

UNIVERZITA KARLOVA V PRAZE
1. LÉKAŘSKÁ FAKULTA



Disertační práce

Modelování prostorového slyšení
Models of binaural hearing

AUTOR: ING. MAREK DRÁPAL
ŠKOLITEL: DOC. MUDR. RNDR. PETR MARŠÁLEK, PHD.
STUDIJNÍ PROGRAM: DOKTORSKÝ
STUDIJNÍ OBOR: LÉKAŘSKÁ BIOFYZIKA

PRAHA, 2011

Prohlášení

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem řádně uvedl a citoval všechny použité prameny a literaturu. Současně prohlašuji, že práce nebyla využita k získání jiného nebo stejného titulu.

Souhlasím s trvalým uložením elektronické verze mé práce v databázi systému meziuniverzitního projektu Theses.cz za účelem soustavné kontroly podobnosti kvalifikačních prací.

V Praze, 15. dubna 2011

Marek Drápal

Identifikační záznam

DRÁPAL, Marek. *Modelování prostorového slyšení. [Models of binaural hearing]*. Praha, 2011. 111 stran, 1 příloha. Disertační práce. Univerzita Karlova v Praze, 1. lékařská fakulta, Ústav patofyziologie. Školitel Maršálek, Petr.

Poděkování

Děkuji vedoucímu své práce Doc. MUDr. RNDr. Petrovi Maršálkovi, PhD., za cenné rady, náměty a připomínky, bez nichž by tato práce nikdy nevznikla a dále za mimořádně pečlivé a časově náročné odborné vedení této práce.

Abstrakt

Tato práce pojednává o modelu prostorového slyšení a srovnává ho s dalšími modely. Podle nejnovějších výsledků experimentů na savcích hraje inhibice velkou roli v určení časového posunu signálu mezi levým a pravým uchem. Tento časový posun je pro nižší frekvence klíčem k určení směru, odkud zvuk přichází. Výsledky experimentů vedou k závěru, že prostorové slyšení savců pracuje na jiném principu než u ptáků. Dnes existuje několik teoretických prací, které se snaží tento jev vysvětlit, ale naprostá většina z nich je založena na mimořádně přesném časování v inhibiční části obvodu. Tento předpoklad je však odtržen od dosavadní znalosti fyziologie. Na druhé straně, modely popsané v této práci jsou založeny na faktu, že každý neuron reaguje na podráždění s jistým náhodným zpožděním. Pokud je tato vlastnost uvážena v obvodu, ve kterém se objevuje inhibice, zpoždění a detektor koincidence, pak lze ukázat, že výstupní frekvence obvodu odpovídá azimutu binaurálního zvuku na vstupu a současně experimentálně získaným datům. Modely jsou podepřeny analytickými výpočty a numerickými simulacemi zahrnujícími i kochleární implantát.

Klíčová slova

prostorové slyšení, stochastický neuronový model, detektor koincidence, kochleární implantát

Anglický abstrakt

In this work is presented stochastic model of binaural hearing in context of another alternative models. According to latest experimental data on mammals, inhibition plays a role in interaural time difference recognition, which is a key for low frequency sound source localization. The outputs of experiments may lead to the conclusion that the binaural hearing works differently in mammals compared to birds. Nowadays there are a few theoretical works addressing this new phenomena, but all of them are relying on a very precise inhibition timing, which was never proved as physiologically valid. On the other hand, models described in this work are based on the fact, that every neuron has a random delay when reacting to an excitation. If this time jitter is taken into account and combined with inhibitory signal, delay in the neuronal circuit and coincidence detection, then the output firing rate corresponds to the azimuth of the sound source. In this work it is shown, that such a neuronal circuits are giving the same output results compared to experimental data. The models are supported by analytical computations and numerical simulations including simulation of cochlear implant.

Keywords

directional hearing, sound localization, stochastic neural model, interaural time delay, coincidence detection, cochlear implant

Obsah

1	Úvod	12
2	Fyzikální principy prostorového slyšení	13
2.1	Kódování azimutu	13
2.2	Kódování elevace	14
2.3	Přesnost určení zdroje	14
3	Morfologie a fyziologie sluchového ústrojí člověka	16
3.1	Zevní ucho	16
3.2	Střední ucho	16
3.3	Morfologie vnitřního ucha	17
3.4	Fyziologie vnitřního ucha	18
3.5	Nervové dráhy	19
4	Hypotézy určení azimutu	21
4.1	Jeffressova zpožďovací linka	21
4.2	Hypotézy reflektující roli inhibice	22
5	Binaurální model prostorového slyšení	24
5.1	Matematická analýza modelu	25
5.2	Numerická simulace	27
6	Kochleární implantát	30
6.1	Hluchota	30
6.2	Hlas	31
6.3	Výběr pacienta pro zavedení kochleárního implantátu	31
6.4	Konstrukce kochleárního implantátu	32
6.4.1	Mikrofon a signálový procesor	32
6.4.2	Přenosový systém	32
6.4.3	Sada elektrod	34
6.5	Zpracování signálu a kódovací strategie	35
6.5.1	Strategie CA a CIS	36
6.5.2	Strategie extrakce formantů	37
6.5.3	Spektrální strategie MPEAK	38

6.5.4	Spektrální strategie SMSP	39
6.5.5	Spektrální strategie SPEAK	39
6.5.6	Strategie kódování ACE	39
6.6	Porozumění řeči u pacientů s kochleárním implantátem . . .	40
6.7	Binaurální slyšení a kochleární implantáty	41
7	Klinické aspekty prostorového slyšení	43
7.1	Problém na koktejlovém večírku	43
8	Model prostorového slyšení s kochleárním implantátem	45
8.1	Numerická simulace	45
9	Diskuse	48
9.1	Oprávněnost použití rozdělení beta	49
9.2	Fyzikální versus psychofyzikální slyšení	50
9.3	Měření s použitím manekýna (fantomu)	51
10	Závěr	52
A	Články editované a prošlé recenzním řízením	59
A.1	Mechanisms of coincidence detection in the auditory brainstem: Examples	59
A.2	Stochastic model explains role of excitation and inhibition in binaural sound localization in mammals	69
A.3	Neural coding in sound localization pathway, roles of timing jitter and random delay	97
A.4	Stochastic model shows how cochlear implants process azimuth in real auditory space	102
B	Seznam příloh	111

Seznam zkratek

ABI	auditory brainstem implant	kmenový sluchový implantát
ACE	advanced combination encoder	zdokonalený kombinační kodér
AP	action potential	akční potenciál (nervový impuls)
AVCN	anterior ventral cochlear nucleus	antero-ventrální kochleární jádro
CA	compressed analog	komprimovaný analogový signál
CIS	continuous interleaved sampling	spojité prokládané vzorkování
DCN	dorsal cochlear nucleus	dorzální kochleární jádro
CD (DK)	coincidence detection	detekce koincidence
DSP	digital signal processing	digitální zpracování signálu
F0/F2	formant 0/ formant 1	strategie extrakce formantů
FFT	fast Fourier transform	rychlá Fourierova transformace
GJB2	gap junction protein beta 2	lidský gen kódující Connexin 26
ILD	interaural level difference	rozdíl intenzit mezi ušima
ITD	interaural time difference	časový rozdíl mezi ušima
KI	cochlear implant	kochleární implantát
LIAF	leaky-integrate and fire	neuron jako RC-obvod (RCC)
LNTB	lateral nucleus of trapezoid body	laterální jádro trapézového tělíska
LSO	lateral superior olive	laterální oliva superior
MNTB	medial nucleus of trapezoid body	mediální jádro trapézového tělíska
MPEAK	multi-peak	vícenásobný spektrální vrchol
MSO	medial superior olive	mediální oliva superior
PVCN	posterior ventral cochlear nucleus	postero-ventrální kochleární jádro
RCC	resistor - capacitor circuit	obvod s rezistancí a kapacitancí
SMSP	spectral maxima sound processor	procesor spektrálních maxim
SPEAK	spectral peak	spektrální vrchol

1 Úvod

Cílem této práce je nabídnout stručný úvod do problematiky prostorového slyšení savců a jeho modelování, na kterém pracuji během svého postgraduálního studia na Ústavu patologické fyziologie 1. lékařské fakulty Univerzity Karlovy.

V práci se nejprve zabývám genezí a fyzikálním popisem prostorového zvukového vjemu. Následuje popis morfologie a fyziologie sluchového ústrojí spolu s popisem hypotéz určení azimutu. Těžiště práce spočívá v popisu vlastního modelu prostorového slyšení, který otevírá nový pohled na možné fungování lokalizace zvuku u savců.

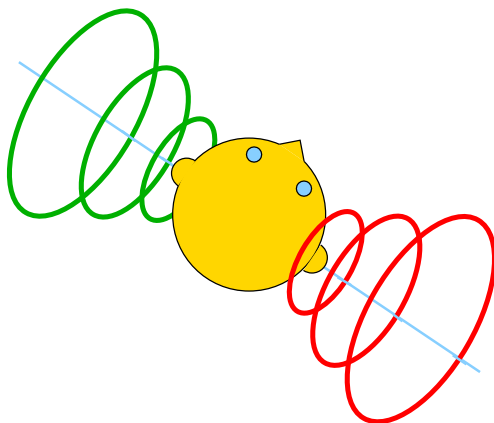
Pro nižší frekvence (u člověka do 1,5 kHz) je určení polohy založeno na detekci fázového posuvu zvukového signálu mezi levým a pravým uchem. Aby tato detekce mohla fungovat s přesností několika málo jednotek úhlových stupňů, je třeba rozlišit časové rozdíly mezi akčními potenciály neuronů v řádu desítek mikrosekund. S takovými nároky na přesnost nervového obvodu se neseťkáme nikde jinde v nervovém systému savců. Nepřekvapí proto, že dosud není známo, jak je tento posuv nervovou soustavou dekodován.

Na rozdíl například od genetiky je naše poznání v oblasti neurobiologie mnohem hůře uchopitelné a při popisu nervového systému nalézáme obrovské množství neprobádaných "bílých míst". Jedním z těchto míst je bezpochyby i oblast prostorového slyšení savců. Zatímco u ptáků se podařilo lokalizovat a popsat neuronální zapojení sloužící k určení zdroje zvuku, u savců na něco takového stále čekáme. Experimenty z konce devadesátých let minulého století ukázaly, že u savců hraje důležitou roli inhibice. Výsledky měření aktivity z těchto *in vivo* experimentů naznačují, že u savců bude zřejmě třeba hledat úplně jiný mechanismus než u ptáků.

Námi vytvořené modely, presentované v této práci, se na problematiku určení zdroje zvuku snaží dívat z pohledu pravděpodobnosti. Ukazuje se, že pokud zohledníme určité náhodné zpoždění vzruchů z vláskových buněk a jednotlivé neurony propojíme ve shodě se známou anatomíí, získáváme na výstupu obdobné výstupy jako experimentátoři. Získané modely jsou přitom velice robustní vzhledem k variaci vstupních parametrů.

2 Fyzikální principy prostorového slyšení

Z fyzikálního hlediska je zvuk mechanické vlnění šířící se v pružném médiu. Lidský sluch je schopen rozpoznat zvuk ve frekvenčním rozsahu přibližně od 20 Hz do 20 kHz. K jednoznačnému určení polohy zdroje zvuku v prostoru je třeba tří detektorů schopných rozlišit fázi a amplitudu vlnění. Při použití dvou detektorů není určení polohy fyzikálně jednoznačné, jak je vidět z obrázku 1. Z pohledu zpracování prostorové informace u člověka je výhodné rozložit úlohu na určení **elevace** a určení **azimutu**.

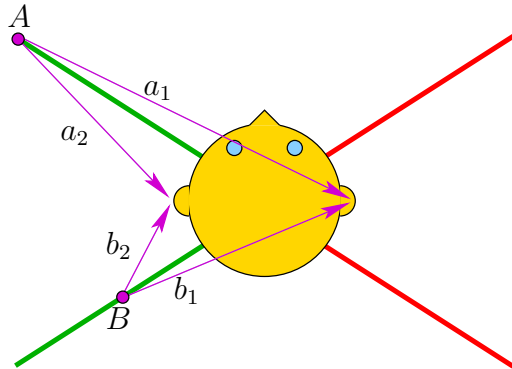


Obrázek 1: **Schematické znázornění neurčitosti v trojrozměrném prostoru.** Při stereofonním poslechu například ze sluchátek jsou všechny body rotačního hyperboloidu (anebo jeho asymptotického kužele = “cone of confusion”) stejně pravděpodobným místem umístění zdroje zvuku.

2.1 Kódování azimutu

Lokalizace zdroje zvuku v horizontální rovině s frekvencí nižší než 1,5 kHz je založena na určení časového posunu mezi signálem levého a pravého ucha – **ITD** (Interaural Time Difference). Pro vyšší frekvence již časový posun nekóduje azimut jednoznačně, podrobněji viz [Syka et al., 1981].

Jednomu časovému posunu odpovídají všechny polohy zdrojů zvuku na větvích hyperboly, které jsou osově souměrné. Osou souměrnosti je osa uší. Na obrázku 2 jsou pro jeden konkrétní úhel zakresleny dvě dvojice větví hyperboly – červené a zelené. Pro každé dva body A a B , ležící na příslušné



Obrázek 2: **Schematické znázornění neurčitosti v dvojrozměrném prostoru.** Vjem prostorového slyšení ze stejně barevných větví hyperboly je ekvivalentní. Například zvuk z bodu A může být vnímán jako zvuk z bodu B .

dvojici větví, a pro vzdálenosti a_1, a_2 a b_1, b_2 (viz obrázek 2) musí platit:

$$a_1 - a_2 = b_1 - b_2. \quad (1)$$

Pokud leží zdroj zvuku na ose uší, splývají tyto dvě větve hyperboly v jedinou polopřímku. Přesnost určení se pohybuje v řádu jednotek stupňů.

Lokalizace zdroje zvuku v horizontální rovině s frekvencí vyšší než 1,5 kHz je založena na určení rozdílu amplitudy mezi signálem levého a pravého ucha – **ILD** (Interaural Level Difference).

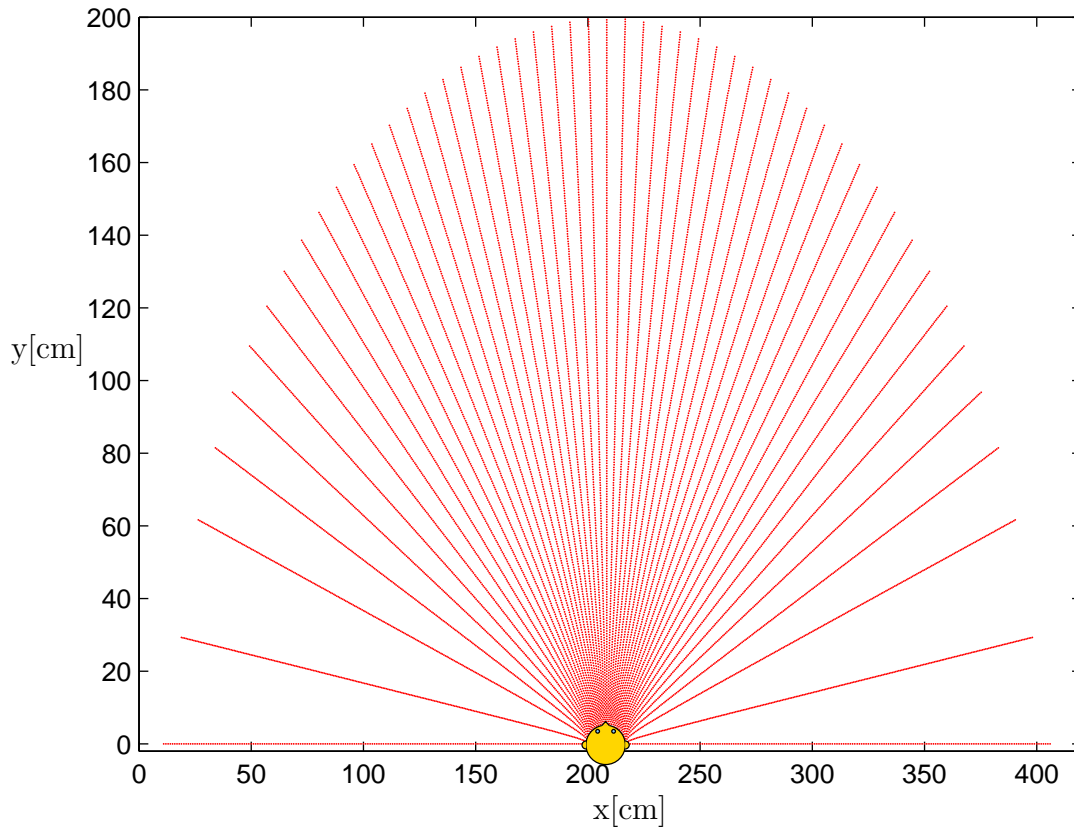
2.2 Kódování elevace

Pro určení polohy zdroje zvuku ve vertikální rovině je využito směrové charakteristiky ušního boltce. Tento mechanismus funguje správně až od frekvence 500 Hz. Jedná se o mechanismus, který pracuje mono-aurálně a přesnost určení se pohybuje v řádu desítek úhlových stupňů. Podrobnosti lze nalézt například v [Middlebrooks and Green, 1991].

2.3 Přesnost určení zdroje

Při určování posunutí dvou signálů uvažujeme vždy nějaký elementární rozlišovací krok. V případě digitálního signálu odpovídá tento krok převrácené hodnotě vzorkovací frekvence. Pro nahrávku v digitální CD kvalitě to jest se vzorkovací frekvencí 44,1 kHz docházíme k rozlišení 22,67 μ s. Na obrázku

3 jsou červenými úsečkami zobrazeny všechny možné zdroje zvuku, které je možno při časovém kroku $22,67 \mu\text{s}$ rozlišit. Zvuky vycházející z jiného bodu se budou jevit jako přicházející z nejbližšího bodu ležícího na červené úsečce.



Obrázek 3: **Rozlišovací kroky v horizontální rovině.** Červené úsečky ukazují všechna rozlišitelná umístění zdrojů zvuku v blízkosti hlavy při vzorkovací frekvenci 44,1 kHz. Rozměry jsou v centimetrech, interaurální vzdálenost činí 17 cm.

3 Morfologie a fyziologie sluchového ústrojí člověka

Sluchové ústrojí člověka dělíme na zevní, střední a vnitřní ucho. Zevní a střední ucho slouží ke vstupu a transformaci zvuku ze zevního vzdušného prostředí do tekutiny vnitřního ucha. Vnitřní ucho pak slouží k zakódování výšky a intenzity zvuku do nervových vzruchů. Podrobný popis fyziologie lze nalézt například v monografii [Syka et al., 1981], popis inhibičních spojení například v [Brand et al., 2002]. Rozdíly mezi slyšením ptáků a savců, včetně krátkého shrnutí evoluce ucha lze nalézt v [Grothe, 2003].

3.1 Zevní ucho

Zvuk přicházející z vnějšího zdroje je pro frekvence vyšší než 500 Hz směřován pomocí **ušního boltce**. Maximální směrový účinek má ušní boltec pro zvuk o frekvenci 5 kHz, přicházející ze směru v úhlu patnácti stupňů odkloněného od interaurální osy a ve frontální rovině. Zvuk o této frekvenci a stejné intenzitě přicházející z jiného směru do daného boltce je vnímán jako slabší. Ušní boltec také vytváří výrazný akustický stín pro zvuky přicházející zezadu a umožňuje tak lokalizovat zvuky v předozadní rovině. Na ušní boltec navazuje zevní zvukovod.

Zevní zvukovod je u člověka dlouhý asi 25 mm a má průměr kolem 7 mm, je ukončen bubínkem. Délka zvukovodu je parametrem určujícím rezonanční vlastnosti zevního zvukovodu. Experimentálně bylo změřeno, že k maximální rezonanci dochází mezi 3 kHz až 4 kHz, kdy dochází ke zvýšení intenzity zvuku asi o 12 dB.

3.2 Střední ucho

Úlohou **středního ucha** je přenos zvukové energie z plynného prostředí vnějšího zvukovodu do prostředí kapalného vnitřního ucha. Převod zvukové energie z jednoho prostředí do druhého je spojen se ztrátou energie, která musí být kompenzována činností středního ucha. O tom, jaká část energie se přenese, rozhoduje poměr impedancí obou prostředí. Účinnost přenosu

akustické energie lze vyjádřit takto:

$$\beta = \frac{4r}{(r + 1)^2}. \quad (2)$$

Koeficient β udává účinnost přenosu, r je poměr akustické impedance vzduchu a perilymfy. Akustická impedance vzduchu je asi $415 \text{ kg}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$, akustická impedance perilymfy je srovnatelná s impedancí mořské vody a činí asi $16,1 \times 10^5 \text{ kg}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$. Z výše uvedeného plyne, že r je rovno 4000. Dosazením do rovnice (2) získáváme tedy $\beta = 0,1\%$ to jest -30 dB . Tato ztráta energie musí být nějak hrazena. V současné době jsou známy tyto mechanismy hrazení ztrát:

1. Zvuk z relativně velké plochy **bubínku** (55 mm^2), který má tvar mělké nálevky (průměr $8 - 10 \text{ mm}$), je přenášen sluchovými kůstkami na oválné okénko, které je téměř 20-krát menší ($3,2 \text{ mm}^2$). Vzhledem k tomu, že není účinná celá plocha bubínku, dochází k zesílení asi o 23 dB .
2. Středoušní kůstky (**kladívko, kovadlinka, třmínek**) tvoří soustavu pák, jejíž zesílení je asi $2,5 \text{ dB}$.
3. Bubínek svým nestejným zakřivením vytváří speciální pákový systém, který rovněž zesiluje signál.

