INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

G.720

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

(07/95)

GENERAL ASPECTS OF DIGITAL TRANSMISSION SYSTEMS

CHARACTERIZATION OF LOW-RATE DIGITAL VOICE CODER PERFORMANCE WITH NON-VOICE SIGNALS

ITU-T Recommendation G.720

(Previously "CCITT Recommendation")

FOREWORD

The ITU-T (Telecommunication Standardization Sector) is a permanent organ of the International Telecommunication Union (ITU). The 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 Conference (WTSC), which meets every four years, establishes the topics for study by the ITU-T Study Groups which, in their turn, produce Recommendations on these topics.

The approval of Recommendations by the Members of the ITU-T is covered by the procedure laid down in WTSC Resolution No. 1 (Helsinki, March 1-12, 1993).

ITU-T Recommendation G.720 was prepared by ITU-T Study Group 15 (1993-1996) and was approved under the WTSC Resolution No. 1 procedure on the 10th of July 1995.

NOTE

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

CONTENTS

SUMMARY

In this Recommendation, an evaluation methodology is presented for the characterization of low-rate digital speech (or more generally source) coder performance with non-voice signals. The types of non-voice signals considered include voiceband data, network signalling, circuit continuity tones, and dual-tone multi-frequency signalling. Also addressed is the measurement of input/output delay and convergence time. Finally, appropriate test configurations, performance metrics, and test procedures are defined.

CHARACTERIZATION OF LOW-RATE DIGITAL VOICE CODER PERFORMANCE WITH NON-VOICE SIGNALS

(Geneva, 1995)

1 Introduction

In selecting speech coding algorithms for network applications, characterization of their performance with a diversity of telephone network signals is normally required. Typical non-voice signals whose in-band transmission over the telephone network is desirable include the following.

- Voiceband Data – Some transparency to voiceband data may be desirable particularly for applications such as thin-route satellite systems, digital mobile services, and circuit multiplication systems, such as DCME and PCME. It is considered that such transparency should be afforded to Public Switched Telephone Network (PSTN) voiceband data modems operating at bit rates up to and including 2400 bit/s. These modems include those conforming to Recommendations V.22, V.22 *bis*, V.21, V.23, V.26 *bis* and V.26 *ter*. It is recognized that other leased line and proprietary modems have found extensive use in the telephone network. Therefore, it is also desirable to use voiceband data signals from other modems which are not standardized by the ITU-T.

- Supervisory and inter-register signals that are incorporated within various signalling systems in the PSTN – These signals include two frequencies (2400 Hz and 2600 Hz) present within Signalling System No. 5 for line signalling, 15 multi-frequencies used by Signalling System No. 5 for register signalling, and all valid combinations (pairs) of the six high-frequency and six low-frequency tones employed by Signalling System R-2.

- Dual-Tone Multi-Frequency (DTMF) signals – These signals may be used independently of a signalling system to provide end-user access to specialized services such as answering machine message retrieval, airline reservations and electronic banking.

- Tones employed by Signalling Systems No. 6 and No. 7 for the verification of end-to-end circuit continuity and the four high-group and four low-group frequencies, conforming to Recommendation Q.23.

- The 2100 Hz tone (including any associated phase reversals) specified in Recommendation V.25, which is employed by the V-Series modems for disabling network echo control devices.

- Finally, the tones utilized in national signalling systems to provide ringing, busy or congestion information (see Recommendation Q.35) could require support depending upon the location of the coder within the network (such as in land-mobile environments) and should therefore be included in these tests.

This methodology has evolved over several years from performance evaluations conducted during the development of Recommendations G.726, G.727 and G.728.

2 Scope

This Recommendation encompasses methods that can be used for quantifying and characterizing the performance of low-rate digital speech codecs with non-voice signals. The signals used are primarily those employed within the telephone network and, consequently, these methods are most appropriate for speech codecs intended for network (as opposed to end-user terminal) applications. The methods presented here can be used to quantify the performance of voice codecs that operate in the telephone band and that employ transmission rates that are less than or equal to 64 kbit/s.

3 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions to this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; all 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 ITU-T Recommendations is regularly published.

- ITU-T V-Series of Recommendations, *Data communication over the telephone network.*
- ITU-T Recommendation Clause 2 Q.141 (1993), *Line signalling – 2.1 signal code for line signalling.*
- CCITT Recommendation Q.151 (1988), *Signal code for register signalling.*
- CCITT Recommendation Q 441 (1988), *Signalling code.*
- CCITT Recommendation Q.271 (1988), *General.*
- CCITT Recommendation Q.724 (1988), *Signalling procedure Specifications of Signalling System No. 7*
- CCITT Recommendation Q.23 (1988), *Technical features of push-button telephone sets.*
- CCITT Recommendation V.25 (1988), *Automatic answering equipment and/or parallel automatic calling equipment on the general switched telephone network including procedures for disabling of echo control devices for both manually and automatically established calls.*
- CCITT Recommendation Q.35 (1988), *Technical characteristics of tones for the telephone service.*
- CCITT Recommendation G.711 (1988), *Pulse Code Modulation (PCM) of voice frequencies.*
- CCITT Recommendation G.726 (1990), 40, 32, 24,16 kbit/s *Adaptive differential pulse code modulation (ADPCM).*
- CCITT Recommendation O.42 (1988), *Equipment to measure non-linear distortion using the 4-tone intermodulation method.*
- CCITT Supplement No. 3.1, *Measuring instrument requirements. Sinusoidal signal generators and level measuring instruments,* Green Book, Vol. IV.2, Supplement to Series-M, -N and -O Recommendations, pp. 530-533, Geneva, 1972.
- CCITT Supplement No. 3.2, *Noise measuring instruments for telecommunication circuits*, Green Book, Vol. IV.2, Supplement to Series-M, -N and -O Recommendations, pp. 534-548, Geneva, 1972.
- ITU-T O-Series of Recommendations, *Specifications for measuring equipment*.
- CCITT Recommendation Q.112 (1988), *Signal levels and signal receiver sensitivity.*
- CCITT Recommendation Q.114 (1988), *Typical transmission requirements for signal senders and receivers.*
- CCITT Recommendation Q.451 (1988), *Definitions.*
- CCITT Recommendation Q.454 (1988), *The sending part of the multifrequency signalling equipment.*
- CCITT Recommendation Q.455 (1988), *The receiving part of the multifrequency equipment. Range, speed, and reliability of interregister signalling.*
- CCITT Recommendation Q.143 (1988), *Line signal sender.*
- CCITT Recommendation G.144 (1988), *Line signal receiver.*
- ITU-T Recommendation G.165 (1993), *Echo cancellers.*
- ITU-T Recommendation O.133 (1993), *Equipment for measuring the performance of PCM encoders and decoders.*
- CCITT Recommendation O.91 (1988), *Phase jitter measuring equipment for telephone-type circuits.*

