AKUSTICKÉ LISTY

České akustické společnosti

www.czakustika.cz

ročník 20, číslo 2 červenec 2014

Obsah

ČESKÁ AKUSTICKÁ SPOLEČNOST

Akustické listy, **20**(2), červenec 2014, str. 3–7

Nejistoty typu A, B a pojem hodnoty veličiny

Type A and Type B Evaluation of Uncertainty, and Concept of Quantity

Jan Obdržálek

MFF UK – ÚTF, V Holešovičkách 2, 180 00 Praha 8
jan.obdrzalek@mff.cuni.cz

The base concepts concerning "uncertainty" and related topics are presented, explained using examples and discussed. Namely, the concept of "true value" passed significant evolution from so called "Error approach" to "Uncertainty approach". The concept of quantity value itself is concistent with the concept of uncertainty, as shown, rather than "exact value" given by one exact real number.
Certain particular schisma in the past Czech terminology has been solved and some general rules for terminology are mentioned.

1. To jisté o nejistotě

1.1. Motivace

„Nobody is perfect“, jak známe mj. z filmu Někdo to rád horké. Taky víme, že „Dvakrát nevstoupíš do téže řeky“, a my bychom ještě fyzikálně dodali „a změříš-li její průtokovou rychlost dvakrát, nenaměříš totéž“. Přitom celá fyzika je založena na měření a souhlas s naměřenými hodnotami je koneckonců základní kritérium pravdivosti a oprávněnosti jakékoli fyzikální teorie. (Od Galilea údajně pochází rada „Co se dá změřit, to změř, a co se nedá změřit, převeď na měřitelné“.) V praktických aplikacích fyziky je měření reality ještě závažnější, neb tam jde zpravidla i o peníze, a to začasté o peníze nemalé. U decibelů si může přijít na své i Zákon.

Podobná situace nastává i v jiných disciplínách. Někdy potřebujeme velký počet číselných hodnot zachytit co nejjednodušeji. Víme samozřejmě, že se tím část informace ztrácí, ale my ji ochotně oželíme, zachytíme-li dostatečně věrohodně vše, co jsme při našich měřeních pokládali za opravdu podstatné.

Přesto nezoufejme. Síla slabých spočívá v tom, že jsou si své slabosti vědomi. Stačí proto vědět si rady, jak si počínat, když opakovaná měření nedávají totéž, a jak tuto skutečnost sdělit všem těm ostatním, které naše výsledky zajímají.

Záměrně jsem zvolil začátek v nezvykle volném slohu; pokusím se totiž ve výkladu v podobném duchu pokračovat a vyhnout se odborným termínům, jen abych zdůvodnil argumenty uvedené na konci – v terminologii.

1.2. Opakované měření jediné neznámé hodnoty; nejistota typu A

Předpokládejme, že měříme veličinu, která má určitou, nám ale a priori neznámou hodnotu; budiž to např. prahové napětí U_{prah}, při kterém zaniká zkoumaný děj.

Měříme opakovaně. Je to sice v různých situacích (minimálně v různých dobách, v různých místech nebo obojí), ale předpokládáme, že tyto situace mají z našeho hlediska stejné vlastnosti. Výsledkem měření je posloupnost n hodnot $\{U_k\}_{k=1}^{n}$. Pokládáme-li všechna měření za rovnoprávná, můžeme určit průměr (neboli aritmetickou střední hodnotu) z n nezávislých hodnot vztahem

$$\overline{U} = \frac{1}{n}\sum_{k=1}^{n} U_k \; , \tag{1}$$

dále varianci s^2 počítanou pouze z $n-1$ nezávislých hodnot $U_k - \overline{U}$; všech hodnot je sice n, ale je mezi nimi vztah, totiž $\sum\limits_{k=1}^{n} (U_k - \overline{U}) = 0$, tedy

$$s^2 = \frac{1}{n-1} \sum\nolimits_{k=1}^{n} \left(U_k - \overline{U}\right)^2, \tag{2}$$

a konečně veličinu[1]

$$u = s_{\overline{U}} = \sqrt{\frac{1}{n(n-1)} \sum_{k=1}^{n} (U_k - \overline{U})^2} \tag{3}$$

(přesně zvanou *výběrová směrodatná odchylka aritmetického průměru nezávislých pozorování*), která představuje *standardní nejistotu určenou metodou typu* A (pro stručnost zvanou často jen *nejistota typu* A). Je pro ni příznačné, že ji můžeme učinit teoreticky libovolně malou tím, že zvětšíme počet n měření. Připomeňme, že při *Gaussově* neboli *normálním* rozdělení sice *roste* maximální absolutní odchylka jednotlivého měření při rostoucím n jako $\sqrt{n}$; měření je však n, takže střední odchylka *klesá,* a to jako $\sqrt{n}/n = 1/\sqrt{n}$.

