
**Acoustics — Methods for calculating
loudness —**

Part 3:
**Moore-Glasberg-Schlittenlacher
method**

STANDARDSISO.COM : Click to view the full PDF of ISO 532-3:2023



STANDARDSISO.COM : Click to view the full PDF of ISO 532-3:2023



COPYRIGHT PROTECTED DOCUMENT

© ISO 2023

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword.....	iv
Introduction.....	v
1 Scope.....	1
2 Normative references.....	1
3 Terms and definitions.....	1
4 General.....	4
5 Input signal.....	4
5.1 Single microphone.....	4
5.2 Two microphones in the ear canals or microphones in a head and torso simulator.....	4
5.3 Earphone presentation.....	4
6 Instrumentation.....	5
7 Description of the method.....	5
7.1 General.....	5
7.2 Transfer of sound through the outer and middle ear.....	6
7.2.1 General.....	6
7.2.2 Free-field transfer function.....	7
7.2.3 Diffuse-field transfer function.....	8
7.2.4 Signal recorded using microphones in the ear canals or using a Head and Torso Simulator.....	8
7.2.5 Earphone presentation.....	8
7.3 Calculation of the running short-term spectrum.....	8
7.4 Calculation of the running short-term excitation pattern.....	9
7.5 Transformation of excitation into specific loudness.....	10
7.5.1 General.....	10
7.5.2 Reference excitation at the reference threshold of hearing.....	10
7.5.3 Gain of the cochlear amplifier for inputs with low sound pressure levels.....	11
7.5.4 Calculation of specific loudness from excitation when $E_{\text{THRQ}}/E_0 \leq E/E_0$	11
7.5.5 Calculation of specific loudness from excitation when $E_{\text{THRQ}}/E_0 > E/E_0$	12
7.5.6 Calculation of specific loudness from excitation when $E/E_0 > 10^{10}$	12
7.6 Calculation of short-term specific loudness.....	13
7.7 Smoothing of short-term specific loudness and application of binaural inhibition.....	13
7.8 Calculation of short-term loudness.....	15
7.9 Calculation of long-term loudness.....	15
7.10 Relationship between loudness level and loudness.....	15
7.11 Calculation of the reference threshold of hearing.....	16
8 Uncertainty of calculated loudness sounds.....	17
9 Data reporting.....	17
Annex A (informative) Software for the calculation of loudness according to the method in this document.....	19
Annex B (informative) Test signals used for verification of this document.....	21
Annex C (informative) Test signals used for verification of equivalence with ISO 532-2.....	24
Bibliography.....	28

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO had/had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents. ISO shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 43, *Acoustics*.

A list of all parts in the ISO 532 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

Loudness and loudness level are two perceptual attributes of sound describing absolute and relative sensations of sound strength perceived by a listener under specific listening conditions. Due to inherent individual differences among people, both loudness and loudness level have the nature of statistical estimators characterized by their respective measures of central tendency and dispersion determined for a specific sample of the general population.

The object of this document is to specify a calculation procedure based on the physical properties of sound for estimating loudness and loudness level of sound as perceived by listeners with otologically normal hearing under specific listening conditions. This procedure seeks numbers that can be used in many scientific and technical applications to estimate the perceived loudness and loudness level of sound without conducting separate human observer studies for each application. Because loudness is a perceived quantity, the perception of which may vary among people, any calculated loudness value represents only an estimate of the average loudness as perceived by a group of individuals with otologically normal hearing.

This document describes a method for calculating the loudness of time-varying sounds from the input signal, which may differ for the two ears. This calculation method is based on Moore-Glasberg-Schlittenlacher loudness calculation algorithms^{[1] to [5]}. The method allows calculation of two quantities:

- a) The short-term loudness, which is the momentary loudness of a short segment of a sound, such as a word in a speech sound or a single note in a piece of music.
- b) The long-term loudness, which is the loudness of a longer segment of sound, such as a whole sentence or a musical phrase.

For most everyday sounds, both the short-term loudness and the long-term loudness vary over time. The loudness of sounds with durations up to 2 s or 3 s is well predicted from the maximum value of the long-term loudness reached during presentation of the sound^{[4][6] to [8]}. For long-duration stationary sounds, the long-term loudness based on the method described in this document is very close to the loudness determined using the method described in ISO 532-2^[9]. Deviations can occur for sounds with strong amplitude fluctuations, such as noises with narrow bandwidth; for such sounds the calculated loudness is more accurate for this document than for ISO 532-2.

The method of loudness calculation described in this standard can be applied to signals of any duration. However, it does not directly give an output corresponding to the overall loudness impression of a sound scene or soundscape over a period of minutes, hours, or days, which is called the “overall loudness” in this standard. The output of the method of loudness calculation described in this standard can be post-processed to estimate the overall loudness of a sound scene.

NOTE Post-processing is outside the scope of this document, but some possible methods have been described^{[10] to [13]}.

This document describes the calculation procedure leading to estimation of the loudness and loudness level of time-varying sounds and provides executable computer programs. The software provided with this document is entirely informative and provided for the convenience of the user. Use of the provided software is not required for conformity with the document.

NOTE Equipment or machinery noise emissions/immissions can also be judged by other quantities defined in various International Standards (see e.g. ISO 1996-1^[14], ISO 3740^[15], ISO 9612^[16], and ISO 11200^[17]).

STANDARDSISO.COM : Click to view the full PDF of ISO 532-3:2023

Acoustics — Methods for calculating loudness —

Part 3: Moore-Glasberg-Schlittenlacher method

1 Scope

This document specifies a method for estimating the loudness and loudness level of both stationary and time-varying sounds as perceived by otologically normal adult listeners under specific listening conditions. The sounds may be recorded using a single microphone, using a head and torso simulator, or, for sounds presented via earphones, the electrical signal delivered to the earphones may be used.

The method is based on the Moore-Glasberg-Schlittenlacher algorithm.

NOTE 1 Users who wish to study the details of the calculation method can review or implement the source code which is entirely informative and provided with the standard for the convenience of the user.

This method can be applied to any sounds, including tones, broadband noises, complex sounds with sharp line spectral components, musical sounds, speech, and impact sounds such as gunshots and sonic booms.

Calculation of a single value for the overall loudness over the entire period of a time-varying signal lasting more than 5 s is outside the scope of this document.

NOTE 2 It has been shown that, for steady tones, this method provides a good match to the contours of equal loudness level as defined in ISO 226:2003^[18] and the reference threshold of hearing as defined in ISO 389-7:2019^[19].

