Joint TSG-S4#10 - SMG11#15 Meeting
February 28-March 3, 2000, Helsinki, Finland

Tdoc S4/SMG11 (00)0147

Title: **Proposed Addition to TS 26.131 and TS 26.132**
Source: **HEAD acoustics, T-NOVA Berkom**
Agenda Item: **12.5**

1. Introduction

TS 26.131 and TS 26.132 cover all types of 3G terminals including hands-free terminals. It can be expected, that mostly all terminals include non linear and/or time variant signal processing, e.g. voice switching, echo cancellation, background noise reduction ... The purpose of this contribution is to provide a minimum set of additional measurements and performance requirements for TS 26.131 and TS 26.132 based on the more fundamental parameters and requirements found in ITU-T Recommendation P.340, P.501 and P.502.

2. Proposed addendum to TS 26.131

5.10 Extended Parameters

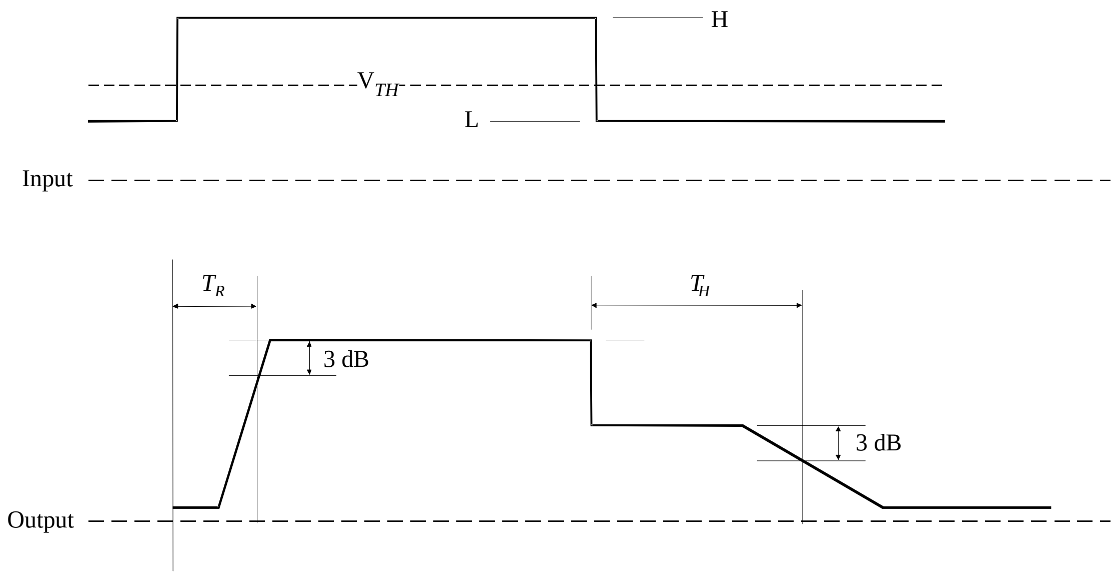
5.10.1 Switching parameters

Definitions

A loss in either the sending or receiving direction may be inserted in various ways. Switching from one direction to the other occurs when a signal above a given threshold is applied from the opposite direction, or when the control circuit, taking into account the relative levels and the nature of the signals in both directions, provides the switching.
The fundamental voice-switching parameters of the attenuation control according to ITU-T Recommendation P.340 are defined as follows:

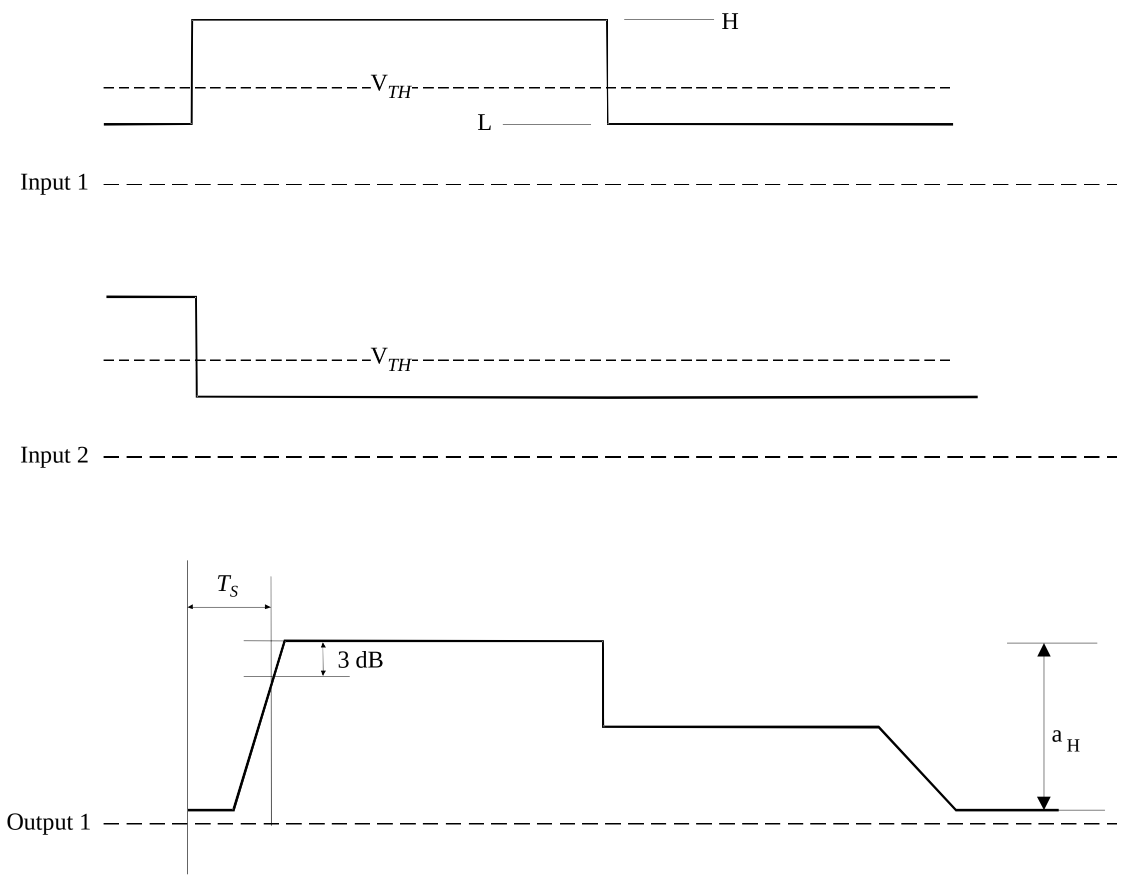
- Threshold level V_{TH}
  Minimum necessary signal level for removing insertion loss.
- Build-up time T_R (switch-on)
  Time from the input signal going above the threshold level until the time at which the output level reaches 3 dB below complete removal of the insertion loss.
- Hang-over time T_H
  Time from the input signal going below the threshold level until 3 dB of the switched loss is inserted in the output signal.
- Switching time T_S (switch-over)
  Time from one transmission direction to the other. T_S is measured from the removal of the signal in the first direction until the level in the second direction reaches 3 dB below its final value.
- Attenuation range a_H
  Attenuation range is determined by the difference in sensitivity response which results when one speech path is activated and when the duplex branch is activated.

Fig.1: Switching parameters (single talk)



R Build-up Time
T_H Hang-over Time
V_{TH} Threshold Level

Fig.2: Switching Time (double talk)



T_S Switching Time

a_H Attenuation Range

Note: These parameters can be determined in both directions. Therefore the letters $_S$ and $_R$ used as additional index indicate that the parameter is measured in sending or receiving direction. The index $_{DT}$ indicates, that the parameter is measured under double talk conditions.

Requirements

5.10.1.1 Threshold levels

Threshold levels should be chosen so that switching is not interrupted by random (environmental) noise sources at either end of the call. In addition, ambient room/network noise effects on threshold should not impair performance. Ambient noise levels can be used to improve threshold performance, as talkers tend to speak louder in a noisy environment than in a quiet one. In any case it should be ensured, that the talker signal is transmitted without any artifacts. No exact requirements are given but it should be insured that the ambient noise characteristic and level, the design of level detection and the threshold level in the terminal consider typical environmental conditions in daily use.

5.10.1.2 Build-up time T_R

In sending direction the build-up time $T_{R,S}$ shall be less than 15 ms, preferably below 10 ms for signal levels in the range of –14,7 dBPa to -4,7 dBPa.

In receiving direction the build-up time $T_{R,R}$ shall be less than 15 ms, preferably below 10 ms for signal levels in the range of –26 dBm0 to -16 dBm0.

5.10.1.3 Hang-over time T_H

Hang-over time should be long enough to cover average pauses in speech so that intermittent unwanted switching does not occur before the initial talker is finished, but short enough to allow reasonable break-in from the second talker.

The hang-over time in sending direction $T_{H,S}$ is measured for input signals of – 4,7 dBPa (at the MRP).
The hang-over time in receiving direction $T_{H,R}$ is measured for input signals of –16 dBm0.

For systems introducing an attenuation range of less than 12 dB, the hang-over time shall be more than 50 ms, preferably more than 100 ms.
For systems introducing an attenuation range of more than 12 dB , the hang-over time shall be more than 250 ms.

5.10.1.4 Switching time T_S (switch-over)

Switching time from one active state to the other should be balanced to best simulate full duplex operation.
Switching time is also dependent on both build-up time and hang-over time.
The switching time T_S shall be between 50 ms and 150 ms.