Vedle výše popsaných struktur středního ucha je třeba zmínit různé **vazy**, kterými jsou vázány sluchové kůstky a dva drobné svaly **musculus tensor tympani** a **musculus stapedii**, které mají shodnou funkci – kontrakcí snižovat přenos zvuku. Tímto je zajištěna ochrana sluchu před příliš intenzivním zvukem. Vyrovnání tlakových poměrů mezi vnějším okolím a středoušní dutinou je uskutečňováno při polknutí **faryngo-tympanickou tubou** (Eustachova trubice).

3.3 Morfologie vnitřního ucha

Úlohou vnitřního ucha je převod akustického vlnění perilymfy na nervové vzruchy, které kódují výšku a intenzitu zvuku. Základem vnitřního ucha je **kostěný hlemýžď**, který je dlouhý asi 35 mm a je stočen ve dva a půl

závitu. Trubice kostěného hlemýždě je rozdělena na dvě poschodí (**scala vestibuli** a **scala tympani**) **kochleární přepážkou**. Poschodí jsou oddělena obdobně jako jednosměrná schodiště rozhledny. Vchod na schodiště vzhůru vede z předsíně (**vestibulum**) s **oválným okénkem**, na které dosedá třmínek. Tato scala vestibuli prochází pod stropem kostěného hlemýždě do jeho vrcholu, kde přechází na dolní schodiště. Tato scala tympani je zakončena **okrouhlým okénkem**.

Scala vestibuli je oddělena Reissnerovou membránou od **scala media**, která je naplněna endolymfou. Od scala tympani je scala media oddělena bazilární membránou.

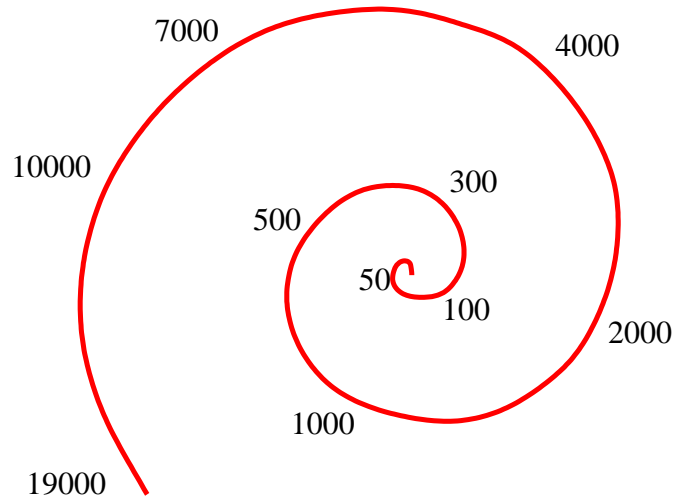
Z bazilární membrány vyrůstají různé nosné elementy Cortiho orgánu, které vytvářejí pravidelnou strukturu nesoucí **retikulární membránu**. Tato porézní struktura je protkána různými tunely a mikrotunely. V této struktuře rovněž vedou nervy **vláskových buněk**, které jsou umístěny na povrchu retikulární membrány. Vláškové buňky se dotýkají membrány tektoriální, která je zavěšena planparalelně nad retikulární.

Ohnutím stereocilií (vlásků) vláskové buňky na stranu bazálního tělíska dochází k depolarizaci, směrem opačným k hyperpolarizaci. Rozlišujeme vláskové buňky vnitřní, kterých je asi 3500 a jsou uspořádány do jedné řady, a vnější, kterých je 15-18 tisíc ve třech řadách. Od vnitřních vláskových buněk vede asi 95% všech aferentních (dostředivě vedoucích) nervových vláken, od vnějších zbylých 5% těchto vláken. Vnitřní vláskové buňky mají tedy úlohu cití, zatímco vnější vláskové buňky úlohu zpětnovazební, kdy jejich napnutím může být tlumen posun tektoriální membrány vůči membráně retikulární.

3.4 Fyziologie vnitřního ucha

Z bubínku je zvukový tlak převeden přes sluchové kůstky na oválné okénko. Akustické vibrace se odtud šíří nestlačitelnou nitroušní kapalinou. Tlak se až do frekvence 16 Hz vyrovnává plně přes okrouhlé okénko, které kmitá v protifázi. Při vyšších frekvencích se ale tlak nestačí vyrovnat, aniž by pohnul kochleární přepážkou. Dochází tedy k rozkmitání kochleární přepážky. Čím je kmitočet vyšší, tím se tlak vyrovná blíže ke třmínku (**tonotopie**, viz obrázek 4). K tonotopii kochleární přepážky přispívají i její mechanické

vlastnosti. Blízko třmínku je přepážka užší (0,05 mm) a postupně se rozšiřuje až na 0,5 mm. V hlemýždi tak dochází k základní frekvenční analýze.



Obrázek 4: **Tonotopické uspořádání hlemýždě.** Jednotlivé údaje na obrázku odpovídají maximálnímu prohnutí kochleární přepážky v závislosti na frekvenci zvuku. Frekvence je uvedena v Hz.

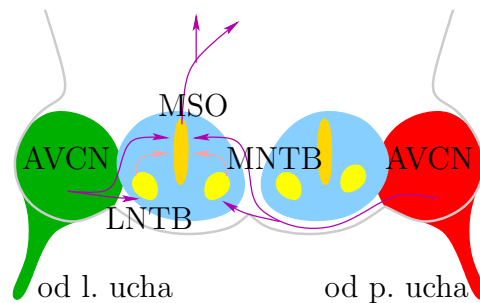
Rozkmitáním kochleární přepážky dojde k posunu mezi tektoriální a retikulární membránou. Posunutí zhruba odpovídá posunutí dvou plan-paralelních rovin přitlačených k povrchu koule. To má za následek to, že na jedné straně posunu dochází k hyperpolarizaci vnitřních vláskových buněk, na straně druhé k depolarizaci. Citlivost tohoto posunutí je nesmírná. Udává se [Syka et al., 1981], že při intenzitách blízkých sluchovému prahu člověka, dochází k registraci výchylky o velikosti 0,1 nm.

3.5 Nervové dráhy

Sluchový nerv vycházející z vnitřního ucha má u člověka asi 30 tisíc vláken. Obsahuje vedle vláken **aferečních** (dostředivě vedoucích) asi 500 vláken **eferčních** (odstředivě vedoucích), která sestupují z olivy superior a zakončují synapsí na vláskových buňkách Cortiho orgánu. Počet synapsí v jednotlivých úsecích hlemýždě není konstantní a dosahuje svého maxima (1400 vláken/mm) [Spoendlin and Schrott, 1989] zhruba v polovině hlemýždě.

Aferentní vlákna sluchového nervu vstupují do komplexu **kochleárního**

jádra, které je součástí prodloužené míchy. Vlákna sluchového nervu se uvnitř jádra rozdělují na část vzestupnou, která zásobuje **AVCN** (antero-ventrální kochleární jádro) a část sestupnou, která směřuje do **PVCN** (postero-ventrální kochleární jádro) a **DCN** (dorzální kochleární jádro). Z obou AVCN vedou nervová vlákna přímo do mediální olivy superior **MSO** a laterální olivy superior **LSO**. Do MSO, respektive do LSO vedou nervové dráhy rovněž přes **MNTB** a **LNTB** (mediální a laterální jádro trapézového tělíska, medial and lateral nucleus of trapezoid body), ve kterých je nervový vzruch transformován z excitačního na inhibiční. Výše popsaná nervová dráha je zachycena v obrázku 5.



Obrázek 5: Část inervace kochleárního jádra. Excitační vstupy (fialové) z antero-ventrálního kochleárního jádra (AVCN) a inhibiční (glycinergické) vstupy (růžové) z MNTB a LNTB vstupují do mediální olivy superior (MSO).

Jak je vidět, MSO a LSO jsou první místa, kde se setkávají vzruchy z levého a pravého ucha, mají proto velký význam pro prostorové slyšení. V MSO jsou zpracovávány nižší frekvence (do 1 až 1,5 kHz), v LSO pak frekvence vyšší. Z hlediska morfologického v nich však u savců nebyla objevena biologická struktura připomínající zpoždovací linku. Přesná neuronální spojení jsou zatím zčásti neznámá.

Z kochleárního jádra pokračují nervové dráhy přes několik dalších neuronů do sluchové kůry (**Heschlův závit**), která je ukryta v hloubi temporální oblasti mozkové kůry. Sluchová kůra na mnoha místech hraničí s místy nesluchovými, zvláště pak s asociačními oblastmi mozku.

4 Hypotézy určení azimutu

Tato práce se zabývá určováním azimutu pro nižší frekvence (do 1,5 kHz), kdy k vyhodnocování dochází na základě rozdílu fázového respektive časového rozdílu signálů (ITD) v mediální olivě superior. Intenzitním zpracováním zvuku v laterální olivě superior (pracujícím od 1,5 kHz) se zabývají například práce [Tollin, 2003], či [Park et al., 2004].

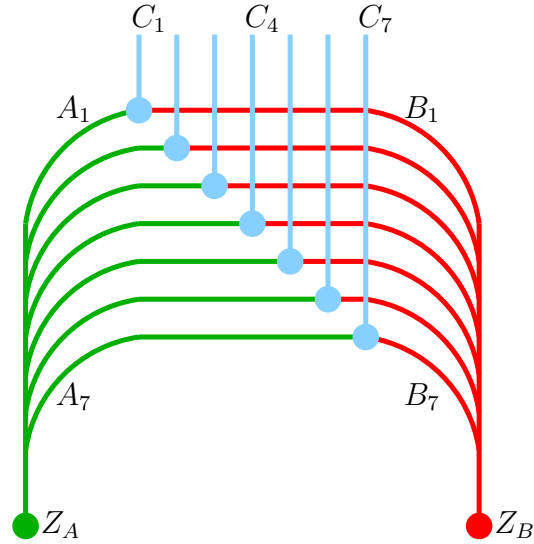
Od roku 1948 až do nedávné minulosti se předpokládalo, že určení azimutu na základě ITD je provedeno v takzvané Jeffressově zpoždovací lince [Jeffress, 1948], [Joris et al., 1998].

Nové práce [Batra et al., 1997], [Recanzone, 2000], [McAlpine et al., 2001], [Brand et al., 2002], [Grothe, 2003], [McAlpine and Grothe, 2003], nebo [Magnusson et al., 2005], či [Fitzpatrick et al., 2000] však ukazují, že chování neuronů v MSO neodpovídá Jeffressově teorii. V uvedených pracích se ukazuje, že silnou roli při určování azimutu hraje inhibice. Proti Jeffressově teorii rovněž svědčí to, že zpoždovací linka nebyla u savců dosud morfologicky prokázána. Vedle nových teorií operujících s inhibicí se objevují i teorie operující s tonotopicky posunutým zapojením binaurálních neuronů [Joris et al., 2006] či teorie pracující na základě nástupů (on-set) akčních potenciálů [Smith, 2001].

4.1 Jeffressova zpoždovací linka

Jeffressova zpoždovací linka je soustava vhodně zapojených detektorů koincidence (DK). Uvažujme nyní pro názornost zpoždovací linku složenou ze sedmi DK, tak jak je uvedeno na obrázku 6. Všechny DK v rámci zpoždovací linky mají své vstupy A_{1-7} připojeny na zdroj signálu Z_A a své vstupy B_{1-7} na zdroj signálu Z_B . Nechť Z_A představuje signál z levé strany, vstup Z_B signál ze strany pravé. Jediné, čím se jednotlivé DK od sebe liší, jsou vzdálenosti $|Z_A A_n|$, respektive $|Z_B B_n|$, pro vedení akčních potenciálů.

Pokud jsou poměry drah uspořádány tak, že $|Z_A A_1|$ je minimální a $|Z_B B_1|$ maximální u prvního DK a $|Z_A A_n|$ se lineárně zvětšuje, respektive $|Z_B B_n|$ zmenšuje směrem k poslednímu DK, kde naopak $|Z_A A_7|$ je maximální a $|Z_B B_7|$ minimální, pak získáváme axonální zpoždovací linku. Ta je zobrazena na obrázku 6. Pokud se objeví akční potenciál (AP) na Z_A a



Obrázek 6: **Schematické znázornění zpoždovací linky.** Z_A značí signál z levého ucha, Z_B signál z ucha pravého. A_{1-7} respektive B_{1-7} jsou vstupy do detektorů koincidence (DK). Výstupy z DK jsou označeny jako C_{1-7}

Z_B současně, dojde ke generování AP na C_4 , to je prostřední DK, kde $|Z_B B_4| = |Z_A A_4|$. Toto odpovídá vybuzení zvukem, který je přímo před námi, nebo přímo za námi (více o této neurčitosti v oddíle 2). Naopak, pokud zachytíme zvuk, který pochází přímo z naší levé strany, bude aktivován neuron první, protože $|Z_A A_1| \ll |Z_B B_1|$.

4.2 Hypotézy reflektující roli inhibice

Experiment vyřazující strychninem inhibici v MSO [Brand et al., 2002] ukázal její důležitost pro binaurální lokalizaci zdroje zvuku. Publikovaná vysvětlení operují s různě řízeným posunem výstupů detektorů koincidence, v práci [Brand et al., 2002] je tohoto posunu dosaženo velice rychlou inhibicí. Další in vivo experimenty (viz dále) však ukazují, že uvažovaná rychlost inhibice je až příliš vysoká.

V pracích [McAlpine and Grothe, 2003], [Magnusson et al., 2005] nebo také [Grothe, 2003] se operuje s rychlostmi nižšími, přičemž požadované naladění na konkrétní zpoždění se ustavuje až v několika prvních dnech po té, co začne pracovat sluch. Otázkou však zůstává jak k takovému naladění

dochází a jak je řízeno, respektive odkud se bere zpětná vazba správného určení lokalizace.

Zajímavou teorii založenou na anatomických zvláštностech buněk v MSO a podloženou detailním neuronálním modelem nabízí [Zhou et al., 2005]. Tyto hypotézy a experimentální měření však ale neodporují i jinému vysvětlení, které je popsáno v následujícím oddílu.

5 Binaurální model prostorového slyšení

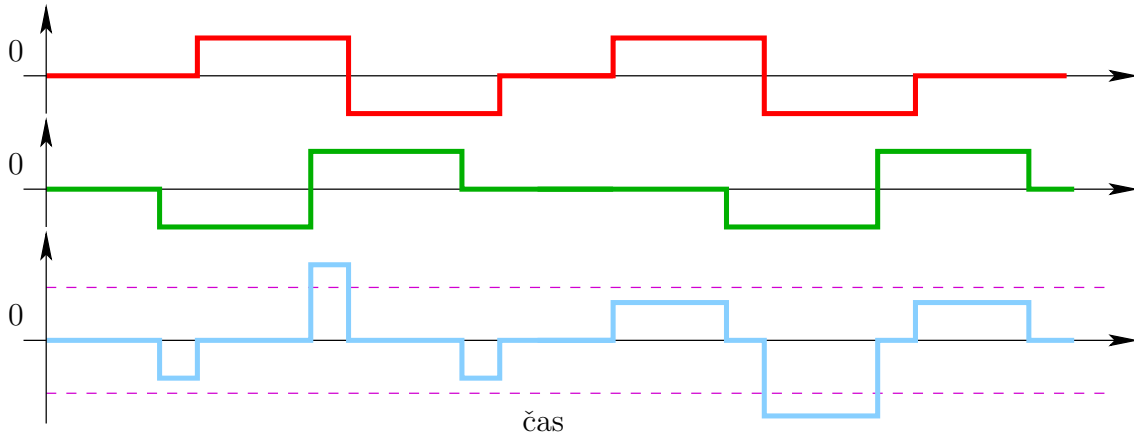
Náš model prostorového slyšení [Drapal and Marsalek, 2008] respektive [Drapal and Marsalek, 2011] se snaží nabídnout možné vysvětlení principu funkce binaurální lokalizace zdroje zvuku u člověka pro nižší frekvence. Inspirací a základem se staly publikace [Marsalek and Kofranek, 2004], [Marsalek and Lansky, 2005], [Marsalek and Kofranek, 2005] a v neposlední řadě [Marsalek and Drapal, 2007].

Model je koncipován tak, aby zjednodušeně nahradil biologické struktury sluchové dráhy od výstupu z vláskové buňky po neuron MSO kódující azimut (směr odkud zvuk vychází) frekvencí akčních potenciálů na výstupu. Otázka transformace zvukového signálu na sled akčních potenciálů je diskutována v oddíle 8. Schéma modelu je zachyceno na obrázku 8.

Pro nízké frekvence (zhruba do 1,5 kHz) je vzruch na výstupu z vláskových buněk přesně synchronizován s fází zvuku vstupujícího do ucha. Předpokládejme nyní, že šířící se vzruch z vláskové buňky je náhodně zpožděn například na další synapsi. Toto náhodné časové zpoždění modelujeme pomocí beta rozdělení inspirování experimentálním měřením populace neuronů [Marsalek and Drapal, 2007]. Není to však podmínkou, obdobného chování modelu lze dosáhnout i s normálním či rovnoměrným rozdělením zpoždění. Pro jednoduchost operujeme s pojmy levá a pravá strana, model je však symetrický, jak je vidět na obrázku 8 a tak jsou tyto pojmy záměnné.

Náhodně zpožděný signál z levého a pravého ucha se setká na detektoru koincidence (LDK/ PDK). Akční potenciál (AP) na výstupu detektorů koincidence se objeví pouze tehdy, setkají-li se signály z levého a pravého ucha v krátkém časovém okně, jak je vidět z obrázku 7.

AP z LDK jsou zpožděny. Toto zpoždění je z biologického pohledu realizováno již před vstupem do LDK v mediálním a laterálním jádře trapézovitěho tělesa, následně je signál invertován, dochází ke změně z vedení excitačního na inhibiční. Z pohledu modelu nezáleží na tom, kdy dojde ke změně z excitace na inhibici nebo kdy dojde ke zpoždění, proto je tento krok v modelu realizován až za detektorem koincidence. AP z LDK jsou vedeny na inhibiční synapsi (INH) v MSO. AP z PDK jsou přes excitační synapsi (EXC) sečteny (respektive odečteny od) s výstupy z INH a se spontánní aktivitou (ŠUM) na neuronu SOUČET.



Obrázek 7: **Koincidenční detektor využívající excitaci a inhibici.** Složením signálu z levého ucha (zelená křivka) a pravého ucha (červená křivka) získáváme signál zobrazený modrou barvou. Tento signál je prahován jak v excitační, tak v inhibiční větvi nervového obvodu (prahy jsou vyznačeny fialově).

Poloha zdroje zvuku je kódována frekvenčně, to znamená, že nízká frekvence akčních potenciálů ukazuje na vlevo umístěný zdroj zvuku, vysoká frekvence naopak na zvuk vycházející zprava. Výsledná aktivita neuronu SOUČET tedy odpovídá azimutu.

5.1 Matematická analýza modelu

Výstup z levých, respektive z pravých vláskových buněk je náhodně posunut v čase pomocí beta rozdělení

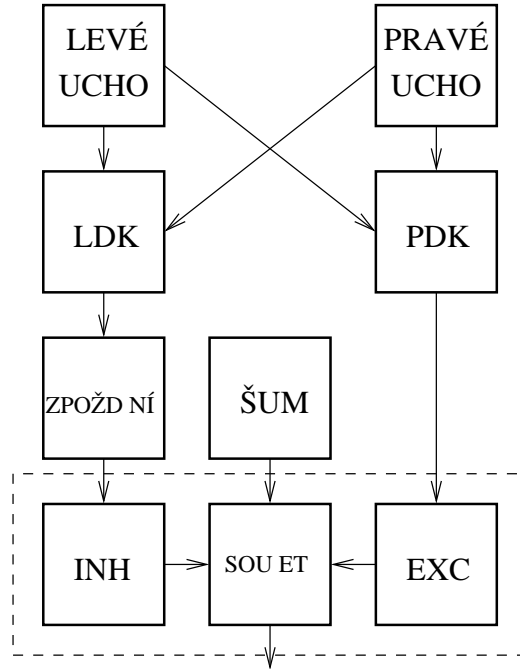
$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} \quad (3)$$

s parametry $\alpha_r = 2$ a $\beta_r = 4$ respektive $\alpha_l = 4$ a $\beta_l = 2$, $\Gamma()$ je gamma funkce. Označíme-li jednotkový skok v čase t jako $H(t)$ (Heavisidova funkce), pak můžeme signál pro levé ($l(t)$) respektive pravé ($r(t)$) ucho zapsat

$$l(t) = H(t) \cdot H(1-t) \cdot 20t(1-t)^3, \quad (4)$$

$$r(t) = H(t) \cdot H(1-t) \cdot 20(1-t)t^3. \quad (5)$$

Graficky je beta rozdělení zachyceno na obrázku 9. Pravděpodobnost aktivity na výstupu z detektorů koincidence pro všechna možná zpoždění



Obrázek 8: **Blokové schéma modelu.** Náhodně posunutý vstup z levých a pravých vláskových buněk (ucha) je veden do detektorů koincidence (LDK a PDK). Výstup z LDK je zpožděn a veden inhibiční větví (INH). Celkový výstup z modelu je SOUČTEM spontánní aktivity (označené jako ŠUM) s výstupy z INH a EXC (excitační větev).

mezi levým a pravým uchem je rovna konvoluci obou signálů. Podržíme-li se vyjádření beta rozdělení pomocí rovnic (4) a (5), můžeme využít vlastností Laplaceovy transformace a získat konvoluci zpětnou transformací součinu obou rovnic. Pro vyjádření této zpětné transformace jsme využili manipulátor symbolických výrazů v programu Matlab. Výsledný výraz zde nevyepisujeme, protože má více než dvacet členů.