– CCITT Recommendation O.81 (1988), *Group-delay measuring equipment for telephone-type circuits.*

– CCITT Recommendation O.131 (1988), *Quantizing distortion measuring equipment using a pseudo-random noise test signal.*

– CCITT Recommendation O.132 (1988), *Quantizing distortion measuring equipment using a sinusoidal test signal.*

– ITU-T Recommendation E.450 (1993), *Facsimile quality of service on PSTN – General aspects.*

– ITU-T Recommendation E.451 (1993), *Facsimile call cut-off performance.*

– ITU-T Recommendation E.452 (1993), *Facsimile modem speed reductions and transaction time.*

4 Abbreviations

For the purposes of this Recommendation, the following abbreviations are used:

ASCII American Standard Code for Information Interchange

BER Bit Error Ratio

CEPT European Committee on Post and Telecommunications

CHER Character Error Ratio

DCME Digital Circuit Multiplication Equipment

DTMF Dual-Tone Multi-Frequency

FIR Finite Impulse Response

MBER Modem Bit Error Ratio

MBLER Modem Block Error Ratio

MSUR Modem Missed Start-Up Ratio

PCM Pulse Code Modulation

PCME Packet Circuit Multiplication Equipment

PSTN Public Switched Telephone Network

S/N Signal-to-Noise Ratio

SNR Signal-to-Noise Ratio

TLP Transmission Level Point

5 Voiceband data performance

In this clause, the procedure to be used for evaluating a speech codec's performance with 2- and 4-wire modems is described.

5.1 Test Set-up Architecture

An architecture to be employed for voiceband data modem measurements should be centred around a Bit Error Ratio (BER) test system which can provide the source data for the transmit direction of the modems (Figure 1). This system should permit the analysis of the receive modem data in order to compute single errors, blocks, block lengths, block errors, test duration, and error-free seconds, all of which are parameters that can be used in the interpretation of the received signal quality.

In addition to the modems and BER test set, a reference 64 kbit/s Pulse Code Modulation (PCM) codec, the speech codec under evaluation, and an analogue impairment channel should be employed. All measurements shall then be performed with the analogue channel in tandem with the digitally (speech) encoded channel in the forward direction of transmission, with the input to the digitally encoded channel defined as the 0 Transmission Level Point (0 Transmission Level Point – TLP). In this Figure 1, the signal levels are defined with the analogue channel impairments disabled, and 0 = 0 dBm0, which makes the data level equal to –15 dBm or –15 dBm0.

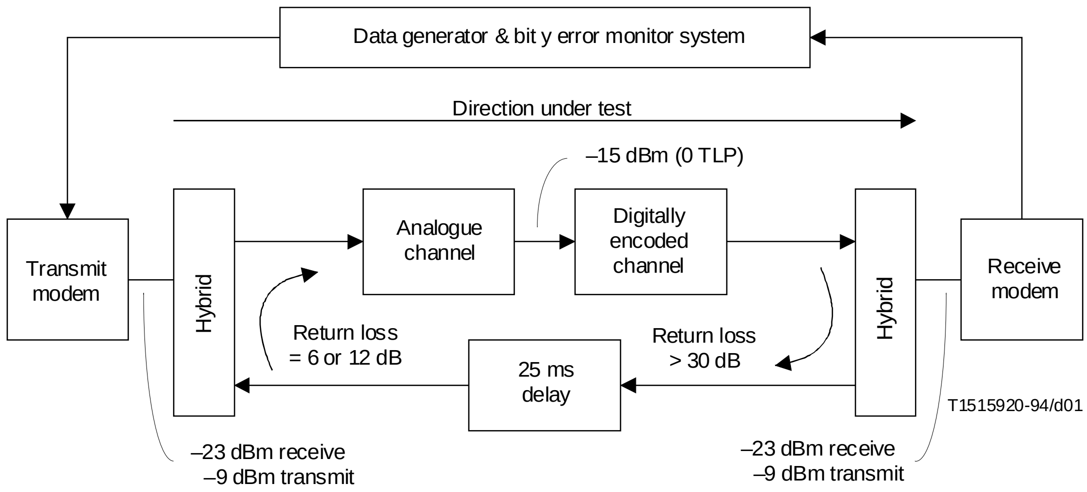

FIGURE 1/G.720
Modem test configuration architecture

5.2 Impairments

The analogue channel shall permit the introduction of two types of network impairments: variable and fixed. Variable impairments consist of the linear addition of controlled amounts of band-limited analogue white noise prior to the digitally encoded channel. Fixed impairments consist of group and attenuation distortion, phase jitter, and second and third order non-linear distortion. The impairments to be injected in the analogue channel are in accordance with the configuration shown in Figure 2.