Student v praktiku má zpravidla za to, že je ideální (být líný a) měřit jen jednou. Pak má jediný výsledek a neváhá, kolik vlastně vyšlo. Asistent mu ovšem vzápětí ochotně vysvětlí, že pak je sice opravdu $U_k - \overline{U} = 0$, ale

[1] Výběrový parametr se obvykle značí x (zde je to napětí U), jeho rozptyl pak u, což ovšem nesouvisí se značkou U užitou zde pro napětí. Výběrovou směrodatnou odchylku průměru značíme $s_{\overline{x}_k}$, výběrový rozptyl $s^2_{\overline{x}_k}$ a výběrovou kovarianci $s_{\overline{x}_{j,k}}$.

Přijato 23. dubna 2014, akceptováno 7. července 2014.

nulu má i ve jmenovateli před sumou. Čili ať to bez odmluv naměří ještě devětkrát. Dejme tomu, že postupně naměří hodnoty (v mV): $x/\mathrm{mV} = 284$; 284; 283; 283; 288; 285; 286; 282; 284; 281; pak vyjde $\overline{x} = 284$; $u = 0{,}632$ a je zřejmé, že po dalších měřeních může standardní nejistota dále klesat.

Potud tedy nejistota typu A. Odchylky jednotlivých dílčích hodnot pokládáme za náhodné a jejich příčiny za neznámé. Proto také zvětšováním počtu měření můžeme nejistotu typu A snižovat. Tím se nejistota typu A liší od nejistoty typu B.

1.3. Nejistoty typu B

Bohužel, naše úvodní tvrzení „nobody is perfect" platí i o měřicím přístroji. Ten je pak *zdrojem nejistot typu* B, které je rovněž nutno zahrnout do výsledku.[2] Student se od asistenta dozví, že své deseticiferné výsledky (z kalkulačky) musí upravit s přihlédnutím na třídu přesnosti Avometu. Ale i kdyby měl použitý voltmetr třídu přesnosti 0,1 a plný rozsah 1 V, přičemž údaj se může lišit od měřené hodnoty o 0,001 V, je nutno tuto nejistotu také započítat. Jde-li o digitální údaj limitovaný počtem cifer, můžeme předpokládat, že např. 0,284 znamená se stejnou pravděpodobností cokoli v intervalu délky 0,001 kolem této hodnoty (někdy symetricky, jindy ne). Konečný počet míst pochopitelně není jediným zdrojem nejistot; digitální měřicí přístroj obsahuje převodník, jehož činnost může být ovlivňována např. teplotou prostředí.

Nejistoty těchto typů nesnížíme opakovaným měřením.[3] Příspěvky těchto nejistot, předem zpracované s přihlédnutím k typu statistického rozdělení (normální, rovnoměrné, ...) pro každou z nejistot, je ovšem rovněž nutno do výsledku zahrnout.

1.4. Skládání nejistot

Uvažme, že máme několik vlivů, z nichž každý může ovlivnit údaj měřicího přístroje, aniž se mění měřená veličina. Jak zjistíme jejich úhrnný vliv na výsledek?

Klíčová otázka je, zda lze tyto rušivé vlivy pokládat za nezávislé (nepřesnost tabulkových hodnot a pracovní teplota), nebo zda jsou spolu provázány (teplota ovlivní jak pozorovaný děj, tak i činnost převodníku v měřiči). Vzájemné souvislosti – korelací těchto vlivů – lze po jejich zjištění popsat *kovarianční maticí*; stačí totiž lineární přiblížení, protože odchylky způsobené rušivými vlivy jsou zpravidla malé, a matice popisuje nejobecnější lineární transformaci mezi vektory, kterými popisujeme jak jednotlivé hodnoty měřené veličiny, tak i hodnoty použitých konstant apod.

Uvažujme dva rušivé vlivy s účinky Δx_1 a Δx_2. V nejnepříznivějším případě se jejich účinky sečtou: $\Delta x = \pm\left(|\Delta x_1| + |\Delta x_2|\right)$, působí-li v souhlasném směru; v případě nejpříznivějším se odečtou $\Delta x = \pm\left(|\Delta x_1| - |\Delta x_2|\right)$; jejich směry budou protilehlé. Oba případy signalizují nejsilnější korelaci obou vlivů. Naopak nezávislost vlivů se projeví tak, že vektory zobrazující jejich účinek budou navzájem kolmé. Výsledná odchylka bude tedy podle Pythagorovy věty rovna

$$\Delta x = \sqrt{\Delta x_1^2 + \Delta x_2^2}\,. \tag{4}$$

Geometrická představa usnadní pochopit, proč nezávislost vektorů zobrazujeme jejich kolmostí.