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

IEC 60318-7, *Electroacoustics – Simulators of human head and ear – Part 7: Head and torso simulator for the measurement of sound sources close to the ear*

IEC 61672-1, *Electroacoustics - Sound level meters - Part 1: Specifications*

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

- ISO Online browsing platform: available at <https://www.iso.org/obp>
- IEC Electropedia: available at <https://www.electropedia.org/>

3.1 sound pressure level

L_p
ten times the logarithm to the base 10 of the ratio of the square of the sound pressure, p , to the square of a reference value, p_0 , expressed in decibels

$$L_p = 10 \lg \frac{p^2}{p_0^2} \text{ dB}$$

where the reference value, p_0 , in air is 20 µPa

Note 1 to entry: Because of practical limitations of the measuring instruments, p^2 is always understood to denote the square of a frequency-weighted, frequency-band-limited or time-weighted sound pressure. If specific frequency and time weightings as specified in IEC 61672-1 and/or specific frequency bands are applied, this should be indicated by appropriate subscripts; e.g. $L_{p,AS}$ denotes the A-weighted sound pressure level with time weighting S (slow). Frequency weightings such as A-weighting should not be used when specifying sound pressure levels for the purpose of loudness calculation using the current procedure.

Note 2 to entry: This definition is technically in accordance with ISO 80000-8:2020, 8-22 [20].

3.2 filter

any device or mathematical operation which, when applied to a complex signal, passes energy of signal components of certain frequencies while substantially attenuating energy of signal components of all other frequencies

3.3 band-pass filter

filter (3.2) that passes signal energy within a certain frequency band and rejects most of the signal energy outside of this frequency band

3.4 sound spectrum

representation of the magnitudes (and sometimes of the phases) of the components of a complex sound as a function of frequency

3.5 auditory filter

filter (3.2) within the human cochlea describing the frequency resolution of the auditory system, whose characteristics are usually estimated from the results of masking experiments

3.6 ERB_n

equivalent rectangular bandwidth of the auditory filter for otologically normal persons
width of an idealised rectangular *band-pass filter* (3.3) that has the same peak transmission as the *auditory filter* (3.5) at the same centre frequency and that passes the same power for a white noise input (in Hz)

Note 1 to entry: The subscript n indicates that the value applies for listeners with otologically normal hearing.

Note 2 to entry: The unconventional use of a multiletter abbreviated term presented in italics and with a subscript is used here in the place of a symbol to maintain the use of an established notation and to avoid confusion.

3.7

 ERB_n -number scale

equivalent rectangular bandwidth number scale

transformation of the frequency scale constructed such that an increase in frequency equal to one ERB_n (Hz) (3.6) leads to an increase of one unit on the ERB_n -number scale

Note 1 to entry: The unit of the ERB_n -number scale is the Cam. For example, the value of ERB_n for a centre frequency of 1 000 Hz is approximately 132 Hz, so an increase in frequency from 934 Hz to 1 066 Hz corresponds to a step of one Cam. The equation relating ERB_n -number to frequency is given in 7.4.

3.8

loudness level

sound pressure level of a frontally incident, sinusoidal plane progressive wave, presented binaurally at a frequency of 1 000 Hz that is judged by otologically normal persons as being as loud as the given sound

Note 1 to entry: Loudness level is expressed in phons.

3.9

loudness

perceived magnitude of a sound, which depends on the acoustic properties of the sound and the specific listening conditions, as estimated by otologically normal listeners

Note 1 to entry: Loudness is expressed in sones.

Note 2 to entry: Loudness depends primarily upon the sound pressure although it also depends upon the frequency, waveform, bandwidth, and duration of the sound.

Note 3 to entry: One sone is the loudness of a sound whose loudness level is 40 phon.

Note 4 to entry: A sound that is twice as loud as another sound is characterized by doubling the number of sones.

3.10

short-term loudness

loudness of an individual brief segment of sound, such as a syllable in speech, a single musical note, or a short burst of a sound, typically lasting up to 500 ms

3.11

long-term loudness

loudness of a long sound, such as a whole sentence, a musical phrase, or a continuous noise, typically lasting up to 5 s

Note 1 to entry: The overall loudness of a sound or soundscape lasting longer than 5 s can be estimated by post-processing of the long-term loudness as a function of time. Such post-processing is outside the scope of this standard, but some possible methods are described in References [10] to [13].

3.12

excitation

E

output of an *auditory filter* (3.5) centred at a given frequency, specified in units that are linearly related to power

Note 1 to entry: An excitation of 1 unit is produced at the output of an auditory filter centred at 1 000 Hz by a tone with a frequency of 1 000 Hz with a sound pressure level of 0 dB presented in a free field with frontal incidence.

3.13

excitation level

L_E

ten times the logarithm to the base 10 of the ratio of the *excitation* (3.12) at the output of an *auditory filter* (3.5) centred at the frequency of interest to the reference *excitation* (3.12), E_0

$$L_E = 10 \lg \frac{E}{E_0} \text{ dB}$$

where the reference excitation E_0 is the excitation produced by a 1 000 Hz tone with a sound pressure level of 0 dB presented in a free field with frontal incidence

3.14 specific loudness

N'

calculated loudness evoked over a frequency band with a bandwidth of 1 ERB_n centred on the frequency of interest

4 General

The method described in this document specifies a method for calculating loudness and loudness level of any sound based on the Moore-Glasberg-Schlittenlacher procedure.

The method involves a sequence of stages. Each stage is described below. However, it is envisaged that those wishing to calculate loudness using this procedure will use one of the computer programs (see [Annex A](#)) provided with this document that implements the described procedure. It is not expected that the procedure will be implemented “by hand”. Such computations would be very time consuming. The source code provided in [Annex A](#) gives an example of the implementation of the method. Other implementations using different software are possible.

NOTE 1 The computational procedure described in this document is an updated version of procedures published earlier elsewhere in References [1] to [5].

NOTE 2 Uncertainties are addressed in [Clause 8](#).

5 Input signal

The signal that is used as input to the algorithm is the waveform for each ear (left and right), sampled using a 32 kHz sampling rate. If the Matlab®¹⁾ code described in [Annex C](#) is used, higher sampling rates for the signal are allowed. These are automatically converted by the Matlab® software to a 32 kHz sampling rate. The signal can be obtained in three ways.

5.1 Single microphone

The sound can be recorded using a single microphone placed at the centre of the position of the listener's head, after the listener has been removed from the sound field. In this case, the sound would be diotic (the same at the two ears) and the single recorded signal would be presented to both input channels of the algorithm.

5.2 Two microphones in the ear canals or microphones in a head and torso simulator

The sound can be recorded using two small probe microphones with the tips placed close to each ear drum (left and right) or using the two ear simulators (left and right) in a head and torso simulator.

5.3 Earphone presentation

If the sound is delivered via earphones, the input signals for the algorithm correspond to the electrical signals delivered to the earphones, but with allowance for the transfer function from each earphone to the eardrum; see [7.2.4](#).

1) Matlab® is a trademark of MathWorks. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results

6 Instrumentation

Measuring instrumentation used to acquire a signal to be used as an input for method [5.1](#) and [5.2](#) shall conform to IEC 61672-1. The microphone(s) used for method [5.1](#) shall have an omnidirectional characteristic or a free-field characteristic. If a head and torso simulator is used it shall conform to IEC 60318-7. For signals acquired using a head and torso simulator, the transfer function of the simulator as supplied by the equipment manufacturer or acquisition software shall be allowed for.