The values of these impairments are given in Tables 1 and 2.

These impairments correspond to stringent network requirements. However, they provide a frame of reference against which the speech coding algorithms can be compared with other ITU-T speech coding Recommendations. It is noted that in measuring second and third order non-linear distortion (Table 1), a four-tone method shall be used which employs two pairs of equi-level tones with a composite power equal to that of the voiceband data modem signal. In the first tone pair (denoted by A), the tones shall be 6 Hz apart and centred at 860 Hz. In the second pair (denoted by B), the tones shall be 16 Hz apart and centred at 1380 Hz. The second order non-linear distortion shall be determined by measuring the power of the four B + A and four B – A intermodulation products, and the third order shall be determined by measuring the six 2B – A products.

NOTE – The non-linear distortion unit provides, for every 1 dB increase in the input signal, an increase in the output second-order component of 2 dB, and an increase in the output third-order component of 3 dB.

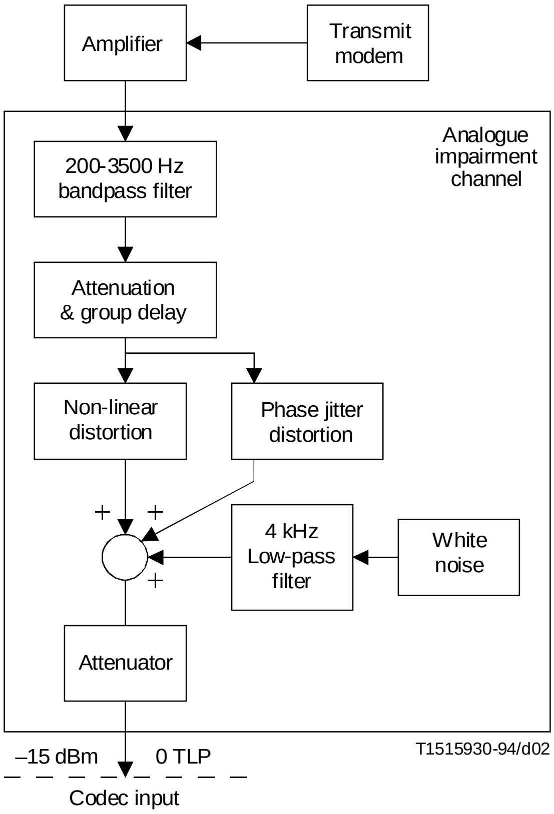

FIGURE 2/G.720

Analogue impairment channel

TABLE 1/G.720

Analogue impairments for voiceband data tests

Analogue impairments	Impairment level
Group delay distortion	See Table 2
Attenuation distortion	See Table 2
Second order non-linear distortion	36 dB
Third order non-linear distortion	38 dB
Phase jitter (degrees) at 120 Hz	3 peak to peak
Additive noise (measured C-message weighted)	Varied as a parameter

TABLE 2/G.720

Added attenuation and group distortion

Frequency (Hz)	Attenuation (dB relative to 1000 Hz)	Group delay (µs relative to 1800 Hz)
200	8.5	4645
300	3.0	3442
400	0.9	2436
500	0.2	1703
600	0.1	1200
700	0.0	857
800	0.0	613
900	0.0	440
1000	0.0	312
1100	0.0	223
1200	0.1	153
1300	0.2	101
1400	0.3	61
1500	0.5	33
1600	0.6	14
1700	0.7	3
1800	0.9	0
1900	1.1	8
2000	1.3	27
2100	1.6	60
2200	2.0	109
2300	2.3	174
2400	2.7	262
2500	3.0	376
2600	3.3	517
2700	3.6	697
2800	3.9	934
2900	4.3	1247
3000	4.8	1616
3100	5.2	1969
3200	5.6	2263

5.3 Measurement Procedures

In order to assure consistency in the manner in which measurements are collected and reported, the following definitions shall be adopted:

- *Codec Input Data Level* – This is defined as the root mean square (r.m.s.) unweighted level of the data signal at the input to the coder. The data level is determined with the analogue noise injection disabled.
- *Noise Level* – This is provided by an r.m.s. psophometrically weighted measurement of the channel noise at the input to the codec. This measurement shall be performed with the noise injection enabled and the modern data signal disconnected.

- *Data Signal-to-Noise Ratio* (SNR) – This shall be computed as the ratio of the codec input data level to the measured noise level.

- *Bit Error Ratio (BER)* – This is defined as the ratio of the number of bits intentionally corrupted in the coder digital channel in a fixed time interval, over the total number of bits transmitted over the same coder channel and over the same time period.

- *Modem Bit Error Ratio (MBER)* – This is defined as the ratio of the number of bits received in error by a modem receiver in a fixed time interval, over the total number of bits received by the same modem receiver over the same time period.

- *Modem Data Block* – A block of modem data is defined as a contiguous stream of 511 data bits.

- *Modem Block Error Ratio (MBLER)* – This is defined as the number of blocks received with at least 1 bit in error by a modem receiver in a fixed time interval, over the total number of blocks received by the same modem receiver over the same time period.

