INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.810

(02/96)

TELEPHONE TRANSMISSION QUALITY

METHODS FOR OBJECTIVE AND SUBJECTIVE ASSESSMENT OF QUALITY

MODULATED NOISE REFERENCE UNIT (MNRU)

ITU-T Recommendation P.810

(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 P.810 was revised by ITU-T Study Group 12 (1993-1996) and was approved under the WTSC Resolution No. 1 procedure on the 6th of February 1996.

NOTE

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

CONTENTS

SUMMARY

This Recommendation describes the Modulated Noise Reference Unit (MNRU), a standalone unit for introducing controlled degradations to speech signals. As such, the MNRU has been used extensively in subjective performance evaluations of digital processes, both in conventional telephone bandwidth and in wideband (e.g. 70-7000 Hz) applications.

Historically, the MNRU has been implemented in analogue hardware. The revisions encompassed in this version of the Recommendation are the inclusion of descriptions of digital implementations of the MNRU. These descriptions are suitable for implementation in software or on digital hardware. One further revision is to note the need for a high-pass filter (removal of any DC component of the input speech material) for all implementations. Existing analogue hardware implementations of the MNRU will continue to meet the specifications in this Recommendation, provided such filtering is applied externally.

INTRODUCTION

The MNRU was originally devised to produce distortion subjectively similar to that produced by logarithmically companded PCM systems. This approach was based on the views:

1) that network planning would require extensive subjective tests to enable evaluation of PCM system performance over a range of compandor characteristics, at various signal levels and in combination with various other transmission impairments (e.g. loss, idle circuit noise, etc.) at various levels; and

2) that it would be as reliable and easier to define a reference distortion system, itself providing distortion perceptually similar to that of PCM systems, in terms of which the performance of PCM systems could be expressed. This requires extensive subjective evaluation of the reference system when inserted in one or more simulated telephone connections, but leads to the possibility of simplified subjective evaluation of new digital processing techniques.

KEYWORDS

Analogue MNRU, controlled degradation, digital MNRU, Modulated Noise Reference Unit (MNRU), subjective performance evaluation.

Recommendation P.810[1)]

MODULATED NOISE REFERENCE UNIT (MNRU)

(Malaga-Torremolinos, 1984; amended Melbourne, 1988;
Helsinki, 1993; revised in 1996)

1 Scope

The Modulated Noise Reference Unit (MNRU) is a standalone unit that is intended to introduce controlled degradations to speech signals. Various organizations (Administrations, scientific/industrial organizations), as well as the ITU-T itself, have made extensive use of the MNRU concept for evaluating the subjective performance of digital processes (in arriving at Recommendations G.722 and the 32 kbit/s ADPCM algorithm of Recommendation G.726, for example). A modified version for use in evaluating codecs of wider bandwidth (70-7000 Hz) is now common practice. However, the actual devices used, while based on common principles, may have differed in detail, and hence the subjective results obtained may also have differed. (Differences in subjective testing methodology are also relevant.) The purpose of this Recommendation is to define the narrow-band and wideband versions of the MNRU as completely and in as much detail as possible in order to minimize the effects of the device, and of its objective calibration procedures, on the results of subjective tests.

The need for a device, implemented in both hardware and software, that introduces controlled degradations to speech signals is exemplified by a number of factors:

a) the use of digital processes (64 kbit/s PCM A-law or μ-law, A/D/A encoder pairs, A-/μ-law or μ-/A-law converters, digital pads based on 8-bit PCM words, 32 kbit/s ADPCM, etc.) in the international telephone network is now widespread;

b) new digital processes are being standardized, e.g. 64 kbit/s 7 kHz wideband ADPCM;

c) there is a need for standard tools to measure the quantization distortion performance of digital processes [for example, 32 kbit/s ADPCM (Recommendation G.726) and 64 kbit/s 7 kHz wideband codec (Recommendation G.722)], so that the tools can be used for estimating the subjective transmission performance of international connections containing digital processes;

d) an objective speech quality assessment method has not yet been established;

e) at the present time, subjective tests incorporating reference system conditions represent the only suitable method for measuring the speech transmission performance of digital processes;

f) expressing results in terms of a common reference system may facilitate comparison of subjective test results obtained at different laboratories.