7 Description of the method

7.1 General

The procedure involves a sequence of processing operations, as illustrated in [Figure 1](#).

For each ear, the processing operations are:

- a) a filter to allow for the effects of transfer of sound through the outer and middle ear;
- b) a short-term spectral analysis of the sound spectrum with greater frequency resolution at low than at high frequencies;
- c) calculation of an excitation pattern, representing the magnitudes of the outputs of the auditory filters as a function of centre frequency;
- d) application of a compressive nonlinearity to the output of each auditory filter to transform excitation to specific loudness;
- e) smoothing over time of the resulting instantaneous specific loudness pattern using an averaging process resembling an automatic gain control (AGC) to give short-term specific loudness.

Subsequent stages are:

- f) the short-term specific loudness patterns for each ear are used to calculate broadly-tuned binaural inhibition functions, the amount of inhibition depending on the relative short-term specific loudness at the two ears;
- g) the inhibited specific loudness patterns are summed across frequency to give an estimate of the short-term loudness for each ear;
- h) the binaural short-term loudness is calculated as the sum of the short-term loudness values for the two ears;
- i) the long-term loudness for each ear is calculated by smoothing the short-term loudness for that ear, again by a process resembling AGC;
- j) the binaural long-term loudness is obtained by summing the long-term loudness across ears.

These steps are described sequentially in [7.2](#) to [7.9](#).

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040

©

The transfer of sound through the outer and middle ear is modelled using one of three finite impulse response (FIR) filters with 4 097 coefficients. Different filters are used depending on the method by which the sound was picked up and the method by which the sound was delivered to the listeners. Each filter represents the combined effect of the outer ear and the middle ear. The transfer function for the middle ear is the same as for ISO 532-2:2017^[9], 7.3 and is specified in column 4 of [Table 1](#). This transfer function is referred to as “middle ear only”.

[illegible]

Frequency	Difference between the sound pressure level at the tympanic membrane and the sound pressure level measured in the free field (in the absence of a listener)	Difference between the sound pressure level at the tympanic membrane and the sound pressure level measured in the diffuse field (in the absence of a listener)	Scaled transfer function value for the middle ear
Hz	dB	dB	dB
20	0,0	0,0	-39,6
25	0,0	0,0	-32,0
31,5	0,0	0,0	-25,85
40	0,0	0,0	-21,4
50	0,0	0,0	-18,5
63	0,0	0,0	-15,9

^a Values are in a range that has not been validated.

Table 1 (continued)

Frequency	Difference between the sound pressure level at the tympanic membrane and the sound pressure level measured in the free field (in the absence of a listener)	Difference between the sound pressure level at the tympanic membrane and the sound pressure level measured in the diffuse field (in the absence of a listener)	Scaled transfer function value for the middle ear
Hz	dB	dB	dB
80	0,0	0,0	-14,1
100	0,0	0,0	-12,4
125	0,1	0,1	-11,0
160	0,3	0,3	-9,6
200	0,5	0,4	-8,3
250	0,9	0,5	-7,4
315	1,4	1,0	-6,2
400	1,6	1,6	-4,8
500	1,7	1,7	-3,8
630	2,5	2,2	-3,3
750	2,7	2,7	-2,9
800	2,6	2,9	-2,6
1 000	2,6	3,8	-2,6
1 250	3,2	5,3	-4,5
1 500	5,2	6,8	-5,4
1 600	6,6	7,2	-6,1
2 000	12,0	10,2	-8,5
2 500	16,8	14,9	-10,4
3 000	15,3	14,5	-7,3
3 150	15,2	14,4	-7,0
4 000	14,2	12,7	-6,6
5 000	10,7	10,8	-7,0
6 000	7,1	8,9	-9,2
6 300	6,4	8,7	-10,2
8 000	1,8	8,5	-12,2
9 000	-0,9	6,2	-10,8
10 000	-1,6	5,0	-10,1
11 200	1,9	4,5	-12,7
12 500	4,9	4,0	-15,0
14 000	2,0	3,3	-18,2 ^a
15 000	-2,0	2,6	-23,8 ^a
16 000	2,5	2,0	-32,3 ^a

^a Values are in a range that has not been validated.

7.2.2 Free-field transfer function

For sounds presented in a free field from a frontal direction, it is assumed that the transformation from free-field sound pressure (measured in the absence of the listener at the position corresponding to the centre of the listener's head) to eardrum sound pressure is as specified in Shaw^[21]. This transfer function is the same as for ISO 532-2:2017^[9], 7.3 and is specified in column 2 of Table 1. The overall transfer function for this option is the sum of the transfer functions in columns 2 and 4 of Table 1.

The free-field option can be used for sound picked up by a single microphone placed at the centre of the position of the listener's head, after the listener has been removed from the sound field. In this case, the sound would be diotic (the same at the two ears).

7.2.3 Diffuse-field transfer function

For diffuse-field presentation, the transfer function is derived by averaging the sound-field-to-eardrum transfer function over many directions of incidence. The values used are based on the average of measurements given in the literature^{[21] to [23]}. The diffuse-field transfer function is the same as for ISO 532-2:2017^[9], 7.3 and is specified in column 3 of [Table 1](#). The overall transfer function for this option is the sum of the transfer functions in columns 3 and 4 of [Table 1](#).

The diffuse-field option can be used for diffuse-field listening conditions using sound picked up by a single microphone placed at the centre of the position of the listener's head, after the listener has been removed from the sound field. In this case, the sound would be diotic (the same at the two ears).

7.2.4 Signal recorded using microphones in the ear canals or using a Head and Torso Simulator

The sound waveforms can be recorded using small microphones placed in each ear canal close to the eardrum or they can be recorded using a Head and Torso simulator (also called a dummy head or acoustic manikin) that mimics the acoustic properties of the torso, head, pinna and ear canal, as specified in IEC 60318-7. In this case, the middle-ear only option is used.

7.2.5 Earphone presentation

For sounds presented via earphones, the input to the algorithm can be based on the electrical signals delivered to the earphones. The sensitivity level of the earphones (the sound pressure level produced for a given applied voltage) at a given frequency (e.g. 1 000 Hz) shall be specified. The transfer function chosen depends on the characteristics of the earphones and on the way that the electrical input to the earphones is processed:

- a) When the earphones have a free field response, i.e. when the transfer function from the electrical signal at the input to the earphones to the sound pressure at the eardrum approximates the transfer function in column 2 of [Table 1](#), the free-field option is used.
- b) When the waveforms of the sounds have been pre-processed using a free-field equalizer prior to delivery to the earphones^[24], the free-field option is used.
- c) When the earphones have a diffuse-field response, i.e. when the transfer function from the electrical signal at the input to the earphones to the sound pressure at the eardrum approximates the transfer function in column 3 of [Table 1](#), the diffuse-field option is used.
- d) When the waveforms of the sounds have been pre-processed using a diffuse-field equalizer prior to delivery to the earphones, the diffuse-field option is used.
- e) When sounds are delivered via earphones with a "flat" response at the eardrum or when the electrical signal delivered to the earphones is digitally filtered to simulate the response of the earphone at the eardrum prior to being used as input to the model, the middle-ear only option is used.

