International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

G.168

(04/2015)

SERIES G: TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

International telephone connections and circuits – Apparatus associated with long-distance telephone circuits

---

Digital network echo cancellers

Recommendation ITU-T G.168

ITU-T G-SERIES RECOMMENDATIONS

TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

INTERNATIONAL TELEPHONE CONNECTIONS AND CIRCUITS	G.100–G.199
Transmission planning and the E-model	G.100–G.109
General Recommendations on the transmission quality for an entire international telephone connection	G.110–G.119
General characteristics of national systems forming part of international connections	G.120–G.129
General characteristics of the 4-wire chain formed by the international circuits and national extension circuits	G.130–G.139
General characteristics of the 4-wire chain of international circuits; international transit	G.140–G.149
General characteristics of international telephone circuits and national extension circuits	G.150–G.159
Apparatus associated with long-distance telephone circuits	G.160–G.169
Transmission plan aspects of special circuits and connections using the international telephone connection network	G.170–G.179
Protection and restoration of transmission systems	G.180–G.189
Software tools for transmission systems	G.190–G.199
GENERAL CHARACTERISTICS COMMON TO ALL ANALOGUE CARRIER-TRANSMISSION SYSTEMS	G.200–G.299
INDIVIDUAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON METALLIC LINES	G.300–G.399
GENERAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON RADIO-RELAY OR SATELLITE LINKS AND INTERCONNECTION WITH METALLIC LINES	G.400–G.449
COORDINATION OF RADIOTELEPHONY AND LINE TELEPHONY	G.450–G.499
TRANSMISSION MEDIA AND OPTICAL SYSTEMS CHARACTERISTICS	G.600–G.699
DIGITAL TERMINAL EQUIPMENTS	G.700–G.799
DIGITAL NETWORKS	G.800–G.899
DIGITAL SECTIONS AND DIGITAL LINE SYSTEM	G.900–G.999
MULTIMEDIA QUALITY OF SERVICE AND PERFORMANCE – GENERIC AND USER-RELATED ASPECTS	G.1000–G.1999
TRANSMISSION MEDIA CHARACTERISTICS	G.6000–G.6999
DATA OVER TRANSPORT – GENERIC ASPECTS	G.7000–G.7999
PACKET OVER TRANSPORT ASPECTS	G.8000–G.8999
ACCESS NETWORKS	G.9000–G.9999

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T G.168

Digital network echo cancellers

Summary

Echo has a major effect on voice quality in telecommunication networks. The objectionable effect of echo results from a combination of reflections from network components such as 2- to 4-wire converters, together with signal processing and transmission delay. Echo may cause users difficulty in talking or listening over a telephone connection. It may also affect the transmission of voiceband data, fax, and text telephones.

Digital network echo cancellers are designed to eliminate echo for the user and to allow successful transmission of voiceband data and fax. Recommendation ITU-T G.168 describes the characteristics of an echo canceller, including the requirement for in-band tone disabling and other control mechanisms. It also describes a number of laboratory tests that should be performed on an echo canceller to assess its performance under conditions likely to be experienced in the network.

This revision of ITU-T G.168 introduces new Annex E and Appendix VIII, as well as corrections and clarifications to the previous edition of the Recommendation.

New Annex E applies to embedded echo cancellers (EECs). Its purpose is to define performance requirements for embedded echo cancellers, for example where access to control signals such as h-register freeze is not available.

The various tests in the main body of ITU-T G.168 require external access to certain echo canceller control signals so that the individual parts of an echo canceller function can be tested in the laboratory. These control signals include h-register freeze and NLP on/off. Access to 4-wire TDM ports is also required for injection of test signals and to obtain the test results. Embedded echo cancellers often have packet or 2-wire interfaces, thus placing additional limitations on testing.

In its current form, this annex defines performance requirements for embedded echo cancellers including a test set-up that may be used to perform tests.

New Appendix VIII defines suggested test methodologies which may be used to test compliance or partial compliance to Annex E.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T G.168	1997-04-08	15	11.1002/1000/4016
2.0	ITU-T G.168	2000-04-04	15	11.1002/1000/5056
3.0	ITU-T G.168	2002-06-13	15	11.1002/1000/6049
4.0	ITU-T G.168	2004-08-06	15	11.1002/1000/7318
5.0	ITU-T G.168	2007-01-13	16	11.1002/1000/9018
6.0	ITU-T G.168	2009-03-16	16	11.1002/1000/9685
7.0	ITU-T G.168	2012-02-29	16	11.1002/1000/11536
7.1	ITU-T G.168 (2012) Amd. 1	2012-05-11	16	11.1002/1000/11681
8.0	ITU-T G.168 (V8)	2015-04-29	16	11.1002/1000/12451

* To access the Recommendation, type the URL http://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID. For example, http://handle.itu.int/11.1002/1000/11830-en.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

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. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at http://www.itu.int/ITU-T/ipr/.

Table of Contents

Recommendation ITU-T G.168

Digital network echo cancellers

1 Scope

Echo cancellers are voice operated devices placed in the 4-wire portion of a circuit (which may be an individual circuit path or a path carrying a multiplexed signal) and are used for reducing the echo by subtracting an estimated echo from the circuit echo (see Figure 1).

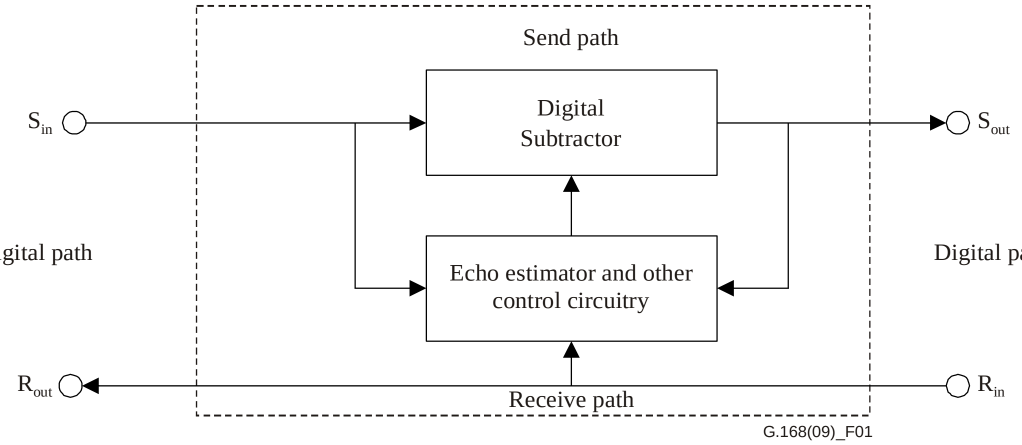


Figure 1 – Digital transmission echo canceller using digital subtraction

NOTE – Functionally, a digital echo canceller (DEC) interfaces at 64 kbit/s. However, 24 or 30 digital echo cancellers, for example, may be combined corresponding to the primary digital hierarchy levels of 1544 kbit/s or 2048 kbit/s, respectively.

This Recommendation is applicable to the design of echo cancellers using digital techniques, and intended for use in circuits where the delay exceeds the limits specified by [ITU-T G.114] and [ITU-T G.131]. It is necessary for all echo control devices used on international connections to be compatible with each other. Echo cancellers designed to this Recommendation will be compatible with each other, with echo cancellers designed in accordance with [ITU-T G.165], and with echo suppressors designed in accordance with [ITU-T G.164]. Compatibility is defined as follows:

Given:

1) that a particular type of echo control device (say Type I) has been designed so that satisfactory performance is achieved when any practical connections is equipped with a pair of such devices; and

2) that another particular type of echo control device (say Type II) has been likewise designed,

then Type II is said to be compatible with Type I if it is possible to replace an echo control device of one type with one of the other type, without degrading the performance of the connection to an unsatisfactory level. In this sense, compatibility does not imply that the same test apparatus or methods can necessarily be used to test both Type I and Type II echo control devices.

Freedom is permitted in design details not covered by the requirements. This Recommendation is for the design of digital echo cancellers and defines tests that ensure that echo canceller performance is adequate under wider network conditions than specified in [ITU-T G.165], such as performance on voice, fax, residual acoustic echo signals, and mobile networks.

This Recommendation does not apply to echo cancellation through active 2-wire/4-wire hybrids or 2-wire repeaters. This Recommendation does not cover acoustic echo cancellation as per [ITU-T P.340].

This Recommendation defines objective tests that if passed will ensure (but will not guarantee) a minimum level of performance when installed in the network. An echo canceller which passes these tests should not harm equipment nor degrade transmission performance of voiceband signals and services below acceptable limits. These tests are lab-type tests and are not designed to be run in-service. Also, these tests are objective tests and do not replace or eliminate the need for subjective tests to measure the perceived quality of echo cancellers. Echo cancellers are complex devices with multiple parameters, and the correlation of these parameters and their interactions to the subjective quality of an echo canceller is difficult to specify. Clause I.7.5 gives some guidelines on how subjective test results were used in order to develop objective tests. [ITU-T P.851] describes methods and procedures for conducting subjective performance evaluation of network echo cancellers. Thus, this Recommendation does not specify nor imply selection criteria; however, guidelines are provided herein, and Administrations have the freedom to specify criteria in their selection process. This set of criteria may include some or all of the thresholds and/or tests in this Recommendation.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T E.180] Recommendation ITU-T E.180/Q.35 (1998), *Technical characteristics of tones for the telephone service.*
<http://www.itu.int/rec/T-REC-E.180/>

[ITU-T G.114] Recommendation ITU-T G.114 (2003), *One-way transmission time*.
<http://www.itu.int/rec/T-REC-G.114/>

[ITU-T G.122] Recommendation ITU-T G.122 (1993), *Influence of national systems on stability and talker echo in international connections*.
<http://www.itu.int/rec/T-REC-G.122/>

[ITU-T G.131] Recommendation ITU-T G.131 (2003), *Talker echo and its control*.
<http://www.itu.int/rec/T-REC-G.131/>

[ITU-T G.161] Recommendation ITU-T G.161 (2012), *Interaction aspects of signal processing network equipment*.
<http://www.itu.int/rec/T-REC-G.161/>

[ITU-T G.164] Recommendation ITU-T G.164 (1988), *Echo suppressors*.
<http://www.itu.int/rec/T-REC-G.164/>

[ITU-T G.165] Recommendation ITU-T G.165 (1993), *Echo cancellers*.
<http://www.itu.int/rec/T-REC-G.165/>

[ITU-T G.711] Recommendation ITU-T G.711 (1988), *Pulse code modulation (PCM) of voice frequencies*.
<http://www.itu.int/rec/T-REC-G.711/>

[ITU-T G.712] Recommendation ITU-T G.712 (2001), *Transmission performance characteristics of pulse code modulation channels*.
<http://www.itu.int/rec/T-REC-G.712/>

[ITU-T P.50] Recommendation ITU-T P.50 (1999), *Artificial voices*.
<http://www.itu.int/rec/T-REC-P.50/>

[ITU-T P.56] Recommendation ITU-T P.56 (2011), *Objective measurement of active speech level*.
<http://www.itu.int/rec/T-REC-P.56/>

[ITU-T P.310] Recommendation ITU-T P.310 (2009), *Transmission characteristics for telephone band (300-3400 Hz) digital telephones*.
<http://www.itu.int/rec/T-REC-P.310/>

[ITU-T P.340] Recommendation ITU-T P.340 (2000), *Transmission characteristics and speech quality parameters of hands-free terminals*.
<http://www.itu.int/rec/T-REC-P.340/>

[ITU-T P.501] Recommendation ITU-T P.501 (2012), *Test signals for use in telephonometry*.
<http://www.itu.int/rec/T-REC-P.501/>

[ITU-T P.800] Recommendation ITU-T P.800 (1996), *Methods for subjective determination of transmission quality*.
<http://www.itu.int/rec/T-REC-P.800/>

[ITU-T P.851] Recommendation ITU-T P.851 (2003), *Subjective quality evaluation of telephone services based on spoken dialogue systems.*
<http://www.itu.int/rec/T-REC-P.851/>

[ITU-T Q.23] Recommendation ITU-T Q.23 (1988), *Technical features of push-button telephone sets.*
<http://www.itu.int/rec/T-REC-Q.23/>

[ITU-T Q.24] Recommendation ITU-T Q.24 (1988), *Multifrequency push-button signal reception.*
<http://www.itu.int/rec/T-REC-Q.24/>

[ITU-T Q.143] Recommendation ITU-T Q.143 (1988), *Specifications of signalling system No. 5 – Line signal sender*.
<http://www.itu.int/rec/T-REC-Q.143/>

[ITU-T Q.144] Recommendation ITU-T Q.144 (1993), *Specifications of signalling system No. 5 – Line signal receiver*.
<http://www.itu.int/rec/T-REC-Q.144/>

[ITU-T Q.271] Recommendation ITU-T Q.271 (1988), *Specifications of signalling system No. 6 – Continuity check of the speech path: General*.
<http://www.itu.int/rec/T-REC-Q.271/>

[ITU-T Q.552] Recommendation ITU-T Q.552 (2001), *Transmission characteristics at 2-wire analogue interfaces of digital exchange*.
<http://www.itu.int/rec/T-REC-Q.552/>

[ITU-T Q.724] Recommendation ITU-T Q.724 (1988), *Specifications of signalling system No. 7 – Telephone user part: Signalling procedures*.
<http://www.itu.int/rec/T-REC-Q.724/>

[ITU-T T.24] Recommendation ITU-T T.24 (1998), *Standardized digitized image set*.
<http://www.itu.int/rec/T-REC-T.24/>

[ITU-T T.30] Recommendation ITU-T T.30 (2005), *Procedures for document facsimile transmission in the general switched telephone network*.
<http://www.itu.int/rec/T-REC-T.30/>

[ITU-T V.2] Recommendation ITU-T V.2 (1988), *Power levels for data transmission over telephone lines*.
<http://www.itu.int/rec/T-REC-V.2/>

[ITU-T V.8] Recommendation ITU-T V.8 (2000), *Procedures for starting sessions of data transmission over the public switched telephone network*.
<http://www.itu.int/rec/T-REC-V.8/>

[ITU-T V.18] Recommendation ITU-T V.18 (2000), *Operational and interworking requirements for DCEs operating in the text telephone mode*.
<http://www.itu.int/rec/T-REC-V.18/>

[ITU-T V.21] Recommendation ITU-T V.21 (1988), *300 bits per second duplex modem standardized for use in the general switched telephone network*.
<http://www.itu.int/rec/T-REC-V.21/>

[ITU-T V.22 *bis*] Recommendation ITU-T V.22 *bis* (1988), *2400 bits per second duplex modem using the frequency division technique standardized for use on the general switched telephone network and on point-to-point 2-wire leased telephone-type circuits.*
<http://www.itu.int/rec/T-REC-V.22bis/>

[ITU-T V.23] Recommendation ITU-T V.23 (1988), *600/1200-baud modem standardized for use in the general switched telephone network*.
<http://www.itu.int/rec/T-REC-V.23/>

[ITU-T V.25] Recommendation ITU-T V.25 (1996), *Automatic answering equipment and general procedures for automatic calling equipment on the general switched telephone network including procedures for disabling of echo control devices for both manually and automatically established calls*.
<http://www.itu.int/rec/T-REC-V.25/>

[ITU-T V.27 *ter*] Recommendation ITU-T V.27 *ter* (1988), *4800/2400 bits per second modem standardized for use in the general switched telephone network*.
<http://www.itu.int/rec/T-REC-V.27ter/>

[ITU-T V.29] Recommendation ITU-T V.29 (1988), *9600 bits per second modem standardized for use on point-to-point 4-wire leased telephone-type circuits*.
<http://www.itu.int/rec/T-REC-V.29/>

[IEC 61672-1] IEC 61672-1 (2013), *Electroacoustics – Sound level meters – Part 1: Specifications*.
<https://webstore.iec.ch/publication/5708>

[TIA-825-A] TIA-825-A (2003), *A Frequency Shift Keyed Modem for Use on the Public Switched Telephone Network.*
<https://global.ihs.com/search_res.cfm?RID=TIA&org_code=TIA&INPUT_DOC_NUMBER=TIA-825>

3 Terms and definitions

In the definition and text, L will refer to the relative power level of a signal, expressed in dBm0 (as defined by [ITU-T G.711]) and A will refer to the attenuation or loss of a signal path expressed in dB. These definitions assume that non-linearities are not present in the echo path and that the signal at S_{in} is purely echo. It is recognized that non-linearities may be present in a network.

This Recommendation defines the following terms:

3.1 acoustic echo

Acoustic echoes consist of reflected signals caused by acoustic environments. In these acoustic environments, an echo path is introduced by the acoustic path from the loudspeaker or earpiece to the microphone, e.g., echo created from hands-free speakerphones.

3.2 cancelled end

The side of an echo canceller which contains the echo path on which this echo canceller is intended to operate. This includes all transmission facilities and equipment (including the hybrid and terminating telephone set) which is included in the echo path. In previous versions of ITU-T G.168, this was defined as the near end.

3.3 combined loss (A_{COM})

The sum of echo return loss, echo return loss enhancement and non-linear processing loss (if present). This loss relates L_{Rin} to L_{RET} by:

$$L_{RET} = L_{Rin} - A_{COM}, \text{ where:}$$

$$A_{COM} = A_{ECHO} + A_{CANC} + A_{NLP}$$

3.4 comfort noise

Insertion of pseudo-random noise during the silent interval when the NLP operates or allowance of some of the background or idle channel noise to pass through the NLP in order to prevent the annoyance of intervals of speech with background noise followed by intervals of silence.

3.5 composite echo

Composite echoes consist of the electric echoes and acoustic echoes caused by reflected signals at hybrids and acoustic environments, e.g., analogue hands-free telephones.

3.6 convergence

The process of developing a model of the echo path which will be used in the echo estimator to produce the estimate of the circuit echo.

3.7 convergence time

For a defined echo path, the interval between the instant a defined test signal is applied to the receive-in port of an echo canceller with the estimated echo path impulse response initially set to zero, and the instant the returned echo level at the send-out port reaches a defined level.

3.8 double-talk

The simultaneous application of signals at R_{in} and S_{gen}.

3.9 echo canceller

A voice-operated device placed in the 4-wire portion of a circuit and used for reducing the cancelled end echo present on the send path by subtracting an estimation of that echo from the cancelled end echo (see Figure 2).

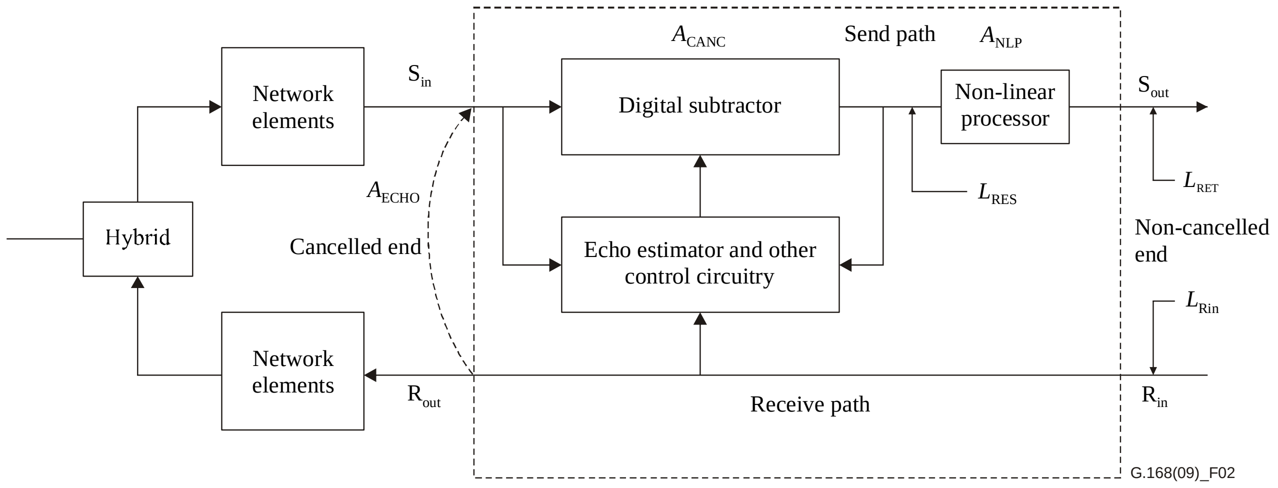


Figure 2 – Location of levels and loss of an echo canceller

3.10 echo cancellers in tandem

Multiple echo cancellers in a connection that are meant to cancel echo from the same echo source.

3.11 echo path

The transmission path between R_{out} and S_{in} of an echo canceller. This term is intended to describe the signal path of the echo.

3.12 echo path capacity (Δ)

The maximum echo path delay for which an echo canceller is designed to operate.

3.13 echo path delay (t_d)

The delay from the R_{out} port to the S_{in} port due to the delays inherent in the echo path transmission facilities *including* dispersion time due to the network elements. In case of multiple echo paths, all delays and dispersions of any individual echo path are included. The dispersion time, which varies with different networks, is required to accommodate the band-limiting, and hybrid transit effects (see Figure 3).

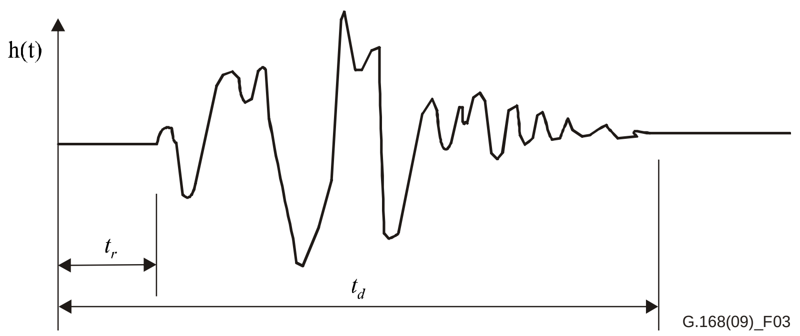


Figure 3 – Example of an impulse response of an echo path

3.14 echo return loss (ERL) (A_{ECHO})

The attenuation of a signal from the receive-out port (R_{out}) to the send-in port (S_{in}) of an echo canceller, due to transmission and hybrid loss, i.e., the loss in the (cancelled-end) echo path.

NOTE – ERL is not the same as echo loss defined in clause 2.2 of [ITU-T G.122] since echo loss is a weighted integral of the loss/frequency function over the band 300-3400 Hz, and so has a fixed value for any particular echo path characteristic. Also, echo loss applies to loss of the a-t-b path viewed from the virtual switching point of the international circuit. The echo canceller may be located closer to the echo reflection point.

3.15 echo return loss enhancement (ERLE) (A_{CANC})

The attenuation of the echo signal as it passes through the send path of an echo canceller. This definition specifically excludes any non-linear processing on the output of the canceller to provide for further attenuation.

3.16 electric echo

Electric echoes consist of reflected signals caused by the cancelled-end impedance mismatch, e.g., at a 2-wire/4-wire conversion unit (hybrid).

3.17 H register

The register within the echo canceller which stores the impulse response model of the echo path.

3.18 H register reset

Clear (zero) the H register and initialize any other appropriate internal variables of the echo canceller. This action is meant to simulate what happens at the start of a call for those cancellers externally controlled by the switch.

3.19 leak time

The interval between the instant a test signal is removed from the receive-in port of a fully-converged echo canceller and the instant the echo path model in the echo canceller changes such that, when a test signal is reapplied to R_{in} with the convergence circuitry inhibited, the returned echo is at a defined level.

This definition refers to echo cancellers employing, for example, leaky integrators in the convergence circuitry.

3.20 non-cancelled end

The side of an echo canceller which does not contain the echo path on which this echo canceller is intended to operate. In previous versions of ITU-T G.168, this was defined as the far end.

3.21 non-linear processor (NLP)

A device having a defined suppression threshold level and in which:

a) signals having a level detected as being below the threshold are suppressed; and

b) signals having a level detected as being above the threshold are passed although the signal may be distorted (for example see Annex B).

This Recommendation assumes an echo canceller is equipped with an NLP function that can be enabled or disabled when performing the tests defined in this Recommendation. An NLP function can be enabled or disabled by the user (for the purpose of performing a particular test), or may also be disabled upon detection of an appropriate disabling tone (e.g., 2100 Hz) as described in clause 7.1.

NOTE 1 – The precise operation of an NLP depends upon the detection and control algorithm used.

NOTE 2 – An example of an NLP is an analogue centre clipper in which all signal levels below a defined threshold are forced to some minimum value.

3.22 non-linear processing loss (A_{NLP})

Additional attenuation of residual echo level by an NLP placed in the send path of an echo canceller.

NOTE – Strictly, the attenuation of a non-linear process cannot be characterized by a loss in dB. However, for purposes of illustration and discussion of echo canceller operation, the careful use of A_{NLP} is helpful.

3.23 open echo path

An echo path with infinite echo return loss.

NOTE – In tests where an open echo path is used, it is important to break the path at a specific location. This location is stated in each individual test.

3.24 pure delay (t_r)

The delay from the R_{out} port to the S_{in} port due to the delays inherent in the near-end echo path transmission facilities, not including dispersion time due to the network elements. In this case, the transit time directly across the hybrid is assumed to be zero (see Figure 3).

3.25 residual echo level (L_{RES})

The level of the echo signal which remains at the send-out port of an operating echo canceller after imperfect cancellation of the circuit echo. It is related to the receive-in signal L_{Rin} by:

$$L_{RES} = L_{Rin} - A_{ECHO} - A_{CANC}$$

Any non-linear processing is not included.

3.26 returned echo level (L_{RET})

The level of the echo signal at the send-out port of an operating echo canceller which will be returned to the talker. The attenuation of an NLP is included, if one is normally present. L_{RET} is related to L_{Rin} by:

$$L_{RET} = L_{Rin} - (A_{ECHO} + A_{CANC} + A_{NLP})$$

If non-linear processing is not present, note that $L_{RES} = L_{RET}$.

3.27 signal processing device (SPD)

A hardware device that contains functions such as electric or acoustic echo control, noise reduction or automatic level control.

3.28 signal processing function (SPF)

A software function that performs voice processing such as electric or acoustic echo control, noise reduction or automatic level control.

3.29 signal processing network equipment/element (SPNE)

Type of network equipment or element which contains one or more signal processing functions or devices.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

ADPCM	Adaptive Differential Pulse Code Modulation
ATME	Automatic Test and Measurement Equipment
CED	Called station Identification
CNG	CalliNG Tone
CPE	Customer Premises Equipment
CSI	Called Subscriber Identification
CSS	Composite Source Signal
DCME	Digital Circuit Multiplication Equipment
DCS	Digital Command Signal
DEC	Digital Echo Canceller
DIS	Digital Identification Signal
DT	Double Talk
DTDT	Double Talk Detection Threshold
DTMF	Dual Tone Multi-Frequency
EC	Echo Canceller
ERL	Echo Return Loss
ERLE	Echo Return Loss Enhancement
FAX	Facsimile
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
HDLC	High-level Data Link Control
ISDN	Integrated Services Digital Network
MLS	Maximal Length Sequence
MOS	Mean Opinion Score
NLMS	Normalized Least-Mean Square
NLP	Non-Linear Processor
NSF	Non-Standard Facilities
NSS	Non-standard Set-up
PCM	Pulse Code Modulation
PCME	Packet Circuit Multiplication Equipment
PN	Pseudo Noise
PSTN	Public Switched Telephone Network
RMS	Root Mean Square

SPD	Signal Processing Device
SPF	Signal Processing Function
SPNE	Signal Processing Network Equipment/Element
TCLw	Weighted Terminal Coupling Loss
TSI	Transmitting Subscriber Identification
TTL	Transistor-Transistor Logic

5 Test signals

The tests in this Recommendation use special signals such as noise, tones, Group 3 facsimile signals, and a subset of the composite source signals (CSS) consisting of the bandlimited CSS with speech like power density spectrum (pseudo noise signal generated using 8192 pt. FFT) and the bandlimited CSS for double talk (see Annex C and [ITU-T P.501]). The CSS emulates the characteristics of speech, and its use as a test signal improves the ability to measure echo canceller performance for speech signals.

Furthermore, network echo cancellers should perform adequately on many non-speech signals, e.g., voiceband data, as well as under non-ideal network scenarios. Tests are included to test performance for Group 3 facsimile signals, residual acoustic echoes (optional), and non-linearities in the echo path such as may arise with low bit rate encoding in the echo path (optional).

6 Characteristics of echo cancellers

6.1 General

This Recommendation is applicable to the design of echo cancellers (ECs). The echo cancellers are assumed to be "half" echo cancellers, i.e., those in which cancellation takes place only in the send path due to signals present in the receive path. A full echo canceller is possible consisting of two half echo cancellers, with each pointing in the opposite direction.

6.2 Purpose, operation and environment

Echo cancellers have the following fundamental requirements:

1) rapid convergence;
2) low returned echo level during single talk;
3) low divergence during double talk and cancelled end speech;
4) assured double talk detection and cancelled end speech detection;
5) proper operation during facsimile and low speed (< 9.6 kbit/s) voiceband data transmissions.

Echo cancellers may remain active for several non-voice signals as well, in particular, Group 3 facsimile and low speed (< 9.6 kbit/s) voiceband data transmissions. Tests 10 and 14 address these issues.

If appropriate signalling to disable echo cancellers in tandem is not implemented, then it is possible that echo cancellers can appear in tandem in a connection. Suitable tests for ensuring adequate performance for tandem echo cancellers are currently under study in Test 11. When echo cancellers are located on the subscriber side of the international signalling equipment, signalling tones do not pass through the cancellers so no special action is necessary. When cancellers are on the international side of the signalling equipment, they are normally disabled by the switch during the active signalling exchange intervals in order to prevent distortion of the signalling tones by the echo canceller. When signalling tones simultaneously appear at the canceller receive and send ports (double talk), the receive signal will be processed through the echo path model contained in the canceller. The signal estimate produced by the canceller may sufficiently distort the send side signal so that it will not be properly recognized by the signalling receive unit (see Note 1).

An echo canceller should be disabled during the transmission of the ITU-T No. 6 and No. 7 continuity check signal (Note 2). If an echo canceller conforming to this Recommendation is located on the international side of a circuit with ITU-T No. 6 or No. 7 signalling and is not externally disabled by the switch, it will not corrupt the return of the continuity check tone only if it is able to pass the optional Test No. 8. Similarly, if an echo canceller conforming to this Recommendation is located on the international side of ITU-T No. 5 signalling units and is not disabled by the switch, it will not corrupt the continuously compelled line signalling exchange only if it is able to pass the optional Test No. 8.

NOTE 1 – For some echo cancellers, this problem may not occur when the send and receive frequencies are different.

NOTE 2 – [ITU-T Q.271] on ITU-T No. 6 and [ITU-T Q.724] on ITU-T No. 7 both include the following statement: "As the presence of active echo suppressors in the circuit would interfere with the continuity check, it is necessary to disable the suppressors during the check and to re-enable them, if required, after the check has been completed." This consideration also applies to echo cancellers.

To improve the subjective performance when conference bridge circuits are in use, it is recommended that each 2-wire network connection to a conference bridge be equipped with a dedicated echo canceller that is oriented away from the bridge, towards the hybrid of the terminating end in order to prevent echo entering the summing node of the conference bridge. In such an arrangement, the talker echo and listener echo (listener hearing multiple copies of distorted talker voice) will be avoided and an EC will not see multiple hybrid reflections, resulting in more natural communication for all users of the bridge.

The following tests address convergence properties of network echo cancellers under various conditions of operation:

- Test No. 2 – Convergence and Steady state residual and returned echo level tests.
  - Test 2A: Convergence and re-convergence test with NLP enabled.
  - Test 2B: Convergence and re-convergence test with NLP disabled.
  - Test 2C: Convergence test in the presence of background noise.
- Test No. 5 – Infinite return loss convergence test.
- Test No. 10 – Facsimile test during call establishment phase.
  - Test 10A: Canceller operation on the calling station side.
  - Test 10B: Canceller operation on the called station side.
- Test No. 15 – PCM offset test.

The term "convergence time" is defined in clause 3.7; however it is recognized that a single number for "convergence time" is not sufficient to provide meaningful information regarding overall convergence properties of an echo canceller. If a value of convergence time is specified, the following information should be defined accordingly:

- The hybrid model, ERL, and echo path delay (t_d) used for the echo path.
- The type of test signal used (e.g., CSS, White Gaussian Noise), and the level at the receive-in port.
- Whether the NLP is enabled or disabled.

• The defined level of returned echo at the send-out port, for which a sustained amount of cancellation is achieved after the convergence time.

NOTE 3 – Although convergence time may be tested with echo path models that include an echo path delay (t_d), the value of t_d may be subtracted from the value of convergence time.

6.3 External enabling/disabling

Certain digital echo cancellers may be disabled directly by a digital signal (e.g., see Recommendation [b-ITU-T Q.55]). These echo cancellers should provide 64 kbit/s bit sequence integrity (i.e., if integrated, the A-law to µ-law or µ-law to A-law conversion will also be disabled) in the externally disabled state.

6.4 Tests and requirements for performance with input signals applied to the send and receive paths

6.4.1 Transmission performance

The performance characteristics apply, unless otherwise noted, when steady state signals are separately applied to the send and receive paths.

A digital network echo canceller inserted between codecs into a digital transmission path meeting the performance characteristics of [ITU-T G.712] shall not alter performance such that the requirements of [ITU-T G.712] are violated.

The appropriate transmission performance requirements as noted below also apply.

6.4.1.1 Delay

The maximum delay experienced at any frequency in the voiceband in the send path should be kept to a minimum and should not exceed 1 ms. No significant delay should occur in the receive path.

NOTE 1 – The creation of frame slips in the echo path can lead to an occasional degradation of the echo cancellation. If a delay is necessary to synchronize the digital send and receive paths, the global admissible delay on the send path, including the group delay mentioned above, should not exceed 1 ms and on the receive path 250 µs.

NOTE 2 – A method to measure one directional delay through an echo canceller is for further study.

6.4.1.2 Measuring input and output levels

For testing purposes, the method defined for measuring the input level of the composite source signals is a RMS method. Unless otherwise specified within a test, the RMS method should also be used for measuring the output levels at S_{out}. Other methods that would give equivalent results are possible (see Annex C). For the RMS method, specifically, CSS is measured using:

$$S(k) = 3.14 + 20\log\left[\frac{\sqrt{\frac{2}{n}\sum_{i=k}^{k-n+1} e_i^2}}{4096}\right] \quad \text{(A-law encoding)}$$

$$S(k) = 3.17 + 20\log\left[\frac{\sqrt{\frac{2}{n}\sum_{i=k}^{k-n+1} e_i^2}}{8159}\right] \quad (\mu\text{-law encoding})$$

where:

$S(k)$ = signal level in dBm0

e_i = linear equivalent of the PCM encoded signal at time i

k = discrete time index

n = number of samples over which the RMS measurement is made, and $n = \alpha\tau$ with $\alpha \geq 1$ (an integer) and τ = period of CSS (5600 for the single-talk portion and 6400 for the double talk portion of CSS)

Some tests in this Recommendation, e.g., Test No. 2, use the RMS value that is measured over the active portion of the CSS (i.e., excluding the pause in the CSS) only. The subscript "act" in a signal level is used to denote the level measurement in this case. For example, if L_{Rin} is the RMS level of R_{in} including the pause, then $L_{\text{Rin,act}}$ is the RMS level of R_{in} excluding the pause. The RMS level of the CSS excluding the pause is larger than that including the pause. The difference is 1.49 dB for single talk CSS and 1.66 dB for double talk CSS. In other words, given a CSS signal level L_{x} in the unit of dBm0, $L_{\text{x,act}}$ can be obtained from:

$$L_{\text{x,act}} = L_{\text{x}} + 1.49 \text{ dB, for single talk CSS}$$

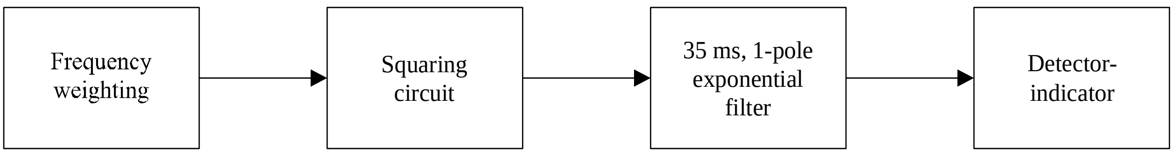
$$L_{\text{x,act}} = L_{\text{x}} + 1.66 \text{ dB, for double talk CSS}$$

Refer to Annex C for the details regarding the single talk CSS and double talk CSS.

6.4.1.2.1 Level measurement device

For some of the tests in this Recommendation, e.g., Test No. 2, it is necessary to measure the short-term level of the signal. This is achieved using the following level measurement device.

The measurement device comprises a frequency-weighting network, a squaring circuit, an exponential filter, and a detector-indicator. This device may use either analogue or digital methods. The impulse response of the frequency-weighting network is listed in Table 1. The table is read in columns. The measurement device will have the characteristics that correspond to the following block diagram.



G.168(09)_F6.4.1.2.1

This frequency weighting network is used to provide a greater attenuation of the frequencies outside the band of interest that can be achieved by the filters identified in [IEC 61672-1]. This filter is required because of D.C. effects due to the bias induced by A-law encoding. The filter is a 101-element finite impulse response bandpass filter with the impulse response shown in Figure 4 and the frequency response shown in Figure 5. The coefficients of the filter are:

Table 1 – Coefficients of bandpass filter for level measurement device

f_0, f_{100}	0.0000	f_{17}, f_{83}	–0.0019	f_{34}, f_{66}	0.0092
f_1, f_{99}	0.0006	f_{18}, f_{82}	–0.0033	f_{35}, f_{65}	0.0000
f_2, f_{98}	0.0005	f_{19}, f_{81}	–0.0047	f_{36}, f_{64}	0.0164
f_3, f_{97}	0.0004	f_{20}, f_{80}	–0.0000	f_{37}, f_{63}	–0.0210
f_4, f_{96}	0.0011	f_{21}, f_{79}	–0.0068	f_{38}, f_{62}	0.0161
f_5, f_{95}	–0.0000	f_{22}, f_{78}	0.0036	f_{39}, f_{61}	–0.0375
f_6, f_{94}	0.0015	f_{23}, f_{77}	–0.0057	f_{40}, f_{60}	0.0000
f_7, f_{93}	–0.0003	f_{24}, f_{76}	0.0054	f_{41}, f_{59}	–0.0406
f_8, f_{92}	0.0012	f_{25}, f_{75}	0.0000	f_{42}, f_{58}	–0.0357
f_9, f_{91}	–0.0002	f_{26}, f_{74}	0.0044	f_{43}, f_{57}	–0.0267
f_{10}, f_{90}	0.0000	f_{27}, f_{73}	0.0095	f_{44}, f_{56}	–0.0871
f_{11}, f_{89}	0.0002	f_{28}, f_{72}	0.0017	f_{45}, f_{55}	–0.0000
f_{12}, f_{88}	–0.0020	f_{29}, f_{71}	0.0188	f_{46}, f_{54}	–0.1420
f_{13}, f_{87}	0.0005	f_{30}, f_{70}	0.0000	f_{47}, f_{53}	0.0289
f_{14}, f_{86}	–0.0040	f_{31}, f_{69}	0.0225	f_{48}, f_{52}	–0.1843
f_{15}, f_{85}	0.0000	f_{32}, f_{68}	0.0024	f_{49}, f_{51}	0.0475
f_{16}, f_{84}	–0.0047	f_{33}, f_{67}	0.0163	f_{50}	0.8006

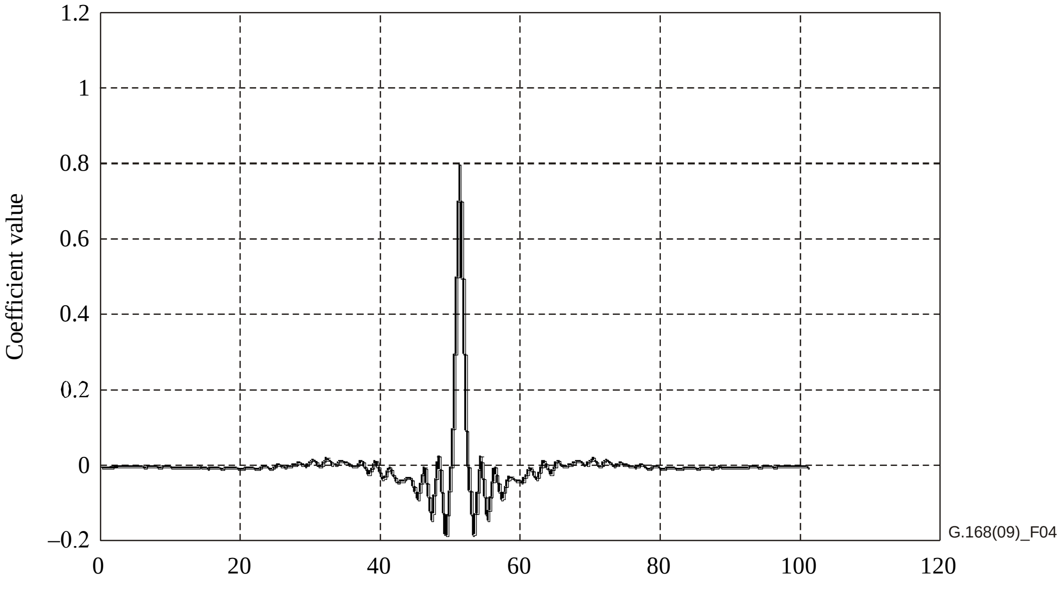


Figure 4 – Impulse response of frequency weighting network

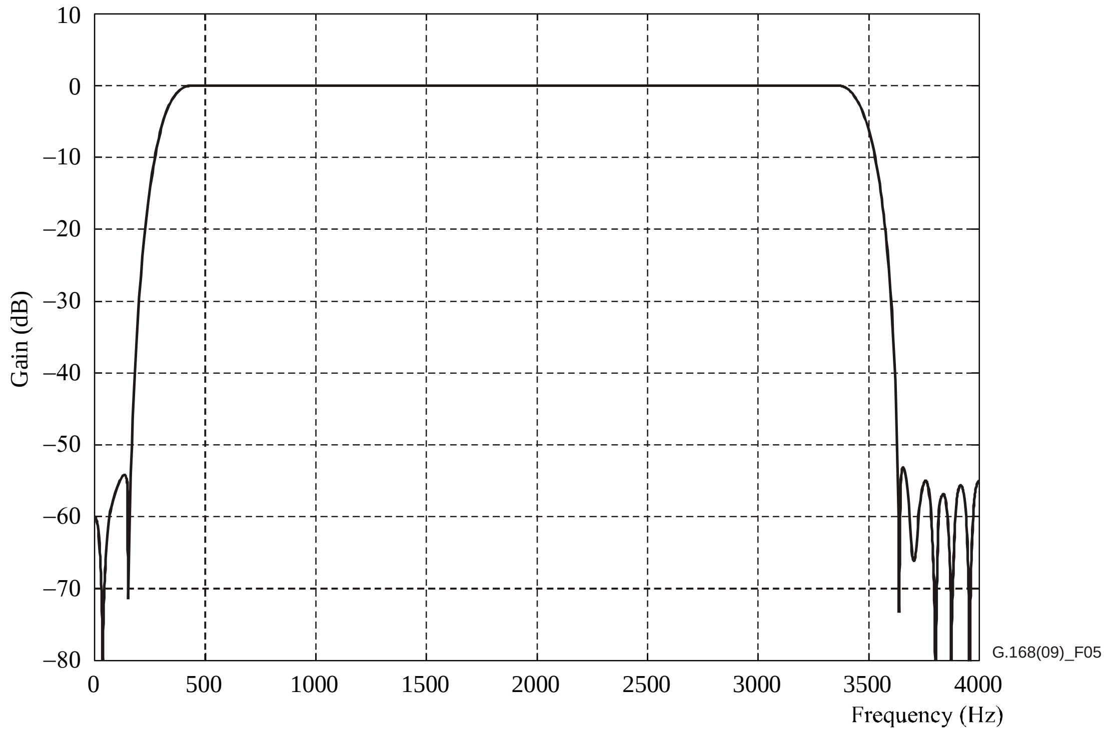


Figure 5 – Magnitude response of frequency weighting network

The magnitude response given in Figure 5 was generated using the following equation:

$$\left|H(\omega)\right| = 10\log_{10}\left(\left|\sum_{n=0}^{100} h_n e^{-j\omega n}\right|^2\right) \ [\text{dB}], \quad 0 < \omega < \pi$$

and the frequency in the x-axis is formed by:

$$f = \frac{\omega}{\pi} 4000 \quad \text{Hz}$$

6.4.1.2.2 Level measurement device for peaks

For tests that have requirements on the peaks at S_{out}, the measurement method used is a 35 ms rectangular sliding window in place of the 35 ms one-pole exponential filter of clause 6.4.1.2.1.

6.4.2 Echo canceller performance

The performance requirements that follow are for echo cancellers that include NLPs.

For testing purposes, it is required that the NLP can be disabled, that the echo path impulse response store (H register) can be reset (set to zero) and that adaptation can be inhibited.

The tests in this Recommendation require a bearer signal interface and a control signal interface to the echo canceller under test as depicted in Figure 6.

The bearer signal interface is for the four-port access to signals, R_{in}, R_{out}, S_{in}, and S_{out}. The digital signals are coded in [ITU-T G.711] μ-law/A-law at a bit rate of 64 kbit/s and are usually carried over T1/E1 links. If the echo canceller under test does not support T1/E1 links for bearer signals, some interconnection equipment should be provided to make the echo canceller under test accessible by the testing unit through an appropriate interface.

The control signal interface is for echo canceller control functions such as reset H register and enabling/disabling adaptation. The network echo canceller should be able to be controlled by a control link in such a way that the H register of the echo canceller can be reset or the adaptation inhibited/enabled at the right time synchronized with the test signals according to the specifications defined in the test requirements. The interface to the control link shall be well documented and provide for ease of implementation. Two examples of a control interface are given in Appendix VI.

The requirements are described in terms of tests made by applying signals to R_{in} and S_{in} of an echo canceller, and measuring the S_{out} signals. The test set-up is as shown in Figures 6 and 7. The ports are assumed to be at equal relative level points. For all values of R_{in}, and for all tests in this Recommendation, the level at R_{out} should be equal to the level at R_{in}. Any optional processing included in the echo canceller which may affect level transparency between R_{in} and R_{out} should be disabled during all tests in this Recommendation. The composite source signals, which consist of the receive-input test signal and send-input test signal (see Annex C and [ITU-T P.501]) are used as the test signals, unless otherwise indicated.

For multiple channel implementations, including implementations where other signal processing functions may share common processing resources, sufficient resources shall be provided for each channel of echo cancellation to meet the requirements of this Recommendation. Tests to ensure that the implementation fulfils this requirement are for further study. When performing the tests described in this Recommendation, for enhanced repeatability, all channels not being tested should have idle code (e.g., 01111111 for 1544 kbit/s systems or 01010100 for 2048 kbit/s systems as outlined in Recommendation ITU-T Q.522) applied to the inputs. It is generally recognized that some operators may wish to apply simulated traffic loading to the untested channels. This type of channel loading is for further study.

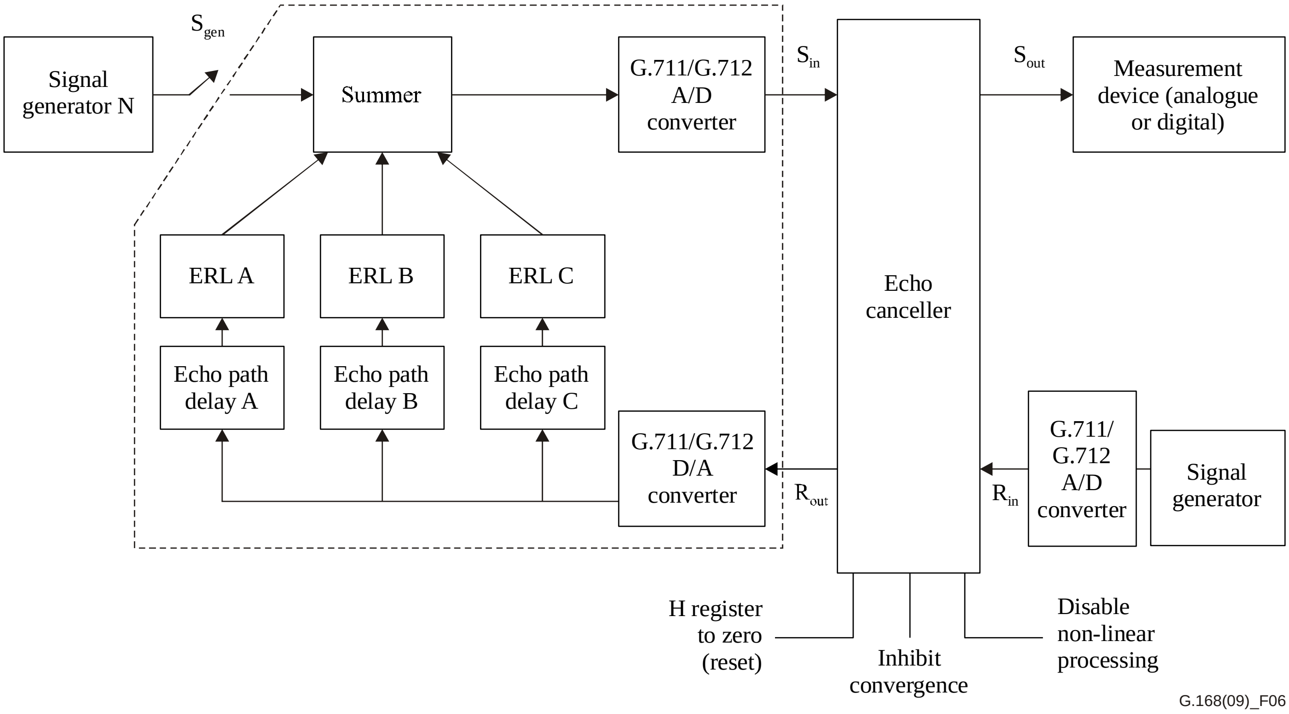
The ERL used in these tests have a minimum value of 6 dB. It should be noted that 6 dB is a typical worst-case value encountered for most networks, and most current networks have typical ERL values better than this.

Also, it should be noted that the test configurations specified in this Recommendation are artificial for purposes of test and result repeatability, and do not fully represent all conditions that would be expected in real networks.

The requirements in clause 6.4.2 are based on the use of the composite source signals, noise, tones, FAX signals, and voiceband data signals as the test signals.

Two echo path models should be used for the tests in this Recommendation (as denoted in Figures 6 and 7). With the exception of Test No. 13, the tests in this Recommendation do not address performance requirements for echo cancellers when the echo path manifests non-linearities beyond the ones incurred in the course of ITU-T G.711/ITU-T G.712 companding as shown in Figures 6 and 7.

Comfort noise should be disabled unless otherwise indicated.



NOTE – The sum of the absolute values of the gains G_A, G_B, G_C that correspond to ERL A, B, C, respectively, taken in dB, should be less than or equal to –6 dB (i.e., $20\log(|G_A| + |G_B| + |G_C| \le -6$ dB), and echo path delay A ≤ Δ ms, echo path delay B ≤ Δ ms, and echo path delay C ≤ Δ ms.

Figure 6 – Functional diagram for echo canceller performance measurements

An echo path model which allows more realistic end paths to be modelled can be realized by replacing the dotted box in Figure 6 with Figure 7. The characteristics of the end path, which include A to μ law converters, can be modelled as an impulse response g(k).

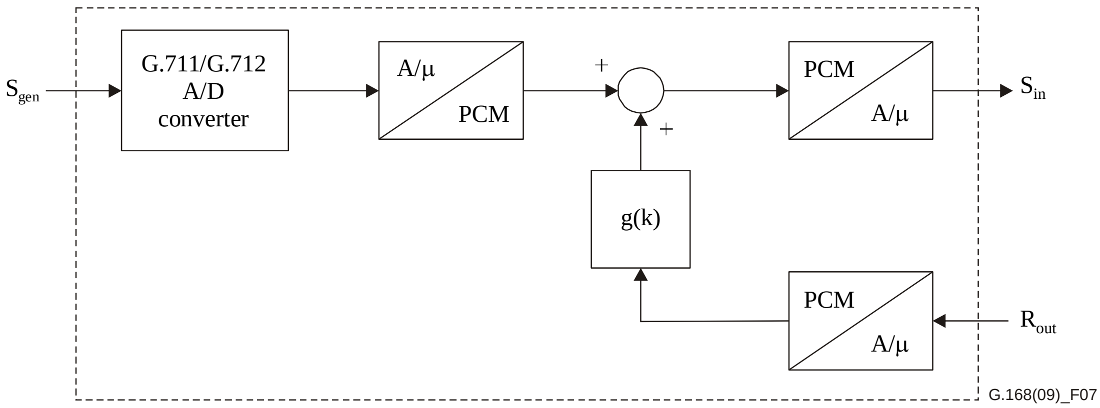


Figure 7 – Echo path model using g(k)

The primary purpose of an echo canceller is to control the echo of a speech signal. This is done by synthesizing a replica of the echo path impulse response and using it to generate an estimate of the echo which is subtracted from the actual circuit echo. The synthesis should be accomplished using a speech input signal. Because of the difficulty of defining a speech test signal, the following tests are type tests and rely upon the use of a composite source signal primarily for convenience and repeatability. These tests should be performed on an echo canceller only after the design has been shown to properly synthesize a replica of the echo path impulse response from a speech input signal and its corresponding echo. Speech signals are not used in the tests in this clause. Additionally, the NLP in the echo canceller should be designed to minimize and potentially avoid undesirable effects such as double talk clipping, gaps in transmitted speech signals, and noise contrast (see Test 9 described later in this Recommendation for noise contrast, and see Appendix I for further discussion on double talk clipping).

Different echo cancellers may be designed to work satisfactorily for different echo path delays depending on their application in various networks. Thus Δ, whenever it appears in this Recommendation, represents the maximum echo path delay for which the echo canceller is designed.

In some cases, a delay is present between the test equipment and echo canceller at the non-cancelled end side. This delay occurs for example if the echo canceller under test can interface only a packet network (such as an IP network) at the non-cancelled end side and is the consequence of packetization, etc. This Recommendation primarily addresses echo cancellers that have 64 kbit/s digital interfaces on both sides of the canceller (see Figure 1). However, if the Recommendation is used to test echo cancellers deployed in gateway devices that interface with packet networks, then any such additional delay should be taken into account by appropriately shifting the test requirements in time.

See Appendix IV for some guidelines on the use of parameters for testing echo cancellers.

6.4.2.1 Echo path models for g(k)

See Annex D for the echo path models that may be used as the g(k) in Figure 7 for the tests in this Recommendation. This does not represent an exhaustive set, and other models may be used provided that they meet the echo path requirements for each individual test. Note that the digital version of Figure 6 where three echo path reflection points are present may also be represented by a g(k). A specific model including this is not described in Annex D. A more complex model that includes realistic dispersion and other effects is for further study. See also Appendix III for additional information.

To ensure that the magnitude response of the echo path g(k) does not exceed 0 dB over the frequency range of some of the models contained in Annex D, the minimum ERL values must be greater than 6 dB. See Annex D for the exact minimum ERL value for each model. Note that the minimum ERL values provided in Annex D overrule the requirement of ERL ≥ 6 dB given in the following tests if there is a conflict between the two.

6.4.2.2 Test No. 1 – Steady state residual and returned echo level test (deleted)

This test has been incorporated into Test 2.

6.4.2.3 Test No. 2 – Convergence and steady state residual and returned echo level tests

This test is meant to ensure that the echo canceller converges rapidly for all combinations of input signal levels, echo paths, and certain echo path changes (but not including continuously varying echo paths)[1], and that the returned echo level is sufficiently low. This test is also meant to ensure that the steady state cancellation (A_{CANC}) is sufficient to produce a residual echo level which is sufficiently low to permit the use of non-linear processing without undue reliance on it. In general, given that all other variables are equal, a higher value of ERLE or lower values of L_{RES} will allow for less dependence on the NLP functionality.