5.10.2 Duplex behaviour

In duplex conditions the system performance is subjectively determined by mostly two paramerters: talker echo loudness rating and attenuation range. In order to achieve a MOS ≥ 4.0 the $TELR_{DT}$ in double talk conditions should be ≥ 37 dB, the attenuation range in sending ≤ 3dB and in receiving ≤ 3dB as well.
For terminals which are not intended to provide double talk capabilities the total echo loss may be achieved by a correspondingly high echo attenuation based on level switching.

For discussion:

We may select the appropriate limits for 3G for attenuation range and TELR from this table:

Assuming a nominal SLR + RLR 10 dB for the 3G terminal where the listener perceives his own echo, the TCL should be at minimum 27 dB in double talk condition for the far end terminal which produces the echo for its volume control set at maximum.

Assuming the listeners terminal where the user perceives his own echo being equipped with a volume control according to 3G specifications, set to maximum volume, the requirement should be >42 dB (15 dB higher RLR).

TABLE A.1/P.340

Values for parameters determining double talk performance ($TELR_{DT}$, a_{Hrdt}, a_{Hsdt}) as a function of correlated MOS scores derived from LOT (HFT-handset, judgement at the handset side)

MOS	***≥ 4.0***	***4.0 - ≥ 3.5***	***3.5 - ≥ 3.0***	***3.0 - ≥ 2.5***	***2.5 - ≥ 2.0***	***< 2.0***
$TELR_{DT}$ [dB]	*≥37*	*≥33*	*≥27*	*≥21*	*≥13*	*<13*
a_{HSdt} [dB]	*≤3*	*≤6*	*≤9*	*≤12*	*≤15*	*>15*
a_{HRdt} [dB]	*≤3*	*≤5*	*≤8*	*≤10*	*≤12*	*>12*

5.10.2.1 Terminal coupling loss in double talk situation

From the subjective tests it appears that the echo loss is also a parameter with a great influence on the quality perceived in double talk situation. Based on subjective tests described in Annex A of ITU-T Recommendation P. 340, the following requirements apply:

To be discussed

Behaviour 1: $TELR_{DT} \geq 37$ dB

Behaviour 2a: 37 dB > $TELR_{DT} \geq 33$ dB

Behaviour 2b: 33 dB > $TELR_{DT} \geq 27$ dB

Behaviour 2c: 27 dB > $TELR_{DT} \geq 21$ dB

Behaviour 3: $TELR_{DT}$ < 21 dB

=> TCL_{DT} > 27 dB or 42 dB ????

5.10.2.2 Attenuation range in double talk situation

to be discussed, see above
the limits shown below may be chosen:

TABLE 4/P.340

Sending direction		***Receiving direction***	
Behaviour 1:	$a_{H,S,DT} \leq 3$ dB	*Behaviour 1:*	$a_{H,R,DT} \leq 3$ dB

Behaviour 2a: 3 dB < $a_{H,S,DT} \leq 6$ *dB* *Behaviour 2b: 6 dB <* $a_{H,S,DT} \leq 9$ *dB* *Behaviour 2c: 9 dB <* $a_{H,S,DT} \leq 12$ *dB* *Behaviour 3:* $a_{H,S,DT} > 12$ *dB*	*Behaviour 2a: 3 dB <* $a_{H,R,DT} \leq 5$ *dB* *Behaviour 2b: 5 dB <* $a_{H,R,DT} \leq 8$ *dB* *Behaviour 2c: 8 dB <* $a_{H,R,DT} \leq 10$ *dB* *Behaviour 3:* $a_{H,R,DT} > 10$ *dB*

5.10.3 Background noise transmission

The sources of the transmitted noise may be:

- Acoustical sources
  - Ambient noise in the test room which should be avoided;
  - Noise intentionally produced in the test room (e.g. babble noise, car noise, …).
- Electrical sources
  - Noise generated by the components of the terminals or/and network;
  - Noise intentionally generated (e.g. comfort noise, …).

Background noise transmission refers to the "noise intentionally produced in the test room" picked up by the microphone(s) of the terminal.

Parameters associated to background noise transmission, impacting on the quality

Parameters	
Absolute level	The transmitted background noise level should be low, but the noise should not be completely suppressed. The background noise transmitted signal should not be interrupted from time to time. For hands-free terminals tests are currently under study but no limits are given yet, for handset terminals the requirements of D-value apply.
Level fluctuations	If the level of the transmitted signal is referred to the original test signal level, the level fluctuations should not be more than ±3 dB, compared to steady state conditions NOTE 1 - This does not apply if the residual background transmitted noise is masked (e.g. by the speech signal). NOTE 2 - No values for the initial adaptation of noise reduction algorithms or any initial noise reduction due to the insertion of loss can be given yet.
Additional parameters	Artifacts of noise reduction algorithms esp. Musical tones and level fluctuations need to be avoided. Tests are currently under study but no exact requirements are given yet.

