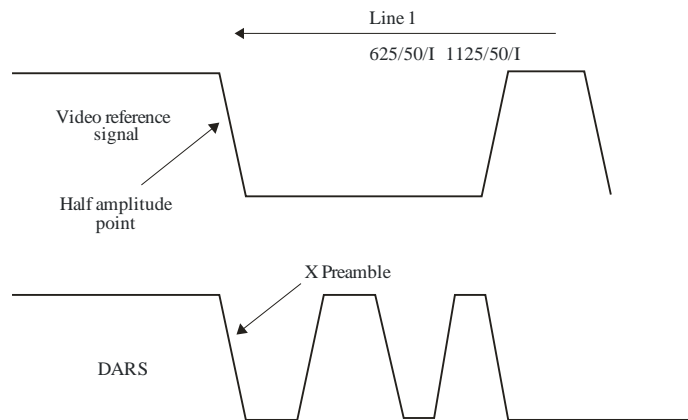**Line numbering of 1125/50/I, 1125/25/PsF, 1125/25/P, and 625/50/I formats**

BS.2032-01

FIGURE 2

DARS X preamble reference

BS.2032-02

The DARS X Preamble shall be referenced to the half-amplitude point of the leading edge of the synchronization pulse of line 1 of the television signal on every video frame.
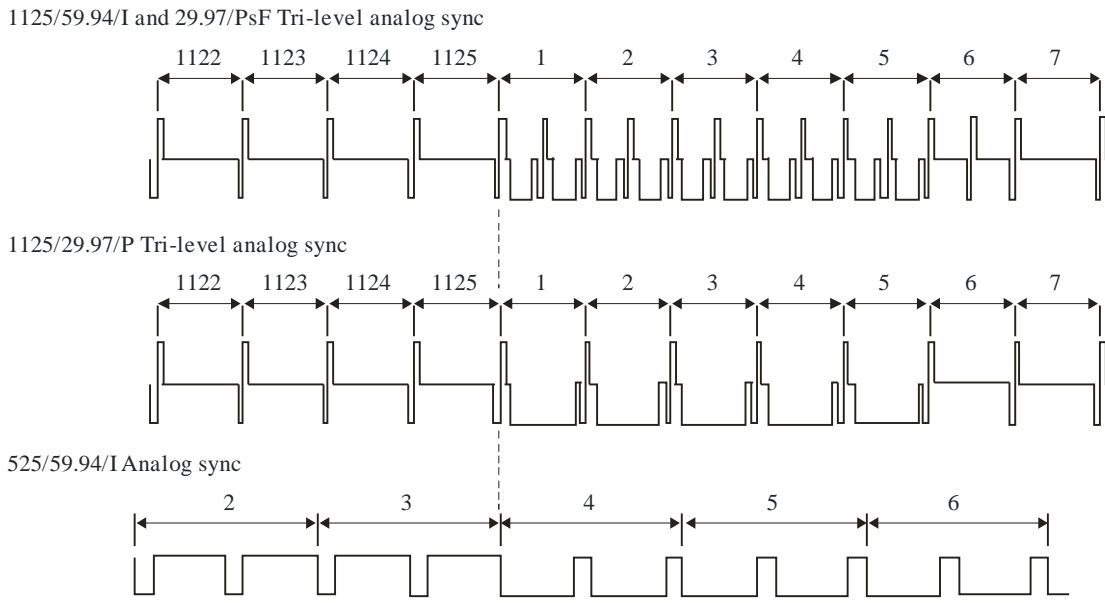
2.4.2 30 or 60[3] Hz reference

V and H sync phase relationship between 1125/59.94/I and 1125/29.97/P tri-level sync, and 525/60/I analogue sync is shown in Fig. 3.