[1] Convergence requirements for continuously changing echo path impulse responses (non-abrupt change tracking of the echo path impulse response) are not covered by this Recommendation.

At the beginning of a call, the convergence should be fast enough to be subjectively unnoticeable. In general, the convergence should be fast enough to handle changes in the echo path in a subjectively transparent fashion. Faster convergence than required in Figures 10 and 12 is desirable, but only if no degradation is observed during single or double talk and the stability of the canceller can be maintained in all network conditions (e.g., various echo path conditions, including various hybrids) and for all voiceband signals.

In Tests 2A-2C, the level at R_{in} is $L_{Rin,act}$. It is the signal level measure using the RMS method over the active portion of the CSS only (i.e., excluding the pause of the CSS) as described in clause 6.4.1.2.

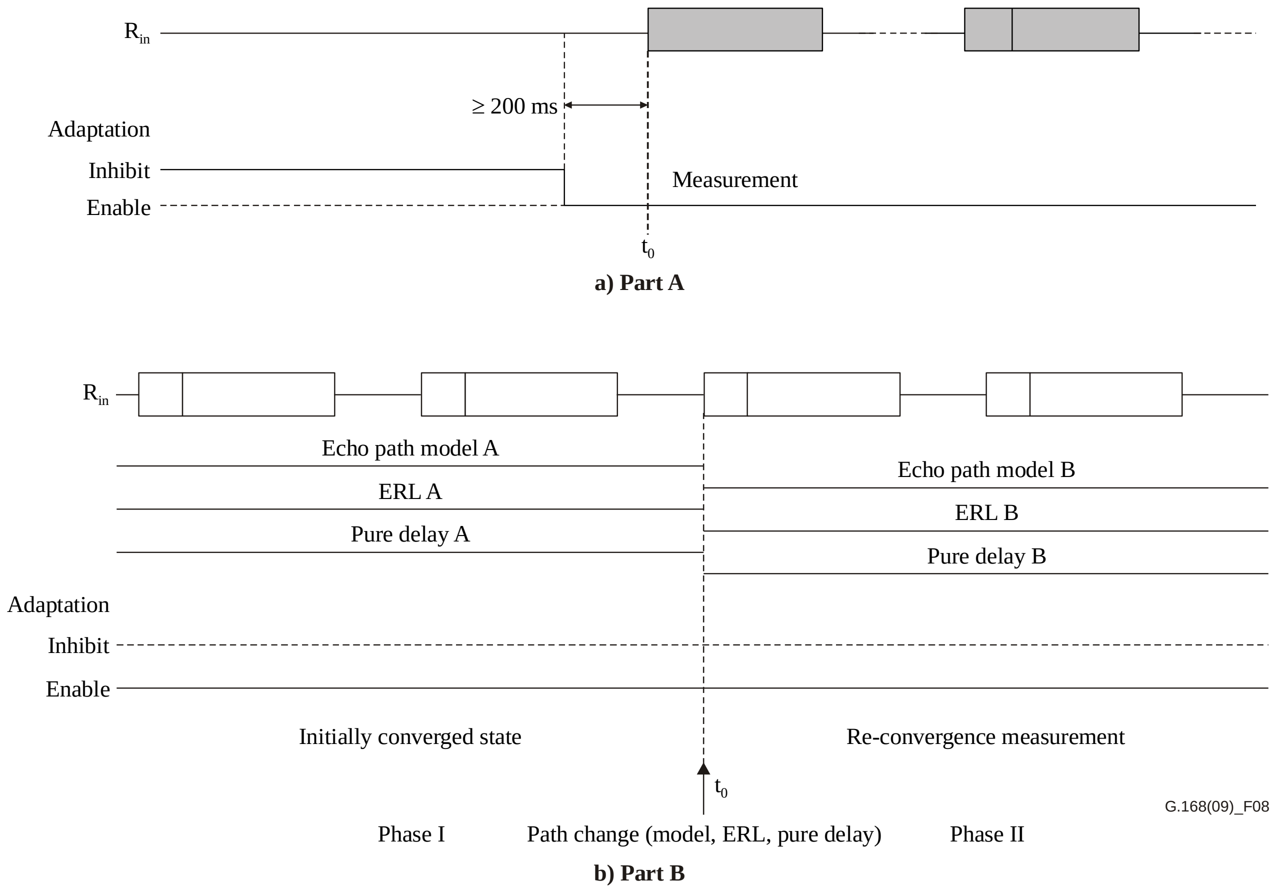


Figure 8 – Signal and time relationships for Test No. 2A/2B

6.4.2.3.1 Test 2A: Convergence and re-convergence test with NLP enabled

a) Convergence test with NLP enabled

Requirement

With the H register initially reset, or alternatively, with an open echo path resulting in $S_{in} = 0$ and the H register content converged to 0, and the NLP enabled, for all values $L_{Rin,act} \geq -30$ dBm0 and ≤ 0 dBm0 and for all values of ERL ≥ 6 dB and echo path delay, $t_d \leq \Delta$ ms, the combined loss ($L_{Rin,act} - L_{RET}$) should be greater than or equal to that shown in Figure 10a. After $1 + t_d + t_0$ s, the combined loss should be greater than or equal to that in Figure 9. The level at S_{out} is measured using a meter conforming to the characteristics of clause 6.4.1.2.1. In addition, no peaks (see clause 6.4.1.2.2) are allowed that exceed 5 dB above the requirements in Figure 9.

Adaptation is enabled at least 200 ms before the start of a CSS burst (see Figure 8a). This period is to allow for the latency time in the adaptation control of the canceller. The residual or returned echo level is then measured as a function of time to reveal the convergence and steady-state properties of the echo canceller.

The variable $L_{Rin,act} – L_{RET}$ in Figure 10a may be replaced by the variable $L_{Sin} – L_{Sout}$ + ERL, where L_{Sin} and L_{Sout} are the levels of S_{in} and S_{out} respectively. The signal levels L_{Sin} and L_{Sout} are measured using the measurement device in clause 6.4.1.2.1, and should be synchronized. The ERL is the value chosen in the test. This method may also be used to observe convergence as a continuous plot over time.

NOTE – The method stated in the preceding paragraph takes into account any dispersion in the echo path, but does not take into account any dispersion present between the S_{in} and S_{out} ports of the echo canceller.

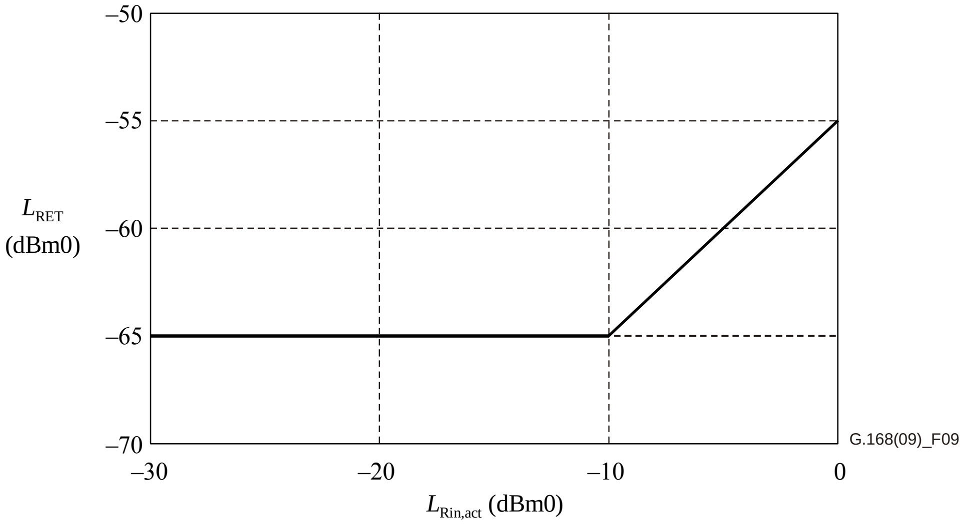


Figure 9 – Relationship between receive input level ($L_{Rin,act}$) and return echo level (L_{RET}) with NLP enabled

The requirements in Figure 9 may not be met with echo cancellers containing a comfort noise feature, if enabled, and so, for the purposes of this test, comfort noise is disabled. For R_{in} signal levels exceeding –5 dBm0, CSS will be clipped. This does not, however, imply that the requirements of Figures 9 and 10 need not be met at R_{in} levels of –5 dBm0 or above. For this range, special care should be taken to ensure that the echo path is linear. Non-linearities in the real network may result in performance less than indicated in the figure.

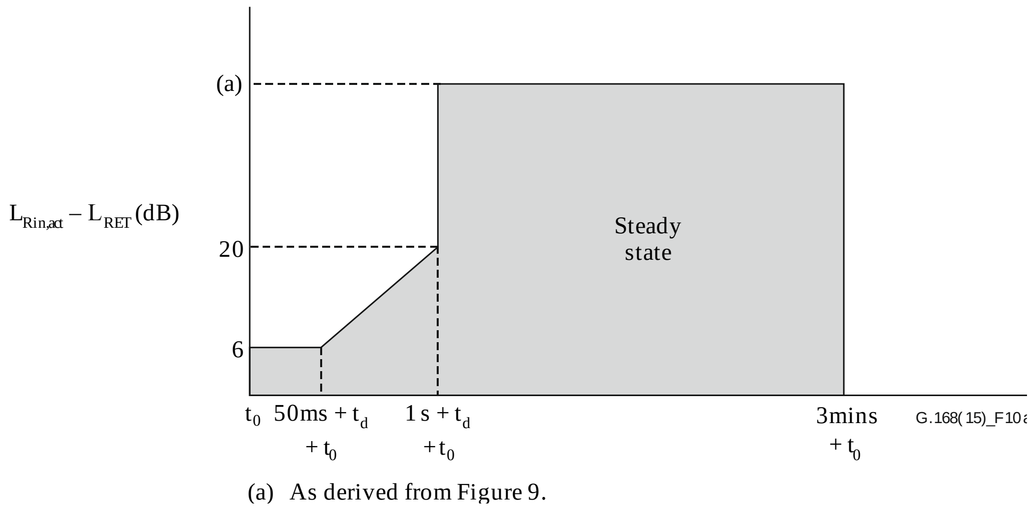


Figure 10a – Convergence characteristics with NLP enabled

b) Re-convergence test with NLP enabled

Requirement

With the H register initially converged to any echo path, at time t_0 the echo path is changed to any other different echo path (which includes echo path model change and/or ERL change (up or down) and/or pure delay change (up or down)). The adaptation enable/inhibit control should remain enabled throughout the test. For all values $L_{Rin,act} \geq -30$ dBm0 and ≤ 0 dBm0 and for all values of ERL ≥ 6 dB and echo path delay, $t_d \leq \Delta$ ms (for both echo path models), the combined loss ($L_{Rin,act} - L_{RET}$) should be greater than or equal to that shown in Figure 10b. That is, the return echo level (L_{RET}) after $1 + t_d + t_0$ s, should be lower than or equal to that shown in Figure 9. The level at S_{out} is measured using a meter conforming to the characteristics of clause 6.4.1.2.1. In addition, during steady state, no peaks (see clause 6.4.1.2.2) are allowed that exceed 5 dB above the requirements in Figure 9. The NLP should be enabled for this test.

With the following terminology regarding characteristics of Phases I and II (see Figure 8b):

– Phase I: ERL_A, echo path model A (in short, $hybrid_A$), pure delay A (in short, $delay_A$)
– Phase II: $ERL_B = ERL_A - \Delta ERL$ [dB], $hybrid_B$, $delay_B$,

it is recommended that, at minimum, the following combinations be tested:

– $hybrid_B \neq hybrid_A$ and $ERL_B = ERL_A$, with $delay_B = delay_A$
– $hybrid_B = hybrid_A$ and $ERL_B < ERL_A$ ($\Delta ERL = 10$ dB or greater), with $delay_B = delay_A$
– $hybrid_B \neq hybrid_A$ and $ERL_B < ERL_A$ ($\Delta ERL = 10$ dB or greater), with $delay_B = delay_A$

For Phase II, it is recommended that, at minimum, the models which have the longest dispersion time, i.e., models m_4, m_6, m_7 and m_8 and the largest pure delay, model m_6, be used.

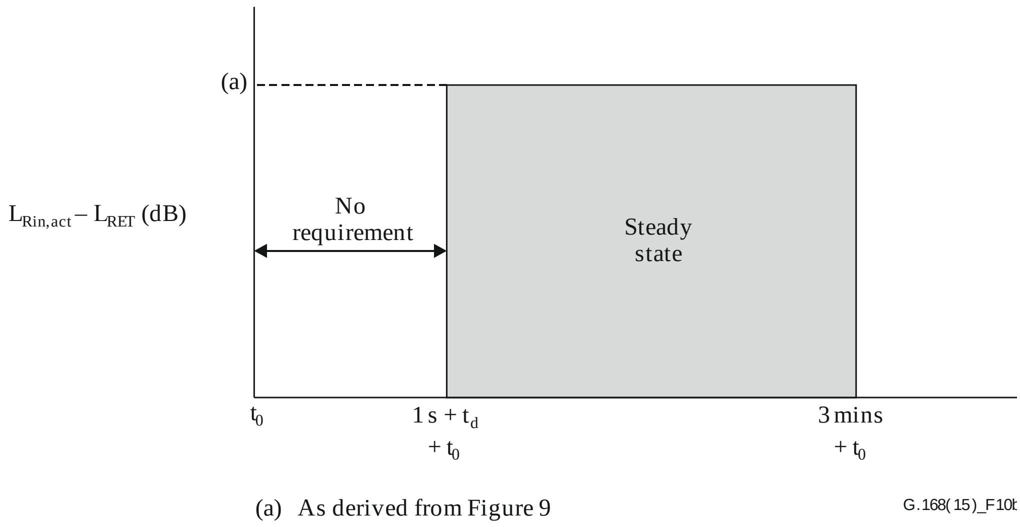


Figure 10b – Re-convergence characteristics with NLP enabled

6.4.2.3.2 Test 2B: Convergence and re-convergence test with NLP disabled

a) Convergence test with NLP disabled

Requirement

With the H register initially reset, or alternatively, with an open echo path resulting in $S_{in} = 0$, and the H register content converged to 0, and the NLP disabled, for all values $L_{Rin,act} \geq -30$ dBm0 and ≤ 0 dBm0 and for all values of ERL ≥ 6 dB and echo path delay, $t_d \leq \Delta$ ms, the loss $L_{Rin,act} - L_{RES}$ should be greater than or equal to that shown in Figure 12a. After $10 + t_d + t_0$ s, the loss ($L_{Rin,act} - L_{RES}$) should be greater than or equal to that in Figure 11. The level at S_{out} is measured using a meter conforming to the characteristics of clause 6.4.1.2.1. In addition, no peaks (see clause 6.4.1.2.2) are allowed that exceed 5 dB above the requirements in Figure 11.

Adaptation is enabled at least 200 ms before the start of a CSS burst (see Figure 8a). This period is to allow for the latency time in the adaptation control of the canceller. The residual or returned echo level is then measured as a function of time to reveal the convergence and steady-state properties of the echo canceller.

The variable $L_{Rin,act} - L_{RES}$ in Figure 12a may be replaced by the variable $L_{Sin} - L_{Sout}$ + ERL, where L_{Sin} and L_{Sout} are the levels of S_{in} and S_{out} respectively. The signal levels L_{Sin} and L_{Sout} are measured using the measurement device in clause 6.4.1.2.1, and should be synchronized. The ERL is the value chosen in the test. This method may also be used to observe convergence as a continuous plot over time.

NOTE 1 – The method stated in the preceding paragraph takes into account any dispersion in the echo path, but does not take into account any dispersion present between the S_{in} and S_{out} ports of the echo canceller.

NOTE 2 – Some echo cancellers employ a supplementary NLP function which cannot be disabled. For information covering this case, see clause 8.2.6, Testing of NLPs.

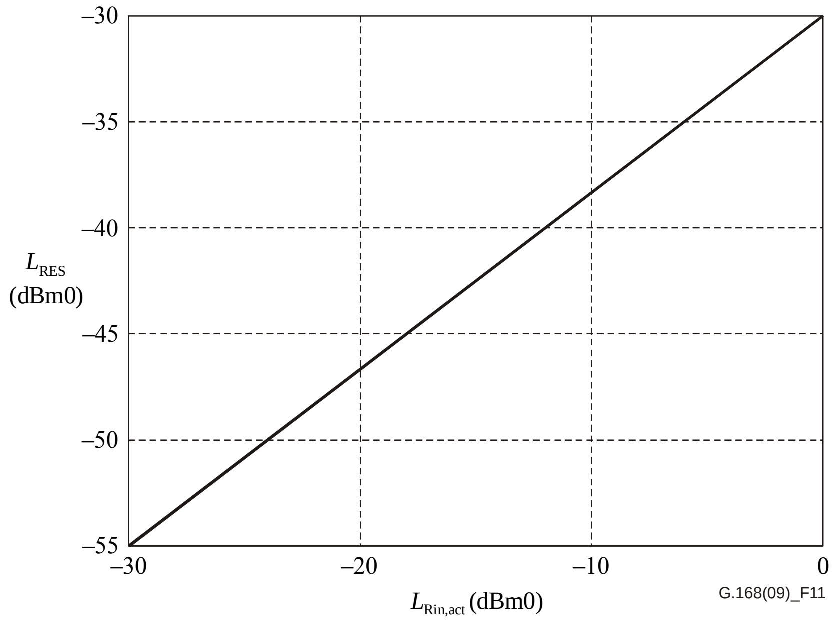


Figure 11 – Relationship between receive input level ($L_{Rin,act}$) and residual echo level (L_{RES}) with NLP disabled

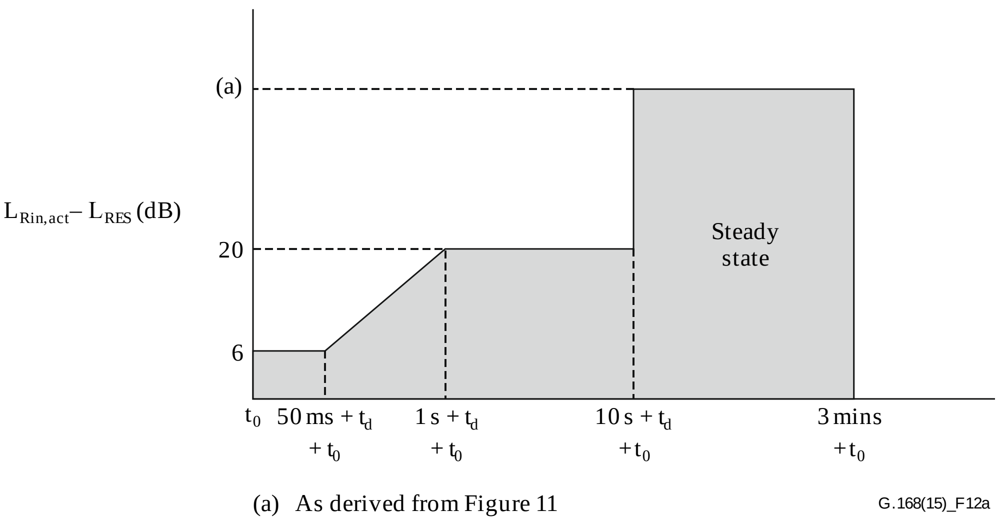


Figure 12a – Convergence characteristics with NLP disabled

b) Re-convergence test with NLP disabled

Requirement

With the H register initially converged to any echo path, at time t_0 the echo path is changed to any other different echo path (which includes echo path model change and/or ERL change (up or down) and/or pure delay change (up or down)). The adaptation enable/inhibit control should remain enabled throughout the test. For all values $L_{Rin,act} \geq -30$ dBm0 and ≤ 0 dBm0 and for all values of ERL ≥ 6 dB and echo path delay, $t_d \leq \Delta$ ms (for both echo path models), the loss $L_{Rin,act} - L_{RES}$ should be greater than or equal to that shown in Figure 12b. That is, the loss $L_{Rin,act} - L_{RES}$ after 1 + t_d + t_0 s, should be greater than or equal to 20 dB. After 10 + t_d + t_0 s the loss $L_{Rin,act} - L_{RES}$ should be greater than or equal to that derived from Figure 11. The level at S_{out} is measured using a meter conforming to the characteristics of clause 6.4.1.2.1. In addition, during steady state, no peaks (see clause 6.4.1.2.2) are allowed that exceed 5 dB above the requirements in Figure 11. The NLP should be disabled for this test.

With the following terminology regarding characteristics of Phases I and II (see Figure 8b):

– Phase I: ERL_A, echo path model A (in short, $hybrid_A$), pure delay A (in short, $delay_A$)
– Phase II: $ERL_B = ERL_A – \Delta ERL$ [dB], $hybrid_B$, $delay_B$,

it is recommended that, at minimum, the following combinations be tested:

– $hybrid_B \neq hybrid_A$ and $ERL_B = ERL_A$, with $delay_B = delay_A$
– $hybrid_B = hybrid_A$ and $ERL_B < ERL_A$ (ΔERL = 10 dB or greater), with $delay_B = delay_A$
– $hybrid_B \neq hybrid_A$ and $ERL_B < ERL_A$ (ΔERL = 10 dB or greater), with $delay_B = delay_A$

For Phase II, it is recommended that, at minimum, the models which have the longest dispersion time, i.e., models m_4, m_6, m_7 and m_8 and the largest pure delay, model m_6, be used.

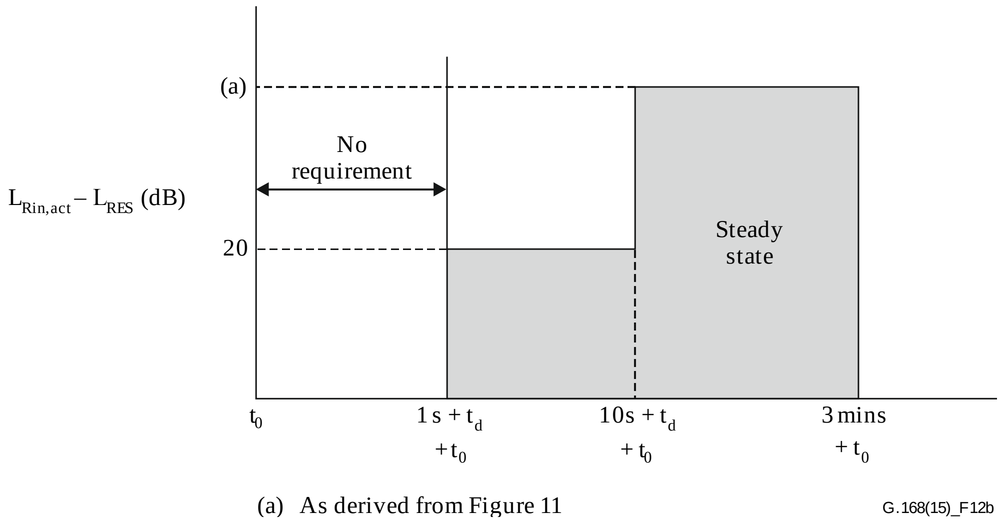


Figure 12b – Re-convergence characteristics with NLP disabled

NOTE – Test 5B (Re-convergence during call transfer test) and Test 12 (Residual acoustic echo test) have similarities to Test 2B-b. However, Test 5B is specifically designed to exercise real-world scenarios where significant ERL changes are presented to the network during call transfers.

6.4.2.3.3 Test 2C: Convergence test in the presence of background noise

Test No. 2C is meant to ensure that the steady state cancellation is sufficient to produce an echo level that is sufficiently low and that the echo canceller converges rapidly for all combinations of input signal levels and echo paths in the presence of background noise.

The test procedure is to reset the H register and inhibit adaptation. A Hoth noise source (see [ITU-T P.800]) with level L_{Sgen} is applied as the signal S_{gen}. Adaptation is enabled at least 200 ms before the start of a CSS burst (see Figure 13). After the adaptation interval, inhibit adaptation, remove S_{gen} and measure the residual echo level.

The level at S_{out} is measured using a meter conforming to the characteristics of clause 6.4.1.2.1.

The variable $L_{Rin,act} – L_{RES}$ in Figure 16 may be replaced by the variable $L_{Sin} – L_{Sout}$ + ERL, where L_{Sin} and L_{Sout} are the levels of S_{in} and S_{out} respectively. The signal levels L_{Sin} and L_{Sout} are measured using the measurement device in clause 6.4.1.2.1, and should be synchronized. The ERL is the value chosen in the test. This method may also be used to observe convergence as a continuous plot over time.

NOTE 1 – The method stated in the preceding paragraph takes into account any dispersion in the echo path, but does not take into account any dispersion present between the S_{in} and S_{out} ports of the echo canceller.

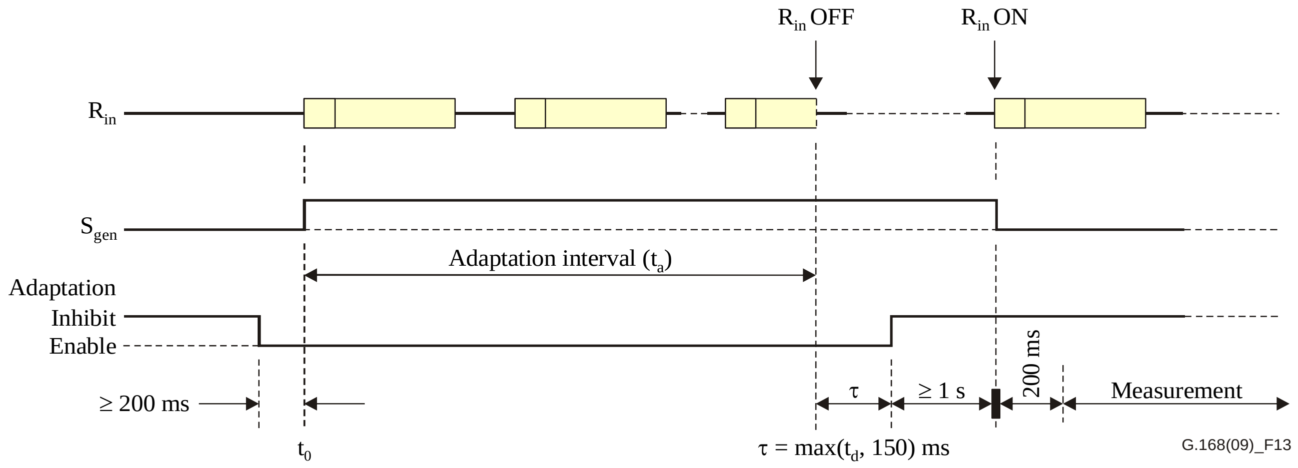


Figure 13 – Test No. 2C signal and time relationships

NOTE 2 – At the end of this test, a time period of ≥ 1 s is allowed for possible latency in the adaptation inhibit command. R_{in} is switched off to prevent further adaptation before the inhibit command is effective. A time period of $\tau = \max(t_d, 150)$ ms is allowed before the adaptation inhibit command to clear the impulse response of an echo path. R_{in} is switched on again to perform the measurement. The 200 ms waiting time before measurement is not needed, if the contents of the bandpass filter and exponential filter in the measurement device are set to zero (flushed) before any measurement is performed.

a) Convergence test with NLP enabled

Requirement

With the H register initially reset and the NLP enabled, for all values of $L_{Rin,act} \geq -30$ dBm0 and ≤ 0 dBm0, $L_{Sgen} = L_{Rin,act} - 15$ dB but no higher than –30 dBm0, ERL ≥ 6 dB and echo path delay, $t_d \leq \Delta$ ms, convergence should occur within 1.0 s (t_a) and L_{RET} should be $\leq L_{Sgen}$ (see Figure 14).

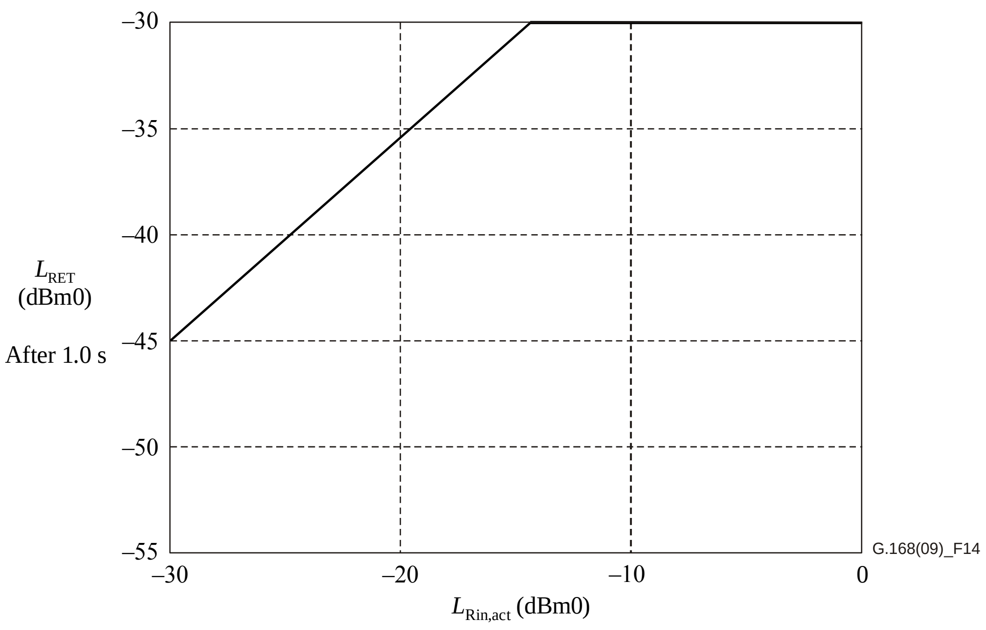


Figure 14 – Test No. 2C requirements NLP enabled

b) Steady state cancellation test with NLP disabled

Requirement

With the H register initially reset and the NLP disabled, for all values of LRin,act ≥ –30 dBm0 and ≤ 0 dBm0, L_{Sgen} as given in Figure 15, ERL ≥ 6 dB, echo path delay $t_d \le \Delta$ ms, and adaptation interval $t_a \ge 10$ s, L_{RES} should be less than that shown in Figure 15 for the corresponding value of L_{Sgen}.

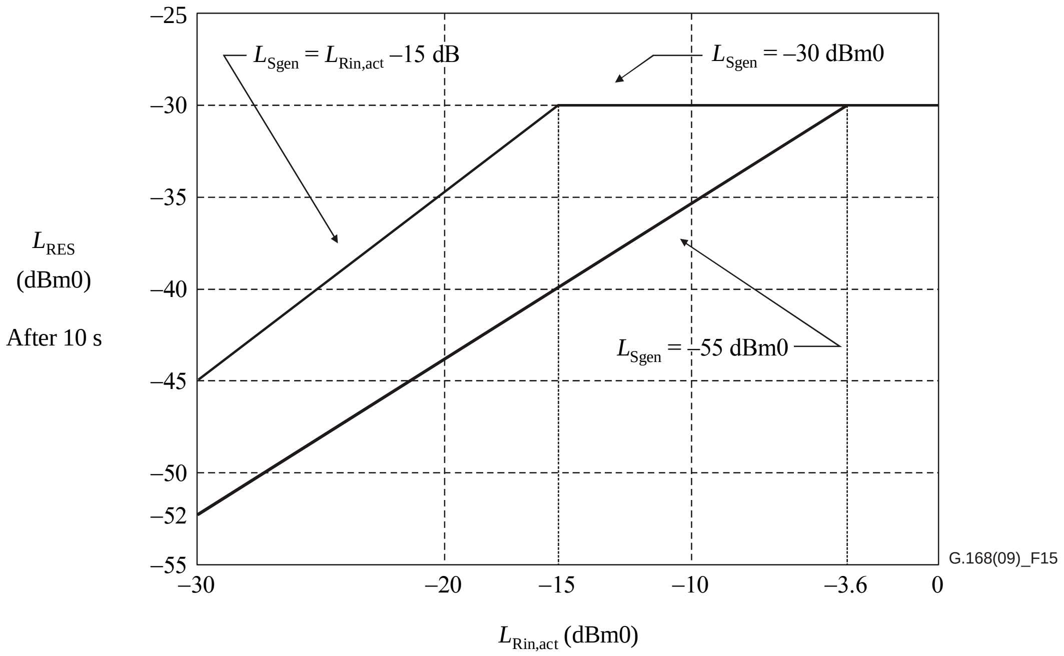


Figure 15 – Test No. 2C steady state requirements NLP disabled

c) Convergence test with NLP disabled

This test differs from parts (a) and (b) above, in that adaptation is not inhibited during the test, the noise applied as the S_{gen} signal is not discontinued during the test and, once started, CSS is applied continuously at R_{in} throughout the test.

Requirement

With the H register initially reset and the NLP disabled, for all values of $L_{\text{Rin,act}} \ge -30$ dBm0 and ≤ 0 dBm0, $L_{\text{Sgen}} = L_{\text{Rin,act}} - 15$ dB but no higher than –30 dBm0, ERL ≥ 6 dB and echo path delay $t_d \le \Delta$ ms, the loss $L_{\text{Rin,act}} - L_{\text{RES}}$ within an adaptation interval of $t_a = 10$ s should be greater than or equal to that shown in Figure 16. The value X is given below:

$$X = \left(L_{Rin,act} - L_{Sgen}\right) - 6\,\text{dB}, \quad \text{for } X \le 17\,\text{dB}$$

$$X = 17\,\text{dB, otherwise}$$

NOTE 1 – The method used to measure the noise in this test, results in fluctuating levels and can result in difficulty when performing this test.

NOTE 2 – In this case, L_{RES} contains the noise component, L_{Sgen}.

NOTE 3 – The requirement of $L_{\text{Rin,act}} - L_{\text{RES}}$ in Figure 16 is 5 dB from t_0 to 50 ms + t_d + t_0 although the minimum ERL is 6 dB. This is to accommodate the presence of background noise at S_{in} because the background noise increases the fluctuations and consequently the value of L_{RES}.

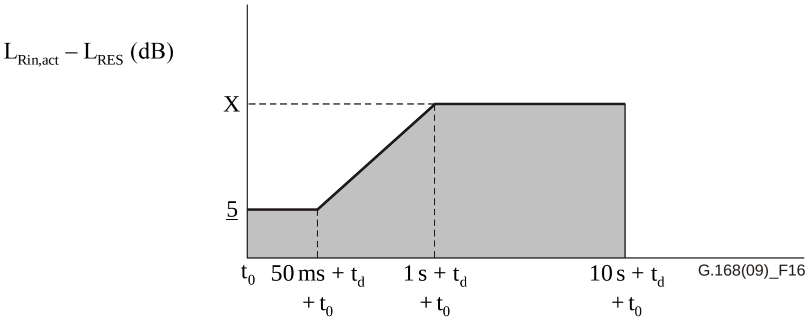


Figure 16 – Test No. 2C Convergence requirements NLP disabled

6.4.2.4 Test No. 3 – Performance under conditions of double talk

The three parts of this test are meant to test the performance of the canceller under various conditions of double talk. During conditions of double talk, the echo canceller can give rise to unwanted artefacts such as clipping, distortion, and noise contrast (see Appendix I). The tests make the assumption that, upon detection of double talk, measures are taken to prevent or slow adaptation in order to avoid excessive reduction in cancellation.

For this test, the Rin signal is CSS and the Sgen signal is the double talk CSS. While CSS is used for this test, it is recognized that it is only a statistical approximation of real speech. Double talk tests performed with actual speech samples may produce results somewhat different than those shown in this test. This test is intended to provide a guideline on how the double talk performance of an echo canceller should be measured. It is possible that this test and its requirements may change as the correlation between CSS and real speech is better understood. Use of different languages has been shown to provide considerable variation in the results for tests 3A and 3B (see [b-ITU-T GSTP.CSS] in clause I.8 for more details).

See clause I.7.4 for guidelines on other double talk test methods for tests 3A and 3B.

6.4.2.4.1 Test 3A: Double talk convergence test with low cancelled-end levels

Test No. 3A is meant to ensure that the double talk detection is not so sensitive that echo and low level cancelled end speech falsely cause operation of the double talk detector to the extent that adaptation does not occur. The test procedure is to reset the H register; then for some value of echo path delay and ERL, a signal is applied to R_{in}. Simultaneously (see Figure 17) an interfering signal (double talk CSS), which is sufficiently low in level to not seriously hamper the ability of the echo canceller to converge, is applied as the signal S_{gen}. This signal should allow adaptation and cancellation to occur. After the allowed adaptation interval, the adaptation is inhibited and the residual echo measured. The NLP should be *disabled*.

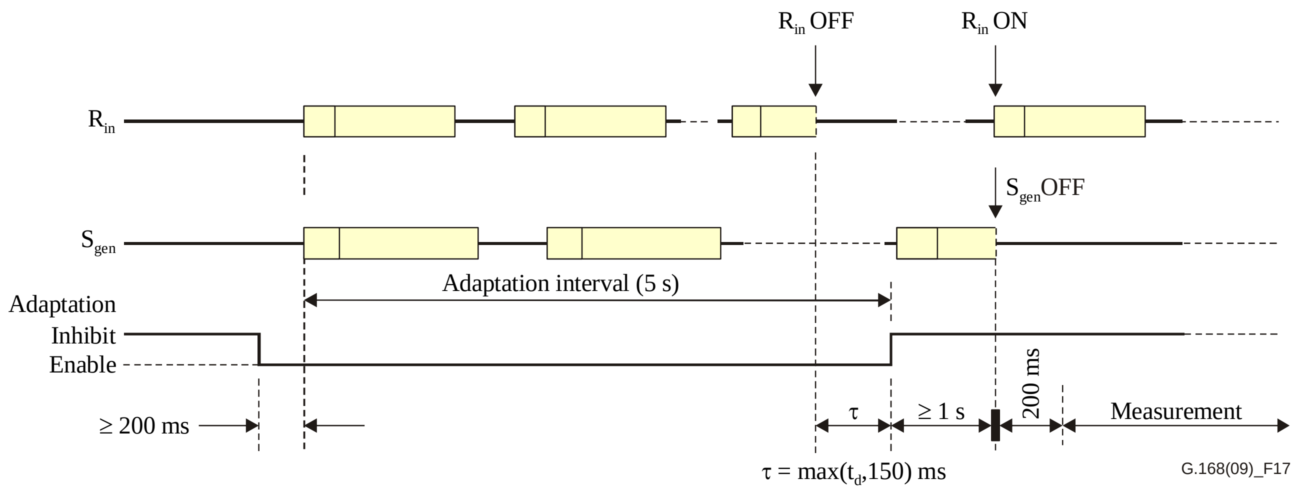


Figure 17 – Test No. 3A signal and time relationships

NOTE – At the end of this test, a time period of ≥ 1 s is allowed for possible latency in the adaptation inhibit command. R_{in} is switched off to prevent further adaptation before the inhibit command is effective. A time period of $\tau = \max(t_d, 150)$ ms is allowed before the adaptation inhibit command to clear the impulse response of an echo path. R_{in} is switched on again to perform the measurement. The 200 ms waiting time before measurement is not needed, if the contents of the bandpass filter and exponential filter in the measurement device are set to zero (flushed) before any measurement is performed.

Requirement

With the H register initially reset for all values of $L_{\mathrm{Rin,act}} \geq -25$ dBm0 and ≤ 0 dBm0, $L_{\mathrm{Sgen,act}} = L_{\mathrm{Rin,act}} - 15$ dB, ERL ≥ 6 dB and echo path delay, $t_d \leq \Delta$ ms, convergence should occur within 5 s and L_{RES} should be $\leq L_{\mathrm{Sgen,act}}$. The level measurement device of clause 6.4.1.2.1 should be used to measure the signal at S_{out}.

6.4.2.4.2 Test 3B: Double talk stability test

Test No. 3B is meant to ensure that the double talk detector is sufficiently sensitive and operates fast enough to prevent large divergence during double talking.

The test procedure is to fully converge the echo canceller for a given echo path by applying CSS to R_{in}. After the canceller is fully converged (see Figure 18), a signal S_{gen} is applied that has a level L_{Sgen} at least that of R_{in}. This should cause the double talk detector to operate. After any arbitrary time, $\delta_t > 0$, the adaptation is inhibited, the S_{gen} signal is removed, and the residual echo measured. The NLP should be *disabled*.

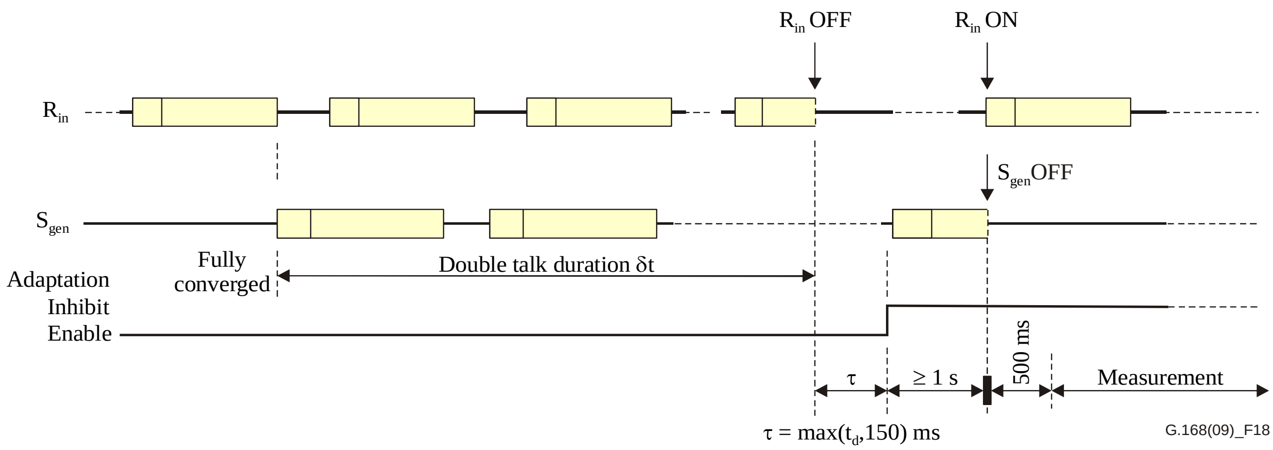


Figure 18 – Test No. 3B signal and time relationships

NOTE – At the end of this test, a time period of ≥ 1 s is allowed for possible latency in the adaptation inhibit command. R_{in} is switched off to prevent further adaptation before the inhibit command is effective. A time period of τ = max(t_d,150) ms is allowed before the adaptation inhibit command to clear the impulse response of an echo path. R_{in} is switched on again to perform the measurement. The 500 ms waiting time before measurement is not needed, if the contents of the bandpass filter and exponential filter in the measurement device are set to zero (flushed) before any measurement is performed.

[b-ITU-T GSTP.CSS] in clause I.8 shows a wide variation of performance for different languages for this test. Note that test 3B is even more sensitive to real speech variations and CSS may not provide adequate approximation of real speech for this test.

a) Double talk stability test with high cancelled-end levels

Requirement

With the echo canceller initially in the fully converged state for all values of $L_{\text{Rin,act}} \geq -30$ dBm0 and ≤ 0 dBm0, and for all values of $L_{\text{Sgen,act}} \geq L_{\text{Rin,act}}$ and for all values of ERL ≥ 6 dB and echo path delay $t_d \leq \Delta$ ms, the residual echo level after the simultaneous application of R_{in} and S_{gen} for any time period should not increase more than 10 dB over the steady state requirements of Figure 11. The level measurement device of clause 6.4.1.2.1 should be used to measure the signal at S_{out}.

b) Double talk stability test with low cancelled-end levels

Requirement

The test definition is identical to part (a) except that $L_{\text{Sgen,act}} = L_{\text{Rin,act}}$ – X dB. For all values 6 dB ≤ X ≤ 30 dB, the steady state requirement of Figure 11 + 3 dB would apply.

6.4.2.4.3 Test 3C: Double talk test under simulated conversation

Test No. 3C is meant to ensure that the echo canceller does not produce undesirable artefacts during and after periods of double talk (see clause I.7).

The test procedure is to reset the H register. Then for some value of echo path delay, $t_d \leq \Delta$ ms, and ERL ≥ 6 dB, a signal is applied to R_{in}. Simultaneously (see Figure 19), a signal S_{gen} is applied that has a level L_{Sgen} at least that of R_{in}. After a time t_1, S_{gen} is removed and S_{out} is measured. During time t_4 and t_5, S_{gen} is reapplied and the output is evaluated for artefacts. R_{in} is removed during period t_5. The NLP should be enabled for this test.

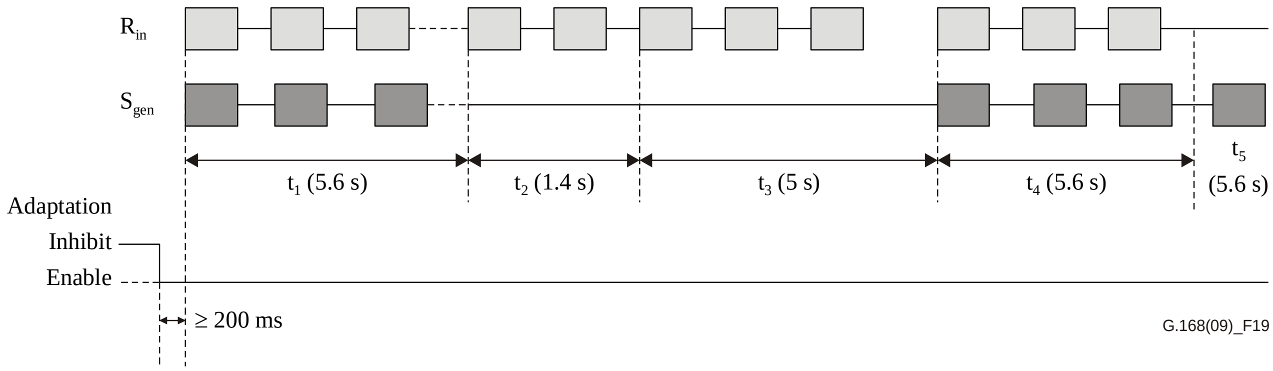


Figure 19 – Test No. 3C signal and time relationships

Requirement

With the H register initially reset, for all values of $L_{\text{Rin,act}} \geq -25$ dBm0 and ≤ 0 dBm0, and for all values of $L_{\text{Sgen,act}} \geq L_{\text{Rin,act}}$ and for all values of ERL ≥ 6 dB and echo path delay $t_d \leq \Delta$ ms, any peaks (see clause 6.4.1.2.2) during period t_2 should not exceed the level $L_{\text{Sgen,act}}$ during period t_1. The returned echo level (measured using the level measurement device of clause 6.4.1.2.1) during time period t_3 should meet the requirements of Figure 9 with NLP enabled. During t_4 and t_5, no peaks (see clause 6.4.1.2.2) should exceed the level of $L_{\text{Sgen,act}}$ + 6 dB.

Level offsets between L_{Rin} and L_{Sgen} can cause inappropriate operation of the NLP and can cause speech degradation and is for further study. Variation of CSS may be useful for this purpose.

6.4.2.5 Test No. 4 – Leak rate test

This test is meant to ensure that the leak time is not too fast, i.e., that the contents of the H register do not go to zero too rapidly.

The test procedure is to fully converge the echo canceller using CSS for a given echo path and then to remove all signals from the echo canceller. After two minutes the contents of the H register are frozen, CSS is reapplied to R_{in} and the residual echo measured (see Figure 20). The NLP should be *disabled*.

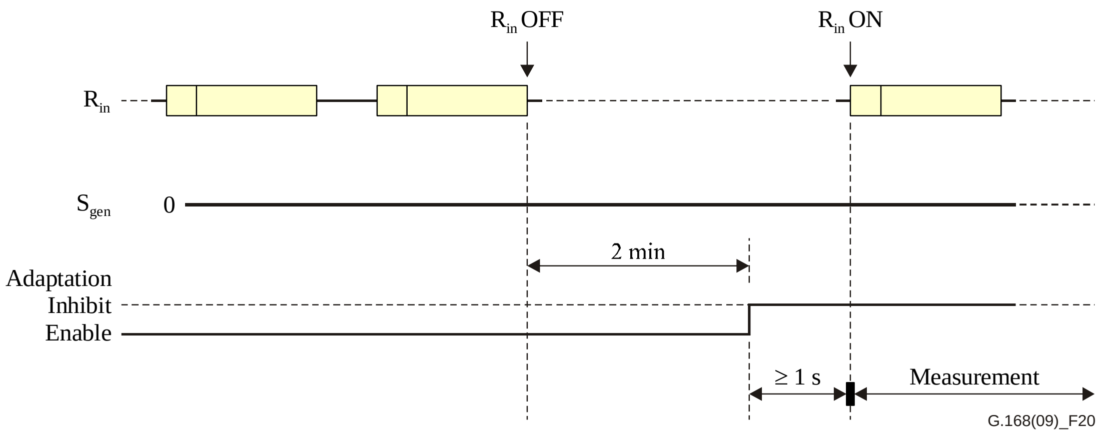


Figure 20 – Test No. 4 signal and time relationships

NOTE – At the end of this test, a time period of ≥ 1 s is necessary to allow for possible latency in the adaptation inhibit command. R_{in} is switched on to perform the measurement.

Requirement

With the echo canceller initially in the fully converged state for all values of $L_{\text{Rin,act}} \geq -30$ dBm0 and ≤ 0 dBm0, two minutes after the removal of the R_{in} signal, the residual echo level should not increase more than 10 dB over the steady state requirement of Test No. 2 (Figure 11). The level measurement device of clause 6.4.1.2.1 should be used to measure the signal at S_{out}.

6.4.2.6 Test No. 5 – Infinite return loss convergence test

These tests are meant to ensure that the echo canceller has some means to prevent the unwanted generation of echo during significant echo path model changes. This may occur when the H register contains an echo path model, either from a previous connection or the current connection, and the echo path is significantly changed (for example: circuit echo vanishes or reappears) while a signal is present at R_{in}.

The tests are divided into two parts and are designed to test real-world scenarios as follows:

– **Test No. 5A** is the same as Test No. 5 in versions of ITU-T G.168 up to 03/2009, and is designed to ensure that the echo canceller has some means to prevent the unwanted generation of echo when the echo path is opened (infinite ERL). This may occur due to switch action or a transmission fault in the echo path.

– **Test No. 5B** is designed to simulate a call transfer scenario where a call is transferred from destination A to destination B and then back to destination A, where A and B have echo paths that are significantly different (e.g., a call to a 2-wire telephone is transferred to a 4-wire ISDN telephone and vice-versa)

6.4.2.6.1 Test No. 5A: Infinite return loss convergence test

The test procedure is to fully converge the echo canceller using CSS for a given echo path. The echo path is then interrupted at R_{out} while a CSS is applied to R_{in}, and the output at S_{out} measured (see Figure 21a). The NLP should be disabled.

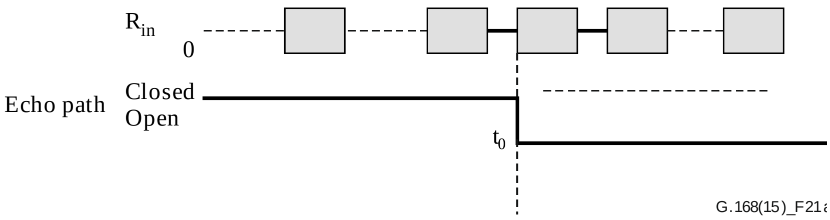


Figure 21a – Test No. 5A signal and time relationships

The level at R_{in} in the test requirement is $L_{\text{Rin,act}}$. It is the signal level measured using the RMS method over the active portion of the CSS only (i.e., excluding the pause of the CSS) as described in clause 6.4.1.2.

Requirement

With the echo canceller initially in the fully converged state for all values of ERL ≥ 6 dB, and for all values of $L_{\text{Rin,act}} \geq -30$ dBm0 and ≤ 0 dBm0, and at time t_0 the echo path is interrupted with an open echo path, the combined loss $L_{\text{Rin,act}} - L_{\text{RES}}$ should meet the requirements of Figure 12a, as measured using the method of clause 6.4.1.2.1.

The level at S_{out} is measured using a meter conforming to the characteristics of clause 6.4.1.2.1.

6.4.2.6.2 Test No. 5B: Re-convergence during call transfer test

The test procedure is to fully converge the echo canceller using CSS for a given echo path. The ERL of the echo path is then changed while a CSS is applied to R_{in}, and the output at S_{out} measured (see Figures 21b and 21c below). The NLP should be disabled for this test.

The following echo path model changes represent various call transfer scenarios:

1) Low-high-low ERL transitions as shown in Figure 21b, and

2) High-low-high ERL transitions as shown in Figure 21c.

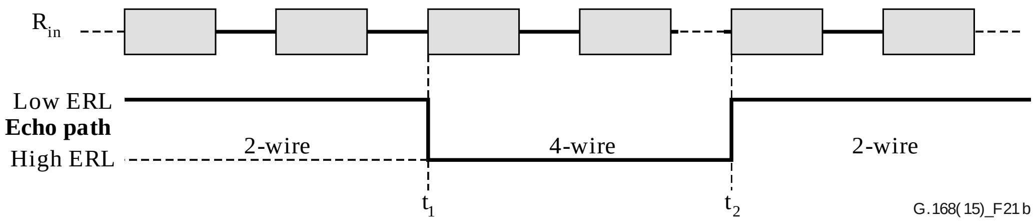


Figure 21b – Test No. 5B signal and timing relationships for low-high-low ERL transitions

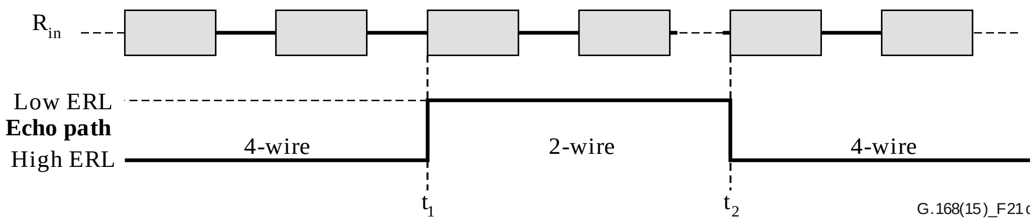


Figure 21c – Test No. 5B signal and timing relationships for high-low-high ERL transitions

The level at R_{in} in the test requirement is $L_{\mathrm{Rin,act}}$. It is the signal level measured using the RMS method over the active portion of the CSS only (i.e., excluding the pause of the CSS) as described in clause 6.4.1.2.

Requirement

With the echo canceller initially in the fully converged state for 30 dB ≥ ERL ≥ 6 dB (for low ERL echo path models), and ERL ≥ 46 dB (for high ERL echo path models), and for all values of $L_{\mathrm{Rin,act}} \geq -30$ dBm0 and ≤ 0 dBm0, the combined loss $L_{\mathrm{Rin,act}} - L_{\mathrm{RES}}$ should meet the requirements of Figure 12b at times t_1 and t_2 (where $t_2 \geq t_1+14$ s, in multiples of the 700 ms CSS period, and both t_1 and t_2 in Figures 21b and 21c correspond to t_0 in Figure 12b). The level at S_{out} is measured using a meter conforming to the characteristics of clause 6.4.1.2.1.

NOTE 1 – The value of $t_2 \geq t_1+14$ s in Test 5B is chosen so that at least 10 s is allowed for re-convergence, and the echo path change aligns with the start of a CSS burst.

NOTE 2 – Test 5B is designed to represent a real-world call transfer scenario and the example given in Figures 21b and 21c refers to the transfer between a 2-wire telephone and a 4-wire ISDN telephone. In the case of 4-wire ISDN telephones, the electrical ERL would normally be infinite, but the requirement is for an ERL ≥ 46 dB. This is to allow for the possibility of acoustic echo from the handset of the 4-wire telephone as given in Recommendation [ITU-T P.310]. Note that the 46 dB Weighted Terminal Coupling Loss (TCLw) given in Recommendation [ITU-T P.310] is a frequency-weighted measurement. However, for the purpose of this test, the ERL definition in ITU-T G.168 (see clause 3.14) should be used. Suggested values of high ERL to use for this test are 46 dB, 55 dB and infinity. Note also that the acoustic echo is modelled using a linear echo path.