It is recommended:

1) that a narrow-band MNRU be used as the reference system in terms of which subjective performance of telephone bandwidth digital processes are expressed;

2) that a wideband MNRU be used as the reference system in terms of which subjective performance of wideband digital processes are expressed.

NOTES

1 The MNRU was originally constructed using analogue circuitry. Currently, however, it tends to be either constructed using digital circuitry or simulated on computers. Further information on the effects of the MNRU parameters is given in the bibliographic sources listed at the end of this Recommendation.

2 The listening-only method presently proposed when using the MNRU in subjective tests is described in Recommendation P.830. See 2.2/P.80 for precautions concerning the use of listening-only tests.

[1)] Previously P.81.

3 Objective measurement methods that suitably reflect subjective quantization distortion performance of various types of digital processes do not exist at present. (For example, the objective techniques of Recommendation G.712, based on sine wave and band-limited noise measurements, are designed for PCM and do not measure appropriately the distortion induced by other systems such as ADPCM.) The artificial voice described in Recommendation P.50 may be relevant. Even if an objective method is developed, subjective tests will be required to establish correlation of subjective results/objective results for particular digital process types.

2 References

The following Recommendations contain provisions that, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated are 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 listed below. A list of the currently valid ITU-T Recommendations is regularly published.

– ITU-T Recommendation G.191 (1993), *Software tools for speech and audio coding standardization*.

– ITU-T Recommendation P.50 (1993), *Artificial voices.*

– ITU-T Recommendation P.56 (1993), *Objective measurement of active speech level*.

– ITU-T Recommendation P.80 (1993), *Methods for subjective determination of transmission quality*.

– ITU–T Recommendation P.830 (1996), *Subjective performance assessment of telephone-band and wideband digital codecs*.

3 Definitions

For the purposes of this Recommendation, the following definitions apply:

3.1 **dBov**: dB relative to the overload of a digital system.

3.2 **signal path**: The path through the Modulated Noise Reference Unit that includes only the input signal.

3.3 **noise path**: The path through the Modulated Noise Reference Unit that generates the modulated noise.

3.4 **combined path**: The combined signal and noise paths through the Modulated Noise Reference Unit.

3.5 **Q**: The ratio, in dB, of speech power to modulated noise power.

3.6 $\mathbf{Q_N}$: Q for a narrow-band Modulated Noise Reference Unit.

3.7 $\mathbf{Q_W}$: Q for a wideband Modulated Noise Reference Unit.

4 Abbreviations

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

ADPCM	Adaptive Differential Pulse Code Modulation
MNRU	Modulated Noise Reference Unit
PCM	Pulse Code Modulation
RMS	Root Mean Square
SNR	Signal-to-Noise Ratio

5 Conventions

Implementations of the MNRU in analogue hardware and on digital hardware or in software are described in this Recommendation. Throughout this Recommendation, references will be made to analogue hardware and to digital implementations. Wherever the digital implementation is mentioned, it should be understood that the description is applicable either to software or to digital hardware.

6 General description

Simplified arrangements of the MNRU are shown in Figure 1 a) for the analogue narrow-band version, Figure 1 b) for the analogue wideband version, Figure 2 a) for the digital narrow-band version, and Figure 2 b) for the digital wideband version. Speech signals entering from the left are split between two paths, a signal path and a noise path. The signal path provides an undistorted (except for bandpass filtering) speech signal at the output. In the noise path, the speech signal instantaneously controls a multiplier with an applied gaussian noise "carrier" that has a uniform spectrum between the cutoff frequencies shown for the noise source. The output of the multiplier, consisting of the noise modulated by the speech signal, is then added to the speech signal to produce the distorted signal.

The attenuators and switches in the signal and noise paths allow independent adjustment of the speech and noise signal levels at the output. Typically, the system is so calibrated that the setting of the attenuator (in dB) in the noise path represents the ratio of instantaneous speech power to noise power, when both are measured at the output of the bandpass filter (terminal OT). Specifically, when both are set to 0 dB, the noise level measured at terminal OT, with separate resistive terminations at the terminals T1 and T2 (unlinked), should be the same as the speech level measured at terminal OT, with separate resistive terminations at the terminals T5 and T6 (unlinked): as a check on this it should be established that the ratio of speech-plus-modulated-noise power (measured at terminal OT) to the power of the input speech (measured at terminal IT) is 3 dB (see 7.3.1).