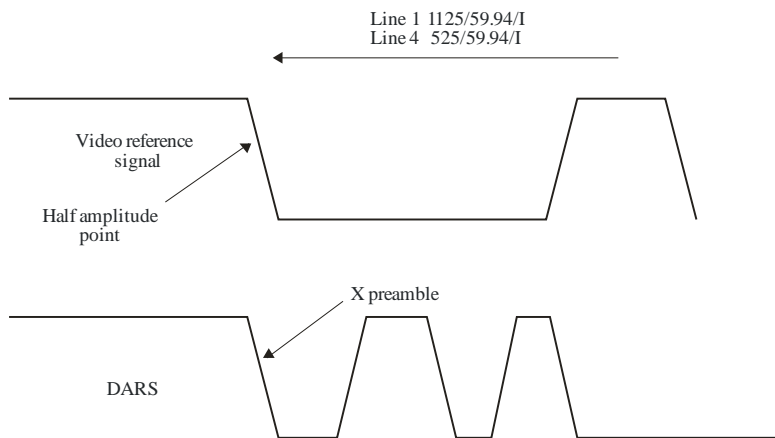### 2.4.2.1 Video interface format timing

FIGURE 3

Line numbering of 1125/59.94/I, 1125/29.97/PsF, 1125/29.97/P and 525/59.94/I

1125/59.94/I and 29.97/PsF Tri-level analog sync

1125/29.97/P Tri-level analog sync

525/59.94/I Analog sync

BS.2032-03

FIGURE 4

DARS X preamble reference

Line 1 1125/59.94/I
Line 4 525/59.94/I

Video reference signal

Half amplitude point

X preamble

DARS

Drawing not to scale

BS.2032-04

---

[3]   Precise values are 30/1.001 and 60/1.001.

The DARS X Preamble shall be referenced to the half-amplitude point of the leading edge of the synchronization pulse of line 1 for 1080 systems and line 4 for 525 line systems. Alignment of the DARS X preamble should be timed for every 5th frame. There is a possibility that there could be a ±1 sample offset. The 525 line reference signal with 5 frame identification is shown in Appendix 3.

2.4.3   Overall phase tolerance

To aid practical implementations, there shall be a phase tolerance of ±5% of the Recommendation ITU-R BS.647 frame period between the DARS signals and the video reference, in addition to the ±5% tolerance defined for digital audio synchronization at the system outputs in § 2.3.1.1.

2.5   System practice

Good engineering practice requires that timing differences between signal paths be minimized, to avoid timing errors accumulating with a risk of loss of synchronism.

2.6   Clock jitter

Jitter noise may be either random or in the form of modulation, which at frequencies less than sample rate will cause a timing error to accumulate according to the amplitude and frequency of the modulation waveform.

NOTE – Recommendation ITU-R BS.647 defines limits for jitter on the digital audio interface.
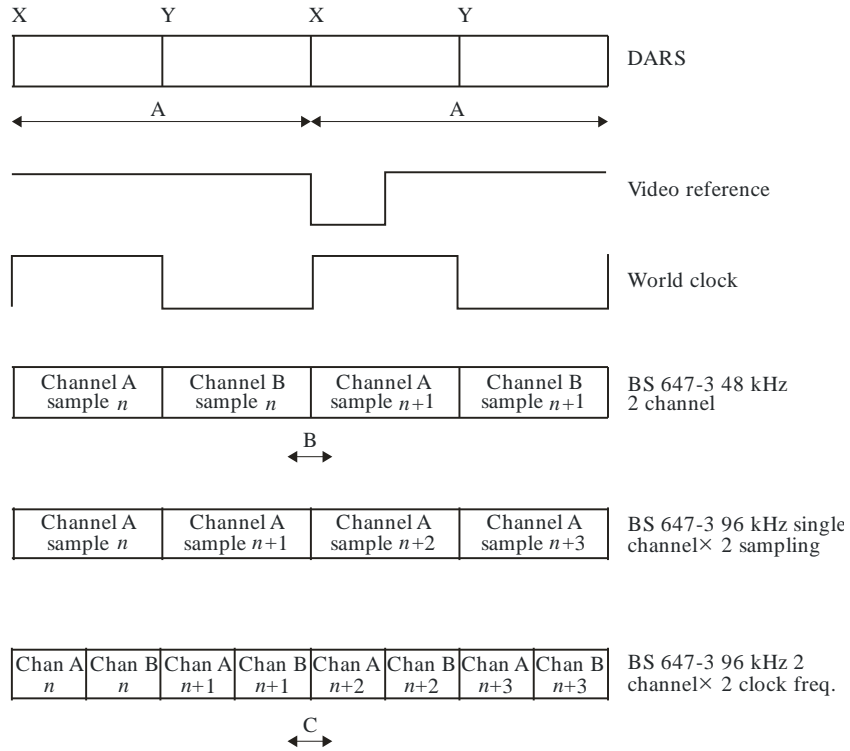
3   Date and time

Flagging of date and time in channel status is specified in § 2.1.5. This may take a convenient form for transfer to Recommendation ITU-R BS.647 metadata stream.