NOTE 3 – Test 5A requirements refer to Figure 12a and Test 5B requirements to Figure 12b. This is intentional, since the transition to an infinite ERL in Test 5A means that the theoretical loss $L_{\text{Rin,act}} - L_{\text{RES}}$ (due to the combined signals from the echo path and echo canceller) cannot be less than 6 dB. This is not the case for Test 5B, where the ERLs are finite and a theoretical loss $L_{\text{Rin,act}} - L_{\text{RES}}$ of less than 6 dB immediately after the ERL transition is possible.

NOTE 4 – The requirement of Test 5B specifies only a change to the ERL. However, the call transfer scenarios that this test is designed to emulate would normally also include changes to the echo path model and the echo path delay. The user is therefore encouraged to consider these aspects when performing this test.

6.4.2.7 Test No. 6 – Non-divergence on narrow-band signals

This test has the object of verifying that the echo canceller will remain converged for subscriber-originated narrow-band signals after having converged on a wideband signal. The residual echo level is measured before and after the application of a sinusoidal wave or a wave composed of two frequencies.

The method consists of completely converging the echo canceller as in Test 2. The sequence of tones of Table 2 is then applied at R_{in}. After the sequence is completed, the adaptation is inhibited and the residual echo is measured with the signal of Test 2. The period of ≥ 1 s allowed following the adaptation inhibit command should be applied. The NLP should be disabled.

Requirement

The echo canceller is fully converged as in Test 2 for all values of ERL ≥ 6 dB, and echo path delay $t_d \leq \Delta$ ms. Once chosen, the echo path remains the same throughout the test. This test incorporates both CSS and tones. The scaling factors defined in Table D.1a should be used to scale the coefficients of the echo path models defined in Annex D. The sequence of tones of Table 2 (with the higher frequency 2 dB higher than the lower frequency for the dual tone combinations) is then applied for 5 s each to R_{in} such that $L_{\text{Rin}} \geq -30$ dBm0 and ≤ 0 dBm0. After the application of the tone sequence, the adaptation is inhibited. Using the same signal as in Test No. 2, the measured residual echo should not degrade more than 10 dB from the requirements of Figure 11 of Test No. 2. The NLP is disabled for this test. The level measurement device of clause 6.4.1.2.1 should be used to measure the signal at S_{out}.

Table 2 – Sequence of tones

697
941
1336
1633
697 and 1209
770 and 1336
852 and 1477
941 and 1633

6.4.2.8 Test No. 7 – Stability test

The object of this test is to verify that the echo canceller will remain stable for narrow-band signals. The residual echo is measured throughout the application of a mono-frequency sinusoidal wave.

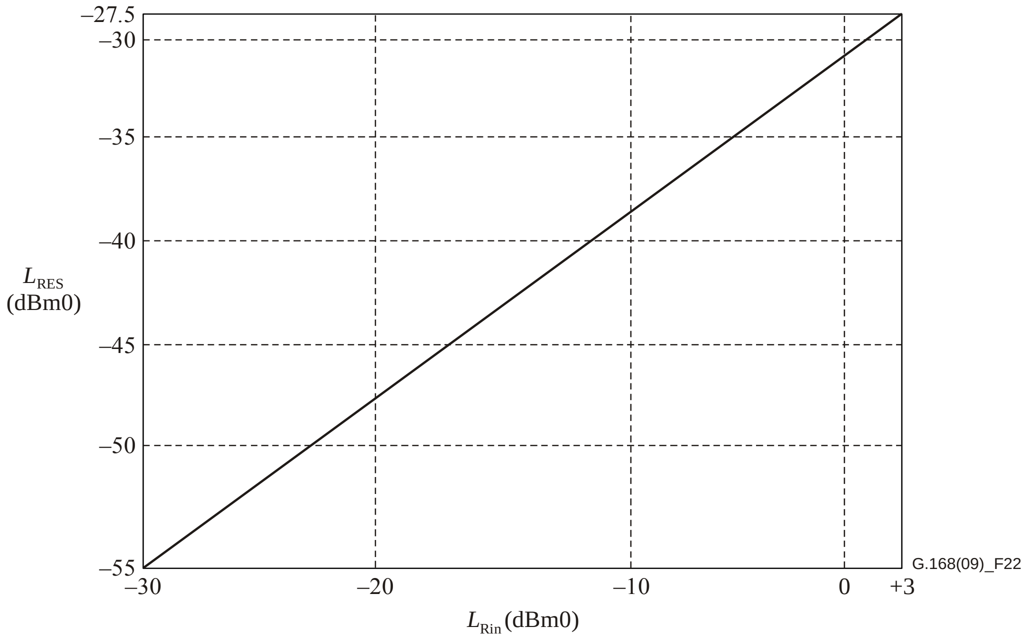
The test method is as follows: with the H register initially reset, and the NLP disabled, the echo canceller is converged on the sinusoidal signal for two minutes. The residual echo level is measured continuously throughout the two minutes that the input signal is applied.

Requirement

With the echo canceller H register initially reset, apply a mono-frequency signal, except for those identified in Table 3 of Test No. 8, for two minutes at R_{in}. The residual echo level, measured continuously throughout the two minutes that the input signal is applied, should be less than or equal to that shown in Figure 22 after an initial adaptation period of 10 s. The level of R_{in} for this test should be in the range $L_{Rin} \geq -30$ dBm0 and $\leq +3$ dBm0, with an echo path consisting of two ERL values of 6 dB and 20 dB, with an echo path delay $t_d \leq \Delta$ ms.

NOTE – t_d is assumed to be negligible relative to the adaptation period allowed in this test. It is therefore not accounted for explicitly in the requirement.

The residual echo level is measured using a meter conforming to the characteristics of clause 6.4.1.2.1.



Note that the equation for the line in Figure 22 is: $L_{Res} = 0.83L_{Rin} - 30$.

Figure 22 – Performance requirements for Test No. 7

6.4.2.9 Test No. 8 – Non-convergence of echo cancellers on specific ITU-T Nos. 5, 6 and 7 in-band signalling and continuity check tones (optional)

Echo cancellers, which are not externally disabled by the switch and which are located on the line side of Signalling System Nos. 5, 6 and 7 in international exchanges or are associated with national exchanges, should minimize the impairment, if any, on specific in-band signalling and continuity check tones. The following tests simulate a possible network scenario where a call is terminated after an echo canceller has converged during the call, and then signalling tones are applied. The tests are meant to ensure that echo cancellers will not remove or cancel a mono or bi-frequency signal transmitted in a handshaking protocol in the transmit direction either before or after receiving the same signal (except for amplitude and phase) in the receive direction. This is intended to allow a correct transmission of specific signalling or continuity check tones without externally disabling the echo canceller. The NLP should be enabled.

For an echo canceller equipped with this optional capability, the echo canceller is initially converged using CSS to a given echo path with an ERL ≥ 6 dB, and an echo path delay $t_d \leq \Delta$. For simplification, the fully converged state may be achieved with an ERL of 6 dB. Then in sequence the CSS is removed, the echo path is opened at R_{out}, and one of the signals from Table 3 for which the echo canceller is optioned is applied at S_{in}. Within 90 ms either before or after the application of the signal at S_{in}, the same signal (may differ in amplitude and phase) is applied at R_{in}. After the detection time, the level at S_{out} is measured.

The power level M, of each signal applied should be within the following limits:

System 5: –16 ≤ M ≤ –2 dBm0

NOTE 1 – This range applies to the single frequency signals f1 and f2. The level of the individual signals in the compound signal may differ from each other by not more than 5 dB. Note that when the levels of the individual signals approach –2 dBm0, the compound signal may be clipped. The nominal transmit level is specified as –9 ± 1 dBm0.

See [ITU-T Q.143] and [ITU-T Q.144].

System 6: –18 ≤ M ≤ –6 dBm0

NOTE 2 – The nominal transmit level is –12 ± 1 dBm0.

See [ITU-T Q.271].

System 7: –18 ≤ M ≤ –6 dBm0

NOTE 3 – The nominal transmit level is –12 ± 1 dBm0.

See clause 7 of [ITU-T Q.724].

The above levels are designed to ensure that the echo canceller will operate with signals that occupy the entire range of levels given in the appropriate signalling ITU-T Recommendations.

The echo canceller response should be sufficiently long to provide immunity from false operation due to voice signals. Also, the echo canceller response should not be so long as to needlessly extend the time for the signal to appear on the network.

Requirement

The level at S_{out}, as measured using the method of clause 6.4.1.2.1, should not vary more than 2 dB with respect to the level at S_{in} after an allowed 1 second detection time.

Table 3 – Applicable signalling tones

System 5	System 6	System 7
2400 ± 15 Hz	2000 ± 30 Hz	2000 ± 30 Hz
2600 ± 15 Hz		
2400 ± 15 Hz and 2600 ± 15 Hz		

6.4.2.10 Test No. 9 – Comfort noise test

6.4.2.10.1 Test No. 9A – Comfort noise performance with step change of background noise level

This test is meant to ensure that the echo canceller is able to provide a comfort noise signal on S_{out} which matches noise received on S_{in}. It also tests whether the canceller is able to adjust the level of this comfort noise signal to compensate for changes in the level of input noise. As this test is not intended as a test of echo cancellation capability, an ERL of 12 dB is used for the entire test. The steps of this test should be applied in sequence. They consist of setting the level of S_{gen} to a value between –50 dBm0 and –40 dBm0, lowering the level of S_{gen} by δ dB, and then raising the level of S_{gen} back to the initial level. Consequently, the test covers a range of operation for L_{Sgen} between -60 dBm0 and –40 dBm0. The level of R_{in} will be at silence or at –10 dBm0. White Gaussian Noise, band-limited 300 Hz – 3400 Hz, with a crest factor of 11 dB ± 1 dB is used for all input signals for this test. The level at S_{out} is measured using the RMS method of clause 6.4.1.2 (i.e., no filtering should precede the RMS calculation). The NLP and comfort noise feature should be enabled.

a) Part 1 (matching)

1) Set L_{Sgen} to a level between –50 dBm0 and –40 dBm0.
2) Set L_{Rin} to silence (< –60 dBm0) and hold for 10 s.
3) L_{Rin} to –10 dBm0.
4) Measure L_{RET} after 2 s.

Requirement

For all values of L_{Sgen}, L_{RET} should be within 2.0 dB of L_{Sgen} when measured over a 700 ms window. Also, this value should hold as long as noise level L_{Sgen} remains constant, and sequential 700 ms measurements should be made over a minimum period of 7.7 s to check this.

b) Part 2 (adjustment down)

The following 8 steps should be performed in sequence for δ = 5 dB and δ =10 dB:

1) Lower L_{Sgen} by δ dB from the level in Part 1.
2) Set L_{Rin} to silence (< –60 dBm0) and hold for 10 s.
3) L_{Rin} to –10 dBm0.
4) Measure L_{RET} after 2 s.

Requirement

L_{RET} should be within 2.0 dB of L_{Sgen} when measured over a 700 ms window. Also, this value should hold as long as noise level L_{Sgen} remains constant, and sequential 700 ms measurements should be made over a minimum period of 7.7 s to check this.

c) Part 3 (adjustment up)

5) Raise L_{Sgen} by δ dB from the level in Part 2.
6) Set L_{Rin} to silence (< –60 dBm0) and hold for 10 s.
7) Set L_{Rin} to –10 dBm0.
8) Measure L_{RET} after 2 s.

Requirement

L_{RET} should be within 2.0 dB of L_{Sgen} when measured over a 700 ms window. Also, this value should hold as long as noise level L_{Sgen} remains constant, and sequential 700 ms measurements should be made over a minimum period of 7.7 s to check this.

6.4.2.10.2 Test No. 9B – Comfort noise performance with dynamic change of background noise level

This test examines the dynamic behaviour of the comfort noise generator by investigating its performance during the continuous change of background noise level. The test is divided into two parts:

– **Part 1 – Mandatory:** This part is meant to ensure that the echo canceller is able to provide a comfort noise signal on S_{out} which tracks the changing level of a background noise signal at S_{gen}. This part has requirements that are derived from Test No. 9A and is performed over the same range of noise levels at S_{gen} (–60 to –40 dBm0).

– **Part 2 – Optional:** This part is meant to examine the behaviour of the comfort noise generator for changing levels of background noise over the range from –70 to –40 dBm0.

Altogether, the two parts of this test provide an analysis of comfort noise stability over a wide range of varying background noise levels. The test is performed with different values of ERL to simulate different network configurations, including 2-wire and 4-wire terminations.

NOTE 1 – This test is designed to investigate the overall performance of the NLP and the Comfort Noise Generator algorithm during the continuous change of background noise level. For example, the 2 seconds long active high level R_{in} parts usually trigger the NLP and with that the generation of comfort noise, whereas the 2 second long silent periods allow the Comfort Noise Generator to learn the level and characteristic of the new background noise. However this behaviour it not ensured. It depends on the NLP design. This means that the S_{out} signal may not be generated by the Comfort Noise Generator, but it can be the background noise itself. Care should be taken if someone wants to investigate the performance of the Comfort Noise Generator algorithm itself with this test.

Test signals

The R_{in} test signal duration is 350 seconds long. The first 10 seconds of this test vector consist of 1 second of silence, followed by 6 cycles (12 bursts) of single-talk CSS at –10 dBm0 followed by silence. The remaining duration of this test vector consists of 2 seconds of continuous CSS (continuous CSS is defined as CSS bursts with the silence intervals removed) followed by 2 seconds of silence with a logarithmically increasing continuous CSS test signal level from –40 to –10 dBm0 for 170 seconds followed by a logarithmically decreasing continuous CSS test signal level from –10 to –40 dBm0 for the remaining 170 seconds.

The S_{gen} test signal duration is also 350 seconds long. The first 10 seconds of this test vector consist of 5.5 seconds of silence followed by 6 cycles (12 bursts) of single-talk CSS at –40 10 dBm0 followed by silence. The remaining duration of this test signal consists of White Gaussian Noise, band-limited 300 Hz – 3400 Hz, with a crest factor of 11 dB ± 1 dB increasing logarithmically in level from –86 to –20 dBm0 for 170 seconds and then logarithmically decreasing in level from –20 to –86 dBm0 for the remaining 170 seconds.

NOTE 2 – The first 10 seconds of this test that include 12 bursts of CSS signal in R_{in} and 12 bursts of CSS in S_{gen} are optional and they only serve as a simple pre-conditioning to allow consistent test results for the CNG measurements.

NOTE 3 – A logarithmic rate of increase and decrease in signal level is used, since the natural sensitivity of the human ear is logarithmic with respect to noise level perception. Furthermore, the use of logarithmically increasing and decreasing signal levels for testing allows for linear presentation on a dB scale.

NOTE 4 – For simplicity, a constant signal level at R_{in} may be used, but this may result in crosstalk when using analogue interfaces, especially for S_{out} signal levels between –80 and –60 dBm0. If using a constant level at R_{in}, a value of –20 dBm0 is recommended, although anywhere in the range –30 to –10 dBm0 is acceptable.

Test procedure

With the H register initially reset, or alternatively, with an open echo path resulting in S_{in} = 0 and the H register content converged to 0, the echo canceller must then be fully converged using a CSS test signal at R_{in} for ≥30 seconds for a given value of echo path delay, *TD* ≤ Δ ms, and ERL ≥ 6 dB. The NLP and comfort noise feature should be enabled for this test.

Optionally, Test 2A-a may be conducted to allow optimum echo canceller convergence prior to Test No. 9b execution. This optional pre-conditioning allows sufficient time for echo canceller convergence and minimizes any variance in the test results when performing this test.

Apply the R_{in} and S_{gen} signals simultaneously and record S_{out} (and optionally R_{in} and/or R_{out}) for the duration of the test. The level at R_{in}, S_{out} and S_{gen} should be measured using the method of clause 6.4.1.2.1.

Requirement

Part 1 – Mandatory: L_{RET} should be within +/–6 dB of L_{Sgen} when L_{Sgen} is within the range –60 to –40 dBm0.

Part 2 – Optional: L_{RET} should be within +/–6 dB of L_{Sgen} when L_{Sgen} is within the range –70 to –40 dBm0.

As a minimum, this test should be performed with ERL values of 40, 15 and 6 dB.

Example results for this test are shown in Figures 23a, 23b and 23c. These tests were performed at different values of ERL.

NOTE 5 – For each of the following figures, the green trace labelled Source is used to represent the R_{in} test signal, and the blue trace labelled Ch1 is used to represent the S_{out} signal.

NOTE 6 – In Figure 23a, the test has been performed with a 4-wire network configuration (ERL ≥ 40 dB) to obtain a baseline performance of the echo canceller comfort noise generation before repeating the test for 6 and 15 dB ERLs. Optionally, the echo canceller may be disabled to verify similar baseline performance to the 4-wire network configuration. It is highly recommended to investigate any significant inconsistencies.

NOTE 7 – To visualise the requirements of this test, Figures 1 to 3 show the approximate upper and lower limits of L_{RET} relating to both the *Mandatory* and *Optional* test requirements. These limits are relative to the average level of L_{Sgen} and not to the values of L_{Sgen} as measured at the output of the 35 ms filter using the method of clause 6.4.1.2.1. The latter approach, which is used in the requirements, results in a measurement uncertainty of up to 2 dB (see Appendix III of [b-ITU-T G.160]). This uncertainty has been taken into account in the +/–6 dB tolerance allowed within the requirement. This effect is for further study.

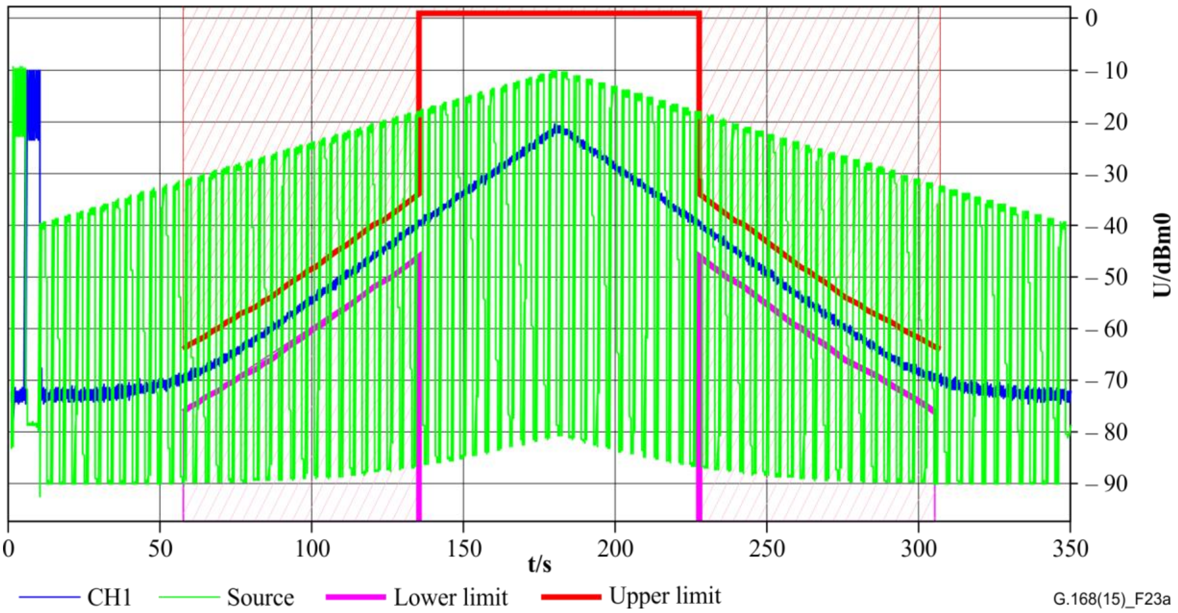


Figure 23a – Dynamic comfort noise level matching at ≥40 dB ERL

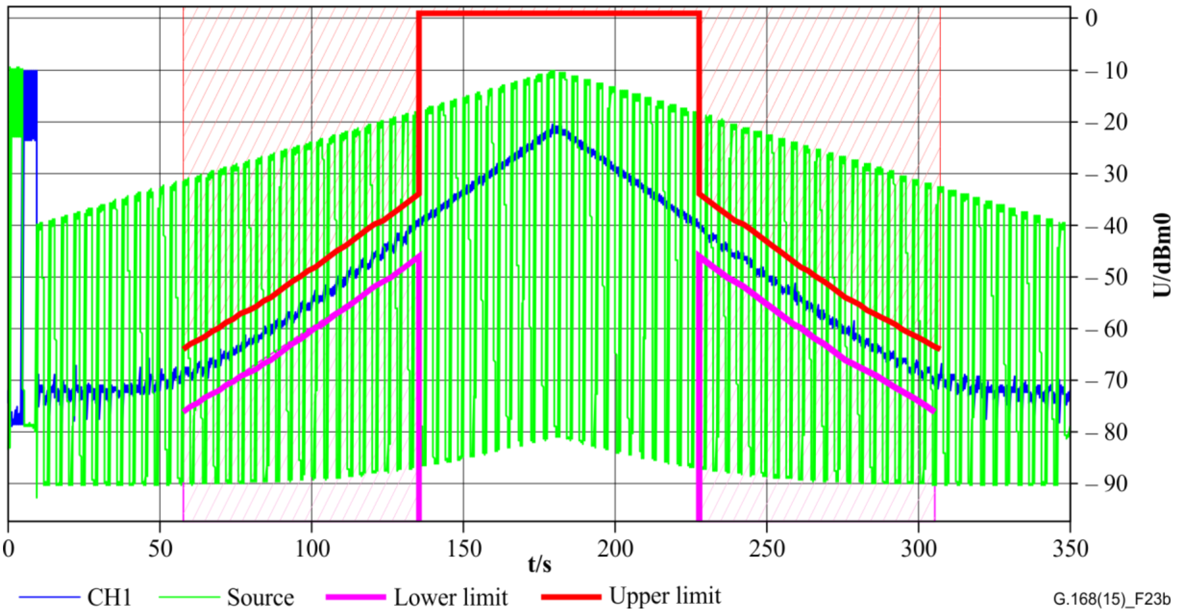


Figure 23b – Dynamic comfort noise level matching at 15 dB ERL

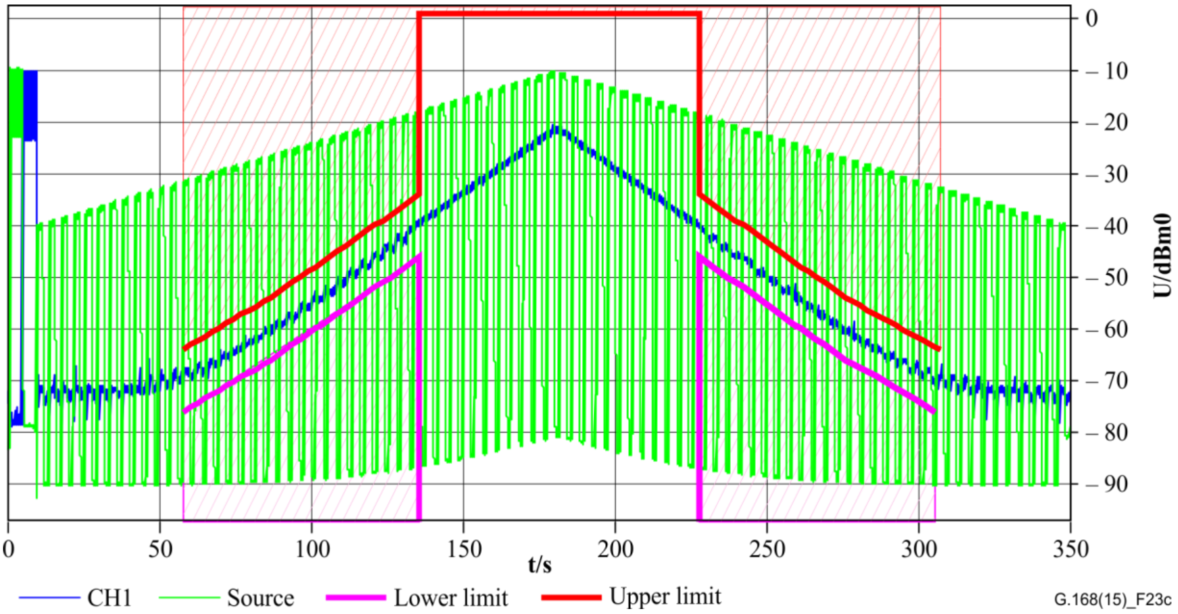


Figure 23c – Dynamic comfort noise level matching at 6 dB ERL

6.4.2.11 Test No. 10 – Facsimile test

This test is meant to ensure that the echo cancellers located at each end of a connection converge rapidly on the initial handshaking sequences of a facsimile call and have some means to prevent the unwanted generation of echo. The test and requirements were originally developed to overcome problems in the network due to the turnaround of fax handshaking signals (see clause 5.2.1 of [ITU-T G.161]: "Interaction of echo cancellers with facsimile transmission"). Customer complaints were reduced following the introduction of this test. The test is split into three parts. Test 10A looks at the performance of the echo canceller located on the calling station side, and Test 10B looks at the performance of the echo canceller on the called station side. Test 10C looks at the performance of the echo canceller on the calling station side during page transmission.

NOTE 1 – The vast majority of echo cancellers use some form of the least-squares algorithm to model the echo path impulse response. When a sinusoidal input signal or a narrow-band signal like ITU-T V.21 is used as the input for the echo canceller, an exact model of the echo path is not possible. This response to a sinusoidal input signal is a well-known and understood consequence of the least-squares algorithm, and can cause a problem with Test 10A and Test 10B.

NOTE 2 – In Test 10A, Region II starts during the application of the CED signal, and extends to the beginning of Sequence No. 1. Because of the response of least-squares-based algorithms to sinusoidal inputs, during that portion of Region II from the cessation of the CED signal to the beginning of Region III, the echo canceller may produce a residual echo (or echo bursts) that exceeds the recommendation limits at the end of the CED signal. The user should be aware that this may not be a failure of the echo canceller, but merely a response to the sinusoidal CED input signal. When this response occurs, additional investigation is warranted.

The test has been designed to run in a laboratory environment using an echo canceller and a fax simulator. The tests should be run separately.

NOTE 3 – When using the hybrid models given in Annex D, the K_i factors for tone signals given in Table D.1b should be used for this test.

The tests should be performed with [ITU-T G.165] or this Recommendation tone disabler switched on. The requirements with NLP disabled are now more relevant following the introduction of the option to disable the NLP during fax transmission as described in clause 7.1.

The provisioned state of the NLP for Test 10A is enabled. However, if the NLP is reliably disabled automatically by means as described in clause 7.1, then only the requirements for NLP disabled are relevant. If NLP can be enabled or disabled as a configurable feature, then Test 10A shall exercise both modes of operation. This statement may also be applied to Tests 10B and 10C if the NLP is guaranteed to remain disabled throughout the fax call. If the NLP is likely to re-enable due to sufficiently long gaps during sequences of message exchange (during a facsimile call, there are a number of silent periods that may be long enough to permit the re-enabling of the NLP, see clause 5.2.1 of [ITU-T G.161]), then the requirements for NLP enabled in Tests 10B and 10C apply. Test results showed that the periods of silence during sequences of message exchange may extend to more than the maximum release time 400 ms. A specific test for this mode of operation is for further study.

For this purpose, the following signals should be applied (bits are transmitted left to right). The initial flag is repeated 37 times for each sequence.

FAX test sequences:

Calling tone (CNG)

Conditions

Signal 1100 Hz ± 38 Hz

Duration On for 0.5 s, Off for 3 s (±15%)

Called station identification (CED)

Conditions

Signal 2100 Hz ± 15 Hz

Duration 2.6-4 s

Binary coded sequences

Sequence No. 1 (called station):

Non-standard facilities (NSF) frame:

Flag	HDLC address field	HDLC control field	Control field NSF	Information field, 8 octets (country, manufacturer, additional code number)	Frame check sequence	Flag
0111 1110	1111 1111	1100 1000	0000 0100	0101 0101, 0101 0101, 0101 0101, ...	1010 1010	0111 1110

Called subscriber identification (CSI) frame:

Flag	HDLC address field	HDLC control field	Control field CSI	Information field, 20 octets (receiver code number)	Frame check sequence	Flag
0111 1110	1111 1111	1100 1000	0000 0100	0101 0101, 0101 0101, 0101 0101, ...	1010 1010	0111 1110

Digital identification signal (DIS) frame:

Flag	HDLC address field	HDLC control field	Control field DIS	Information field, 3 octets	Frame check sequence	Flag
0111 1110	1111 1111	1100 1000	0000 0001	0101 0101, 0101 0101, 0101 0101	1010 1010	0111 1110

Sequence No. 2 (calling station):

Non-standard set-up (NSS) frame:

Flag	HDLC address field	HDLC control field	Control field NSS	Information field, 3 octets	Frame check sequence	Flag
0111 1110	1111 1111	1100 1000	1100 0100	0101 0101, 0101 0101, 0101 0101	1010 1010	0111 1110

Transmitting subscriber identification (TSI) frame:

Flag	HDLC address field	HDLC control field	Control field TSI	Information field, 20 octets (Transmitter code number)	Frame check sequence	Flag
0111 1110	1111 1111	1100 1000	1100 0010	0101 0101, 0101 0101, 0101 0101, ...	1010 1010	0111 1110

Digital command signal (DCS) frame:

Flag	HDLC address field	HDLC control field	Control field DCS	Information field, 20 octets (Transmitter code number)	Frame check sequence	Flag
0111 1110	1111 1111	1100 1000	1100 0001	0101 0101, 0101 0101, 0101 0101, ...	1010 1010	0111 1110

Data transmission conditions

The transmission of sequences No. 1 and No. 2 in the telephone channel is obtained by means of frequency shift (see [ITU-T V.21]).

Conditions

Data signalling rate, synchronous	300 bit/s
Centre frequency	1750 Hz
Frequency deviation	±100 Hz
Characteristic frequencies	1650/1850 Hz
Tolerances of the characteristic frequencies	±6 Hz

The higher characteristic frequency corresponds to a binary "0".

6.4.2.11.1 Test No. 10A: Canceller operation on the calling station side

The convergence test procedure is to reset the H register and to inhibit adaptation. Then adaptation is enabled while CNG, CED and sequence No. 1 are applied (see Figure 24). During the adaptation interval, the residual/returned echo level is measured. This test should be performed with the NLP both enabled and disabled.

Requirement

With the H register initially reset and the value L_{Rin} = –13 dBm0 for the entire test, the following specific requirements apply for all values of ERL ≥ 6 dB and echo path delay $t_d \le \Delta$ ms. The test should run by applying CED and repeating sequence No. 1 four times.

Region I (converging on CED tone)

– the peaks (see clause 6.4.1.2.2) of L_{RES} should be ≤ max(–37 dBm0, (–13 – A_{ECHO}) dBm0);

– the time to enter Region II should be ≤ (0.15 s + t_{d}) with respect to the start of the CED signal.

Region II (converged on CED tone)

– the peaks (see clause 6.4.1.2.2) of LRES should be ≤ –37 dBm0.

Region III (converging on sequence No. 1)

– the peaks (see clause 6.4.1.2.2) of LRES should be ≤ max(–37 dBm0, (–13 – AECHO) dBm0);

– the time to enter Region IV should be ≤ (1.3 s + td) with respect to the start of the Sequence No. 1 signal.

Region IV (converged on sequence No. 1)

– the peaks (see clause 6.4.1.2.2) of L_{RES} should be ≤ –24 dBm0.

If the NLP is enabled, the peaks (see clause 6.4.1.2.2) of L_{RET} should be ≤ –37 dBm0 in Regions II and IV.

NOTE – The –37 dBm0 upper limit on the L_{RES} requirement in Regions I and III is to address an issue with the quantization error that may occur with low level A-law signals and high values of ERL.

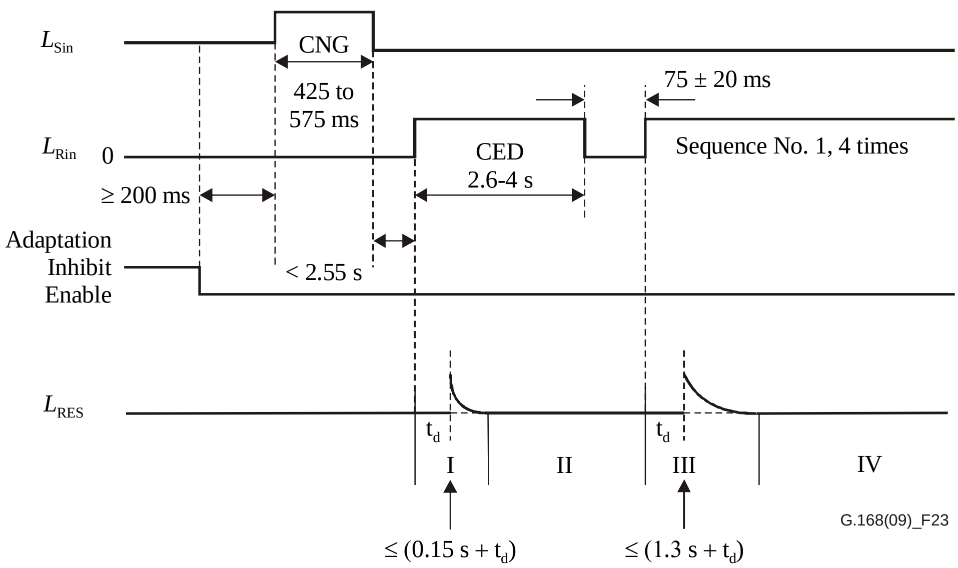


Figure 24 – Test No. 10A signal and time relationships

6.4.2.11.2 Test No. 10B: Canceller operation on the called station side

The convergence test procedure is to reset the H register and to inhibit adaptation. Then adaptation is enabled, while sequence No. 2 is applied (see Figure 25). During the adaptation interval, the residual/returned echo level is measured. This test should be performed with the NLP both enabled and disabled.

Requirement

With the H register initially reset and the value L_{Rin} = –13 dBm0 for the entire test, the following specific requirements apply for all values of ERL ≥ 6 dB and echo path delay t_d ≤ Δ ms. The test should run by repeating sequence No. 2 four times.

Region I (converging on sequence No. 2)

– the peaks (see clause 6.4.1.2.2) of L_{RES} should be ≤ max(–37 dBm0, (–13 – A_{ECHO}) dBm0);

– the time to enter Region II should be ≤ (1.3 s + t_d) with respect to the start of the Sequence No. 2 signal.

Region II (converged on sequence No. 2)

– the peaks (see clause 6.4.1.2.2) of L_{RES} should be ≤ –24 dBm0.

If the NLP is enabled, the peaks (see clause 6.4.1.2.2) of L_{RET} should be ≤ –37 dBm0 in Region II.

NOTE – The –37 dBm0 upper limit on the L_{RES} requirement in Region I is to address an issue with the quantization error that may occur with low level A-law signals and high values of ERL.

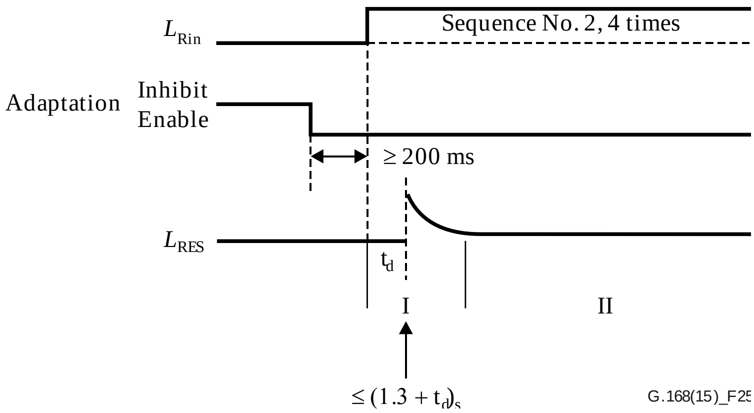


Figure 25 – Test No. 10B signal and time relationships

6.4.2.11.3 Test No. 10C: Canceller operation on the calling station side during page transmission and page breaks (for further study)

Figure 26 shows the sequence of message exchange for a typical fax transmission consisting of two pages. The sequence begins with an ITU-T V.21 message handshake procedure. Operation and performance of the echo cancellers at each end of the link are tested during this period by Tests 10A and 10B. This test is designed to check the operation and performance of the echo canceller at the calling station side during page transmission and page breaks as shown in Figure 26.

Test 10C uses data files A and B and software implementations of FSK and PSK modems. The test described uses ITU-T V.29 modulation for the transmission of image data, but can be further extended to cover different types of modems such as ITU-T V.17 and ITU-T V.27 *ter* as well as different timing relationships between handshake signals (represented by the ITU-T V.21 modem) and page transmission (represented by the ITU-T V.29 modem). The timing relationship shown in Figure 27 has been designed to mimic the real facsimile T.30 protocol.

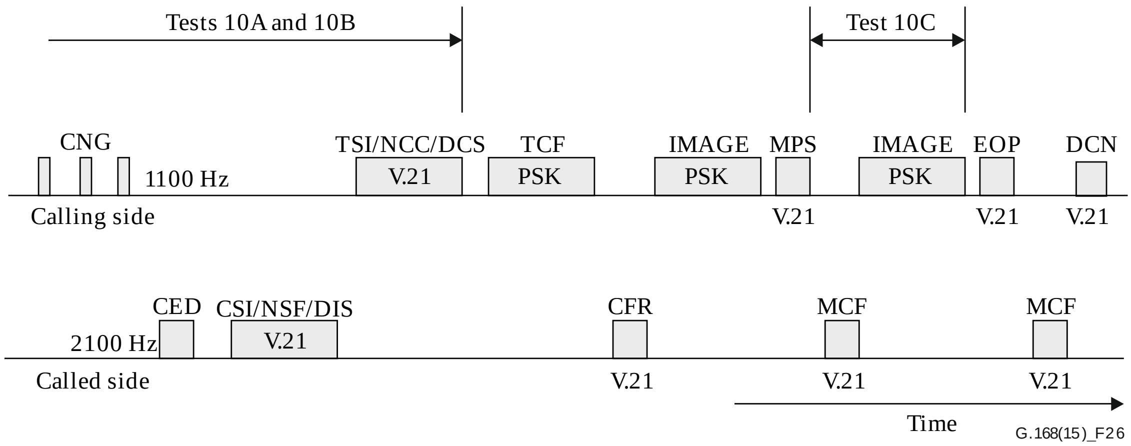


Figure 26 – Sequence of message exchange for a typical 2-page fax transmission

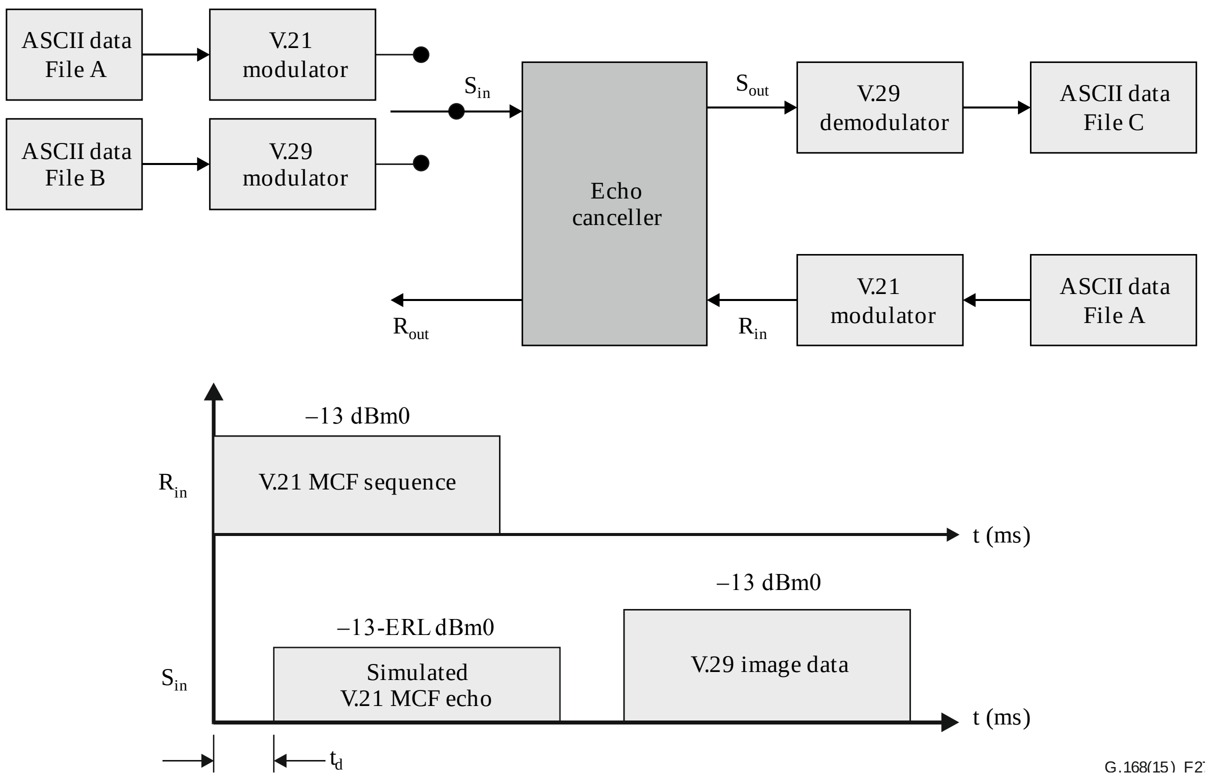


Figure 27 – Canceller operation during page transmission and page breaks

The test procedure is to reset the H register and inhibit adaptation. Adaptation is then enabled and the canceller is converged by sending data file A via the ITU-T V.21 modulator, into R_{in} of the echo canceller (see Figure 27). This is meant to simulate the MCF signal transmitted by the called fax machine at the end of the first page. To simulate the echo signal, data file A is also transmitted (using the ITU-T V.21 modulator) to the echo canceller's S_{in} input, at a level of $\geq$ 6 dB below R_{in} and with a timing offset of t_d. Following this, data file B is transmitted (using the ITU-T V.29 modulator) to the echo canceller's S_{in} input.

The ITU-T V.29 data from the echo canceller's S_{out} output is first collected, stored and demodulated off-line. The recovered file C is then analysed. The data from the echo canceller's R_{out} output is discarded (on the assumption that no impairments are introduced).

Note that in the latest version of T.30, an echo protect tone (EPT) is specified before all ITU-T V.29 transmissions (TCF and image data). This EPT signal was originally intended to "turn around" echo suppressors and prevent any front-end clipping of TCF or image data. In the case of echo cancellers, it helps to open the NLP before the transmission of TCF or image data to avoid corruption. However, many of the fax implementations currently in the field do not incorporate this EPT signal, so it has not been included in this test. In this respect, this test therefore represents a worst-case scenario.

The NLP should be enabled during this test. Note that some echo cancellers will automatically disable the NLP on detection of a 2100 Hz tone.

Data File A

In the following table, all data is given in hex. The initial flag is repeated 37 times.

Flag	HDLC address field	HDLC control field	Control field MCF	Frame check sequence	Flag
7E	FF	C8	B1	D4, 07	7E

Data File B

Data File B should be taken from the following ITU-T CD-ROM:

– ITU-T Recommendation T.24 (1998), *Standardized digitized image set,* Document No. 1, "slerexe" letter.

Requirement

The intent is to compare the results with the echo canceller first disabled and then enabled. The actual requirements are for further study. The requirements should be based on a comparison of files B and C.

6.4.2.12 Test No. 11 – Tandem echo canceller test (for further study)

This test is meant to ensure that echo cancellers operating in tandem will not degrade the quality of a call below acceptable limits. See [ITU-T G.161] for further discussion on this issue. The following points summarize the philosophy of the approach to this test:

- The echo canceller "under test" should be used in both positions (cancelled end and tandemed – see Figure 28). This is the simplest case and avoids the necessity to test with combinations of other canceller types. The philosophy behind this approach is that if a canceller is designed to perform well in tandem with itself, then it has a good chance of performing well with other types too. It also makes it easier to demonstrate "compliance" to this Recommendation since it means only one set of tests needs to be performed.
- Neither of the cancellers in the test are controlled (h-reset, etc.). This reduces the complexity of the test and is also representative of what happens in practice (one of the reasons that tandem echo cancellers will appear in the field is if they are not controlled by the switch on call set-up).
- Only steady state performance is tested. There may be convergence artefacts, but the dominant effect on voice quality as perceived by the users is what happens during steady state.

Figure 28 shows the proposed set-up for the tandem echo canceller test and the conventions used.

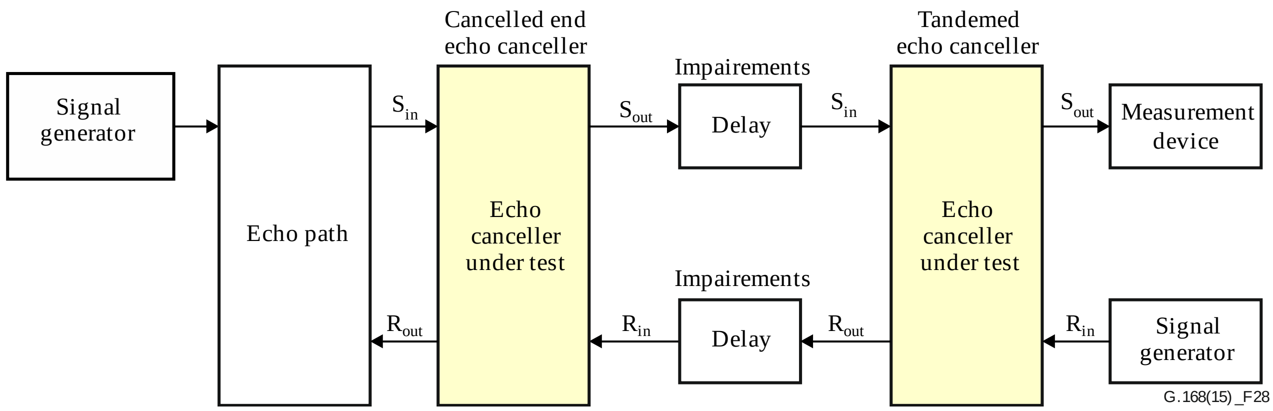


Figure 28 – Test No. 11 configuration

Note that the terms *cancelled end* and *tandemed* are used to distinguish between the two echo cancellers. The *cancelled end* echo canceller is the one nearest the echo source and therefore doing most of the work. The *tandemed* echo canceller is the one in tandem furthest from the echo source and doing least work.

The impairments block shows delay. The test should be performed with values of delay both inside and outside the tail capacity of the *tandemed* echo canceller.

Initial conditions – To be determined. For example, the canceller either converged to 0 or any other echo path; NLP enabled; etc.

Test signals – To be determined.

Measurement methods – To be determined.

Requirements – TBD.

6.4.2.13 Test No. 12 – Residual acoustic echo test

See clauses I.6.3, I.6.3.1 and [ITU-T G.161] for further discussion on this issue. This test is meant to check the performance of the echo canceller in the presence of residual acoustic echo and to ensure that its performance with electric echo is not overly degraded in the presence of acoustic echo.

The following example of a call to a loud-speaking telephone illustrates the rationale behind this test:

A user makes a call to a loud-speaking telephone. The call is answered first using the handset, and is then switched to loud-speaker mode to share the conversation with other listeners in the room. The call is then returned to handset mode for a private conversation before clearing down.

This scenario is summarized in Figure 29. Phases 1 and 3 correspond to the presence of electric echo only. Phases 1 and 3 have the same "echo" properties, i.e., the same ERL and the same echo path. Phase 2 corresponds to the presence of electric and acoustic echoes. In terms of echo properties, Phase 2 is associated with a lower ERL (presence of residual acoustic echo + electrical echo) and an echo path that is not the same as Phases 1 and 3. An echo canceller involved in such a scenario is required to converge three times during the call.

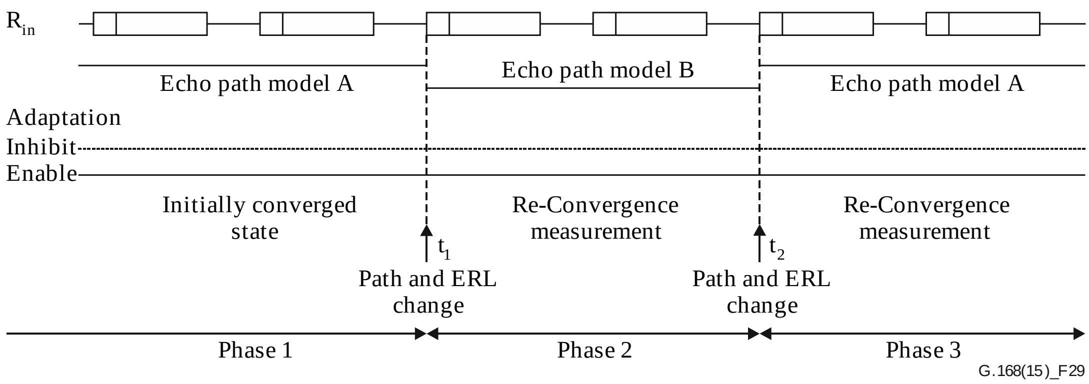


Figure 29 – A conversation scenario with acoustic echo

Test procedure

The test procedure of Tests 2A should be used for this test. The following echo path characteristics should be used for each phase:

Phase 1: ERL_A and Echo Path Model A, where $ERL_A \geq 16$ dB

Phase 2: $ERL_B = ERL_A - 10$ dB and Echo Path Model B (B not equal to A)

Phase 3: ERL_A and Echo Path Model A

NOTE – Echo Path Model A may be any model represented by Figures 6 and 7. Echo Path Model B should be g(k) models m_4, m_7 and m_8 defined in Annex D, since these models represent echo paths with long dispersion times.

Requirement

The convergence requirement of Test 2A(a) should apply to Phase 1 of Figure 29 and the re-convergence requirement of Test 2A(b) should apply to Phases 2 and 3 of Figure 29, without resetting the H-register.

NOTE 1 – The test procedure and requirements defined above do not consider the possibility of non-linear echo paths or echo paths that extend beyond the tail capacity of the echo canceller that may be present in real conversation scenarios.

NOTE 2 – The requirement above is a sub-set of the existing requirements for Test 2.

NOTE 3 – This test may be enhanced by the future addition of test vectors that represent real acoustic echo signals. This enhancement is for further study.

6.4.2.14 Test No. 13 – Performance with ITU-T low-bit rate coders in echo path (optional)

This test is meant to ensure that the echo canceller will not degrade the performance of the network beyond acceptable limits when any ITU-T low-bit rate coder, as stated in the requirement, is included in the echo path of the echo canceller. The user should test the echo canceller using the stated ITU-T low-bit rate coders that may appear in the echo path of the echo canceller in the user's network. See clauses I.5.3, I.5.4 and [ITU-T G.161] for further discussion of ITU-T low-bit rate coders and non-linearities in the echo path.

The test method is to place the echo canceller in the test configuration of Figure 30. Tests 2A and 2B are performed with the echo canceller disabled, and the residual echo level measured using a meter conforming to the characteristics of clause 6.4.1.2.1. The highest measured level is used as a baseline for the next stage of the test. Tests 2A and 2B are then repeated with the echo canceller enabled and the performance is compared with the baseline.

The impact on speech quality performance by low-bit-rate coders in the echo path is for further study.

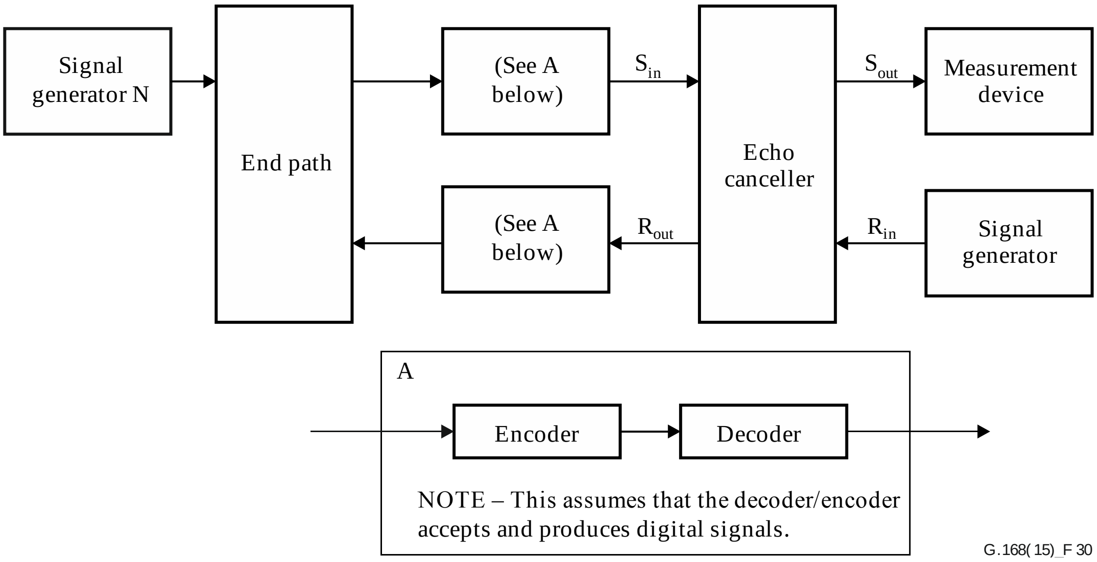


Figure 30 – Test No. 13 configuration

NOTE 1 – The interface to the echo canceller shown in Figure 30 conforms to [ITU-T G.711].

Requirement

For the stated ITU-T coders, the residual echo power measured (see clause 6.4.1.2.1) with the echo canceller enabled should be no more than 0.5 dB greater than the measured baseline with the echo canceller disabled. In addition, no peaks (see clause 6.4.1.2.2) are allowed that exceed 5 dB above the measured baseline with the echo canceller disabled. The coders for which this test is applicable are to be stated by the user.

NOTE 2 – This test includes non-linear echo path. See clauses I.5.3, I.5.4, I.6.3 and [ITU-T G.161] for further discussion on this issue.

NOTE 3 – A low bit-rate coder exhibits a finite throughput delay. This throughput delay will be present in the echo path of the echo canceller and should be taken into account when performing this test, such that the overall delay in the echo path does not exceed the tail capacity of the echo canceller.

NOTE 4 – The 0.5 dB tolerance is to allow for statistical variations in the measurement process. The intent is to allow a practical tolerance that is effectively inaudible.

6.4.2.15 Test No. 14 – Performance with V-series low-speed data modems

This test is meant to ensure that echo cancellers will not impair the performance of data modems and text telephones that do not disable, or allow the re-enabling of the network echo cancellers. This includes the following two cases:

i) the network echo cancellers are not disabled at the start of the call and remain enabled for the duration of the call;

ii) the network echo cancellers are disabled at the start of the call, but are re-enabled at some point or points during the call.

NOTE 1 – Text telephones commonly use different modes of operation, including switching back and forth between voice and text during the call. The switching back and forth between text and voice is the basic operational mode of textphones of type TIA 825, ITU-T V.21 110 bits/s half duplex and DTMF. [ITU-T V.18] describes an automoding procedure that collects many textphone types into one connection procedure in order to provide backwards compatibility. [ITU-T V.18] describes briefly the modulations and operating characteristics of the text telephone types. Different types dominate in different regions, but users may travel and expect to be able to use the textphones anywhere.

The scope of the test is to cover modems and text telephones that operate over the global PSTN. In practice, the set of modem types given in Table 4 represents a practical sub-set that should be used for testing. For each application listed in Table 4, a number of different modem types are also given. This test should be performed with each application and its associated modem types listed in Table 4. The intent is to allow the echo canceller vendor to demonstrate acceptable performance with a range of different modem types and for the network operator to gain a level of confidence that existing modem traffic will not be impaired.

Table 4 represents a practical sub-set of modems that are mandatory for this test. In practice, developers are encouraged to test with as wide a range of different modems as possible. A more comprehensive list (although not all inclusive) is given in Appendix V.