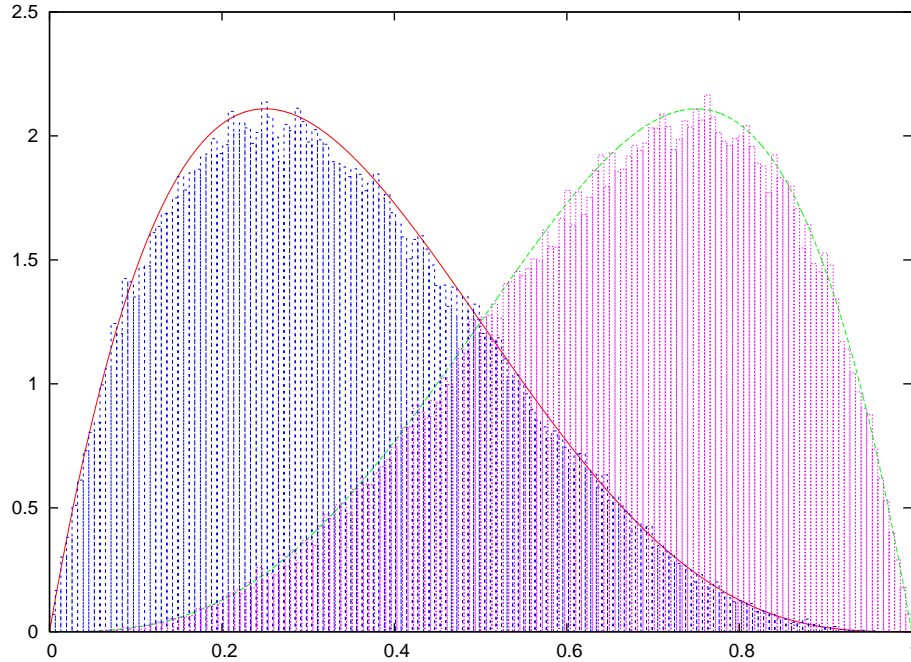
Ke stejnému výsledku dojdeme níže, v rovnici (8), při použití integrálního počtu. Opět uvažujme konvoluci dvou rozdělení

$$q(z) = \int_{-\infty}^{\infty} l(x)r(x-z)dx. \quad (6)$$

Dosazením mezí, které vyplývají z ohraničení hodnot polynomu jednotkovými skoky, získáme

$$q(z) = \int_{\max(z-1,0)}^{\min(z,1)} l(x)r(z-x)dx. \quad (7)$$

Úpravou rovnice (7) získáváme po zaokrouhlení pro interval od nuly do



Obrázek 9: **Grafy beta rozdělení.** Graf beta rozdělení s parametry (2,4) a (4,2). Křivky odpovídají analytickému výstupu, histogramy ukazují výstup ze simulace.

jedné:

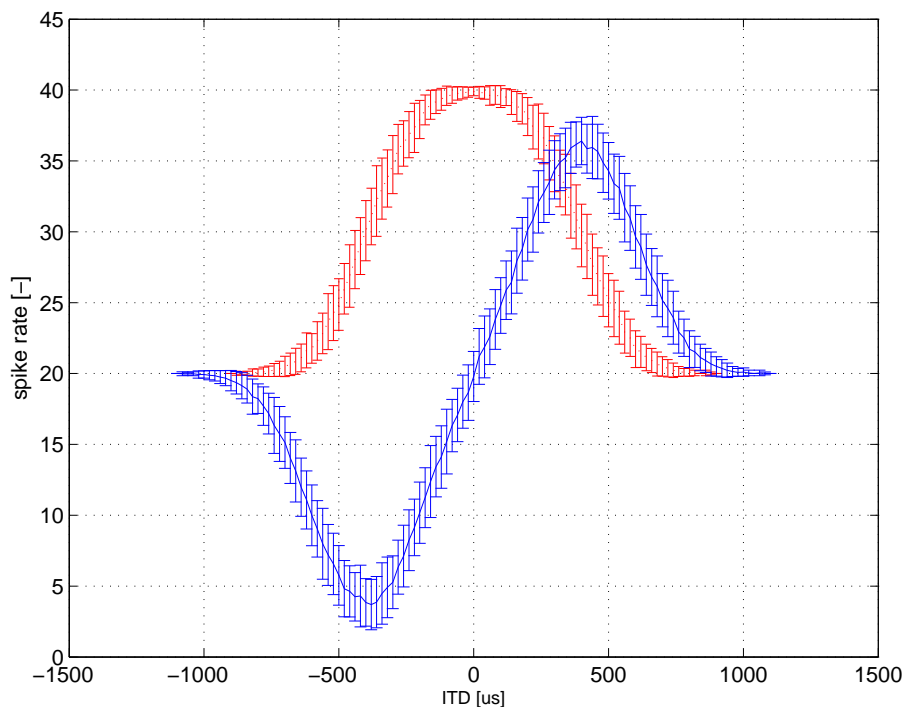
$$q(z) = q_9 z^9 + q_7 z^7 - q_4 z^4 + q_3 z^3 - q_2 z^2 + q_0, \quad (8)$$

kde $q_9 = -0,6349$, $q_7 = 2,8571$, $q_4 = -20$, $q_3 = 33,33$, $q_2 = -17,1429$, $q_0 = 1,5873$.

Funkce $q(z)$ je lichá, proto zde neuvádíme tvar pro interval od mínus jedné do nuly, pro podrobnosti lze nahlédnout do [Marsalek and Drapal, 2007]. Nicméně výsledek rovnice (8) a výsledek zpětné Laplaceovy transformace získaný výše je shodný. Prostým posunem výstupu z LDK v čase a součtem s PDK získáváme závislost aktivity na výstupu na zpoždění (ITD) na vstupu. Průběh tohoto signálu je zachycen na obrázku 10.

5.2 Numerická simulace

Pro ověření teoretických výsledků a porovnání s experimentálními daty jsme provedli numerickou simulaci. Jednotlivé komponenty modelu byly naprogramovány v Matlabu. Jako vstup diskrétní simulace byl vždy použit zvuk



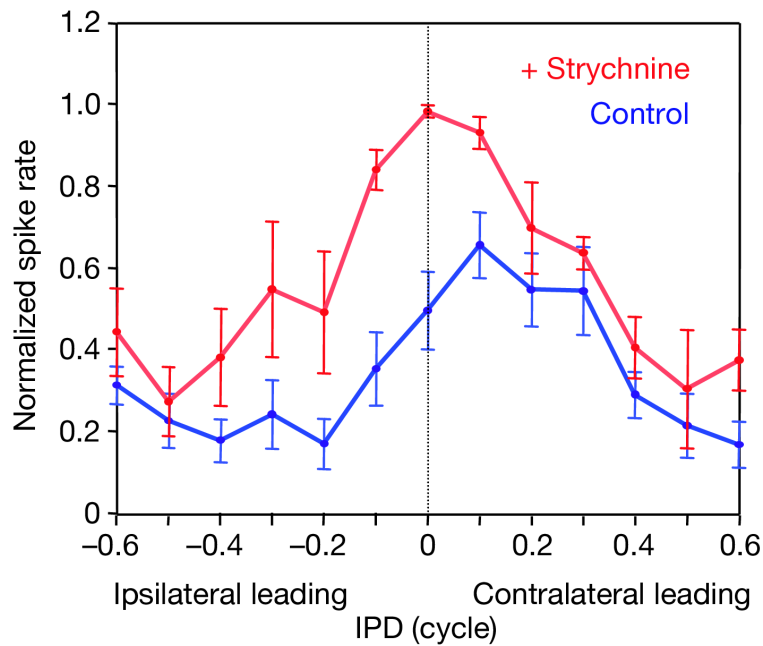
Obrázek 10: **Výstupní aktivity z modelu.** Červená křivka s maximem v $ITD=0$ je výstup z PDK. V případě vyřazení INH (to jest inhibičního vedení) odpovídá výstupu ze SOUČTU tato křivka. Modrá křivka s maximálním sklonem v $ITD = 0$ je výstup z plně funkčního modelu. Aktivita je zanesena v relativních jednotkách, časové zpoždění signálů (ITD) je v μs . Parametry simulace: je použito 100 paralelních vláken, frekvence zvuku je 1 kHz, trvání zvuku je 20 ms, maximální časové rozostření (jitter) na synapsi je $600 \mu s$, délka koincidenčního okna je $400 \mu s$ a velikost kroku ITD je $40 \mu s$.

o jedné frekvenci. Pro každé zpoždění bylo provedeno sto realizací, tomu odpovídá rozptyl na obrázku.

Výsledek simulace (výstup ze SOUČTU) je na obrázku 10 znázorněn modrou čarou. Křivka má největší sklon v blízkosti nulového ITD. Tento výsledek koresponduje s experimentálně naměřenými daty na pískomilech [Brand et al., 2002], která jsou zobrazena na obrázku 11.

Červená křivka na obrázku 10 odpovídá výstupu z detektoru koincidence PDK respektive vstupu EXC. Průběh na vstupu INH je obdobný, ale invertovaný a posunutý v čase o zpoždění ZPOŽDĚNÍ.

Provedeme-li na modelu obdobný experiment jako A. Brand s pískomily, vyřadíme-li inhibiční část, čehož bylo dosaženo v experimentu použitím glycinového antagonisty strychninu, získáme na výstupu výše popsanou



Obrázek 11: **Experimentální data.** Průměrovaná ITD funkce pro MSO neurony pískomila při normální funkci (modrá křivka) a po aplikaci strychninu (červená křivka). Převzato z [Brand et al., 2002].

červenou křivku, která odpovídá výstupu z PDK. Opět získáváme výstup obdobný experimentálním datům.

6 Kochleární implantát

Kochleární implantát (KI) je jedna z nejdokonalejších smyslových náhrad, která zpřístupňuje sluchové vjemy osobám jinak zcela neslyšícím. Principem KI je přímá elektrická stimulace sluchového nervu elektrodou.

Historie kochleárních implantátů se začíná psát v padesátých letech dvacátého století. Jednalo se o jednoelektrodové KI, které byly v začátcích velice poruchové, jen velice málo uživatelů bylo schopno rozumět slovům, spíše tento KI umožnil vnímat rytmus řeči. Přesto i jen to málo, co KI zprostředkoval, zlepšilo schopnost odezírání. Toto hmatatelné zlepšení však podpořilo další výzkum v této oblasti, který nebyl pouze technicky obtížný, ale narážel dokonce i na etické hranice. Přímá stimulace nervů elektrodami KI a jejich umístění do hlavy pacienta nebyla v tehdejší době pro řadu vědců a lékařů přijatelná.

Postupem času došlo ke zvýšení počtu elektrod, heuristickému předzpracování signálu, úpravě elektrod na základě zkušeností s vývojem kardios-timulátoru, miniaturizaci budící části a podobně. V následujících podkapitolách vycházím především z těchto zdrojů [Loizou, 1998], [Clark, 2003], [Swanson, 2001] a [Vondrášek, 2008].

6.1 Hluchota

Bilaterální hluchota či velice silná nedoslýchavost (více než 90 dB) je základním předpokladem pro implantaci KI. Zavedením elektrod KI do hlemýžďe dojde ke zničení případných posledních zbytků sluchu. Strategie unilaterální implantace KI bývá proto někdy obhajována z toho důvodu, že se očekává možnost nové léčby hluchoty, kterou bude moci provést v budoucnosti.

Hluchota může mít řadu značně variabilních příčin. Může být geneticky podmíněná (GJB2 mutace, dědičná otoskleróza), může vzniknout při nitroděložním vývoji (syfilis matky), předčasným porodem, vlivem infekce (meningitida, spalničky, AIDS, HIV, chlamydióza), extrémně silným hlukem či užíváním některých ototoxických léků (aminoglykozidy).

Z pohledu schopnosti naučení se rozumět řeči je důležité rozlišovat prelingvální a postlingvální hluchotu. Pacienti postlingválně hluší jsou často schopni slyšet okamžitě po aktivaci KI, na rozdíl od prelingválně hluchých,

kteří se musí nejprve řeč naučit a to ve ztížených podmínkách omezených spektrem zvuků zprostředkovaných KI.

Schopnost mozku naučit se využívat zvukovou informaci je časově omezena a trvá maximálně do adolescence. Nejlepších výsledků dosahují prelingválně hluché děti, kterým byl KI implantován časně (maximálně do čtyř let). Pokud je implantace provedena až kolem devátého roku či později, lze očekávat, že zpracování zvukové informace nebude úplně funkční.

Implantace prelingválně hluchých v dospělosti nemá díky omezenému trvání vývojové plasticity mozku již žádný efekt. Sluchová centra jsou obsazena jinými funkcemi a pacienti se slyšet nenaučí.

6.2 Hlas

Hlavní funkcí KI je zprostředkovat lidské hlasové projevy. Maximum informace v řeči leží ve frekvenčním pásmu mezi 300 Hz - 3,4 kHz (standard mezinárodní telekomunikační unie).

Aby přenos informace byl maximální, musí KI zachovávat charakteristické znaky mluvené řeči. Jde především o základní frekvenci generovanou hlasivkami společně s příslušnými rezonančními maximy tzv. formanty, které vznikají rezonancemi v ústní dutině. KI musí tedy provádět frekvenční analýzu zvuku a příslušná maxima (formanty) zachovávat.

6.3 Výběr pacienta pro zavedení kochleárního implantátu

Aby pacient po implantaci KI s velikou pravděpodobností slyšel, je třeba, aby byly splněny následující podmínky:

- Sluchový nerv musí být neporušený a správně pracující. Elektrody KI totiž proudově stimulují tento nerv. Nicméně i osoby s porušeným sluchovým nervem mohou využívat zařízení pracující na obdobném principu jako KI, jedná se například o ABI (auditory brainstem implants).
- Pacient musí být schopen podstoupit operaci, která probíhá v celkové anestezii a trvá jeden a půl až osm hodin.
- Sociální zázemí pacientů, především prelingválně hluchých, musí být

dostatečně kvalitní, aby mohl probíhat nácvik porozumění řeči a mluvení.

6.4 Konstrukce kochleárního implantátu

Ve světě existuje několik komerčních výrobců KI, například Med-El (Rakousko), Cochlear (Austrálie), Advanced Bionics (USA) nebo Neurelec (Francie). Nicméně každý KI se skládá z těchto částí:

- **mikrofon** – slouží k převodu zvuku (tlaku) na elektrický signál
- **digitální signálový procesor (DSP)** – mikroprocesor, který analyzuje zvuk a konvertuje jej na elektrické signály
- **přenosový systém**, který přenáší elektrické signály z DSP do vlastního implantátu; toto se děje většinou indukční cestou
- **sada elektrod** implantovaných do kochley, které proudově budí sluchový nerv

Toto uspořádání je zobrazeno na obrázku 12.

6.4.1 Mikrofon a signálový procesor

Mikrofon společně s digitálním signálovým procesorem a napájecí baterií bývá většinou umístěn ve společném pouzdře. Toto pouzdro bývá buď miniaturního rázu a bývá zavěšeno za uchem, podobně jako naslouchadla. Druhá varianta pouzdra bývá větší a je počítáno s tím, že bude nošena přichycena někde na těle či v oblečení. Výhodou větší varianty je možnost umístění kapacitnějších baterií (což vede k prodloužení intervalu údržby) a zvýšení robustnosti zařízení.

Hlavní úlohou signálového procesoru je rozklad zvukového signálu na jednotlivé frekvenční složky, jejich klasifikace, výběr a převod na proudové hodnoty jednotlivých elektrod.

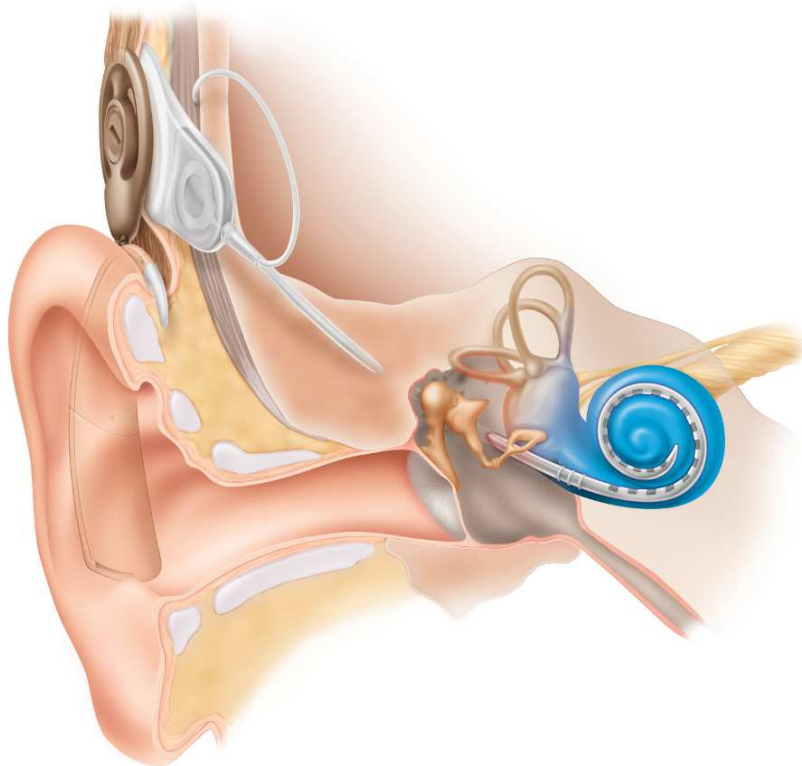
6.4.2 Přenosový systém

Přenos energie a informace do implantátu může být buď transkutánní nebo perkutánní. Naprostá většina systémů používá k přenosu transkutánní in-

dukční systém, kdy vysílací část vytváří elektromagnetické pole, které indukuje v cívce v přijímací části elektrický proud, který je použit na napájení vlastního implantátu. Systém pracuje většinou na dvou frekvencích, kdy jedna slouží k přenosu energie a druhá k přenosu informace.

Výhodou tohoto uspořádání je úplné uzavření implantátu do těla, kdy přijímací cívka je překryta kůží a je tak zabráněno infekci. V přijímací části je umístěn permanentní magnet, který přidržuje vysílací cívku. Určitou nevýhodou je následná nemožnost podstoupit vyšetření hlavy magnetickou rezonancí. Nicméně novější KI tento nedostatek odstraňují použitím vyjímatelného magnetu, po jehož odstranění je možné podstoupit magnetickou rezonanci do síly 1,5 T.

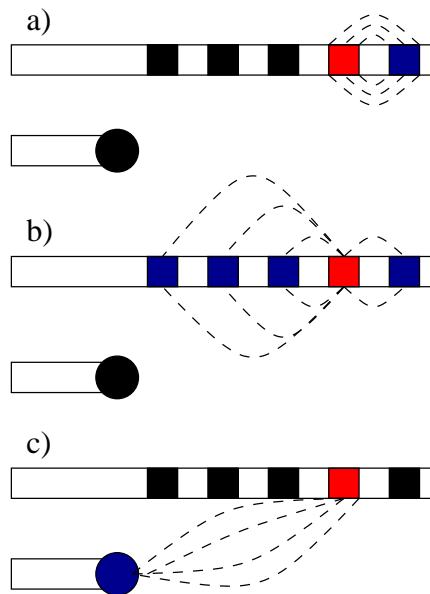
Perkutánní systém je jinou výjimkou, sestává pouze ze zásuvky a elektrod a je proto vhodný k výzkumným účelům, kdy je možné elektrody ovládat přímo, bez jakýchkoliv omezení.



Obrázek 12: **Kochleární implantát.** Umístění KI v těle. [Převzato z volně šiřitelného propagačního materiálu od National Institutes of Health].

6.4.3 Sada elektrod

První kochleární implantáty byly pouze jednoelektrodové [Loizou, 1998]. Jedna elektroda nemohla samozřejmě stimulovat sluchový nerv tak, aby pacient rozlišoval různě vysoké tóny a tak se postupně vyvinuly systémy víceelektrodové. Modernější systémy využívají dvanáct až dvacet čtyři elektrod, jejichž vzdálenosti jsou zhruba od 2,5 do půl milimetru. Další zvýšení počtu je omezeno počtem zdravých neuronů, které mohou být excitovány a především rozptylem elektrické stimulace. Vzhledem k tomu, že šíření proudu od zdroje je téměř symetrické dochází k podráždění více neuronů najednou. Oblast excitace je možné zmenšit bipolární stimulací. Zatímco při unipolární stimulaci teče proud z jedné elektrody směrem k elektrodě referenční, při bipolární stimulaci teče proud většinou mezi sousedícími nebo blízkce sousedícími elektrodami. Jednotlivé strategie jsou zobrazeny na obrázku 13. Přesto nelze předem určit, která strategie bude konkrétnímu pacientovi vyhovovat více.



Obrázek 13: **Různé stimulační módy elektrod.** a) bipolární stimulace, b) stimulace se společnou zemí, c) monopolární stimulace; v horní části je vždy zobrazeno pole intrakochleárních elektrod, v dolní části kuličková referenční elektroda, která někdy bývá nahrazena kovovým pouzdem stimulatoru. Upraveno podle [Loizou, 1998].

Elektrody stimulují sluchový nerv proudově. Jedná se o proud v řádu

desítek mikroampérů (μA). Čím je proud silnější, tím je aktivováno více vláken sluchového nervu a pacient vnímá zvuk jako hlasitější. Na základě polohy elektrody v hlemýždi pak pacient vnímá rozdílnou výšku tónu – viz obrázek 4. Některé komerční implementace využívají takzvané virtuální elektrody, které vznikají buzením sousedních elektrod nižším proudem, který dosáhne díky superpozici nejvyšší hodnoty mezi elektrodami a vyvolá tak zvukový vjem o frekvenci f_3

$$f_3 = \frac{f_1 + f_2}{2} \quad (9)$$

kde f_1 a f_2 je frekvence zvukového vjemu z první a druhé elektrody.

Stimulace může být buď analogová, kdy dochází k stimulaci spojitým signálem, který odpovídá zvuku vyfiltrovanému frekvenční propustí (tonotopicky odpovídající poloze elektrody v hlemýždi), nebo pulsní, kdy informace je přenášena řadou impulsů o vysoké frekvenci, tak aby jejich složením vznikla podobná frekvenční obálka jako na vstupu. Výhodou pulsní stimulace je možnost sériového buzení jednotlivých elektrod a tím pádem zaizení interakcí mezi kanály. Tato serializace však může mít i negativní dopady např. při určování časového rozdílu při binaurálním slyšení bilaterálně implantovaných pacientů.