For sounds presented via earphones, the sounds can be identical at the two ears (diotic) or they can differ at the two ears (dichotic).

7.3 Calculation of the running short-term spectrum

A running estimate of the sound spectrum at the output of the FIR filter for each ear is obtained by calculating six Fast Fourier Transforms (FFTs) in parallel, using signal segment durations that decrease with increasing centre frequency. This is done to give sufficient spectral resolution at low frequencies and sufficient temporal resolution at high frequencies. The six FFTs are based on Hann-windowed

segments with durations of 2 ms, 4 ms, 8 ms, 16 ms, 32 ms and 64 ms, all aligned at their temporal centres. The windowed segments are zero padded and all FFTs are based on 2 048 sample points. All FFTs are updated every 1 ms. Hence, the overlap of successive segments is at least 50 %.

The use of a Hann window reduces the effective level relative to what would be obtained without a window by 4,26 dB. However, this effect is partly offset by the fact that for a sinusoidal input signal, the calculated short-term spectrum has significant energy in more than one bin. To allow for these effects the level of each component in the calculated spectrum is increased by 3,32 dB. This leads to a calculated loudness of exactly 1 sone for a 1 000 Hz sinusoid presented binaurally in free field with frontal incidence with a sound pressure level of 40 dB.

Each FFT is used to calculate spectral magnitudes over a specific frequency range; values outside that range are discarded. These ranges are 20 Hz to 80 Hz, 80 Hz to 500 Hz, 500 Hz to 1 250 Hz, 1 250 Hz to 2 540 Hz, 2 540 Hz to 4 050 Hz and 4 050 Hz to 15 000 Hz, for segment durations of 64 ms, 32 ms, 16 ms, 8 ms, 4 ms and 2 ms, respectively.

7.4 Calculation of the running short-term excitation pattern

An excitation pattern is calculated from the short-term sound spectrum at 1-ms intervals, using the same method as described in ISO 532-2:2017^[2], 7.4, but with slightly coarser spacing of points along the frequency axis. The excitation pattern is defined as the output of the auditory filters, plotted as a function of centre frequency. The bandwidths and shapes of the filters depend on both the input sound pressure level and the centre frequency, f_c , of the filter. The excitation pattern of a given sound is defined as the output of the auditory filters represented as a function of f_c . The output can be specified either as excitation ratio, E/E_0 , or as excitation level, L_E , as defined in 3.12 and 3.13.

The equivalent rectangular bandwidth, ERB_n (in Hz), of the auditory filter for otologically normal persons and for an input sound pressure level to the cochlea of 51 dB, is specified as a function of the centre frequency of the band pass auditory filter, f_c (in Hz), by [Formula \(1\)](#):

$$ERB_n = 24,673(0,004368f_c + 1) \quad (1)$$

The characteristics of the auditory filter for other input levels are derived as described below.

For a given auditory filter (with a specific f_c), the excitation is calculated by summing the power of the output in response to all of the different frequency components in the input. A first stage in this process is to sum the powers of the components of the input spectrum in $1-ERB_n$ -wide bands, where the width of the bands is defined by [Formula \(1\)](#). The resulting power, converted to decibels (using the reference excitation defined in 3.13), is referred to as the level per ERB_n and is denoted X . It is assumed that the sharpness of the auditory filter depends on X .

The value of X is calculated using a rounded-exponential weighting function (called hereafter a filter) rather than a rectangular weighting function. The rounded-exponential filter is defined by [Formula \(2\)](#)

$$W(g, f_c) = (1 + pg) \exp(-pg) \quad (2)$$

where the value of g at frequency f is given by [Formula \(3\)](#):

$$g = |(f - f_c)| / f_c \quad (3)$$

and p is a dimensionless parameter determining the bandwidth and slope of the filter. For calculating X , the value of p in [Formula \(2\)](#) is set to $4f_c/ERB_n$. The output power of the rounded-exponential filter is calculated over the following ranges given by [Formula \(4\)](#):

$$\text{For } f < f_c \quad g = 0 \text{ to } 1 \quad (4)$$

For $f > f_c$ $g = 0$ to 4

The calculation of X is performed with the filter centred in turn on every component in the input spectrum.

To calculate the output of a given auditory filter in response to a given group of frequency components, it is first necessary to specify the shape of the filter. Each side of the filter is specified to have the form given in [Formula \(2\)](#). The value of p for the lower side of the filter (frequencies below the centre frequency) is denoted p_l , while the value of p for the upper side of the filter (frequencies above the centre frequency) is denoted p_u . The value of p_u is invariant with level and is equal to $4f_c / ERB_n$. The value of p_l is calculated as follows.

The value of p_l for $X = 51$ dB is set equal to $4f_c / ERB_n$. Let $p_l(X, f_c)$ denote the value of p_l at level X and centre frequency f_c . Then [Formula \(5\)](#)

$$p_l(X, f_c) = p_l(51, f_c) - D[p_l(51, f_c) / p_l(51, 1000)](X - 51) \quad (5)$$

where $p_l(51, f_c)$ is the value of p_l at centre frequency f_c for $X = 51$ dB, $p_l(51, 1000)$ denotes the value of p_l at 1 000 Hz for $X = 51$ dB and D is a constant with unit 1/dB and a value of 0,35.

The final excitation pattern is plotted with the scale of centre frequency, f_c , transformed to an ERB_n -number scale. An increase in frequency equal to 1 ERB_n corresponds to a step of one unit on the ERB_n -number scale. For brevity, the unit of the ERB_n -number scale is denoted the Cam, and the scale is denoted the Cam scale.

EXAMPLE According to [Formula \(1\)](#), the value of ERB_n for $f_c = 1\,000$ Hz is approximately 132 Hz, so an increase in frequency from 934 Hz to 1 066 Hz corresponds to a step of 1 Cam.

The relationship of ERB_n -number i to f_c is given by an equation derived from [Formula \(1\)](#)

$$i = 21,366 \lg(0,004\,368 f_c + 1) \quad (6)$$

The excitation pattern is calculated for ERB_n -numbers i from 1,75 to 39 in steps of 0,25. The corresponding centre frequencies f_c are calculated by inverting [Formula \(6\)](#).

7.5 Transformation of excitation into specific loudness

7.5.1 General

The short-term excitation at each centre frequency is transformed to specific loudness, using exactly the same processing steps as described in ISO 532-2:2017^[9], 7.5. The specific loudness pattern for a given ear at this stage is what would occur if there were no input to the other ear.