3. Proposed addendum to TS 26.132

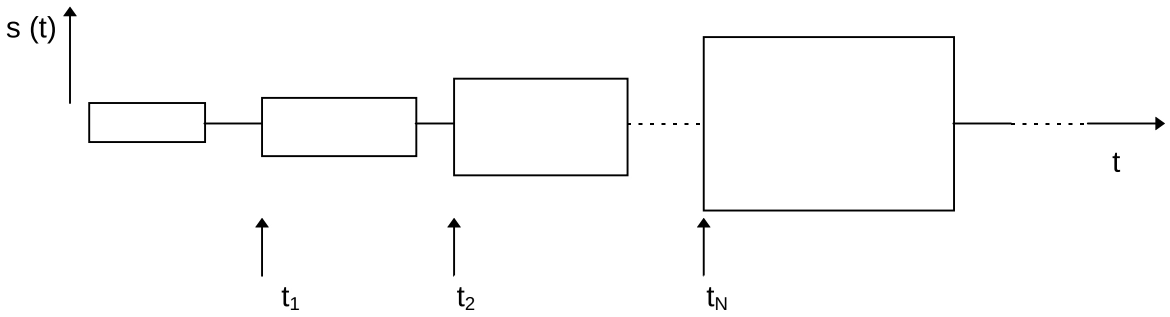
5.10 Extended Parameters

5.10.1 Switching parameters

5.10.1.1 Build Up Time

The signal structure as given through Fig. 2 represents signal parts with increasing levels. Periods of the CSS (as a simulation of speech) with increasing levels are suited for this signal.

Fig. 2: Structure of test signal to determine the build up time



Note:

The dotted line indicates the repetition or elongation of the test signal to achieve the suitable length for the measurement

Typical settings can be chosen as follows:

	Active duration / pause duration	level of the first period	Level difference between two periods
CSS for switching in **Sending direction**	248.62 ms / 451.38 ms	-24,7 dB_{Pa}	1 dB
CSS for switching in **Receiving direction**	248.62 ms / 451.38 ms	-36 dB_m	1 dB

It is assumed that the pause length of 451.38 ms is longer than the hangover time so that the test object mode is the idle mode after each signal burst independent if is was activated or not.
If the transmitted signals are measured and referred to the original measurement signal, the minimum activation level can be determined. The activation can be analyzed at the beginning of each signal burst (t_1, t_2, ..., t_N). The parameters which can be determined using this signal are

build up time in sending direction $T_{R,S}$
The measurement is conducted simply by a level versus time analysis derived from the measured output signal in sending direction. The analysis time constant to be chosen for this measurement is 1 ms. The build-up time then is determined by evaluating the level versus time graph.

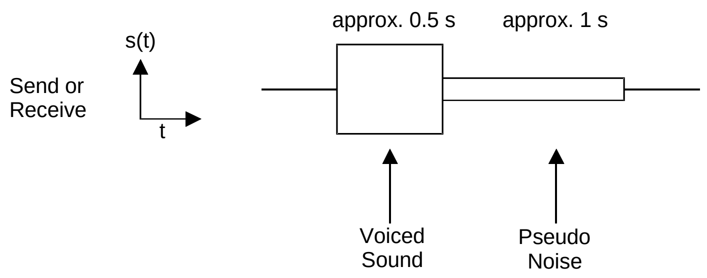
build up time in receiving direction $T_{R,R}$

The measurement is conducted simply a the level versus time analysis derived from the measured output signal in receiving direction. The analysis time constant to be chosen for this measurement is 1 ms. The build-up time then is determined by evaluating the level versus time graph.

5.10.1.2 Hang-over time

The test signal structure is given below.

Fig. 3: Structure of test signal to determine the hang-over time



The transition from activation to idle can be represented by feeding in an activation signal, a voiced sound of CSS in one direction, followed by a second signal in the same direction but of lower level, which does not activate the terminal, typically a random noise or pseudo noise signal signal (see Figure 4). The second part of the signal measured thus indicates the attenuation, from which the hang-over time (switch-off time) can be determined.
The duration of the voiced sound is 0.5 s in order to reach a final stable system condition. The second part of the signal has a duration of 1 s. The level must be selected low enough so as not to activate the equipment.

Hang-over time in sending direction $T_{H,S}$
The level of the activating signal is –4,7 dBPa, measured at the MRP. The level of the noise signal is -34.7 dBPa, measured at the MRP.

The measurement is conducted by a level versus time analysis derived from the measured output signal in sending direction. The analysis time constant to be chosen for this measurement is 1 ms. The hang-over time then is determined by evaluating the level versus time graph.