For this Recommendation, the decibel representation of the ratio is called Q_N for the narrow-band version and Q_W for the wideband version.

Digital implementations of the MNRU will be facilitated by noting that, when the signal path has unity gain, output speech-plus-modulated-noise generation can be expressed by:

$$y(i) = x(i)\left[1 + 10^{-Q/20}\, N(i)\right] \qquad (6\text{-}1)$$

where $x(i)$ is the input speech, $N(i)$ is the random noise, Q is the ratio of speech power to modulated noise power (determined by the noise path gain), and $y(i)$ is the output speech-plus-modulated-noise.

7 Performance specifications

7.1 General

The specifications in this subclause apply both to analogue hardware implementations and to digital implementations. Analogue MNRU hardware and digital implementations of the MNRU are described in parallel. Existing analogue hardware meeting the specifications in this Recommendation and digital implementations of the MNRU were tested extensively by Study Group 12 during the 1989-1992 study period. These implementations were found to be equivalent in their subjective effects. As the MNRU is intended only as a reference degradation for subjective evaluation of digital processes, the experts in Study Group 12 have deemed subjective equivalence of MNRU implementations to be of paramount importance.

NOTE – The implementations of the wideband MNRU described here were not tested. Comments on the wideband MNRU found in clause 1, above, are noted here.

For practical implementations, the actual signal levels and noise levels may be increased or decreased to meet special needs. In such cases, the level requirements detailed below will have to be modified accordingly. In particular, digital implementations will typically require speech material with a RMS level of –26.15 dBov (i.e. dB relative to the overload point of a digital system). However, if the language used in subjective testing has a peak-to-average level greater than 23 dB, the RMS level of input speech should be reduced appropriately.

It can be shown that a DC component in the input signal will generate an additive noise component on the output signal. This additive component of noise is not accounted for when one specifies a Q value. Thus, rather than having only a multiplicative noise at the output with SNR of Q, the output will have another additive noise component, this component increasing the overall noise in the output, and decreasing the total signal to (additive-plus-multiplicative) noise ratio. Thus, MNRU implementations should include (or provide external) high-pass filtering at the input.