The excitation ratio E/E_0 is transformed to specific loudness N' in sone/ ERB_n . The calculation of specific loudness depends on two properties of the cochlea:

- excitation at the reference threshold of hearing, and
- gain of the cochlea for inputs with low sound pressure levels,

which are described in [7.5.2](#) and [7.5.3](#). Subclauses, [7.5.4](#) to 7.5.7, describe the calculation procedure based on the excitation ratio E/E_0 .

7.5.2 Reference excitation at the reference threshold of hearing

The reference threshold of hearing is the lowest detectable sound pressure level of a sound in the absence of any other sounds. The function relating the excitation level at the reference threshold of hearing to frequency for monaural listening is specified in [Table 2](#) (binaural listening is considered in [7.7](#)). Interpolation is used to determine values at frequencies between those shown in the table. The

values given in [Table 2](#) are values at the peak of the excitation pattern for sinusoidal signals, i.e., values at the output of the auditory filter centred at the signal frequency. Above 500 Hz, the excitation ratio at the reference threshold of hearing is constant. The peak excitation ratio produced by a sinusoidal signal at threshold (for monaural listening) is denoted E_{THRQ}/E_0 . For frequencies of 500 Hz and above, the value of E_{THRQ}/E_0 is 2,307 (equivalent to an excitation level, L_E , of 3,63 dB).

Table 2 — Excitation level and value of $10\lg G$ at the reference threshold of hearing for monaural listening

Centre frequency Hz	Excitation level at reference threshold (constant for all values above 500 Hz) dB	$10\lg G$ at reference threshold dB
50	28,18	-24,55
63	23,90	-20,27
80	19,20	-15,57
100	15,68	-12,05
125	12,67	-9,04
160	10,09	-6,46
200	8,08	-4,45
250	6,30	-2,67
315	5,30	-1,67
400	4,50	-0,87
500	3,63	0,00
630	3,63	0,00
750	3,63	0,00
800	3,63	0,00
1 000	3,63	0,00

7.5.3 Gain of the cochlear amplifier for inputs with low sound pressure levels

The term G represents the gain of the cochlea for inputs with low sound pressure levels at a specific frequency, relative to the gain at 500 Hz and above (which is assumed to be constant). The product of G and E_{THRQ}/E_0 at a specific frequency is independent of frequency. Column 3 of [Table 2](#) shows the value of G , expressed in decibels, for different frequencies.

EXAMPLE If E_{THRQ}/E_0 is a factor of ten higher than the value at 500 Hz and above, then G is equal to 0,1. More generally, if E_{THRQ}/E_0 is a factor K higher than the value at 500 Hz and above, then G is equal to $1/K$.

7.5.4 Calculation of specific loudness from excitation when $E_{\text{THRQ}}/E_0 \leq E/E_0$

When the excitation evoked by the signal of interest at a specific centre frequency is greater than or equal to the value of E_{THRQ}/E_0 for that frequency, but less than or equal to 10^{10} , which covers the range of most practical applications, the specific loudness is calculated by [Formula \(7\)](#):

$$N' = C \left[(GE/E_0 + A)^\alpha - A^\alpha \right] \quad (7)$$

where $C = 0,063$ sone/Cam. The quantity E/E_0 is dimensionless. For frequencies of 500 Hz and above, the value of α is equal to 0,2 and the value of A is equal to $2 E_{\text{THRQ}}/E_0$. A is dimensionless. Below 500 Hz, the values of α and A are related to the value of G . The relationship of α to G is specified in [Table 3](#). The relationship of A to G is specified in [Table 4](#). In these tables, G has been converted to decibel units. Interpolation is used to determine values at frequencies between those shown in the table.

7.5.5 Calculation of specific loudness from excitation when $E_{\text{THRQ}}/E_0 > E/E_0$

When the excitation evoked by the signal of interest at a specific centre frequency is less than the value of E_{THRQ}/E_0 for that frequency, the specific loudness is calculated as shown in [Formula \(8\)](#):

$$N' = C \left(\frac{2E}{(E + E_{\text{THRQ}})} \right)^{1,5} \left[\left(G \frac{E}{E_0} + A \right)^\alpha - A^\alpha \right] \quad (8)$$

7.5.6 Calculation of specific loudness from excitation when $E/E_0 > 10^{10}$

When the excitation ratio evoked by the signal of interest at a specific centre frequency is greater than 10^{10} , the specific loudness is calculated by [Formula \(9\)](#):

$$N' = C \left(\frac{E/E_0}{1,0707} \right)^{0,2} \quad (9)$$

Note that this corresponds to sound pressure levels greater than 100 dB, and there are few data on loudness perception for such high levels. The procedure has not been validated for such high levels.

Table 3 — Value of the parameter α as a function of the parameter G , converted to decibels

10 lg G dB	α
-25,0	0,266 92
-20,0	0,250 16
-15,0	0,236 79
-10,0	0,222 28
-5,0	0,210 55
0,0	0,200 00

Table 4 — Value of the parameter A as a function of the parameter G , with G expressed in decibels

10 lg G dB	A
-25,0	8,792 3
-24,5	8,658 4
-24,0	8,524 5
-23,5	8,390 6
-23,0	8,256 7
-22,5	8,132 4
-22,0	8,009 5
-21,5	7,886 6
-21,0	7,763 7
-20,5	7,640 8
-20,0	7,517 9
-19,5	7,426 8
-19,0	7,336 6
-18,5	7,246 8
-18,0	7,156 2
-17,5	7,066 1

10 lg G dB	A
-12,0	6,100 2
-11,5	6,016 9
-11,0	5,933 6
-10,5	5,850 4
-10,0	5,767 1
-9,5	5,699 8
-9,0	5,632 8
-8,5	5,570 5
-8,0	5,508 2
-7,5	5,445 9
-7,0	5,383 7
-6,5	5,321 4
-6,0	5,259 1
-5,5	5,196 9
-5,0	5,134 6
-4,5	5,080 6

Table 4 (continued)

10 lg G dB	A
-17,0	6,975 9
-16,5	6,885 7
-16,0	6,798 4
-15,5	6,715 3
-15,0	6,632 2
-14,5	6,542 0
-14,0	6,451 8
-13,5	6,361 6
-13,0	6,271 4
-12,5	6,183 4

10 lg G dB	A
-4,0	5,028 7
-3,5	4,976 8
-3,0	4,924 9
-2,5	4,873 0
-2,0	4,821 1
-1,5	4,769 2
-1,0	4,717 3
-0,5	4,665 4
0,0	4,613 5

7.6 Calculation of short-term specific loudness

The specific loudness as a function of centre frequency is called the specific loudness pattern. The pattern calculated from a single short-term spectral estimate is called the instantaneous specific loudness pattern. This is smoothed over time to give the short-term specific loudness pattern. The short-term specific loudness is calculated from a running average of the instantaneous specific loudness, separately for each centre frequency, using a process resembling the way that a control signal is generated in an automatic gain control (AGC) circuit, with an attack time, T_a , and a release time, T_r . This is implemented in the following way. N'_n is defined as the running short-term estimate of specific loudness at the time corresponding to the n^{th} frame, N_n as the instantaneous specific loudness at the n^{th} frame, and N'_{n-1} as the short-term specific loudness at the time corresponding to frame $n-1$.