Stejnou představu a euklidovskou vzdálenost můžeme přenést do n-rozměrného prostoru; střední odchylku způsobenou n nezávislými vlivy (pro stručnost pišme x namísto Δx) pak určíme jako

$$x = \sqrt{\sum\nolimits_{k=1}^{n} x_k^2}\,. \tag{5}$$

Tímto způsobem budeme tedy skládat vlivy, které pokládáme za nezávislé: sčítat čtverce jejich účinku, a součet pak odmocníme. Všimněme si, že toto pravidlo je vnitřně konzistentní v tom směru, že takto můžeme i dodatečně přidávat další členy, stejně jako můžeme rozdělit nezávislé vlivy do skupin a ty uvedeným způsobem skládat. Jestliže je např.

$$_{[1..3]}x = \sqrt{\sum\nolimits_{k=1}^{3} x_k^2}; \quad _{[4..9]}x = \sqrt{\sum\nolimits_{k=4}^{9} x_k^2}\,,$$

pak jejich současné působení je opravdu popsáno účinkem

$$_{[1..9]}x = \sqrt{\left(_{[1..3]}x\right)^2 + \left(_{[4..9]}x\right)^2} = \sqrt{\sum\nolimits_{k=1}^{9} x_k^2}.$$

1.5. Další zobecnění nejistot

K úplnosti (nerozebíráme-li ovšem jen letmo zmíněné korelované vlivy) zbývá dodat jen dvě skutečnosti.

Především, zatím jsme uvažovali všechna měření za „stejně závažná" (stejně kvalitní, hodnotná apod.). Můžeme však mít rozumný důvod „oznámkovat" jednotlivé údaje x_k tím, že jim připíšeme různé váhyw_k. Nejjednodušší představa je, že prostě příslušnou hodnotu započteme nikoli jednou, ale – „pro veliký úspěch" – třeba sedmkrát; pak bude mít váhu 7. Pochopitelně se příslušně zvětší počet údajů ve jmenovateli zlomku před sumou ve střední hodnotě. Budeme tedy mít *vážený průměr* (s váhou w) podle vzorce

$$\overline{U} = \sum_{k=1}^{n} w_k x_k \Big/ \sum_{k=1}^{n} w_k \tag{6}$$

a snadno nahlédneme, že vzorec (1) vyjadřuje totéž, zvolíme-li $w_k = 1$, tedy zvolíme-li všechny váhy stejné a rovné jedné.

[2] A platí to i o měřicí metodě, o podmínkách, za nichž měříme (teplota, tlak, ...), o tabulkových konstantách, a dokonce i o použitých vzorcích, pokud totiž obsahují empirické vztahy a empirické konstanty. Toto vše přispívá k nejistotě typu B.

[3] Lze si představit jiné provedení měření, které by toto omezení údaje obešel, např. rozmítáním měřené hodnoty a následným středováním. To ale teď není předmětem našeho zájmu.

Poslední zobecnění se týká *rozšířené nejistoty*. Při některých rozděleních (rovnoměrné, trojúhelníkové) je už v podstatě dáno, že odchylka konkrétní hodnoty od střední hodnoty má svou nepřekročitelnou mez. Při jiných, zejména při normálním, v praxi nejčastěji užívaném, taková principiální mez není. Pravděpodobnost odchylky, pravda, klesá exponenciálně s její velikostí, nicméně alespoň teoreticky zde pevná mez není. Význam parametru σ normálního rozdělení s hustotou pravděpodobnosti

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{7}$$

(čím větší σ, tím plošší průběh) pro pravděpodobnost p nalezení konkrétní hodnoty v intervalu $(-k\varphi; +k\varphi)$ udává tato tabulka:

p	0,8	0,9	0,95	0,98	0,99	0,999
k	1,282	1,645	1,960	2,326	2,576	3,291

V praxi proto někdy zadáme namísto nejistoty σ hodnotu $k\sigma$ s $2 \leq k \leq 3$, tzv. rozšířenou nejistotu; tím zajistíme dostatečnou spolehlivost výsledku pro všechny případy.