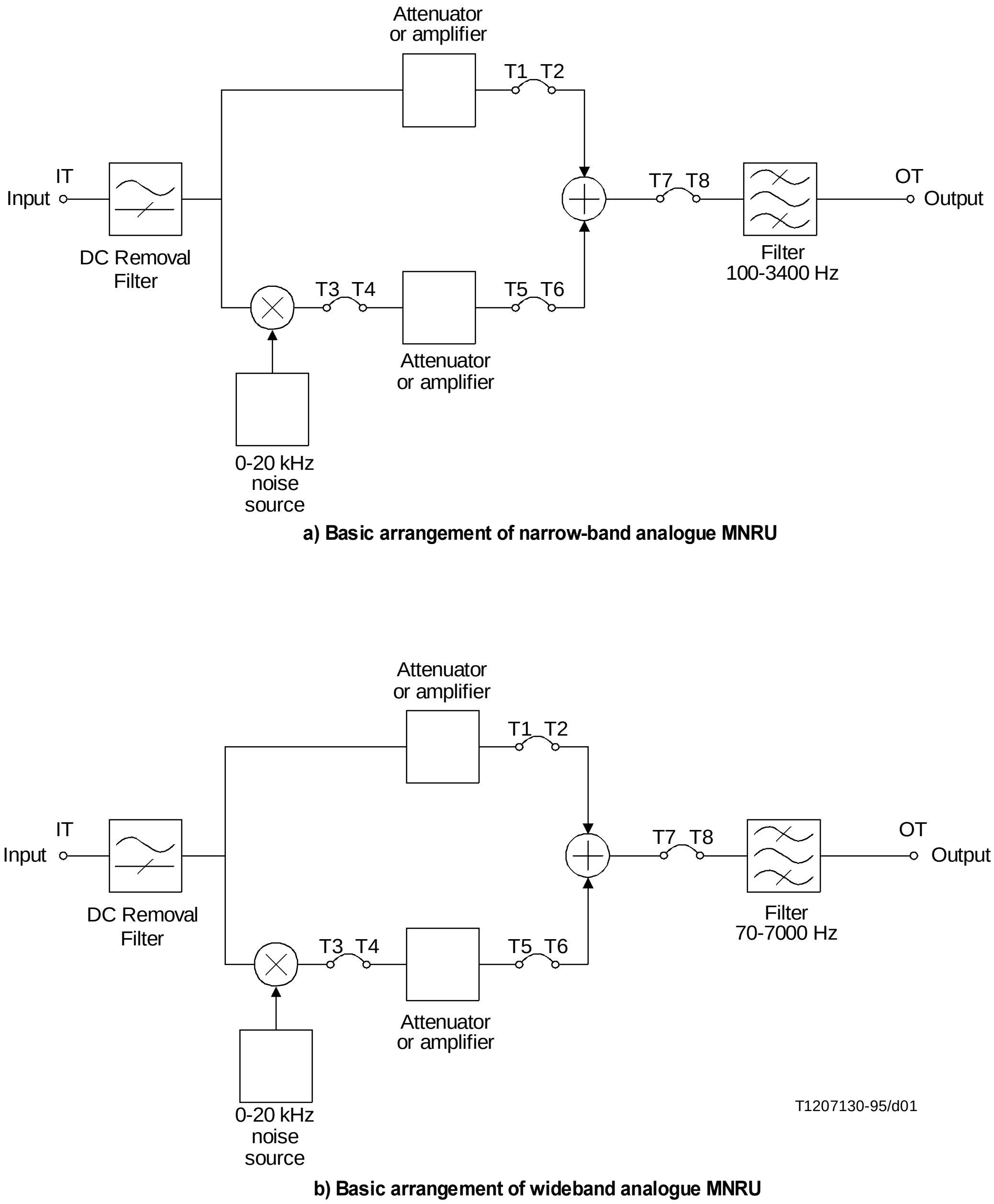
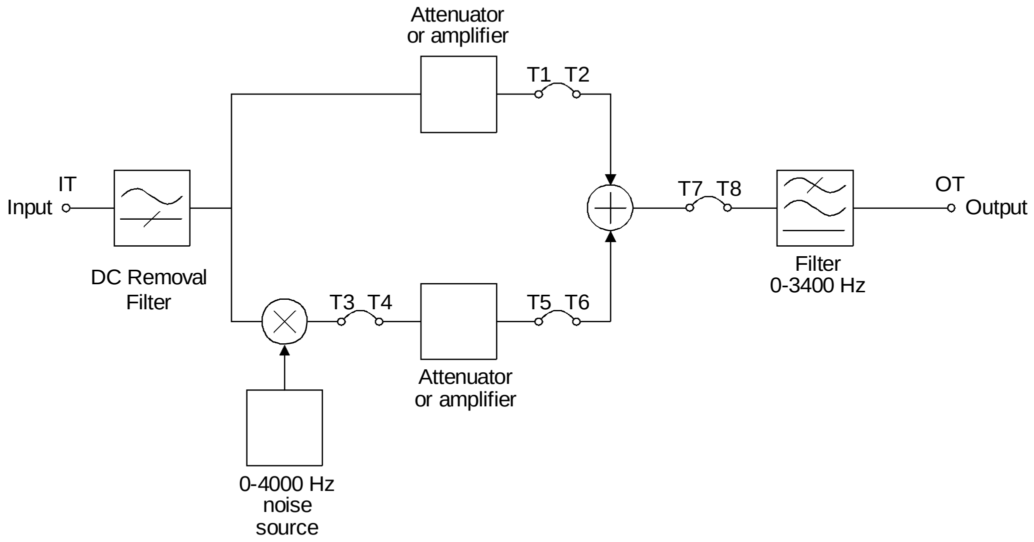