Stimulační proud musí mít nulovou střední hodnotu, aby nedocházelo k elektrolýze tkáně a nesmí být příliš veliký, aby nedošlo i z dlouhodobého hlediska ke korozi elektrod či poškození tkáně. Elektrody jsou umísťovány nejčastěji do scala tympani, protože jsou tak nejblíže zakončením sluchového nervu a jednoduše sledují tonotopické uspořádání. Elektrody jsou zavedeny asi 22–30 milimetrů hluboko (maximální průměrná hloubka hlemýždě je 35 mm).

6.5 Zpracování signálu a kódovací strategie

Zpracování signálu pro víceelektrodové KI můžeme rozdělit do dvou základních kategorií: strategie zachovávající průběh signálu a strategie založené na extrakci spektrálních vlastností signálu (na příklad zachování formantů).



Obrázek 14: **Kochleární implantát.** Přijímací cívka a sada elektrod kochleárního implantátu SONATA společnosti MED-EL. [Převzato z tiskové zprávy společnosti MED-EL.]

6.5.1 Strategie CA a CIS

Mezi strategie zachovávající průběh patří CA (Compressed–Analog) strategie, která pouze filtruje signál pásmovými propustmi (střední frekvence odpovídá umístění elektrody v kochleě) a upravuje zesílení [Loizou, 1998].

Ukázalo se, že při použití CA strategie dochází k ovlivnění mezi jednotlivými kanály (rušení). Proto byla vyvinuta CIS (Continuous Interleaved Sampling) strategie, která stimuluje impulsem vždy jen jednu elektrodu a cyklicky elektrody střídá. U CIS strategie, při které dosahují pacienti vyššího skóre ve všech testech porozumění, můžeme měnit tyto parametry:

- **frekvenci a šířku impulsů** – nejlepšího porozumění dosahují pacienti s impulsy o frekvenci 800 až 2500 Hz a délce impulsu kolem 30 μs .
- **pořadí stimulace elektrod** – může být tonotopické to jest stimulace elektrody číslo 1, 2, 3, 4, 5, ... nebo nějakým způsobem zamíchané: 6, 3, 5, 2, 4, ... Opět platí, že u každého pacienta vyhovuje jiná strategie.
- **kompresní funkce obálky vstupního signálu** – převádí akustickou amplitudu zvuku na amplitudu jednotlivých impulsů.

Funkce komprese je nezbytná vzhledem k omezené škále amplitud elektrických impulsů oproti vstupnímu akustickému signálu. V praxi se jedná zhruba o poměr 30 dB ku 5 dB. Ke kompresi bývá většinou použita logaritmická funkce (obrázek 15), jejíž parametry je možné ladit individuálně dle potřeb pacienta. Druhou využívanou komprimační funkcí je funkce mocninná, která má výhodu v tom, že tvar a strmost je možné měnit jediným parametrem, kterým je exponent p . Je-li výstupní proud roven

$$I = A \log(x) + B \quad (10)$$

kde $p > 0$ a

$$A = \frac{MCL - THR}{x_{\max}^p - x_{\min}^p} \quad (11)$$

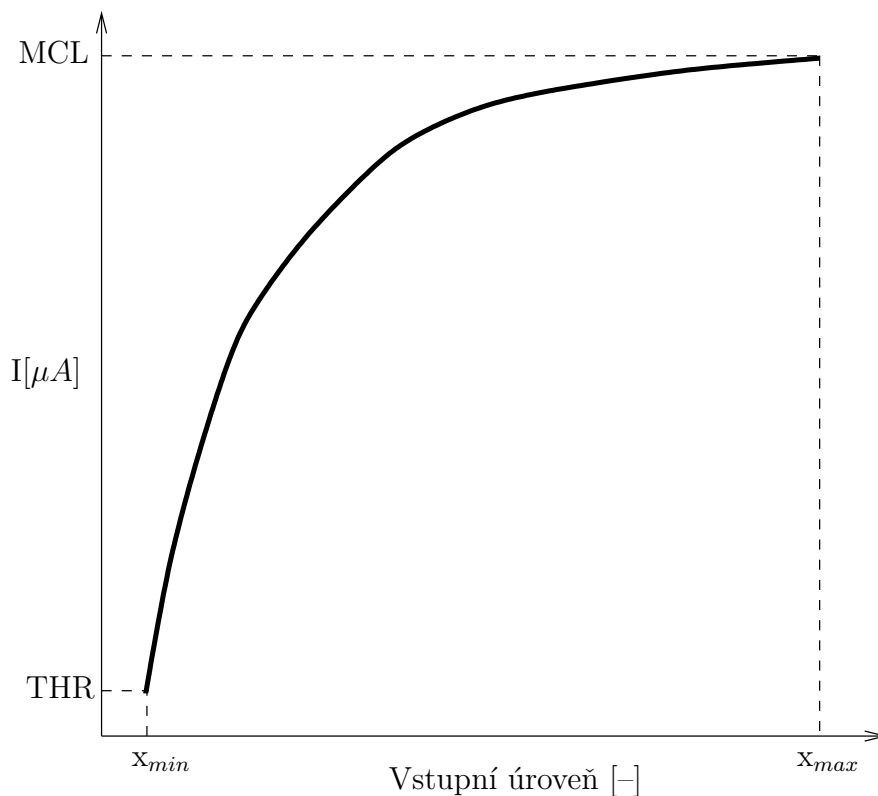
$$B = THR - Ax_{\min}^p \quad (12)$$

x_{\min} resp. x_{\max} je minimální resp. maximální vstupní úroveň, THR je prahová úroveň, MCL je maximální ještě příjemná velikost budícího proudu.

6.5.2 Strategie extrakce formantů

Strategie F0/F2 byla první extrakční strategie vyvinutá pro KI Nucleus na začátku osmdesátých let. Jak napovídá název pracovala na principu extrakce fundamentální frekvence formantu (F0) a druhého formantu (F2). Přičemž frekvence F0 byla hledána v signálu filtrovaném dolní propustí (270 Hz), F2 obdobně s použitím pásmové propusti (1–4 kHz). Tato strategie měla u některých pacientů unikátně dobré výsledky [Loizou, 1998].

V polovině osmdesátých let byla strategie F0/F2 rozšířena o první formant F1, který je velice důležitý pro srozumitelnost i pro normálně slyšící osoby (označení F0/F1/F2). Na základě frekvence F1 a F2 jsou buzeny příslušné elektrody, které odpovídají dané frekvenci. V KI Nucleus WSP byly použity bifázické impulsy, kdy každá fáze trvala $200 \mu s$ a jednotlivé impulsy byly odděleny $800 \mu s$ mezerou, aby nedocházelo k interakcím mezi kanály. Jako ve strategii F0/F2 byly elektrody stimulovány frekvencí F0 impulsů za sekundu u znělých hlásek a zhruba 100Hz u neznělých. Není překvapivé, že porozumění při strategii F0/F1/F2 zlepšilo srozumitelnost ve slovním testu o více než 30% ve srovnání s F0/F2.



Obrázek 15: **Logaritmická komprese** používaná v různých kódovacích strategiích KI. Kompresní funkce mapuje vstupní akustický signál z rozsahu $[x_{\min}, x_{\max}]$ na výstupní elektrický signál do rozsahu $[\text{THR}, \text{MCL}]$. x_{\min} resp. x_{\max} je minimální respektive maximální vstupní úroveň, THR je prahová úroveň, MCL je maximální ještě příjemná velikost budícího proudu.

6.5.3 Spektrální strategie MPEAK

Strategie MPEAK (zkratka pro Multiple PEAK) vznikla evolucí strategie F0/F1/F2, oproti které došlo k mírnému rozšíření frekvenčního okna pro formant F2 a z tohoto pásma pak bylo vybráno více spektrálních maxim ve třech podpásmech pevně svázaných se třemi elektrodami. Toto rozšíření bylo motivováno jednak snahou o zlepšení reprezentace formantu F2 a na druhé straně snahou o přenos informace o vyšších frekvencích, které jsou důležité pro souhlásky. Ukázalo se, že rozšíření přineslo zlepšení ve slovním testu asi o 30% ve srovnání s F0/F1/F2. Ačkoliv se strategie MPEAK ukázala jako efektivní v extrakci důležité informace z mluveného slova, má jedno hlavní omezení, stejně jako F0/F2 a F0/F1/F2 strategie. Tímto

omezením je chybovost algoritmu extrakce formantů zejména v zašuměném signálu.

6.5.4 Spektrální strategie SMSP

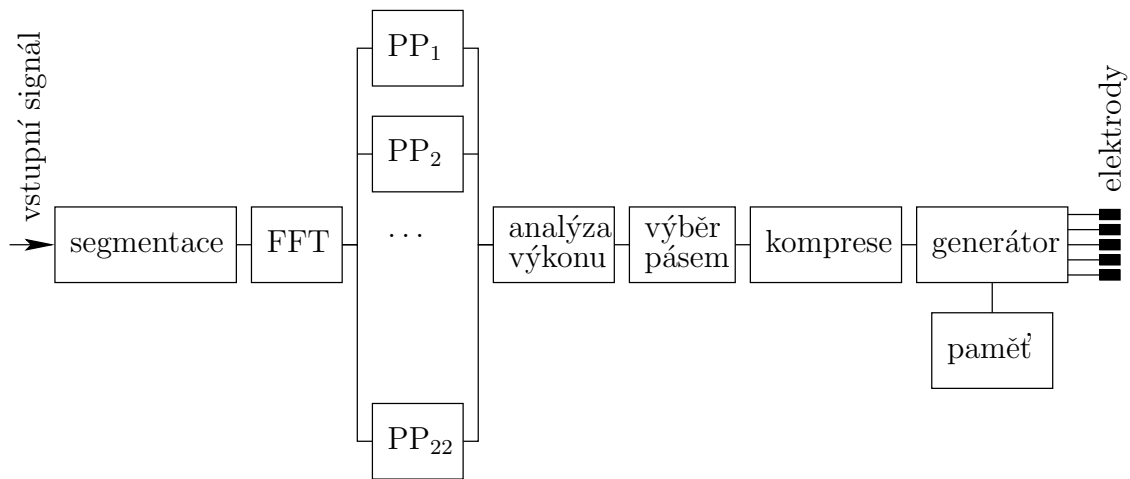
Na začátku devadesátých let byl vyvinut nový procesor SMSP (Spectral Maxima Sound Processor), který opustil cestu extrakce formantů, kvůli problému popsanému výše a vydal se cestou detekce spektrálních maxim. Vstupní signál byl pásmovými propustmi rozdělen na 16 nezávislých pásem s centrálními frekvencemi od 250 Hz do 5,4 kHz a v těchto pásmech byla nalezena spektrální maxima. Z těchto 16 maxim je v rámci 4ms okna vybráno 6 s největší amplitudou. Po logaritmické kompresi (viz 6.5.1) je signál přiveden do 6 odpovídajících elektrod jako bifázické impulsy o frekvenci 250 Hz. V klinických testech SMSP strategie byla opět lepší ve srovnání s MPEAK.

6.5.5 Spektrální strategie SPEAK

Strategie SPEAK (Spectral PEAK), která je využita i v dále uvedeném modelu prostorového slyšení, je v mnohém podobná SMSP strategii. Oproti SMSP používá 20 pásem od 250 Hz do 10 kHz, ze kterých procesor dynamicky vybírá 5–10 v závislosti na vstupním signálu. Buzení elektrod probíhá na frekvenci 180 až 300 Hz v závislosti na individuálním rozhodnutí pacienta a množství buzených elektrod. Při vyšším počtu aktivovaných elektrod je frekvence buzení nižší. Dynamický výběr maxim oproti pevně danému časovému oknu 4 ms u SMSP umožňuje zachování nejen spektrálních vlastností zvuku, ale i časových, což je mimořádně důležité pro binaurální slyšení. Ve srovnání se strategií MPEAK je SPEAK úspěšnější ve všech druzích testů, především pak v těch, které jsou zatíženy šumem.

6.5.6 Strategie kódování ACE

Strategie ACE (Advanced Combination Encoder) je podobná strategii SPEAK, kombinuje však vyšší počet pásem (22) a vyšší stimulační frekvenci. Podrobný popis algoritmu je uveden v popisu obrázku 16.



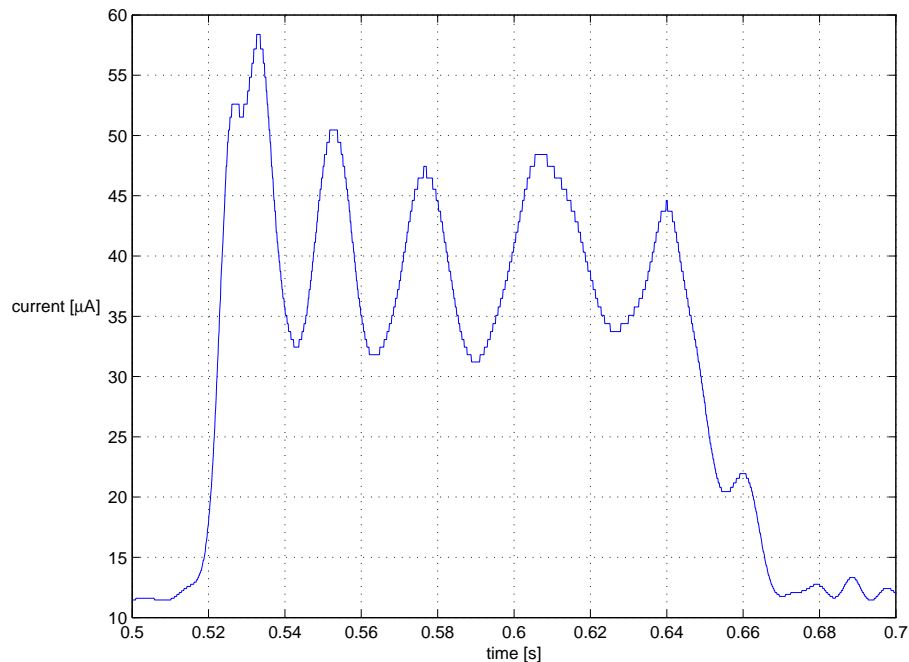
Obrázek 16: **Strategie ACE.** Vstupní signál z mikrofonu je vzorkován frekvencí 16 kHz a rozdělen na segmenty o délce 128 vzorků s tím, že dochází k 75% překryvu sousedních segmentů. Z každého segmentu je spočítáno spektrum pomocí algoritmu rychlé Fourierovy transformace (FFT). Signál je následně rozdělen pomocí 22 pásmových propustí. Pro každé pásmo je vypočítána jeho energie. Na základě nastavení je vybráno několik nejsilnějších pásem, maximálně však dvacet. Informace obsažená v nevybraných pásmech se dále nikam nepřenáší. Signál ve vybraných pásmech se zkomprimuje pomocí logaritmické funkce. Generátor převede zkomprimovaný signál na proudové impulsy při dané stimulační frekvenci. V paměti je uložen: počet vybíraných energetických maxim, stimulační frekvence, parametry komprese, maximální a minimální budící úrovně (pro každý kanál zvlášť), délka impulsu, mezery mezi impulsy a druh stimulace. Upraveno podle [Vondráška, 2008].

6.6 Porozumění řeči u pacientů s kochleárním implantátem

Schopnosti porozumět mluvené řeči se u pacientů s KI dramaticky liší a to i v případě, že jsou vybaveni stejným typem KI. Lze však najít několik faktorů, které tyto schopnosti ovlivňují:

- **Celková doba hluchoty** má silně negativní efekt na výkonnost porozumění mluvené řeči. Pacienti, kteří trpěli hluchotou kratší dobu vykazují lepší výsledky než ti s delší periodou hluchoty.
- **Doba začátku hluchoty** rovněž velice ovlivňuje výkonnost porozumění. Postlingválně hluší dosahují lepších výsledků než prelingválně hluší.
- **Věk, kdy je KI implantován** je zásadním faktorem u prelingválně hluchých. Zjednodušeně lze říci, čím později tím hůře.

- **Délka používání KI** pozitivně ovlivňuje výkonnost.
- **Fyziologické faktory** jako počet funkčních synapsí sluchového nervu, inteligence a komunikativnost.
- **Technické faktory** jako budící dynamický rozsah elektrod, strategie zpracování zvuku, umístění a hloubka zasunutí elektrod.



Obrázek 17: **Stimulační signál jedné elektrody KI.** Nejsilnější signál jedné z dvaceti elektrod KI při použití strategie SPEAK a komplexního zvuku – mluveného slova.

6.7 Binaurální slyšení a kochleární implantáty

Jedním z cílů při vyvíjení modelu binaurálního slyšení je aplikovat objasněné mechanismy při konstrukci respektive programování kochleárních implantátů. V současné době existuje ve světě řada pacientů se dvěma kochleárními implantáty (KI). Na první pohled by se mohlo zdát, že správné funkci prostorového slyšení nic nebrání. V praxi se však ukazuje, že mnoho pacientů je schopno jen obtížně schopno lokalizovat zdroj zvuku. Jak uvádí [Majdak et al., 2006], hlavní problémy způsobuje chybějící synchronizace

mezi oběma kochleárními implantáty (současná technická řešení se synchronizací nepočítají), nestejná frekvence obou procesorů, nízká vzorkovací frekvence, sériová stimulace elektrodami a výběr pouze nejsilnějších signálů ze spektra.

Podle nedávných výzkumů téže vídeňské skupiny [Laback et al., 2007] se ukazuje, že umělé zavedení zpoždění do signálu, podobně jako ve výše popsaném modelu, může u pacientů s KI zlepšit schopnost prostorové lokalizace.

Odpověď na bipolární stimulaci elektrodou kochleárního implantátu silně závisí na modulaci signálu [Litvak et al., 2001]. Za účelem modelování lze odpověď vláskové buňky dobře aproximovat modelem popsáním v publikaci [Bruce et al., 1999b] a [Bruce et al., 1999a]. Pokud by se podařilo zlepšit schopnost lokalizace zdroje zvuku u pacientů se dvěma KI, vedlo by to k zlepšení slyšení v situaci nazývané “cocktail-party problem”, neboli problém koktejlového večírku [Nie et al., 2005].

7 Klinické aspekty prostorového slyšení

Základní funkcí binaurálního prostorového slyšení savců je rozpoznání azimutu zdroje zvuku. Tato lokace pracuje na fyzikálním principu. Na základě znalosti frekvence zvuku je využit buď časový či intenzitní klíč. Pokud se jedná o známý zvuk (například hlasový projev zvířete) je možné rozpoznat i vzdálenost zdroje zvuku. Určení vzdálenosti není možné jen na základě fyzikálního principu poklesu intenzity zvuku se vzdáleností. Je třeba znát hlasitost zdroje zvuku, například typický hlasový projev určitého živočišného druhu. Bez této znalosti nelze rozpoznat blízký tichý zdroj od vzdáleného hlasitého.

V případě člověka žijícího ve středoevropském regionu není rozpoznání azimutu a intenzity zvuku otázkou přežití. Na druhé straně tomuto tvrzení částečně odporuje nařízení z USA, které ukládá výrobcům elektromobilů, které jsou jinak prakticky bezhlučné, aby generovaly nějaký zvuk. Stejně tak lidové označení moderních motorových železničních souprav termínem "tichá smrt" otvírá prostor k polemice.

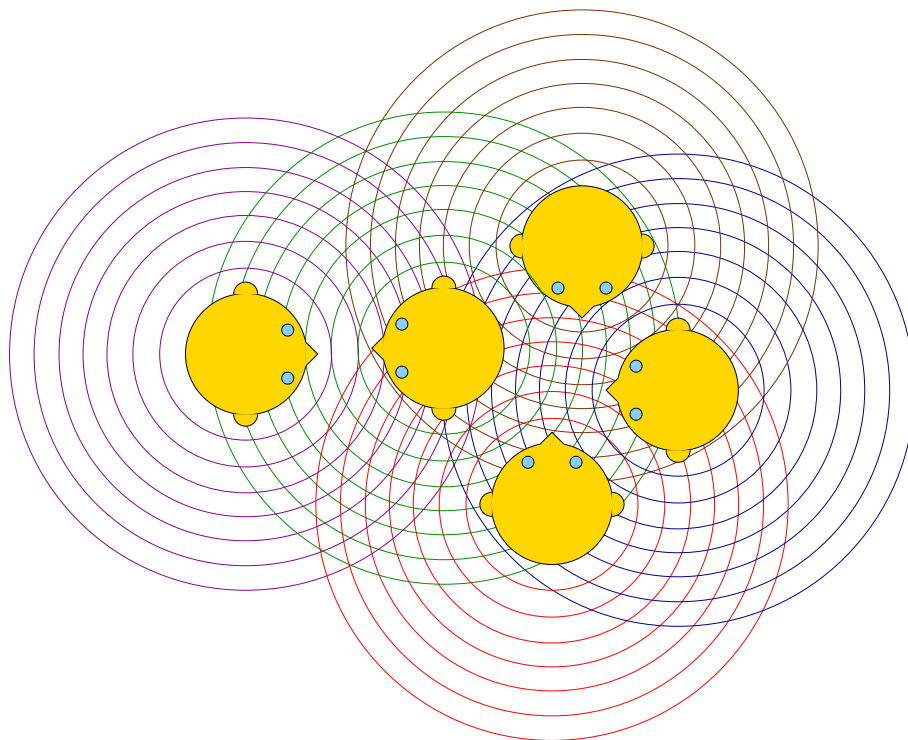
Prostorové slyšení hraje určitou roli i v orientaci v prostoru, což je zejména důležité pro nevidomé osoby, kteří si vytvářejí na základě echa jistý druh prostorové mapy. S binaurálním slyšením souvisí rovněž jev předcházení. Více o tomto jevu je v kapitole 9. Jako klinicky nejdůležitější se ovšem jeví efekt nazývaný binaurální hladina maskování, známý rovněž pod anglickým názvem cocktail party problem.