2. Pojem hodnoty veličiny

2.1. Opravdu přesná hodnota?

Lehce pikantní se zdá, že nám zde zatím nikde nevystupuje to, co nás vlastně zajímá: veličina U_{prah}, kvůli které měření provádíme. Její hodnota prostě není známa. Navíc je jak alternativa, že je zcela přesná (např. počet pulzů), ale taky nemusí striktně vzato existovat: pokud je totiž dána racionálním číslem (a tím spíše číslem reálným), pak zpravidla *není* onou exaktní hodnotou čísla z matematiky, ale je stejně jen přiblížením, jakýmsi rámcem, kam ji situujeme.

2.2. Jak přesně?

K představě „matematicky přesné“ hodnoty veličiny uveďme dvě velmi kritické poznámky:

1. Jakou mám výšku? Podle vojenské knížky 184 cm, což je rozumný údaj. Protože se ale přes noc mé obratle uvolní a přes den stlačují a sesedají, mění se má výška prakticky průběžně a např. přesnost v milimetrech je iluzorní. I kdybychom se však omezili na jediný okamžik měření, pak při vyšší přesnosti – třeba na nanometry – zjistíme, že měřená veličina není dobře definována (a přitom to v praxi nevadí).

2. Jediné racionální číslo ξ je schopno obsahovat veškerou moudrost lidstva (a ještě hodně místa zbude). Např. soubor tohoto článku je menší než 83 kB, tedy při zápisu 1 B trojcifernými čísly 000 až 255 jde o necelý čtvrtmilion cifer. Ale číslo $\xi = 0{,}\text{xxx}\ldots$, kde za nulou je uveden čtvrtmilion cifer, je jediné (leží mezi nulou a jedničkou) a z jeho přesné hodnoty lze tento článek přesně rekonstruovat. (Možností je zřejmě $10^{250\,000}$.) Totéž lze ovšem provést třeba s celou Britskou encyklopedií včetně všech obrázků; cifer bude více, ale výsledné číslo ξ' bude opět jediné a $0 \leq \xi' < 1$. A nakonec i se všemi knihami ve všech knihovnách světa dohromady...

S ohledem na tato fakta musíme chápat třeba hodnotu U_{prah}, s níž jsme začali. Proto zavádíme dohodou jistou hodnotu (*conventional true value*) jakožto hodnotu pro daný účel dostatečně vyhovující. Např. konvenční hodnota Avogadrovy konstanty leží v rámci $\mathrm{N_A} = 6{,}022\,141\,29(27) \times 10^{23}\,\mathrm{mol^{-1}}$; uvedená přesnost nám při současném stavu techniky postačuje. (Také se ovšem může stát, že Avogadrovu konstantu nově zadefinujeme prostě její číselnou hodnotou, a nikoli pomocí uhlíku jako dosud. Pak bude ovšem – definitoricky – zcela přesná, ale na přesnosti měření se stane závislý počet atomů ^{12}C ve 12 g uhlíku ^{12}C.)

3. Terminologie

3.1. V čem byl vlastně problém?

Přehlédneme-li velmi volný vyjadřovací styl a ignorování korelovaných vlivů, byla (snad) celá předchozí část jasná. To, co jasné není, bylo to, jak se vlastně správně řekne česky „true value“: máme říct, že je to hodnota skutečná? pravá? nebo snad „doopravdická“? (Poslední návrh podávám proto, abychom se všichni shodli – když nic jiného, tak že toto *doopravdicky* ne!)

Zatímco u většiny anglických termínů se (my Češi) shodneme, jak je budeme překládat, u několika málo termínů nebyly názory jednotné, resp. vykrystalovaly různé v různých oborech. Historické důvody jsou zřejmé: Jde o pojmy vyskytující se v různých, často navzájem značně odlehlých oblastech vědy i techniky. Přitom však ne všechny tyto pojmy se vyskytují v příslušných oborech stejně často; ne každý obor také potřebuje vedle základního tvaru také další odvozeniny. A ovšem – jako obvykle – každý návrh měl své výhody i nevýhody

Situace byla zejména neudržitelná tam, kde byl týž český termín (přesnost, správnost) používán v různých oblastech v jiných významech. Proto se ÚNMZ na základě rozboru rozhodl a zavedl terminologii společnou (viz tabulka níže), za vydatné součinnosti všech zúčastněných. A podařilo se to.

3.2. A jak si počínat příště?

Samotná změna termínů, resp. náhrada jednoho (historického) termínu novým je přirozený a nutný jev. Rozvoj techniky i vědy vede k odlišování a rozlišování okolností, které dříve nebyly známy nebo které dosud nebylo potřeba rozlišovat. Příkladem par excellence jsou názvy prvků vzácných zemin, konkrétně didym, který se při podrobnějším rozboru „rozpadl“ na dva další prvky praseodym a neodym.