FIGURE 1/P.810

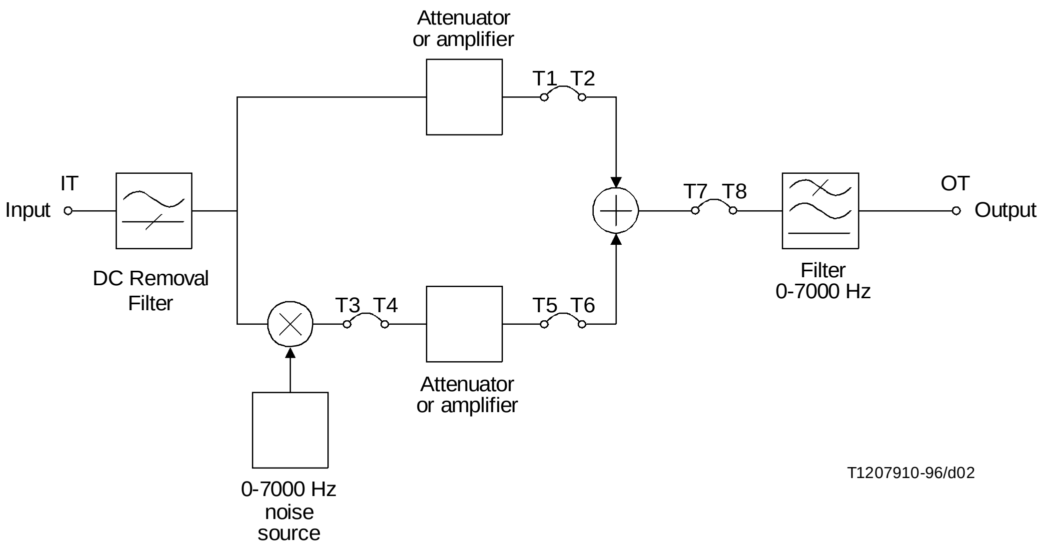
Basic arrangement of analogue MNRU

7.2 Signal path

The requirements under this heading refer to the MNRU with infinite attenuation in the noise paths of Figures 1 and 2. For analogue hardware, separate resistive terminations at the terminals T5 and T6 (unlinked) will achieve this. Digital implementations may provide an option to disable the noise path, allowing a filtered version of the original signal to be produced as output.