Hang-over time in receiving direction $T_{H,R}$
The level of the activating signal is –16 dBm0. The level of the noise signal is –50 dBm0.

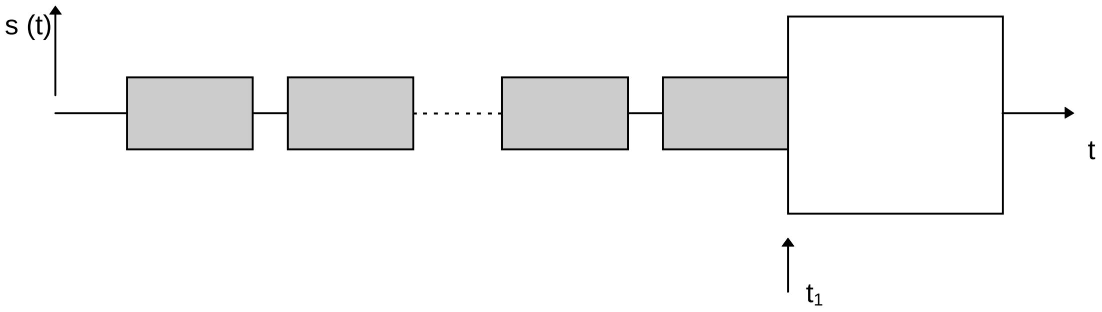
The measurement is conducted simply a the level versus time analysis derived from the measured output signal in receiving direction. The analysis time constant to be chosen for this measurement is 1 ms. The hang-over time then is determined by evaluating the level versus time graph.

5.10.1.3 Attenuation Range

Before the actual measurement of the switching times the attenuation range in sending and receiving direction has to be measured in order to determine the limits for the switching times T_S.

The test signal structure is as follows:

Fig. 4: Structure of test signal for attenuation range measurement



Note:

The dotted line indicates the repetition or elongation of the test signal to achieve the suitable length for the measurement

A periodical repetition of CSS bursts as a simulation of speech is used to activate one transmission path (gray color). At the end of one CSS burst, indicated by t_1 on the time scale (or at the end of the voiced part of the CSS), the measurement signal is applied in the opposite path (white color). This signal consists of a periodical repetition of a voiced sound.

Attenuation Range in sending direction, $a_{H,S}$

The activating signal is inserted in the receiving direction, immediately followed by the measurement signal in the sending direction. The level of the receive activating signal shall be –16 dBm0, the level of the measurement signal shall be –4,7 dBPa, measured at the MRP.
The level of the measured output signal in sending direction is represented versus time (time constant = 5 ms). The attenuation range is obtained from the difference between the maximum level at full activation and the minimum level obtaining immediately after switch-over.

Attenuation Range in receiving direction, $a_{H,R}$

The activating signal is inserted in the sending direction, immediately followed by the measurement signal in the receiving direction. The level of the send activating signal shall be –4,7 dBPa, measured at the MRP, the level of the measurement signal shall be –16 dBm0.
The level of the measured output signal in receiving direction is represented versus time (time constant 5 ms). The attenuation range is obtained from the difference between the maximum level at full activation and the minimum level obtaining immediately after switch-over.

5.10.1.4 Switching time

The measurement signals and levels are the same as for the determination of the attenuation range.

Switching time in sending direction, $T_{S,S}$

The activating signal is inserted in the receiving direction, immediately followed by the measurement signal in the sending direction. The level of the receive activating signal shall be –16 dBm0, the level of the measurement signal shall be –4,7 dBPa, measured at the MRP.
The level of the measured output signal in sending direction is represented versus time (time constant 1 ms). The switching time is obtained from the difference between time t_1 (see Fig. 4) and the time until the output signal in sending direction reaches 3 dB below its final value.

Switching time in receiving direction, $T_{S,R}$

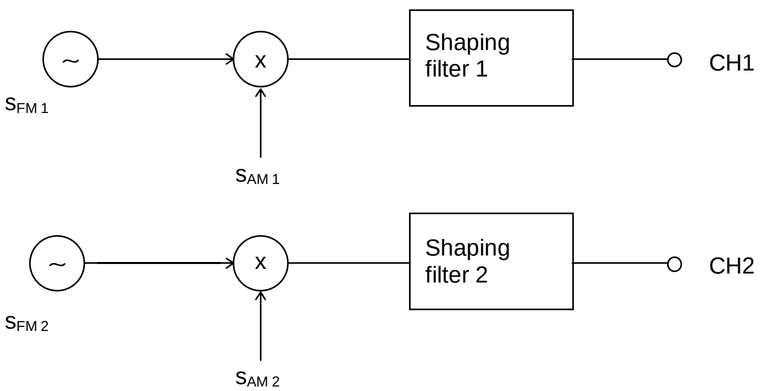
The activating signal is inserted in the sending direction, immediately followed by the measurement signal in the receiving direction. The level of the receive activating signal shall be –16 dBm0, the level of the measurement signal shall be –4,7 dBPa, measured at the MRP.
The level of the measured output signal in receiving direction is represented versus time (time constant = 1 ms). The switching time is obtained from the difference between time t_1 (see Fig. 4) and the time until the output signal in sending direction reaches 3 dB below its final value.