The performance criteria for this test depend on the type of modem being tested. In general, for low speed data modems, bit error rate (BER) is used to measure performance (measured synchronously where possible). For text telephones, however, character error rate is more appropriate. The performance criteria that should be used for each type of modem are given in Table 4.

Test method

The sequence for this test is to measure the performance of each modem type with the echo canceller DISABLED and compare this performance with the echo canceller ENABLED. The philosophy is that when enabled, the echo canceller should not degrade performance compared to that when disabled.

The echo canceller is placed in the test configuration of Figure 31. The test procedure is to first disable the echo cancellers and allow the modems to train. The modems are then operated for the time it takes to send the minimum volume of test data as defined for each modem type given in Table 4 (e.g., bit error rate, character error rate, etc.). This performance is noted and the test is then repeated with the echo cancellers enabled. First reset the H-register and ensure the NLP is enabled before allowing the modems to train.

In the test set-up, the artificial 2-wire lines and the hybrids should simulate the actual range of echo paths that the echo canceller under test is intended to cope with. For the hybrid, this means a specification of the equivalent balance network.

NOTE 2 – Examples of typical balance networks are given in Figure 11 of [ITU-T Q.552].

For the artificial line, this means a specification of the fundamental cable parameters, e.g., ohms/km and nF/km for unloaded cables. The length of the artificial lines should be variable. Test cases should include minimum and maximum lengths as well as that length for which the highest weighted echo loss, calculated according to [ITU-T G.122], is obtained.

The hybrid and artificial line arrangements should be equal at each side of the test set-up.

NOTE 3 – When evaluating the suitability of an echo canceller for a specific network, the test configuration of Figure 31 is well-suited to the purpose of Test 14. However, for generalized evaluation across a range of networks, the number of different real hybrids, balancing networks and proprietary modems makes full and thorough testing complex. In this context, 4-wire software emulations of the modems and the ranges of echo path models defined in Figures 6 and 7 provide an efficient test environment, and can be used to supplement the 2-wire modems and real hybrids of Figure 31. For these cases, line emulations would be used in place of the 2-wire artificial lines.

The values of the settings should be as follows:

R1, R2 = 6 dB to simulate access/egress loss.

T1 = 3 dB to 9 dB (3 dB is the nominal level, 9 dB simulates a 6 dB level offset).

T2 = 3 dB.

DR1, DR2 = echo path delay ≤ Δ ms.

N1, N2 = set to produce signal-to-noise ratios of not less than 25 dB, and, no noise.

D1, D2 = set to produce a round trip delay of up to 520 ms, with D1 = D2.

The transmit level at the 2-wire interface of a modem is determined by the regulations applying to the particular country in which the modem is used. This level is usually factory pre-set and is not adjustable by the user. For Test 14, where it is desired to vary the levels at the input to the network echo cancellers, this can be achieved by the use of different lengths of artificial line (W1 and W2 in Figure 31) together with the transmit attenuators T1 and T2.

Requirement

With the H register initially reset and the NLP enabled, for the conditions specified above, the percentage of data errors should not increase when the echo canceller is enabled, compared with when the echo canceller is disabled, when data is exchanged between the two terminals for the given minimum volume of data transmitted.

Table 4 – Minimum set of modems to be used for Test 14

Application	Modulation type	Data rate (bits/s)	Answer tone	Performance criteria	Minimum volume of data	Comments
Dial-up data	ITU-T V.21 ITU-T V.22 *bis*	ITU-T V.21: 300 ITU-T V.22 *bis*: 1200 and 2400	ANS	BER	ITU-T V.21: 20 kbits ITU-T V.22 *bis*: 100 kbits	
Point of sale terminal	ITU-T V.21 ITU-T V.22 *bis* Bell 103 Bell 212A	ITU-T V.21 and Bell 103: 300 ITU-T V.22 *bis*: 2400 Bell 212A: 1200	Proprietary	BER	ITU-T V.21 and Bell 103: 20 kbits ITU-T V.22 *bis* and Bell 212A: 100 kbits	
Telemetry	ITU-T V.23	600 and 1200	ANS (optional)	BER	20 kbits	
Text telephone	ITU-T V.18 (Note 7) DTMF (Note 9) ITU-T V.21 110 half duplex low channel (Note 8) ("EDT") ITU-T V.21 300 bits/s full duplex TIA 825A (Baudot) ITU-T V.23	ITU-T V.18 300 bit/s DTMF rate n/a ITU-T V.21: 110 half duplex low channel ("EDT") ITU-T V.21: 300 full duplex TIA 825A: 45.45 and 50 ITU-T V.23 1200/75	ITU-T V.18 and ITU-T V.21: ANS (optional) DTMF, TIA 825A and ITU-T V.23: None	Character error rate	ITU-T V.18, ITU-T V.21, TIA 825A and ITU-T V.23: 1000 characters DTMF: 300 characters	
Security/Alarm system	SIA DC-02 SIA DC-05	Pulse and DTMF: 10 to 33 pulses/s Bell 103: 300	Various handshake tones	Character error rate	300 characters	The SIA standards specify a number of different pulse, DTMF and FSK formats for use by security/alarm systems. In addition, SIA DC-03 specifies a standard protocol for use with Bell 103 systems.

Table 4 – Minimum set of modems to be used for Test 14

Application	Modulation type	Data rate (bits/s)	Answer tone	Performance criteria	Minimum volume of data	Comments
Leased line dial back-up	Bell 208 A/B Bell 202	Bell 208 A/B: 4800 Bell 202: 1200	Unknown	BER	100 kbits	

Key to abbreviations:

ANS 2100 Hz Answer Tone as defined in ITU-T V.25

BER Bit Error Rate

NOTE 1 – Where possible, synchronous testing should be performed, since this more accurately detects bit errors due to transmission problems. The exceptions to this are ITU-T V.21 and Bell 103, which are usually tested in asynchronous mode. In general, error correction should be switched off.

NOTE 2 – Point of sale modems may use standard and proprietary versions of modulations.

NOTE 3 – Security and Alarm system modems may use standard and proprietary versions of modulations.

NOTE 4 – SIA DC-02 Security Industry Association (SIA) General Protocols – Available from www.siaonline.org

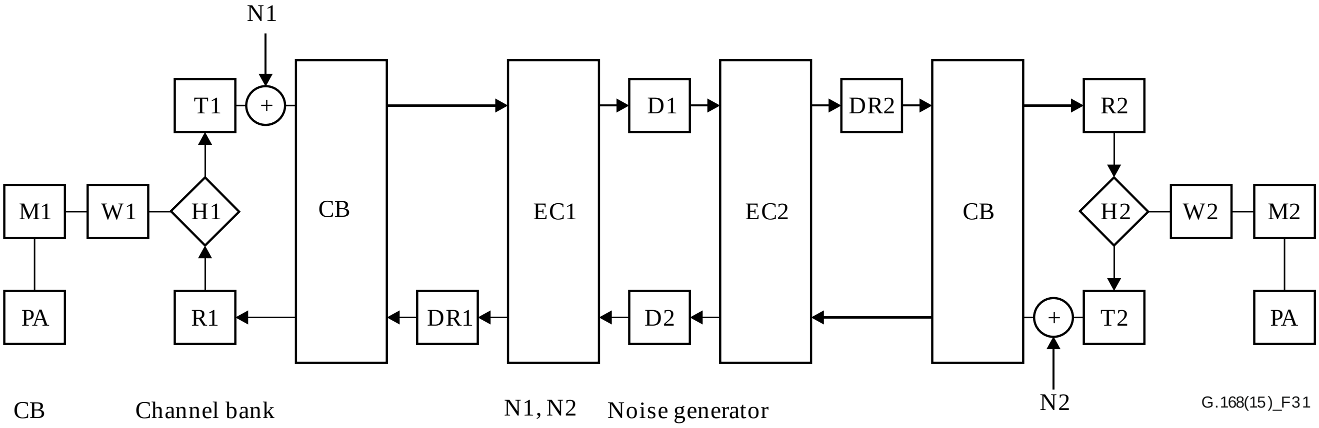
NOTE 5 – SIA DC-03 Security Industry Association (SIA) SIA Format – Available from www.siaonline.org

NOTE 6 – SIA DC-05 Security Industry Association (SIA) Ademco Contact ID Standard – Available from www.siaonline.org

NOTE 7 – ITU-T V.18 is an automoding procedure using ITU-T V.8 and all other textphone modulations. ITU-T V.8 uses ITU-T V.21 300 bits/s one channel at a time both before and after ANSAM. Full information is found in ITU-T V.18 and ITU-T V.8.

NOTE 8 – ITU-T V.21 110 bits/s half duplex is used for EDT, a legacy sub-mode of ITU-T V.18. Only lower channel is used. Pauses between text transmissions are silent or used for voice. EDT is used for text telephony in Germany, Italy, Spain and Switzerland.

NOTE 9 – DTMF is for textphone based on DTMF coded text.



CB	Channel bank
D1, D2	Long-haul delay
DR1, DR2	Echo path delay
EC1, EC2	Echo canceller
H1, H2	Hybrid
M1, M2	Modem under test
N1, N2	Noise generator
PA	Protocol analyzer
R1, R2	Receive attenuator
T1, T2	Transmit attenuator
W1, W2	Wire simulator

Figure 31 – Test No. 14 configuration

6.4.2.16 Test No. 15 – PCM offset test (Optional)

This test is meant to ensure that the echo canceller will operate properly in the presence of PCM offset in the speech signal applied to S_{in} or in the speech signal applied to R_{in}.

PCM offset is an unwanted, fixed DC level in the S_{in} signal relative to the R_{out} signal, or in the R_{in} signal itself. This can be caused by some network equipment, such as A/D converters, and can result in degraded performance of echo canceller and other speech processing equipment. The methods consist of performing Test 2B, but with a DC offset applied at S_{in} or R_{in}.

Part 1

Apply a PCM offset error to the S_{in} signal relative to the R_{in} signal as indicated in Figure 32. Depending on the preference by the users, the PCM offset can be injected at point A – the linear domain, or at point B – the A/μ law domain.

NOTE 1 – Offset injection at point B is not a linear operation.

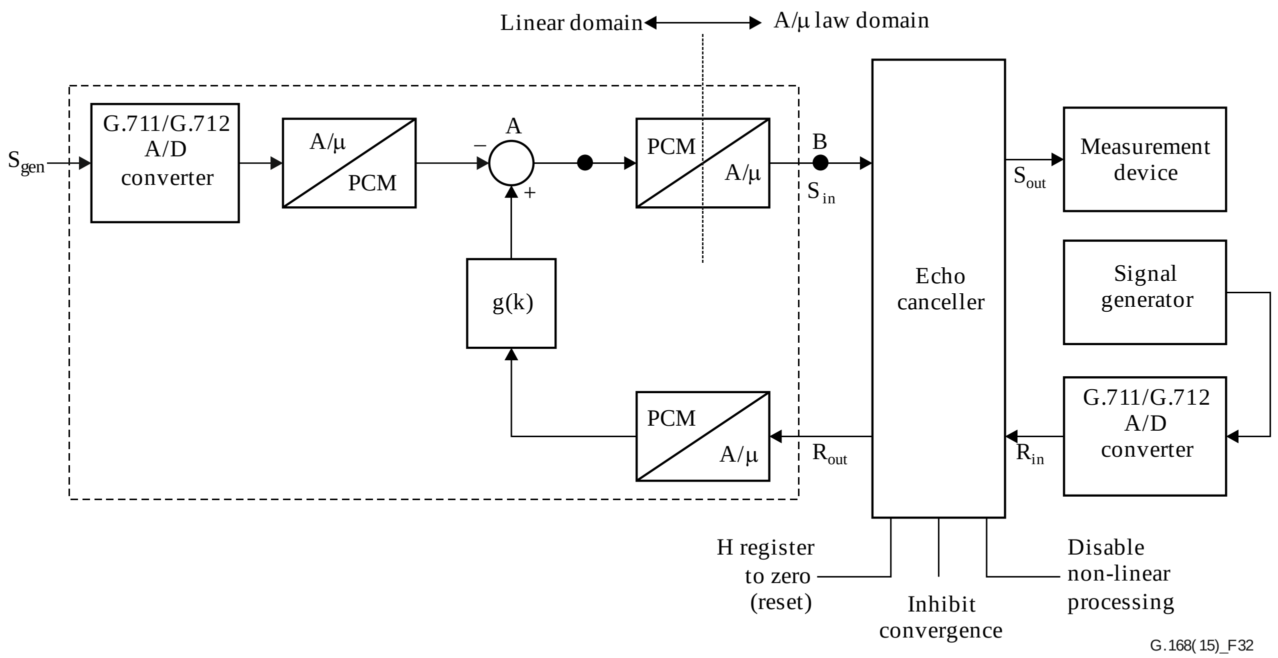


Figure 32 – Test configuration for S_{in} PCM offset test

Part 2

Apply a PCM offset error to the R_{in} signal as indicated in Figure 33. Depending on the preference by the users, the PCM offset can be injected at point A – the linear domain, or at point B – the A/μ law domain.

NOTE 2 – Offset injection at point B is not a linear operation.

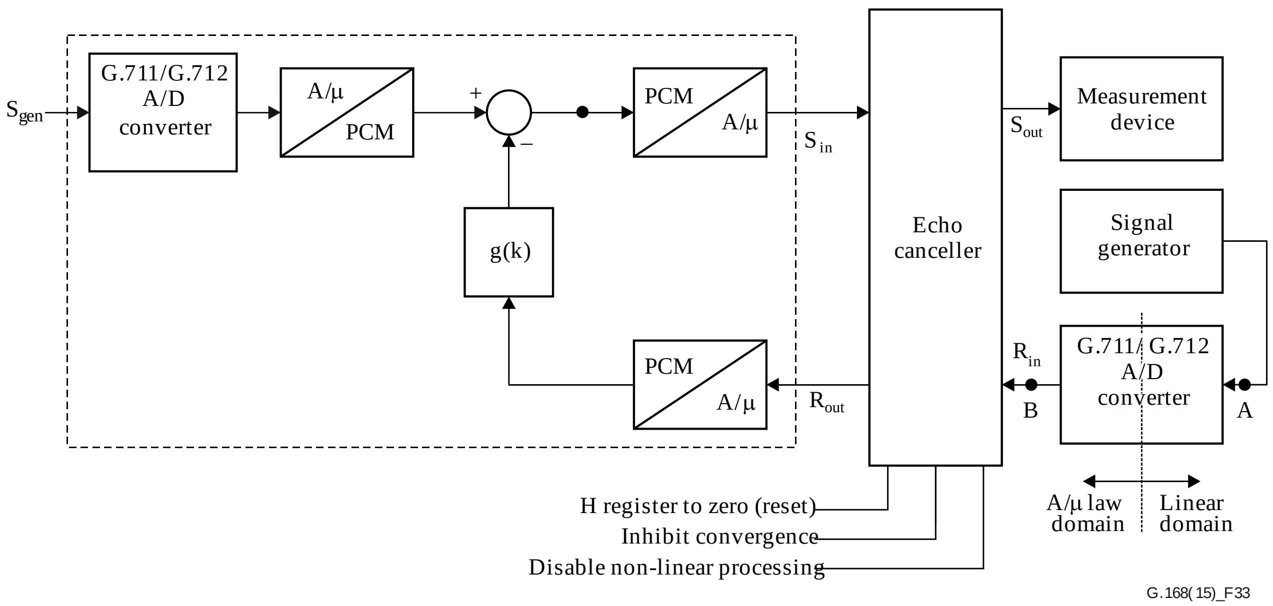


Figure 33 – Test configuration for Rin PCM offset test

The level at R_{in} in the test requirement is $L_{Rin,act}$. It is the signal level measure using the RMS method over the active portion of the CSS only (i.e., excluding the pause of the CSS) as described in clause 6.4.1.2.

Requirement

For both Part 1 and Part 2 above, with the H register initially reset, or alternatively, with an open echo path resulting in $S_{in} = 0$, the H register content is converged to 0, and then the adaptation is inhibited. Adaptation is then enabled at least 200 ms before a DC offset at a level of –40 dBm0 is injected at either point A or point B in Figure 32 (for the Part 1 test) or Figure 33 (for the Part 2 test). The first CSS burst is applied to R_{in} at the same instant the DC offset is applied. With the NLP disabled, for all values $L_{Rin,act} \geq -30$ dBm0 and ≤ 0 dBm0 and for all values of ERL $\geq$ 6 dB and echo path delay, $t_d \leq \Delta$ ms, the loss $L_{Rin,act} - L_{RES}$ should be greater than or equal to that shown in Figure 12a. The level at S_{out} is measured using a meter conforming to the characteristics of clause 6.4.1.2.1. In addition, no peaks (see clause 6.4.1.2.2) are allowed that exceed 5 dB above the requirements in Figure 11.

The variable $L_{Rin,act} - L_{RES}$ in Figure 12a may be replaced by the variable $L_{Sin} - L_{Sout}$ + ERL, where L_{Sin} and L_{Sout} are the levels of S_{in} and S_{out} respectively. The signal levels L_{Sin} and L_{Sout} are measured using the measurement device in clause 6.4.1.2.1, and should be synchronized. The ERL is the value chosen in the test. This method may also be used to observe convergence as a continuous plot over time.

NOTE 3 – The method stated in the preceding paragraph takes into account any dispersion in the echo path, but does not take into account any dispersion present between the S_{in} and S_{out} ports of the echo canceller.

NOTE 4 – As the level measurement device includes a band-pass filter, this test does not imply that the echo canceller must remove the DC offset.

NOTE 5 – This test includes a non-linear echo path. See clauses I.5.3, I.5.4 and [ITU-T G.161] for further discussion on this issue.

6.4.2.17 Test No. 16 – Dual Tone Multi-Frequency (DTMF) transparency test (optional)

This test has the objective of verifying that the echo canceller will transmit, without excessive distortions, narrow-band signals (specifically, DTMF digits) applied at S_{gen}, during the presence of the dial tone or human voice (emulated by the CSS for testing purposes) which is applied at R_{in}. The echo canceller's adaptation is enabled throughout the entire test. The NLP should be enabled.

Both parts of this test are optional, depending on the nature of the echo canceller deployment. Details of when to perform each test are given in the individual test descriptions.

6.4.2.17.1 Test 16A: DTMF transparency in the presence of a dial tone (optional)

The method consists of enabling adaptation prior to applying the dial tone. After the dial tone is applied, a DTMF digit sequence is applied at S_{gen}. The sequence is to continue so that at least one DTMF digit is present after the dial tone is terminated.

The optional nature of this test is motivated by the fact that the echo cancellers that are controlled by central office or gateway actions may not need to conform to this test.

Test method

The echo canceller is enabled. NLP is enabled. The test is to be performed for all values of ERL $\geq$ 6 dB. Once chosen, the echo path (along with its ERL as well as t_r) remains the same throughout the test. The dial tone levels are to be within the range of [–15 dBm0, –5 dBm0], see [ITU-T E.180]. The DTMF levels and the timing are to be within the range specified in Table A.1 of [ITU-T Q.24]. The timing of the dial tone and DTMF tones has to be such that the first DTMF digit applied at S_{gen} has to be delayed with respect to the start of the dial tone by T_{ad} ($T_{ad} > 400$ ms) (see Figure 34). The dial tone shall last for $T_{dto} > 40$ ms (40 ms represents a typical DTMF detection time) beyond the start of the first DTMF digit. At least one digit should be applied after the dial tone is removed (see Figures 35 and 36).

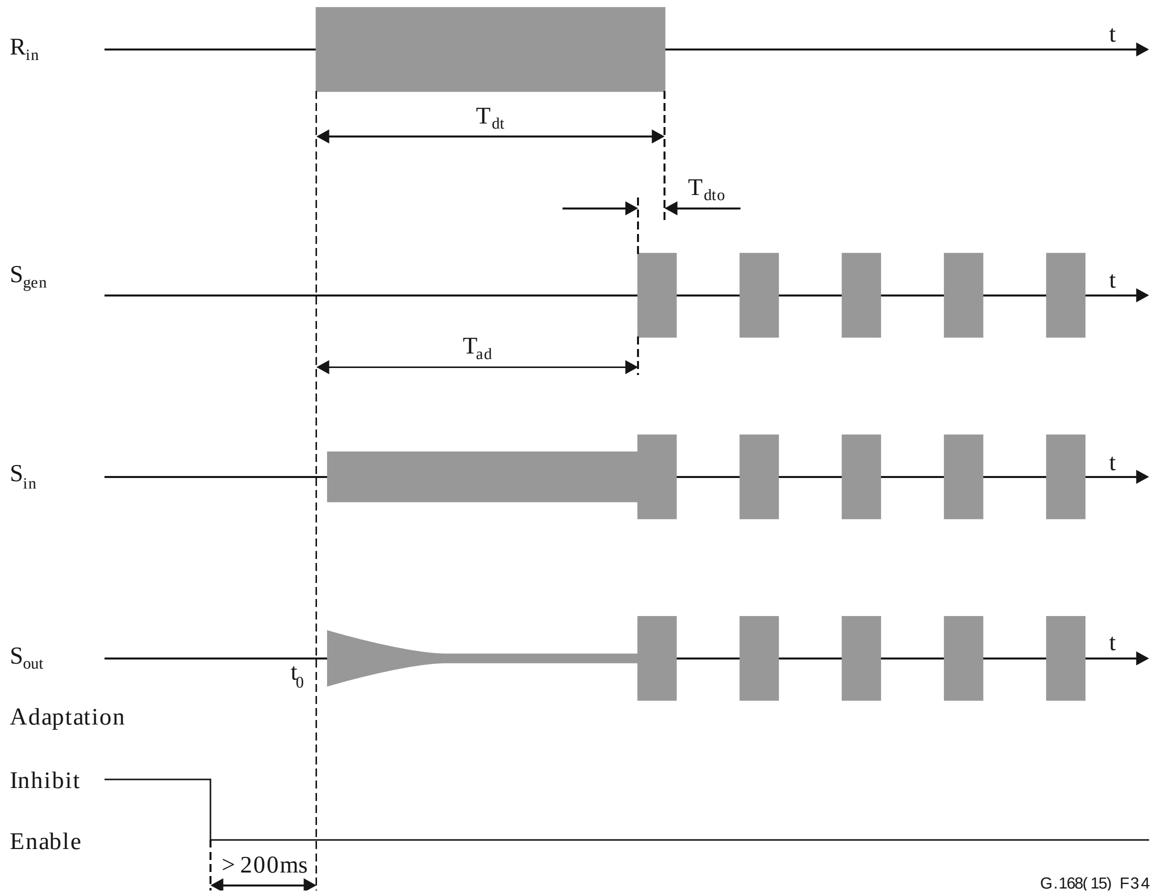


Figure 34 – Illustration of requirements for DTMF transparency in the presence of dial tone

Requirements

Specific requirements for DTMF digit/sequence selection

The DTMF test sequence should consist of 5 different DTMF digits selected from the set of {0,...,9, #, *, A, B, C, D}, see [ITU-T Q.23]. The minimum sub-set of test conditions is proposed to include 16 DTMF sequences that are to be tested. The sequences are built by placing all possible digits at the front of the sequences while the remaining four digits in each sequence should be taken from the full set such that every digit in the sequence is different. Table 5 defines the sequences.

Table 5 – DTMF test sequences (as per minimum sub-set of test conditions)

Sequence No.	1st digit	2nd digit	3rd digit	4th digit	5th digit
1	0	1	2	3	4
2	1	2	3	4	5
3	2	3	4	5	6
4	3	4	5	6	7
5	4	5	6	7	8
6	5	6	7	8	9
7	6	7	8	9	#
8	7	8	9	#	*

Table 5 – DTMF test sequences (as per minimum sub-set of test conditions)

Sequence No.	1st digit	2nd digit	3rd digit	4th digit	5th digit
9	8	9	#	*	A
10	9	#	*	A	B
11	#	*	A	B	C
12	*	A	B	C	D
13	A	B	C	D	0
14	B	C	D	0	1
15	C	D	0	1	2
16	D	0	1	2	3

Specific requirements for DTMF versus dial tone timing

Two versions of timing constitute the minimum sub-set of test conditions.

The first version of timing includes the full overlap of four DTMF digits with the dial tone (see Figure 35). The duration T_{dt} of the dial tone is defined as 1000 ms. The first DTMF digit starts at T_{ad} = 450 ms after the start of the dial tone. The length of each DTMF digit is 50 ms. The interval between the DTMF digits is 100 ms.

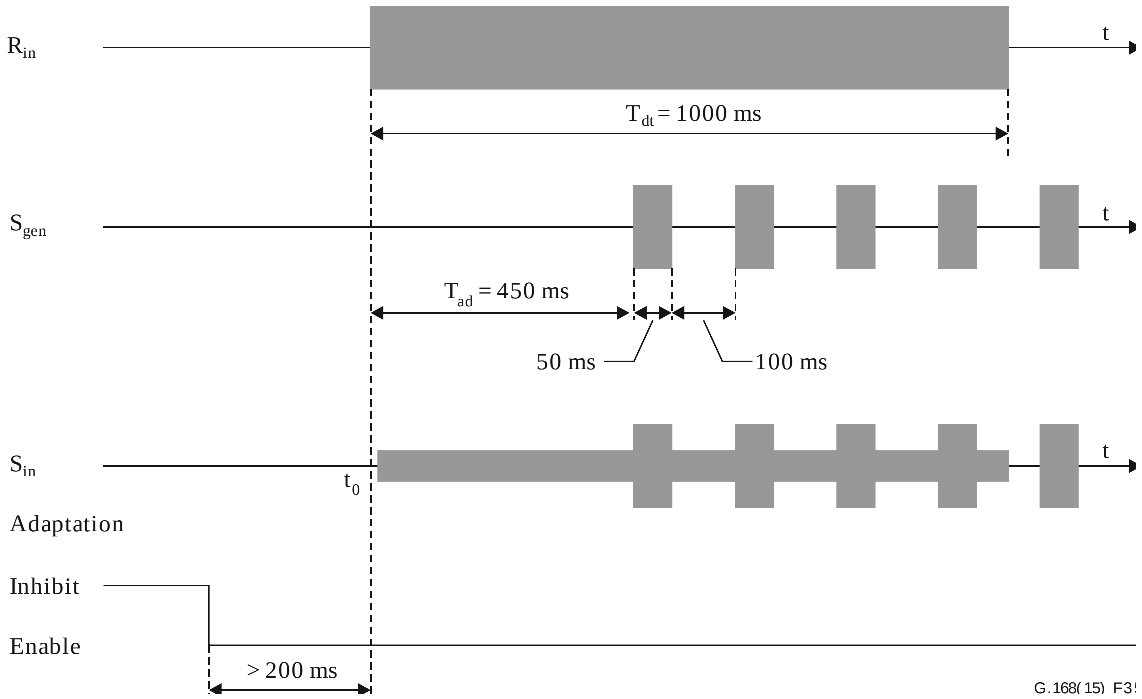


Figure 35 – Test timing version 1: "fast" DTMF reaction and a "slow" removal of the dial tone

The second version of timing includes a partial overlap of the first DTMF digit with the dial tone (see Figure 36). The duration T_{dt} of the dial tone is defined as 2030 ms. The first DTMF digit starts at

T_{ad} = 2000 ms after the start of the dial tone. The length of each DTMF digit is 50 ms. The interval between the DTMF digits is 100 ms.

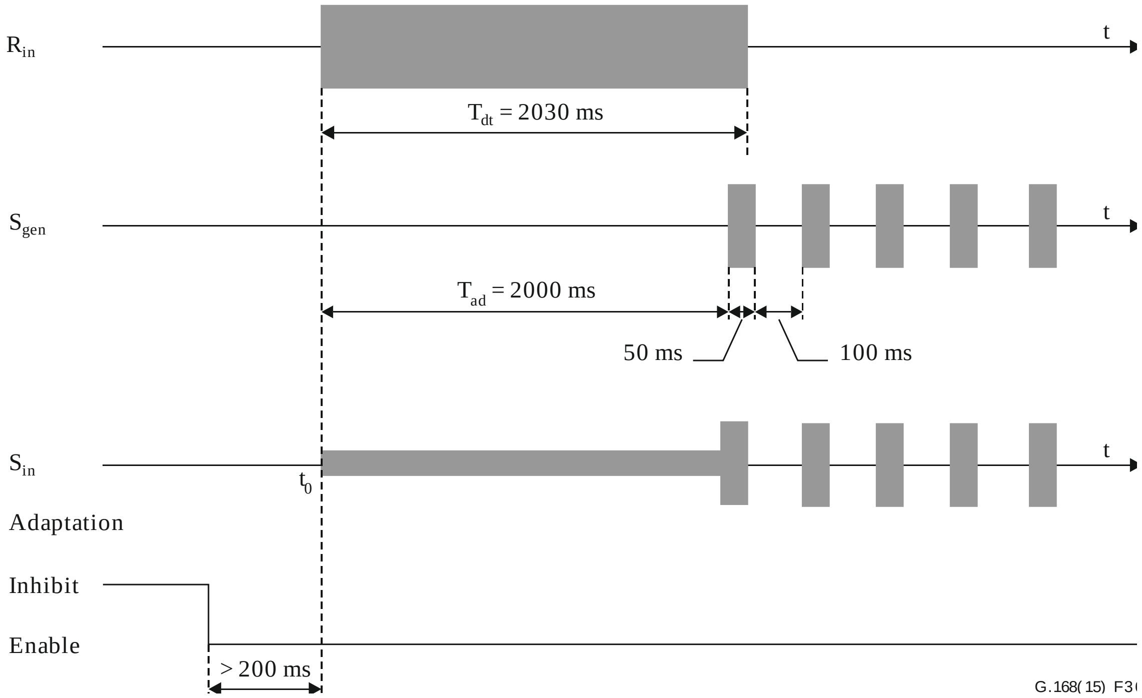


Figure 36 – Test timing version 2: a "typical" DTMF reaction and an almost instantaneous removal of the dial tone upon detection of the first DTMF digit by the system

Specific requirements for the levels of DTMF digits and dial tone as well as the echo path models and ERL values

While the proposed requirements for this test include all valid dial tone and DTMF tone parameters combined with all ERL ≥ 6 dB and all echo path models, the minimum sub-set of test conditions consists of:

– Dial tone levels: range of [–15 dBm0, –5 dBm0] see [ITU-T E.180], with steps of 5 dB;

– DTMF levels: range of [–27 dBm0, –6 dBm0] see [ITU-T Q.24], Table A.1, with steps of 3 dB;

– ERL values: range of [6 dB, 30 dB], with steps of 3 dB;

– Echo path model: ITU-T G.168 model m_1;

– Pure delay: $t_r = 0$

NOTE – The minimum sub-set of test conditions is limited to the normative frequencies (DTMF and dial tone) and no twist (DTMF). No DTMF distortions and/or electrical noise simulation make part of the minimum sub-set of test conditions.

Specific requirements for DTMF at S_{out}: DTMF detection scores

The method of determining whether the echo canceller distorts the DTMF digits beyond recognition uses a DTMF tone detector compliant with [ITU-T Q.24]. Signals S_{out} (for echo canceller enabled) and S_{out} (for echo canceller disabled), as generated using the test vectors defined in Table 5, should be applied to the DTMF tone detector. The DTMF detection score for the case with echo canceller enabled should not be lower than for the case of echo canceller disabled.

6.4.2.17.2 Test 16B: DTMF transparency in the presence of speech (optional)

The test method includes enabling adaptation prior to applying the CSS. After the CSS is applied, a DTMF digit sequence is applied at S_{gen}. The sequence is to continue so that at least one DTMF digit is present after the CSS is terminated.

The optional nature of this test is motivated by the fact that there are TDM switches/gateway architectures where a DTMF detector component follows the echo canceller component (i.e., where the S_{out} signal becomes the input signal to the DTMF detector). In such a case, this test is necessary. Otherwise (i.e., where the DTMF detectors receive the S_{in} signal), this test is not required.

Test method

The echo canceller is enabled. NLP is enabled. Once chosen, the echo path (including its ERL as well as the t_r) remains the same throughout the test. The test is to be performed for all values of ERL $\geq$ 6 dB. The CSS levels are to be within the range of [–15 dBm0, –5 dBm0]. The DTMF levels and the timing are to be within the range specified in Table A.1 of [ITU-T Q.24]. The timing of the CSS and DTMF tones has to be such that the first DTMF digit applied at S_{gen} has to be delayed with respect to the start of the CSS by T_{ad} (T_{ad} > 1500 ms) (see Figure 37). At least one digit (but not more than two digits) shall be applied following the time when the CSS is removed.

The convergence characteristic during the time the CSS is present at R_{in} and S_{gen} is silence (the time duration of T_{ad}) should conform to Test No. 2A(a) requirements, as illustrated in Figure 10a with pure delay $t_r = 0$.

Requirements

Specific requirements for DTMF digit/sequence selection

The DTMF test sequence should consist of DTMF digits selected from the set of {0,...,9, #, *, A, B, C, D}. The minimum sub-set of test conditions is proposed to include any 4 of 16 DTMF sequences that cover the entire DTMF digit set (for example, Sequences No. 1, 6, 11 and 16). Note that the sequences are built by placing all possible digits at the front of the sequences, while the remaining four digits in each sequence should be taken from the full set, such that every digit in the sequence is different. Table 6 defines the sequences.

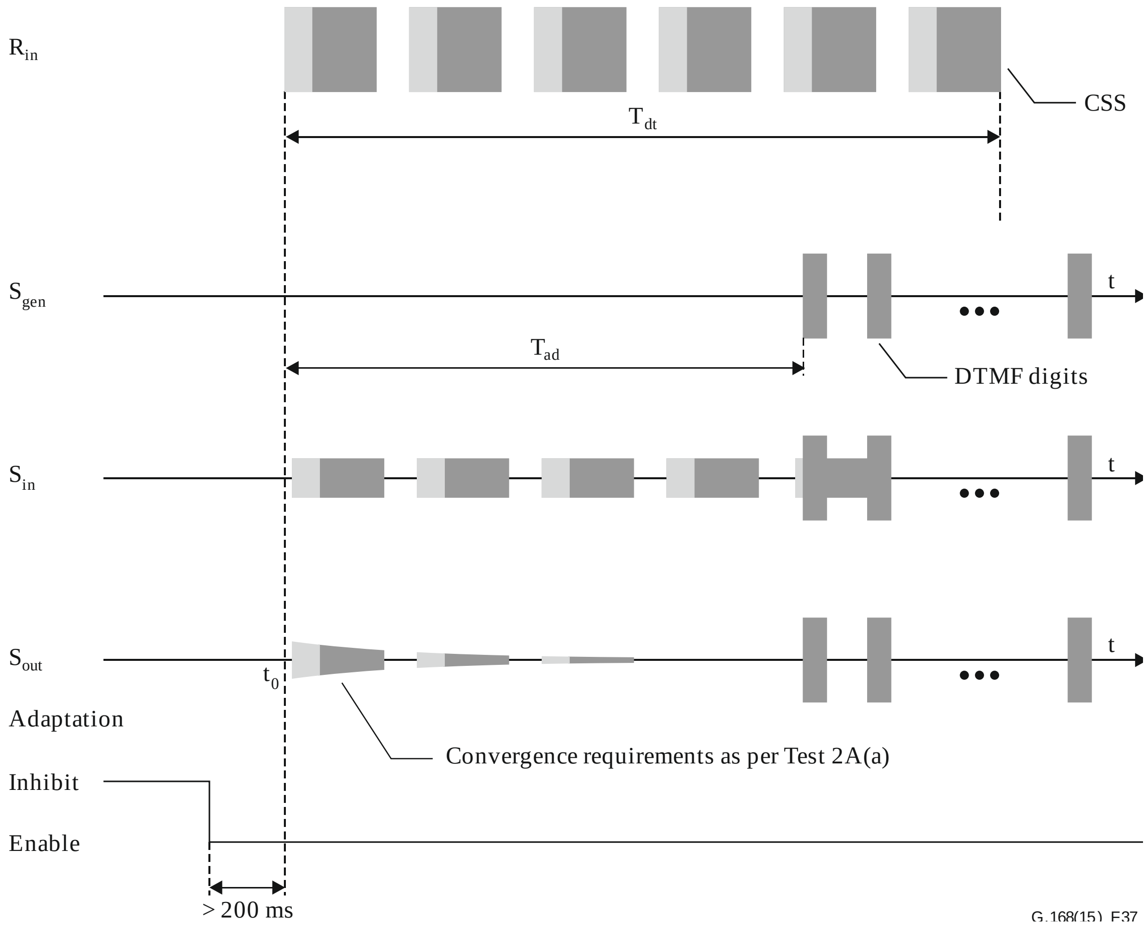


Figure 37 – Illustration of requirements for DTMF transparency in the presence of CSS

Specific requirements for DTMF versus announcement/speech (CSS) timing

The timing relationship specified here constitutes a minimum sub-set of test conditions.

Mandatory: The first five digits in Table 6 (1st digit to 5th digit) are used. The timing relationship includes at least a partial overlap of four DTMF digits with the CSS (see Figure 37), while the fifth digit does not overlap the CSS. The duration T_{dt} of the CSS is defined as 2000 ms. The first DTMF digit starts at T_{ad} = 1500 ms after the start of the CSS. The length of each DTMF digit is 50 ms. The interval between the DTMF digits is 100 ms.

Optional: All ten DTMF digits in Table 6 are used in order to provide more adequate means for statistical processing of the test results. In the timing relationship, the duration T_{dt} of the CSS is changed to 2700 ms so that the last two DTMF digits do not overlap with CSS pulses. Other timing details remain unchanged per the five DTMF digits case.

Specific requirements for the levels of DTMF digits and CSS as well as the echo path models, ERL and pure delay values

While the proposed requirements for test setup include all valid CSS and DTMF tone parameters combined with all ERL ≥ 6 dB and all echo path models, the minimum sub-set of test conditions consists of:

CSS levels: range of [–15 dBm0, –5 dBm0], with steps of 5 dB;

DTMF levels: range of [–27 dBm0, –6 dBm0], see [ITU-T Q.24], with steps of 3 dB;

ERL values: range of [6 dB, 30 dB], with steps of 3 dB;

Echo path model: ITU-T G.168 m_1;

Pure delay: $t_r = 0$.

NOTE – The minimum sub-set of test conditions is limited to the normative frequencies and twist (DTMF). No DTMF distortions and/or electrical noise simulation make part of the minimum sub-set of test conditions.

Specific requirements for DTMF at S_{out}: DTMF detection scores

The method of determining whether the echo canceller distorts the DTMF digits beyond recognition uses a DTMF tone detector compliant with [ITU-T Q.24]. Signals S_{out} (for echo canceller enabled) and S_{out} (for echo canceller disabled), as generated using the test sequences defined in Table 6, should be applied to the DTMF tone detector. The DTMF detection score for the case with echo canceller enabled should not be lower than for the case of echo canceller disabled.

Table 6 – DTMF test sequences (as per minimum sub-set of test conditions)

Set No.	1st digit	2nd digit	3rd digit	4th digit	5th digit	6th digit	7th digit	8th digit	9th digit	10th digit
1	0	1	2	3	4	5	6	7	8	9
2	1	2	3	4	5	6	7	8	9	#
3	2	3	4	5	6	7	8	9	#	*
4	3	4	5	6	7	8	9	#	*	A
5	4	5	6	7	8	9	#	*	A	B
6	5	6	7	8	9	#	*	A	B	C
7	6	7	8	9	#	*	A	B	C	D
8	7	8	9	#	*	A	B	C	D	0
9	8	9	#	*	A	B	C	D	0	1
10	9	#	*	A	B	C	D	0	1	2
11	#	*	A	B	C	D	0	1	2	3
12	*	A	B	C	D	0	1	2	3	4
13	A	B	C	D	0	1	2	3	4	5
14	B	C	D	0	1	2	3	4	5	6
15	C	D	0	1	2	3	4	5	6	7
16	D	0	1	2	3	4	5	6	7	8

7 Characteristics of an echo canceller tone disabler

7.1 General

The echo cancellers covered by this Recommendation should be equipped with a tone detector that conforms to this clause. This tone detector should disable the echo canceller only upon detection of a signal which consists of a 2100 Hz tone with periodic phase reversals inserted in that tone, and not disable with any other in-band signal, e.g., speech or a 2100 Hz tone without phase reversals. The tone disabler should detect and respond to a disabling signal which may be present in either the send or the receive path.

To improve the operation of the echo canceller for fax signals and low-speed voiceband data, it may be beneficial for some echo cancellers to disable the NLP for such calls. In this case, the echo canceller may optionally detect any 2100 Hz tone without phase reversals. If 2100 Hz tone without phase reversal is detected, the echo canceller shall remain enabled, and the NLP may optionally be disabled. The frequency characteristics of the tone detector are given in Figure 38. The tone disabler characteristics as specified in clauses 7.4 through 7.9 also apply for this NLP disabling detector. Note that if the 2100 Hz tone contains phase reversals, then the echo canceller shall be disabled as defined elsewhere in this clause.

The term disabled in this clause refers to a condition in which the echo canceller is configured in such a way as to no longer modify the signals which pass through it in either direction. Under this condition, no echo estimate is subtracted from the send path, the non-linear processor is made transparent, and the delay through the echo canceller still meets the conditions specified in clause 6.4.1. However, no relationship between the circuit conditions before and after disabling should be assumed. The impulse response stored in the echo canceller prior to convergence (and prior to the disabling tone being sent) is arbitrary. This can lead to apparent additional echo paths which, in some echo canceller implementations, remain unchanged until the disabling tone is recognized. Also note that echo suppressors could be on the same circuit and there is no specified relationship between their delay in the enabled and disabled states. In spite of the above, it is possible, for example, to measure the round-trip delay of a circuit with the disabling tone but the trailing edge of the tone burst should be used and sufficient time for all devices to be disabled should be allotted before terminating the disabling tone and starting the timing.

It should be noted that the echo canceller should provide 64 kbit/s bit-sequence integrity when disabled.

7.2 Detector characteristics

The tone detector shall detect a tone in the frequency range of 2100 Hz ± 21 Hz (see [ITU-T V.21]). The detection channel bandwidth should be chosen wide enough to encompass this tone (and possibly other disabling tones used within national networks). At the same time, the detection channel bandwidth should be such that, in conjunction with guard action and timing, adequate protection is provided against false operation of the detector by speech signals. The detector channel sensitivity (threshold level) should be such that the detector will operate on the lowest expected power of the disabling tone. The band characteristics shown in Figure 38 will permit disabling by the 2100 Hz disabling tone as well as others used in North America. The figure indicates that in the frequency band 2079 Hz to 2121 Hz detection **must** be possible whilst in the band 1900 Hz to 2350 Hz detection **may** be possible.

Providing that only the recommended 2100 Hz disabling tone is used internationally, interference with signalling equipment will be avoided.

The dynamic range of the detector should be consistent with the input levels as specified in [ITU-T V.2] with allowances for variation introduced by the public switched telephone network.

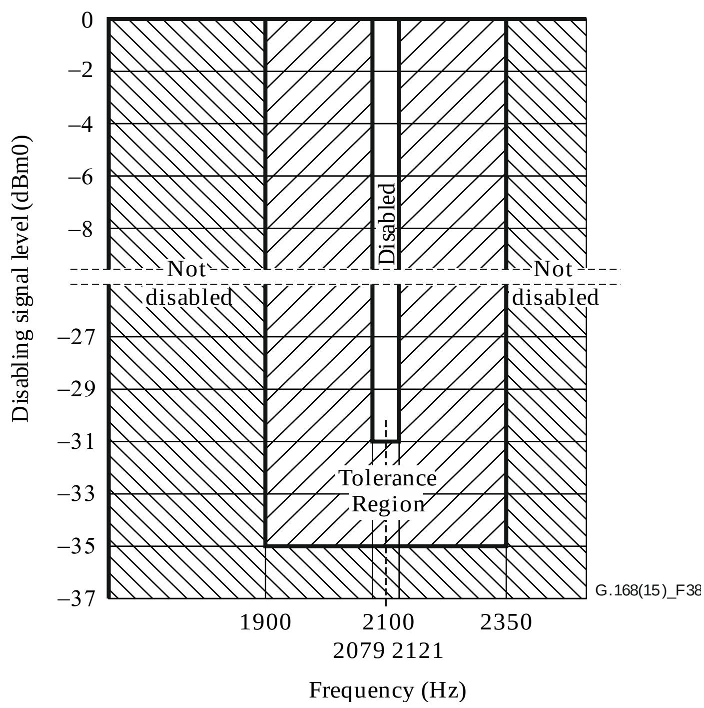


Figure 38 – Required disabling band characteristics

7.2.1 Phase reversal detection

The echo canceller tone disabler requires the detection of a 2100 Hz tone with periodic phase reversals which occur every 450 ± 25 ms. The characteristics of the transmitted signal are defined in [ITU-T V.25] and [ITU-T V.8]. Phase variations in the range of 180° ± 25° should be detected while those in the range of 0° ± 110° should not be detected. This restriction is to minimize the probability of false disabling of the echo canceller due to speech currents and network-induced phase changes. The ±110° range represents the approximate phase shift caused by a single frame slip in a PCM system.

7.3 Guardband characteristics

Energy in the voice band, excluding the disable band, must be used to oppose disabling so that speech will not falsely operate the tone disabler. The guard band should be wide enough and with a sensitivity such that the speech energy outside the disabling band is utilized. The sensitivity and shape of the guard band must not be such that the maximum idle or busy circuit noise will prevent disabling. In the requirement, white noise is used to simulate speech and circuit noise. Thus, the requirement follows:

Given that white noise (in a band of approximately 300-3400 Hz) is applied to the tone disabler simultaneously with a 2100 Hz signal, the 2100 Hz signal is applied at a level 3 dB above the midband disabler threshold level. The white noise energy level required to inhibit disabling should be no greater than the level of the 2100 Hz signal and no less than a level 5 dB below the level of the 2100 Hz signal. As the level of the 2100 Hz signal is increased over the range of levels to 30 dB above the midband disabler threshold level, the white noise energy level required to inhibit disabling should always be less than the 2100 Hz signal level. These requirements, together with the noise tolerance requirements given in clause 7.3.1 are illustrated in Figure 39.

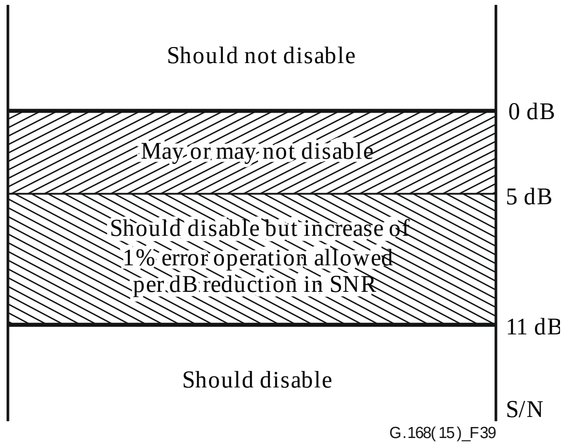


Figure 39 – Guardband and noise tolerance requirements

NOTE – The possibility of interference during the phase reversal detection period has been taken into account. One potential source of interference is the presence of calling tone as specified in [ITU-T V.25]. If the calling tone interferes with the detection of the phase reversal, the entire disabling detection sequence is restarted, but only one time. [ITU-T V.25] ensures at least one second of quiet time between calling tone bursts.

7.3.1 Noise tolerance

The detector should operate correctly with white noise less than or equal to 11 dB below the level of the 2100 Hz signal. No definitive guidelines can be given for the range between 5 and 11 dB because of the variations in the test equipment used. In particular, performance may vary with the peak-to-average ratio of the noise generator used. As a general guideline, however, the percentage of correct operation (detection of phase variations of 180° ± 25° and non-detection of phase variations of 0° ± 110°) should fall by no more than 1% for each dB reduction in the signal-to-noise ratio below 11 dB. It is noted that it is possible to design a detector capable of operating correctly at 5 dB signal-to-noise ratio.

7.4 Holding-band characteristics

The tone detector, after disabling either the NLP or the echo canceller, should hold the NLP or echo canceller in the disabled state for tones in a range of frequencies specified below. The release sensitivity should be sufficient to maintain disabling for the lowest level data signals expected, but should be such that the detector will release for the maximum idle or busy circuit noise. Thus the requirement follows:

The tone detector should hold the NLP or echo canceller in the disabled state for any single-frequency sinusoid in the band from 390-700 Hz having a level of –27 dBm0 or greater, and from 700-3000 Hz having a level of –31 dBm0 or greater. The tone disabler should release for any signal in the band from 200-3400 Hz having a level of –36 dBm0 or less.

7.5 Operate time

The operate time should be sufficiently long to provide immunity from false operation due to voice signals, but not so long as to needlessly extend the time to disable. The tone disabler is required to operate within one second of the receipt of the disabling signal. The one second operate time permits the detection of the 2100 Hz tone and ensures that two-phase reversals will occur.

7.6 False operation due to speech currents

It is desirable that the tone disabler should rarely operate falsely on speech. To this end, a reasonable objective is that, for an echo canceller installed on a working circuit, usual speech currents should not on the average cause more than 10 false operations during 100 hours of speech. In addition to the talk-off protection supplied by the disabling channel bandwidth, by guard band operation and by the operate time, talk-off protection can be supplied by recycling. That is, if speech which simulates the disabling signal is interrupted because of inter-syllabic periods, before disabling has taken place, the operate timing mechanism should reset. However, momentary absence or change of level in a true disabling signal should not reset the timing.

7.7 False operation due to data signals

It is desirable that the tone disabler should rarely operate falsely on data signals from data sets that would be adversely affected by disabling the echo canceller. To this end, a reasonable objective is that, for an echo canceller installed on a working circuit, usual data signals from such data sets should not, on the average, cause more than 10 false operations during 100 hours of data transmissions.

To this end, in the reference tone disabler described in Annex B of [ITU-T G.165], which meets the above requirements, the tone disabler circuitry becomes inoperative if one second of clear (i.e., no phase reversals or other interference) 2100 Hz tone is detected. The detector circuit remains inoperative during the data transmission and only becomes operative again 250 ± 150 ms after a signal in the holding band falls at least 3 dB below the maximum holding sensitivity. Thus the possibility of inadvertent disabling of the echo canceller during facsimile or low speed (< 9.6 kbit/s) voiceband data transmission is minimized.

7.8 Release time

The disabler should not release for signal drop-outs less than the ITU-T recommended value of 100 ms. To cause a minimum of impairment upon accidental speech disabling, it should release within 250 ± 150 ms after a signal in the holding band falls at least 3 dB below the maximum holding sensitivity in both directions of signal transmission.

7.9 Other considerations

Both the echo of the disabling tone and the echo of the calling tone may disturb the detection of the echo canceller disabling tone. As such, it is not recommended to add the receive and transmit signal inputs together to form an input to a single detector.

Careful attention should be given to the number of phase reversals required for detection of the disabling tone. Some Administrations favour relying on 1 to improve the probability of detection even in the presence of slips, impulse noise, and low signal-to-noise ratio. Other Administrations favour relying on 2 to improve the probability of correctly distinguishing between non-phase-reversed and phase-reversed 2100 Hz tones, and to reduce the likelihood of false triggering of the tone disabler by speech or data signals.

8 Non-linear processors (NLPs) for use in echo cancellers

8.1 Scope

For the purpose of this Recommendation, the term "NLP" is intended to mean only those devices which fall within the definition given in clause 3.21 and which have been proven to be effective in echo cancellers. It is possible to implement such NLPs in a number of ways (centre clippers being just one example), with fixed or adaptive operating features, but no recommendation is made for any particular implementation. General principles and guidelines are given in clause 8.2. More detailed and concrete information requires reference to specific implementations. This is done in Annex B for the particular case of a "reference NLP". The use of this term denotes an implementation given for guidance and illustration only. It does not exclude other implementations nor does it imply that the reference NLP is necessarily the most appropriate realization on any technical, operational or economic grounds.

8.2 General principles and guidelines

8.2.1 Function

8.2.1.1 General

The NLP is located in the send path between the output of the subtractor and the send-out port of the echo canceller. Conceptually, it is a device which blocks low level signals and passes high-level signals. Its function is to further reduce the residual echo level (L_{RES} as defined in clause 3.25) which remains after imperfect cancellation of the circuit echo so that the necessary low returned echo level (L_{RET} as defined in clause 3.26) can be achieved.

8.2.1.2 Network performance

Imperfect cancellation can occur because echo cancellers which conform to this Recommendation may not be capable of adequately modelling echo paths which generate significant levels of non-linear distortion (see [ITU-T G.161]). Such distortion can occur, for example, in networks conforming to Recommendation [b-ITU-T G.113] in which up to five pairs of PCM codecs (conforming to [ITU-T G.712]) are permitted in an echo path. The accumulated quantization distortion from these codecs may prevent an echo canceller from achieving the necessary L_{RET} by using linear cancellation techniques alone. It is therefore recommended that all echo cancellers capable only of modelling the linear components of echo paths but intended for general network use should incorporate suitable NLPs. In specific network environments with low delay or high ERL, it may be possible to disable the NLP in an echo canceller with a sufficiently high ERLE. This may result in higher overall speech quality, as NLPs sometimes cause speech degradation.

8.2.1.3 Limitations

This use of NLPs represents a compromise in the circuit transparency which would be possible by an echo canceller which could achieve the necessary L_{RET} by using only modelling and cancellation techniques. Ideally, the non-linear processor should not cause distortion of cancelled end speech. In practical devices, it may not be possible to sufficiently approach this ideal. In this case, it is recommended that NLPs should not be active under double talk or cancelled end single speech conditions. From this it follows that excessive dependence should not be placed on the NLP and that L_{RES} should be low enough to prevent objectionable echo under double talk conditions.

8.2.1.4 Data transmission

NLPs may affect the transmission of data through an enabled echo canceller. This is under study.

8.2.2 Suppression threshold

8.2.2.1 General

The suppression threshold level (T_{SUP}) of an NLP is expressed in dBm0 and is equal to the highest level of a sine-wave signal at a given moment that is just suppressed. Either fixed or adaptive suppression threshold levels may be used.

8.2.2.2 Fixed suppression threshold

With a fixed suppression threshold level, the appropriate level to use will depend upon the cancellation achieved and the statistics of speech levels and line conditions found in the particular network in which the echo canceller is to be used. Values of fixed suppression threshold levels to be used are under study – see Notes 1 and 2.

NOTE 1 – As an interim guide, it is suggested that the suppression threshold level should be set a few decibels above the level that would result in the *peaks* of L_{RES} for a "2s-talker" and a "2s-ERL" being suppressed.

NOTE 2 – Results of a field trial reported by one Administration indicated that a fixed suppression threshold level of –36 dBm0 gave a satisfactory performance. A theoretical study, by another Administration, of an echo path containing five pairs of PCM codecs showed that for an L_{Rin} of –10 dBm0, the quantization noise could result in an L_{RES} of –38 dBm0.

8.2.2.3 Adaptive suppression threshold

A good compromise can be made between using a high T_{SUP} to prevent it being exceeded by loud talker residual echo and using a low T_{SUP} to reduce speech distortion on break-in by making T_{SUP} adaptive to the actual circuit conditions and speech levels. This may be achieved in a number of ways and no recommendation is made for any particular implementation. General guidelines applicable to the control algorithm and suppression threshold levels are under study.

8.2.3 Control of NLP activation

8.2.3.1 General

To conform to the recommendation made in clause 8.2.1.3, it is necessary to control the activation of the NLP so that it is not active when cancelled end speech is likely to be present. When "active", the NLP should function as intended to reduce L_{RES}. When "inactive", it should not perform any non-linear processing on any signal passing through the echo canceller.

8.2.3.2 Control guidelines

It is recommended that the following two guidelines should govern control of the activation of an NLP. First, because they are intended to further reduce L_{RES}, they should be active when L_{RES} is at a significant level. Second, because they should not distort cancelled end speech, they should be inactive when cancelled end speech is present. Where these two guidelines conflict, the control function should favour the second.