Při tvorbě nových termínů se projevují různou měrou různé aspekty:

- Složitější konstrukce se v praxi zjednodušují, jde-li o pojmy často se vyskytující. Tak se z původně slangové elektronky a obrazovky staly řádné termíny namísto původní elektronové lampy a obrazové elektronky.
- Univerbálnost (jednoslovnost) je výhodná nejen stručností (kterou prosazuje hovorový jazyk), ale i snadnější tvorbou odvozenin, popř. složenin: Tak je výhodnější *zkrat* než *krátké spojení*: lze *zkratovat*, měřit *zkratový* proud atp.; *interakce* oproti *vzájemnému působení* umožňuje tvary *interagovat, interakční, interakčně* atp.
- Možnost odvozování dalších slov je též žádoucí. Zde často vadí konzervativnost uživatelů a (zbytečná) obava ze zatím nezvyklého slova. Tvar *energiový* od energie je stejně správný jako *sériový* od série, nezvyklé ť v *napěťový* má paralelu ve schodišťovém přepínači apod. Slova s koncovkou -nost jsou schopna dalších úprav a odvozování, viz *hmotnostní* defekt.
- Termín nemá být výklad. (Oč je snadnější *flip-flop* než *bistabilní klopný obvod*!) Zejména angličtina je velmi tolerantní vůči hovorovému stylu, náznakům, ba i žertu: jednotka *barn* pochází ze sportovního prostředí z úsloví *big as a barn*, dosl. velký jako stodola, zpravidla velký jako vrata do stodoly.
- Analogie s jinými jazyky je vítána, protože usnadní srozumitelnost (*polovodič* jako *Halbleiter* či *semiconductor*). Není však nutno vždy překládat (zvlášť doslova), má-li čeština jiný jazykový prostředek (centre of weight = těžiště).
- Systémovost v tvoření je jedním z hlavních pilířů terminologie. Ale zvyk může být neobyčejně tuhý. Zkuste prosadit, aby se říkalo *teplotoměr*, a ne teploměr, když měří teplotu, a ne teplo!
- Je velmi nebezpečné *zaměnit* význam slova; jednodušší a bezpečnější je užít slovo úplně jiné, třeba i internacionální.

Každý lingvista nás však přesvědčí o tom, že jazyk (a to i jazyk terminologie!) je živý organismus. Vyvíjí se, ať už chceme, či ne, a „mocenské zásahy“ jsou zpravidla neúspěšné, nemají-li vnitřní podporu uživatelů. Zde se ale bohužel nejspíš projevují i dozvuky totality: argumentace není vždy korektní, je neochota připustit argumenty druhé strany („porážka“ na jednom poli bývala porážkou se vším všudy), a tak se používají pseudoargumenty k obhajobě stůj co stůj, např.:

- takhle se to říká odjakživa (opravdu? A od kdy je to doloženo?)
- to se používá všude (anebo jen tam, kam já chodím?)
- tohle je směšné slovo (anebo jen pro mne nezvyklé?)

Každý terminolog ví, jak nevděčná je jeho práce: termín, který se ujme, je později samozřejmý a bezproblémový (třeba vodík, spin). Termín, který se neujme, je později směšný a trapný (ďasík = kobalt, vrt = spin). A přesto je nutné se terminologii věnovat, vytvářet ji, udržovat – stejně jako jazyku, kterým mluvíme a kterým se dorozumíváme s ostatními.

4. ZÁVĚR

Bylo vyloženo nové pojetí nejistot typu A a B a připomenuty základní pojmy z oblasti měření a jeho vyhodnocení. Záměrně volný sloh a jazyk umožnil sice vysvětlit věcnou podstatu problému, zároveň však poukázal na potřebu jistého sjednocení terminologie. Konkrétní vyřešený terminologický problém z problematiky představuje úspěšný výsledek.

Poděkování

Tento článek vznikl s podporou Ministerstva školství, mládeže a tělovýchovy (program INGO II, projekt LG13026).