- *Modem Missed Start-Up Ratio (MSUR)* – This is defined as the total number of switched carrier messages received with at least one bit in error by a modem receiver in a fixed time interval, over the total number of switched carrier messages received by the same modem receiver and over the same time period. The switched carrier messages should consist of the three characters: SYN, SYN, and EOT.

- *Character Error Ratio (CHER)* – This is defined as the total number of asynchronous ASCII characters received with at least one bit in error by a modem receiver in a fixed time interval, over the total number of asynchronous ASCII characters received by the same modem receiver and over the same time period.

5.4 Transmit and echo levels

In order to provide a realistic two-wire network environment, the loss of the transmit side hybrid shall be adjusted so that the echo return loss is equal to 6 dB (12 dB for the V.26 *ter* modem), measured at the output of the analogue channel's bandpass filter. The return loss across the receive side hybrid shall be greater than 30 dB. For modems employing echo cancellation, 800 ms of transmission delay shall be introduced into the receive side path so that any echo control flat delay compensation can be exercised.

5.5 Performance measurements

Once the configuration and calibration of the facility has been completed, a series of measurements can be taken to quantify the performance of various modems over the digitally speech encoded channel of interest.

5.5.1 Switched carrier mode

- Each Measurement shall be made until either 20 000 messages are transmitted without any messages missing or being in error, or 100 messages are missing or are in error, or 100 000 messages are transmitted with less than 100 messages missing or being in error, or if it is apparent that the Modem Missed Start-Up Ratio (MSUR) versus SNR relationship is approaching an asymptote.

- The message itself shall consist of three ASCII characters (SYN, SYN, EOT) separated by a predetermined time gap ms, preceded by a modem-dependent training sequence. The value of the time gap is for further study.

- Tests shall be operated with all modems in a multipoint mode.

5.5.2 Continuous carrier mode

- Each measurement shall be made until either 20 000 blocks are received without any blocks in error, or 100 blocks are received in error, or 100 000 blocks are received with less than 100 errors, or if it becomes apparent that the MBLER versus S/N relationship is approaching asymptote.

5.5.3 Character mode

- In order to generate a character sequence for transmission over the evaluated link, a number of files need to be generated off-line using suitable computer equipment. The first file should be generated to contain randomized sequences of all printable keyboard characters. Subsequently a “word” template file is to be constructed containing randomized grouping information of characters into words, with 1 to 10 characters per word. Finally, a third file containing inter-word delay information is generated, wherein delays are randomly varied between 16 and 320 ms with a granularity of 16 ms. All three files are generated only once and then used for all character mode tests.

- The keyboard character file is then transmitted over an RS-232 interface to the asynchronous transmit modem with program control of word delimiting and inter-word delay as specified in the word and delay template files respectively. In addition, a fixed 48 ms delay shall be inserted between each character within a word. Characters received over the RS-232 interface from the receiving modem will be stored in an output file. Upon reception, the characters are read sequentially, ignoring word delimiting and delays. As a result, a raw output file is written and its contents can be compared against the original input file for errors.

- In order to assure sufficient statistical significance in the results obtained, at least 20 000 characters should be transmitted over the link under evaluation.

5.6 Calibration

The calibration procedures need to be further defined. The procedures used in the characterization of Recommendations G.726 and G.728 can be used as a basis in the interim.

6 Narrow-band signal performance

As identified in the introduction, supervisory and interregister signals (that is, signals that are used to monitor the status of a call and signals that are used to carry service addresses across the network) are incorporated within various PSTN signalling systems. In view of the potential use of speech coding within the PSTN, transparency to such signals may be desirable. This clause describes the methods to be adopted in characterizing the performance of speech coders with narrow-band signals. Although not specifically addressed in this contribution, all instrumentation employed shall meet or exceed the requirements identified in Supplements 3.1 and 3.2 to the Series-M, -N and -O Recommendations.

One of the first problems to be addressed is the selection of an acceptable method for determining whether acceptable transparency to narrow-band signals can be achieved (and, if so, at what margin). Two principal philosophies are identified, one based on the use of network sending and receiving equipment (method A), and the other (method B) based on an analytical comparison of the input and output waveforms transmitted through the digitally (speech) encoded channel. In order to permit some flexibility both methods are defined wherever possible. Which of the two methods is to be selected subsequently depends on whether senders and receivers are available, and/or on whether satisfactory methods for waveform analysis can be identified and implemented. In general, waveform analysis methods are considered preferable when suitable received signal tolerance specifications could be identified.

Since the methods that employed signal senders and receivers comprised the same test architecture as that employed for voiceband data modem measurements, the architecture employed for network signalling is not repeated here.

6.1 System No. 5 inter-register signalling

One concern with Signalling System No. 5 is the possible violation of the transmitter's tolerance limits of 55 ± 5 ms when low-rate-encoded channels employ block processing frames of duration of the order of 5 ms. In assessing inter-register signalling transparency, the six frequencies of 700, 900, 1100, 1300, 1500 and 1700 Hz shall be used in appropriate combinations to generate the start of pulsing KP1 and KP2 signals, the end of pulsing ST signal, the Code 11 and Code 12 operator response signals, and the ten digits of 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 (Table 2/Q.151). The waveform sequence, whose degradation is to be assessed, shall then be constructed by transmitting the KP signal followed by 20 repetitions of a single digit, followed by the ST signal. This shall be repeated 12 times for each of the digits 1, 2, 3, 4, 5, 6, 7, 8, 9 and Codes 11 and 12. For Signally System No. 5, measurements shall be undertaken using a signal level of –7 dBm per frequency (–4 dBm composite) and by subsequently adjusting the calibration procedures accordingly.