7.1 Problém na koktejlovém večírku

S efektem binaurální hladiny maskování se setkáváme běžně v konverzaci ve společnosti, při sledování televize, či poslechu radia v hlučnějším prostředí.

Jedná se o typický problém pacientů nedoslýchavých na jedno ucho, kteří si stěžují na to, že ve společnosti nejsou schopni konverzovat. V ruchu který je všude kolem nich nejsou schopni rozpoznat ten správný hlas, kterému chtějí naslouchat. Právě neschopnost konverzace ve společnosti bývá často spouštěčem, který přivede slaběji nedoslýchavé pacienty k lékaři.

Na základě experimentálních měření bylo určeno, že užitečný signál (zvuk) v šumu (ruchy okolí), který má výraznou směrovou charakteristiku je vnímán



Obrázek 18: **Problém na koktejlovém večírku.** Na obrázku je zachyceno pět mluvících lidí. Barevná soukruží schematicky představují vlnoplochy tvořené mluveným projevem. Zkoumaná osoba se zelenými vlnoplochami (zhruba uprostřed obrázku) hovoří s osobou vlevo (fialové vlnoplochy), ale tento hovor je rušen hlasy zbývajících osob, jak je patrné z průniků uší a vlnoploch.

o 10 až 15 dB silnější než ve skutečnosti je [Syka et al., 1981]. Tato směrová selektivita tedy umožňuje rozeznat hlas řečníka i při jeho značné vzdálenosti a silném maskování šumem. Schématické zobrazení typického uspořádání při cocktail-party je na obrázku 18.

8 Model prostorového slyšení s kochleárním implantátem

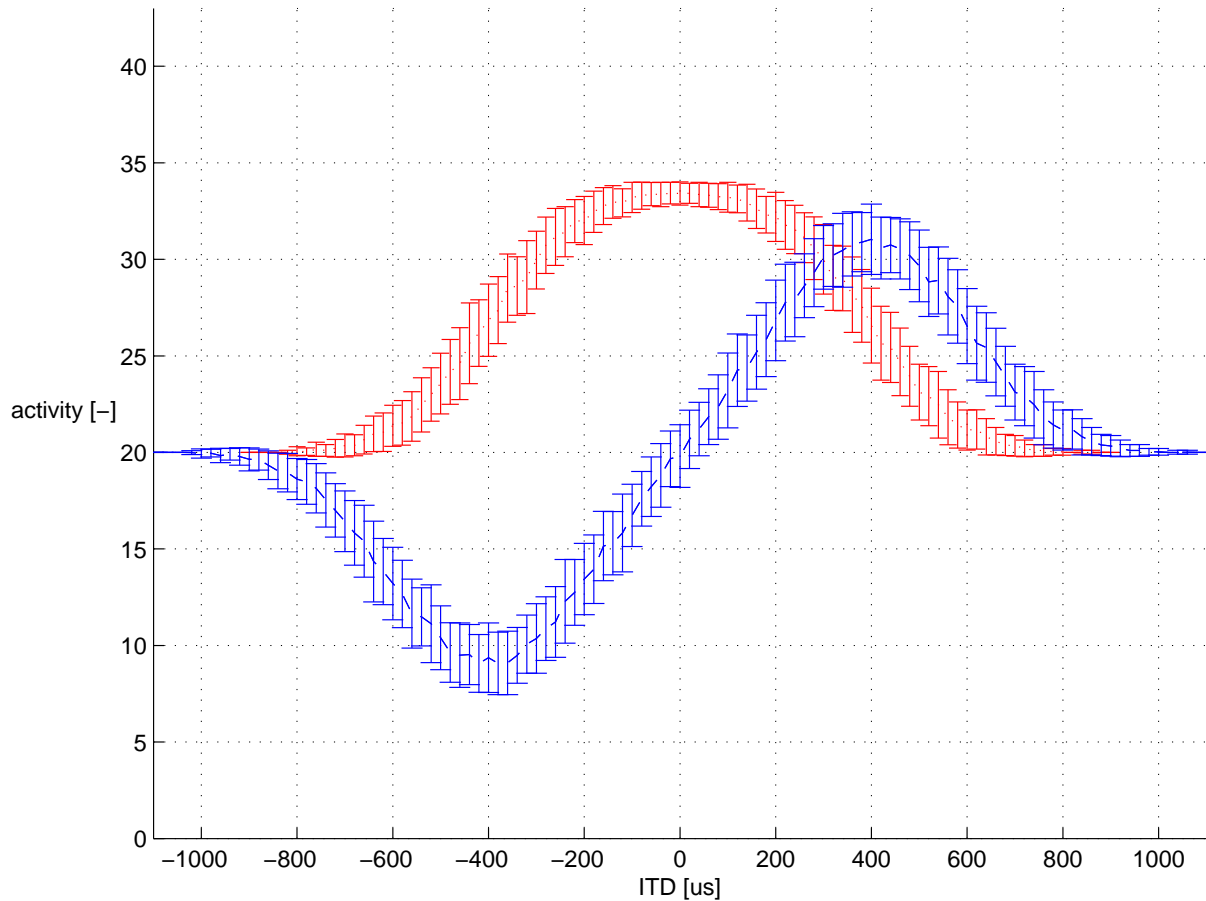
Model popsáný v oddíle 5 používá k přeměně zvukového signálu na nervové vzruchy pouze jednoduchý derivační obvod. Jednotlivé části věrohodnější transformace zvuku na aktivitu vnitřních vláskových buněk řeší některé již existující modely např. [Krishna, 2002], [Patterson and Holdsworth, 1991], [Patterson, 1994], [Irino and Patterson, 1997], [Slaney, 1993]. Vzhledem k tomu, že tyto modely řeší přeměnu zvuku na "spike-train" pouze parciálně, a jak již bylo řečeno v předchozím oddíle, že nás zajímá vliv časového rozostření na prostorové slyšení u pacientů s KI [Laback and Majdak, 2008], implementovali jsme tuto část s využitím softwaru používaného v kochleárních implantátech a přenosu z elektrod KI na sluchový nerv [Bruce et al., 1999b], [Bruce et al., 1999a]. Výsledný model popsáný v [Drapal and Marsalek, 2010] je tak komplexní, že není prakticky možné provést matematickou analýzu celého modelu, ale pouze jeho numerickou simulaci.

8.1 Numerická simulace

Vstupem do modelu, který byl naprogramován v prostředí Matlab, je komplexní zvukový signál, který je nejprve zpracován pomocí algoritmu SPEAK používaného v KI. Tento algoritmus je dodnes standardně používán v mnoha různých modifikacích v komerčních KI. Bylo by zajímavé prověřit fungování modelu i s jinými algoritmy, které jsou dnes v KI běžně dostupné, ale brání tomu silný konkurenční boj mezi výrobci KI, který vede k tomu, že algoritmy nejsou veřejně dostupné a jejich využití je vázáno mnoha restriktivními smlouvami.

Nicméně, ať už je použit jakýkoliv algoritmus, výsledkem je tonotopicky uspořádaná matice o typicky dvaceti řádcích, kde každý řádek představuje stimulaci do jedné elektrody KI v μA . Na obrázku 17 je pro představu zobrazen nejsilnější signál z jedné elektrody pro mluvené slovo.

Z matice stimulací je pro zjednodušení výpočetní náročnosti modelu vybrán řádek s nejsilnějším signálem, který je zpracován modelem popsáným v [Bruce et al., 1999a]. Vzhledem k tomu, že jednotlivé elektrody jsou v KI buzeny sériově, neubírá modelu výběr pouze jednoho řádku matice na obec-



Obrázek 19: **Výstupní aktivity z KI modelu.** Výstup z komplexního KI modelu odpovídající výstupu z LIAF modelu na obrázku 10. Ze srovnání obou obrázků je patrné, že model je schopen pracovat i s komplexními zvuky, bez újmy na funkčnosti.

nosti. Model popsáný v [Bruce et al., 1999a] resp. [Bruce et al., 1999b] je experimentálně verifikovaný numerický model, který na základě vstupního proudu generuje pravděpodobnost nervového vzruchu vnitřní vláskové buňky. Ověření modelu probíhalo na kočkách, které měly implantovány elektrody KI a měřila se pravděpodobnost odpovědi sluchového nervu na základě buzení elektrickým proudem stejným mechanismem, jaký je použit u KI.

Vlastní vzruchy v nervovém vlákne generované za pomoci pravděpodobnostního IAF neuronu následně stimulují model popsáný v oddíle 5. Na obrázku 19 je zobrazen výstup z tohoto komplexního modelu. Při srovnání výstupů KI a LIAF modelu prostorového slyšení je patrné, že modely jsou robustní nejen pokud jde o variaci jednotlivých parametrů, ale i v případě

komplexnosti vstupu. Nezávislost na míře komplexity vlastního vstupu je dána tím, že správně pracující model prostorového slyšení musí být vždy nahraditelný korelací dvou signálů – z levého a pravého ucha. Vzhledem k tomu, že korelace ze složitějšího signálu dává přesnější výsledky, není překvapující, že i KI model s komplexnějším vstupem (oproti LIAF modelu) pracuje správně.

9 Diskuse

Binaurální slyšení v živočišné třídě savců včetně člověka je na rozdíl od třídy ptáků stále otevřenou otázkou. Neuronální struktura odpovídající ptačí zpoždovací lince nebyla u člověka z anatomického hlediska lokalizována a stejně tak různá měření v rámci sluchového nervu a navazujících nervových center existenci zpoždovací linky u savců nepotvrzují, ale spíše vyvracejí. I přes podrobné znalosti fyziologické akustiky nejsme schopni řadu sluchových vjemů vysvětlit pouze na základě fyzikálních zákonů. Do hry často vstupuje psychofyzikální vysvětlení, které zatím nemá pevnou oporu ve fyziologii respektive ve znalosti sluchových neuronálních obvodů. I v případě prostorového slyšení tomu není jinak. V následujících odstavcích se tak po stručném úvodu dotýkáme několika problémů, na které jsme v rámci studia prostorového slyšení narazili.

Z hlediska fyziologické akustiky, hovoříme o zvuku jako o mechanickém vlnění ve frekvenčním rozsahu od 16 Hz do 20 kHz. Intenzita zvuku (neboli hladina hlasitosti) se nejčastěji udává v decibelech, což je poměrná logaritmická jednotka. Hodnota hladiny hlasitosti (SPL, sound pressure level) nula decibelů pak odpovídá prahu slyšení. V praxi se setkáme s hodnotou 30 dB v lese za bezvětrí, 90 dB je intenzita hlasitého křiku a od 130 dB (práh bolesti a hluk motoru tryskového motoru) dochází ve vnitřním uchu k rupturám blanitých přepážek a destrukci Cortiho orgánu [Syka et al., 1981].

Prostorové slyšení je primárně založeno na binaurálním slyšení. Pro frekvence do 1,5 kHz je uplatňován časový klíč, pro frekvence vyšší klíč intenzitní. Nejlepšího prostorového rozlišení lze dosáhnout pomocí časového klíče v okolí střední sagitální roviny.

V případě sluchových ztrát kolem 90 dB a více nelze využít ke zlepšení sluchového vjemu principu zesilovače (standardního naslouchátka), ale v případě vitálního sluchového nervu je možno využít kochleární implantát. Jedná se o sadu elektrod, které jsou zavedeny do hlemýždě Cortiho orgánu. Elektrody na základě buzení elektrickým proudem vytvářejí vjem zvuku o výšce, která odpovídá poloze v tonotopicky uspořádaném hlemýždi.

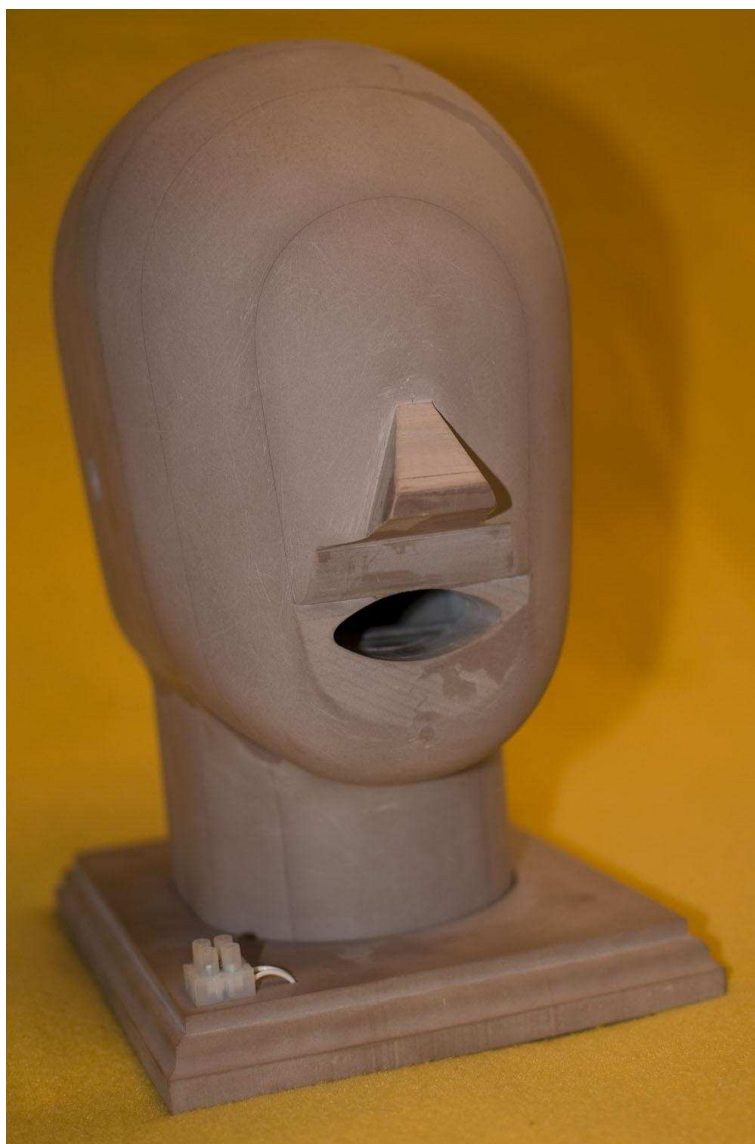
Z hlediska patologické fyziologie sluchu rozlišujeme dvě základní skupiny poruch, podle toho, kde na cestě mezi vzduchovým vedením a zpracováním mozkovou kůrou porucha nastává. Pokud je porucha před vnitřním uchem,

to jest před místem, kde je převáděna mechanická energie na posloupnosti akčních potenciálů, nazýváme tuto poruchu poruchou převodní. Pokud porucha nastává ve sluchové (nervové) dráze včetně hlemýždě, nazývá se tato porucha poruchou percepční. Je samozřejmě možné, že porucha sluchu je i smíšená. Navíc úplná oboustranná ztráta funkce vnějšího a středního ucha, například po opakovaných zánětech středního ucha v dětství s následným oboustranným chyběním bubínku, vede maximálně k oboustrannému zvýšení sluchového prahu o 35 dB. Každé zvýšení sluchového prahu s hodnotou vyšší než tato má tedy nutně i percepční složku.

V článku [Drapal and Marsalek, 2010] je k modelování buzení sluchového nervu použit kochleární implantát. Celý model prostorového slyšení s využitím KI pracuje velice spolehlivě i při použití komplexních (reálných) zvuků. V případě bilaterálních implantací KI, však pacienti prostorové slyšení nemají a pokud jej mají, tak velice špatné. Problém má celou řadu příčin, ale největší roli hraje absence synchronizace taktu DSP procesorů a časových oken, ze kterých je počítáno spektrum budícího signálu, frekvence budících impulsů elektrod a ztráta informace v rámci časového okna.

9.1 Oprávněnost použití rozdělení beta

V článku [Marsalek and Drapal, 2007], ale i v [Drapal and Marsalek, 2010] či [Drapal and Marsalek, 2011] je použito k modelování beta rozdělení. Jak již bylo uvedeno výše, předkládané modely jsou robustní pokud jde o variaci parametrů a komplexnost vstupu, nicméně otázka použitého pravděpodobnostního rozdělení diskutována nebyla. Je použité beta rozdělení opravdu biologicky oprávněné, nebylo by vhodnější použít např. rozdělení Poissonovo? Odpověď na tuto otázku jsme se pokusili najít v zatím nepublikovaných analytických výpočtech a numerických simulacích, ze kterých vyplývá, že model je velmi robustní i pokud jde o volbu pravděpodobnostního rozdělení. Jakmile je tvar pravděpodobnostního rozdělení alespoň vzdáleně podobný beta rozdělení, model pracuje spolehlivě. Toto nastává například pro normální, anebo pro gama rozdělení. Dokonce i pro triviální případy beta rozdělení, kterými jsou trojúhelníkovité a rovnoměrné rozdělení, dává model funkční výstup.



Obrázek 20: **Manekýn.** Hlava a torzo sloužící k měření a testování hypotéz prostorového slyšení. Manekýn je vyroben z umělého dřeva a je opatřen otvory pro přesné měřící kondenzátorové mikrofony.

9.2 Fyzikální versus psychofyzikální slyšení

Hlubší znalosti prostorového slyšení mohou být aplikovány nejen v lékařství ale i při návrhu ozvučení místností či při zpracování zvuku. Zpracování zvukového signálu mozkem není v mnoha případech lineární, ale zahrnuje řadu netriviálních operací. Vedle fyzikálních jevů je tedy třeba zohlednit i specifika psychofyzikální. Příkladem budiž například jev předcházení. Jev

předcházení se běžně uplatňuje v každé normální místnosti. Stěny místnosti většinou pohltí jen 20% intenzity zvuku a zbytek je odražen zpět. Přesto vnímáme pouze vlnu první a vlny následné v určitém časovém okně neslyšíme, ale ovlivňují vjem vlny první a to i pokud jde o vnímání prostorového zvuku. Stejně tak binaurální hladina maskování (cocktail party problem) umožňuje zesílit vjem prostorově lokalizovaného zdroje zvuku až o 15 dB oproti šumu. Výraznou roli hraje zejména pro vyšší frekvence stín hlavy a tvar ušních boltců.

9.3 Měření s použitím manekýna (fantomu)

Aby bylo možné ověřit některé teoretické předpoklady a odlišit psychofyzikální vnímání pokusných osob od vnímání fyzikálního, vybavili jsme naši laboratoř tak zvaným fantomem či manekýnem. Manekýn je na obrázku 20. Jedná se o zjednodušenou kopii hlavy vyrobenou z kompozitního materiálu s definovanými akustickými vlastnostmi nazývaného jednoduše umělé dřevo. Tento materiál se vyznačuje izotropním šířením zvuku. Rychlost šíření zvuku v umělém dřevu je srovnatelná s rychlostí šíření v lidské hlavě. Fantom je vybaven přesně soustruženými otvory pro zasunutí mimořádně citlivých měřících kondenzátorových mikrofونů, které byly společně s fantomem kalibrovány v bezdozvukové komoře na fakultě elektrotechnické ČVUT. Fantom je zcela unikátně vybaven sadou odnímatelných ušních boltců, které umožňují zkoumat za pomoci fantoma nejen binaurální ale i monaurální prostorové slyšení. Výstup z mikrofونů je možné připojit přímo do zvukové karty, která podporuje takzvané fantomové napájení mikrofونů v námi použité elektro-akustické soustavě. Výše uvedený výčet otevřených problémů, na které při studiu prostorového slyšení narážíme není zdaleka úplný, stejně tak jako výčet modelů a experimentů, které by bylo možné k poodhalení fyziologické reality použít.

10 Závěr

Snaha o nalezení přesného biologického popisu fyziologie prostorového slyšení savců v posledních letech zesílila zejména díky objevu zvláštní role inhibice [Brand et al., 2002], nicméně zatím se nepodařilo nalézt uspokojivou odpověď na otázku, jakým mechanismem se inhibice uplatňuje.

Předložená práce ukazuje jedno z možných řešení určení polohy zdroje zvuku, které se opírá o stochastické vlastnosti neuronů sluchového nervu a je velice robustní vzhledem k variabilitě parametrů. Vedle tohoto hlavního tématu, práce také otvírá souvislost mezi časovým rozostřením signálu a přesností lokalizace u pacientů s kochleárními implantáty.

Ačkoliv ztráta prostorového slyšení nemá pro člověka primárně fatální důsledky, přesto může být neschopnost rozpoznat, odkud zvuk přichází (zejména ve městech plných automobilů) důležitá. Neméně důležitou roli hraje prostorové slyšení v rozeznávání užitečného směrového signálu v šumu. Výzkum na tomto poli má rovněž velký význam při konstrukci párově pracujících kochleárních implantátů.

Použitá literatura

- [Batra et al., 1997] Batra, R., Kuwada, S., and Fitzpatrick, D. (1997). Sensitivity to interaural temporal disparities of low-and high-frequency neurons in the superior olivary complex. II. Coincidence detection. *Journal of Neurophysiology*, 78(3):1237–1247.
- [Brand et al., 2002] Brand, A., Behrend, O., Marquardt, T., McAlpine, D., and Grothe, B. (2002). Precise inhibition is essential for microsecond interaural time difference coding. *Nature*, 417(6888):543–547.
- [Bruce et al., 1999a] Bruce, I., Irlicht, L., White, M., O’Leary, S., Dynes, S., Javel, E., and Clark, G. (1999a). A stochastic model of the electrically stimulated auditory nerve: Pulse-Train Response. *IEEE Transactions on Biomedical Engineering*, 46(6):630–637.
- [Bruce et al., 1999b] Bruce, I., Irlicht, L., White, M., O’Leary, S., Dynes, S., Javel, E., and Clark, G. (1999b). A stochastic model of the electrically stimulated auditory nerve: Single-pulse response. *IEEE Transactions on Biomedical Engineering*, 46(6):617–629.
- [Clark, 2003] Clark, G. (2003). *Cochlear implants: fundamentals and applications*. Springer Verlag, New York.
- [Drapal and Marsalek, 2008] Drapal, M. and Marsalek, P. (2008). Model prostorového slyšení. In Kelemen, J., Kvasnička, V., and K., P., editors, *Kognice a umělý život VIII*, pages 93–96. Slezská Univerzita v Opavě.
- [Drapal and Marsalek, 2010] Drapal, M. and Marsalek, P. (2010). Stochastic model shows how cochlear implants process azimuth in real auditory space. *The Chinese Journal of Physiology*, 53(6). In press.
- [Drapal and Marsalek, 2011] Drapal, M. and Marsalek, P. (2011). Stochastic model explains role of excitation and inhibition in binaural sound localization in mammals. *Physiological Research*, 60(3). In press.
- [Fitzpatrick et al., 2000] Fitzpatrick, D., Kuwada, S., and Batra, R. (2000). Neural sensitivity to interaural time differences: beyond the Jeffress model. *Journal of Neuroscience*, 20(4):1605–1615.