Reference

[1] ČSN P ENV 13005, Pokyn pro vyjádření nejistoty měření, listopad 2005.

angl.:	EURACHEM	ČSN 01 0115	ČSN ISO 5725: 1994	ČSN ISO 3534: 1994	**nyní (TNI 01 0115, TNI 01 4109-3)**
accuracy	správnost	přesnost		přesnost; exaktnost	**přesnost**
bias	odchylka; vychýlení	chyba správnosti	strannost; vychýlení	strannost; vychýlení	**vychýlení**
precision	přesnost		shodnost; preciznost	shodnost; preciznost	**preciznost**
true (value)	skutečná, pravá (hodnota)	pravá (hodnota)		pravá; skutečná (hodnota)	**pravá (hodnota)**
trueness	pravdivost	správnost		správnost	**pravdivost**

[2] ČSN ISO 3534-1 a -2, Statistika – Slovník a značky, 2010.

[3] ČSN ISO 5725-1 až -6, Přesnost (správnost a shodnost) metod a výsledků měření.

[4] TNI 01 0115, Mezinárodní metrologický slovník – Základní a všeobecné pojmy a přidružené termíny (VIM).

[5] TNI 01 4109-3, Nejistota měření – Část 3: Pokyn pro vyjadřování nejistoty měření (GUM:1995) (Pokyn ISO/IEC 98-3).

[6] Metrologická terminologie v chemii. Plzák Z., EURACHEM, 2000.

[7] Fyzikální veličiny a jednotky SI, Obdržálek J. AlBra Úvaly, 2004 (Dodatek C – Norma a jazyk, Obdržálek J., Vlková V.).

[8] Závěrečná zpráva o plnění rozborového úkolu RU/0821/06, ÚNMZ.

A Brief Comparison of Loudness Evaluation Models

Stručné srovnání modelů hodnocení hlasitosti

Phil Webster and Ondřej Jiříček

CTU in Prague, Faculty of Electrical Engineering, Technická 2, 162 27 Praha 6
jiricek@fel.cvut.cz

In this paper, the concept of psychoacoustic loudness is discussed and in particular, objective methods for its evaluation. Two prominent methods are considered; the method proposed by Zwicker and the method proposed by Moore and Glasberg. Algorithms for each method are implemented in MATLAB in order to evaluate loudness for a variety of stimuli including pure tones, bandpass noise and industrial noise. Quantifiable differences are observed in the outputs of the two methods and the possible causes of these differences are discussed.

1. Introduction

Loudness is a subjective quantity corresponding to the intensity of a hearing sensation; it is related to, but not strictly depended on, sound pressure level. As a subjective quantity there is no precise method with which to quantify it for any given stimulus although the ability to do this would be hugely beneficial in any number of areas from television broadcasting to the design of manufacturing plants [1,2]. To this end a number of models have been proposed which evaluate the loudness of a sound based on parameters such as the frequency and type of sound field as well as masking and other temporal effects [3]. One of the earliest models was Steven's power law which gave a relationship between the magnitude of a sound stimulus and its perceived intensity although it has since been superseded by more comprehensive models. This paper aims to investigate briefly the two most predominant and relevant methods of objective loudness evaluation today.

2. Description of Loudness Models

The first of these methods is the method proposed by Zwicker [5, 6]. It takes as its input an audio signal and applies a filter approximating that which is applied to a sound by the outer and middle ear. The frequency spectrum of the resulting signal is then transformed into an excitation pattern representing neural activity that would result from this spectrum.

Next, specific loudness is calculated from excitation for each critical band using an equation based on excitation at threshold in quiet, excitation at the specific frequency and excitation at reference intnesity $I_0 = 10^{-12}$ W/m^2

$$N' = \alpha \left(\frac{E_{\mathrm{RQ}}}{E_0}\right)^{\beta} \left[0.5 + 0.5 \left(\frac{E}{E_{\mathrm{RQ}}}\right)^{\beta} - 1\right], \qquad (1)$$

where:
N' is specific loudness (sones/Bark),
E_{RQ} excitation at threshold in quiet,
E the excitation at the specific frequency,
E_0 the excitation at I_0,
α and β are constants.

This gives loudness as a function of frequency and so the area under this pattern can be determined to give a value for the overall loudness.

The second method was proposed by Moore and Glasberg [7]. The model is very similar to Zwicker's but with some different assumptions made and alternative procedures carried out. The predominant differences between Moore and Glasberg's and Zwicker's models are as follows [7]:

- Zwicker held differing opinions from Moore and Glasberg on the way in which sound is filtered by the outer and middle ear. While Zwicker assumed that below 2 kHz, transmission through the outer and middle ear was uniform for all frequencies, Moore and Glasberg postulated that below 1 kHz the transfer function is reflected in the shape of the absolute threshold curve.
- In Zwicker's method the excitation patterns are calculated through the use of critical band theorem. Moore and Glasberg prefer to use equivalent rectangular bandwidth for this purpose claiming that this method will not be affected by beat detection between the signal and a masker or the detection of combination products which are a result of interaction between the signal and a masker.
- Moore and Glasberg's model uses a relationship between excitation and loudness for partially masked sounds which was derived mathematically. In contrast Zwicker developed correction factors through subjective testing which brought his results in line with those of the tests.