If $N_n > N'_{n-1}$ (corresponding to an attack, as the instantaneous specific loudness at frame n is greater than the short-term loudness at the previous frame), then [Formula \(10\)](#) gives:

$$N'_n = \alpha_a N_n + (1 - \alpha_a) N'_{n-1} \quad (10)$$

where α_a is a constant that is related to T_a according to [Formula \(11\)](#):

$$\alpha_a = 1 - \exp(-T_0 / T_a) \quad (11)$$

where T_0 is the time interval (1 ms) between successive values of the instantaneous specific loudness.

If $N_n \leq N'_{n-1}$ (corresponding to a release, as the instantaneous specific loudness at frame n is less than the short-term loudness at the previous frame), then [Formula \(12\)](#):

$$N'_n = \alpha_r N_n + (1 - \alpha_r) N'_{n-1} \quad (12)$$

where α_r is a constant that is related to T_r according to [Formula \(13\)](#):

$$\alpha_r = 1 - \exp(-T_0 / T_r) \quad (13)$$

The values of α_a and α_r are 0,045 and 0,033, respectively.

7.7 Smoothing of short-term specific loudness and application of binaural inhibition

When a sound is presented to both ears, the loudness is less than that predicted from summing the loudness values from the two ears^{[3][4]}. When the sound is identical at the two ears (diotic presentation), the loudness is close to 1,5 times that evoked by the sound at each ear alone^[3]. This effect

is implemented using the concept of inhibitory interactions between the two ears. The application of this concept involves two stages:

- smoothing of the short-term specific loudness pattern across frequency, to implement the broad tuning of the binaural inhibition;
- reduction of short-term specific loudness in each ear at each centre frequency depending on the relative values of the smoothed short-term specific loudness at the two ears.

Both processing operations, smoothing and application of binaural inhibition, are implemented exactly as described in ISO 532-2:2017^[9], 8.1.

Let $N'_L(i)$ and $N'_R(i)$ be the specific loudness values evoked at the left and right ears, respectively, at a given frequency expressed by ERB_n -number i . It is assumed that there are inhibitory interactions between the two ears, such that a signal at the left ear inhibits (reduces) the loudness evoked by a signal at the right ear, and vice versa. It is assumed further that these inhibitory interactions are relatively broadly tuned.

To implement the broad tuning of the inhibition, the specific loudness pattern at each ear is initially smeared or smoothed by a process resembling convolution with a Gaussian-shaped weighting function.

The smoothed result at i Cam for the left ear is calculated as shown in [Formula \(14\)](#):

$$N'_L(i)_{\text{smoothed}} = \sum_{D_i=\max[-18,(i-1,75)]}^{D_i=\min[18,(39-i)]} N'_L(i-D_i) \exp[-(BD_i)^2] \quad (14)$$

where D_i is the deviation from the given i and B is a parameter determining the degree of spread of inhibition along the ERB_n -number scale. The value of B is 0,08. D_i is changed in steps of 0,25 Cam. This equation results in an increase in overall magnitude after smoothing, but this is irrelevant because only ratios of the smoothed specific loudness patterns for the two ears are used subsequently. When $i+D_i$ is less than 1,75 Cam or greater than 39 Cam, N'_L is set to 0.

Similarly for the right ear, the smoothed specific loudness at i Cam is given by [Formula \(15\)](#):

$$N'_R(i)_{\text{smoothed}} = \sum_{D_i=\max[-18,(i-1,75)]}^{D_i=\min[18,(39-i)]} N'_R(i-D_i) \exp[-(BD_i)^2] \quad (15)$$

The values determined by [Formulae 14](#) and [15](#) are determined for $i = 1,75$ to 39 at intervals of 0,25.

Let $INH_L(i)$ denote the factor by which the specific loudness evoked by the signal at the left ear is reduced after inhibition produced by the signal at the right ear. Let $INH_R(i)$ denote the factor by which the specific loudness evoked by the signal at the right ear is reduced after inhibition produced by the signal at the left ear. The inhibition is modelled by [Formulae \(16\)](#) and [\(17\)](#):

$$INH_L(i) = 2 / \left[1 + \left\{ \text{sech}(N'_R(i)_{\text{smoothed}} / N'_L(i)_{\text{smoothed}}) \right\}^\theta \right] \quad (16)$$

$$INH_R(i) = 2 / \left[1 + \left\{ \text{sech}(N'_L(i)_{\text{smoothed}} / N'_R(i)_{\text{smoothed}}) \right\}^\theta \right] \quad (17)$$

where sech represents the mathematical function hyperbolic secant, and $\theta = 1,597\ 8$. To prevent problems associated with dividing by zero when $N'_L(i)_{\text{smoothed}}$ or $N'_R(i)_{\text{smoothed}}$ are zero, a small number (10^{-13}) is added to the values of $N'_L(i)_{\text{smoothed}}$ and $N'_R(i)_{\text{smoothed}}$ prior to entering them in [Formulae \(16\)](#) and [\(17\)](#).

The gain values calculated using [Formulae \(16\)](#) and [\(17\)](#) are applied to the original short-term specific loudness values in each ear to give inhibited short-term specific loudness values. Specifically, the value of $N'_L(i)$ is divided by $INH_L(i)$ and the value of $N'_R(i)$ is divided by $INH_R(i)$.

7.8 Calculation of short-term loudness

The short-term loudness for each ear at frame n , S'_n , is calculated by summing the inhibited short-term specific loudness values for that ear over Cam values, i , from 1,75 to 39 in steps of 0,25 Cam and dividing by 4 (to allow for the 0,25-Cam steps). The binaural short-term loudness at frame n is obtained by summing the short-term loudness values at frame n across the two ears.

7.9 Calculation of long-term loudness

The long-term loudness for each ear is calculated from the short-term loudness for that ear, again using a form of averaging resembling the operation of an AGC circuit. The long-term loudness for a given ear at the time corresponding to frame n is denoted S''_n . If $S'_n > S''_{n-1}$, then [Formula \(18\)](#):

$$S''_n = \alpha_{al} S'_n + (1 - \alpha_{al}) S''_{n-1} \quad (18)$$

where α_{al} is a constant related to the attack time of the averager, T_{al} , according to [Formula \(19\)](#):

$$\alpha_{al} = 1 - \exp(-T_0 / T_{al}) \quad (19)$$

If $S'_n \leq S''_{n-1}$, then [Formula \(20\)](#):

$$S''_n = \alpha_{rl} S'_n + (1 - \alpha_{rl}) S''_{n-1} \quad (20)$$

where α_{rl} is a constant related to the release time of the averager, T_{rl} , according to [Formula \(21\)](#):

$$\alpha_{rl} = 1 - \exp(-T_0 / T_{rl}) \quad (21)$$

The values of α_{al} and α_{rl} are 0,01 and 0,001 33, respectively.