5.10.2 Duplex behaviour

5.10.2.1 Terminal coupling loss in double talk situation

The test signal is constructed as described below:
Orthogonal sequences are generated by a set of voice like modulated sinewaves, spectrally shaped. The general construction principle which in detail can be found in ITU-T Recommendation P.501 is shown below.

Fig. 5: Two channel test signal generation for double talk evaluations based on AM - FM signals



$$s_{FM1,2}(t) = \sum A_{FM1,2} * \cos(2\pi t n * F_{01,2}) \; ; \qquad n = 1,2,....$$
$$s_{AM1,2}(t) = A_{AM1,2} * \cos(2\pi t F_{AM1,2}) \; ;$$

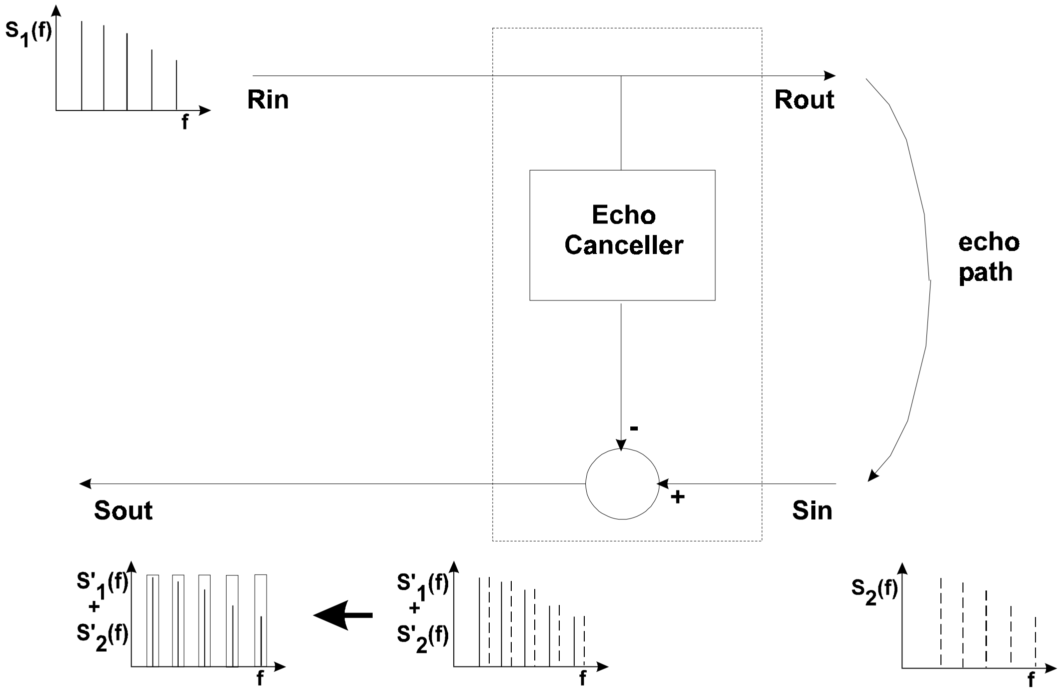
Typical settings are given in the following table:

Receiving direction			Sending direction		
f_m[Hz]	$f_{mod(fm)}$[Hz]	F_{am}[Hz]	f_m[Hz]	$f_{mod(fm)}$[Hz]	F_{am}[Hz]
250	± 5	3	270	± 5	3
500	± 10	3	540	± 10	3
750	± 15	3	810	± 15	3
1000	± 20	3	1080	± 20	3
1250	± 25	3	1350	± 25	3
1500	± 30	3	1620	± 30	3
1750	± 35	3	1890	± 35	3
2000	± 40	3	2160	± 35	3
2250	± 40	3	2400	± 35	3
2500	± 40	3	2900	± 35	3
2750	± 40	3	3150	± 35	3
3000	± 40	3	3400	± 35	3
3250	± 40	3	3650	± 35	3
3500	± 40	3	3900	± 35	3
3750	± 40	3			

Parameters of the shaping filter: LP, 5 dB/oct.

Fig. 6 shows how to determine the terminal coupling loss (TCL).