- [Grothe, 2003] Grothe, B. (2003). New roles for synaptic inhibition in sound localization. *Nature Reviews Neuroscience*, 4(7):540–50.
- [Irino and Patterson, 1997] Irino, T. and Patterson, R. (1997). A time-domain, level-dependent auditory filter: The gammachirp. *The Journal of the Acoustical Society of America*, 101(1):412–419.
- [Jeffress, 1948] Jeffress, L. (1948). A place theory of sound localization. *Journal of Comparative & Physiological Psychology*, 41:35–39.
- [Joris et al., 1998] Joris, P., Smith, P., and Yin, T. (1998). Coincidence detection minireview in the auditory system: 50 years after Jeffress. *Neuron*, 21:1235–1238.
- [Joris et al., 2006] Joris, P., Van de Sande, B., Louage, D., and van der Heijden, M. (2006). Binaural and cochlear disparities. *Proceedings of the National Academy of Sciences of the United States of America*, 103(34):12917–12922.
- [Krishna, 2002] Krishna, B. (2002). A unified mechanism for spontaneous-rate and first-spike timing in the auditory nerve. *Journal of Computational Neuroscience*, 13(2):71–91.
- [Laback and Majdak, 2008] Laback, B. and Majdak, P. (2008). Binaural jitter improves interaural time-difference sensitivity of cochlear implantees at high pulse rates. *Proceedings of the National Academy of Sciences of the United States of America*, 105:814–817.
- [Laback et al., 2007] Laback, B., Majdak, P., and Baumgartner, W. (2007). Lateralization discrimination of interaural time delays in four-pulse sequences in electric and acoustic hearing. *The Journal of the Acoustical Society of America*, 121:2182–2191.
- [Litvak et al., 2001] Litvak, L., Delgutte, B., and Eddington, D. (2001). Auditory nerve fiber responses to electric stimulation: Modulated and unmodulated pulse trains. *The Journal of the Acoustical Society of America*, 110:368–379.

- [Loizou, 1998] Loizou, P. (1998). Mimicking the human ear. *IEEE Signal Processing Magazine*, 15(5):101–130.
- [Magnusson et al., 2005] Magnusson, A., Kapfer, C., Grothe, B., and Koch, U. (2005). Maturation of glycinergic inhibition in the gerbil medial superior olive after hearing onset. *The Journal of Physiology*, 568(2):497–512.
- [Majdak et al., 2006] Majdak, P., Laback, B., and Baumgartner, W. (2006). Effects of interaural time differences in fine structure and envelope on lateral discrimination in electric hearing. *The Journal of the Acoustical Society of America*, 120:2190–2201.
- [Marsalek and Drapal, 2007] Marsalek, P. and Drapal, M. (2007). Mechanisms of coincidence detection in the auditory brainstem: Examples. In Deutsch, A., Bravo de la Parra, R., de Boer, R., Diekmann, O., Jagers, P., Kisdi, E., Kretzschmar, M., Lansky, P., and Metz, H., editors, *Mathematical Modeling of Biological Systems*, volume 2, pages 255–264. Birkhaeuser.
- [Marsalek and Drapal, 2009] Marsalek, P. and Drapal, M. (2009). Neural coding in sound localization pathway, roles of timing jitter and random delay. *Neuronal Coding, 8th International Workshop*, pages 105–109.
- [Marsalek and Kofranek, 2004] Marsalek, P. and Kofranek, J. (2004). Sound localization at high frequencies and across the frequency range. *Neurocomputing*, 58–60:999–1006.
- [Marsalek and Kofranek, 2005] Marsalek, P. and Kofranek, J. (2005). Spike encoding mechanisms in the sound localization pathway. *Biosystems*, 79(1-3):191–198.
- [Marsalek and Lansky, 2005] Marsalek, P. and Lansky, P. (2005). Proposed mechanisms for coincidence detection in the auditory brainstem. *Biological Cybernetics*, 92(6):445–451.
- [McAlpine and Grothe, 2003] McAlpine, D. and Grothe, B. (2003). Sound localization and delay lines—do mammals fit the model? *Trends in Neurosciences*, 26(7):347–350.

- [McAlpine et al., 2001] McAlpine, D., Jiang, D., and Palmer, A. (2001). A neural code for low-frequency sound localization in mammals. *Nature Neuroscience*, 4:396–401.
- [Middlebrooks and Green, 1991] Middlebrooks, J. and Green, D. (1991). Sound localization by human listeners. *Annual Review of Psychology*, 42(1):135–159.
- [Nie et al., 2005] Nie, K., Stickney, G., and Zeng, F. (2005). Encoding frequency modulation to improve cochlear implant performance in noise. *IEEE Transactions on Biomedical Engineering*, 52(1):64–73.
- [Park et al., 2004] Park, T., Klug, A., Holinstat, M., and Grothe, B. (2004). Interaural level difference processing in the lateral superior olive and the inferior colliculus. *Journal of Neurophysiology*, 92(1):289–301.
- [Patterson, 1994] Patterson, R. (1994). The sound of a sinusoid: Spectral models. *The Journal of the Acoustical Society of America*, 96:1409–1418.
- [Patterson and Holdsworth, 1991] Patterson, R. and Holdsworth, J. (1991). A functional model of neural activity patterns and auditory images. *Advances in Speech, Hearing and Language Processing*, 3:547–563.
- [Recanzone, 2000] Recanzone, G. (2000). Spatial processing in the auditory cortex of the macaque monkey. *Proceedings of the National Academy of Sciences of the United States of America*, 97:11829–11835.
- [Slaney, 1993] Slaney, M. (1993). An efficient implementation of the Patterson-Holdsworth auditory filter bank. *Perception Group, Apple Computer, Tech. Rep.*, 35:1–41.
- [Smith, 2001] Smith, L. (2001). Using depressing synapses for phase locked auditory onset detection. *Proceedings of the International Conference on Artificial Neural Networks*, pages 1103–1108.
- [Spoendlin and Schrott, 1989] Spoendlin, H. and Schrott, A. (1989). Analysis of the human auditory nerve. *Hearing Research*, 43(1):25–38.
- [Swanson, 2001] Swanson, B. (2001). *Nucleus® Reference manual*. Cochlear Ltd., Australia.

- [Syka et al., 1981] Syka, J., Voldrich, L., and Vrabec, F. (1981). *Fyziologie a patofyziologie zraku a sluchu*. Avicenum, Praha.
- [Tollin, 2003] Tollin, D. (2003). The lateral superior olive: a functional role in sound source localization. *The Neuroscientist*, 9(2):127–143.
- [Vondrášek, 2008] Vondrášek, M. (2008). *Speech preprocessing for cochlear implants, doctoral thesis*. Faculty of electrical engineering, Czech technical university in Prague.
- [Zhou et al., 2005] Zhou, Y., Carney, L., and Colburn, H. (2005). A model for interaural time difference sensitivity in the medial superior olive: interaction of excitatory and inhibitory synaptic inputs, channel dynamics, and cellular morphology. *Journal of Neuroscience*, 25(12):3046–3058.

Seznam obrázků

1	Schematické znázornění neurčitosti ve 3D	13
2	Schematické znázornění neurčitosti ve 2D	14
3	Rozlišovací kroky v horizontální rovině	15
4	Tonotopické uspořádání hlemýždě.	19
5	Část inervace kochleárního jádra.	20
6	Schematické znázornění zpoždovací linky.	22
7	Koincidenční detektor využívající excitaci a inhibici.	25
8	Blokové schéma modelu.	26
9	Grafy beta rozdělení.	27
10	Výstupní aktivity z modelu.	28
11	Experimentální data.	29
12	Kochleární implantát.	33
13	Různé stimulační módy elektrod.	34
14	Kochleární implantát.	36
15	Logaritmická komprese	38
16	Strategie ACE.	40
17	Stimulační signál jedné elektrody KI.	41
18	Problém na koktejlové party.	44
19	Výstupní aktivity z KI modelu.	46
20	Manekýn	50

A Články editované a prošlé recenzním řízením

A.1 Mechanisms of coincidence detection in the auditory brainstem: Examples

Marsalek, P. and Drapal, M. (2008). Mechanisms of coincidence detection in the auditory brainstem: Examples. In Deutsch, A., Bravo de la Parra, R., de Boer, R., Diekmann, O., Jagers, P., Kisdi, E., Kretzschmar, M., Lansky, P., and Metz, H., editors, *Mathematical Modeling of Biological Systems*, volume 2, pages 255–264. Birkhaeuser.

Mechanisms of Coincidence Detection in the Auditory Brainstem: Examples

Petr Marsalek^{1,2} and Marek Drapal¹

¹ Charles University Prague, Department of Pathological Physiology, U nemocnice 5, CZ-128 53, Praha 2, The Czech Republic; marsalek@cesnet.cz, drapal@email.cz

² Corresponding author. Czech Technical University in Prague, Faculty of Biomedical Engineering, nam. Sitna 3105, CZ-272 01, Kladno, The Czech Republic

Summary. The auditory brainstem in mammals contains a neural circuit for sound localization. The exact functioning of this circuit is still under controversy. Two spike generation mechanisms studied previously, excitatory coincidence detection and inhibitory coincidence detection, are studied here regarding the input-output relationship of the spike time densities. We propose that synchronous binary multiplication operation on spikes is the underlying process of these two variants of coincidence detection. A derivation of time to the spike is shown, which enables us to estimate the contribution of the neural circuit in the auditory brainstem to the overall reaction time of sound localization. The brainstem contribution is minute compared to the conduction delays in the mammalian neocortex. Finally, the skewness of the resulting output spike time densities is discussed in both the excitatory and inhibitory cases and the inhibitory case is shown to be close to the normal density with a standard goodness-of-fit test for the normal probability density function.

Key words: Coincidence detection, directional hearing, interaural time delay, stochastic neuronal model.

22.1 Introduction

The auditory brainstem in mammals contains a neural circuit for sound localization. When the direction of incoming sound changes, so do the differences of some of its parameters between the two ears (interaural parameters). Sound localization is based on these sound parameters: interaural spectral difference, interaural intensity difference and interaural time difference (ITD). In this chapter we deal only with the ITD. The ITD changes when there is a change of the angle of the line from the sound source to the center of the head with the left-right symmetry plane. The discrimination threshold of this angle change in humans corresponds to a time difference in the range of 10 microseconds. Alternatively, such a short time difference can be described as a sound phase difference. Of course, such short time differences are not perceived subjectively. They are only employed in the sound localization circuit in the brainstem.

In the brainstem of birds the ITD is processed by a delay line [16], as it was suggested more than half a century ago in a visionary work of Jeffress [5]. However, such a delay line has not been found in mammals and recent experiments [1,4,12] and [15] indicate that the mechanism of the ITD processing in mammals does not work as the delay line. Yet in both birds and mammals the microsecond ITD has to be somehow detected by a unit (by a neuron, or by a neural circuit) acting as a coincidence detector. Without the coincidence detection none of the circuits in both birds and mammals would work.

This chapter presents a stochastic model of the coincidence detection (CD) in the auditory brainstem neural circuits. We have studied the feed-forward processing in the brainstem previously [8,10,11]. Here we present a not yet published example of CD using the beta density as a probability density of input spike occurrences in time. We stated previously that in the brainstem circuitry two types of CD take place, based on the polarity of the synaptic transmission. We call these two types of CD the excitatory CD and the inhibitory CD [10].

In [14] the CD was proposed as a mechanism for multiplication of firing rates within a limited spiking frequency range of asynchronous spike trains in a compound eye of a crab. Since in our synchronous cases the frequency is a sub-harmonic of sound, in our versions of spike operations only binary and not continuous multiplication with spikes is implemented.

22.2 Results

22.2.1 Conditions of Coincidence Detection

Following [14], we denote input from one side A, and from the other side B. The CD is realized when input spikes from both sides are closer in time than the fixed value of Δ , which is a constant of the neural circuit. During neuronal relaying, spikes are subject to a delay with a random component. We assume that the random delays on sides A and B are mutually independent and identically distributed non-negative random variables D with a maximum value of δ :

$$0 \leq D \leq \delta. \quad (22.1)$$

For relevant sound frequencies and delays the length of the window for CD Δ is not longer than the maximum delay δ and the delay δ is not longer than the sound period T :

$$0 \leq \Delta \leq \delta \leq T. \quad (22.2)$$

In the excitatory CD the spike is generated only when the two spikes from sides A and B meet in a time interval shorter than Δ , in other words when the two spike delays D_A and D_B are closer in time than Δ :

$$|D_A - D_B| \leq \Delta. \quad (22.3)$$

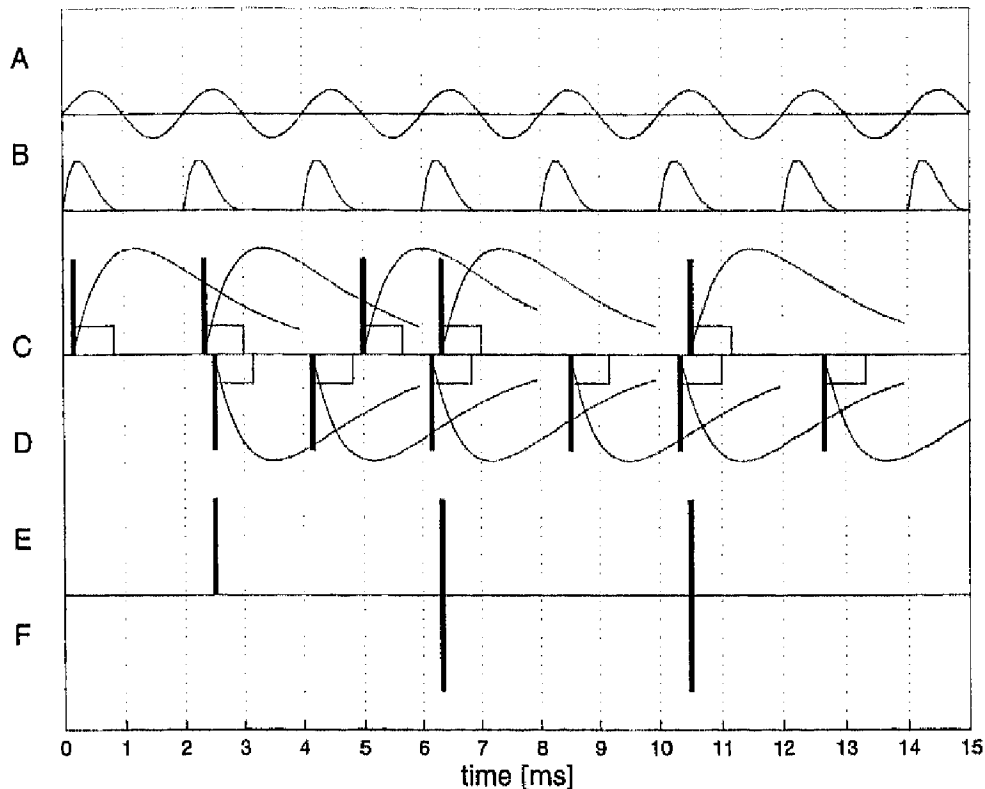


Fig. 22.1. Generation of the output spike trains of the CD mechanisms. Trace A: Sound with the frequency 500 Hz. Trace B: Series of spike time densities $f(x)$ phase locked to the sound rising phase. Traces C and D: The windows of length Δ (thin rectangles) are triggered by the leading edges (thick horizontal bars) of the spikes. These spikes (thin smooth functions with cut off tails) are postsynaptic potentials of the two inputs. The D trace is flipped vertically to show the overlaps of the coincidence detection windows. Trace E: The spike train of the excitatory CD is generated by placing the output spike aligned with the later of the two spikes of C and D, if the two windows of length Δ overlap in time. Trace F: For the output spike in the spike train of the inhibitory CD, all the conditions for the excitatory CD are necessary and furthermore the (inhibitory) spike from the contra-lateral side (trace D) must arrive sooner in time than the spike from the ipsi-lateral side (trace C). The F trace is also flipped vertically.

In the inhibitory CD the condition expressed in equation (22.3) is modified. Spikes must arrive in the right succession. The excitatory spike from side A must come after the inhibitory spike from side B. This is formulated as

$$0 \leq D_A - D_B \leq \Delta. \quad (22.4)$$

Mechanisms showing how these excitatory CD and inhibitory CD are generated are shown in Fig. 22.1. The example sound frequency is 500 Hz in the figure. The postsynaptic potentials of the two inputs from cochlear nuclei are the series of alpha functions, $x \exp(-x/\tau)$, with $\tau = 1$ ms. The window duration is $\Delta = 0.66$ ms, mean delays are $E(D_A) = E(D_B) = 0.33$ ms and their standard deviation is 0.18 ms. Note the longer duration of these alpha functions, compared to the CD windows, even though their time constant τ is shorter than typical biological time constants of postsynaptic potentials. Durations of all the events depicted in Fig. 22.1 are only loosely to scale.

Further, we denote the probability of events described by inequalities (22.3) and (22.4) as $P(\Delta)$. For $\Delta = 0$ it holds that $P(\Delta) = 0$. In the case of $\Delta = \delta$, $P(\Delta) = 1$, and since $P(\Delta)$ monotonically increases with Δ , $P(\Delta)$ as a function of Δ is a cumulative distribution function. Random variables D are distributed in accordance with a given probability density function $f(x)$. Then for the excitatory CD, $P(\Delta) = \text{Prob}(|D_A - D_B| \leq \Delta)$ can be calculated. Analogously for the inhibitory CD, $P(\Delta) = \text{Prob}(0 \leq D_A - D_B \leq \Delta)$. The output spike is emitted at the moment of the latter input spike. The output time density for specific examples of D and $f(x)$ and the rationale behind these examples are given in Section 22.2.3. In Fig. 22.1, beta density B_{24} discussed in Section 22.2.3 is shown as the example of the density $f(x)$.

22.2.2 Further Sources of Probabilistic Spiking

Under conditions (22.3) or (22.4), respectively, when $\Delta < \delta$, the spike generation probability is lower than 1. In [11] we showed that lowered spiking probability is also due to the process of sub-harmonic spike generation, which progressively lowers the spiking probability towards higher frequencies. We also gave a formula for dependence of this probability on the main sound frequency in that paper. Let us denote the net spiking probability set by the processes mentioned above as p . Further, one could ask: How many sound cycles and how much time does it take before the output spike is generated with desired reliability? Let us study the 50% and the 95% reliability. The processing time will determine the processing speed and therefore the reaction time for the sound localization circuit as a whole.

The procedure to calculate the processing time is shown in Fig. 22.2 for the two example values of $p = 0.4$ and 0.8 . N is the number of sound cycles from 1 up to the cycle in which spikes at both neurons A and B are generated together at some time. Let $P(n)$ denote the probability that $N \leq n$. $P(n)$ is a discrete cumulative distribution function of n . Using the spike delivery probability p we calculate a probability of how many cycles it will take for the mechanisms of the excitatory CD and of the inhibitory CD to generate their output. From this we will characterize the processing time by the two example values of N : by the number of cycles at which there is a 50% reliability of generating the output spike and by the number of cycles needed to reach a 95% reliability that the output spike is generated. The processing times to the first spike of the 50% and 95% reliabilities are NT , for the sound period T and for the two example values of N . The larger of the last two delays D in the last sound cycle is not included in these formulas, see [11] for details. In calculating these processing times we also set the coincidence detection probability in (22.3) to $P(\Delta) = 1$. Other cases, with $P(\Delta) < 1$, would be calculated analogously.