8.2.3.3 Static characteristics

A conceptual diagram showing the two operational states of an NLP is shown in Figure 40. The L_{Sin} L_{Rin} plane is divided into two regions, W and Z by the threshold WZ (T_{WZ}). In the W region the NLP is inactive while in the Z region it is active. Proper control of the NLP to ensure operation in the appropriate region requires recognition of the double talk condition or the presence of cancelled end speech. Imperfect detection of double talk combined with a high suppression threshold level will result in distortion of cancelled end speech. The echo canceller then exhibits some of the characteristics of an echo suppressor. A low suppression level will permit easy double talking, even if a detection error is made because the cancelled end speech will suffer only a low level of non-linear distortion. If the suppression threshold level is too low, then peaks of residual echo may be heard.

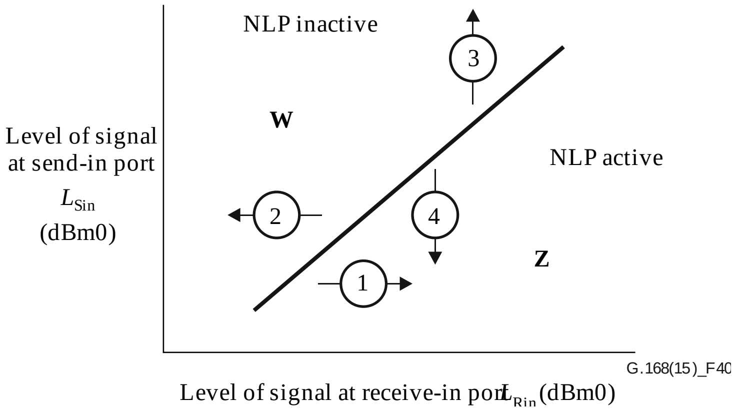


Figure 40 – NLP operating regions

8.2.3.4 Dynamic characteristics

The dynamic characteristics can be specified by stating the time that elapses when the signal conditions pass from a point in one area to a point in the other area before the state appropriate to the second area is established. Four such transitions are shown by arrows in Figure 40.

Transition No. 1 – W to Z, L_{Sin} constant, L_{Rin} increasing

In this case, the L_{Sin} signal occurred first and the L_{Rin} is increasing to a sufficiently high level to override the L_{Sin} signal in the control path and cause the NLP to change from the inactive to the active state. Since this will cause distortion of the L_{Sin} signal (cancelled end talker speech in this case), the action should not be initiated too quickly.

Transition No. 2 – Z to W, L_{Sin} constant, L_{Rin} decreasing

In this case, the L_{Rin} signal has overridden the L_{Sin} signal in the control path and the NLP is in the active state. The L_{Rin} signal is now decreasing. The NLP should remain in the active state sufficiently long to prevent echo, which is stored in the echo path, from being heard by the far talker.

Transition No. 3 – Z to W, L_{Rin} constant, L_{Sin} increasing

This transition is replicating the onset of double talk. As soon as possible after the L_{Sin} signal is detected, the NLP should be switched to the inactive state in order to minimize any distortion of the cancelled end talker speech.

Transition No. 4 – W to Z, L_{Rin} constant, L_{Sin} decreasing

In this case, L_{Sin} has been recognized but is decreasing. Any action which is taken should favour continuing to permit the L_{Sin} signal to pass. This implies there should be some delay in switching the NLP back to the active state.

8.2.4 Frequency limits of control paths

Under study.

NOTE – Depending on the particular implementation of the NLP, the considerations and frequency response limits given in clause 3.2.4.2 of [ITU-T G.164] for the suppression and break-in control paths of echo suppressors may also be applicable to similar control paths used in NLPs. These control paths may include the activation control and adaptive suppression threshold level control.

8.2.5 Signal attenuation below threshold level

The attenuation of signals having a level below that of the suppression threshold level of an NLP in the active state should be such that the requirements of clause 6.4.2.3.1 are met.

8.2.6 Testing of NLPs

The NLP may be considered as a special case of an echo suppressor which is limited to suppressing only low-level signals. The types of test required to determine the NLP performance characteristics are very similar to the echo suppressor tests given in [ITU-T G.164]. However, depending on the specific implementation of an NLP, the transitions between areas W and Z of Figure 40 may not be as sharply defined as is the case for echo suppressors. Signals observed at the send-out port of the echo canceller may be distorted for short periods when transitions between the W and Z operating regions occur. Although [ITU-T G.164] may be used as a guide to the testing of NLPs, it may be necessary to introduce unique test circuit modifications in order to make measurements on some specific NLP implementations. In particular, it is known that some echo cancellers employ a supplementary NLP technique independent of the NLP enable/disable control. Those tests of this Recommendation which call for NLP disabled do not strictly apply to such cancellers unless some further interpretation of results is made. The following clause suggests a technique for identifying the presence of such supplementary NLP functions.

8.2.6.1 Testing for the presence of an NLP

Set up Test 2B of this Recommendation (Convergence without NLP). However, modify the test by mixing uncorrelated noise at various known levels below the returned CSS echo in the Send path. The noise level establishes a floor below which noise plus residual echo due to cancellation should not fall. Observing residual echo levels below the noise floor, then, is an indication of NLP action.

Annex A

Description of an echo canceller reference tone disabler

Former Annex A of 2000 edition has been deleted.
Any information or requirements have been transferred
to clause 7 of this Recommendation

Annex B

Description of a reference non-linear processor

(This annex forms an integral part of this Recommendation.)

B.1 General

This annex, which is for the purposes of illustration only and not intended as a detailed design (see clause 8.1), describes a reference non-linear processor (NLP) based upon concepts that are as simple as possible but having included in it a sufficient number of features to give guidance for a wide range of possible implementations. To this end, two variants of the reference NLP are included. Both are based on a centre clipper having either of the idealized transfer functions illustrated in Figure B.1. The suppression threshold level (determined, in this case by the clipping level) in the first variant is adaptive, adaptation being by reference to L_{Rin}. Activation control is by reference to the difference between L_{Rin} and L_{Sin}. In the second variant, the suppression threshold is fixed. It is assumed that the reference NLP is used in an echo canceller which can achieve a cancellation of the linear components of any returned echo of at least N dB. The value of N is under study.

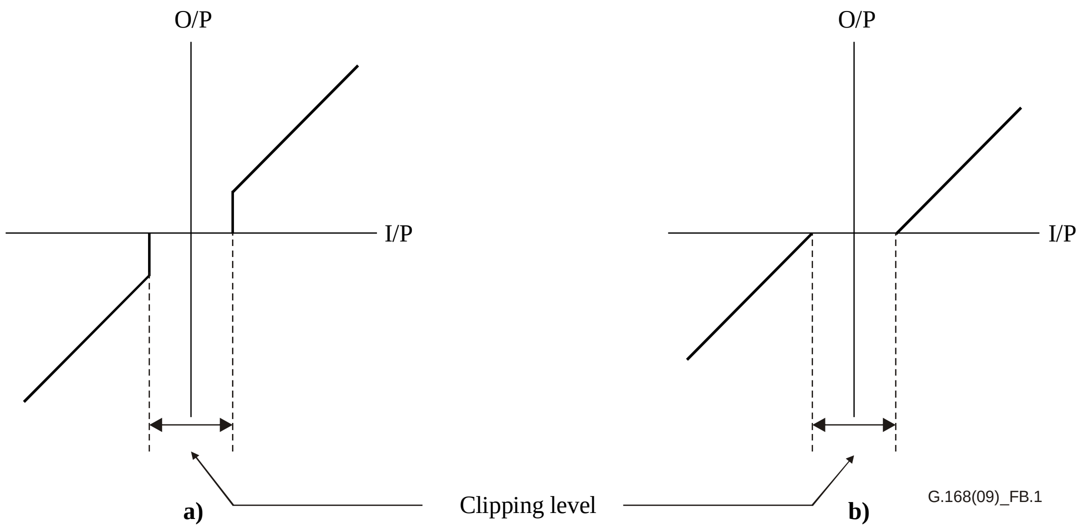


Figure B.1 – Two examples of idealized centre clipper transfer function

B.2 Suppression threshold (T_{SUP})

Adaptive $T_{\text{SUP}} = (L_{\text{Rin}} - x \pm 3)$ dBm0 for $-30 \leq L_{\text{Rin}} \leq -10$ dBm0

Fixed $T_{\text{SUP}} = x'$ dBm0

NOTE – Values of x and x' are under study. Values of 18 for x and –36 for x' have been suggested but confirmation is required that these values are appropriate for use in all networks.

B.3 Static characteristics of activation control

$T_{\text{WZ}} = (L_{\text{Rin}} - y \pm 3)$ dBm0 for $-30 \leq L_{\text{Rin}} \leq -10$ dBm0

NOTE 1 – T_{WZ} is as defined in clause 8.2.3.3.

NOTE 2 – The value of y may be different for each variant, and this is under study. Values of x dB in the case of the adaptive T_{SUP} and ≥ 6 dB for y in the case of the fixed T_{SUP} seem reasonable.

B.4 Dynamic characteristics of activation control

Dynamic characteristics of the activation control are given in Tables B.1 and B.2. Also see Figure 40.

Table B.1 – NLP hangover times

Boundary		Initial signal		Final signal		Recommended value (ms)	Test No. [ITU-T G.164]	Excursion (see Figure 40)	Test circuit, Figure:	Oscilloscope trace
		Send L_{Sin} (dBm0)	Receive L_{Rin} (dBm0)	Send L_{Sin} (dBm0)	Receive L_{Rin} (dBm0)					
	Fixed	–25	–10	–25	–30	15-64				
Z/W	Adaptive	–55 –40 –30	–20 –15 –5	–55 –40 –30	–40 –40 –30	Δ	5	Transition 2	14/ITU-T G.164	Trace 1 and trace 2 of Figure B.3 (β)
	Fixed	–15	–25	–40	–25	16-120				
W/Z	Adaptive	–40 –40 –25	–50 –30 –15	–55 –55 –40	–50 –30 –15	30-50	6	Transition 4	17/ITU-T G.164	Trace 1 and trace 2 of Figure B.2 (β)

Table B.2 – NLP operate times

Boundary		Initial signal		Final signal		Recommended value (ms)	Test No. [ITU-T G.164]	Excursion (see Figure 40)	Test circuit, Figure:	Oscilloscope trace
		Send L_{Sin} (dBm0)	Receive L_{Rin} (dBm0)	Send L_{Sin} (dBm0)	Receive L_{Rin} (dBm0)					
	Fixed	–25	–30	–25	–10	16-120				
W/Z	Adaptive	–55 –40 –30	–40 –40 –30	–55 –40 –30	–20 –15 –5	15-75	4	Transition 1	14/ITU-T G.164	Trace 2 of Figure B.3 (α)
	Fixed	–40	–25	–15	–25	≤ 1				
Z/W	Adaptive	–55 –55 –40	–50 –30 –15	–40 –40 –25	–50 –30 –15	≤ 5	6	Transition 3	17/ITU-T G.164	Trace 2 of Figure B.2 (α)

B.5 Frequency limits of control paths

See clause 8.2.4.

B.6 Testing

Tables B.1 and B.2 indicate, by reference to [ITU-T G.164], how the dynamic performance of NLP activation control may be checked using sine wave signals. Figures B.2 and B.3 show the traces obtained on an oscilloscope for these tests.

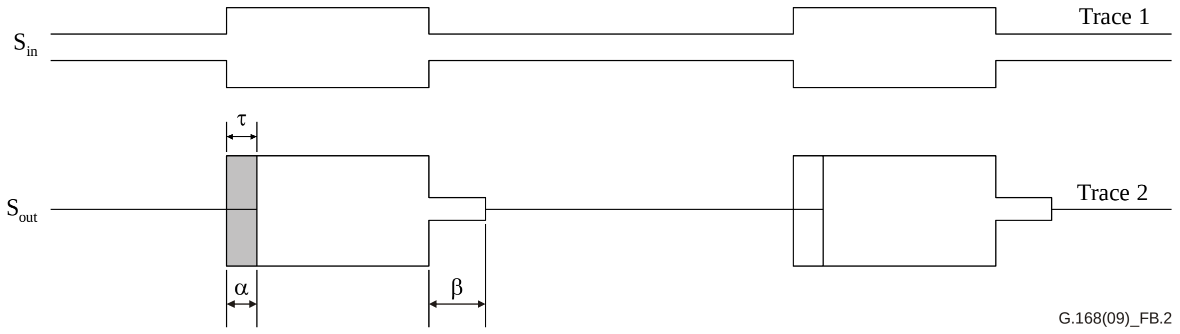
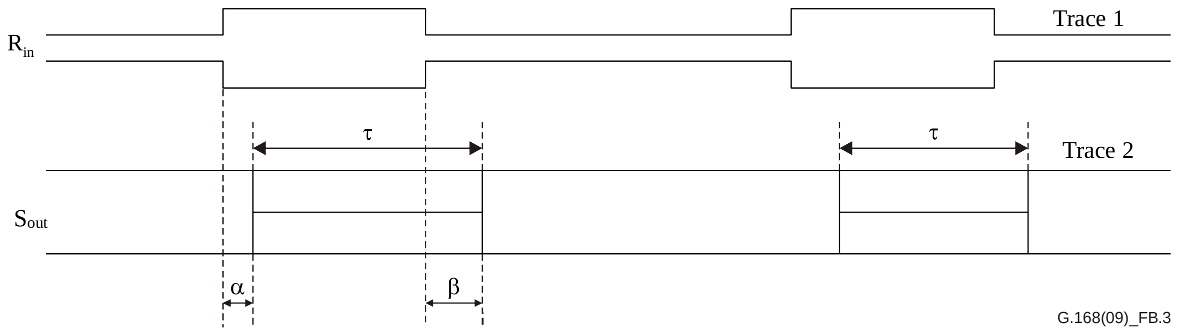


Figure B.2 – Traces for NLP operate and hangover time, L_{Rin} constant



α Operate time
β Hangover time
τ Time interval in which the distorted signal may be observed

Figure B.3 – Traces for NLP operate and hangover times, L_{Sin} constant

Annex C

Composite source signals for testing of speech echo cancellers – Signal, description and analysis

(This annex forms an integral part of this Recommendation.)

C.1 Introduction

This annex describes the subset of composite source signals that are used for testing speech echo cancellers in the network under single and double talk conditions. The exact definition of these signals is part of [ITU-T P.501] on test signals for use in telephonometry (see [b-ITU-T GSTP.CSS] in clause I.8 for more details). First, a general description of composite source signals is given. The following clauses give the exact definition of both signals for testing speech echo cancellers under single and double talk conditions. Moreover, kinds of analysis are considered and described to test the specific parameters of echo cancellers especially for the tests of this Recommendation.

C.2 Composite source signal – General considerations

C.2.1 General description of the different sequences

Composite source signals, in general, consists of different sequences including voiced and unvoiced sounds as well as pauses.

Voiced signal produced from the "artificial voice" signal according to [ITU-T P.50]

The voiced signal part of CSS is the conditioning signal intended to activate possible speech detectors in voice-controlled systems and to reproduce voiced sounds of real speech in general. As the duration, beginning and end of the voiced signal are known exactly, this signal can also be used to measure the switching time for the direction of transmission under test. By means of the signal shape in the time domain, the switching time and delay time of the entire system can be determined. The duration of the signal amounts to 50 ms approximately.

Pseudo noise signal

The signal presented after the voiced artificial speech sound is the Pseudo Noise (PN) signal. This signal has certain noise-like features. The magnitude of its Fourier transform is initially constant with frequency while the phase is changing. For tests usually only the magnitude of the transfer function is of interest, the phase is not that important but can be determined as well.

The signal is produced as follows:

First a complex spectrum is produced in the frequency domain according to the following equation:

$$H(k) = W(k)e^{j \cdot i_k \cdot \pi}; k = -\frac{M}{2}, \dots, \frac{M}{2} \text{ without } 0; i_k \{+1,0\}, i_k = -i_{-k} \text{ random} \qquad \text{(C.2-1)}$$

The index M is adjusted to the chosen FFT size (e.g., 2048, 4096 or 8192 points). The equation shows that the amount of the produced complex spectrum is constant for all frequencies if W(k) is chosen equal to 1 for all frequencies, whereas the phase may be π or 0 for each frequency, corresponding to a random sequence. However, to produce a different weighting in the frequency domain, W(k) can easily be adjusted in order to produce different spectra for the duration of the PN-sequence. Then, this spectrum will be transformed into the time domain by means of the inverse Fourier transform producing the following signal:

$$S(n)=\frac{1}{M}\sum_{k=-M/2,k\neq 0}^{M/2} H(k)\cdot e^{j2\pi\cdot n\cdot k/M},\ n=-\frac{M}{2},\ \ldots\ ,\frac{M}{2}-1 \qquad \text{(C.2-2)}$$

NOTE 1 – Thus, a signal is produced which is limited in time (corresponding to the chosen length of the Fourier transform) and which is adjusted to the chosen FFT size correctly. If a longer time sequence is wanted, the signal can be cycled. This method permits time sequences of any length. The duration of this measurement signal amounts to about 200 ms by appropriate choice of M, the sampling rate and numbers of repetitions.

The pseudo noise sequence of the composite source signal for measurements of speech echo cancellers is calculated in that way that W(k) is chosen constant and the corresponding signal S(n) (calculated by inverse Fourier transform) is filtered with a transfer function which is given below in clause C.3.1.

NOTE 2 – Typically, the length of the FFT should be short for systems with highly time variant parameters such as companding techniques in order to get a good short time estimation of the time variant transfer function. For systems incorporating adaptive techniques such as echo cancellers or noise cancellers a higher number of M (close to 200 ms signal duration) may be appropriate in order to have the autocorrelation function of the measurement signal not periodically within the processing window of the device under test.

Pause

The third part of the composite source signal is a pause. Regarding the composite source signal as a measurement signal that reproduces important characteristics of real running speech, the pause has the purpose to provide suitable amplitude modulation to the composite signal. Moreover it reproduces real speech pauses that occur in running speech signals as well. This also means a certain period without excitation signal, which gives the possibility to analyse noise or artefacts produced by the system under test. The length of the pause is chosen between 100 ms and 150 ms.

In order to achieve a long term offset free sequence, the repeated CS-sequence should be inverted in amplitude (phase shift by 180°).

C.2.2 Calculation and analysis using a composite source signal

When using CSS for measurements, the sequence of voiced sound, pseudo noise signal and pause can be cycled. This means that after the pause the sequence starts again beginning with a voiced sound. Using this procedure sequences of any length may be produced.

Having created a sequence as described above, this signal can be handled like a standard measurement signal, e.g., a white noise signal or a switched pink noise. The level calibration (acoustical and electrical) is done using the whole sequence including voiced sounds, PN-sequences and pauses. In principle a standard RMS meter with a bandwidth of 20 kHz operating with "fast" averaging can be used. Another method is to use a FFT analysis for level calculations. The parameters for the FFT based calculation are:

- sampling rate according to the one chosen for signal generation (preferred 44.1 kHz or 48 kHz);
- FFT length according to the one chosen for signal generation;
- rectangular windowing;
- no overlap;
- averaging over the whole (cycled) sequence, including voiced sounds, PN-sequences, pauses;
- calculation of the level from the power density spectrum derived by the FFT calculation (integration of the levels over all frequency components).

C.3 Bandlimited composite source signal with speech like power density spectrum – Practical realization for measurements of echo cancellers

Both composite source signals described below in this annex have a speech-like power density spectrum. This means that the noise sequences of both signals (the measurement signal and the signal to simulate double talk) are shaped with a decrease of 5 dB/octave towards higher frequency. The convergence characteristics of speech echo cancellers largely depend on the power density spectrum of the input signal. Therefore, these composite source signals were adapted in this way to reproduce the power density spectrum of real speech.

C.3.1 Composite source signal for single talk

Figure C.1 shows the principle construction of the composite source signal for single talk.

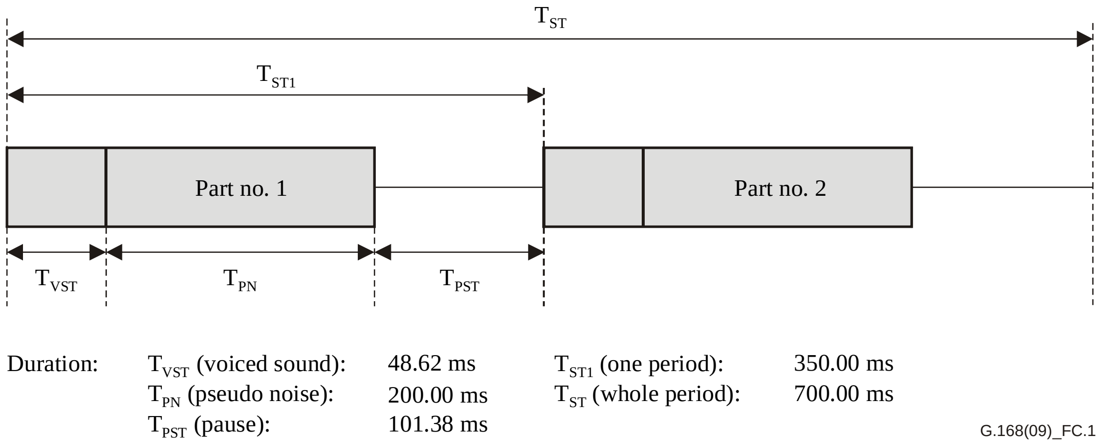


Figure C.1 – Composite source signal for measuring echo cancellers (schematic)

Bandlimited voiced signal

In Table C.1 the 16-bit word values for the voiced signal, bandlimited between 200 Hz and 3.6 kHz can be found. According to a sampling rate of 44.1 kHz, the 134 16-bit word values amount to 3.04 ms. The values are to be read in columns:

Table C.1 – 16-bit word values of the bandlimited voiced signal

–155	948	3224	4000	3129	1440	241	–888	–1853	–6137	–3474
276	1362	3370	4043	3043	1310	190	–957	–2121	–6560	–2508
517	1741	3500	4034	2914	1146	103	–1034	–2414	–6948	–1595
578	2043	3569	3974	2750	965	–9	–1103	–2707	–7301	–802
491	2276	3603	3862	2560	776	–138	–1146	–3017	–7568	
302	2422	3603	3724	2353	603	–267	–1181	–3319	–7732	
86	2500	3595	3577	2155	448	–388	–1190	–3612	–7758	
–103	2552	3586	3439	1991	345	–491	–1198	–3913	–7620	
–207	2595	3595	3336	1853	276	–569	–1215	–4224	–7310	
–198	2655	3638	3267	1750	250	–638	–1259	–4560	–6810	
–60	2758	3724	3224	1672	250	–698	–1327	–4922	–6155	
190	2896	3819	3198	1603	267	–759	–1457	–5301	–5344	
543	3060	3922	3172	1534	267	–813	–1629	–5715	–4439	

The values of the voiced signal in the frequency range 200 Hz-3.6 kHz again are calculated such that the RMS value of the voiced signal and the PN-sequence are equal. The sequence is repeated 16 times to achieve a length of 48.62 ms.

Pseudo noise signal generated using 2048 pt. FFT

The parameters for the PN-sequence are:

Sampling rate 44.1 kHz, 16-bit word length, length of Fourier transform 2048 points.

$$H(k) = \begin{cases} W(k) \cdot e^{j \cdot i_k \cdot \pi}; k = -928, \ldots, +928 \; except \; 0, i_k \{+1, 0\}, random, i_k = -i_{-k} \\ 0 \; else \end{cases} \quad \text{(C.3-1)}$$

According to the above-described Formula C.2-2, the time signal is calculated by inverse Fourier-Transformation. This sequence is repeated 4.307 times to achieve a length of 200 ms for the PN sequence. The crest factor of the PN-sequence is 11 dB ± 1 dB.

According to the frequency resolution of 21.5 Hz (44.1 kHz/2048), there are 928 FFT values in the frequency range between 0 and 20 kHz. Each value W(k) (before filtering) is 152 680. It is calculated such that levels within a bandwidth of 20 kHz are the same for the voiced signal and the PN sequence.

Pseudo noise signal generated using 8192 pt. FFT

According to the above-described Formula C.2-2, the time signal is calculated by inverse Fourier-Transformation. This sequence is repeated 1.077 times to achieve a length of 200 ms for the PN-sequence. The crest factor of the PN sequence is 11 dB ± 1 dB.

According to the frequency resolution of 5.4 Hz (44.1 kHz/8192), there are 3715 FFT values in the frequency range between 0 and 20 kHz. Each value W(k) before filtering is 305 360. It is calculated such that levels within a bandwidth of 20 kHz are the same for the voiced signal and the PN-sequence.

In order to achieve the same RMS value for the bandlimited PN-sequence the filter function shown in Figure C.2 should be applied. The filter is chosen such, that the levels of the filtered and the unfiltered PN-sequence are equal. The filter owner frequencies are shown in Table C.2.

NOTE – By appropriate up- or down-sampling other sampling rates for the described sequence can be achieved. The interpolation filter used for up- and down-sampling should be close to an ideal rectangular filter. The stopband attenuation should be > 60 dB, the passband ripple < ±0.2 dB.

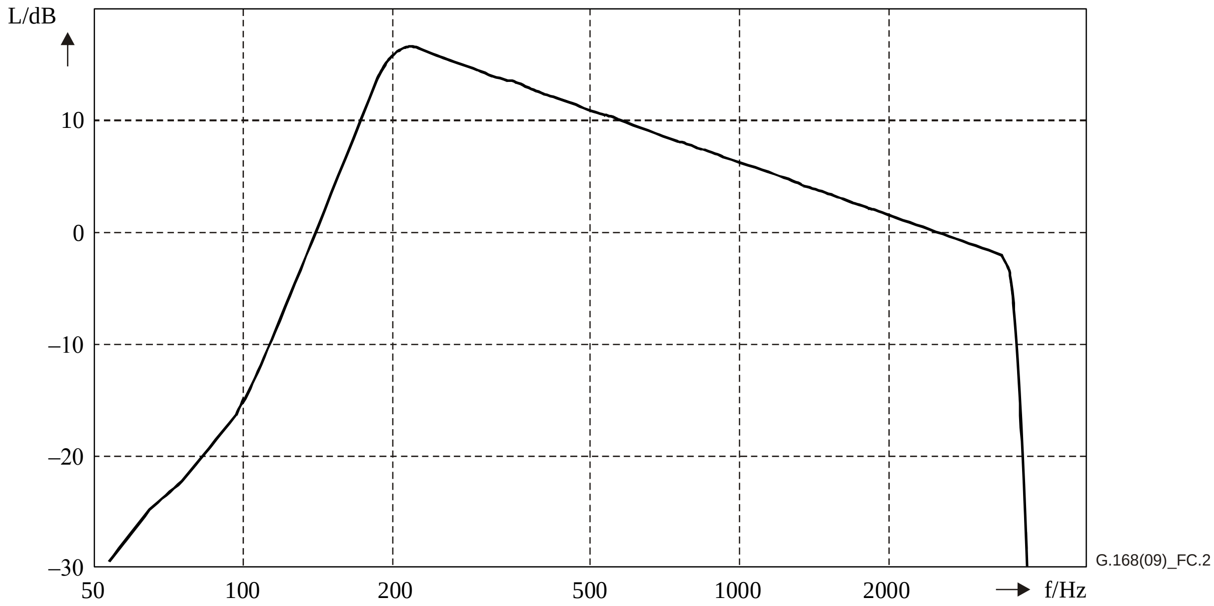


Figure C.2 – Transfer function of the filter for bandlimiting the PN-sequence

Table C.2 – Table of filter corner frequencies

50 Hz	100 Hz	200 Hz	215 Hz	500 Hz	1 kHz	2.85 kHz	3.6 kHz	3.66 kHz	3.68 kHz
–25.8 dB	–12.8 dB	17.4 dB	17.8 dB	12.2 dB	7.2 dB	0 dB	–2 dB	–20 dB	–30 dB

For adaptive systems such as echo cancellers, a longer PN sequence may be preferable in order not to have correlated measurement signals within the adaptation window. For those systems the FFT length should be extended to 8192 points when using 44.1 kHz sampling rate as described above.

Pause

The length of the pause is chosen to 101.38 ms in order to achieve a complete length of 350 ms for the voiced sound, the pseudo noise sequence and the pause.

To achieve a long term offset free sequence, this CS-sequence of 350 ms is repeated and inverted in amplitude (phase shift by 180°). The complete length amounts to 700 ms.

C.3.2 Bandlimited composite source signal to simulate double talk

The double talk sequence is generated in the same way as the single talk signal. Figure C.7 shows the principle construction of the double talk signal. However, the times of the voiced signal and the pause are slightly different in order to achieve a typical double talk condition with two signals applied the same time, signal present only in one channel, voiced signals present on both sides as well as voiced signals and unvoiced signals present the same time in the different channels. The correlation between single talk signal and double talk signal is low. This is achieved by choosing a different voiced signal with a different pitch frequency and a random noise signal instead of the PN sequence. The duration of the voiced signal is 72.69 ms, the duration of the random noise signal is 200 ms and the duration of the pause amounts to 127.31 ms.

Voiced signal

The voiced signal for double talk was chosen to have a different base frequency than the signal talk voiced signal. The values for the voiced signal for double talk can be found in Table C.3. The level of this sound again is the same as the one for single talk. Using a sampling rate of 44.1 kHz 229 16-bit word values represent 5.19 ms. The table is to be read in columns:

Table C.3 – 16-bit word values for the bandlimited double talk voiced signal

–198	1146	–8292	4827	5853	1422	–1293	–810	–690	–1052	–621
–112	871	–8715	5094	5715	1224	–1302	–793	–724	–1043	–560
–9	560	–9077	5344	5560	1026	–1293	–767	–767	–1043	–509
103	233	–9370	5594	5387	819	–1267	–741	–793	–1052	–457
233	–121	–9542	5827	5215	603	–1250	–698	–819	–1060	–397
388	–491	–9542	6043	5043	388	–1233	–672	–845	–1060	–345
543	–871	–9361	6215	4879	181	–1224	–638	–853	–1060	–276
724	–1250	–8956	6344	4732	9	–1224	–603	–871	–1052	–207
896	–1638	–8327	6413	4586	–181	–1224	–595	–879	–1034	–112
1060	–2043	–7465	6422	4439	–328	–1224	–586	–888	–1017	
1233	–2465	–6396	6379	4276	–448	–1215	–595	–896	–991	
1388	–2896	–5163	6310	4086	–543	–1198	–603	–922	–957	
1517	–3345	–3827	6215	3870	–629	–1172	–621	–948	–931	
1638	–3819	–2448	6120	3629	–707	–1129	–629	–974	–905	
1747	–4310	–1103	6051	3370	–784	–1077	–938	–1009	–888	
1810	–4810	155	6000	3086	–871	–1026	–638	–1026	–862	
1845	–5319	1293	5991	2801	–948	–974	–638	–1052	–845	
1845	–5836	2241	5991	2534	–1026	–922	–638	–1069	–819	
1802	–6353	3034	6000	2267	–1112	–888	–638	–1077	–793	
1707	–6853	3655	6008	2034	–1181	–871	–638	–1069	–767	
1569	–7353	4138	5991	1819	–1241	–845	–647	–1060	–724	
1379	–7836	4517	5939	1612	–1276	–828	–664	–1060	–672	

In order to achieve the required length of 72.69 ms, the values are to be repeated 14 times.

Random noise

The random noise is chosen as a White Gaussian Noise bandlimited at 20 kHz. The crest factor of the signal is 12 ± 1 dB. The RMS value of the bandlimited random noise is chosen to be the same as the one for the voiced signal.

In order to bandlimit the random noise between 200 Hz and 3.6 kHz, the filter function shown in Figure C.2 is used. This ensures the same RMS value for the bandlimited random noise.

Pause

The pause is chosen to 127.31 ms in order to achieve a length of 400 ms for the voiced sound, the random noise sequence and the pause.

Again, in order to achieve a long-term signal which is free of offset, this sequence of 400 ms is repeated and inverted in amplitude (phase shift by 180°). Thus the resulting length of the double talk signal is 800 ms.

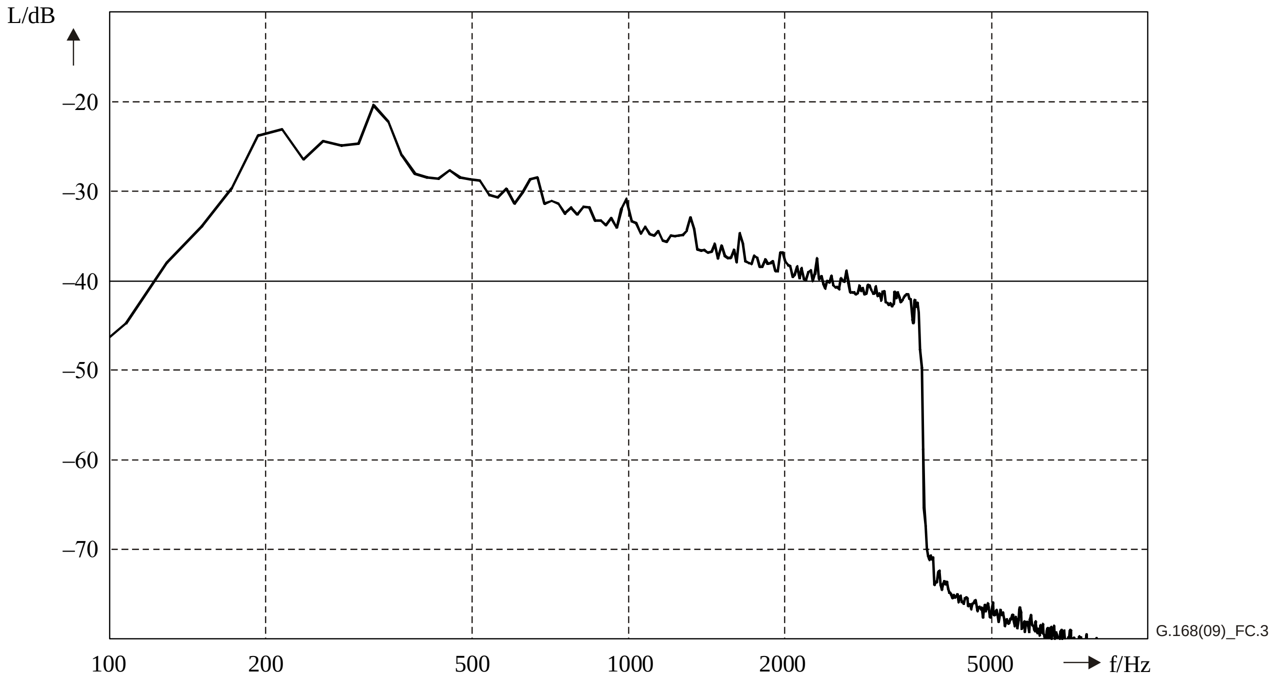


Figure C.3 – Power density spectrum of the bandlimited CSS (single talk signal, analysis window: Hanning)

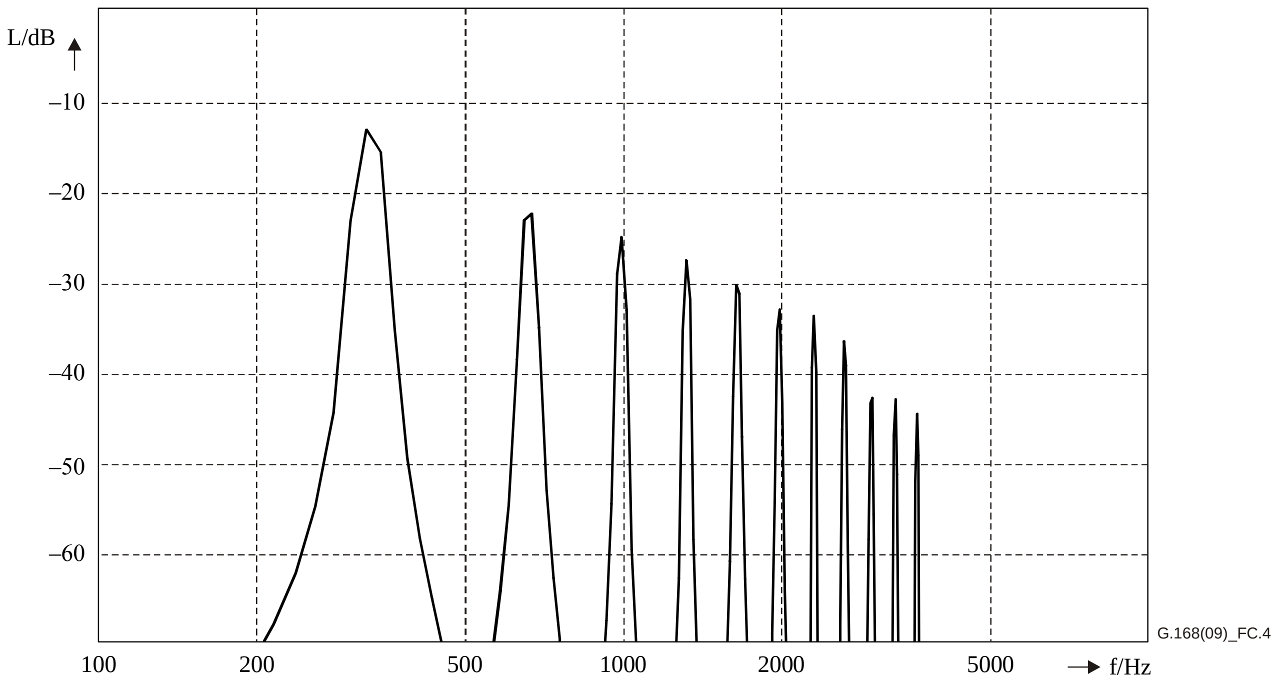


Figure C.4 – Power density spectrum of the bandlimited voiced signal (single talk signal, analysis window: Hanning)

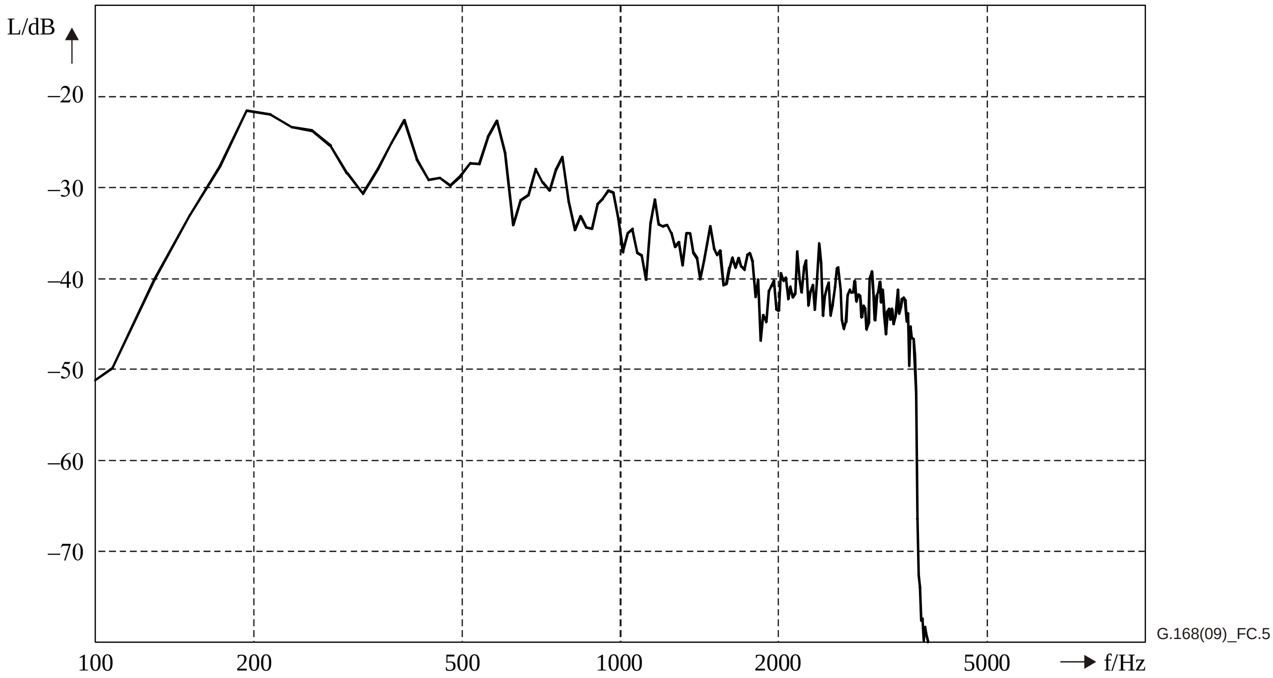


Figure C.5 – Power density spectrum of the bandlimited double-talk CSS (analysis window: Hanning)

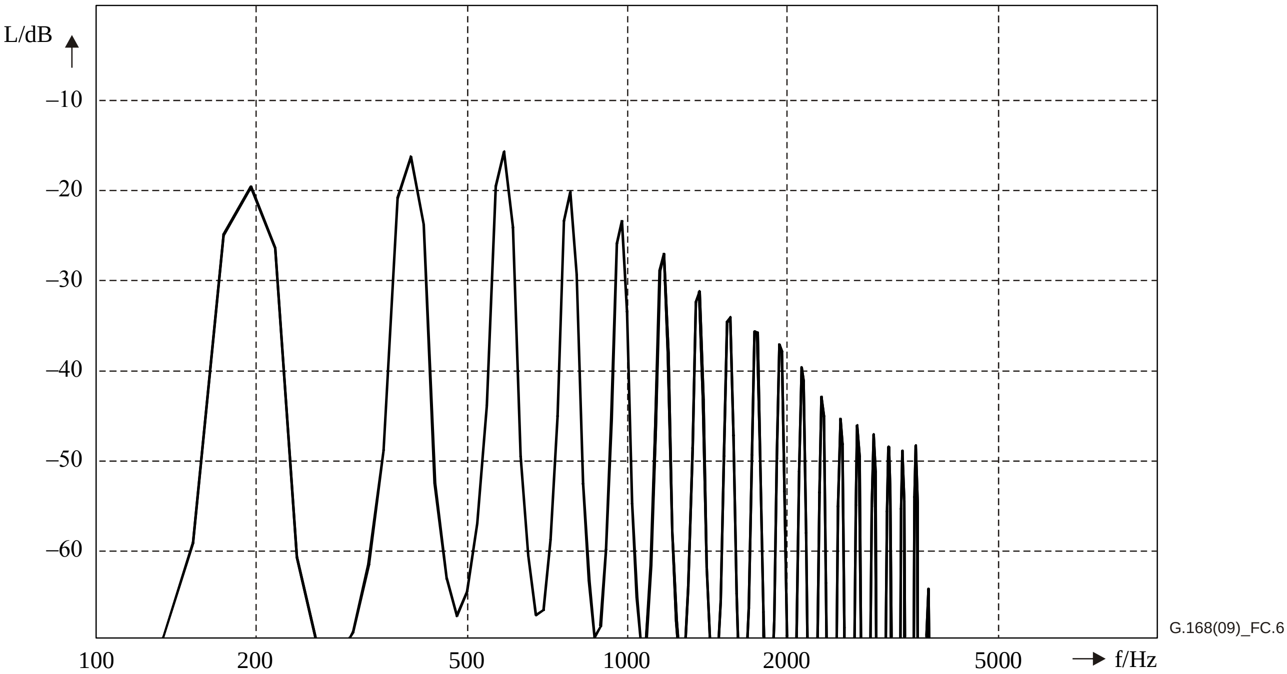
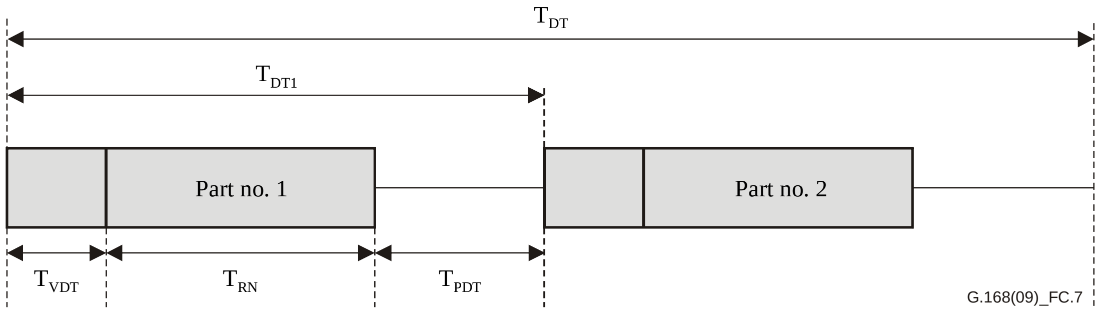


Figure C.6 – Power density spectrum of the bandlimited double talk-voiced signal (analysis window: Hanning)

NOTE – By appropriate up- or down-sampling, other sampling rates for the described sequence can be achieved. The interpolation filter used for up- and down-sampling should be close to an ideal rectangular filter. The stopband attenuation should be > 60 dB, the passband ripple < ±0.2 dB.



Duration:	T_{VDT} (voiced sound):	72.69 ms	T_{DT1} (one period):	400.00 ms
	T_{RN} (random):	200.00 ms	T_{DT} (whole period):	800.00 ms
	T_{PDT} (pause):	127.31 ms		

Figure C.7 – Composite source signals to simulate double talk (schematic)

Application

The application of the bandlimited composite source signals for single talk as well as for double talk is for all testing where bandlimited systems need to be tested working non-linear and time variant and requiring the typical long-term power density spectrum of speech. The typical application is the testing of speech echo cancellers in the network. For all one directional tests, the bandlimited CSS for single talk tests should be used. In case of tests in double talk conditions, the double talk signal should be used in double talk direction (S_{gen}), whereas the single talk signal is fed in the far-end direction (R_{in}).

C.4 Appropriate analyses to determine convergence characteristics of speech echo cancellers using the composite source signal

The composite source signal for testing speech echo cancellers and the second composite source signal to simulate double talk are described above. If the echo signal level should be measured, there are several possibilities of analysis technique. Calculations can be made in the time or frequency domain.

C.4.1 Calculation in the frequency domain

The signal level can be determined by calculations in the frequency domain, after the time sequence has been transformed by Fourier Transformation. This allows the level calculations in a certain frequency range, i.e., the telephone bandwidth of 300 Hz to 3.4 kHz. Another advantage is that the Fourier Transformation gives the possibility to analyse further characteristics of the echo signal in the frequency range, for example the echo attenuation versus frequency. For the composite source signal, a rectangle window should be used before calculating the Fourier Transformation. The Pseudo Noise sequence is generated with a 8192 points FFT. The sampling rate should be 44.1 kHz as described above in clauses C.3.1 and C.3.2 for generating the composite source signals. The sequence length used for transformation should be the complete length of 700 ms including the voiced sound, the pseudo noise sequence and the pause. Various measurements showed that due to signal delay or noise produced by the circuit under test, additional artefacts may appear during the pauses (e.g., switched residual echo signal or modulated background noise). Therefore it is suitable to analyse the echo signal over a sequence length of 700 ms, i.e., one whole period of the composite source signal. A disadvantage of level calculations from the frequency range is due to the fact that this gives only a limited time resolution of one Fourier Transformation length. The level calculation in the frequency domain should be used to determine signal levels and residual echo levels after full convergence of after inhibiting adaptation.

C.4.2 Calculation in the time domain

The echo signal level calculation from the time domain is necessary for analysis of echo attenuation versus time because of its high resolution in the time domain. A suitable method is given through [IEC 61672-1], sound level meters. It describes the sound level measurement and recommends three different time constants, "Slow" (1000 ms), "Fast" (125 ms) and "Impulse" (35 ms). If measurement results of different laboratories should be compared, an agreement about the measurement procedure is necessary. A short time constant has advantages because of the highest possible resolution in the time domain, whereas longer time constants have the advantage that the results obtained with this kind of calculation demonstrate more the average level of the time sequence that is analysed. Especially if several measurements calculated, for example, as the level versus time are represented in the same picture, very short time constants may lead to confusing representations. This is due to the fact that using a very short time constant of, for example, 35 ms ("Impulse"), the calculation is more sensitive to even very small signal variations. For this reasons the use of the time constant "Fast" (125 ms) according to [IEC 61672-1] is more suitable for level calculations versus time.

This is a suitable method to analyse the convergence speed of speech echo cancellers at the beginning of adaptation. The echo signal level is calculated using the time constant "Fast" according to [IEC 61672-1]. Level fluctuations due to input signal fluctuations can be eliminated if the echo signal level is referred to the input signal level. This represents the echo return loss enhancement (ERLE) versus time. A disadvantage is that no further analysis is possible in the frequency domain. When using the meters of [IEC 61672-1], any peak detection or decay time constants referenced in [IEC 61672-1] should not be incorporated for measurements in this Recommendation.

C.4.3 Level calculations according to [ITU-T P.56] active speech level

Level calculations can also be done according to [ITU-T P.56]. This calculation is made from the time domain as well. It delivers one value and a percentage of speech activity. It may be suitable to calculate the residual echo level but there are more parameters that have to be defined to guarantee the same implementation of this algorithm. Difficulties may appear if echo signals with a very low level are analysed. It may fall below the recognition level for active speech. Another disadvantage is, although this is a calculation in the time domain it delivers only one value. It is not possible to achieve the level variation versus time, as it is important for convergence measurements. Therefore it is more suitable to analyse residual echo signal level using the Fourier Transformation as described in clause C.4.1 or the level calculation in the time domain for time varying echo signals (e.g., the convergence of echo cancellers) based on [IEC 61672-1] as described in clause C.4.2.

Annex D

Echo-path models for testing of speech echo cancellers

(This annex forms an integral part of this Recommendation.)

D.1 Introduction

The following echo path models can be used for the tests in this Recommendation. The echo path is simulated by a linear digital filter with the impulse response g(k). To account for various delays, different ERLs and different dispersion characteristics and time widths, g(k) is chosen as a delayed and attenuated version of any of the sequences $m_i(k)$, i = 1, 2, ..., 8, that are given in tables in clauses D.2 and D.3.

$$g(k) = (10^{-ERL/20} K_i) m_i (k - \delta) \quad \text{(D-1)}$$

The sequences $m_i(k)$ represent echo paths with various dispersion characteristics and different time widths. The delay δ should be chosen such that the non-zero values of g(k) can be captured by the H register of the echo canceller. The scaling factor K_i depends on the input signals used in the tests. The value of K_i in Table D.1a results in an ERL that is equivalent to the value used in Equation D-1 when measured with CSS. The value of K_i in Table D.1b is used to limit the maximum of the magnitude response to the chosen ERL value. The K_i factors for CSS or white noise input, and for tone input are given below. Specifically for Test 10 Fax test, the K_i factors for tone signal given in Table D.1b should be used. Caution should be used when using Table D.1a for white noise, as the measured ERL may not be equal to the ERL value used in Equation D-1. Only Test 9 uses white noise as the input signal.

D.1.1 CSS or white noise input

For the tests that use CSS or white noise as the input signals, the values of K_i are given in the Table D.1a for the eight sequences $m_i(k)$:

Table D.1a – Scaling factors K_i for the eight digital echo path models

Echo path model #(i)	Scaling factor K_i	Minimum ERL for CSS (dB) (Note)
1	1.39×10^{-5}	6
2	1.44×10^{-5}	6.55
3	1.52×10^{-5}	6
4	1.77×10^{-5}	6
5	9.33×10^{-6}	6
6	1.51×10^{-5}	6
7	2.33×10^{-5}	11.06
8	1.33×10^{-5}	9.27
NOTE – A minimum ERL value of 6 dB should be used in the tests for echo path models m_1, m_3, m_4, m_5 and m_6. For echo path models m_2, m_7 and m_8, the minimum ERL values used in the tests should be, respectively, 6.55 dB, 11.06 dB and 9.27 dB. This is to ensure that the magnitude response of the scaled echo-path g(k) does not exceed 0 dB over the appropriate frequency range.		

D.1.2 Tone(s) input

For the tests that use tone(s) as input signals, the values of K_i should be computed using the formula

$$K_i = \frac{1}{\max\limits_{f} |M_i(f)|} \tag{D-2}$$

where:

$$M_i(f) = \sum_{k=0}^{L-1} m_i(k) \exp\left(-\frac{j2\pi fk}{8192}\right) \tag{D-3}$$

with: i = 1, 2, ..., 8

f = 0, 1, ..., 4095

L = length of $m_i(k)$

The values of K_i computed from Equations D-2 and D-3 are given in Table D.1b for the eight sequences $m_i(k)$:

Table D.1b – Scaling factors K_i for the eight digital echo path models for tones

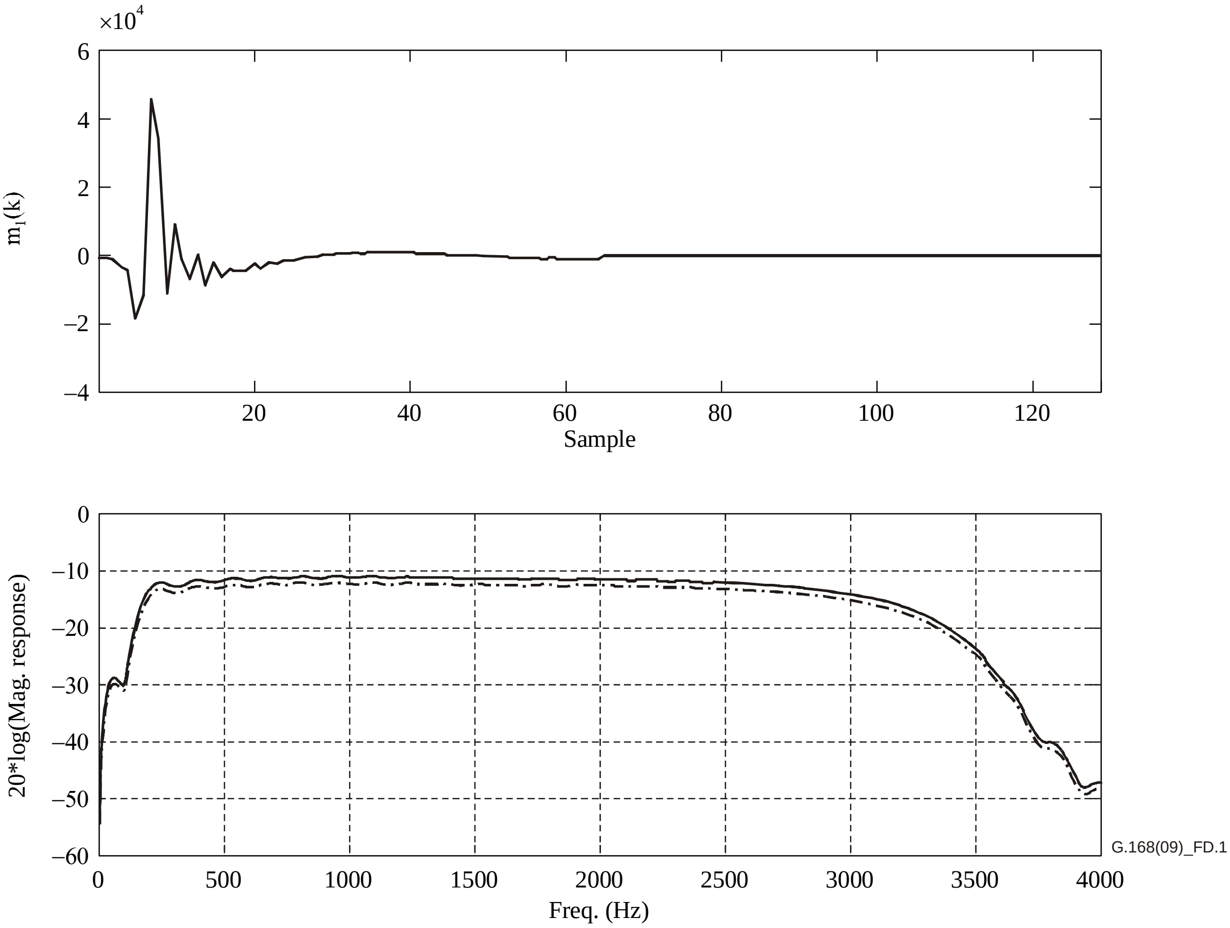
Echo path model #(i)	Scaling factor K_i	Minimum ERL for Tones (dB) (Note)
1	1.22×10^{-5}	6
2	6.78×10^{-6}	6
3	9.66×10^{-6}	6
4	1.07×10^{-5}	6
5	7.05×10^{-6}	6
6	8.60×10^{-6}	6
7	6.58×10^{-6}	6
8	4.58×10^{-6}	6
NOTE – A minimum ERL value of 6 dB should be used in the tests for all 8 echo path models. Each scaling factor is used to limit the maximum of the magnitude response to the chosen ERL value.		