Fig. 6: Principle of double talk TCL measurement



The methodology is explained below:

In case echo cancellers are involved, the set is reset, and converged using at least a 10 s sequence. The test signal in receiving direction from Fig. 7 can be used or other test signals which guarantee the adaptation of the echo canceller (e.g. artificial voice according to ITU-T Recommendation P.50).The signal level is –16 dBm0. After activation the actual test signal as described above is inserted. The test signal level in receiving direction is between -16 dBm0 and 0 dBm0 , the maximum level possible for the system under test should be chosen in order to achieve a sufficient signal to noise ratio for the echo loss dtermination. The test signal level in sending direction is –4,7 dBPa.

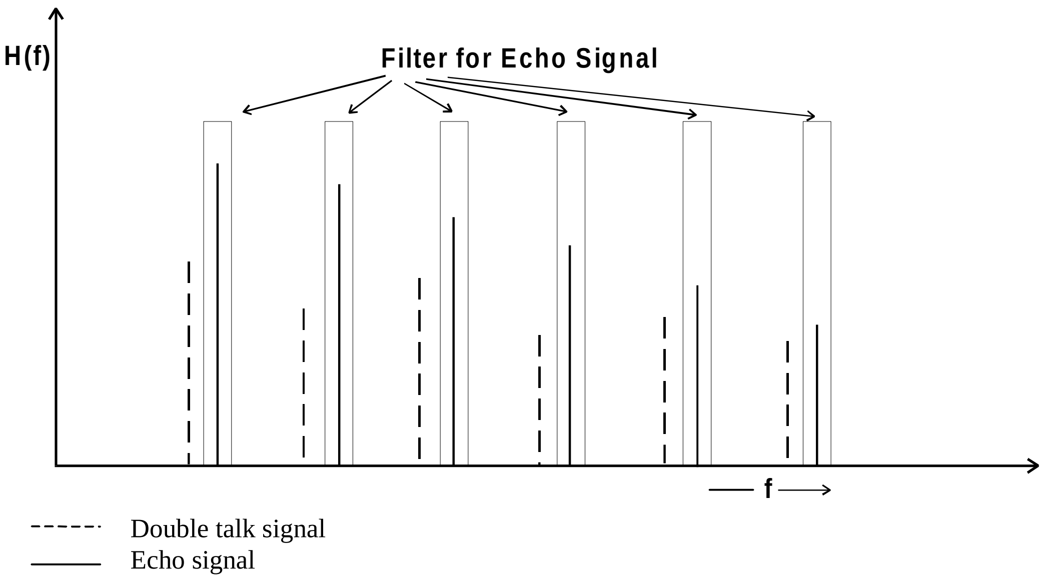
In order to extract the echo signal from the double talk either a specific filter setting or a specific post processing of the FFT analysis is required, since the spectrum of the signal as well as of the double talk signal is a kind of combfilter spectrum where a specific modulation is applied. The mid frequency f_{mid} of any frequency component, the according frequency modulation f_{mod} as well as the filter shapes or the windowing function of the Fourier transformation need to be taken into account. If the filter approach is used the bandwidth of each filter should be constructed that way that:

$f_u = f_{mid} - f_{mod\ (fm)}$
$f_o = f_{mid} + f_{mod\ (fm)}$

The stopband attenuation should be at least 10 dB higher than the minimum level to be measured within the passband. The same applies for analysis derived from Fourier transformations of the measured echo signal. Here the frequency “smearing“ effect of the windowing function needs to be taken into account. In order to have a sufficient separation between the echo signal and the double talk signal in the low frequency domain, a minimum

FFT lengths of 8 k (sampling rate 44.1 or 48 kHz) which amounts to a time window of about 170 ms should be chosen.

Figure 7: Extraction of the echo components of the double talk signal (schematic).



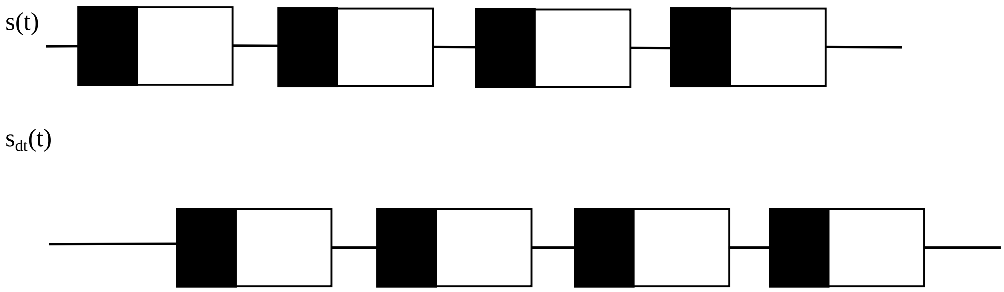
The timing of the measurement must be fine tuned knowing the echo path delay. This delay properly aligns the source and echo.