3. Comparison of Models

Source codes were obtained for use in Matlab for both Zwicker's time-varying model and Moore and Glasberg's time-varying model online from Genesis; a French acoustics company [8]. Using these codes, the loudness of a series

 Přijato 20. května 2014, akceptováno 7. července 2014.

of artificial sounds created using an online .wav file generator [9] and other sounds (discussed later) was objectively determined. Evaluations of both instantaneous loudness and instantaneous loudness level were plotted and maximum values for each were recorded. The numerical results of these tests are shown in full in Table 1.

Sounds of the following types were applied to the loudness models for comparison.

- ***Tone Pairs*** – Pairs of tones of the same level which did not lie in the same critical band were tested first. Zwicker's model consistently evaluated higher loudness level by as much as 8 Phons for a tone combination of 100 Hz and 350 Hz. This may be due to the differeing assumptions made by Zwicker and Moore regarding perception of lower frequency sounds.

  The tendency of Zwicker's model to predict higher loudness levels was seen for pairs of tones which lay in adjacent critical bands, with low frequency tones yielding the greatest differences.

- ***Individual Tones*** – The corresponding individual tones which formed the first three tone pairs were also tested. Again, Zwicker's model estimated higher values of loudness level for all stimuli with the difference in level decreasing with an increase in frequency.

- ***Bandpass Noise*** – Generated bandpass noise was the next stimulus tested. In all cases of white bandpass noise, Moore and Glasberg's method was found to give higher values of loudness level with the difference ranging from 2.2 to 4.5 Phons. The biggest differences were found in the noise bands above 2 kHz. Pink Noise was also tested and although Moore and Glasberg's method once again predicted the higher loudness level, the difference was only 1.7 Phons. This also suggests that the difference in values generated by the models is due to the way in which they assess the higher frequency noise components.

- ***Industrial Noise*** – Three signals with noise generated by traffic, a bus and a car were tested to represent broadband industrial noise. Consistently across these three stimuli, Moore and Glasberg's method estimated the loudness level to be between 1 and 3 Phons greater than Zwicker's did.

- ***Speech*** – Two speech signals were applied as stimuli; both taken from a single radio podcast created by the British Broadcasting Company (BBC) featuring an interview between a male and a female. Two 15 second clips were created from this source with only male or female speech in each. In the case of these speech signals, Zwicker's model evalutes higher values of average loudness level for male and female speech than Moore and Glasberg's with the greatest difference being seen for male voice at 2.4 Phons.

  This test further indicates the tendency of Zwicker's model to estimate higher values for signals of low frequency. Interestingly, there is no notable difference between the loudness levels evaluated for male and female speech by Moore and Glasberg's model however a difference of 1.5 Phons is seen in the results of Zwicker's calculation.

- ***Music*** – Finally, two short pieces of music were tested. The first was a clip of a rock song which was found to give almost identical values of loudness level by both models with a difference of only 0.1 Phons, where Moore and Glasberg's method evaluated a slightly higher level. The second piece of music, a short clip of classical music, gave similar results with Moore and Glasberg's model giving a value for level just 0.3 Phons greater.

4. Conclusion

From the results of these tests it is clear that there are some non-trivial differences in the results of Zwicker's model and that of Moore and Glasberg's. In the evaluation of the loudness of tones of all types, Zwicker's method gives higher values. However in the evaluation of noise of all types it is Moore and Glasberg's method which yields the higher values. Speech is an example of a sound which it is desirable to evaluate the loudness and so too is industrial noise; it is hence interesting that Zwicker's model gives larger values for speech whilst Moore and Glasberg's method gives larger values for industrial noise.

The most similar results were generated for testing of audio signals in the form of clips of music. The complex nature of the signal which represents a clip of music more closely resembles the types of sounds which objective loudness evaluation will be useful for. It is unlikely that the models being developed today will be required to assess the loudness of pure tones or high level bandpass noise as these are not sounds that appear very often in our everyday lives.

The results of this testing are not enough in themselves to indicate which of the two methods is the best for evaluating the loudness of a sound; for this purpose, extensive psychoacoustic experimentation should be carried out to compare the results of each method to the loudness perceived by human subjects.