The binaural long-term loudness at frame n is obtained by summing the long-term loudness values across the two ears at that n .

For sounds like speech and music, the calculated long-term loudness fluctuates slightly even when the sound lasts several seconds. The loudness of sounds lasting up to 5 s can be predicted from the maximum value of the long-term loudness [\[4\]](#) [\[6\]](#) to [\[8\]](#).

NOTE For sounds lasting longer than 5 s, the loudness often changes over time and it is difficult or impossible to assign a single loudness value to the sound. For example, a piece of orchestral music can have loud passages and soft passages. For this reason, the prediction of overall loudness for sounds with durations above 5 s is beyond the scope of this standard. Nonetheless several applications and regulations can require a single representative loudness value, for example the overall loudness near a busy street or close to an airport. It is likely that summary statistics that are based on long-term loudness give a far more accurate estimate of overall loudness than other metrics, for example weighted sound levels. A suitable metric for calculating a representative overall loudness value can be obtained by converting the running long-term loudness level values to running intensity-like values, averaging the resulting intensity-like values over time, and then converting the average back to a loudness level. The overall loudness calculated in this way is highly correlated (correlation greater than 0,9) with subjective evaluations of the loudness of the sounds used by Schlittenlacher et al. [\[13\]](#), which had a duration of 10 s and high fluctuations in loudness over time, for example produced by hammer blows.

7.10 Relationship between loudness level and loudness

The procedure can be used to determine the relationship between loudness level in phons and loudness in sones. This is done by using as input to the procedure a long-duration (at least 2 s) 1 000 Hz sinusoid and specifying the conditions as binaural listening in a free field with frontal incidence. The input level of the 1 000 Hz tone is then equal to its loudness level in phons and the maximum value of the long-term loudness predicted by the procedure gives the corresponding calculated loudness, N , in sones. The relationship between loudness level in phons and loudness in sones is specified in [Table 5](#). The sone values in [Table 5](#) were obtained using a signal duration of 5 s including raised-cosine shaped rise/

fall times of 100 ms and using the reference Matlab® implementation of the procedure described in [Annex A](#) (TV2018.m). The use of other signal durations above 2 s and the use of other rise/fall times, including no rise/fall time at all, may lead to slightly different outcomes, with a range of approximately $\pm 0,5$ %.

NOTE Values in sones in the following tables are presented to an approximately constant percentage precision, not to a constant absolute precision. The values are therefore presented to three significant figures, except for very low sone values or for random noises, for which the values are presented to one or two significant figures.

Table 5 — Relationship between loudness level in phons and calculated loudness in sones, N

Loudness level phon	Calculated loudness, N sone
0,00	0,001
2,20	0,002
4,00	0,004
5,00	0,006
7,50	0,014
10,00	0,025
15,00	0,066
20,00	0,138
25,00	0,252
30,00	0,422
35,00	0,664
40,00	1,00
45,00	1,46
50,00	2,09
55,00	2,95
60,00	4,11
65,00	5,71
70,00	7,92
75,00	11,0
80,00	15,4
85,00	21,7
90,00	31,1
95,00	44,7
100,00	64,8
105,00	94,3
110,00	138
115,00	205
120,00	306

7.11 Calculation of the reference threshold of hearing

The reference threshold of hearing of a sound is taken to correspond to the level at which the procedure gives a calculated loudness of 0,002 sone. This value corresponds to a loudness level of 2,2 phon. The procedure can be used to calculate the reference threshold of hearing of a given sound by determining the input level that leads to a calculated loudness level of 2,2 phon. Thus calculated, the reference

thresholds of hearing for sinusoids are within 0,2 dB of those specified in ISO 389-7:2019^[19] for all frequencies from 0,05 kHz to 12,5 kHz.

8 Uncertainty of calculated loudness sounds

The use of the method described in this document will produce results with uncertainty comparable to the uncertainty of perceived loudness assessed in a psychoacoustic experiment, assuming that the loudness data are averaged for a sample of at least 10 otologically normal persons.

Since sufficient information needed to calculate uncertainty of the loudness calculation described in this document by applying the otological approach described in ISO/IEC Guide 98-3^[25] is not currently available, the determination of uncertainty has been based on the reproducibility of the data obtained across a large number of loudness assessment studies conducted in various laboratories.

The uncertainty of loudness calculation for stationary sounds described in this document has been based on reproducibility of loudness matching judgments and judgments relating loudness to loudness level for a number of simple and complex input sounds. This uncertainty is based on the results of several studies conducted with $n \geq 10$ otologically normal persons.

Since the perceptual loudness data are normally distributed, the maximum deviation of the data has been defined as the $[-3\sigma; +3\sigma]$ range corresponding to 99,6 % of all the data, where σ is the standard deviation of the sample distribution. Based on a number of studies, this maximum deviation is approximately equivalent to a change in loudness level of 4 phon (e.g. References ^[13]^[26]^[27]). This uncertainty indicates that sounds that differ in “true” loudness level by ± 4 phon may result in the same loudness estimate. Alternatively, it indicates that a sound of a given loudness may cause loudness level judgments that differ up to ± 4 phon.

The standard deviation (standard uncertainty) of such reproducibility of the data is 1,4 phon and the expanded uncertainty for a stated coverage probability of 95 % is two times the standard deviation of the reproducibility, that is, 2,8 phon. The expanded uncertainty of loudness calculation is the maximum uncertainty permitted for demonstration of conformance to the requirements of this document. It is the uncertainty of the loudness calculation that the user needs to take into account.

Additionally, the random error estimated by expanded uncertainty can be compounded by a systematic error resulting from methodological differences between a given study and the studies that were used as the basis for the development of the loudness calculation procedure described in this document. Such systematic errors, if present, can increase the overall deviation of the loudness calculated according to this document from the perceived loudness but cannot be a priori estimated, since they can be different in each study.

9 Data reporting

The following information shall be reported:

- a) the sound under consideration;
- b) a reference to this document, i.e. ISO 532-3:2023;
- c) the method used to record the sound and the assumed presentation conditions (free field, diffuse field, earphone);
- d) whether the sound is presented monaurally or binaurally;
- e) the long-term loudness in sones as a function of time;
- f) the peak long-term loudness in sones;
- g) the long-term loudness level in phons as a function of time;
- h) the peak long-term loudness level in phons;

- i) if required, the short-term loudness as a function of time in sones;
- j) if required, the short-term loudness level as a function of time in phons.