Method A

This method necessitates the use of a System No. 5 inter-register signalling generator and the use of controllable impairments identical to the ones to be used for voiceband modem testing. The procedures to be adopted shall be as follows:

- Measurements shall be obtained with the codec in series with the impairment configuration. The level of impairments shall be equal to that used for modem testing (Tables 1 and 2). The non-linear distortion simulator shall be calibrated using a 4-tone signal at a level equal to that of the composite signalling tone-pair. Injected analogue noise shall be linearly added as the last impairment to the impairment configuration and prior to the codec input. The level of noise shall be varied and received inter-register signalling error rate as a function of injected noise shall be assessed at a composite signalling tone level of –4 dBm0 and a composite nominal codec input level of –4 dBm0.

- Measurements shall also be performed as above, but with the analogue noise injection disabled. This time the peak-to-peak jitter shall be varied and the received inter-register signalling error rate as a function of jitter shall be assessed.

Method B

The second technique addresses the analysis of the input and output waveforms transmitted through the coder channel under evaluation (with and without the analogue impairments of Tables 1 and 2 – injected noise disabled). Here, the following measurements are required:

- Mean duration and mean amplitude of the active and silent signalling intervals.

- Variance of the duration and variance of the amplitude of the active and silent signalling intervals.

- A silent interval is characterized by the fact that its waveform energy is less than a predetermined percentage of the peak waveform energy during the active intervals. The value of the predetermined percentage is for further study. A sliding 2 ms window is used for this determination.

- An active interval is characterized by the fact that its waveform energy is within predetermined percentage of the peak waveform energy during the active intervals. The value of the predetermined percentage is for further study. A sliding 2 ms window is used for this determination.

- Distortion in the attack and fall times of each of the tones leading to the non-temporally coincident start and termination of each frequency from a register tone pair. It is noted that a limit of 1 ms is specified as the maximum signal sender tolerance for the non-temporarily coincident start and termination of the tone pair Q.153.

- Level differences between the two tones. A limit of 4 dB in absolute level differences between the two unmodulated signal frequencies is imposed at signal senders (Recommendation Q.154).

6.2 R-2 Signalling

Signalling System R-2 is designed to use six signalling frequencies (1380, 1500, 1620, 1740, 1860 and 1980 Hz) in the forward direction and six signalling frequencies (1140, 1020, 900, 780, 660 and 540 Hz) in the backward direction of transmission. A 2-out-of-n code (Table 5/Q.441) then composes an inter-register signal consisting of the simultaneous transmission of a pair of frequencies.

In conducting an assessment of inter-register signalling, the 15 forward and ten backward multi-frequencies [Table 5/Q.441, Note a)] shall be used. These frequencies shall be employed in appropriate combinations to generate the ten digits of 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9 with nominal duration values in accordance to those defined in Recommendation Q.451. Finally, the signal sequence whose degradation is to be assessed shall consist of 20 repetitions of a single digit. This process was repeated ten times to generate each of the digits 1, 2, 3, 4, 5, 6 ,7, 8, 9 and 0. Unlike Signalling System No. 5, where the signal levels are equal to –7 dBm per frequency, the nominal signal level at the input to the digitally encoded channel is defined to be equal to –11.5 dBm per frequency (or –8.5 dBm for the composite signal).

Method A

This method necessitates the use of a System R-2 inter-register signalling generator and the use of controllable impairments identical to the ones defined in conjunction with voiceband modem testing. The procedures to be applied shall be the same as in method A of Signalling System No. 5 inter-register testing with appropriate calibration adjustments to account for differences in signal levels. Thus, the level of noise shall be varied and the received inter-register signalling error rate as a function of injected noise shall be assessed at a composite signalling tone level of –8.5 dBm0.

Method B

The second technique addresses the analysis of the received signal which is processed through the digitally (speech) encoded channel (with and without the analogue impairments of Tables 1 and 2 – injected noise disabled). Here the following measurements are required:

- Level between differences between the two tones. A limit of 1 dB in absolute level differences between the two unmodulated signal frequencies is imposed at signal senders (Recommendation Q.454) and a limit of 3 dB is imposed for signal receivers (Recommendation Q.455).

- Distortion in the attack and fall times of each tone leading to the non-temporally coincident start and termination of each frequency from a register tone pair. It is noted that a limit of 1 ms is specified as the maximum signal sender tolerance for the non-temporally coincident start and termination of the tone pair Recommendation Q.454.

6.3 DTMF Signals

DTMF signals are often used independently of a signalling system and after a call has been established to provide end-user access to specialized services. The DTMF signals which should be used to characterize the performance of a speech coder under evaluation shall be transmitted at a composite level of –4 dBm employing a minimum digit separation of 40 ms, and comprise pairs of tones selected from a high- and a low-group frequency set. The low-group frequencies are 697, 770, 852 and 941 Hz. The high-group frequencies are 1209, 1336, 1477 and 1633 Hz. Appropriate combinations of these frequencies shall then be used to derive the ten digits 1, 2, 3, 4, 5, 6, 7, 8, 9, and 0, and the characters *, #, A, B, C and D (see Figure 1/Q.23).

In measuring the performance of the speech coder under evaluation with DTMF signalling, the correct recognition percentage, as well as the variation of signal and signal-gap duration as a function of SNR were measured (the noise was injected in the analogue channel prior to codec input). It is noted that satisfactory performance can typically be assured when the total distortion products (resulting from harmonics and inter-modulation) does not exceed 20 dB below the fundamental frequencies.