D.2 Echo path models from network hybrid simulator

This clause contains four echo path models that are generated from a network hybrid simulator. Tables D.2, D.3, D.4 and D.5 show the values for $m_i(k)$ for i = 1, 2, 3, 4. They are to be read in columns. Echo path model $m_1(k)$ has short dispersion, $m_2(k)$ has median-short dispersion, $m_3(k)$ has median-long dispersion and $m_4(k)$ has long dispersion. The corresponding impulse responses and magnitude responses are shown in Figures D.1, D.2, D.3 and D.4. As an example, an ERL of 12 dB is chosen for the plots of magnitude responses.

Table D.2 – $m_1(k)$: impulse response of echo path model 1

–436	46150	390	–3948	–1098	745	1033	899	73	–512	–772
–829	34480	–8191	–2557	–618	716	1091	716	–119	–580	–820
–2797	–10427	–1751	–3372	–340	946	1053	390	–109	–704	–839
–4208	9049	–6051	–1808	–61	880	1042	313	–176	–618	–724
–17968	–1309	–3796	–2259	323	1014	794	304	–359	–685	
–11215	–6320	–4055	–1300	419	976	831	304	–407	–791	

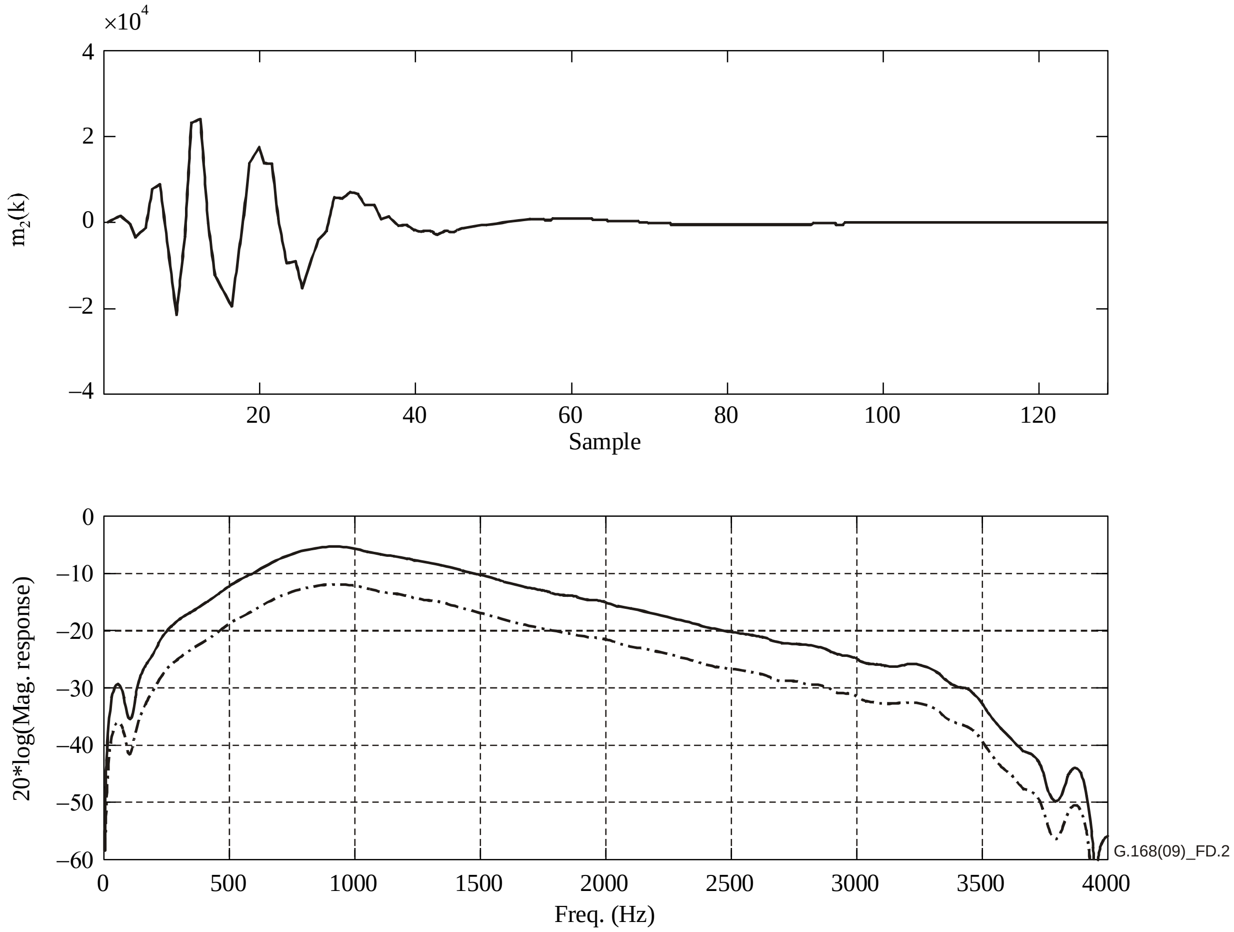


NOTE – 12 dB ERL, CSS input with scaling factor from Table D.1a (solid line), tones(s) input with scaling factor from Table D.1b (dashed line).

Figure D.1 – Impulse response and magnitude response of echo path model 1

Table D.3 – $m_2(k)$: impulse response of echo path model 2

–381	–21370	13509	–3858	1316	–1468	789	658	–331	–479	–249
658	–5307	17115	–1979	–693	–1221	954	476	–347	–479	–216
1730	23064	13952	6029	–759	–842	756	377	–430	–512	–249
–51	24020	13952	5616	–1517	–463	839	377	–314	–479	–265
–3511	1020	97	7214	–2176	–298	872	262	–430	–397	–166
–1418	–12374	–9326	6820	–2028	–68	1020	97	–463	–430	–232
7660	–16296	–9046	3935	–2654	64	789	–68	–463	–397	
8861	–19524	–15208	3919	–1814	493	822	–183	–414	–298	
–8106	–7480	–9853	921	–2077	723	558	–232	–381	–265	

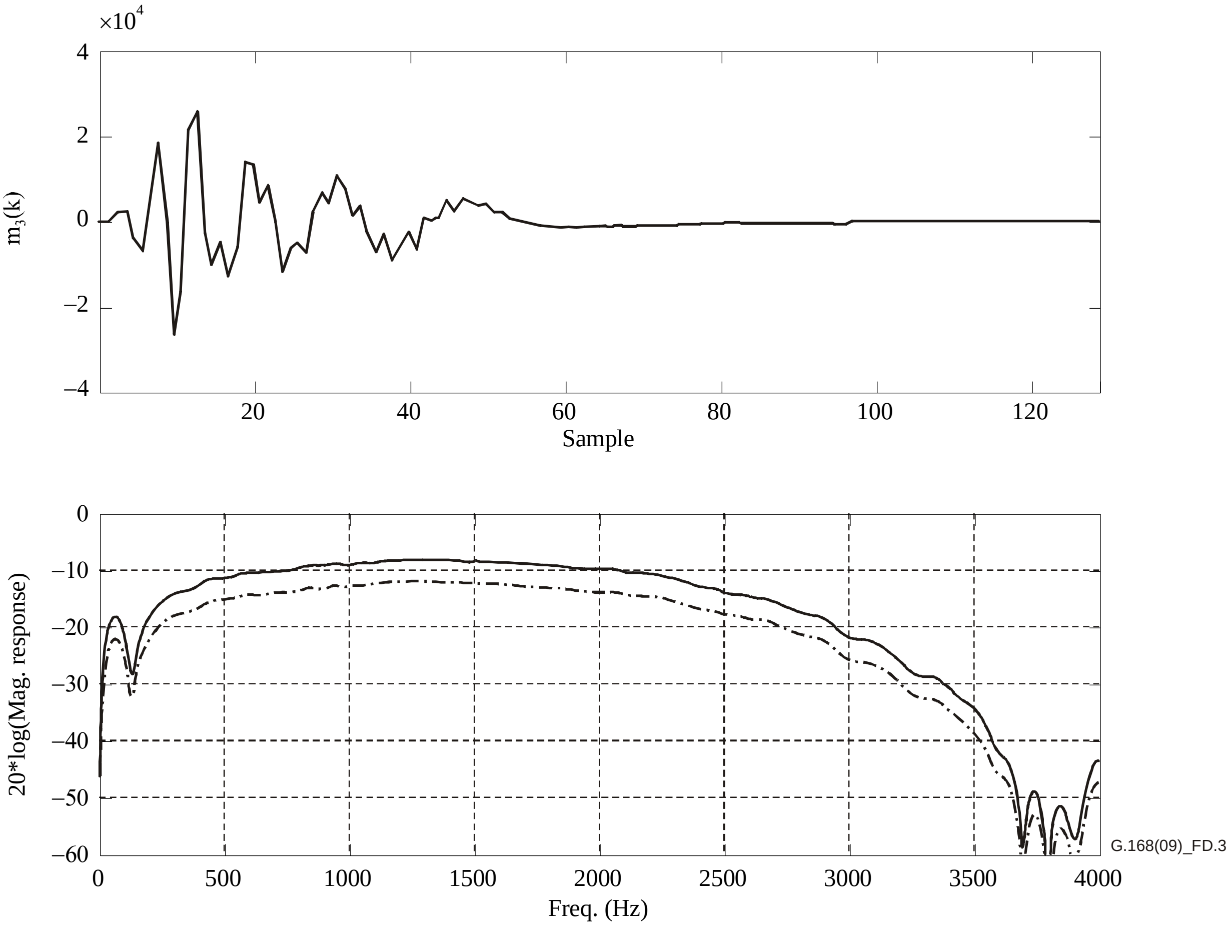


NOTE – 12 dB ERL, CSS input with scaling factor from Table D.1a (solid line), tones(s) input with scaling factor from Table D.1b (dashed line).

Figure D.2 – Impulse response and magnitude response of echo path model 2

Table D.4 – $m_3(k)$: impulse response of echo path model 3

–448	–26261	14164	3271	–3101	2172	–139	–1066	–814	–233	–390
–436	–16249	13467	6566	–9269	5387	–573	–1020	–871	–333	–482
2230	21637	4438	4277	–6146	4598	–1100	–1100	–734	–356	–459
2448	25649	8627	11131	–2553	3535	–1157	–1008	–642	–390	–482
–4178	–2267	456	7562	–6272	4004	–1180	–1077	–562	–310	–551
–7050	–10311	–11879	1475	811	2311	–1455	–1088	–356	–265	–573
5846	–4693	–6352	3728	124	2150	–1123	–917	–379	–368	
18581	–12690	–5104	–3525	788	1017	–1386	–917	–345	–310	
2322	–7428	–7496	–7301	5147	330	–1123	–963	–230	–310	

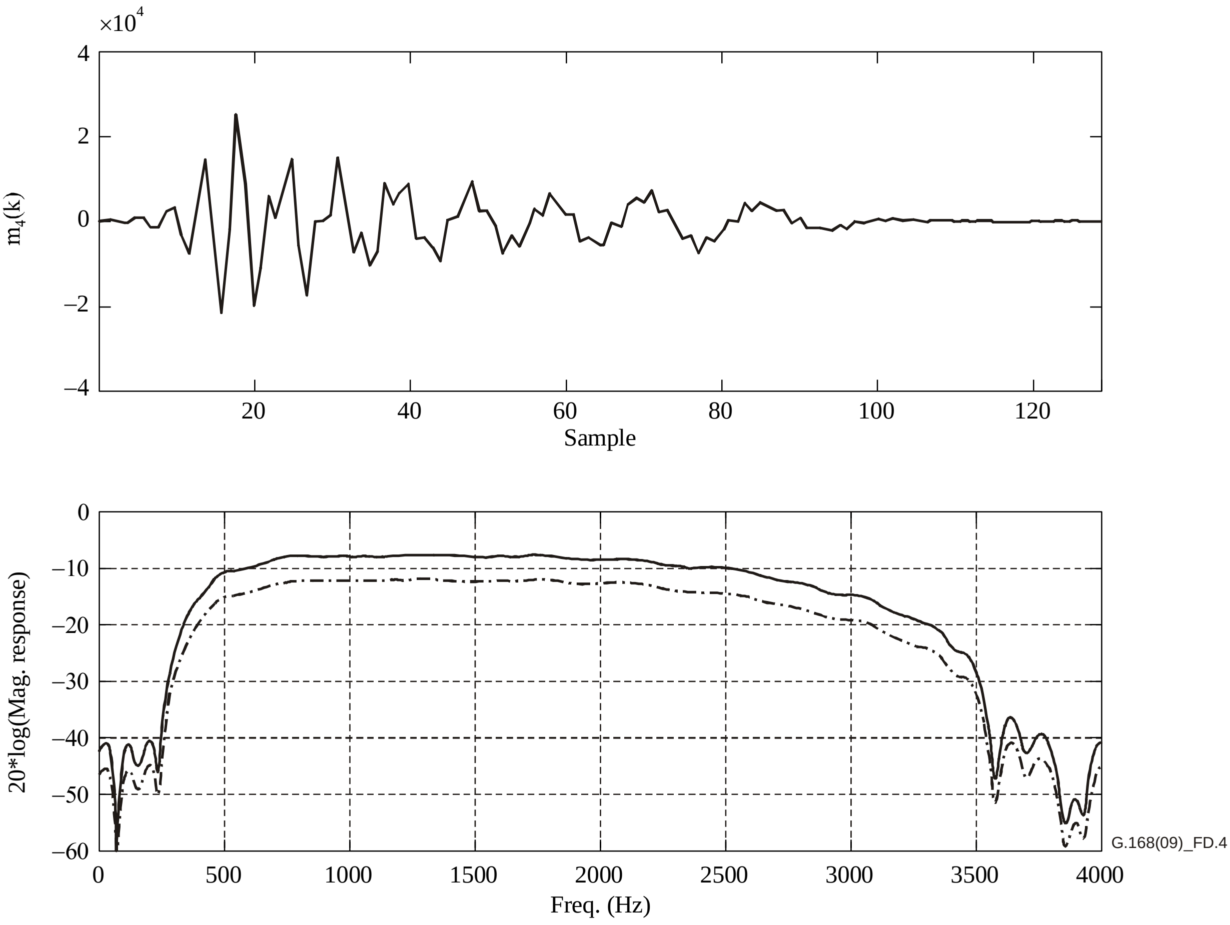


NOTE – 12 dB ERL, CSS input with scaling factor from Table D.1a (solid line), tones(s) input with scaling factor from Table D.1b (dashed line).

Figure D.3 – Impulse response and magnitude response of echo path model 3

Table D.5 – $m_4(k)$: impulse response of echo path model 4

160	4041	14624	6850	2393	1592	2654	4617	–220	203	–57
312	14484	–6975	3944	2784	–4752	–881	3576	–306	–111	–24
–241	–1477	–17156	6969	–892	–3646	–4113	2382	257	95	30
–415	–21739	–187	8694	–7366	–5207	–3244	2839	615	–79	–68
897	–4470	149	–4068	–3376	–5577	–7289	–404	225	30	84
908	25356	1515	–3852	–5847	–501	–3830	539	561	84	–155
–1326	11458	14907	–5793	–2399	–1174	–4600	–1803	8	–13	–68
–1499	–19696	4345	–9371	3011	4041	–2508	–1401	344	–68	19
2405	–11800	–7128	453	1537	5647	431	–1705	127	–241	
3347	5766	–2757	1060	6623	4628	–144	–2269	–57	–68	
–3624	789	–10185	3965	4205	7252	4184	–783	182	–24	
–7733	6633	–7083	9463	1602	2123	2372	–1608	41	19	



NOTE – 12 dB ERL, CSS input with scaling factor from Table D.1a (solid line), tones(s) input with scaling factor from Table D.1b (dashed line).

Figure D.4 – Impulse response and magnitude response of echo path model 4

D.3 Echo path models measured from telephone networks in North America

This clause presents four realistic digital echo path models measured from telephone networks in North America. The measurement method for the echo path models is described in clause II.2.2.1.

Clause II.4 also includes the statistical characteristics (magnitude response, dispersion time width and the number of reflections) of echo paths measured in North America.

Tables D.6, D.7, D.8 and D.9 below tabulate the four echo path models. The numbers are read in columns. The impulse responses as well as the magnitude responses of the models are shown in Figures D.5, D.6, D.7 and D.8 respectively.

The echo path model shown in Table D.6 has a single reflection with a dispersion width of about 6 ms. This model has a frequency domain characteristic very close to the mean of the measured echo path magnitude responses. This kind of echo path occurs most often in the measurements.

The echo path model shown in Table D.7 has a single reflection. It has a longer dispersion width of about 9 ms. The long dispersion width is due to the spectral peak at about 250 Hz.

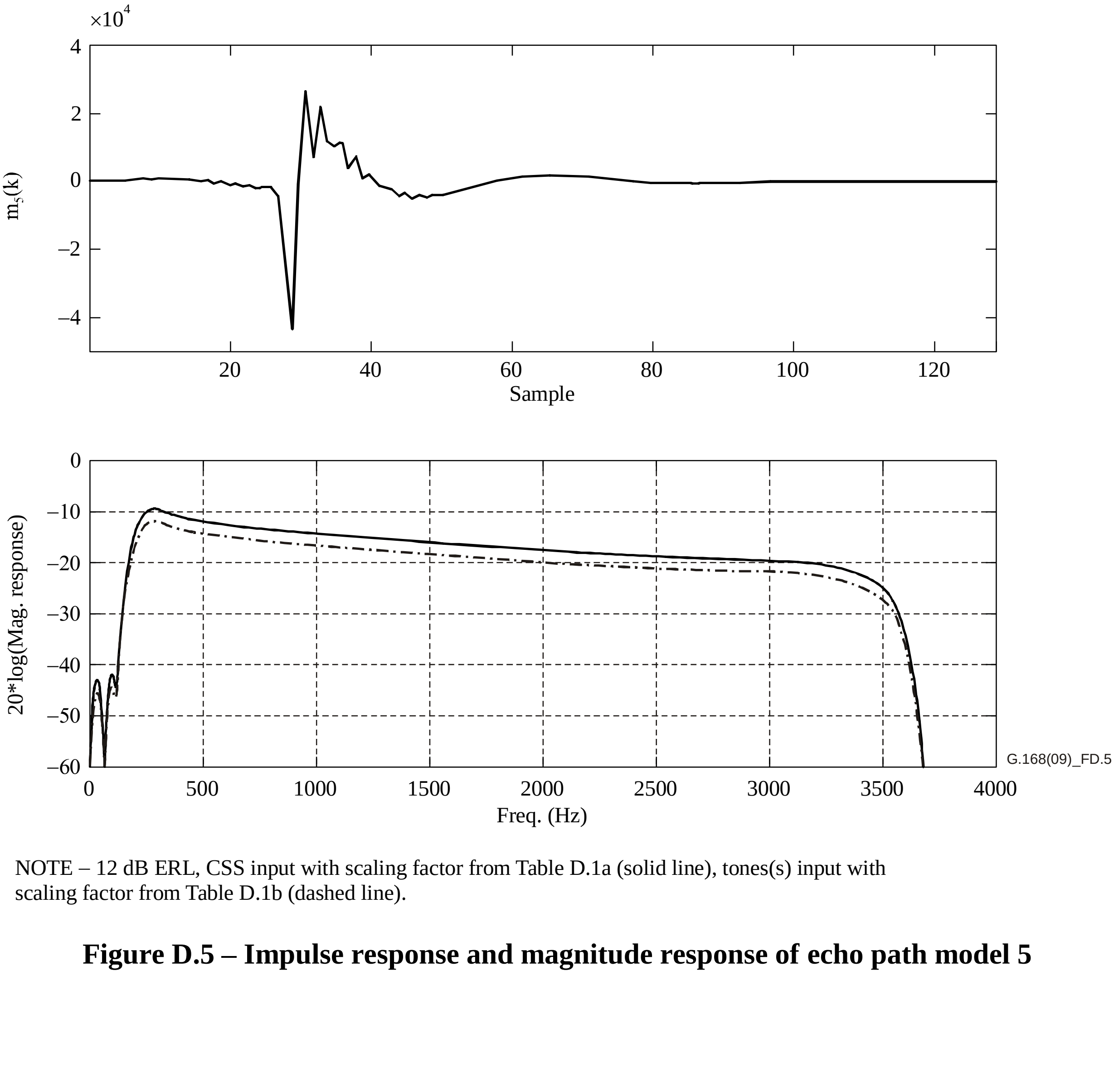
The echo path model shown in Table D.8 has double reflections. The impulse responses from the two reflectors are overlapped with each other. The dispersion width of this model is about 6 ms.

Note that due to the specific magnitude response of echo path model m_7, it may not be an effective model for Tests 10A and 10B when the NLP is disabled. This is, coincidentally, due to notches in the magnitude response at the frequencies used in the tests. Model m_8 in Table D.9 is an alternative double-reflection model that may be used for Tests 10A and 10B, since it does not contain notches in the magnitude response at the frequencies used in these tests.

The impulse responses and magnitude responses of the four echo path models are shown in Figures D.5, D.6, D.7 and D.8. As an example, an ERL of 12 dB is chosen for the plots of magnitude responses.

Table D.6 – $m_5(k)$: impulse response of the echo path model 5

293	896	20	–22548	3889	–5022	–1608	1640	733	–513	–404
268	604	–938	–43424	7241	–4039	–645	1901	665	–473	–344
475	787	–523	2743	925	–4842	–495	1687	323	–588	–290
460	561	–1438	25897	2018	–4104	279	1803	221	–612	–202
517	538	–1134	7380	–821	–4089	471	1543	–14	–652	–180
704	440	–1887	21499	–2068	–3582	947	1566	–107	–616	–123
581	97	–1727	11983	–2236	–2978	1186	1342	–279	–566	
879	265	–1698	10400	–4283	–2734	1438	1163	–379	–515	
573	–385	–4266	11667	–3406	–1805	1669	963	–468	–485	

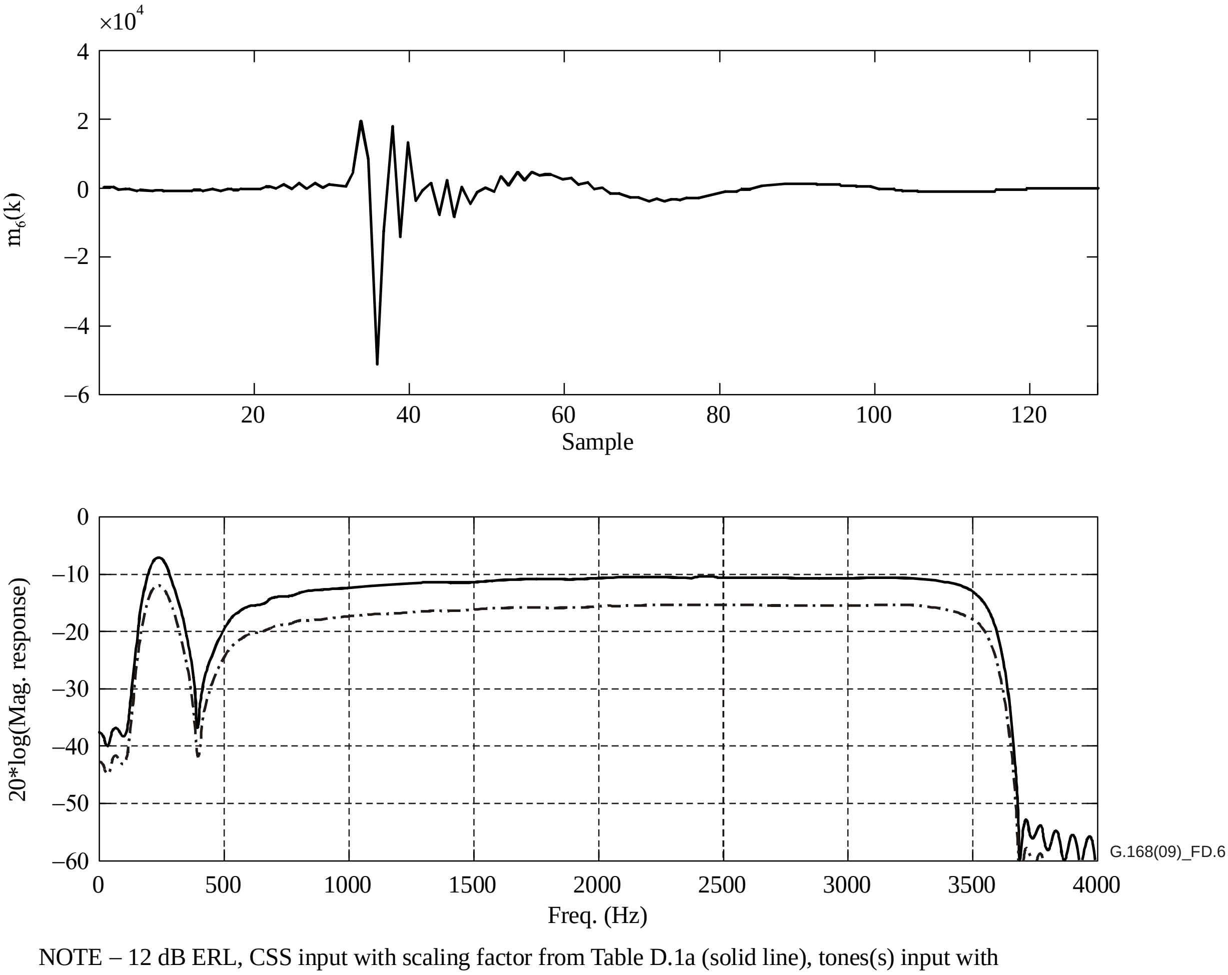


NOTE – 12 dB ERL, CSS input with scaling factor from Table D.1a (solid line), tones(s) input with scaling factor from Table D.1b (dashed line).

Figure D.5 – Impulse response and magnitude response of echo path model 5

Table D.7 – m6(k): impulse response of the echo path model 6

29	–450	158	19522	2469	5025	–1117	–1956	1724	482	–838
109	–105	1341	8421	–7994	3946	–2134	–1539	1871	289	–837
–83	–503	195	–50953	490	4414	–2547	–1239	1767	54	–834
198	145	1798	–9043	–3860	4026	–2589	–570	1802	–137	–740
–294	–490	344	18046	–837	3005	–3310	–377	1630	–321	–673
–135	267	1845	–13553	490	3380	–2778	251	1632	–490	–581
–415	–231	629	13336	–636	1616	–3427	331	1379	–638	–493
–202	340	1604	–3471	3682	2007	–2779	964	1271	–764	–436
–444	77	1182	–107	1141	158	–3116	1177	1063	–836	–327
–337	343	940	1788	5019	388	–2502	1449	856	–800	–201
–313	783	5163	–7409	2635	–1198	–2399	1564	711	–859	



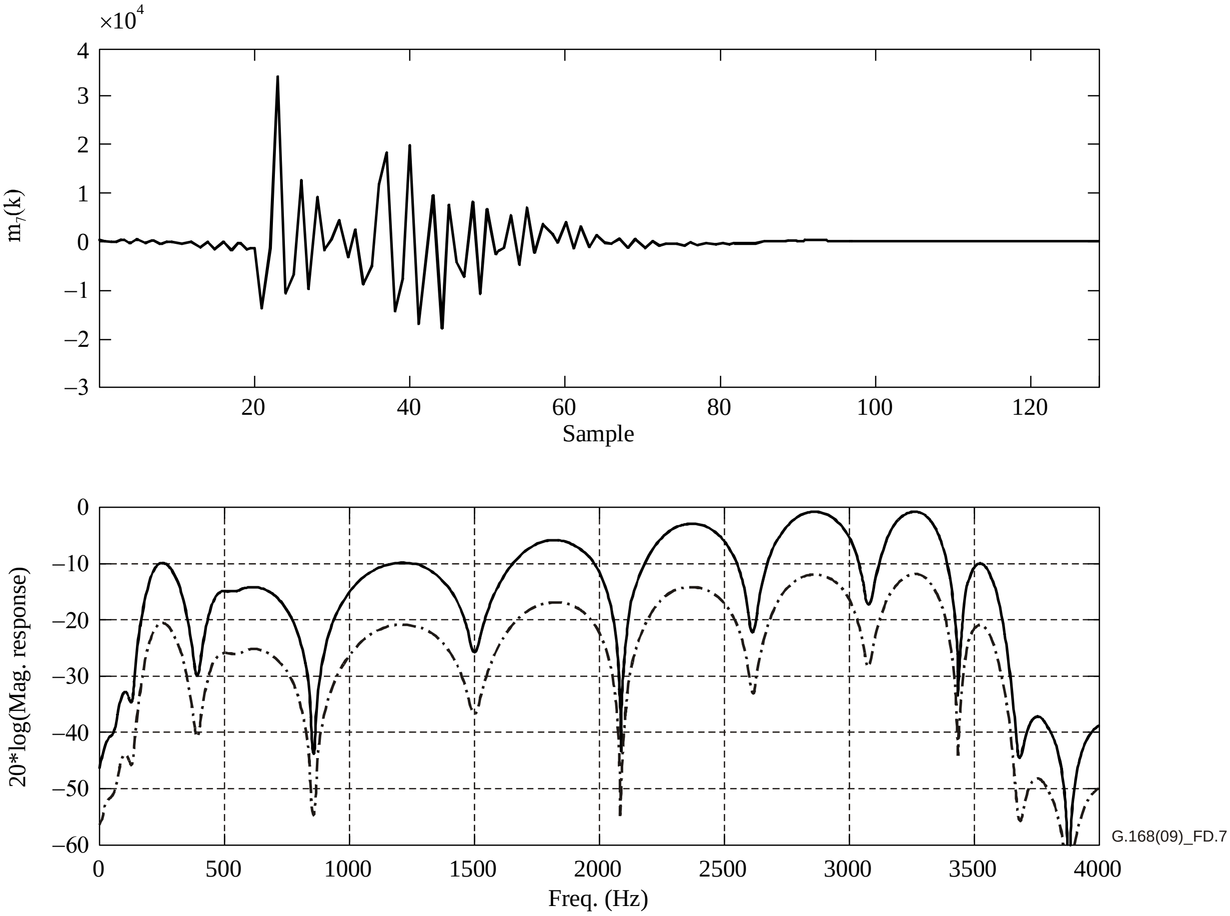
NOTE – 12 dB ERL, CSS input with scaling factor from Table D.1a (solid line), tones(s) input with scaling factor from Table D.1b (dashed line).

Figure D.6 – Impulse response and magnitude response of echo path model 6

Note that due to the specific magnitude response of echo path model m_7, it may not be an effective model for Tests 10A and 10B when the NLP is disabled. This is, coincidentally, due to notches in the magnitude response at the frequencies used in the tests.

Table D.8 – $m_7(k)$: impulse response of the echo path model 7

258	–343	–1601	8950	18072	–4342	6868	1239	–415	–325	365
–111	–596	–1389	–1574	–14410	–7415	–2195	2	–372	–245	303
337	–177	–13620	758	–7473	7929	3425	–427	–769	–255	251
–319	–1187	–720	3526	19836	–10726	1969	596	–183	–60	230
347	–52	33818	–3118	–16854	6239	–109	–1184	–785	35	209
–434	–1781	–10683	2421	–3115	–2526	3963	551	–270	218	179
192	–147	–6742	–8966	9483	–1317	–1275	–1244	–659	149	
–450	–1959	12489	–4901	–17799	5345	3087	141	–377	340	
–108	–326	–9862	11385	7399	–4565	–892	–743	–523	233	

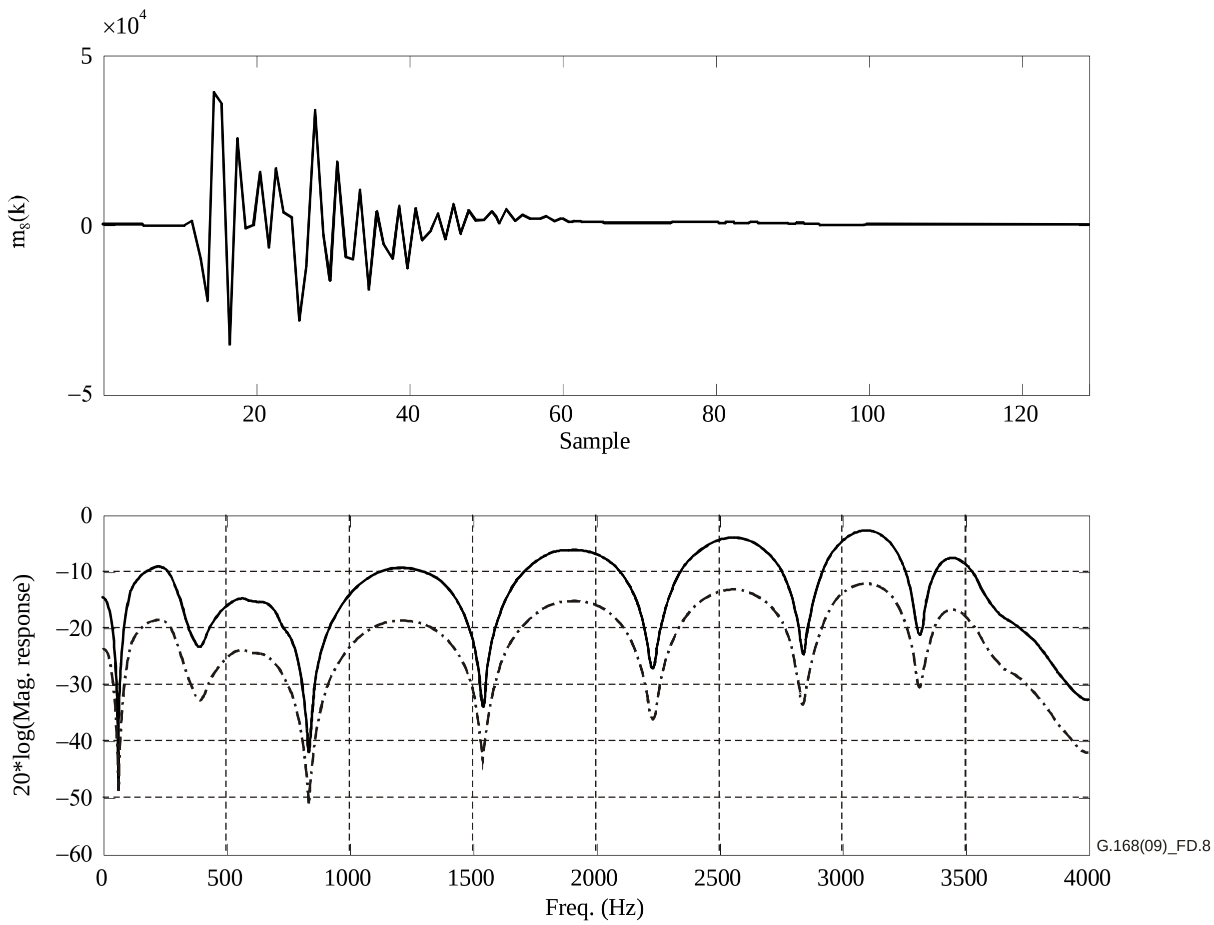


NOTE – 12 dB ERL, CSS input with scaling factor from Table D.1a (solid line), tones(s) input with scaling factor from Table D.1b (dashed line).

Figure D.7 – Impulse response and magnitude response of echo path model 7

Table D.9 – $m_8(k)$: impulse response of the echo path model 8

80	–102	–1457	33871	–5907	5553	3083	806	452	669	356
31	–26	–229	–176	–10257	–2596	1917	869	538	619	147
4	1002	15659	–16421	5336	3992	1756	471	717	500	107
42	–9250	–6786	18173	–12933	1255	2478	646	723	650	–50
42	–22562	16791	–9669	4348	1450	1027	438	850	615	–88
–61	39321	3860	–10163	–4802	4079	1871	449	756	516	–59
–81	35681	2239	9941	–1791	324	845	432	753	492	–238
–64	–35289	–28730	–19365	3035	4340	1284	473	899	427	–165
–121	25312	–11885	3592	–4433	1059	813	394	555	291	–183



NOTE – 12 dB ERL, CSS input with scaling factor from Table D.1a (solid line), tones(s) input with scaling factor from Table D.1b (dashed line).

Figure D.8 – Impulse response and magnitude response of echo path model 8

D.4 Echo path models measured from telephone networks in Europe

Measurements were made in the telephone networks in France (see clause II.2.2.2 for measurement method and clause II.4.2 for typical impulse responses observed). Three different types of echo path impulse responses were observed and they have equivalent characteristics of echo path models $m_1(k)$, $m_5(k)$ and $m_6(k)$.

Annex E

Embedded Echo Cancellers (EECs)

(This annex forms an integral part of this Recommendation.)

E.1 Scope

This annex applies to embedded echo cancellers (EECs). Its purpose is to define performance requirements for an embedded echo canceller function, together with a test set-up that may be used to perform tests.

Characteristics of an embedded echo canceller include:

– Limited or no access to the control signals used to perform ITU-T G.168 tests (e.g., h-register freeze, NLP on/off).

– Limited or no access to conventional 4-wire TDM signal ports. Examples include echo cancellers embedded in access gateways where the interface at the cancelled-end is 2-wire analogue and echo cancellers embedded in IP gateways where at least one interface is packet-based.

E.2 Definitions

This annex defines the following terms:

E.2.1 Embedded echo canceller (EEC)

An echo canceller function that may not provide access to control signals (e.g., h-register freeze) that are necessary to perform some of the tests in ITU-T G.168.

NOTE – In addition, some EECs may have no direct access to some or all of the input-output ports (R_{in}, R_{out}, S_{in}, S_{out}). An example of this is an EEC integrated with other voice processing functions in a single device.

E.3 EEC requirements

An embedded echo canceller that complies with this annex will pass all of the mandatory tests in the body of ITU-T G.168 and any optional tests that are applicable to the intended application of the device containing the EEC. Compliance or partial compliance may be tested using one of the test methodologies defined in Appendix VIII.

E.3.1 EEC performance under load

The performance requirements given in the paragraph regarding multiple channel implementations in clause 6.4.2 also apply to EECs. Specific tests for EEC performance under load are for further study.

E.4 EEC testing

Figure E.1 shows the test set-up that may be used for performing ITU-T G.168 tests according to this annex. This test set-up requires clear channel access via ITU-T G.711 coding to R_{in}, R_{out}, S_{in} and S_{out} ports of the EEC.

NOTE 1 – If an echo canceller is embedded with other signal processing functions, the other functions should be disabled while performing these tests. It is also recommended that further testing is performed with the other signal processing functions enabled to evaluate their effect on overall performance.

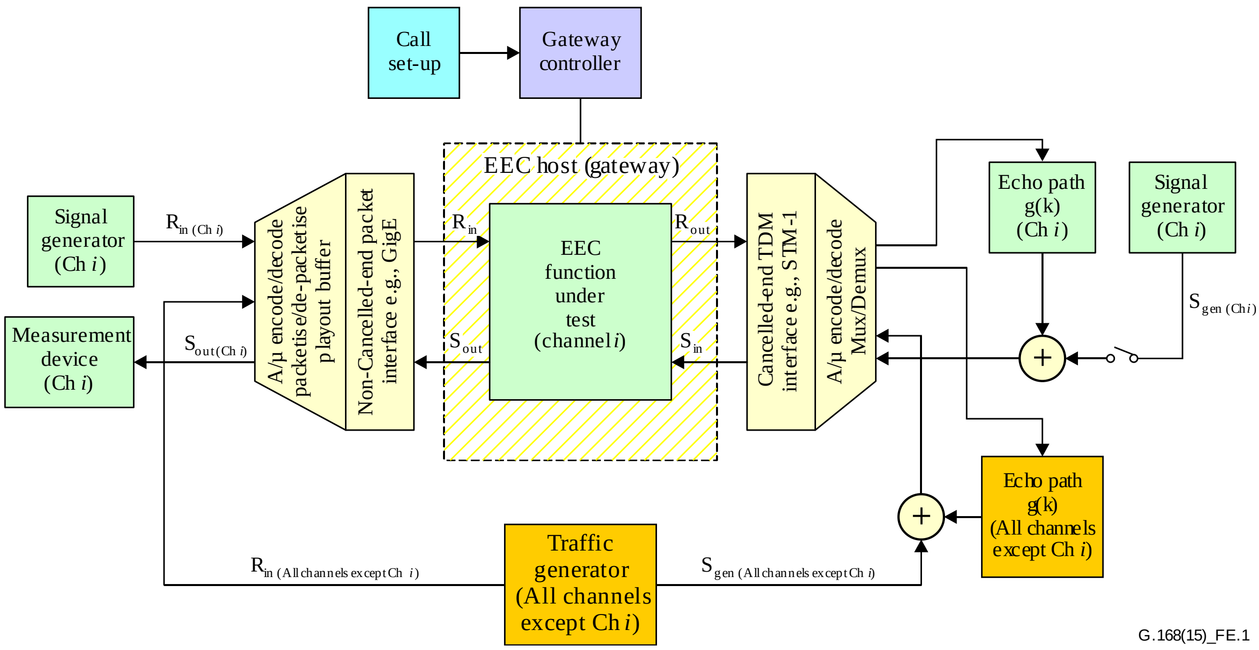


Figure E.1 – Test set-up for use with EECs

NOTE 2 – The test set-up in Figure E.1 is not applicable to access gateways, where the interface is 2-wire analogue.

NOTE 3 – Call set-up may either be performed via the gateway controller (as shown) or by any other means (e.g., using permanently "nailed-up" connections via a special "Test Mode").

NOTE 4 – ITU-T G.168 tests are performed on a single channel only. The other channels are loaded using the Traffic Generator.

NOTE 5 – g(k) represents any of the echo path models depicted in Figure 6 or Figure 7 of this Recommendation.

Possible approaches are given in Appendix VIII for testing an embedded echo canceller, depending on what control signals are available.

E.5 Special considerations

For the purpose of evaluating echo cancellation performance, the embedded speech codec should be set to the highest coding rate or the best quality mode when generating the R_{in} signal for the testing of the embedded echo canceller. Any voice activity detection (VAD), discontinuous transmission (DTX) or other silence compression functions should also be disabled.

Appendix I

Guidance for application of echo cancellers

(This appendix does not form an integral part of this Recommendation.)

I.1 Scope

Echo cancellers are adaptive signal processors used to control echo; they are expected to replace echo suppressors in modern telecommunication networks. Echo cancellers are increasingly present on nearly every long distance connection and may be encountered singly or in tandem on a given connection. The purpose of this appendix is to:

- explain the general principles of operation of echo cancellers;
- identify a limited set of application rules and the constraints under which echo cancellers operate;
- identify how echo cancellers may affect the perceived quality of speech and the quality of voiceband data;
- explain the effects of high-level speech on echo cancellers;
- provide network and service evolutionary considerations;
- offer some considerations of echo canceller performance during double talk;
- provide some guidelines on the use of parameters for testing echo cancellers.

I.2 Echo control in the public switched telephone network

In the past, echo suppressors were used to control echo in long distance networks. Today, however, the echo canceller is the device of choice. While public switched telephone network (PSTN) planners and designers typically deploy the most current and modern technologies, it should be understood by modem designers, end users, and others, that for the foreseeable future the worldwide embedded plant may include some older echo control technologies on some connections. For example, connections through the PSTN may include some combinations of ITU-T G.164 analogue or digital echo suppressors, ITU-T G.165 analogue or digital echo cancellers equipped with ITU-T G.164 tone disablers, and ITU-T G.165 analogue or digital echo cancellers and digital echo cancellers of this Recommendation equipped with ITU-T G.165 or ITU-T G.168 tone disablers. The following two clauses summarize the reasons for the use of echo cancellers instead of echo suppressors in modern telephone networks.

I.2.1 Echo suppressors

The principle of echo suppressors is well-known; it is summarized as follows: When speech is detected on the receive path, a very high attenuation is inserted in the send path. When double talk is detected, the send path is closed and a receive loss is inserted in the receive path. Thus, during double talk, there is no echo suppression in the send path, but the echo is more attenuated than the direct speech. Other refinements are possible, as indicated in [ITU-T G.164].

Generally, echo suppressors do not provide the same level of performance for speech, voiceband data, or facsimile as echo cancellers. Many problems can occur in the operation of echo suppressors. First, low level double talk speech could be mutilated if the echo level is high. Second, the tandeming of echo suppressors is not recommended. Third, echo suppressors may create problems for facsimile transmission. Clause 5.2 of [ITU-T G.161] contains the detailed explanation about the disadvantages of echo suppressors.

I.2.2 Echo cancellers

Echo cancellers are devices that use adaptive signal processing to reduce or eliminate echoes. Echo cancellers are placed in the 4-wire portion of a circuit, and reduce (or cancel) the echo by subtracting an estimate of the echo from the returned echo signal. Echo cancellers may operate on a single circuit or on a multiplexed facility, e.g., echo cancellers operate on a 64 kbit/s speech facility that is multiplexed into a primary rate link.

Echo cancellers are designed to:

- cancel linear echo path signals;
- refrain from cancelling the echo when requested to do so by an in-band disabling signal;
- return to an operational mode after being disabled when the in-band signal power level drops below a specified level for a specified period of time. This design allows some networks to transport voiceband data on the same speech channels. It also allows the echo canceller to re-enable during a voice call after it has been turned off erroneously (talkoff).

Echo cancellers are characterized by whether the interface path is analogue or digital, and/or whether the subtraction of the echo is by analogue or digital means. This appendix is limited to echo cancellers that have a digital input and digital subtractors.

Echo cancellers have the following main advantages over echo suppressors:

- send path transparency is improved;
- NLP hangover introduces fewer impairments;
- there is no receive insertion loss;
- echo cancellation continues during double talk;
- tandeming is possible (for well-designed echo cancellers).

Some echo cancellers are optioned to disable on the 2100 Hz tone specified in [ITU-T G.164] for echo suppressors, and some are disabled with a 2100 Hz tone with periodic phase reversals of 180° ± 25°, as specified in [ITU-T G.165] and in this Recommendation for echo cancellers. Use of ITU-T G.165 or ITU-T G.168 tone is intended to allow echo cancellers to be disabled independently of echo suppressors. Echo cancellers that respond to ITU-T G.165 or ITU-T G.168 disabling tone are not disabled by the 2100 Hz tone without phase reversal.

I.2.3 Responsibilities of modem manufacturers and end users

It is the responsibility of the modem manufacturers and end users to understand the characteristics of the network-based echo canceller fully and decide whether the echo cancellers should be enabled or disabled. If the modem manufacturers and end users decide that the network-based echo canceller functionality should be disabled, they should ensure that the terminal uses the appropriate approved methods, defined in Recommendations, to disable cancellers. Additionally, it is the end user's responsibility to ensure that terminals and private networks are designed to operate in a fashion compatible with the PSTN network-based echo cancellers. For example:

- Digital telephone sets are expected to control their own echoes, see [ITU-T G.122], [ITU-T G.131], [ITU-T P.340] and [ITU-T P.310] (the PSTN network is not responsible for cancelling acoustic echoes).
- Terminals and private networks should be designed to provide circuit extensions compatible with the design intent of the PSTN, e.g., echo paths outside the PSTN-network should be linear and time-invariant or the terminal should control its own echo.

- Either the delay of the terminal or private network should be within the operational limits of the network-based echo canceller, or the terminal/private network should control its own echo.

I.3 Application rules and operational constraints

The following items should be considered for application rules and operational constraints of echo cancellers:

- public network transmission planning;
- transmission delay and echo return loss;
- echo-path characteristics and echo-path capacity of an echo canceller;
- end user/manufacturer/private network transmission planning.

The detailed descriptions of these items can be found in ITU-T Recommendations on transmission planning aspects of echo cancellers (see [b-ITU-T G.108.2]).

I.4 Effect of cancellers on voice and data services

Network-based echo cancellers are present on connections that experience long delays. They should be designed to allow a speech channel to support voiceband data, including facsimile. This means that they should retain the capability of being disabled upon an appropriate request from customer terminal equipment. However, the modem manufacturer is responsible for determining if network-based echo cancellers should be enabled or disabled.

Full-duplex data transmission in the voiceband can occur, depending on the modem modulation scheme. New modulation schemes are being introduced, and manufacturers should determine the optimal state in which the echo canceller should be when the modem is operating, i.e., if the canceller should be enabled or disabled, or whether the call should be routed on a connection that never has an echo canceller functionality present.

The designers of facsimile terminals generated these terminals with the understanding that network providers were installing network-based echo control devices as per [ITU-T G.164], [ITU-T G.165] and this Recommendation. Thus, PSTN network planners were expected to continue to evolve the network in such a way that it would not knowingly prevent the continued carriage of a permissive voiceband data/facsimile service.

ITU-T V.32 modems, in contrast, use the same band of frequencies in both directions and achieve duplex operation through the use of an integrated echo canceller. The echo canceller integrated in this voiceband data modem is not to be confused with the network echo cancellers that conform to [ITU-T G.165] and to this Recommendation because the performance requirements for each are very different.

Details of the interaction of echo control with voiceband data and facsimile transmission can be found in clause 5.2.1 of [ITU-T G.161].

I.5 High-level speech

I.5.1 Introduction

A number of sources could produce high speech levels in the network. In hands-free telephones, for example, the microphone may allow high speech levels to be generated. With this perspective in mind, [ITU-T G.165] was modified in 1993 to include an overload test (Test No. 8) at levels exceeding 0 dBm0 and to increase the maximum test levels from –10 dBm0 to 0 dBm0.

The presence of high speech levels may cause increased non-linearities that would degrade the performance of some echo cancellers, especially echo cancellers that have not been implemented in a fully digital manner. Another area in which high signal levels may cause difficulty is in the double talk detection and non-linear processor control circuits. These are discussed in the following two clauses.

I.5.2 Double talk detection and activity detection

The performance of echo cancellers is very dependent on the activity detection and double talk detection algorithms used. For example, if double talk is not recognized quickly, the cancelled end speech masks the residual echo that is used to update the impulse response model of the echo canceller.

The following items are for further study:

- the effect of activity detection algorithms for low bit rate coders;
- the effect of double talk detection in the presence of high signal levels.

New echo canceller requirements for echo canceller design may result.

I.5.3 Effects of low bit rate coders

For network planning purposes it is useful to know what degradation low-bit rate coders in the echo path of an echo canceller may cause. An echo canceller may provide a certain amount of echo return loss enhancement.

This topic is for further study.

I.5.4 Effects of a non-linear echo path

The theory of echo cancellation assumes that the echo path is linear and time-invariant. An echo canceller will have limited ability to cancel echo in the presence of a non-linear echo path resulting from clipping and non-linear distortion in the echo path between R_{out} and S_{in}. More information about the effects of non-linear echo path can be found in clause 5.2.5.1 of [ITU-T G.161].

I.5.5 Guidelines for R_{out} usage in echo cancellers

The configuration in which the same signal feeds both R_{in} and the echo path may result in degraded performance if R_{out} is not digitally equivalent (bit for bit) to R_{in} under all signal conditions. The signal R_{rcv} used internally by the echo canceller after passing through the R_{in} port can be used as the source signal for the echo path. Therefore, it is recommended that R_{out} (which is used to drive the echo path) should be digitally equivalent to R_{rcv}.

I.6 Network and service evolutionary considerations

I.6.1 Bit transparency of echo cancellers

[ITU-T G.165] was amended in 1993 to make it clear that a 2100 Hz disabling tone with phase reversals should cause the echo canceller to disable and provide an analogue clear-channel signal path. In other words, a tone between 300 Hz and 3400 Hz should pass with its power level and frequency unaltered through the echo canceller, but 64 kbit/s bit-transparency is not guaranteed (see clause 3.3 of [ITU-T G.165]). It is noted that 64 kbit/s transparency is achievable and is implemented in some echo cancellers, but to remain in that state, the in-band power level should remain above a predefined power level.

If cancellers are to be applied to trunks and disabled by use of a "switch to echo canceller signalling channel", the canceller should support a 64 kbit/s clear channel capability, if such capability is to be provided.

I.6.2 Convergence speed

High speed of convergence is desirable to reduce echo during initial acquisition, and to minimize echo when the echo path is changing. Some echo cancellers generate noise in trying to continuously adapt to the echo path. This may be related to adaptation speed. The effect is very noticeable and annoying, especially during double talk, if the adaptation process is not suspended. For some echo canceller implementations, as the speed of adaptation is increased beyond the optimum speed, the accuracy of the transfer function after adaptation becomes poorer. High speed of convergence is desirable for initial acquisition, while lower convergence may be needed for subsequent tracking, since the echo transfer function changes very slowly. The need of high convergence speed when time varying components are in the echo path is still under study.

I.6.3 Acoustic echo control and environments

Acoustic echo control is becoming an important issue due to the increase in hands-free telephone sets. Although there is some commonalty between issues encountered for acoustic echo cancellation and network echo cancellation, there are also many differences. The issues of level points, natural echo path loss (or gain), degree of loss-switching, as well as level and/or type of singing (howling) protection are all important to a study of acoustic echo cancellers. In addition, it is important that an acoustic echo canceller is capable of working in harmony with a network-based electric echo canceller.

I.6.3.1 References to acoustic echo control

The following materials are extracted from the Recommendations related to acoustic echo control.

Information about acoustic echo controllers and speech enhancement devices can be found in clause 10 of [ITU-T P.340].

I.6.3.1.1 Weighted terminal coupling loss (TCLw) (from clause 6.1 of [b-ITU-T P.341])

The TCLw measured from the digital input to digital output shall be at least 35 dB when corrected to the nominal values of SLR and RLR as specified in clauses 4.1 and 5.1 of [b-ITU-T P.341], respectively. If a volume control is provided, the requirement applies at a setting as close as possible to the nominal value of RLR as specified in clause 5.1 of [b-ITU-T P.341].

I.6.3.1.2 Terminal coupling loss (from clause 6.1 of [b-ITU-T P.342])

The weighted terminal coupling loss (TCLw) should be greater than 40 dB when measured under field conditions and with SLR normalized to SLR = +13 dB and RLR = +2 dB. For example, if the measured TCLw is 42 dB, the measured SLR is +16 dB and the measured RLR is 0 dB, then the normalized value of TCLw = 42 dB + (13 – 16) dB + (2 – 0) dB = 41 dB.

However, in order to meet ITU-T G.131 talker echo objective requirements, a TCLw greater than 45 dB is desirable and should be striven for.

NOTE – The perceived echo impairment, by the person at the opposite end of the connection from a telephone set that has a TCLw of less than 45 dB, is a function of the magnitude of the talker echo signal as well as the talker echo path delay. A telephone set that has a TCLw of less than 45 dB will provide an echo signal that becomes more disturbing as the talker echo path delay increases. Thus, a telephone set that has a TCLw of less than 45 dB may provide satisfactory performance on low delay connections while the same may not be true for connections that have a long delay.

It is assumed that this requirement is met if TCL and TCLw, respectively, meet the values of Table I.1 with the receive volume control in its maximum setting.

Table I.1

TCL (1/3-octave band)	TCLw
> 25 dB	> 35 dB
NOTE – These values assume no other echo control in the connection.	

If information is available in the terminal about the one-way transmission time of the connection, and if the terminal operates in double talk, then the limits defined in Table I.2 may apply.

Table I.2

	One-way transmission time	TCLw
Single talk	≤ 10 ms	≥ 25 dB
Double talk	≤ 10 ms	≥ 19 dB[a)]
a) To achieve MOS ≥ 4. Further information is found in [ITU-T P.340].		

I.6.4 Comfort noise

As the telephone network migrates to more digital connections, it becomes more likely that the echo path will be analogue while the long distance connections path will be digital. One consequence is that the long distance path has a low idle channel noise while the echo path has a higher idle channel noise. This in turn leads to a situation called "noise modulation". When the NLP operates, the talker "hears" the idle channel noise of the digital long distance path, but when the NLP releases, the talker "hears" the idle channel noise of the echo path and the far-end environmental noise. Thus, the talker hears intervals of speech with background noise followed by intervals of silence, which can be very annoying in some instances.