a) Basic arrangement of narrow-band digital MNRU

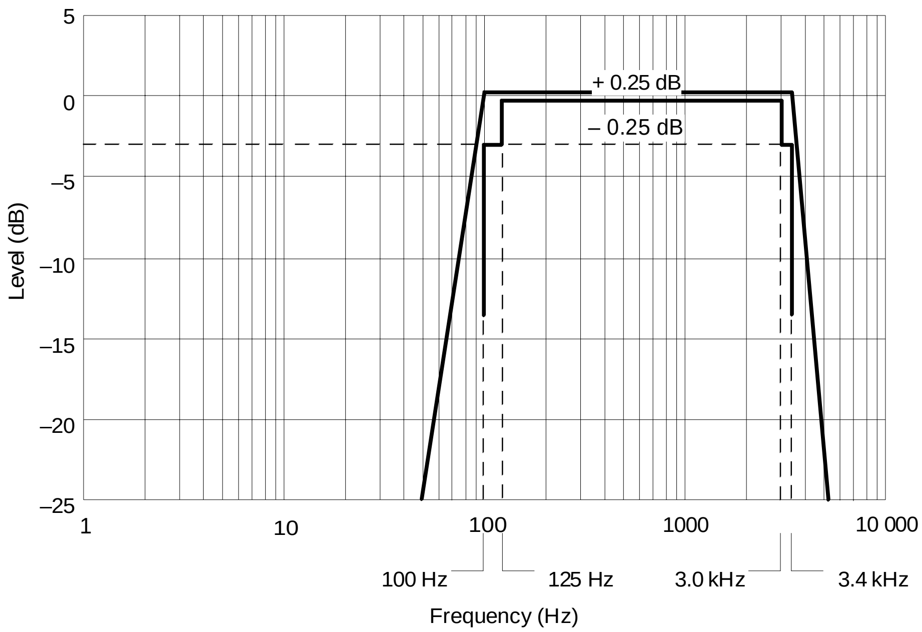


b) Basic arrangement of wideband digital MNRU

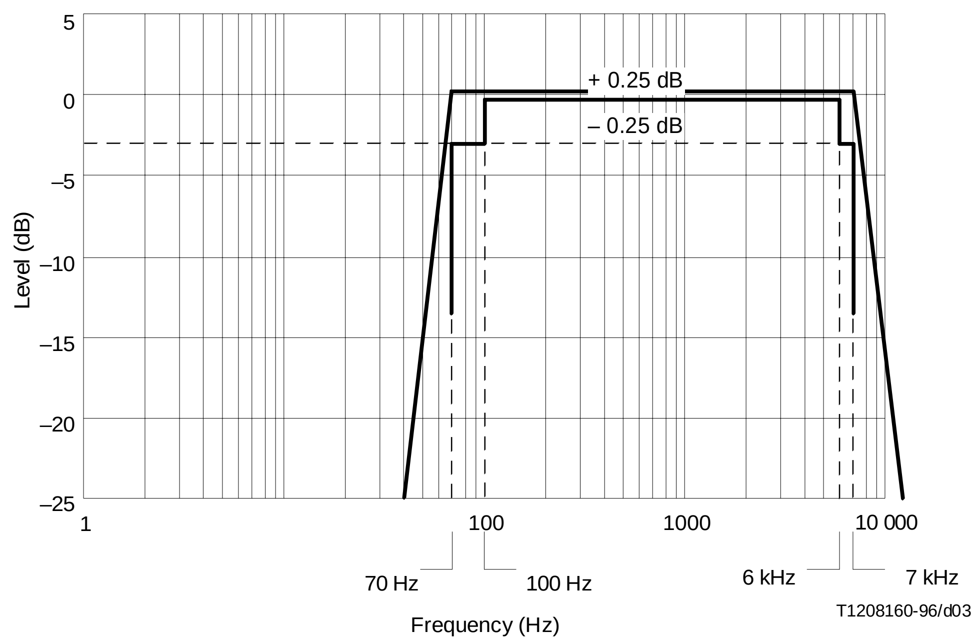
FIGURE 2/P.810

Basic arrangement of digital MNRU

The frequency response of the signal path (i.e. between terminals IT and OT) of Figures 1 a), 1 b), 2 a) and 2 b) should be within the limits of Figure 3 a) for the circuit of Figure 1 a), within the limits of Figure 3 b) for the circuit of Figure 1 b), within the limits of Figure 4 a) for the circuit of Figure 2 a), and within the limits of Figure 4 b) for the circuit of Figure 2 b). The frequency responses of digital implementations have the same low-pass characteristic as the corresponding analogue hardware implementation.



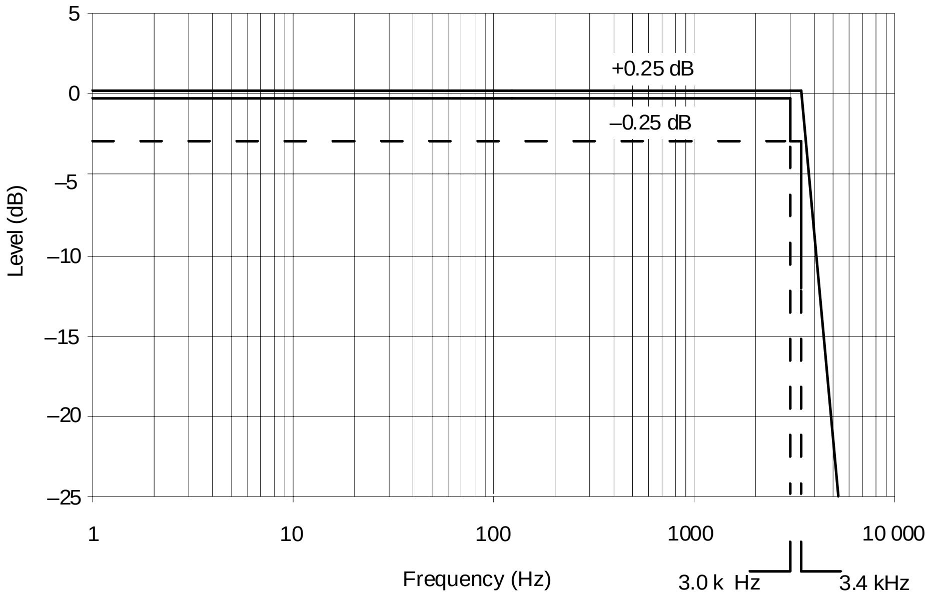
a) Requirements for output filter of the narrow-band analogue MNRU