6.3.1 Balanced pair

The signal duration for each DTMF tone pair shall equal 50 ms and the separation between adjacent tone pairs shall equal 40 ms. In addition, the start and termination of the two tones, in a given DTMF tone pair, shall be temporally coincident. In assessing performance with DTMF tones, the following shall be measured (both with and without analogue impairments, and with analogue noise and bit error injection disabled).

- Total distortion products (resulting from inter-modulation), which must not exceed 20 dB below the fundamental frequencies.

- The correct recognition percentage of DTMF tone pairs when analogue noise is injected prior to the codec input so that the reference input signal to noise ratio is equal to 20 dB.

6.3.2 Twisted pair

- This method of DTMF signal pair generation employs the same timing relationship for the "on" and "off" cycles as above. In this case however, one of the two tones shall be 6 dB higher than the other. Subsequently, the following shall be measured.

- The correct recognition percentage of DTMF tone pairs when analogue noise is injected prior to the codec input so that the reference input signal to noise ratio is equal to 22.5 dB.

6.4 System No. 5 line signalling

In addition to inter-register signalling which was addressed above, line signalling is employed in the network. Signalling System No. 5 line signals are comprised of two frequencies: 2400 and 2600 Hz, used either separately or combined.

In assessing single-tone performance, measurements should be undertaken with and without analogue impairments. No bit errors shall be injected into the coder channel under evaluation. The analogue impairment channel shall be adjusted (using the output attenuator) so that the level of each tone at the encoder input is equal to –9 dBm0. The following shall then be measured:

- frequency offset; and

- total distortion.

It is noted that the requirement regarding frequency offsets impose a tolerance mask of ± 6 Hz for signal senders (Recommendation Q.143) and ± 15 Hz for signal receivers (Recommendation Q.144). Recommendation Q.143 also imposes a signal to total distortion ratio limit of 40 dB for signal senders. (Recommendation O.132 deals with relevant methods for performing these measurements.)

Compound tone

This is defined by the simultaneous presence of the 2400 and 2600 Hz tones and is used to signify a clear forward and a release guard line condition. Performance with this compound tone shall be assessed without bit errors injected into the coder link under evaluation. The analogue impairment channel shall be adjusted so that the level of each tone is the same and equal to –9 dBm0 per frequency at the input to the encoder (this calibration is made without analogue impairments). The start and termination of the two tones shall be temporarily coincidental. The following shall then be measured:

- any distortion in the attack and fall times of each tone leading to the non-temporally coincidental start and termination of this tone pair;

- total distortion; and

- level differences between the two tones.

It is noted that a limit of 5 ms is specified as the maximum signal sender tolerance for the non-temporally coincident start and termination of the tone pair Q.141. Further, a limit of 5 dB in absolute level differences between the two unmodulated signal frequencies is imposed at signal receivers (Recommendation Q.144). The requirements for total distortion (in this case this will include inter-modulation components) are the same as above.

6.5 Circuit Continuity Tones

Because Signalling Systems Nos. 6 and 7 are transmitted out of band, facilities are normally provided for making a continuity check of the speech path prior to making an end-to-end connection available to the end-users. This check is undertaken by transmitting a 2000 ± 20 Hz tone at a –12 dBm level. For these tests, a tone duration of 30 ms is defined, and the following shall be measured:

- Total distortion of the channel output signal.

- Tone duration of the channel output signal. A tone is determined as being present if its energy is within predetermined percentage of the peak waveform energy of the tone during the active signal period. The value of the predetermined percentage is for further study. A 2.5 ms window shall be employed for this determination.

6.6 Echo canceller disabling tone

A special narrow-band signal (which is often associated with voiceband data and Group 3 facsimile transmissions) may be transmitted over the network after a call has been established in order to disable echo canceller devices. This tone comprises a 2100 Hz signal with 180° phase reversals every 400 ms, as defined in Recommendation G.165.

In assessing the ability of a digitally (speech) encoded channel to transmit the 2100 Hz signal and associated phase reversals, the requirements defined in Annex B/G.165 are noted, particularly as regards the fact that phase reversals are only considered valid if they fall within ± 70 of the transmitter's 180° phase change.

Method A

This method uses a sufficient number of signal receivers to assess the performance degradation due to processing through a coder channel with and without bit slips in the digital channel.

- Use will be made of an echo canceller employing a phase sensitive tone disabler. Subsequently the V.25 tone degradation will be assessed by observation of the number of times the tone disabler fails to function properly. In order to obtain statistical significance in these results, it is important to replicate these measurements a minimum of 50 times. Furthermore, in obtaining each measurement, the system under evaluation should be reset, or should this capability not be provided, the tone should be disconnected and reconnected for each measurement.

- These measurements should be repeated with and without the fixed analogue impairments defined in Tables 1 and 2. It will be important to employ at least four different tone disabling cancellers, each being supplied from a different manufacturer.

Method B

This method depends on the recording and analysis of the waveform transmitted through the coder channel under evaluation in order to measure the phase reversals that should be occurring every 400 ms. In practice, since the 180° phase change will eventually always be reached, the parameter measured can be the time taken to settle to the 180° phase state following the actual transition in the transmitted waveform.

In order to assess this type of degradation, a modulated 2100 Hz sine wave consisting of a sequence of 180° phase transitions occurring at 400 ms intervals shall be processed over the digitally (speech) encoded channel. In addition, this same waveform shall be recorded (after phase alignment) through a direct (unprocessed) recording path. The two waveforms may then be compared, and the settling time of the processed waveform after the phase transition measured.