STANDARDSISO.COM : Click to view the full PDF of ISO 532-3:2023

Annex A (informative)

Software for the calculation of loudness according to the method in this document

The software implementing the procedure described here is available in two forms: as “C” code and as Matlab® code. The executable C code runs on a personal computer (PC) under a “DOS” prompt, also called a “command” prompt. The Matlab® software should run on a PC or Mac with Matlab® installed. The C code is more computationally efficient and may be preferred when the input waveform file has a long duration or when there are many files to be processed. The Matlab® code is called TV2018.m. This is referred to as the reference implementation, and it is intended to be easy to understand, especially for those familiar with Matlab®. Test signals are described in [Annex B](#) and are the same as those in Annex B of ISO 532-1 with four exceptions as described in the Annex. These can be used to compare the outputs using the various methods.

Both versions of the software require a stereo “wav” file as input. The Matlab® version converts a mono file to a stereo file with the same signal in both channels, i.e. it assumes diotic presentation if a mono file is specified as input. The two channels are referred to as “channel 1” and “channel 2”. The C code requires the sample rate to be 32 kHz and the resolution to be 16 bit. The Matlab® code automatically converts the sample rate to 32 kHz if it is not already 32 kHz. For free-field frontal incidence or diffuse-field listening conditions, the signal may be picked up by a single microphone at the position corresponding to the centre of the listener’s head, giving a single-channel wav file. In this case the C code requires a two-channel (stereo) wav file to be created, with the same signal in each channel.

Each program starts by dividing the two-channel input wav file into two mono signal files (stored as .wav for the C code and as 64-bit double-precision vectors for the Matlab® code). The program then filters each mono file with the filter selected to account for transmission through the outer and middle ear (df.32k, ff.32k, or midealr.32k). The level of the resulting files can be higher after filtering than before, especially when the input signal has strong components close to 3 kHz. This is not important when using the Matlab® code, as the files are not represented in .wav format. However, for the C code the files are represented in .wav format and it is recommended to have at least 8 dB of “headroom” in the original input file to avoid the signal amplitude after filtering exceeding the full-scale value (overload). In other words, the largest peak in the input file should have a level 8 dB or more below the full-scale level. If overload occurs, i.e. the full-scale level is exceeded after filtering, the output text file containing the results of the loudness computations will give a message about “overflows” and “underflows” and will indicate how often they occurred. The files “mon_left_fir.wav” and “mono_right_fir.wav”, which represent the signal after filtering to account for transmission through the outer and middle ear, are saved by the C code and can be inspected using suitable software. Note that the wav files generated by the C code are over-written each time the program is re-run, so if it is desired to retain any of the files they should be re-named.

A second Matlab program called TV2018b.m is provided to allow faster computations. Its code for computation of the excitation patterns is optimized for Matlab®^[28] but is less easy to understand than that for the reference version. The other parts of TV2018b.m are identical to those for the reference method. The loudness levels calculated using TV2018b.m are within 0,05 phons of those for the reference implementation.

This program code and software can be freely downloaded from the following URN:

<https://standards.iso.org/iso/532/-3/ed-1/en/>

The software is entirely informative and provided for the convenience of the user. Use of the provided software is not required for conformance with this document. The software is not warranted to be appropriate for any particular purpose. No warranties or guarantees are provided.

This program is made available with permission of the authors (Brian R. Glasberg, Brian C. J. Moore, and Josef Schlittenlacher, Department of Psychology, University of Cambridge, UK).

STANDARDSISO.COM : Click to view the full PDF of ISO 532-3:2023

Annex B (informative)

Test signals used for verification of this document

B.1 General

The test signals described below are the same as those in ISO 532-1:2017, Annex B, with four exceptions:

- a) the sample rate is 32 kHz;
- b) the .wav files are two-channel (stereo) files;
- c) for the synthetic signals 2-9, the duration is 5 s, including 100-ms raised-cosine shaped rise/fall ramps;
- d) signal 1 in ISO 532-1:2017, Annex B is omitted, since that signal is not specified as a time-domain waveform.

The waveform files of the signals listed below are included as part of the standard. They may be used to verify a user's implementation of the procedure. [Table B.1](#) shows the loudness and loudness level for each test signal calculated from the maximum value of the long-term loudness given by the reference Matlab® implementation of the procedure described in [Annex A](#) (TV2018.m).

B.2 Steady synthetic signals

Test signal 2:

A 250 Hz tone presented binaurally in free field with a sound pressure level of 80 dB.

Test signal 3:

A 1 kHz tone presented binaurally in free field with a sound pressure level of 60 dB.

Test signal 4:

A 4 kHz tone presented binaurally in free field with a sound pressure level of 40 dB.

Test signal 5:

A pink noise presented binaurally in free field with an overall sound pressure level of 60 dB.

B.3 Time varying synthetic signals

Test signal 6:

A 250 Hz tone presented binaurally in free field with a time varying sound pressure level starting at 30 dB and increasing linearly to 80 dB.

Test signal 7:

A 1 kHz tone presented binaurally in free field with a time varying sound pressure level starting at 30 dB and increasing linearly to 80 dB.

Test signal 8:

A 4 kHz tone presented binaurally in free field with a time varying sound pressure level starting at 30 dB and increasing linearly to 80 dB.

Test signal 9:

A pink noise presented binaurally in free field with a time varying sound pressure level starting at 30 dB/third octave band and increasing linearly to 80 dB/third octave band.

Test signal 10:

A 1 kHz tone pulse presented binaurally in free field with a peak rms sound pressure level of 70 dB and a duration of 10 ms.

Test signal 11:

A 1 kHz tone pulse presented binaurally in free field with a peak rms sound pressure level of 70 dB and a duration of 50 ms.

Test signal 12:

A 1 kHz tone pulse presented binaurally in free field with a peak rms sound pressure level of 70 dB and a duration of 500 ms.

Test signal 13:

A combination of two 1 kHz tone pulses presented binaurally in free field, the first pulse with a peak rms sound pressure level of 60 dB and a duration of 100 ms and the second pulse with a peak rms sound pressure level of 80 dB and a duration of 10 ms.

B.4 Technical time varying signals

Test signal 14:

Propeller-driven airplane noise measured with a single microphone. Free field presentation with frontal incidence is assumed.

Test signal 15:

Vehicle interior noise measured at 40 km/h with a single microphone. Diffuse field presentation is assumed.

Test signal 16:

Hairdryer noise measured with a single microphone. Free field is assumed for the calculation of loudness.

Test signal 17:

Machine gun firing 6 bursts of automatic fire, of differing lengths per burst. Free field presentation with frontal incidence is assumed.

Test signal 18:

Hammering sound, banging a small picture-hanging type nail into a typical hollow wooden door. Free field presentation with frontal incidence is assumed.

Test signal 19:

Rattling slow creak from a door opening. Free field presentation with frontal incidence is assumed.

Test signal 20:

Shaking coins sound. Free field presentation with frontal incidence is assumed.