5.10.2.2 Attenuation range in double talk situation

The test signal used for the attenuation range test is shown in Fig. 8. This test signal consists of a series of uncorrelated composite source signals (ITU-T Recommendation P.501) which are fed in sending and receiving direction simultaneously. The signal levels are chosen as follows:

receiving direction: -16 dB_{m0}
sending direction: -4.7 dB_{Pa}

Fig. 8: Measurement sequence with detailed view on the overlap of sending and receiving direction signal, principle arrangement

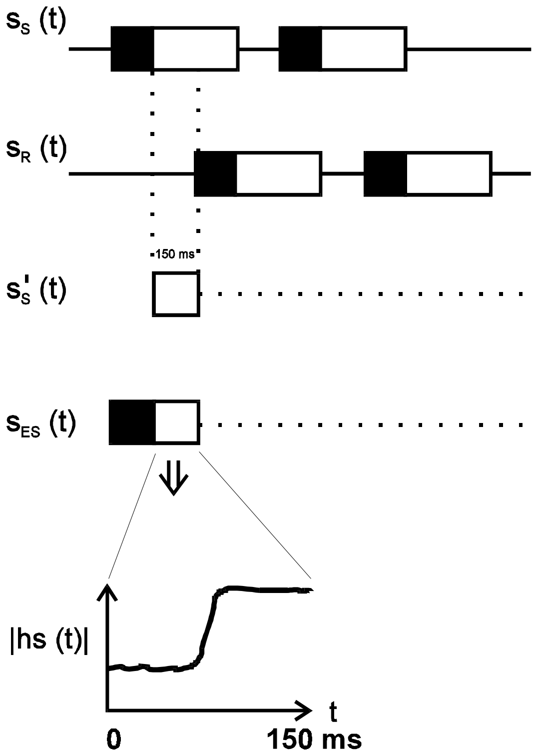


s(t) - signal in one direction
$s_{dt}(t)$ - double talk signal

From Fig. 8 it can be seen that the overlap of the sequences is only partially. Always the voiced sound (black) overlaps with the end of the pseudo random noise sequence (white) of the opposite channel. The sequence is constructed in that way, that during the pauses in receiving direction, the sending direction can be measured; during the pauses in sending direction, the receiving direction can be evaluated.
In the same way a sequence can be constructed which starts with high level exitation in receiving direction and low level exitation in sending direction, in case that different starting points of levels should be evaluated.

The attenuation range is evaluated determining the level versus time with an adequate short time constant, typically at minimum 10 times shorter than the switching time of the system under test to be evaluated. By monitoring the output signal during the periods, where only one signal is present, switching or level variations can be evaluated in great detail. The general procedure is given in Fig. 9. Although the output signal is referenced to the input signal in this example, this referencing is not required since only the difference between maximum and minimum level is used for measuring the evaluation range. Only in cases, where it is expected that during double talk the maximum level is never achieved (constant attenuation during double talk, the referencing is required. In this case the attenuation range is the difference between maximum (nominal) level when no double talk is present to the minimum level measured during double talk.

Fig. 9: Principle of signal extraction and determination of time constants, the example shows the switching time during double talk.



$s_S(t)$: Measured signal sending direction
$s_R(t)$: Measured signal receiving direction
$s'_S(t)$: Extracted signal sending direction
$s_{ES}(t)$: Exitation signal for sending direction
$|h_S(t)|$: Level vs. time

$$\left|h_s(t)\right| = \left|\int s_s{}^{'}(t)e^{t/10ms}\right| - \left|\int s_{ES}\ (t)e^{t/10ms}\right|$$

Note: The time constant of 10 ms is shown as one example, different time constant may be used depending on the analysis requirements

Attenuation Range during double talk in sending direction, $a_{H,Sdt}$

The double talk test signal activating signal is inserted in sending and the receiving direction. The level of the receive signal shall be –16 dBm0, the level of the measurement signal shall be –4,7 dBPa, measured at the MRP. The level of the measured output signal in sending direction during the pauses of the receiving double talk signal is represented versus time (time constant = 5 ms). The attenuation range is obtained from the difference between the maximum level at full activation and the minimum level obtaining immediately after switch-over.
Care must be taken to properly align the test signals and to adequately compensate any system delay.

Attenuation Range during double talk in receiving direction, $a_{H,Rdt}$

The double talk test signal activating signal is inserted in sending and the receiving direction. The level of the send signal shall be –4,7 dBPa, measured at the MRP, the level of the measurement signal shall be –16 dBm0. The level of the measured output signal in receiving direction during the pauses of the sending double talk signal is represented versus time (time constant = 5 ms). The attenuation range is obtained from the difference between the maximum level at full activation and the minimum level obtaining immediately after switch-over.
Care must be taken to properly align the test signals and to adequately compensate any system delay.

5.10.3 Background noise transmission

to be added