The two waveforms are compared in accordance with the functional configuration shown in Figure 3. Subsequently, the settling time of the processed waveform in the neighbourhood of the phase transition is measured.

6.7 Distortion, frequency and level

In order to assess the ability of a digitally-encoded channel to transcode with 64 kbit/s PCM, a series of measurements aimed at the characterization of non-linear distortion, dynamic range, and quantization distortion may be undertaken. These include the following:

- Second- and third-order non-linear distortion employing a four-tone test signal at a codec input signal level of –4 and –15 dBm.

- Idle channel noise measurements taken psophometrically weighted.

- Input/output gain with an 820 Hz and/or a 1020 Hz and/or four-tone test signals (see bibliography) over a codec input signal range of +3 to –50 dBm (measured at 0 TLP) using the procedures defined in Recommendation O.133.

NOTE – A reference to an 820 Hz and/or a 1020 Hz tone is meant to imply a tone whose frequency is in the neighbourhood of 800 and 1000 Hz but which is not a sub-multiple of 8 kHz.

- Phase jitter, amplitude jitter, second and third order harmonic distortion employing a 1 kHz tone at a codec input signal level of –4 and –15 dBm.

- Group delay and amplitude response variation with frequency at a codec input signal level of –15 dBm over the telephone voiceband spectrum of 200 to 3400 Hz. Recommendations O.81 and O.133 address relevant procedures to be used in taking these measurements.

- Signal-to-quantization distortion over a codec input signal range of +3 to –50 dBm (measured at 0 TLP) employing an 820 Hz test signal measured psophometrically weighted and/or a 1020 Hz test signal measured C-message weighted. Recommendations O.131 and O.132 address relevant procedures to be used in taking these measurements.

- Signal-to-quantization distortion over a codec input signal range of 0 to –50 dBm (measured at 0 TLP) employing a pseudo-random noise generator. Recommendation O.131 addresses relevant procedures to be used in making these measurements.

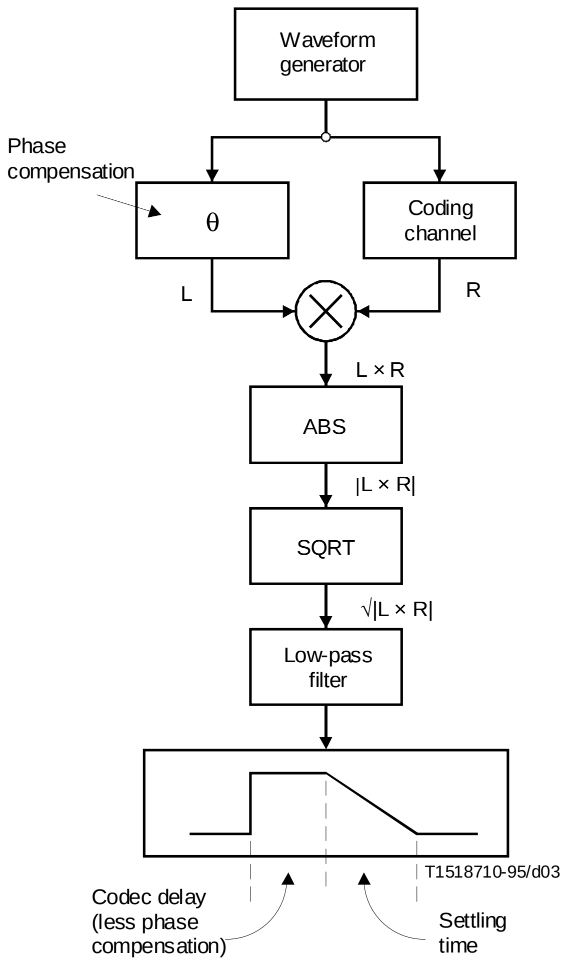

FIGURE 3/G.720

Functional diagram for echo canceller phase distortion measurement configuration

7 Coding delay

In measuring processing delay, it should be noted that in waveform coders, measurements of delay rarely present a problem, because a direct assessment of the elapsed time between the signal onset at the codec input and the signal appearance at the codec output is possible (e.g. using a modulated sine wave). However, in block processing coders (or when little is known about the actual technique employed) this approach may be unreliable. This is because codec performance could vary substantially with the nature and type of signal being encoded, as well as with the precise time at which a signal appears at the codec input (in relation to the coder's frame structure, if there is one).

The technique presented below for measuring the input/output delay of digitally (speech) encoded channels employs artificial speech and overcomes the previously mentioned narrow-band signal limitation.

7.1 Total delay measurement method

This method permits the direct measurement of the total delay introduced by a digitally (speech) encoded channel between its input and output 8-bit companded PCM interfaces. This delay is comprised of a delay introduced by the algorithm encoding and decoding functions, as well as any additional delay introduced by the processing or specific algorithm hardware implementation.

The delay is measured by observing the first peak of the cross-correlation function between the encoder input and decoder output signals. This is determined by examining the peak response of an adaptive Finite Impulse Response (FIR) filter in a joint linear-process estimator configuration, as shown in Figure 4. In order to obtain a delay measurement which represents as closely as possible the delay likely to be experienced in real conversations (and at the same time retain the reproducibility of the measurement), an artificial speech signal shall be employed at a nominal codec input level of –22 dBm.

As can be seen from Figure 4, the signal provided as reference to the adaptive filter is obtained from the input and 64 kbit/s μ-law PCM interface to the speech codec. Similarly, the point where the adaptive filter output is subtracted from the speech codec signal is defined as the codec's 64 kbit/s μ-law PCM output interface.