The N is the minimal n such that the probability P of spike generation in cycle n is $P(n) \geq R$ (where reliability R is 0.5 or 0.95). This N is found as the occurrence of a success in a partial trial of a set of Bernoulli trials. This scheme does not exclude the occurrence of a success in some other partial trial. Consequently, the probability of the event that a success occurs in at least one partial trial of n trials does not equal np^2 , as one might naively expect, but the probability equals to $1 - (1 - p^2)^n$. Therefore we seek the minimal natural number N satisfying inequality

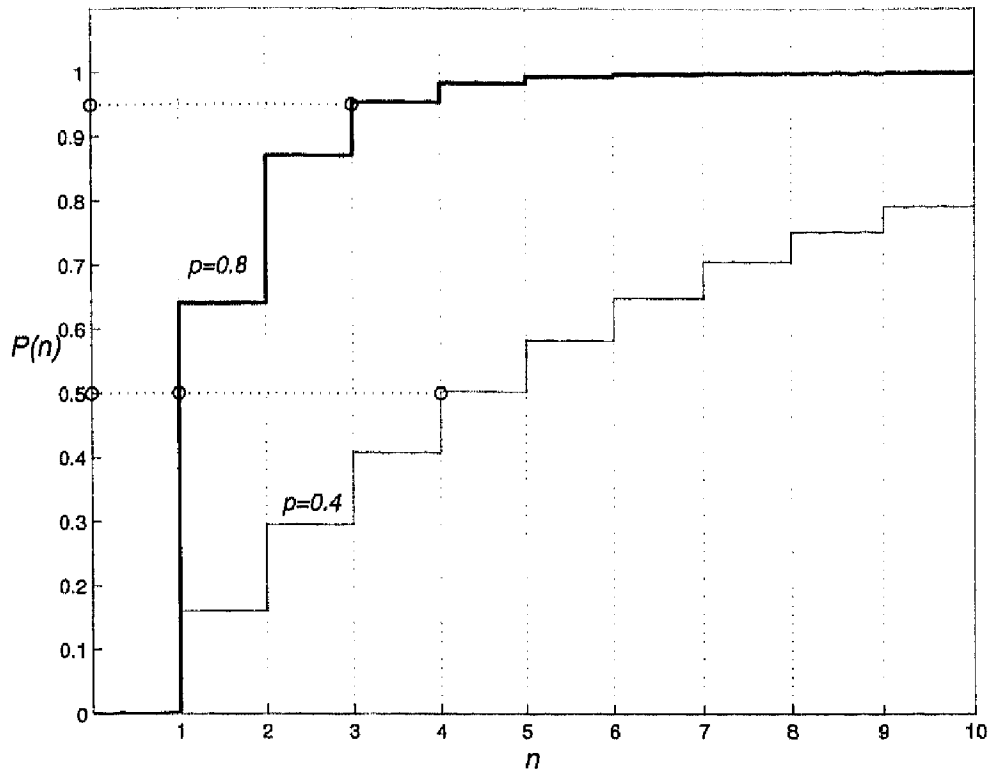


Fig. 22.2. Determining processing time counted in sound cycles from $P(n)$. $P(n)$ is discrete and is discretized by the sound period steps. The values of $P(n)$ are the following: during the first cycle $P(1) = p^2$, during the second cycle $P(2) = p^2 + (1 - p^2)p^2$, in the n th cycle $P(n) = \sum_0^{n-1} (1 - p^2)^i p^2$ and this sum gives the left side in inequality (22.5). $P(n) = 1$ for $n \rightarrow \infty$. Intersection of these staircases with horizontals at $R = 0.5$ and $R = 0.95$, respectively, gives the values of natural number N . Examples with $p = 0.4$, light line, and $p = 0.8$, bold line, are shown, the latter in order to demonstrate that N can even be equal to 1 for the value of $p > \sqrt{2}/2$ and the reliability value of $R = 0.5$.

$$1 - (1 - p^2)^N \geq R \tag{22.5}$$

for unknown N and for the values of R equal to 0.5 and to 0.95. Solving this inequality for N we obtain N as

$$N = \left\lceil \frac{\log(1 - R)}{\log(1 - p^2)} \right\rceil, \tag{22.6}$$

where $\lceil \cdot \rceil$ denotes rounding towards the nearest larger natural number. The base of the logarithm in both the numerator and the denominator can be arbitrary. Some numerical examples are shown in Fig. 22.2 and some others are used as parameters in a table in [11].

22.2.3 Examples of Input and Output Density Functions

For the input spike time probability density function we will use the beta density. We will write the beta density in a standard form, as

$$B_{ab} = \begin{cases} x^{a-1} (1 - x)^{b-1} B(a, b)^{-1}, & \text{for } x \in [0, 1] \\ 0 & \text{otherwise,} \end{cases} \tag{22.7}$$

where the parameters $a, b > 0$ and $B(a, b)$ is the beta function. The beta density with $a, b \geq 1, a < b$, which means nonzero skewness, fits well to experimental data. This is also because its nonzero value is confined to the range of $[0, 1]$ and this guarantees the assumption (22.2), namely that the spike fits into one sound cycle, after the cycle length $[0, T]$ is normalized to $[0, 1]$. We substitute the beta density B_{ab} with parameters a, b for the probability density function $f(x)$ of D_A and D_B in equation (22.3). For $a = 1, b = 1$, the probability density function becomes the uniform density. The uniform density was used as a choice for $f(x)$ in [10]. Next we will use parameters $a = 2$ and $b = 4$ as the example:

$$B_{24}(x) = \begin{cases} 20x(1-x)^3, & \text{for } x \in [0, 1] \\ 0 & \text{otherwise.} \end{cases} \tag{22.8}$$

To calculate the probability of the CD the condition (22.3) has to be fulfilled. Therefore for the excitatory CD we seek a density u of the random variable $W = |D_A - D_B|$, where D_A and D_B are distributed in accordance with equation (22.8). As the intermediary step giving out the density for the inhibitory CD we set

$$Z = D_A - D_B, \tag{22.9}$$

with its density denoted $q(z)$.

We substitute $f(x) = B_{24}(x)$ and $g(x) = B_{24}(x)$ into the standard convolution formula for the sum of two random variables, $q(z) = \int_{-\infty}^{\infty} f(x)g(x - z)dx$. The piecewise evaluated integral $q(z)$ is

$$\begin{aligned} q(z) &= \int_{\max(z-1,0)}^{\min(z,1)} f(x)g(z - x)dx \\ &= \int_{z \in [-1,0]} f(x)g(z - x)dx + \int_{z \in [0,1]} f(x)g(z - x)dx \\ &= \int_0^z f(x)g(z - x)dx + \int_{z-1}^1 f(x)g(z - x)dx. \end{aligned}$$

Obviously, $q(z)$ is an even function. We substitute into these integrals for f and g from (22.8) and we obtain the sum of two ninth-degree polynomials in two variables x and z when starting with a fourth-degree polynomial B_{24} . This would require tedious calculation without the use of a symbolic manipulator software package (we used the Symbolic Math Toolbox of the Matlab package).

In summary, using (22.3) and assuming that D_A and D_B are distributed with density (22.8), then $Z = D_A - D_B$ for the inhibitory CD has the following density:

$$q_{24}(z) = \begin{cases} -0.6349z^9 + 2.8571z^7 - 20z^4 + 33.33z^3 \\ \quad -17.1429z^2 + 1.5873, & \text{for } x \in [0, 1], \\ 0.6349z^9 - 2.8571z^7 - 20z^4 - 33.33z^3 \\ \quad -17.1429z^2 + 1.5873, & \text{for } x \in [-1, 0] \\ 0 & \text{otherwise.} \end{cases} \tag{22.10}$$

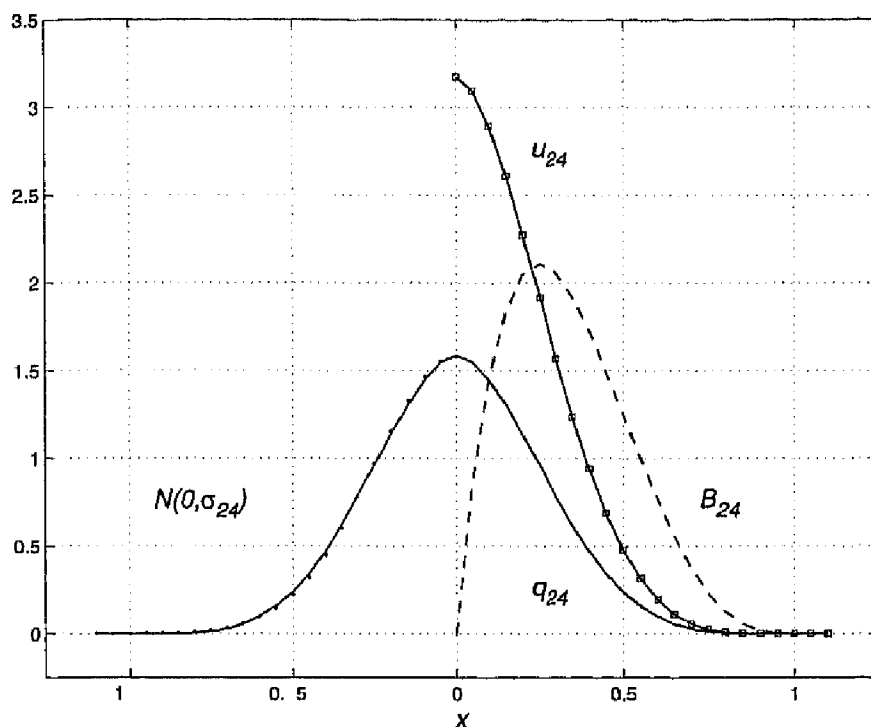


Fig. 22.3. Densities of input and output time in one sound cycle. B_{24} , broken line, is the input beta density with parameters of $a = 2$, $b = 4$. The intermediate result of calculations, q_{24} , is the solid line, in fact consisting of the two separate but aligned polynomials for x in $[-1, 0]$ and in $[0, 1]$, respectively, of (22.10). Dots, superimposed on the left branch, denoted $N(0, \sigma_{24})$, belong to the normal density with the respective mean and standard deviation and are close to the solid lines of q_{24} . Solid line marked with squares is the output density u_{24} , which is zero outside the range of $[0, 1]$.

The variable calculated as the absolute value of $W = |D_A - D_B|$ for the excitatory CD then has the resulting density $u_{24}(w)$, $u_{24}(w) = 2q_{24}(w)$ for $w \in [0, 1]$ and $u_{24}(w) = 0$ otherwise.

Starting from B_{24} , for enough sample data points, the numerically calculated density of $q_{24}(z)$ is close to the normal density. Testing 500 data points of the density $q_{24}(z)$ for the samples of normal density gives the positive answer with a significance level of 10^{-15} . This was in the Jarque–Bera test [2] for goodness of fit to, or departure from, the normal density, based on the sample kurtosis and skewness.

Therefore the normal density $N(\mu, \sigma)$ with zero mean and adjusted standard deviation of $\sigma_{24} = 0.2518$ can replace the density q_{24} . For maximal delay δ the range of x is $x \in [0, \delta]$, in other words we should substitute x/δ for x . This renormalization will match to the B_{24} with nonzero values from $[0, 1]$ and the standard deviation of $N(\mu, \sigma_{24})$ will be equal to $\sigma_{24} = 0.2518/\delta$. The nonzero but close to zero values of $N(\mu, \sigma_{24})$ outside the range of $[-1, 1]$ are negligible, as shown by the above-mentioned significance level. Densities B_{24} , q_{24} , u_{24} and $N(\mu, \sigma_{24})$ are shown in Fig. 22.3. In the next section we comment on the prospective use of the normal density in the description of output density. Even though the input beta density is used specifically because the input spike trains are generated by a sound stimulus, the output density q_{24} for the inhibitory CD is close to the normal density. The signs in the gener-

ation of the excitatory CD do not cancel, the algorithm is not symmetric in this sense. Therefore its output density u_{24} is also not symmetric and has a nonzero skewness.

22.3 Discussion

The wiring of the brainstem circuits for the CD is described in detail in previous works [11] and reviews [5]. We will discuss only the differences between the excitatory and inhibitory CD below. Of all the experimental literature which calls for an alternative to a delay line theory in mammals we list just the representative papers [1, 4, 12].

Concerning the difference in the summation sign between the excitatory and inhibitory CDs, we should note that the polarity change is possible in a neural circuit only by chaining the excitatory and inhibitory synapse in series. In addition, negative and positive amplitudes which are summed at a postsynaptic neuron have exactly two origins: (1) the excitatory and inhibitory synapses at the same postsynaptic neuron and (2) depolarization and hyperpolarization currents at the same neuron [3, 4] and [15].

All spike generation is discussed with respect to one sound cycle. Spikes are generated synchronously with a sub-harmonic oscillation and a limit frequency is imposed to the output frequency. This frequency limit, also mentioned at the beginning of Section 22.2.2, was used in [6] before. The processing time discussed in this section as a contribution to the reaction time would be just a minute contribution to the cortical processing in humans. However, the smaller the experimental animal is, the faster is its cortical processing and then the processing time may be more important. These reaction and processing times may also be of interest for an experimenter, who could devise further experiments elucidating the functioning of the brainstem neural circuitry.

The prediction of the output density type is useful in the description of propagation of the spike timing jitter into higher-order neurons within the auditory pathway. Most authors choose normal density, sometimes because there is no prior knowledge about the density. In the examples of output excitatory and inhibitory densities we show how such knowledge can be derived. Other examples of densities can be found in [7, 9] and [13].

Acknowledgments

Thanks to Michal Andrlík, Nick Dorrell and Petr Lánský. Special thanks to Eduard Kuriščák and Martin Zápotocký. Supported by the GAUK grant no. 32/2005 and by the MSM grant no. 6840770012.

References

1. Batra, R., Kuwada, S., Fitzpatrick, D. C.: Sensitivity to interaural temporal disparities of low- and high- frequency neurons in the superior olivary complex. I. Heterogeneity of responses. II. Coincidence detection. *J. Neurophysiol.*, **78**, 1222–1247 (1997).

2. Bera, A., Jarque, C.: Efficient tests for normality, heteroscedasticity and serial independence of regression residuals: Monte Carlo evidence. *Economics Letter*, **7**, 313–318 (1981).
3. Borisyuk, A., Semple, M. N., Rinzel, J.: Adaptation and inhibition underlie responses to time-varying interaural phase cues in a model of inferior colliculus neurons. *J. Neurophysiol.*, **88**, 2134–2146 (2002).
4. Brand, A., Behrend, O., Marquardt, T., McAlpine, D., Grothe, B.: Precise inhibition is essential for microsecond interaural time difference coding. *Nature*, **417**, 543–547 (2002).
5. Joris, P. X., Smith, P. H., Yin, T. C. T.: Coincidence detection in the auditory system: 50 years after Jeffress. *Neuron*, **21**, 1235–1238 (1998).
6. Kral, A., Majernik, V.: Neural networks simulating the frequency discrimination of hearing for non-stationary short tone stimuli. *Biol. Cybern.*, **74**, 359–366 (1996).
7. Lansky, P.: Sources of periodical force in noisy integrate-and-fire models of neuronal dynamics. *Phys. Rev. E*, **55**, 2040–2043 (1997).
8. Marsalek, P.: Coincidence detection in the Hodgkin-Huxley equations. *Biosystems*, **58**, 83–91 (2000).
9. Marsalek, P., Koch, C., Maunsell, J.: On the relationship between synaptic input and spike output jitter in individual neurons. *Proc. Natl. Acad. Sci. USA*, **94**, 735–740 (1997).
10. Marsalek, P., Kofranek, J.: Sound localization at high frequencies and across the frequency range. *Neurocomputing*, **58–60**, 999–1006 (2004).
11. Marsalek, P., Lansky, P.: Proposed mechanisms for coincidence detection in the auditory brainstem. *Biol. Cybern.*, **92**, 445–451 (2005).
12. McAlpine, D., Jiang, D., Palmer, A. R.: A neural code for low-frequency sound localization in mammals. *Nat. Neuroscience*, **4**, 396–401 (2001).
13. Reed, M. C., Blum, J. J., Mitchell, C. C.: Precision of neural timing: effects of convergence and time windowing. *J. Comput. Neurosci.*, **14**, 34–47 (2002).
14. Srinivasan, M. V., Bernard, G. D.: A proposed mechanism for multiplication of neural signals. *Biol. Cybern.*, **21**, 227–236 (1976).
15. Szalicszyo, K., Zalanyi, L.: Role of hyperpolarization-activated conductances in the auditory brainstem. *Neurocomputing*, **58–60**, 401–407 (2004).
16. Wagner, H., Brill, S., Kempfer, R., Carr, C. E.: Microsecond precision of phase delay in the auditory system of the barn owl. *J. Neurophysiol.*, **94**, 1655–8 (2005).

A.2 Stochastic model explains role of excitation and inhibition in binaural sound localization in mammals

Drapal, M. and Marsalek, P. (2011). Stochastic model explains role of excitation and inhibition in binaural sound localization in mammals. *Physiological Research*, 60(3). Electronic publication released ahead of press. **IF (2008) 1.430.**

Stochastic model explains role of excitation and inhibition in binaural sound localization in mammals

M. DRAPAL¹, P. MARSALEK^{1,2,3,4}

¹ Department of Pathological Physiology, First Medical Faculty, Charles University of Prague, Czech Republic

² Faculty of Biomedical Engineering, Czech Technical University of Prague, Kladno, Czech Republic

³ Institute of Physiology, Academy of Sciences of the Czech Republic, Prague, Czech Republic

⁴ Max Planck Institute for the Physics of Complex Systems, Dresden, Germany

Corresponding author:

Petr Marsalek, Max Planck Institute for the Physics of Complex Systems,
Noethnitzer Str. 38, D-011 87, Dresden, Germany.

Tel: +49-351 871 1217 and +420-910 806 204. E-mail: marsalek@pks.mpg.de

Submitted December 30, 2009

Revision of June 15, 2010

Second revision of August 23, 2010

Short title: Excitation and inhibition in binaural sound localization

Key words: Coincidence detection, Directional hearing, Interaural time delay, Medial superior olive, Stochastic neuronal model.

Summary

Interaural time differences (ITDs), the differences of arrival time of the sound at the two ears, provide a major cue for low-frequency sound localization in the horizontal plane. The first nucleus involved in the computation of ITDs is the medial superior olive (MSO). We model the neural circuit of the MSO using a stochastic description of spike timing. The inputs to the circuit are stochastic spike trains with a spike timing distribution described by a given probability density function (beta density). The outputs of the circuit reproduce the empirical firing rates found in experiment in response to the varying ITD. The outputs of the computational model are calculated numerically and these numerical simulations are also supported by analytical calculations. We formulate a simple hypothesis concerning how sound localization works in mammals. According to this hypothesis, there is no array of delay lines as in the Jeffress' model, but instead the inhibitory input is shifted in time as a whole. This is consistent with experimental observations in mammals.

1 Introduction

The auditory system of humans and many other animals is able to localize sound sources with amazing precision. This ability is partially possible with only one ear (monaural hearing), yet for localization in the horizontal plane two ears are necessary (binaural hearing). Sound source localization can be enhanced when the source and the receiver move relative to each other (Phillips and Brugge 1985). In this paper, however, we limit the task of sound localization to static sources using binaural cues at low frequencies. Binaural cues determine the azimuth defined as follows. The vector from a listener to a sound source is projected perpendicularly onto the horizontal plane. The angle between the projected vector and a reference vector, forming the intersection of the horizontal plane with the plane of head symmetry, oriented to the front, is called the azimuth. One of the parameters influencing binaural sound source localization is the fundamental sound

frequency. In mammals, for low fundamental frequencies (below 1,500 Hz) or for broadband sounds, the interaural time difference (ITD) is the dominant sound localization cue. For high fundamental frequencies, the interaural intensity difference (IID) is used. While this paper deals with lower frequency bands, it is possible that higher frequencies are processed with the use of similar neuronal algorithms, as we proposed in earlier studies by Marsalek and Kofranek (2004 and 2005).

This paper presents a theory of how binaural sound localization for low frequencies might be realized in mammals and particularly in humans. The theory of Jeffress (1948) is one of the first well-known attempts to explain how neuronal circuitry achieves this. His prescient work is still frequently cited (Joris *et al.* 1998). Jeffress' visionary hypothesis asserted that the ITD is converted to a binary signal in a higher order neuron through an array of delay lines of fibers in lower order neurons from both sides. Pioneering experiments by Carr and Konishi (1988) showed that Jeffress was correct in case of birds. As far as we know, the existence of an analogous delay line in mammals remains an open question (Grothe 2003, Joris and Yin 2007, McAlpine and Grothe 2003). What other neural circuit mechanism might be responsible for calculating the azimuth from the ITD? In this paper we propose an alternative to the delay line array model based on recent physiological evidence. This alternative is a stochastic delay of a very small number of broadly tuned channels (McAlpine and Grothe 2003).

The amazing time precision (Joris *et al.* 1998) in the range of tens of microsecond points towards another statement of Jeffress that the neurons of the circuit should be located among the lower order neurons of the auditory pathway. The lowest order suitable neuron is the first binaural neuron.

The information about the sound source location contained in the ITD is implicitly encoded by spike trains of lower order neurons. The first binaural neurons function as encoders of the ITD. The circuit has to make the information accessible, in other words make it explicit within another spike train in higher level neurons of the auditory pathway. The function of the circuit is to convert the information implicit in the ITD into the explicit neural code for the ITD. The definition of implicit and explicit neural coding can

be found in Koch (2004). The binaural neurons of the circuit can function either as a delay line (Jeffress 1948), or as broadly tuned channels (McAlpine and Grothe 2003).

Our model is based on one time delay in the neuronal circuit. The access to a continuum of responses to various azimuth locations is accomplished through stochastic variations of action potential times as processed by the model circuit. Our novel finding demonstrates that stochastic spike timing can be used by neurons as an instrument for computing the sound azimuth. Model circuit connections and properties after mathematical simplification of their connectivity are still consistent with the neuro-anatomical description of the wiring of the medial superior olive (MSO) circuit in mammals (Beckius *et al.* 1999, Young 1998, McAlpine *et al.* 2001).

We have designed and improved a model description of how the neural circuit in the auditory brain stem calculates the direction of incoming sound. This model is an alternative to the classical theory of delay lines. We present a stochastic description of the output spike train and spike timing within the model. Both the analytical calculations and numerical simulations give qualitatively similar results to those of experimental recordings from the rodent auditory brain stem. Our results are also comparable with recordings from brainstems in gerbils by Brand *et al.* (2002). Other authors (McAlpine *et al.* 2001) have found similar tuning curves in response to the changing ITD in different (higher) neurons of the auditory pathway (colliculus inferior) of another animal (guinea pig). We find the results to be robust with respect to variations of the time window size and spike timing jitter.

2 Model

2.1 Anatomical connections and their simplification

The notation of the mathematical formulation of the model follows conventions used in Marsalek and Lansky (2005) and Marsalek and Drapal (2008), where the excitatory-excitatory (EE) interaction is called excitatory coincidence detection (ECD) and the excitatory-inhibitory (EI) interaction is called inhibitory coincidence detection

(ICD). The medial superior olive (MSO) works mostly with low frequencies and the lateral superior olive (LSO) deals mostly with high frequencies.