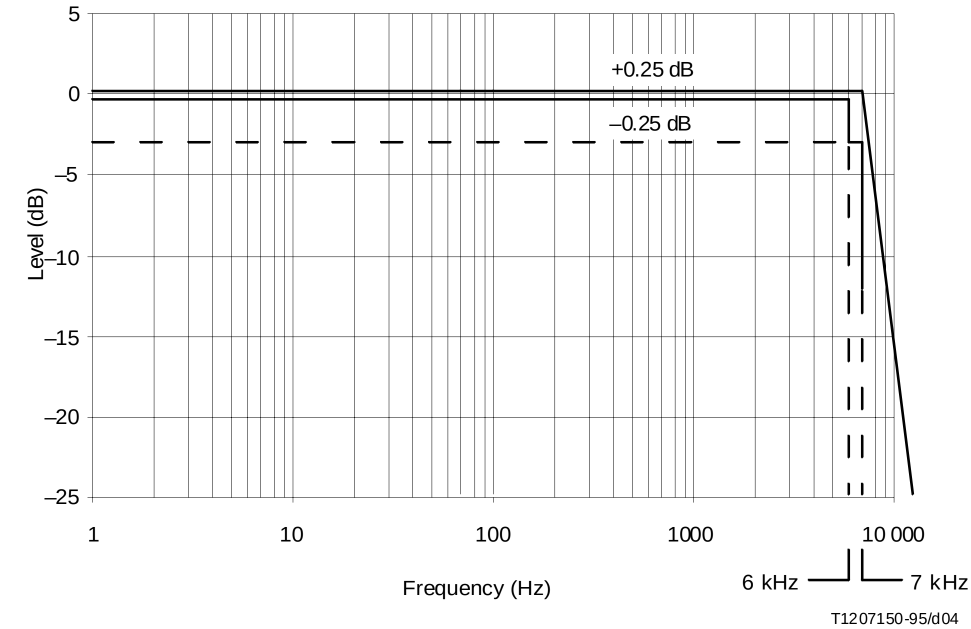
b) Requirements for output filter of the wideband analogue MNRU

FIGURE 3/P.810

Requirements for output filter of analoque MNRU



a) Requirements for output filter of the narrow-band digital MNRU



b) Requirements for output filter of the wideband digital MNRU

FIGURE 4/P.810

Requirements for output filter of digital MNRU

For analogue hardware, the loss between terminals IT and OT for a 0 dBm, 1 kHz input sine wave should be 0 dB. Over the input level range +10 dBm to –50 dBm, the loss should be 0 dB ± 0.1 dB. Digital implementations should guarantee unity gain in the signal path.

Any harmonic component should be at least 50 dB below the fundamental at the system output (terminal OT in Figures 1 a) and 1 b) for any fundamental frequency between 125 Hz and 3000 Hz in a narrow-band system and 100 Hz and 6000 Hz in a wideband system. Digital implementations, the output of which is delivered over high quality 16-bit digital-to-analogue converters, will readily meet these criteria.

The idle noise generated in the signal path must be less than –60 dBm, measured at terminal OT, in order to conform with 7.4.

It is recommended that the level of speech signals applied to the terminals IT of analogue hardware (Figure 1) should be less than –10 dBm (mean power while active, i.e. mean active level according to Recommendation P.56) in order to avoid amplifier peak-clippings of the signal, and be greater than –30 dBm to ensure sufficient speech signal-to-background-noise ratio. Digital speech materials used as input to digital implementations of the MNRU should be quantized with 16-bit linear quantization and have mean active levels that are, at most, –26.15 dB relative to the overload point of a 16-bit digital-to-analogue converter. Otherwise, speech may be clipped. As noted in 7.1, languages having peak-to-average ratios greater than 23 dB should adjust levels appropriately. The software implementation of Recommendation P.56, as found in Recommendation G.191, should be used to determine the level for the purpose of meeting this requirement.

7.3 Noise path

The requirements under this heading refer to the MNRU with infinite attenuation in the signal paths of Figures 1 and 2. For analogue hardware, separate resistive terminations at the terminals T1 and T2 (unlinked) will achieve this. Digital implementations may provide an option to provide only the modulated noise as output.