Appendix 1
(Informative)

Timing relationships

FIGURE 5

Preferred phase relationships and channel usage

BS.2032-05

	Timing values
A	20.5 μs – frame period at 48 kHz sampling frequency
B	±1 μs tolerance
C	±0.5 μs tolerance

Appendix 2
(Informative)

Word clock

It is possible to meet all the timing requirements of this Recommendation by means of a square wave at sampling frequency basic rate, commonly called word clock. It is used between different pieces of equipment to provide sampling frequency locking of various sources.

This signal is not standardized and the parameters quoted are merely examples. The signal is commonly carried on a coaxial cable, so that a single output can synchronize several receiving equipments by looping the signal through each in turn, and possibly terminating the cable with a 75 Ohm resistor at the far end.

The transmitted signal may vary in peak-to-peak amplitude from 1 V up to 5 V, and be either AC- or DC-coupled.

The driver level most likely to provide the functionality required is a full 5 V, DC coupled, and able to drive a 75 Ohm load.

Where new equipment is designed to use a word-clock signal, it is recommended that the rising edge is treated as the timing reference point referred to in § 2.3.

The expression "word clock" is also used at circuit-board level to describe various sampling-frequency logic signals.

Word clock is commonly used with digital audio signals other than Recommendation ITU-R BS.647.

Appendix 3
(Informative)

525 line reference signal with 5 frame identification

FIGURE 6

525 line reference signal with 5 frame identification
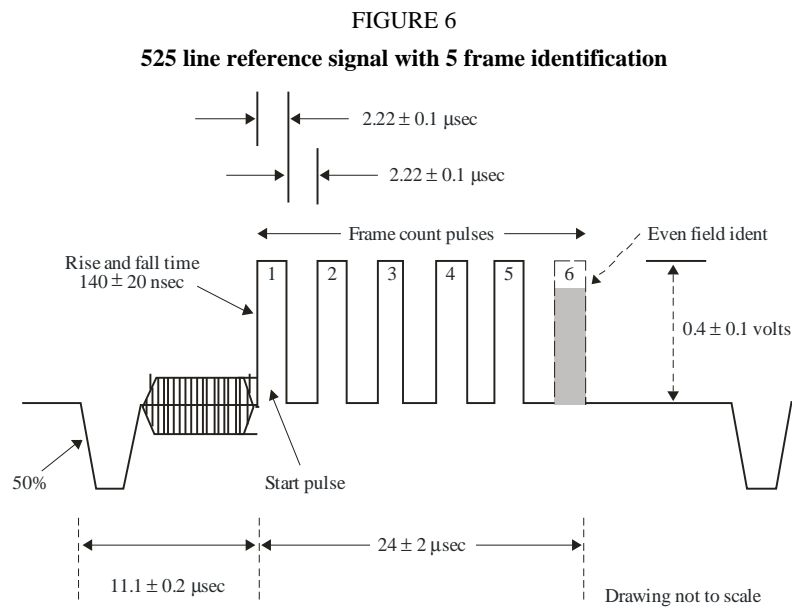
BS.2032-06

The ten-field (five-frame) sequence identification is coded as follows:

–   The identification signal is inserted on Lines 15 and 278.

–   The first pulse is always present, and functions as a start pulse.

–   There follows a string of between zero and four frame count pulses that increase by one on line 15 (each odd field). The sixth pulse is present on line 278 (even field) only. Pulses are separated by spaces of duration equal to the pulses.

The start of the ten-field sequence is unspecified and is not aligned with any time or time code value. More details are covered in Appendix 4.

Appendix 4
(Informative)

Informative references

SMPTE 318M-1999, *Synchronization of 59.94 or 50 Hertz related video and audio systems in analogue and digital areas*: Society of Motion Picture and Television Engineers, White Plains, NY., US.

st0318-1999.pdf