There are two known approaches for comfort noise. The first solution is to insert pseudo-random noise during the silent interval. The second solution is to allow some of the background or idle channel noise to pass through the NLP.

NOTE – Based on the input from ITU-T experts on "end-to-end transmission performance of networks and terminals":

- Artefacts due to comfort noise insertions should be minimized.
- The noise used should match the background noise, both in frequency content and level (discussions indicate that this may not be a good idea for high and low noise levels).
- Level of the inserted noise should match that of the background noise; appropriate level measurements and adjustments should be done using dBm0p.
- The time course of changes in the level of the inserted noise should match, as closely as possible, the level changes that are occurring in the background noise.

I.7 Considerations regarding echo canceller performance during double talk

I.7.1 Introduction

A double talk situation (as the name suggests) could occur when both signals present at the input of an echo canceller have characteristics of active speech.

The CSS, which simulates double talk, consists of a burst (of constant energy) and a real pause. However, it was shown that a better double talk signal could be achieved with a signal in which the two bursts with high signal energy are identical to the original one, while the pause is filled up with a shortened CSS consisting of a voiced sound, a noise sequence and a real pause. Figure I.1 shows the modified double talk signal with the sequence length of 800 ms.

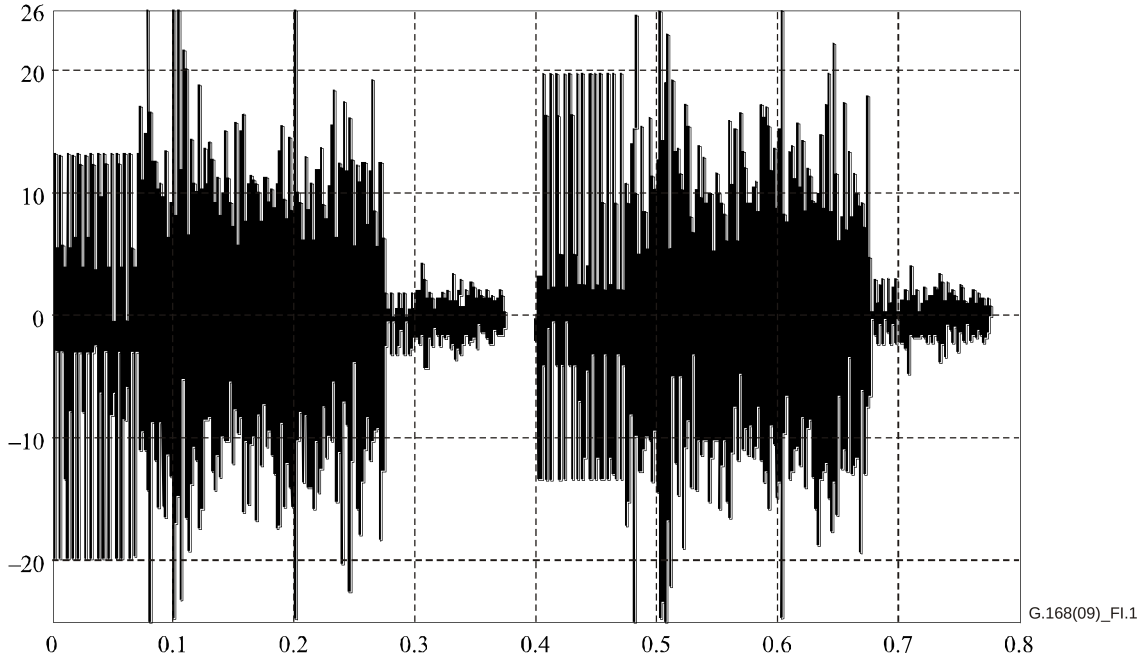


Figure I.1 – Modified double talk signal

I.7.2 Double talk parameters

The performance of echo cancellers under conditions of double talk is determined by many parameters. From recordings and listening tests, the following aspects can be derived:

- One of the most noticeable degradations when listening to the double talk signal is caused by the erroneous insertion of the NLP during continuous speech.
- Speech gaps caused by the NLP as mentioned above during continuous speech seem to be more annoying than clipping at the beginning of a double talk sequence (first word of the first sound).
- The detectability and annoyance of echo signals during double talk depend on the echo signal level and echo signal sound.
- Echo cancellers behave in a different way if double talk occurs at the beginning of adaptation or after full convergence.
- Based on the input from ITU-T experts on "end-to-end transmission performance of networks and terminals": Temporal clipping (i.e., syllable clipping or mutilation) introduced by the NLP should be less than 64 ms and less than 0.1% of active speech.

I.7.3 Analysis of technical parameters that influences performance under double talk conditions

The following parameters need to be taken into account when defining a test signal and the measurement procedure:

- signal levels at the R_{in} and S_{gen} port (receive signal and double talk signal);
- level ratio and time pattern of both signals at the R_{in} and the S_{gen} port;
- time of double talk (convergence status of echo canceller);
- duration of double talk.

The performance of the echo canceller itself is determined by technical parameters such as:

1) sensitivity of double talk detection;
2) threshold level of double talk detection (insertion of NLP, possible adaptive control);
3) reliability of double talk detection;
4) switching time of NLP;
5) double talk detection hangover time;
6) frequency characteristics of the residual echo signal loss measured between the R_{in} and S_{out} port (ERL versus frequency, "sound" of echo signal);
7) divergence during double talk.

Again these influencing parameters can be separated into different groups:

- points 1-3 are determined by the performance of double talk detection (sensitivity, reliability);
- the switching characteristics of the NLP determine points 4 and 5;
- points 6 and 7 (frequency characteristics, i.e., ERL vs. frequency and divergence) depend on the filter algorithm.

A suitable measurement procedure to evaluate double talk performance requires a suitable measurement sequence. A combination of two composite source signals was derived to reproduce typical speech double talk sequences. Both signals are described in [ITU-T P.501]. The length of the measurement CSS is 700 ms, the second CSS, which simulates the double talk fed into the echo path, has a duration of 800 ms. Due to their different sequence length, the level relationships on both echo cancellers inputs R_{in} and S_{gen} (or) S_{in} change, if both signals are periodically repeated. The same relationships can be observed if real speech signals are used. Various measurements on different echo cancellers demonstrate that this signal combination reproduces results under double talk conditions compared to speech.

I.7.4 Conducting the double talk Tests 3A and 3B without inhibiting the adaptation

I.7.4.1 Introduction

In Tests 3A and 3B, the echo canceller is exposed to double talk for some time δt and the convergence C/divergence D is determined thereafter by removing the double talk and freezing the adaptation while the CSS remains active on the receive-in port. Thus, each experiment yields only one single point of the graph $C = C(\delta t)$ or $D = D(\delta t)$. Instead, one could measure the whole graph $C = C(\delta t)$ or $D = D(\delta t)$ in one experiment by subtracting the cancelled end double talk component sgen(k) from the signal e(k) at the send-out port. The difference $e_r(k) = e(k) - sgen(k)$ would be the residual echo, which directly leads to $C = C(\delta t)$ or $D = D(\delta t)$.

I.7.4.2 Test procedure

The test is performed using the test configuration of Figure I.2. For high levels of sgen(k), the magnitude of sum of sgen(k) and the echo, $g(k)*c(k)$, may exceed the linear range of the A/μ-law coder. As a result, the echo canceller sees the saturated value of $sgen(k)+g(k)*c(k)$. The double-talk component at the send-in port becomes:

sgen_sat(k) = codec[sgen(k) + g(k)∗c(k)] – g(k)*c(k)

Thereby the function codec(·) is defined as a linear to A/μ-law conversion followed by an A/μ-law to linear conversion. The signal sgen_sat(k) is computed by the far left blocks of Figure I.2.

Some echo cancellers contain a high- or bandpass filter in the send-path. If that is the case, sgen_sat(k) must be passed through such a filter before it can be subtracted from the send-out signal. This can be achieved in various ways. If the filter is known, it appears simplest to pass sgen_sat(k) directly through the filter. If it is unknown, one can pass sgen_sat(k) through the echo canceller while there is silence on the receive-in port (see Figure I.2). The obtained signal sgen_sat_filt(k) represents the double talk component of the send-out signal, and can be subtracted to compute the residual echo $e_r(k)$.

The requirements on the power level of $e_r(k)$ should be the same as on L_{RES}. This procedure provides a continuous (rather than discrete) view of the echo canceller behaviour throughout the double-talk course. Hence, it does not require freezing the echo canceller adaptation, and it eliminates the possibility of a potential distortion of the echo canceller measured performance. Note that the S_{out} signal is measured during the application of the double talk signal, S_{gen}.

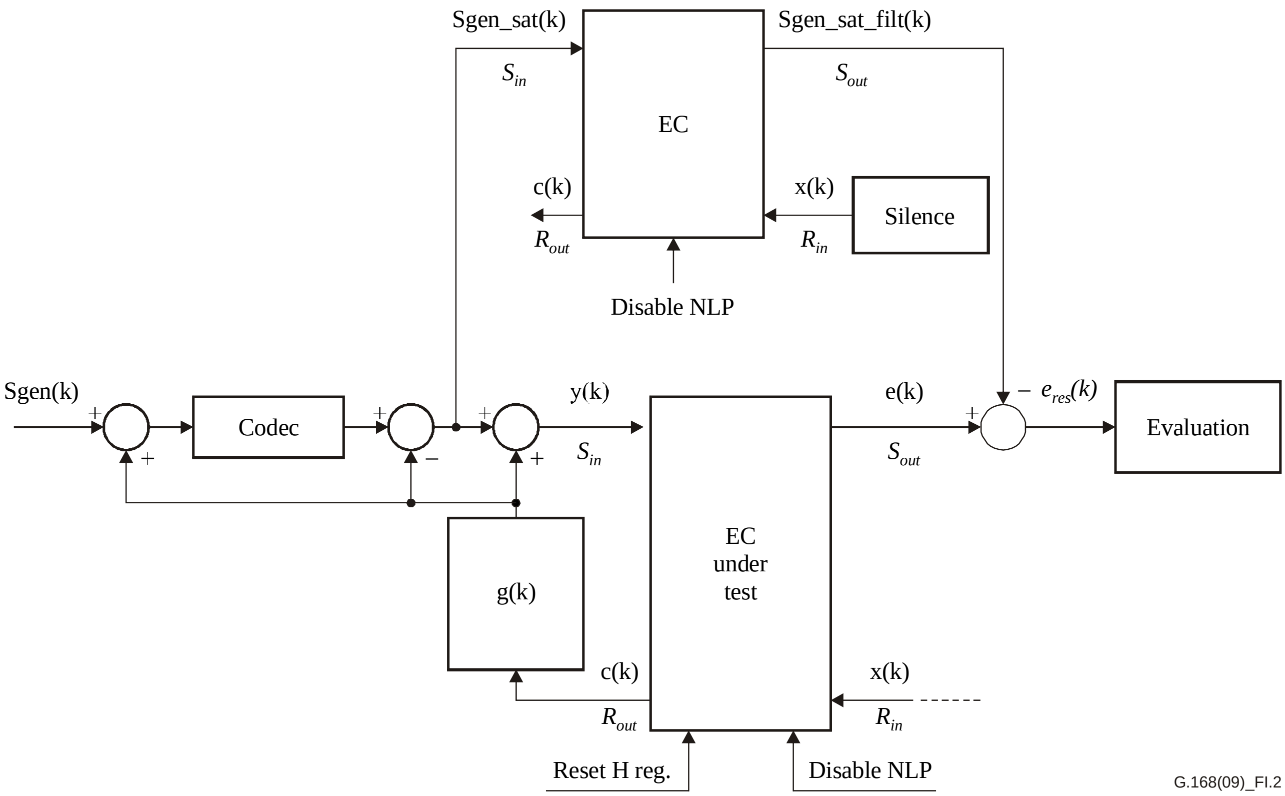


Figure I.2 – Test to subtract the double talk component from the send-out signal

I.7.5 Subjective and objective echo canceller testing

I.7.5.1 Subjective and objective echo canceller testing – Fundamentals

One of the most important points in studying the performance of echo cancellers is to investigate how objective tests correlate with subjective tests.

The subjective test procedures (conversational test, talking and listening tests, listening-only tests) were suggested for standardization in [b-ITU-T P.831]. Based on the data from subjective tests parameters, determining the transmission quality for speech echo cancellers can be identified and in a second step conclusions can be drawn about important test procedures and requirements for laboratory tests.

The following clauses summarize how subjective test results were used in order to develop objective tests.

I.7.5.1.1 Auditory test procedures

Figure I.3 demonstrates the structure and relation of the different subjective test procedures like *conversational tests, talking and listening tests and listening-only tests* together with parameters which can be assessed by these tests. The different procedures were developed as parts of the whole test set-up. The procedures were not performed in isolation but each test for a certain purpose.

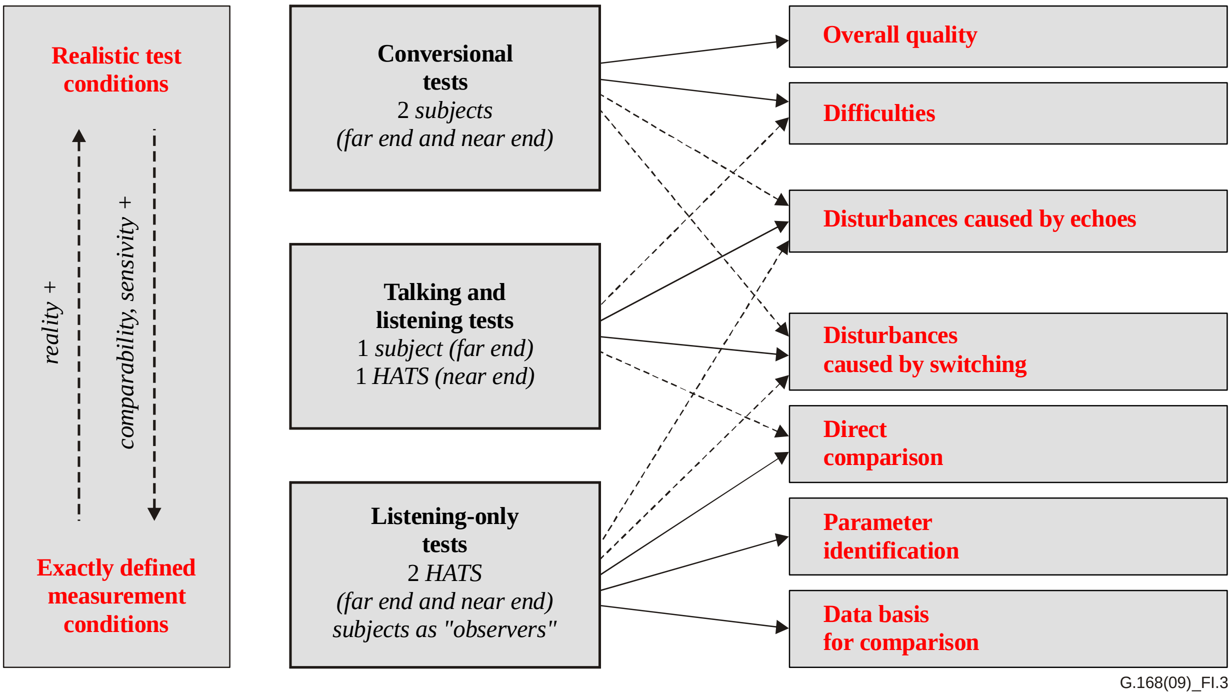


Figure I.3 – Structure of subjective test procedures for speech echo cancellers and a selection of parameters

Conversational tests play an important role in the evaluation of speech echo cancellers, since the performance of the echo canceller influences the perceived quality of the overall telephone connection. During a conversational test, subjects give their rating of the quality of the *overall* connection after a *complete* conversation, thus taking into account all of the aforementioned performance factors. Conversational tests are therefore probably the most important tests designed to evaluate the influence of speech echo cancellers on overall transmission performance.

Furthermore the conversational tests are necessary to identify those parameters, which play an important role concerning the complex parameter "*overall quality*". But on the other hand, such tests are not sensitive enough to evaluate specific transmission performance characteristics. If laboratory tests shall be developed and requirements for specific transmission parameters for echo cancellers shall be fixed (residual echo level, initial convergence, NLP behaviour, etc.), other test procedures are necessary. For this purpose, *talking and listening* tests and *listening-only* tests have been developed. If these tests are properly designed, they provide a powerful and effective method to investigate the effect which specific echo canceller parameters have on overall transmission performance. The results of these tests may therefore be directly correlated to the results of objective measurements.

I.7.5.1.2 Parameter identification by conversational tests

In 1996, conversational tests were carried out with 4 commercially available echo cancellers being at least ITU-T G.165 compliant. After the test was finished, the operators who were present in the test rooms interviewed subjects about the *overall quality rating* and the *percentage difficulty in talking or listening over the connection (% D)*. The echo path characteristics and the corresponding test results are given in Figures I.4 and I.5.

In addition to these recommended questions, all subjects who answered the question about difficulties with "Yes" where additionally asked about *the kind of difficulties* they had. At the end of the interview, each subject was asked about *the most annoying aspect* during the call they just had finished. The answers are of special interests for the operators not only for the subjective tests itself but also for the evaluation of those technical parameters which cause the problems for the subscribers. This information gives an important idea about the relevance of instrumental parameters.

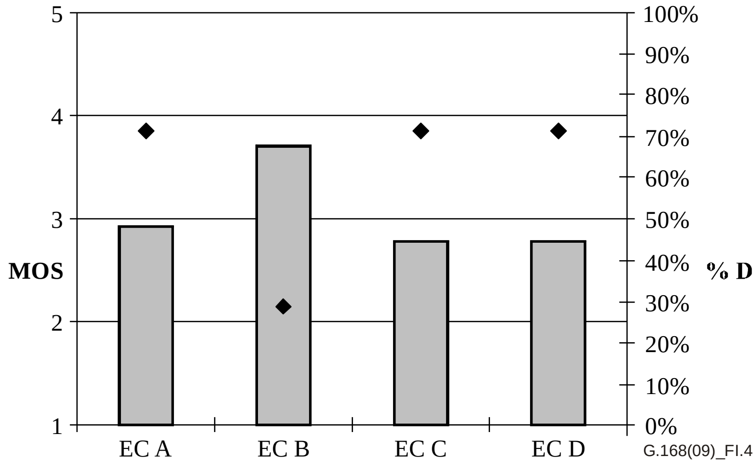


Figure I.4 – Results from *conversational tests,* overall quality mean opinion score (MOS) and % D (rhombus) for the 4 echo cancellers, ERL 7 dB, room noise level 40 dB(A), corresponding level –61 dBm0

A significant difference between echo canceller B and the other cancellers A, C and D could be analysed in Figure I.4. The main focus of these tests was not to compare different echo cancellers but to identify the parameters that determine the subjectively perceived quality. If the comments about the difficulties during the conversation are analysed, most of these quotations concern the following points:

- Audible speech clipping during double talk (implementation of the NLP). This is especially relevant for the design of Test 3C.

  Instrumental measurements based on the composite source signals demonstrate that echo canceller B in Figure I.4 shows a very good double talk performance. In connection with a high echo attenuation this leads to the best rating given through Figure I.4.

- Disturbances caused by echoes (initial convergence and residual echo). This is especially relevant for Tests 2A and 2B.

  The echo signal of EC D itself was typically characterized as "distorted" or "...like whispering". This leads to significant worse MOS values compared to echo canceller B although the double talk performance was good.

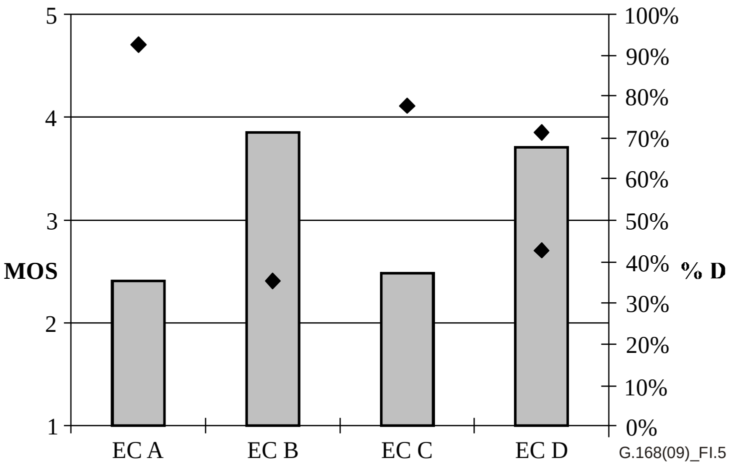


Figure I.5 – Results from *conversational tests*, overall quality MOS and % D (rhombus) for the 4 echo cancellers, 4-wire echo path, room noise level 50 dB(A), corresponding level –55 dBm0

The echo cancellers can be divided into two groups in Figure I.5. Both cancellers A and C are judged significantly worse than the cancellers B and D. Analysing the comments of subjects after the tests expressing their rating of percentage difficulty, it clearly points out two important aspects:

- Clipping during double talk as the most annoying impairment. This is as especially relevant for the design of Test 3C.

  Both echo cancellers B and D with good double talk performances in connection with this kind of echo path realization are judged significantly better.

- Quality of background noise transmission. There is currently no test to examine the quality of background noise transmission in this Recommendation.

  The modulation of background noise is audible and annoying for subjects especially if the noise disappears when they start talking and the echo canceller attenuates the send path.

Conversational tests conducted by another Administration under different network conditions also pointed out the importance of the NLP performance. Significant differences between two of the echo cancellers under test appeared in these evaluations, when the signal levels were either above or below nominal. In one condition, it was concluded that the difference in performance was due to the clipping of speech in time, caused by operation of the non-linear processor and its associated control.

The results of conversational tests published by two Administrations and corresponding results of specific listening-only tests under double talk conditions indicate that the occurrence of clipping and gaps is very important to subjective performance. Objective measurements, concerning the switching characteristics of the NLP, give correlated parameters. The results of the different subjective test procedures clearly indicate that the control of the NLP and the associated switching characteristic under single and double talk conditions is one of the most important parameters. This influences the quality of background noise transmission and double talk performance. In addition, the echo attenuation during initial convergence and steady state conditions (under single and double talk conditions) is important.

I.7.5.1.3 Specific talking and listening tests

In the conversational tests, complaints were given concerning the echo disturbances for some of the echo cancellers under test. This indicates the need for a specific subjective test and corresponding objective test results to fix requirements for laboratory tests. For this purpose, specific *talking and* *listening tests* were designed to evaluate talking related disturbances. Four echo cancellers were tested.

One part of this test considered the initial convergence of the echo cancellers, both with enabled and disabled NLP. The results of the subjective tests and correlating objective measurements are given in Figures I.6 and I.7.

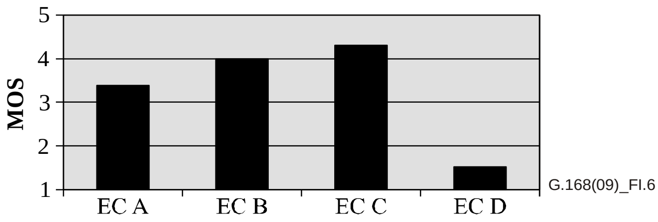


Figure I.6 – Results from *talking and listening tests*, MOS, disturbance caused by echoes, digital echo path, ERL 6 dB, no cancelled end background noise, NLP disabled

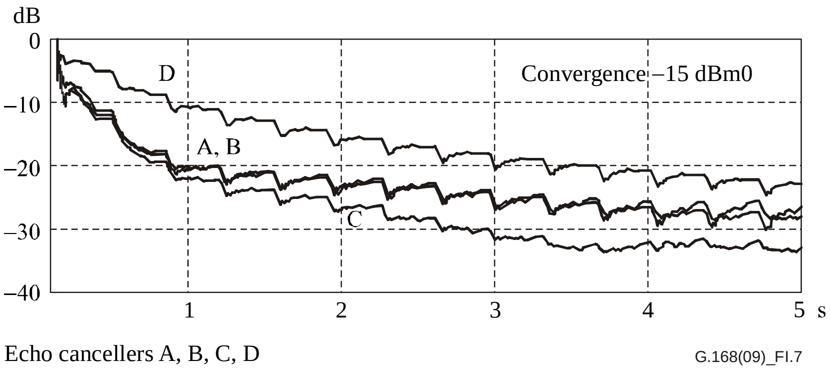


Figure I.7 – Convergence test versus time, digital echo path, ERL 6 dB, no cancelled end background noise, receive level –15 dBm0, NLP disabled

If the NLP is disabled, the results are given in Figure I.6. This test characterizes the convergence of the adaptive filters. The *talking and listening tests* are sensitive enough to point out differences between the tested echo cancellers. These differences could be expected from the test subjects' comments during the conversational tests but the conversational tests were not sensitive enough for a further comparison between the tested echo cancellers. The echo signal, which is produced by EC D causes the highest annoyance. These MOS values correlate to objective test results measured for the same echo cancellers as given in Figure I.7. This is relevant for Test 2B.

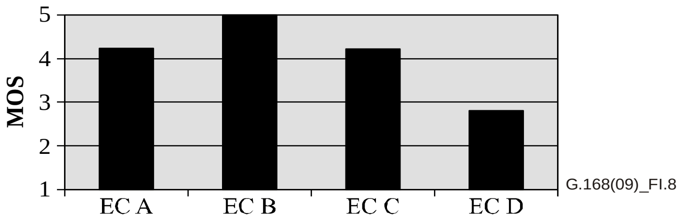


Figure I.8 – Results from *talking and listening tests*, MOS, disturbance caused by echoes, digital echo path, ERL 6 dB, background noise –55 dBm0, NLP enabled

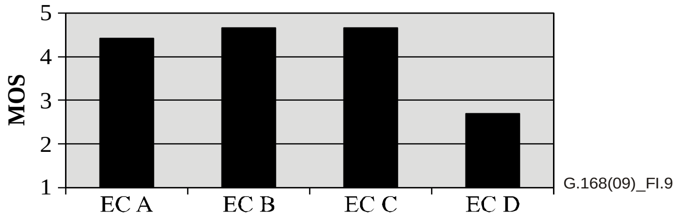
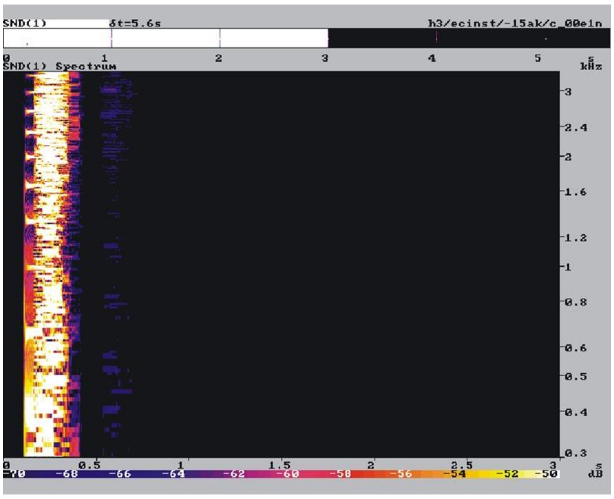


Figure I.9 – Results from *talking and listening tests*, MOS, disturbance caused by audible switching, digital echo path, ERL 6 dB, background noise –55 dBm0, NLP enabled

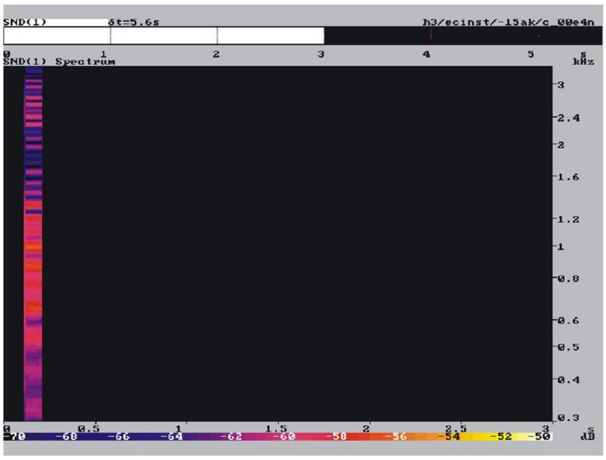
Figures I.8 and I.9 give the results, if the NLP is enabled. Figure I.8 shows the ratings for the echo disturbances, if the NLP is enabled, and Figure I.9 shows the corresponding results for audible clipping. EC D obtains the worse ratings with less than 3 points for both parameters. Echoes can be heard for a longer time until they are suppressed (Figure I.8). Audible switching causes a higher annoyance compared to the other three echo cancellers EC A, EC B and EC C (Figure I.9). The switching characteristic of EC D is more annoying compared to the other echo cancellers.

- Basically, several parameters cause the annoyance at the beginning of convergence: The initial convergence speed, determined as the echo attenuation vs. time is one important aspect as given through the correlated subjective and objective test results in Figures I.6 and I.7. The switching characteristics of the NLP and the echo attenuation vs. frequency influence the annoyance too. Correlated objective measurement results are given in Figure I.10. The pictures show a spectral analysis of the echo signal versus time (x-axis) and frequency (y-axis) during initial convergence for the echo cancellers which were tested subjectively (Figures I.8 and I.9). High echo signal peaks are given in light colours, dark colours represent a better echo attenuation. These measurement results are again a very good example of how subjectively obtained test results can be correlated to objective analyses. The tests clearly point out that the echo cancellers A, B and C which were judged significantly better in the talking and listening tests (Figures I.8 and I.9) attenuate and suppress the residual echo faster than echo canceller D. This is relevant for Tests 2A and 2C. There is currently no test to examine the quality of background noise transmission in this Recommendation.

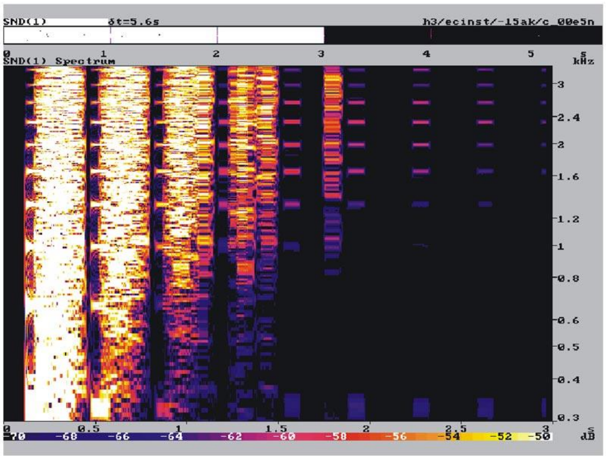


Echo cancellers A

Echo cancellers B



Echo cancellers C



G.168(09)_FI.10

Echo cancellers D

Figure I.10 – Spectral analysis of residual echo during initial convergence

Listening-only tests

A listening-only test is probably the most sensitive test method for the direct comparison of echo cancellers and for the evaluation of single transmission parameters. In addition to conversational tests and talking and listening tests, this test method was used for the evaluation of specific transmission parameters like residual echo levels, initial convergence or double talk performance.

Figures I.11 and I.12 show one example how the subjective test results were used to find values for laboratory tests. The double talk performance of six echo cancellers was accessed by a group of untrained subjects (Figure I.11) and experts (Figure I.12) for an ERL of 24 dB, a receive level of −15 dBm0 and a double talk level of −30 dBm0.

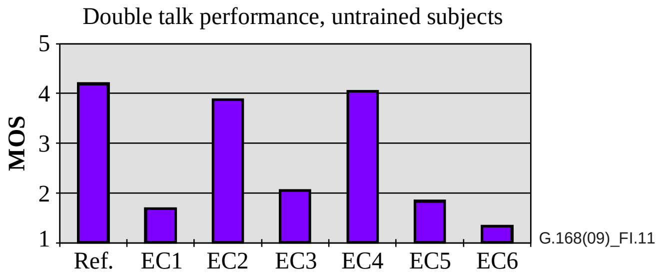


Figure I.11 – Results from listening-only tests, untrained subjects, double talk performance MOS, ERL 24 dB, receive level −15 dBm0, double talk level −30 dBm0

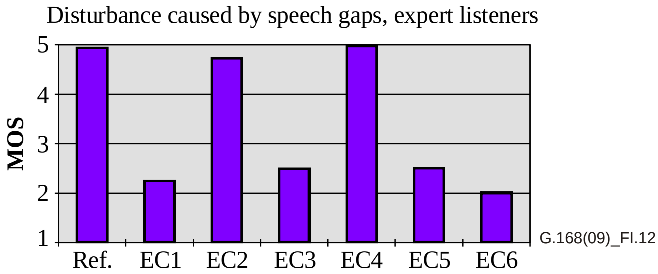


Figure I.12 – Results from listening-only tests, experts, disturbances during double talk caused by speech gaps MOS, ERL 24 dB, receive level −15 dBm0, double talk level −30 dBm0

Although the double talk signal level is very weak, the two echo cancellers EC2 and EC4 reach comparable MOS values like the undisturbed reference in Figure I.11. The untrained subjects assessed the parameter "double talk performance". This demonstrates the sensitivity of listening-only tests and indicates that even under these conditions a good transmission quality can be expected for echo cancellers. The other four speech echo cancellers are judged significantly worse. The ratings of the experts group for the disturbances caused by speech gaps (given in Figure I.12) show the same order of preference of the echo cancellers. This result again correlates to the conversational test result and the corresponding comments. It demonstrates that the annoyance caused by speech gaps seems to be the most important parameter.

Objective tests and test signals could be adapted to measure the correlated parameters. The tests pointed out that the degradation of a transmitted double talk signal is mainly determined by the insertion of the NLP during continuous speech. The correlated parameter is the switching characteristic of the NLP and its control during double talk. This is especially relevant for the design of Test 3C.

I.7.5.2 Subjective tests with the purpose of qualifying those effects of the echo cancellation process that cannot be captured by objective measurements

Subjective tests were performed with the purpose of qualifying those effects of the echo cancellation process that cannot be captured by objective measurements.

The results of the tests, judged by untrained and trained listeners, pointed out that one degradation of a transmitted double talk signal is mainly determined by the insertion of the NLP during continuous speech if the echo cancellers are fully converged. If CSS is used for the objective measurements, the switching characteristics can easily be determined after a burst of the double talk signal, because the time duration of all components is exactly defined for CSS. Subjective tests pointed out that a good double talk performance can be achieved even with double talk signal levels 15 dB lower than the receive input signal levels. If the bursts of the double talk CSS are not completely transmitted, the probability is high that longer speech gaps occur.

Appendix II

Measurement methods for characteristics of echo paths

(This appendix does not form an integral part of this Recommendation.)

II.1 Introduction

Knowledge of an echo path is necessary in order to control voice echo efficiently in telecommunication systems. Two of the main characteristics of an echo path are the dispersion time and frequency response. This appendix summarizes the measurement techniques used to obtain the echo-path impulse responses in telephone networks. It also provides analysis results for the echo paths measured in real networks.

This appendix is organized as follows. Clause II.2 gives details on the existing measurement procedures. The generation of echo-path characteristics is explained in clause II.3. Clause II.4 contains examples of the characteristics of echo paths measured in real networks. Finally, clause II.5 presents the conclusions.

This appendix does not include any reference to echo return loss (ERL) measurements. This is because this Recommendation specifies that all echo cancellers should operate and meet the requirements of all the tests with an ERL ≥ 6 dB.

More information on the characteristics of echo paths may be found on the ITU-T website.

II.2 Measurement procedure

Figure II.1 shows a typical call configuration where $x(k)$ and $y(k)$ are, respectively, the signals to and from the cancelled end. Because of the impedance mismatch in the hybrid that connects the 2-wire and 4-wire circuits, part of the signal from the non-cancelled end will leak through the hybrid and propagate back to the talker to form an echo. If the user in the cancelled end is not talking, $y(k)$ will be the echo signal. Therefore measurement of the signals $x(k)$ and $y(k)$ is performed to obtain the echo-path characteristics.

NOTE – The existence of level control devices, e.g., ALC (automatic level control), HLC (high level compensation), amplifiers, or attenuators along the path from the hybrid to the measurement device, can affect the measured ERL values.

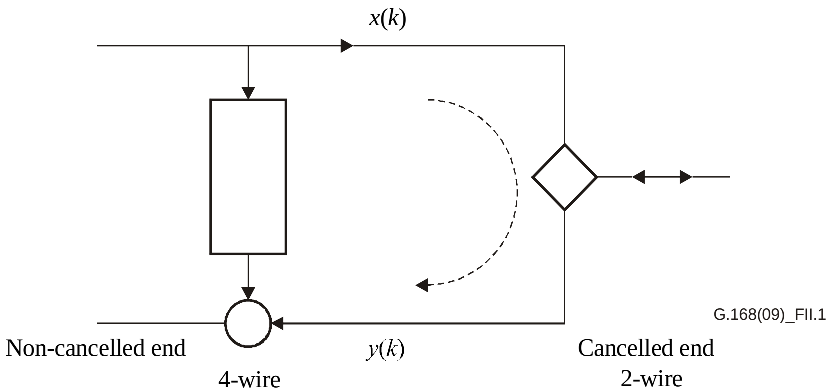


Figure II.1 – A typical call connection

II.2.1 Measurement set-up

II.2.1.1 Intrusive measurement set-up

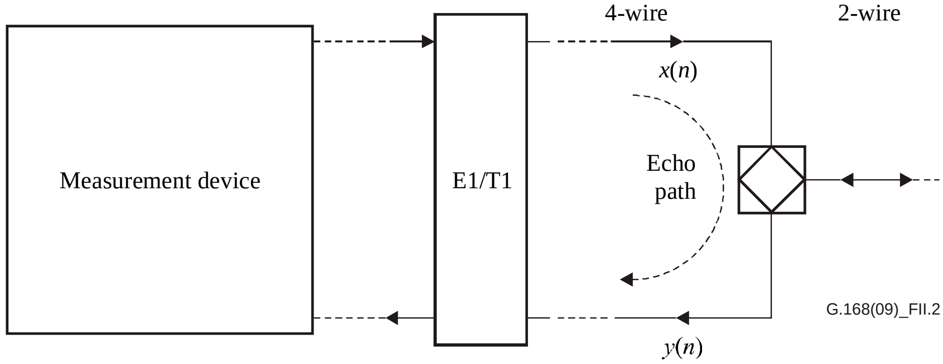


Figure II.2 – Block diagram of intrusive measurement set-up

Figure II.2 is the block diagram of the intrusive measurement set-up. The measurement device generates the test signal and performs the signal recording. The equipment uses a 4-wire connection for sending and receiving signals.

When a T1/E1 interface is not available, the test signal could be generated at a 2-wire point. In any case, the signals (transmit and receive) should be measured at the 4-wire digital point where the signals are well defined (in terms of level).

During measurement, the equipment first dials a telephone number. It then sends out the test signal after the call is established. The test signal goes through the T1/E1 interface, the central office (CO) and the PSTN to a distant user. Part of the transmitted signal is reflected back in the form of an echo. Both the transmitted signal and the returned echo are recorded in the equipment using a synchronous recorder. The two recorded signals are the R_{in} and S_{in} signals for an echo canceller.

NOTE – It must be ensured that neither an echo canceller nor an echo suppressor is present in the connection.

II.2.1.2 Non-intrusive measurement set-up

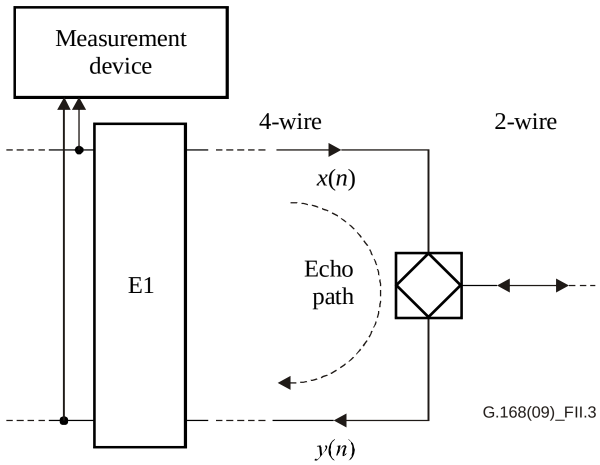


Figure II.3 – Block diagram of non-intrusive measurement set-up

Figure II.3 is the block diagram of the non-intrusive measurement set-up based on the use of an "In-service non-intrusive measurement device" (see [b-ITU-T P.561] for complete information). Prior to any measurements, it must be ensured that no echo canceller or echo suppressor is present in the connection. The measurement device monitors the source signal *x(n)* and echo signal *y(n)* at the T1/E1 interface.

II.2.2 Computation of the echo path impulse response

Computation of the echo path impulse response requires the perfect knowledge of the transmit and receive test signals. Three methods based on three different test signals are described herein.

II.2.2.1 Method 1: LS/LMS algorithms based on white noise injection

The test signal used is shown in Figure II.4. It consists of three segments. The first segment is a 2100 Hz tone with phase reversal. The purpose of the tone is to disable all the echo cancellers and echo suppressors in the link during measurement. The duration of the tone is set to T_1 = 1.35 s. The amplitude of the tone is −12 dBm0. The second segment is a pause. Its purpose is to obtain the background noise characteristics in the returned echo signal. Since a tone-disabled echo canceller, as specified in clause 7 and [ITU-T G.164] and [ITU-T G.165], will enable itself within 250 ms ± 150 ms when the signal level is below −39 dBm0, the duration of the pause is set to T_2 = 80 ms. The third segment is a White Gaussian Noise signal which is used to identify the echo-path impulse response. The power level of the white noise signal is −18 dBm0 and its duration is T_3 = 5 s.

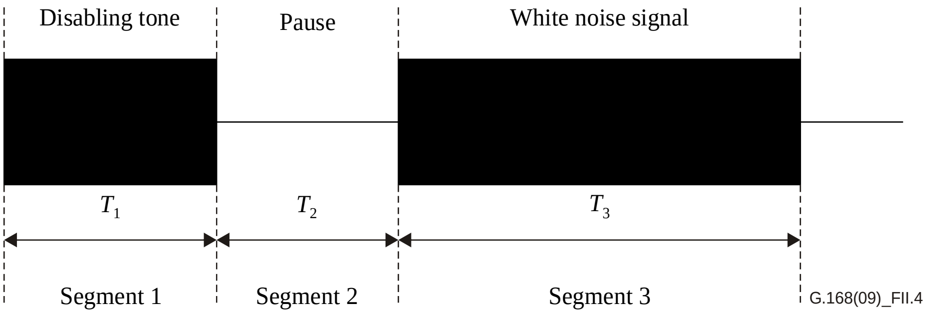


Figure II.4 – Test signal for echo-path measurement

Computation of impulse response

The echo signal can be modelled as:

$$y(n) = x(n) * h^{o}(n) + v(n) = \sum_{i=0}^{N-1} h^{o}(i) x(n-i) + v(n), \ n = 0, 1, 2, \ \ldots, L-1$$

where N is the length of the echo-path impulse response and L is the total number of data samples available. Define vectors:

$$\mathbf{h}^{o} = \left[h^{o}(0), h^{o}(1), \ \ldots, h^{o}(N-1)\right]^{\mathrm{T}}$$

and:

$$\mathbf{x}(n) = \left[x(n), x(n-1), \ \ldots, \mathrm{x}(n-N+1)\right]^{\mathrm{T}}$$

In vector notation,

$$y(n) = \mathbf{x}(n)^{\mathrm{T}} \mathbf{h}^{o} + v(n), \ n = 0, 1, 2, \ \ldots, L-1$$

The estimation problem can be stated as follows: Given L samples of $x(n)$ and $y(n)$, compute the echo-path impulse response $\mathbf{h}^o$. This is a standard system identification problem which can be solved by two common approaches. One is the least-squares (LS) method and the other is the iterative method using the normalized least-mean square (NLMS) algorithm. The first method yields a good solution with short data record. However, it is more computationally intensive than the second method. In most cases, provided that the data length L is large enough, the final solutions from both methods are very close and the difference is insignificant.

LS method

Define error signal:

$$e(n) = y(n) - \mathbf{h}^{\mathrm{T}}\mathbf{x}(n)$$

where $\mathbf{h}$ is the echo-path impulse response estimate. The LS method minimizes:

$$J = \sum_{n=0}^{L-1} e(n)^2$$

to determine $\mathbf{h}$. The corresponding solution can be shown to be:

$$\mathbf{h} = \mathbf{R}^{-1}\mathbf{p}$$

where:

$$\mathbf{R} = \sum_{n=0}^{L-1} \mathbf{x}(n)\mathbf{x}(n)^{\mathrm{T}}$$

$$\mathbf{p} = \sum_{n=0}^{L-1} y(n)\mathbf{x}(n)$$

NLMS method

The NLMS method finds $\mathbf{h}$ iteratively using the following equations:

$$e(n) = y(n) - \mathbf{h}(n)^{\mathrm{T}}\,\mathbf{x}(n)$$

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \frac{\mu}{\delta + \mathbf{x}(n)^{\mathrm{T}}\,\mathbf{x}(n)} e(n)\mathbf{x}(n)$$

where $\mathbf{h}(n)$ denotes the echo-path estimate at time n, δ is a small positive number and μ is a positive constant called step-size. The step-size must be less than two to allow convergence.

II.2.2.2 Method 2: Hadamard transform based on MLS sequence injection

The test signal consists in the use of MLS (maximal length sequence) (binary sequence which instantaneous normalized value is ±1, periodical of period L). The test signal is therefore a pseudo-random signal which has statistical properties close to white noise. Its autocorrelation function is circular and can be expressed as:

$$C_{xx}(n) = \frac{1}{L}\sum_{m=0}^{m=L-1} x(m-n)x(m) = \begin{cases} 1 & \text{if } n \equiv 0 \text{ modulo L} \\ -\dfrac{1}{L} & \text{otherwise} \end{cases}$$

Defining the function $\delta^L(n)$ by:

$$\delta^L(n) = \begin{cases} 1 & \text{if } n \equiv 0 \text{ modulo L} \\ 0 & \text{otherwise,} \end{cases}$$

the autocorrelation function can be rewritten as:

$$C_{xx}(n) = \left(1 + \frac{1}{L}\right)\delta^L(n) - \frac{1}{L}$$

Computation of impulse response

The echo signal can be modelled as: *y(n) = h*x(n)* when the circuit noise is ignored. If we assume that the signal *x(n)* is a white noise, the impulse response *h* is given by the cross-correlation function of signals *y(n)* and *x(n)*:

$$C_{xy}(n) = C_{xx}(n) * h$$

In case of a MLS sequence as described above, this equation is expressed equivalently as:

$$C_{xy}^L(n) = \frac{1}{L}\sum_{m=0}^{m=L-1} x(m-n)y(m) = \left(1 + \frac{1}{L}\right)\delta^L(n) * h - \frac{1}{L}\bar{h}$$

where $\bar{h} = \sum_{m=0}^{L-1} h(m)$ is the DC offset. The resulting estimate of the echo path impulse response is then given by:

$$h(n) = \frac{L}{L+1}[C_{xy}^L(n) + \bar{C}_{xy}^L], n = 0, \cdots, L-1$$

where $\bar{C}_{xy}^L = \sum_{m=0}^{L-1} C_{xy}^L(m) = \bar{h}/L$. In fact, the term $\bar{C}_{xy}^L$ may be ignored since the measured impulse response has zero mean in principle and/or the length of the sequence *L* is chosen to be large. In practice, it is recommended that the length of the sequence is at least 2^{14} samples, since this will ensure good white noise-like properties. Then, this equation shows that the measured impulse response *h*(*n*) may be directly approximated by the cross-correlation function $C_{xy}^L(n)$, for n=0,…, $L - 1$. Since *x*(*n*) is a pseudo-random noise having normalized values ±1, computing the cross-correlation function is easily performed by the Hadamard transform (see [b-Borish]).

II.2.2.3 Method 3: FFT algorithm based on periodic noise injection

The test signal consists of a periodic sequence having unit magnitude at discrete frequency points. It is defined using a superposition of sinusoidal signals having increasing frequencies, constant magnitudes and random phases:

$$x(n) = \frac{1 + (-1)^n}{N} + \frac{2}{N}\sum_{k=1}^{\frac{N}{2}-1}\cos\left(\frac{2\pi kn}{N} + \Phi_k\right), n = 0, \cdots, N-1$$

where Φ_k are independent, identically distributed random variables having uniform distribution between 0 and 2π. The period of the test signal *x(n)*, *N*, is selected to be larger than the expected echo path pure delay (t_r, expressed in samples) and controls the frequency resolution of the test signal. Depending on the noise level affecting the receive signal *y(n)*, an appropriate number of frames, *M*, is selected for periodically transmitting *x(n)*. In this way, the overall test signal will have duration of

$M \times N$ samples. At the echo path receiving end, M received frames of $y(n)$ are averaged to generate $\bar{y}(n), n = 0, \cdots, N-1$, so that the estimated echo path impulse response is then given by:

$$h(n) = IFFT[FFT[\bar{y}(n)] \cdot FFT[x(n)]^*]$$

During the injection of the test signal, the existing echo canceller connected to the echo path must be disabled.

II.3 Analysis of echo-path characteristics

This clause describes methods of echo-path impulse response analysis. The echo-path characteristics considered are dispersion width and magnitude response of echo paths.

Before the analysis of echo-path characteristics, the measured echo-path impulse responses should be post-processed by a bandpass filter which covers the same frequency region as a speech signal. The passband of the filter is from 200 Hz to 3600 Hz, the same frequency range as that of the CS signal specified in clause C.3.1.

II.3.1 Dispersion time

The echo-path impulse response is not a single impulse. It has a finite duration which we shall refer to as dispersion time. The dispersion time is a factor to determine the length of the H register in an echo canceller. A region where the echo reflection occurs is first located from an impulse response measurement. The duration for this region should be long enough to cover the entire echo reflection. For example, in the North American measurements, it was set to 30 ms. This corresponds to 240 samples at 8 kHz sampling rate. The impulse response in this range is denoted as $h(n)$, $n = 0, 1, ..., M$, where $M = 239$. If we truncate $h(n)$ by keeping the impulse response values from $n = N_1$ to N_2, then the relative square-error due to truncation is:

$$ERR = 1 - \frac{\sum_{n=N_1}^{N_2} h(n)^2}{\sum_{n=0}^{M} h(n)^2}$$

Figure II.5 shows such a possible truncation. According to Figure 11, we need the largest amount of echo reduction at 0 dBm0 L_{Rin}. With 6 dB ERL, this translates to 30 – 6 = 24 dB ERLE provided by an echo canceller. Hence the dispersion time is considered to be the shortest length $N_2 - N_1 + 1$ such that the ERR is less than −24 dB.

Note that the method suggested here to compute the dispersion time is for the worst-case scenario. This is because a 6 dB ERL is assumed to compute the ERR threshold. In practice, the ERL is much higher. The dispersion time is shorter when ERL increases.

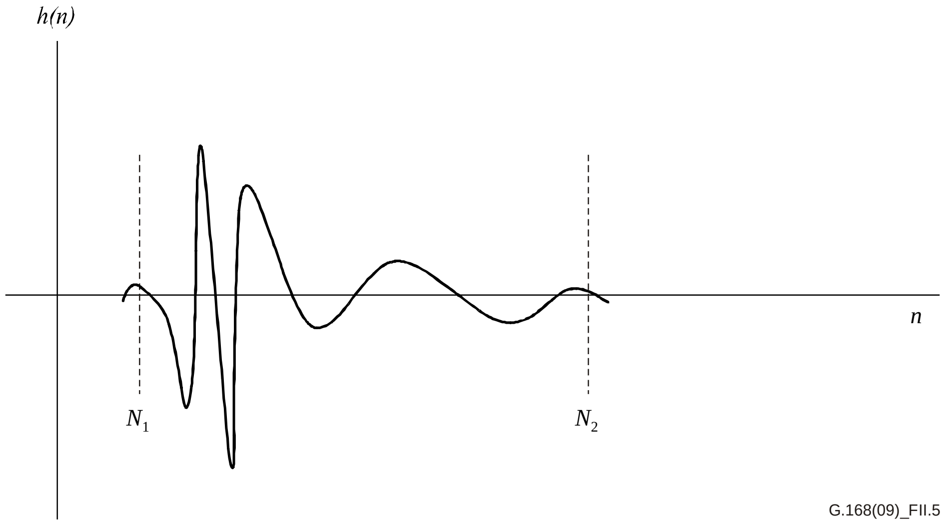


Figure II.5 – Echo-path truncation for dispersion time estimate

II.3.2 Magnitude response of echo path

While the echo-path impulse responses can be quite different in shape, the frequency domain characteristics of the echo paths are expected to have some similarity. The magnitude response characteristics of echo path can be generated by taking the average of the magnitude of the Fourier transform of each echo-path measurement.

II.4 Examples of echo path measurements from real networks

II.4.1 Echo-path characteristics from measurements in North America

During the period from June 1998 to April 1999, 101 long-distance calls were made from Montreal to the following Provinces and States across North America: Arizona, British Columbia, California, Louisiana, Manitoba, Massachusetts, Michigan, Minnesota, Missouri, Nevada, New York, North Carolina, Ontario, Quebec, Saskatchewan, Texas and Wisconsin. The measurements were made using the set-up as shown in Figure II.2. The send-out and returned signals were recorded in each call and the echo-path impulse responses were computed using the method described in clause II.2.2.1. This clause reports the echo-path characteristics generated from the above echo-path measurements.

II.4.1.1 Dispersion time

Figure II.6 is the histogram of the echo-path dispersion time. The largest percentage of dispersion time was between 5-7 ms. Only two calls had dispersion time between 11-12 ms. There was no call with dispersion time over 12 ms.

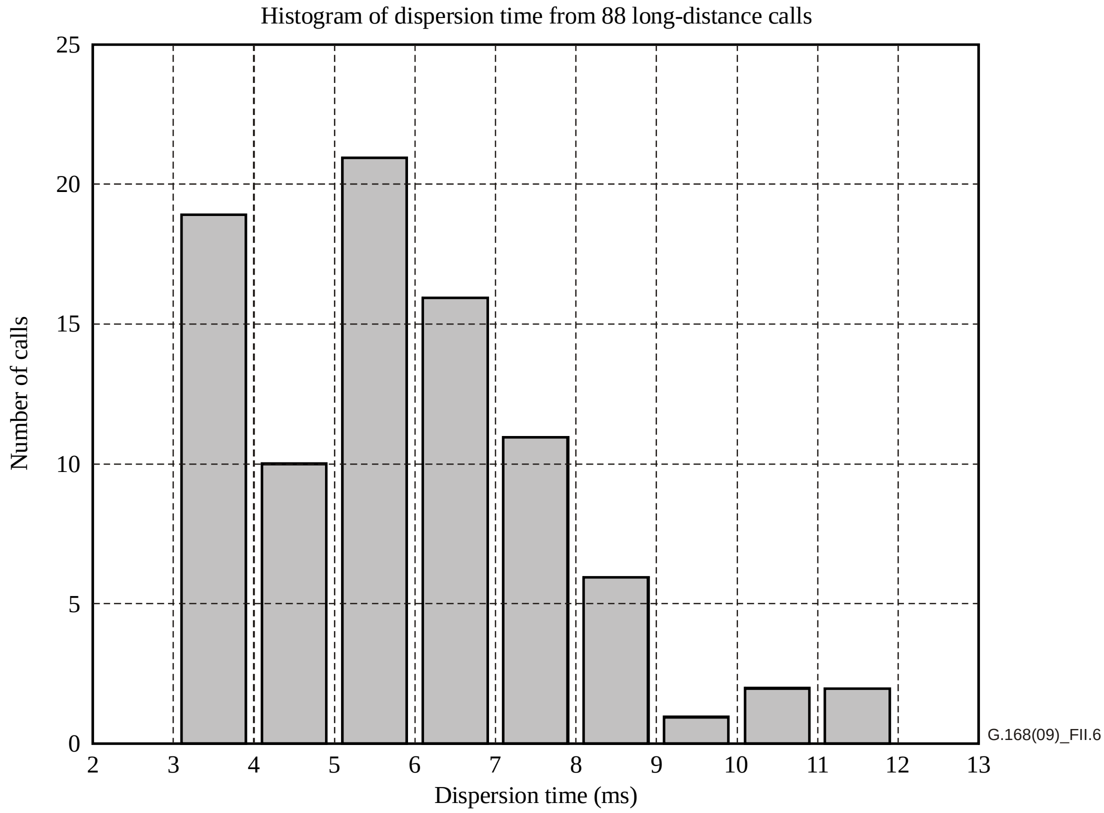


Figure II.6 – Histogram of dispersion time for long-distance calls: Mean = 6.02 ms, StD = 2.26 ms

II.4.1.2 Magnitude response of echo path

Figure II.7 is the average of the magnitude spectra of the measured echo-path impulse responses. The impulse responses were normalized to have unit energy before the spectra were computed. The solid line is the mean and the two dotted lines represent the one standard deviation region. It appears that the variance of magnitude spectra is not large. This indicates that there is a high consistency among the magnitude responses in the echo-paths. The averaged magnitude spectrum has a small peak around 250 Hz. In addition, the magnitude responses are relatively flat.