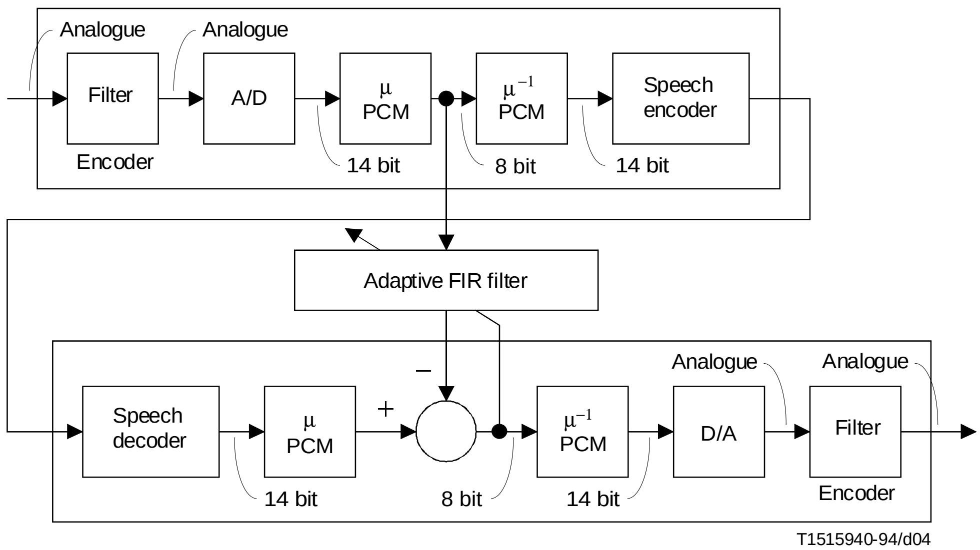

FIGURE 4/G.720

Block diagram of delay measurement architecture

8 Convergence time

Assuming that the encoder input signal characteristics have remained steady for a sufficient amount of time and assuming that disturbances causing loss of block synchronization have not occurred, two distinct phenomena can cause the decoder of an encoder-decoder pair to acquire an abnormal state. In addition, a convergence time is associated with a coder's tracking capability in the case of non-stationary signals. This gives rise to a third case where convergence is an issue. The three cases are as follows.

8.1 Encoder/decoder decorrelation

This case is generated when a burst of digital errors is presented to the encoder to decoder link, that are sufficiently long to cause a decorrelation between the encoder input and the decoder output to occur. Further, these burst errors should not cause loss of block synchronization. In this case, convergence time is determined by measuring the time needed for the "input signal" to "ourput noise" ratio to return to 90% of its steady state value after termination of the error burst. (The length of burst errors, the type of signals to be employed for this measurement, as well as the level of input signal to output noise ratio, under steady state conditions, are for further study.)

8.2 Encoder/decoder state mismatch

This case is generated when a decoder is connected to a transmitting encoder such that the encoder and decoder states are mismatched. This, for example, can occur when a decoder is disconnected or is switched between different encoded channels. In this case, convergence time can be assessed by measurement of the time necessary to achieve 90% of the steady state value of the "input signal" to "output noise" ratio after encoder to decoder reconnection. Should the decoder be provided with an independent reset capability, this case can also be more efficiently generated by resetting the decoder function, while maintaining block synchronization.

8.3 Attack time

This case is relevant to the attack time of the coding process when a new signal is presented or a signal present for some time is terminated at the encoder input. In this case, convergence time can be assessed by measurement of the time necessary for the "input signal" to "output noise" ratio to rise to 90% or fall to 10% of its steady state value after the appearance or after the termination of a sinusoidal signal at the encoder. It is assumed that prior to the appearance and after the termination of this signal, the encoder input signal is comprised of digital zeros. This measurement addresses the attack and fall times of sinusoidal signals. Of particular interest is the fall time of the 2100 Hz tone, since excessive distortion may affect the performance of the V.25 automatic answering sequence and the T.30 facsimile protocols. As a consequence, a 2100 Hz tone should be employed as a minimum in obtaining these measurements.

8.4 Measurement method

Unlike coder delay which is only likely to be impacted during the first frame (if there is one) after a signal has appeared at the encoder input, convergence time may be found to be signal dependent. As a consequence, a sufficiently exhaustive range of signals should be employed for this purpose.

The signals to be used for this measurement should include tones and voice signals. At least ten tones spanning the 200 to 3200 Hz frequency range should be used and input signal power levels of 0, –15 and –30 dBm0 should be employed. Measurement of signal to noise ratio should be accomplished by application of a sliding window with predetermined duration. Since the sliding window will probably be of a duration that is comparable to the convergence time being measured, a large number of replicated measurements (e.g. 50) may be needed in order to obtain statistical significance. The duration of sliding window is for further study.

In conducting these measurements, it is assumed that the input and output codec signals will be recorded and analysed digitally. Further, proper signal alignment will be required, this being accomplished by knowledge of the codec processing delay or by using the method described in clause 7.

9 Facsimile performance

The E.450-Series of Recommendations provide a comprehensive methodology for measuring facsimile performance over the telephone network. These Recommendations shall also be applied, as appropriate, in measuring performance of Group 3 facsimile over digitally (speech) encoded channels. In this case, the equipment configurations presented in clause 5 for voiceband data measurements (Figure 1) continue to apply, with the exception that the voiceband data modems shall be replaced by facsimile terminals.

Bibliography

Bell System Technical Reference: *Transmission parameters affecting voiceband data transmission-measuring techniques*, Publication 41009. Section 3.7 Intermodulation (non-linear) Distortion, May 1975.