There is currently an ISO standard proposal for each of these methods, ISO/CD 532-1 for Zwicker's method [10] and ISO/CD 532-2 for Moore and Glasberg's [11], which are expected to be adopted as true standards in the near future. The introduction of an ISO standard would have a significant impact on the way in which loudness is evaluated across numerous industries. Potentially many years of previous testing with one method would need to be reevaluated in order to make a meaningful comparison with the results of a new standardised method.

Signal		Loudness (Sone)		Loudness Level (Phons)	
		Zwicker	M & G	Zwicker	M & G
Dual Tones	100 Hz 350 Hz	34.4	20.6	91.0	83.4
	500 Hz 800 Hz	35.9	32.0	91.6	89.4
	900 Hz 1250 Hz	33.7	30.7	90.8	88.8
Corresponding Single Tones	100 Hz	19.2	6.1	82.6	65.5
	350 Hz	24.1	16.5	85.9	80.4
	500 Hz	25.4	18.9	86.7	82.1
	800 Hz	24.4	22.7	86.1	84.8
	900 Hz	26.2	22.7	87.1	84.8
	1250 Hz	23.3	20.2	85.4	83.0
Dual Tones	75 Hz 130 Hz	21.3	8.8	84.1	71.1
(in adjacent critical bands)	430 Hz 510 Hz	31.5	24.1	89.8	85.6
	930 Hz 1250 Hz	32.9	30.0	90.4	88.5
Noise	White Noise 100–1000 Hz	28.4	34.8	88.3	90.5
	White Noise 500–1000 Hz	25.0	31.7	86.5	89.3
	White Noise 2–4 kHz	26.5	38.5	87.3	91.8
	White Noise 3–5 kHz	24.9	35.4	86.4	90.7
	White Noise 4–8 kHz	22.5	30.5	84.9	88.7
	Pink Noise	39.2	47.1	92.9	94.6
Industrial Noise	Traffic	20.6	26.8	83.7	86.9
	Bus	11.1	13.1	74.7	76.9
	Car	14.6	15.9	78.7	79.9
Speech	Male Voice	31.3	27.6	89.7	87.3
	Female Voice	28.3	27.4	88.2	87.2
Music	Rock	44.1	47.4	94.6	94.7
	Classical	30.7	32.7	89.4	89.7

Table 1: Table of results for comparison of Zwicker and Moore and Glasberg's methods

Acknowledgement

This work was supported by CTU project No. SGS13/193/OHK3/3T/13 „Monitoring and modeling methods in acoustics“.

References

[1] Florentine, M., Popper, A., Fay, R.: *Loudness*, Springer Science+Business Media, 2010.

[2] Munson, W. A.: The growth of auditory sensation, *J. Acoust. Soc. Am.* 19, 1947.

[3] Skovenborg, E., Nielsen, S.: Evaluation of Different Loudness Models with Music and Speech Material, Presented at *117th Audio Engineering Society Converntion*, 2004.

[4] Stevens, S., S.: On the psychophysical law, *Psychological Review* 64, 1957.

[5] Zwicker, E., Fastl, H.: *Psychoacoustics, Facts and Models*, Springer Verlag, Berlin, Germany, 1990.

[6] Zwicker, E.: Procedure for calculating the loudness of temporally variable sounds. *J. Acoust. Soc. Am.* 62, 675–682 (1977). Erratum: *J. Acoust. Soc. Am* 63, 283 (1978).

[7] Moore, B. C. J., Glasberg, B. R.: A Revision of Zwicker's Loudness Model, *Acta Acoustica* Vol. 82, 1996.

[8] GENESIS: Loudness Online. (2014). Retrieved from www.genesis-acoustics.com/en/loudness_online-32.html

[9] Pigeon, S.: Online Audio Frequency Signal Generator. (2014). Retrieved from www.wavtones.com

[10] First ISO/CD 532-1 *Acoustics – Method for calculating loudness Part 1: Zwicker method*, 2014

[11] First ISO/CD 532-2 *Acoustics – Method for calculating loudness Part 2: Moore-Glasberg method*, 2014

Akustické listy: ročník 20, číslo 2 červenec 2014 ISSN: 1212-4702
Vydavatel: Česká akustická společnost, Technická 2, 166 27 Praha 6 Vytisklo: Nakladatelství ČVUT, výroba
Počet stran: 12 Počet výtisků: 200
Redakční rada: M. Brothánek, O. Jiříček, J. Kozák, R. Čmejla, J. Volín
Jazyková úprava: R. Svobodová

Uzávěrka příštího čísla Akustických listů je 29. srpna 2014.