The results are consistent with the study in the article "Echo Performance of Toll Telephone Connections in the United States," by F.P. Duffy, G.K. McNees, I. Nasell, and T.W. Thatcher, Jr., in Bell System Technical Journal, 1974.

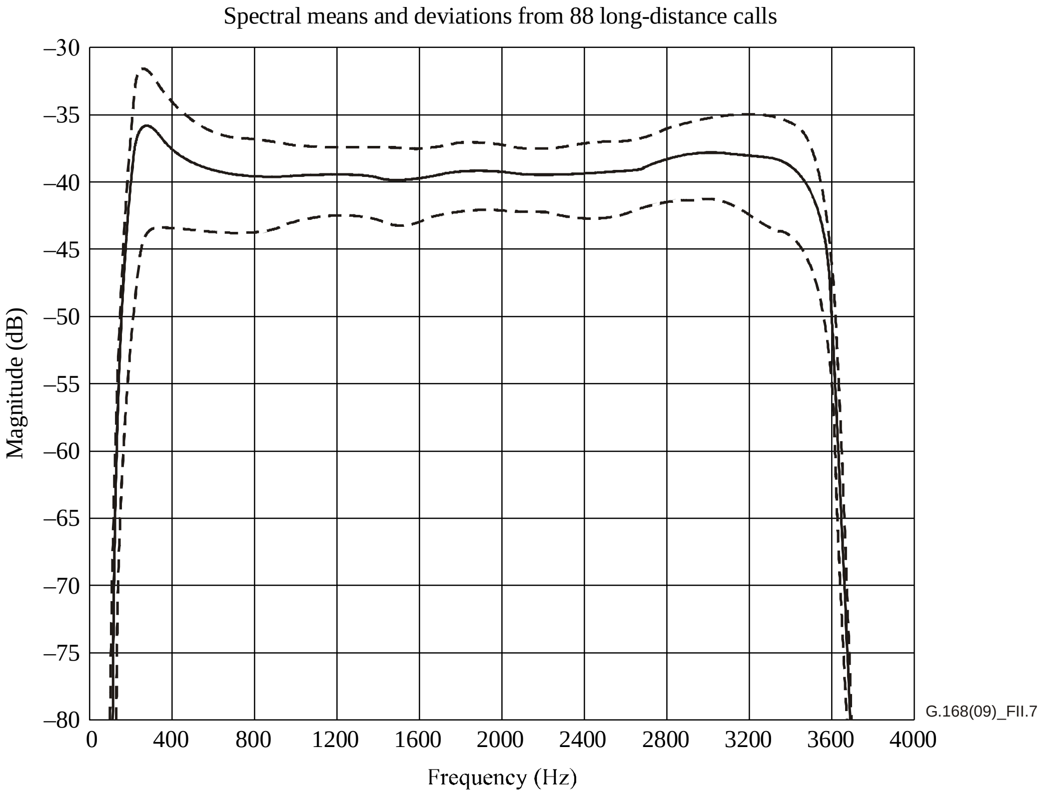


Figure II.7 – Magnitude spectra of echo paths for long-distance calls, echo paths normalized to have unity energy

II.4.1.3 Multiple reflections

We have observed 6 double reflections from the 101 long-distance calls. However, the total dispersion times of these double reflections were all limited to 10 ms. In other words, the two reflectors in each of these cases were closely located.

The number of reflections did not exceed two in the measurements.

II.4.2 Echo path characteristics from measurements in Europe

Measurements were made in the telephone networks in France using the set-up described in clause II.2.1. There are three types of echo path impulse responses. The impulse and frequency responses of these three types of echo paths are shown in Figures II.8 to II.10. They have equivalent characteristics of echo path models $m_1(k)$, $m_5(k)$ and $m_6(k)$ of Annex D.

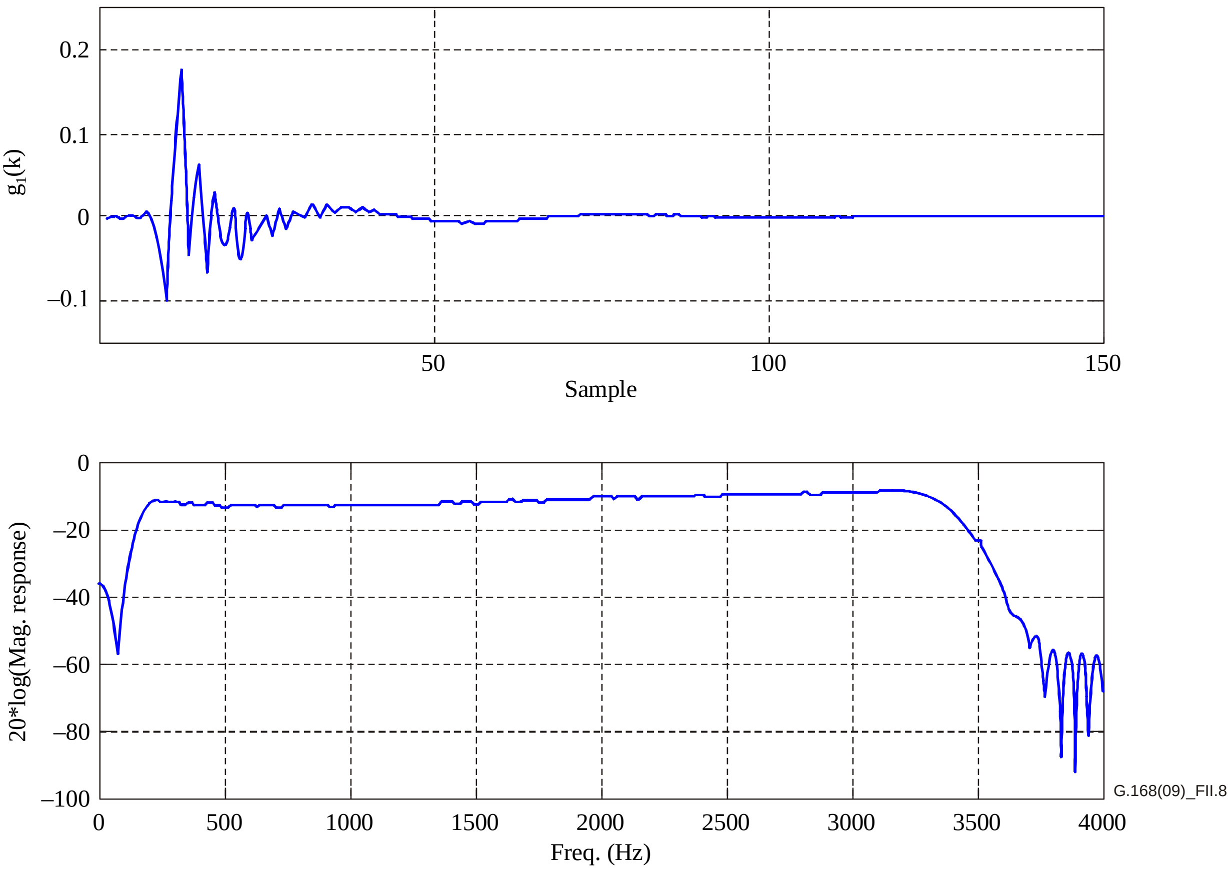


Figure II.8 – Echo path impulse response g_1, an ERL of 12 dB was used in the magnitude response plot

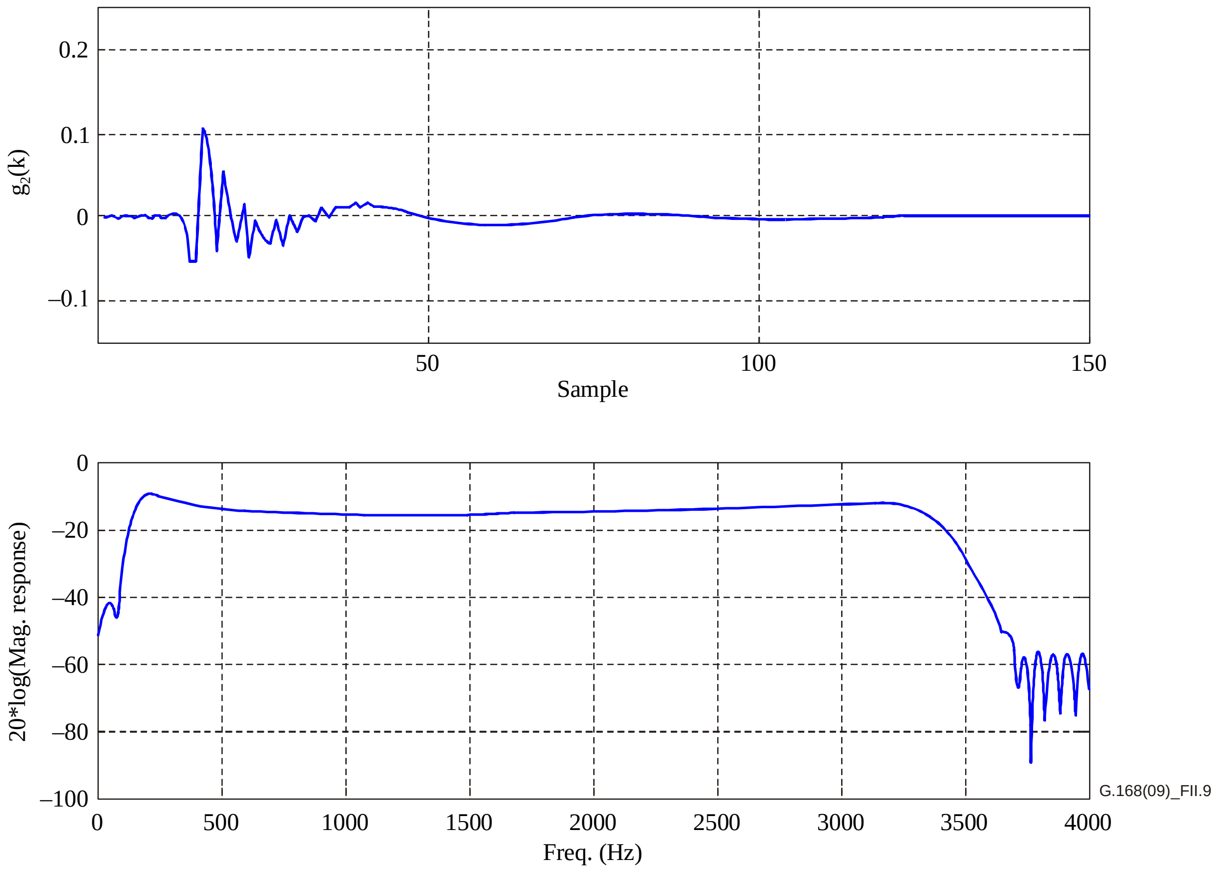


Figure II.9 – Echo path impulse response g_2, an ERL of 12 dB was used in the magnitude response plot

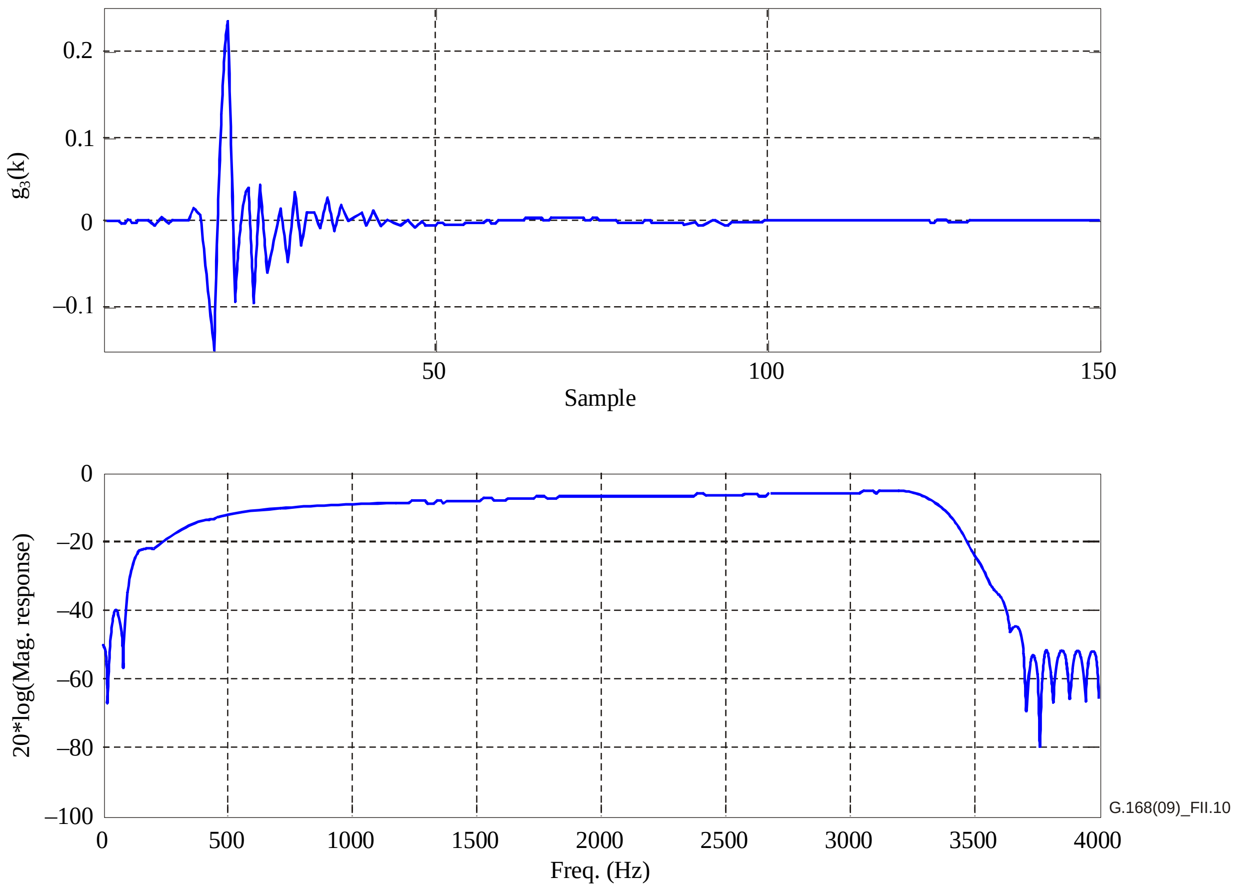


Figure II.10 – Echo path impulse response g_3, an ERL of 12 dB was used in the magnitude response plot

II.5 Conclusions

This appendix summarized the measurement procedures together with the computation of echo-path impulse response and its characteristics. Also included are the results from a number of echo paths measured in North America and Europe. The results can serve as references in designing a digital echo path for the testing of echo cancellers in this Recommendation. Specifically, the following important properties of echo paths are observed:

– The dispersion time of an echo path is within 12 ms.

– The magnitude response of most echo paths is relatively flat, with a small peak around 250 Hz.

– On some occasions double reflections may occur. Three or more reflections, however, rarely occur.

Appendix III

Multiple tail circuits

(This appendix does not form an integral part of this Recommendation.)

In modern networks, it may be the case that a two-party call is modified after the call is initially set up, so that one or more additional parties can participate in the conversation, as illustrated in Figure III.1.

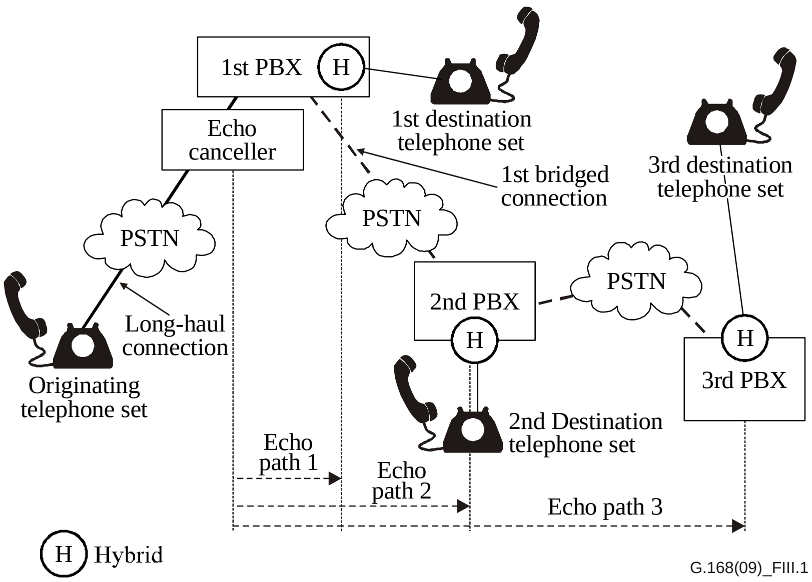


Figure III.1 – Multiple tails in a typical bridged telephone call

In Figure III.1, the originating telephone set initiates a call to the 1st destination set. The echo canceller sees the hybrid associated with the 1st PBX, illustrated as echo path 1. The recipient at the 1st PBX then bridges in a 2nd destination set by using the bridging function found on all modern PBXs. The echo canceller now sees the second hybrid, appended to the first, and delayed in time by the network delay between PBX 1 and 2. This is illustrated as echo path 2. The second destination may well bridge in a third destination, adding another hybrid tail associated with the hybrid in the 3rd PBX, and delayed by the sum of the network delay between PBX 1 and 2, and PBX 2 and 3. This is illustrated as echo path 3.

PBXs may not have any echo cancellation built into them, even though they perform this bridging function. This requires that the network echo canceller be able to support multiple tails up to the echo tail capacity of the canceller.

Network operators and administrators should take this into account when testing echo cancellers for use in their network.

Appendix IV

Guidelines on the use of parameters for testing echo cancellers

(This appendix does not form an integral part of this Recommendation.)

The tests in this Recommendation have many parameters which are specified as ranges. To provide some guidelines on a set of test conditions which should be considered at a minimum, the following is suggested. It is strongly cautioned that these values do not imply a sufficient set for compliance to this Recommendation, which is left for the discretion of the telecommunication providers. Also, it should be noted that all possible combinations of these parameters listed below would be time consuming.

1) *Echo path delay*

   a) Δ – dispersion of g(k) – 4.

   b) Δ/2 – 4.

   c) Δ/10.

2) *Echo return loss*

   a) 6 dB.

   b) 15 dB.

   c) 30 dB.

3) *Receive-in level (L_{Rin})*

   a) Maximum specified input level for the particular test.

   b) –10 dBm0.

   c) –20 dBm0.

   d) Minimum specified input level for the particular test.

4) *Initial condition of H Register*

   a) H register reset.

   b) H register converged to an open echo path.

   c) H register converged to some g(k).

Appendix V

List of low speed modems for optional use with Test No. 14

(This appendix does not form an integral part of this Recommendation.)

This appendix contains a comprehensive list of low speed modems that may optionally be used in Test No. 14. The list has been compiled by operators in different countries and is meant to augment the minimum set that is required to demonstrate ITU-T G.168 compliance given in Test No. 14.

Table V.1 – List of low speed modems for optional use with Test No. 14

Application	Modulation type	Data rate (bits/s)	Answer tone	Performance criteria	Minimum volume of data	Comments
Dial-up data	ITU-T V.21 ITU-T V.22 ITU-T V.22 *bis* ITU-T V.32 ITU-T V.32 *bis*	ITU-T V.21: 300 ITU-T V.22: 1200 ITU-T V.22 *bis*: 1200 and 2400 ITU-T V.32: 4800 and 9600 ITU-T V.32 *bis*: 4800 to 14400	ITU-T V.21, ITU-T V.22 and ITU-T V.22 *bis*: ANS ITU-T V.32 and ITU-T V.32 *bis*: ANS and /ANS	BER	ITU-T V.21: 20 kbits ITU-T V.22 and ITU-T V.22 *bis*: 100 kbits ITU-T V.32 and ITU-T V.32 *bis*: 1 Mbits	ITU-T V.32 and ITU-T V.32 *bis* are generally used with /ANS but can be used with ANS with the internal echo canceller switched off for double satellite hops where the internal echo canceller does not have sufficient tail capacity. Used for text telephony in conjunction with ITU-T V.61.
Point of sale terminal	ITU-T V.21 ITU-T V.22 ITU-T V.22 *bis* ITU-T V.23 Bell 103 Bell 212A	ITU-T V.21 and Bell 103: 300 ITU-T V.22 *bis*: 2400 ITU-T V.22 and Bell 212A: 1200 ITU-T V.23: 75, 600 and 1200	Proprietary	BER	ITU-T V.23 (75): 1 kbits ITU-T V.21, Bell 103 and ITU-T V.23 (600/1200): 20 kbits ITU-T V.22, ITU-T V.22 *bis* and Bell 212A: 100 kbits	
Telemetry	ITU-T V.23	600 and 1200	ANS (optional)	BER	20 kbits	
Text telephone	ITU-T V.21 half duplex low channel ("EDT") ITU-T V.21 full duplex TIA 825A (Baudot) ITU-T V.23 ITU-T V.18 DTMF	ITU-T V.18: 300 bit/s ITU-T V.21: 110 half duplex low channel ("EDT") ITU-T V.21: 300 full duplex TIA 825A: 45.45 and 50 ITU-T V.23: 1200/75 DTMF rate n/a	ITU-T V.18, ITU-T V.21 and ITU-T V.23: ANS (optional) DTMF and TIA 825A: None	Character error rate	ITU-T V.18, ITU-T V.21, ITU-T V.23 and TIA 825A: 1000 characters DTMF: 300 characteristics	

Table V.1 – List of low speed modems for optional use with Test No. 14

Application	Modulation type	Data rate (bits/s)	Answer tone	Performance criteria	Minimum volume of data	Comments
Security/Alarm system	SIA DC-02 SIA DC-05	Pulse and DTMF: 10 to 33 pulses/s Bell 103: 300	Various handshake tones	Character error rate	300 characters	The SIA standards specify a number of different pulse, DTMF and FSK formats for use by security/alarm systems. In addition, SIA DC-03 specifies a standard protocol for use with Bell 103 systems.
Leased line dial back-up	Bell 208 A/B Bell 202	Bell 202: 1200 Bell 208 A/B: 4800	Unknown	BER	100 kbits	

Legend:

ANS 2100 Hz Answer Tone as defined in [ITU-T V.25]

/ANS 2100 Hz Answer Tone with phase reversals as defined in [ITU-T V.25]

BER Bit Error Rate

NOTE 1 – Where possible, synchronous testing should be performed, since this more accurately detects bit errors due to transmission problems. The exceptions to this are ITU-T V.21 and Bell 103, which are usually tested in asynchronous mode. In general, error correction should be switched off.

NOTE 2 – Point of sale modems may use standard and proprietary versions of modulations.

NOTE 3 – Security and Alarm system modems may use standard and proprietary versions of modulations.

NOTE 4 – SIA DC-02 Security Industry Association (SIA) General Protocols – Available from www.siaonline.org.

NOTE 5 – SIA DC-03 Security Industry Association (SIA) SIA Format – Available from www.siaonline.org.

NOTE 6 – SIA DC-05 Security Industry Association (SIA) Ademco Contact ID Standard – Available from www.siaonline.org.

NOTE 7 – ITU-T V.18 is an automoding procedure using ITU-T V.8 and all other textphone modulations. ITU-T V.8 uses ITU-T V.21 300 bits/s one channel at a time both before and after ANSam. Full information is found in [ITU-T V.18] and [ITU-T V.8].

NOTE 8 – ITU-T V.21 110 bits/s half duplex is used for EDT, a legacy sub-mode of ITU-T V.18. Only lower channel is used. Pauses between text transmissions are silent or used for voice. EDT is used for text telephony in Germany, Italy, Spain, and Switzerland.

NOTE 9 – DTMF is for textphones with text coding based on DTMF coded text.

Appendix VI

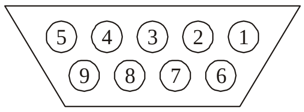
Example control interfaces

(This appendix does not form an integral part of this Recommendation.)

The following are two examples of control interfaces used for the required control functions that include enable/disable NLP, inhibit/enable adaptation, and reset H-register.

VI.1 Parallel transistor-transistor logic (TTL) connection for digital signals

Figure VI.1 is an example of TTL connection on an echo canceller testing unit. A cable needs to be provided by the echo canceller manufacturer for the connection to the echo canceller under test.



TTL DB-9 female connector front view

G.168(09)_FVI.1

Pin	Connection
1	Ground
2	N/A
3	Adaptation inhibit/enable
4	H-register reset
5	N/A
6	Ground
7	N/A
8	NLP enable/disable
9	N/A

Figure VI.1 – Connector wiring for echo canceller cable

VI.2 Serial data link (ITU-T V.24 or RS-232) for software commands

The serial data link is used for the software command control of the echo canceller under test. Besides the serial cable for the connection between the testing unit and the echo canceller under test, a set of commands needs to be provided by the echo canceller manufacturer for the required control functions.

To meet the requirements on the synchronization between the test signals and the operation mode change of the echo canceller under test, the control signals must be well aligned with the test signals. Since it may take certain time from the moment a control signal is transmitted from the testing unit to the moment the echo canceller under test changes its operational mode, a control signal must be transmitted a control-ahead time before the test signals, so that the echo canceller can change its mode at a moment aligned with the test signals. It is required that the echo canceller manufacturer provide the control-ahead time. It is recommended that this control-ahead time be specified in the interface document and it shall be kept to no more than 200 ms.

Appendix VII

Guidance on echo canceller orientation in conference bridge applications

(This appendix does not form an integral part of this Recommendation.)

VII.1 Introduction

In call topologies that include conference bridges where the participants are distant from each other, it is assumed that each conference bridge participant has an echo canceller intended to protect him from echo. The echo cancellers can be oriented in two ways: (a) facing the conference bridge circuit, or (b) facing the conference bridge participants. It is explained below that it is preferred to orient the echo cancellers facing the individual conference bridge participants for the purpose of cancelling the echoes created from their hybrids before the echoes leak to the network and get distributed by the conference bridge to the participants. Such a placement of echo cancellers not only avoids the degradation in Quality of Experience of other conference bridge participants, but also eliminates the occurrence of multiple-reflection echo in a conferencing scenario.

One should avoid orienting echo cancellers to face the conference bridge circuit because the other participants in the call will experience quality degradation caused by listener echo [b-ITU-T G.126] from the hybrid reflections, regardless of the capabilities of the individual echo cancellers to handle multiple reflections. The reason for this is explained in the following clauses.

VII.2 Echo canceller oriented towards the hybrid of the conference bridge participant (recommended)

Figure VII.1 gives the configuration when placing echo cancellers oriented towards the hybrids of conference bridge participants. The conference bridge participant 1 echo canceller will cancel any echo caused by his own hybrid so that the outgoing signal from conference bridge participant 1 entering the bridging network will be free of echo.

When conference bridge participant 1 is talking, his voice will go through the hybrids of different conference bridge participants. However, the cancellers of different conference bridge participants ensure that hybrid reflections are not present in their respective out-going signals to the bridging network. As a result, no participant in the conference bridge will be subjected to listener echo in the incoming voice signal. Furthermore, conference bridge participant 1 is protected from echo indirectly because no echo is present in the bridge connection. An added benefit for such placement of echo canceller is that each canceller is dealing with a single hybrid reflection and it will not encounter multiple reflections caused by the hybrids of different conference bridge participants. Consequently, this will improve the initial echo canceller rate of adaptation and double-talk performance.

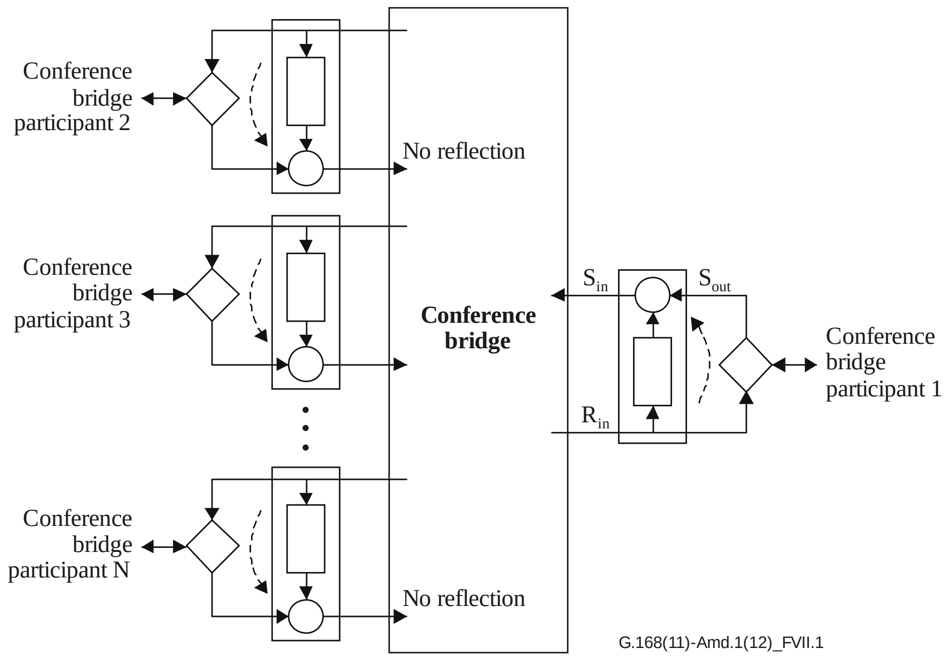


Figure VII.1 – Conference bridge with echo canceller facing the hybrid of conference bridge participant

VII.3 Echo canceller oriented away from the hybrid of the conference bridge participant (not recommended)

Figure VII.2 depicts the block diagram of a conference bridge when the echo cancellers are oriented away from the conference bridge circuit. When conference bridge participant 1 is talking, his voice will be reflected by the hybrids of different conference bridge participants and the echo canceller of conference bridge participant 1 will see multiple-reflection echoes in the send-path for cancellation.

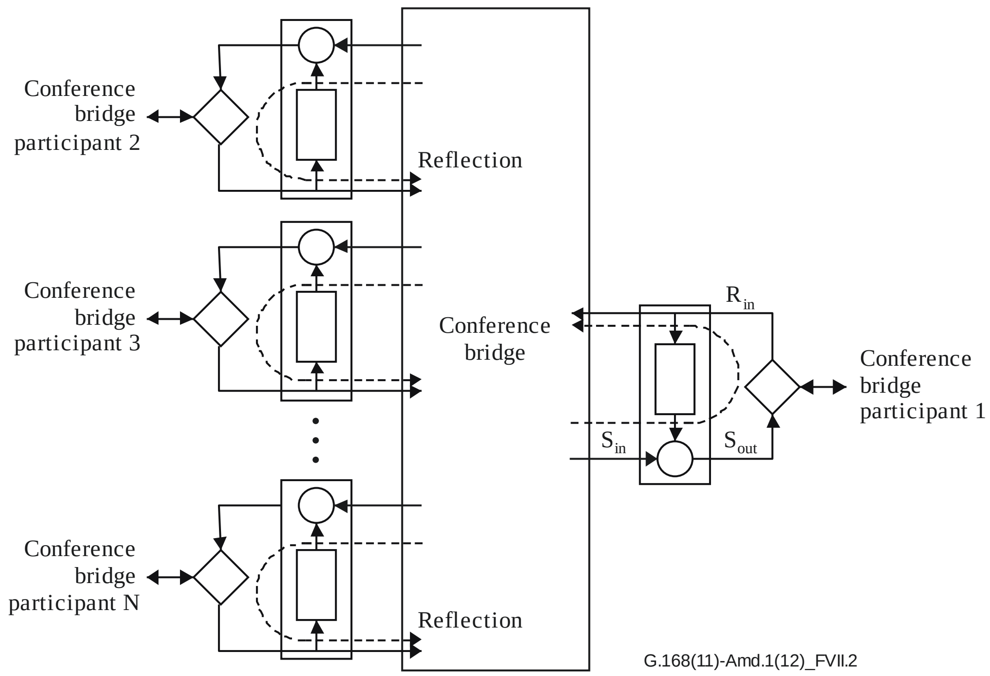


Figure VII.2 – Conference bridge with echo canceller facing the bridging circuit

When echo canceller 1 is operating properly, conference bridge participant 1 will not hear his own echo. However, the reflection of the voice from conference bridge participant 1 caused by the hybrid of conference bridge participant 2 will propagate to other conference bridge participants; conference bridge participant 3 will hear the original conference bridge participant 1 signal plus its attenuated and delayed version from the reflection of the conference bridge participant 2 hybrid. If the delay in the reflected signal is not small relative to the original signal, the echo will be noticeable. Similarly, other conference participants will encounter voice quality degradation due to such reflections. In other words, having the echo canceller oriented towards the network can protect the talker from his/her own echo but cannot avoid the degradation of Quality of Experience perceived by other conference bridge participants.

For echo cancellers configured in this manner, the voice signal perceived by conference bridge participant i can be approximated by:

$$X_i = \sum_{\substack{j=1 \\ j \neq i}}^{N} \left(S_j + \sum_{\substack{k=1 \\ k \neq j}}^{N} m^{(k)} * S_j \right)$$

where N is the number of participants in the conference call, S_j is the speech signal from talker j, $m^{(k)}$ is the hybrid impulse response of conference bridge participant k and $*$ denotes the convolution operation. The second term inside the parenthesis represents listener echo due to hybrid reflections which can significantly degrade the Quality of Experience for conference bridge participant i. Note that the echo canceller of conference bridge participant i is configured to remove his own echo using his originating speech signal as R_{in}. It should be emphasized that no matter how well this echo canceller is designed, it will not be able to remove the listener echo component because it is not caused by the speech of conference bridge participant i.

VII.3.1 Echo canceller orientation for conference bridge participant with small tail end delays (special case)

In conferencing scenarios in which the terminating connections do not have echo cancellers for the reason that the listener echoes are not noticeable due to the short echo delays, the echo canceller may be positioned after the conference bridge to remove the talker echo of the far-end user. This scenario is shown in Figure VII.3, in which the conference bridge participants 2 to N are not too far from each other and they do not have individual echo cancellers. Owing to the small differences in echo delays, the echo canceller in this case will not see separated multiple hybrid reflections (the reflections largely overlap with each other), but could encounter echo path change when some of conference participants 2 to N enter and/or leave during the conference call. To protect conference bridge participants 2 to N against the echo caused by the hybrid of the far-end user (conference bridge participant 1), another echo canceller facing the far-end-user hybrid is needed before the incoming signal is distributed to participants 2 to N. The suggested arrangement for the two echo cancellers is shown in Figure VII.3. However, if technically or economically feasible, it is highly recommended to adopt separate echo canceller arrangements shown in Figure VII.1 to ensure higher voice quality.

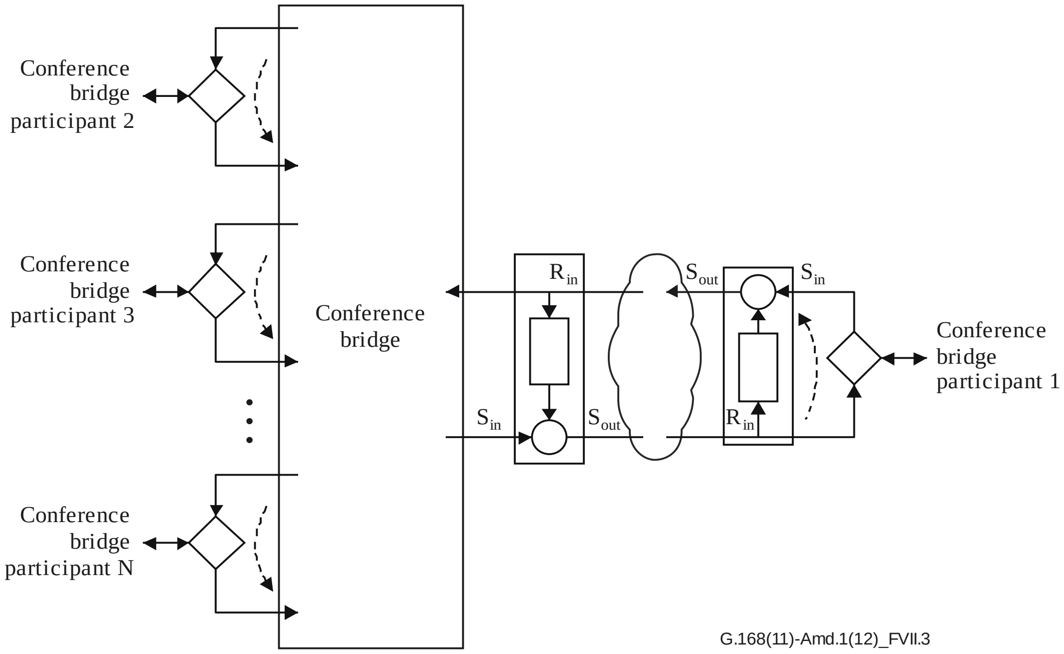


Figure VII.3 – Conference bridge scenario in which the terminating connections for participants 2 to *N* do not experience listener echoes due to short echo delays amongst them

Appendix VIII

Test methodologies for use on embedded echo cancellers

(This appendix does not form an integral part of this Recommendation.)

VIII.1 Introduction

The purpose of this appendix is to define methods for testing embedded echo cancellers (EECs) for conformance with Annex E.

The appendix defines two test methodologies. Test methodology 1 uses a special "Test Mode" to access the various echo canceller control signals required to perform ITU-T G.168 Annex E tests. Test methodology 2 provides a means to perform the tests when h-register freeze is not available. This will allow for testing of an EEC that has conventional 4-wire access (TDM or packet), together with some means to control the NLP. Test methodology 2 may not be applicable to all echo canceller algorithms that are otherwise compliant to ITU-T G.168 (i.e., the algorithm itself is tested as a non-embedded echo canceller for ITU-T G.168 compliance, before it becomes embedded). For example, it requires that the echo canceller inhibits adaptation immediately after the near end signal level is higher than the far end signal level. In order to avoid misleading results with methodology 2, it is important to check its applicability to the echo canceller algorithm. Procedures to do so are for further study. This methodology cannot be used for tests where the NLP is on (for example Test 2C a), thus it can provide only partial compliance to Annex E. Methodology 2 has not been thoroughly tested, therefore it should be considered experimental and a topic for further study. Testing of EECs with 2-wire interfaces and where NLP on/off control is not available is also for further study.

VIII.2 Test Methodology 1: Testing with the aid of a special "Test Mode" (under study)

With this methodology, the EEC under test should be provided with a special "TEST MODE". When this mode is entered the EEC will accept pre-defined control signals (such as NLP on/off, h-register reset, h-register freeze, etc.) using a method to be determined. The "TEST MODE" may be entered by either a hardware or software binary switch provided by the EEC host. When in the "TEST MODE", other functions associated with the EEC host (e.g., voice coding) that would be active during normal operation MUST remain active.

VIII.3 Test Methodology 2: Testing without access to control signals (under study)

This methodology requires that the EEC under test is equipped with a double talk (DT) detector that will freeze h-register adaptation immediately upon detection of a DT signal and keeps the adaptation frozen during the presence of this signal. The inhibition of h-register adaptation is accomplished by injecting a high level DT signal to the S_{gen} port. The artificial DT signal occupies a portion of the frequency range only, so that it can be filtered out by processing the S_{out} signal. The residual echo component that does not reside in the same frequency range as the artificial DT signal will remain in the filter output for EEC performance evaluation.

The method uses two pairs of filters. One pair $(G_{LL}(z), G_{LH}(z))$ is for assessing the echo cancellation performance in the low-frequency portion of the voice spectrum. The other pair $(G_{HL}(z), G_{HH}(z))$ is for assessing the performance in the high-frequency portion. The magnitude responses of the filters should satisfy the following three conditions over the voice frequency range from 0 to 4000 Hz:

$$|G_{LL}(f)|(dB) + |G_{LH}(f)|(dB) \le -55\,dB \qquad \text{(VIII-1)}$$

$$|G_{HL}(f)|(dB) + |G_{HH}(f)|(dB) \le -55\,dB \qquad \text{(VIII-2)}$$

$$-0.1dB \le 10\log\left(|G_{LL}(f)|^2 + |G_{HH}(f)|^2\right)(dB) \le 0.1dB \qquad \text{(VIII-3)}$$

Note that the additions in equations (VIII-1) and (VIII-2) are in log-scale and the sum in equation (VIII-3) is in linear-scale. Condition (VIII-1) guarantees the filter pair $(G_{LL}(z), G_{LH}(z))$ complement each other and have a net gain sufficiently close to zero when a signal goes though the two filters. Condition (VIII-2) guarantees these properties for the pair $(G_{HL}(z), G_{HH}(z))$. Condition (VIII-3) ensures the two filters $G_{LL}(z)$ and $G_{HH}(z)$ together provide nearly unit energy gain over the entire frequency range. An example of the filters is shown in Tables VIII.1 to VIII.4, where the coefficients are specified in fixed-point integer format. Figures VIII.1 to VIII.3 confirm that the filters satisfy the three conditions listed above.

To perform ITU-T G.168 tests in which adaptation of the h-register must be inhibited, the following steps are used:

i) Performance evaluation in the low frequency region:

- Generate white Gaussian noise (WGN) of ±11 dB crest factor
- Pass the WGN through the filter $G_{LH}(z)$ to generate Gaussian noise $w_{HP}(n)$ that has high-frequency content only
- Scale $w_{HP}(n)$ to have a level higher than L_{Rin}
- Add $w_{HP}(n)$ to the S_{gen} port whenever h-register adaptation is required to be disabled during the test
- Process the S_{out} signal using $G_{LL}(z)$ to generate $S_{out,LP}(n)$, the component of S_{out} in the low part of the spectrum
- Measure the level of $S_{out,LP}(n)$, $L_{Sout,\ LP}$

ii) Performance evaluation in the high frequency region

- Generate white Gaussian noise (WGN) of ±11 dB crest factor
- Pass the WGN through the filter $G_{HL}(z)$ to generate Gaussian noise $w_{LP}(n)$ that has low-frequency content only
- Scale $w_{LP}(n)$ to have a level higher than L_{Rin}
- Add $w_{LP}(n)$ to the S_{gen} port whenever h-register adaptation is required to be disabled during the test
- Process the S_{out} signal using $G_{HH}(z)$ to generate $S_{out,HP}(n)$, the component of S_{out} in the high part of the spectrum
- Measure the level of $S_{out,HP}(n)$, $L_{Sout,\ HP}$

iii) Performance evaluation (measuring EEC residual echo during DT)

- Add $L_{Sout,\ LP}$ and $L_{Sout,\ HP}$ (on a power basis) to form L_{Sout} to assess ITU-T G.168 test requirements.

NOTE 1 – The use of other artificial DT signals and filters, such as tones and notch filters, are currently under investigation.

NOTE 2 – Depending on implementation, the w(n) signal level may be attenuated before reaching the S_{in} port of an EEC (e.g., by other signal processing functions such as noise reduction, etc.). To ensure the w(n) signal level is sufficient to activate the DT detector, a level check on w(n) in S_{out} with respect to the original level should be made while the R_{in} signal is removed.

NOTE 3 – The level of w(n) should be set high enough to activate promptly the DT detector, but to avoid possible clipping of the w(n) signal it should not be too high.

Table VIII.1 – Coefficients $g_{LL}(k)$ of the filter $G_{LL}(z)$ (Q15 fixed-point format)

Index *k*	$g_{LL}(k)$	**Index *k***	$g_{LL}(k)$	**Index *k***	$g_{LL}(k)$
0, 64	22	16, 48	208	32	16762
1, 63	−20	17, 47	−370		
2, 62	−36	18, 46	−240		
3, 61	18	19, 45	502		
4, 60	44	20, 44	272		
5, 59	−39	21, 43	−680		
6, 58	−67	22, 42	−301		
7, 57	59	23, 41	930		
8, 56	88	24, 40	327		
9, 55	−92	25, 39	−1305		
10, 54	−115	26, 38	−349		
11, 53	135	27, 37	1950		
12, 52	144	28, 36	365		
13, 51	−193	29, 35	−3394		
14, 50	−175	30, 34	−375		
15, 49	270	31, 33	10402		

Table VIII.2 – Coefficients $g_{LH}(k)$ of the filter $G_{LH}(z)$ (Q15 fixed-point format)

Index *k*	$g_{LH}(k)$	**Index *k***	$g_{LH}(k)$	**Index *k***	$g_{LH}(k)$
0, 64	26	16, 48	238	32	12731
1, 63	2	17, 47	223		
2, 62	−35	18, 46	−486		
3, 61	36	19, 45	90		
4, 60	17	20, 44	563		
5, 59	−65	21, 43	−560		
6, 58	24	22, 42	−299		
7, 57	78	23, 41	982		
8, 56	−100	24, 40	−382		
9, 55	−26	25, 39	−1040		
10, 54	158	26, 38	1394		
11, 53	−89	27, 37	355		
12, 52	−144	28, 36	−2488		
13, 51	234	29, 35	1697		
14, 50	5	30, 34	3334		
15, 49	−313	31, 33	−9778		

Table VIII.3 – Coefficients $g_{HL}(k)$ of the filter $G_{HL}(z)$ (Q15 fixed-point format)

Index *k*	$g_{HL}(k)$	Index *k*	$g_{HL}(k)$	Index *k*	$g_{HL}(k)$
0, 64	26	16, 48	239	32	12 732
1, 63	−1	17, 47	−222		
2, 62	−35	18, 46	−486		
3, 61	−37	19, 45	−91		
4, 60	17	20, 44	563		
5, 59	65	21, 43	560		
6, 58	24	22, 42	−299		
7, 57	−78	23, 41	−982		
8, 56	−100	24, 40	−382		
9, 55	26	25, 39	1039		
10, 54	158	26, 38	1394		
11, 53	90	27, 37	−355		
12, 52	−143	28, 36	−2488		
13, 51	−234	29, 35	−1698		
14, 50	5	30, 34	3334		
15, 49	313	31, 33	9778		

Table VIII.4 – Coefficients $g_{HH}(k)$ of the filter $G_{HH}(z)$ (Q15 fixed-point format)

Index *k*	$g_{HH}(k)$	Index *k*	$g_{HH}(k)$	Index *k*	$g_{HH}(k)$
0, 64	22	16, 48	207	32	16 762
1, 63	20	17, 47	370		
2, 62	−36	18, 46	−240		
3, 61	−18	19, 45	−503		
4, 60	44	20, 44	271		
5, 59	39	21, 43	681		
6, 58	−67	22, 42	−301		
7, 57	−59	23, 41	−930		
8, 56	88	24, 40	327		
9, 55	92	25, 39	1305		
10, 54	−115	26, 38	−348		
11, 53	−135	27, 37	−1950		
12, 52	144	28, 36	364		
13, 51	193	29, 35	3394		
14, 50	−175	30, 34	−374		
15, 49	−270	31, 33	−10403		

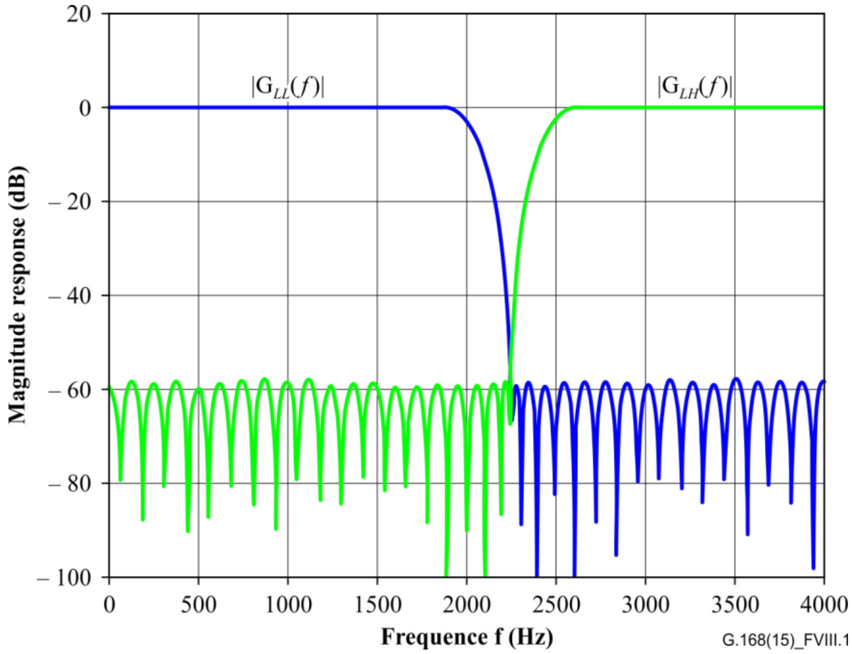


Figure VIII.1 – Magnitude responses of the filter pair ($G_{LL}(z)$, $G_{LH}(z)$) generated from the coefficients specified in Tables VIII.1 and VIII.2

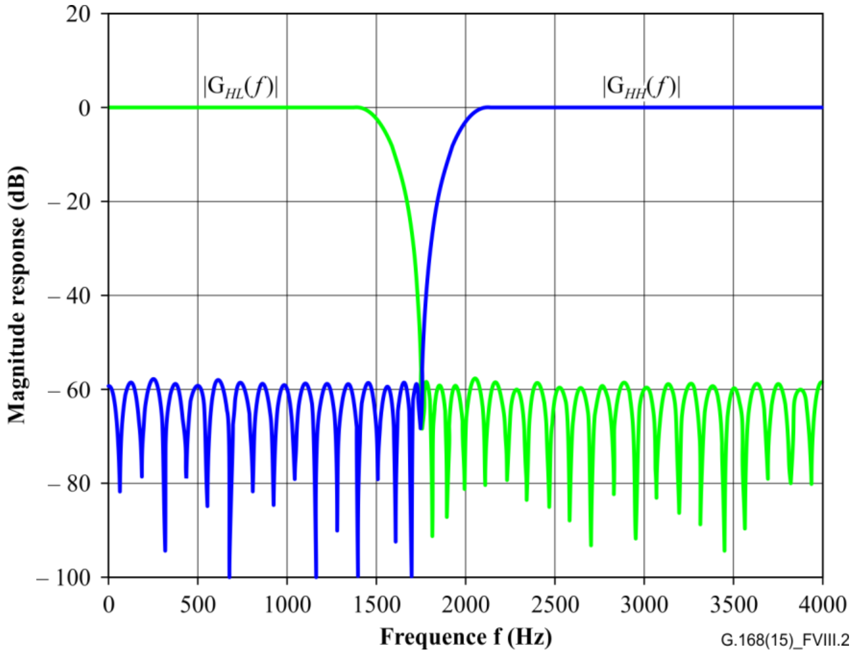


Figure VIII.2 – Magnitude responses of the filter pair ($G_{HL}(z)$, $G_{HH}(z)$) generated from the coefficients specified in Tables VIII.3 and VIII.4

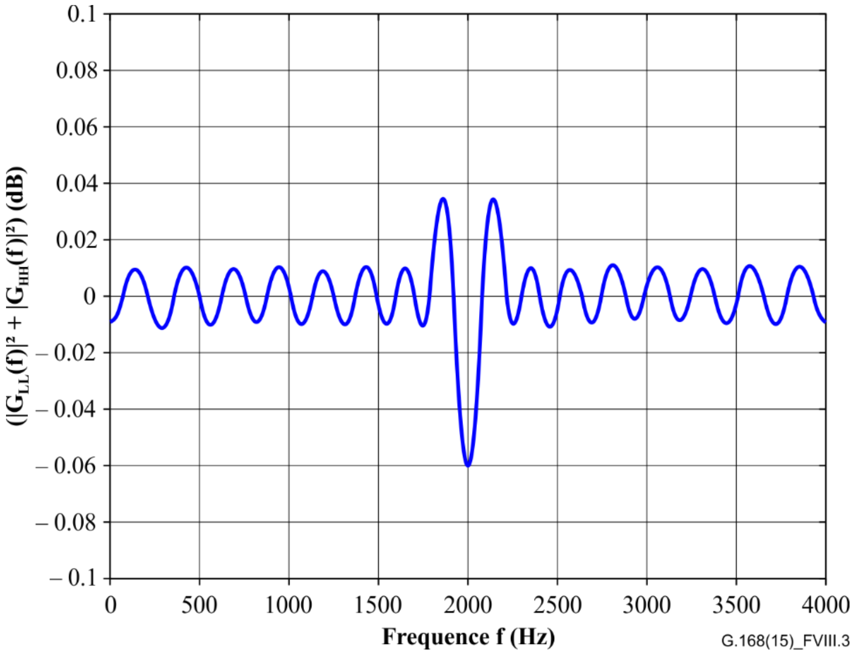


Figure VIII.3 – Total gain resulting from $G_{LL}(z)$, $G_{HH}(z)$

Bibliography

[b-ITU-T G.108.2] Recommendation ITU-T G.108.2 (2007), *Transmission planning aspects of echo cancellers*.

[b-ITU-T G.113] Recommendation ITU-T G.113 (2007), *Transmission impairments due to speech processing.*

[b-ITU-T G.126] Recommendation ITU-T G.126 (1993), *Listener echo in telephone networks*.

[b-ITU-T G.160] Recommendation ITU-T G.160 (2012), *Voice Enhancement Devices*.

[b-ITU-T P.341] Recommendation ITU-T P.341 (2005), *Transmission characteristics for wideband (150-7000 Hz) digital hands-free telephony terminals*.

[b-ITU-T P.342] Recommendation ITU-T P.342 (2000), *Transmission characteristics for telephone band (300-3400 Hz) digital loudspeaking and hands-free telephony terminals*.

[b-ITU-T P.561] Recommendation ITU-T P.561 (2002), *In-service non-intrusive measurement device – Voice service measurements*.

[b-ITU-T P.831] ITU-T Recommendation P.831 (1998), *Subjective performance evaluation of network echo cancellers*.

[b-ITU-T Q.55] Recommendation ITU-T Q.55 (1999), *Signalling between signal processing network equipments (SPNE) and international switching centres (ISC).*

[b-ITU-T V.32] Recommendation ITU-T V.32 (1993), *A family of 2-wire, duplex modems operating at data signalling rates of up to 9600 bit/s for use on the general switched telephone network and on leased telephone-type circuits*.

[b-ITU-T V.32 *bis*] Recommendation ITU-T V.32 *bis* (1991), *A duplex modem operating at data signalling rates of up to 14 400 bit/s for use on the general switched telephone network and on leased point-to-point 2-wire telephone-type circuits.*

[b-ITU-T V.61] Recommendation ITU-T V.61 (1996), *A simultaneous voice plus data modem, operating at a voice plus data signalling rate of 4800 bit/s, with optional automatic switching to data-only signalling rates of up to 14 400 bit/s, for use on the General Switched Telephone Network and on leased point-to-point 2-wire telephone type circuits.*

[b-ITU-T GSTP.CSS] ITU-T Technical Paper GSTP.CSS (2010), *The composite source signal as a measuring signal and a summary of various investigations on speech echo cancellers.*

[b-Borish] Borish, J. and Angel, J.B. (1983), *An efficient algorithm for measuring the impulse response using pseudo-random noise*, Journal of the Audio Engineering Society, Vol. 31, pp. 478-488.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	**Transmission systems and media, digital systems and networks**
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Terminals and subjective and objective assessment methods
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects and next-generation networks
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2015