7.3.1 Linearity as a function of input level

With a Q_N setting of 0 dB in the circuit of Figure 1 a), or a Q_W setting of 0 dB in the circuit of Figure 1 b), as the case may be, the level of the modulated noise at the system output (terminal OT) should be numerically equal to the sine wave level at the input terminal (terminal IT). A correspondence within ± 0.5 dB should be obtained for sine wave input levels from +5 dBm to –45 dBm, and for input frequencies from 125 Hz to 3000 Hz in a narrow-band system and 100 Hz to 6000 Hz in a wideband system. Digital implementations calibrated for a given level (e.g. –26.15 dBov) will readily meet these criteria.

7.3.2 Noise spectrum

For a narrow-band analogue system, when Q_N is set to 0 dB, input sine waves applied to terminal IT in Figure 1 a) with levels from +5 to –45 dBm and frequencies from 125 Hz to 3000 Hz should result in a flat noise system spectrum density at the output of the multiplication device [terminal T5 of Figure 1 a)] within ± 1 dB over the frequency range 75 Hz to 5000 Hz. The spectrum density should be measured with a bandwidth resolution of maximum 50 Hz.

For a wideband analogue system, when Q_W is set to 0 dB, input sine waves applied to terminal IT in Figure 1 b) with levels from +5 to –45 dBm and frequencies from 100 Hz to 6000 Hz should result in a flat noise system spectrum density at the output of the multiplication device [terminal T5 of Figure 1 b)] within ± 1 dB over the frequency range 75 Hz to 10 000 Hz. The spectrum density should be measured with a bandwidth resolution of maximum 50 Hz.

Digital implementations (both narrow-band and wideband) calibrated for a given level (e.g. –26.15 dBov) and where the noise source meets the requirements given in 7.3.3 will readily meet these criteria.

7.3.3 Amplitude distribution

The amplitude distribution of the noise at the system output should be approximately gaussian.

NOTES

1 For analogue hardware, a noise source consisting of a gaussian noise generator followed by a peak clipper with a flat spectrum from near zero to 20 kHz will produce a satisfactory output noise at terminal OT.

2 To ensure sufficient linearity in digital implementations, a noise source consisting of a gaussian random noise generator (followed by a peak clipper) with a flat spectrum from 50 Hz to the cutoff frequency of the low pass portion of the bandpass filter in Figures 2 a) and 2 b) will produce a satisfactory output noise at terminal OT when the clipping level is at least 12 dB above the RMS noise level.

3 For digital implementations, noise that has an amplitude distribution that is approximately gaussian may be generated by any of several methods. If the noise is generated from uniformly distributed random numbers by appealing to the Central Limit Theorem, then the number of such random numbers should be greater than 10.

7.3.3.1 Noise period

In digital implementations, the random noise should be generated using a gaussian random number generator having a period longer than 2^{20} samples.

7.3.4 Noise attenuators

The loss of the noise attenuator(s), i.e. between terminals T4 and T5 in Figures 1 a) and 1 b), should be within ± 0.1 dB of the nominal setting. The attenuator(s) should at least allow Q_N and Q_W settings in the range –5 dB to 45 dB, i.e. a 50 dB range. Digital implementations will readily meet these criteria.

7.4 Combined path

The requirements under this heading refer to the MNRU with both speech and noise paths simultaneously in operation.

With Q_N or Q_W (as the case may be) set to zero, and the input terminated by an equivalent resistance, the idle noise generated in the combined path should be less than –60 dBm when measured at the system output (terminal OT).

Bibliography

LAW (H.B.), SEYMOUR (R.A.): A reference distortion system using modulated noise, *The Institute of Electrical Engineers,* pp. 484-485, November 1962.

CCITT Contribution COM XII-No. 63, *Some considerations on specifications for modulated noise reference unit*, NTT, study period 1981-1984, Japan.

CCITT Contribution COM XII-No. R4, pp. 71-79, study period 1981-1984.

CCITT Contribution COM XII-No. 119: *Description and method of use of the modulated noise reference unit (MNRU/MALT)*, study period 1981-1984, France.

ITU-T Contribution COM 12-8: *Report on comparative experiments for MNRU algorithms*, NTT, study period 1993-1996.