Our model assumes different connections to the MSO neurons than those assumed in the Jeffress model. The division of neuronal fibers based on their excitatory and inhibitory effect is important. We show in the Results section that inhibitory fibers phase shift the tuning curve of ITD, as compared to that obtained without inhibition. The inhibitory connection to the MSO results from the inversion of synaptic polarity in the medial nucleus of the trapezoid body. The MNTB receives excitatory inputs from the contra-lateral cochlear nuclei and sends inhibitory inputs to the ipsi-lateral MSO. The same MSO also receives an inhibitory connection from the ipsi-lateral side. This inhibitory activity originates from the lateral nucleus of trapezoid body (LNTB), which receives its excitation from the ipsi-lateral cochlear nucleus. To complete the picture, the MNTB sends further inhibitory connections to the LSO. This intricate anatomy is nicely summarized by Young (1998).

Though the design of our model is based on knowledge of these anatomical connections, we have to simplify the model wiring to extract the functional core of the neural circuit. The delays in the real system are present in both ipsi- and contra-lateral pathways (Joris 1996, Beckius *et al.* 1999). For the purpose of simplification, however, relative delay on one side suffices. This relative delay represents the net delay difference. Furthermore, one inhibitory branch from one side is enough to model the net inhibition from both sides. This leads us to the schematics shown in Figure 1.

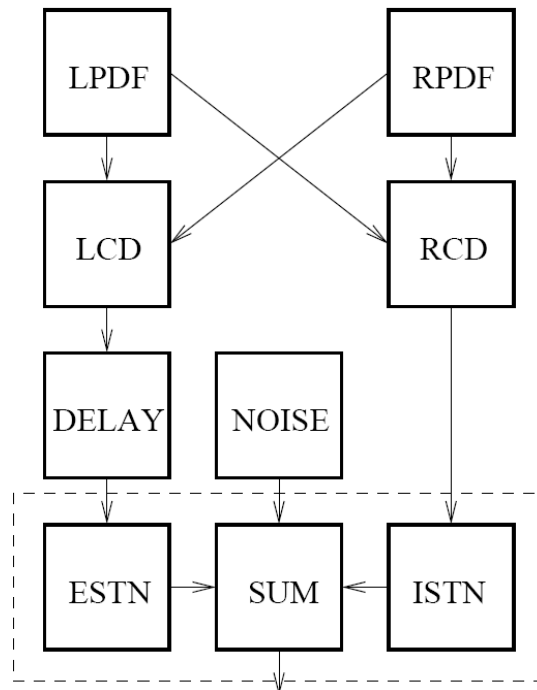


Figure 1: **Flow chart of the neuronal circuitry.** (LPDF) and (RPDF) represent input spikes on left and right sides. Left (Right) Probability Density Function in time governs spike times in the last monaural neuron. (LCD) and (RCD) are Left and Right Coincidence Detectors, realized by the first binaural neurons on the ipsi- and contra-lateral sides, respectively. Next is the (DELAY), which is shown here at the excitatory branch of the circuit only, because only the relative delay matters. Spontaneous spiking acts as an additional (NOISE) source. (ISTN) and (ESTN) represent Inhibitory and Excitatory Synaptic Transmissions, respectively. (SUM) adds excitatory inputs, inhibitory inputs and noise together. From the point of view of functionality, the exact sequence of operations (delay, sign change due to the inhibitory synapse, coincidence detection) in the feed forward pathway does not matter. The dashed line box encompasses processing, which can be performed in one binaural neuron. Note that individual boxes in this chart neither necessarily correspond to individual neurons nor represent the description of the mammalian brain stem circuit in the text.

2.2 Operating conditions and constraints

Before proceeding with a formal description of the variables in the model, we should briefly mention the coincidence detector (CD). Without detection of the leading edges of incoming post-synaptic potentials, extraction of a signal from delays in the microsecond range would not be possible. Regardless of which of the two theories we propose, they all must use this microsecond precision. The element detecting the leading edge is called the coincidence detector. The anatomical substrate of the CD is believed to be within the MSO neurons. The location of right and left CDs in our model circuit in Figure 1 does not reflect all the detailed connections found in anatomy, however. The algorithm realized by a single delay in the model circuit is computationally equivalent to the original circuit, regardless of the actual succession of coincidence detectors, delays and the polarity change from excitatory to inhibitory signals. For details see the Discussion section.

We can shuffle the order of delays among selected points in our model without loss of generality. This is based on observation that selected neural operations are commutative. An example of commutative additions of delays is shown in Equation (1) below. The first processing elements of our model are random delays, which have a specific probability density function (PDF) of synaptic input to neuron in time. Since there is chain of delays in both synaptic chains from the left and from the right ear, we can suppose that we have n ipsilateral delays in the ipsilateral (A) branch of the pathway $\Delta_{A1}, \Delta_{A2}, \dots, \Delta_{An}$ and m contralateral delays in the contralateral (B) branch of the pathway $\Delta_{B1}, \Delta_{B2}, \dots, \Delta_{Bm}$. For the EE interaction $m = 3$ and $n = 3$, for the EI interaction of the MNTB $m = 4$ and $n = 3$ and for the EI interaction of the LNTB $m = 3$ and $n = 4$. Numbers m and n include the first synapses (ribbon synapses from hair cells). The total delay difference between branches (A) and (B), denoted $\Delta_{JA} - \Delta_{JB}$ is given by:

$$\Delta_{JA} - \Delta_{JB} = \sum_{i=1}^n \Delta_{Ai} - \sum_{j=1}^m \Delta_{Bj}. \quad (1)$$

We assume that all these random delays with (timing) jitter (subscript J) on sides A and B (left and right) are mutually independent and identically distributed non-negative random

variables. Δ_{JA} and Δ_{JB} have a maximum of Δ_{max} . The constraints imposed on them are given in Equation (2). The coincidence detection (time) window Δ_w must be shorter than or equal to the maximum delay and the sound period, T , must be greater than or equal to the maximum delay:

$$\begin{aligned} 0 &\leq \Delta_{JA} \leq \Delta_{max}, \\ 0 &\leq \Delta_{JB} \leq \Delta_{max}, \\ 0 &\leq \Delta_w \leq \Delta_{max}, \\ \Delta_{max} &\leq T. \end{aligned} \tag{2}$$

In excitatory coincidence detection, the spike is generated only when the two spikes from sides (A) and (B) meet in a time interval shorter than Δ_w . In other words, the two spike delays Δ_{JA} and Δ_{JB} must satisfy:

$$|\Delta_{JA} - \Delta_{JB}| \leq \Delta_w. \tag{3}$$

To model inhibitory coincidence detection, a modification of the condition expressed above in Equation (3) is used. Spikes must arrive in proper succession. The excitation from side A must come after the inhibition from side B. This is formulated as:

$$0 \leq \Delta_{JA} - \Delta_{JB} \leq \Delta_w. \tag{4}$$

Using the model with one delay, one obtains the continuum of detected ITDs through stochastic variation of the random interaural time delay $\Delta_{JA} - \Delta_{JB}$. The output, the spike rate, clearly depends on the choice of input PDF of the synaptic delays. After summarizing the properties of the model, connections between neurons, and constraints imposed on the parameters and random variables, we can turn to the outputs of the model in the following two subsections.

2.3 Input distribution of the coincidence detector

The output of the model is dependent on the proper choice of a PDF of random variables.

Firstly, the range of the PDF is defined such that its support is over one sound period. This is on the time interval $[0, T]$, which we normalize to the interval $[0, 1]$. Therefore probability densities spanning one or both tails from minus infinity to plus infinity, such as gamma density, where its support is on $[0, \infty)$, or normal density, where

its support is on $(-\infty, \infty)$, are not particularly useful. This justifies the choice of beta density, which is nonzero only within the range $[0, 1]$ and is close to gamma density in this range as well. This circumvents the need to normalize the corresponding cumulative distribution function to unity and makes the calculation more transparent without loss of generality. Another useful property of beta density is its simple polynomial definition. The formula for beta density appears in the Abbreviations and symbols section at the end of this paper. We used beta density for the description of spike timing distribution also in (Marsalek *et al.* 1997).

Secondly, the PDF shape, specifically the skewness and the kurtosis, influences the shape of the output tuning curve. We have experimented with both uniform and triangular densities, (Marsalek and Lansky 2005), which are special trivial cases of the beta density with parameters $a = 1, b = 1$, and $a = 1, b = 2$, respectively, and make calculation simpler. However, the corresponding output functions analogous to Equation (7) are not as satisfactory as the output resulting from a non-trivial beta density. These output functions are not shown here and the resulting function itself is discussed in detail at the beginning of the Results section.

Thirdly, the mean (output) activity of the model obtained with inhibition must not drop below the zero line. This is corrected by adding half the height of the span of the output range to the function. This way all output values are positive. They correspond to the neural spike rates which cannot be negative. After eventual normalization, so that the density integral over the whole function range equals unity, the output spike rates express output probability.

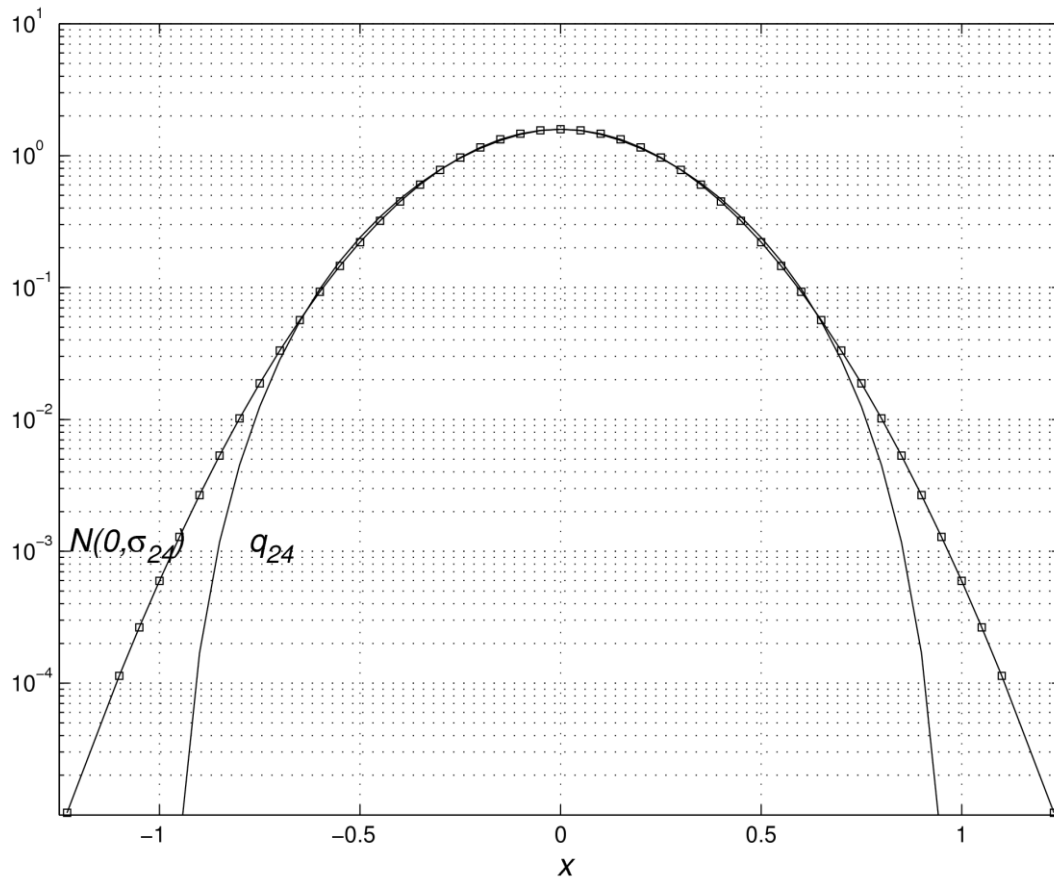


Figure 2: **Resulting probability density function q_{24} .** We use semi-logarithmic scale in the y-axis. On a linear scale (not shown) the resulting function q_{24} cannot be separated from the probability density of the normal distribution $N(0, \sigma_{24})$, where its standard deviation σ_{24} was chosen to match standard deviation of the function q_{24} . The semi-logarithmic scale used here clearly shows the divergence of the two tails of the normal and of the resulting density, when the tail of the normal density $N(0, \sigma_{24})$, marked with squares, is nonzero outside of interval $[-1, 1]$. The other function q_{24} possesses two discontinuities at points $q_{24}(-1) = 0$ and $q_{24}(1) = 0$. On the semi-logarithmic scale in y-axis these cannot be shown. At these points, the tails of the q_{24} function are cut off by the impulse (Heaviside) function. Semi-logarithmic plotting on the y-axis of the PDF shows normal density as parabolic in these coordinates.

Finally, the input beta density has parameters $a = 2$ and $b = 4$ and we denote it B_{24} . To obtain a smooth function on the interval $(0, 1)$ we must have $a > 1$ and $b > 1$, and to obtain nonzero skewness we must have $a \neq b$. The whole numbers $a = 2$ and $b = 4$ are the second smallest non-trivial values of parameters (after $a = 2$ and $b = 3$). Figure 2 shows the result of the analytical calculation of the output, which is the function denoted q_{24} .

2.4 Output distribution from the coincidence detector

The output function q_{24} is obtained as follows. Let us denote the difference of the two delays in Equation (1) as:

$$Z = \Delta_{JA} - \Delta_{JB}. \quad (5)$$

The probability density function of this new, compound random variable Z is obtained with a convolution integral formula for the difference of the two random variables. The PDFs of the Δ_{JA} and the Δ_{JB} are denoted $f(x)$ and $g(y)$, respectively, and the output PDF of the variable Z is denoted $q(z)$. We substitute $f = B_{24}$ and $g = B_{42}$ into the convolution formula $q(z) = \int_{-\infty}^{\infty} f(x)g(x-z)dx$. Now, since $B_{24}(x)$ and $B_{42}(y)$ are nonzero only for x between 0 and 1, we must evaluate the integral piecewise within the respective ranges, such that $q(z)$ becomes:

$$q(z) = \int_{\max(z-1,0)}^{\min(z,1)} f(x)g(z-x)dx = \int_{z \in [-1,0]} f(x)g(z-x) + \int_{z \in [0,1]} f(x)g(z-x). \quad (6)$$

Obviously, $q(z)$ is an even function, satisfying $q(z) = q(-z)$. Therefore, we can change the sign of the argument $z-x$ without loss of validity. We substitute the polynomial densities $B_{24}(x)$ and $B_{42}(y)$ into the integrals. For two 4th degree polynomials, B_{24} and B_{42} , we obtain the sum of two 9th degree polynomials in two variables x and z . We use the Symbolic Math Toolbox in the Matlab software to avoid tedious manual computation.

The source code for the symbolic calculation in the Matlab script language is available upon request.

In summary, using Equation (3) and assuming that Δ_{JA} and Δ_{JB} are distributed with beta densities $B_{24}(x)$ and $B_{42}(y)$ we find that $Z = \Delta_{JA} - \Delta_{JB}$ has density:

$$q_{24}(z) = \sum_{i=0}^{i=9} s_i c_i z^i, \quad (7)$$

where $S = [s_0; s_1; s_2; s_3; s_4; s_5; s_6; s_7; s_8; s_9]$ takes the following values:

$$\begin{aligned} S &= [1; 0; -1; -1; -1; 0; 0; -1; 0; 1] \quad \text{for } x \in [-1; 0], \\ S &= [1; 0; -1; 1; -1; 0; 0; 1; 0; -1] \quad \text{for } x \in [0; 1], \text{ and} \\ S &= [0; 0; 0; 0; 0; 0; 0; 0; 0; 0] \quad \text{otherwise.} \end{aligned} \quad (8)$$

By this prescription, $q_{24}(z)$ is a smooth, piecewise polynomial function. It has four parts defined consecutively on intervals $(-\infty, -1]$, $[-1, 0]$, $[0, 1]$ and $[1, \infty)$. In defining the coefficients $s_i c_i$ of these four parts, we use signum function S valued -1, 0 and 1 listed above in Equation (8). The numerical absolute values of the polynomial coefficients are $c_0 = 1.5873$, $c_2 = 17.1429$, $c_3 = 33.33$, $c_4 = 20$, $c_7 = 2.8571$, $c_9 = 0.6349$ and $c_1 = c_5 = c_6 = c_8 = 0$. Signs of the coefficients confirm that this piecewise polynomial function is an even function $q(z) = q(-z)$. We also recently found an alternative way to obtain the analytical results. A more straightforward way to calculate the indefinite integral $q(z)$ makes use of the Laplace transform. In using the transform, we obtain the convolution by using the inverse transform of the product of the two function images (Drapal and Marsalek 2010). We prefer to show here our original piecewise evaluation, because this integral evaluation method gives more insight into how the output density behaves and how it is obtained. The output function is similar to the normal density, although close inspection reveals subtle differences. Comparison of the output and normal probability density functions is shown in Figure 2. Note the difference between the two tails.

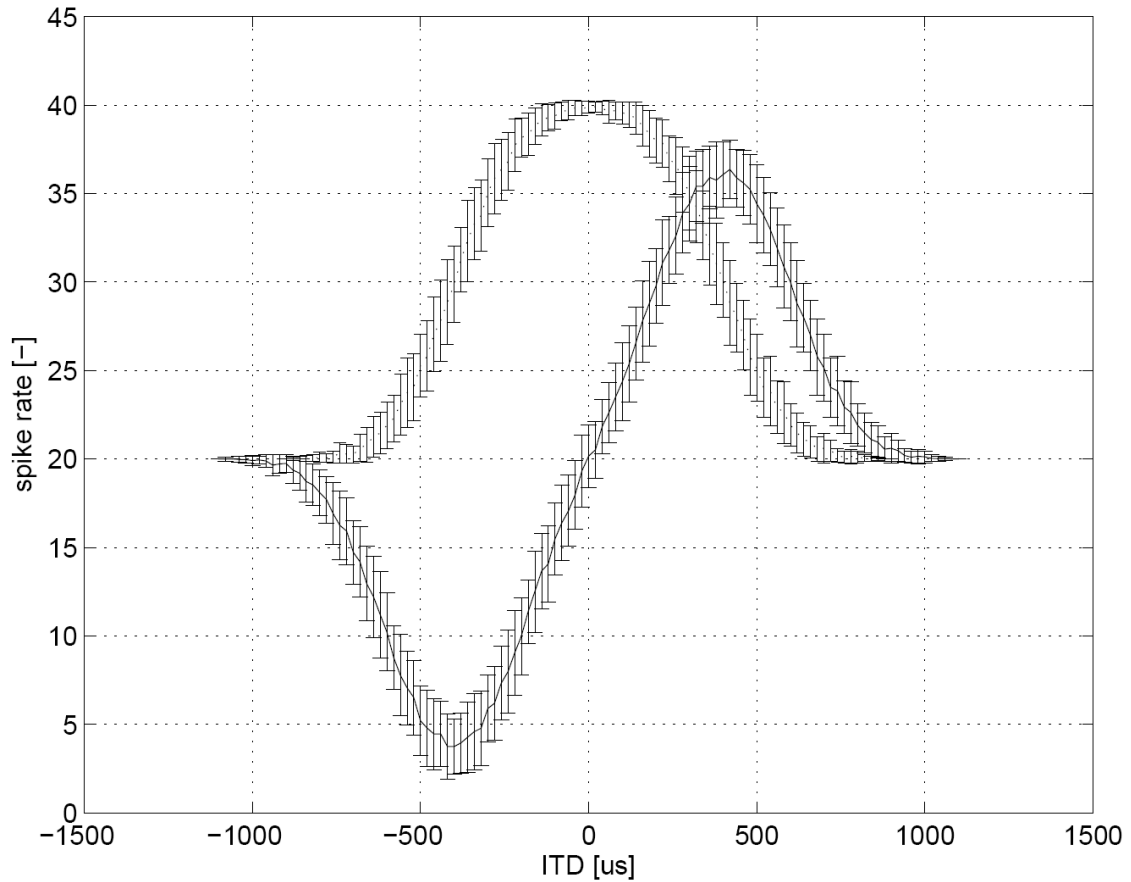


Figure 3: **Model response to pure tone stimulation.** ITD is on the x-axis in microseconds. Neuronal output activity is on the y-axis. The solid curve is the output of the full model with both excitation and inhibition. The dotted curve is the model without inhibition. Only one period of the output is shown here for clarity, but the output is periodic, so the x-axis maps to interaural phase difference. The error bars are sample standard deviations obtained numerically in simulations. In both curves they are the result of 100 trials. In other words, these error bars do not represent the level of the noise in the system. Instead, the noise is introduced into the system via the randomness of random variables and the magnitude of timing jitter.

3 Results

Figure 3 shows the neural response curves of the MSO neuron with varying ITD. The curve in Figure 3 with the maximum at $\text{ITD} = 0$ forms the output of the circuit with only the excitatory branch plugged in. This corresponds to the output of the MSO in experiments, when the inhibitory branch is blocked by the application of strychnine, as in Brand *et al.* (2002). The other curve in Figure 3 with the maximal slope at $\text{ITD} = 0$ shows the output of the full model. This corresponds to the normal, control recordings from the MSO. Brand *et al.* (2002) also modeled some properties of the recorded cells using a detailed biophysical model with explicit representations of the voltage sensitive ion channels of the neuron, obtaining results very similar to those presented here.

Figures 3, 4, and 5 show results of numerical simulations of the model, which are in agreement with the analytical calculation. Both analytical and numerical computations were performed using Matlab software. Sound is represented as discrete samples of pure tones digitized using standard sound digitization (44 kHz sampling rate, 128 intensity levels). Spikes in response to pure tones were triggered by the leading edges of sound waves. In Drapal and Marsalek (2010) we used clicks and complex sounds, such as speech, processed by an cochlear implant emulator, with no qualitative difference from pure tone stimulus (not shown). Here we use pure tones only. The sound period, T , used in the model is that of the fundamental sound frequency. Spike trains were represented as trains of unitary events without any details of action potential shape. Synaptic integration was implemented with the use of time windows. Spontaneous activity is not a free parameter, but is added to the system to obtain positive values of responses. The activity in this and subsequent figures is a dimensionless quantity, which can be interpreted as the spike rate multiplied by a constant.

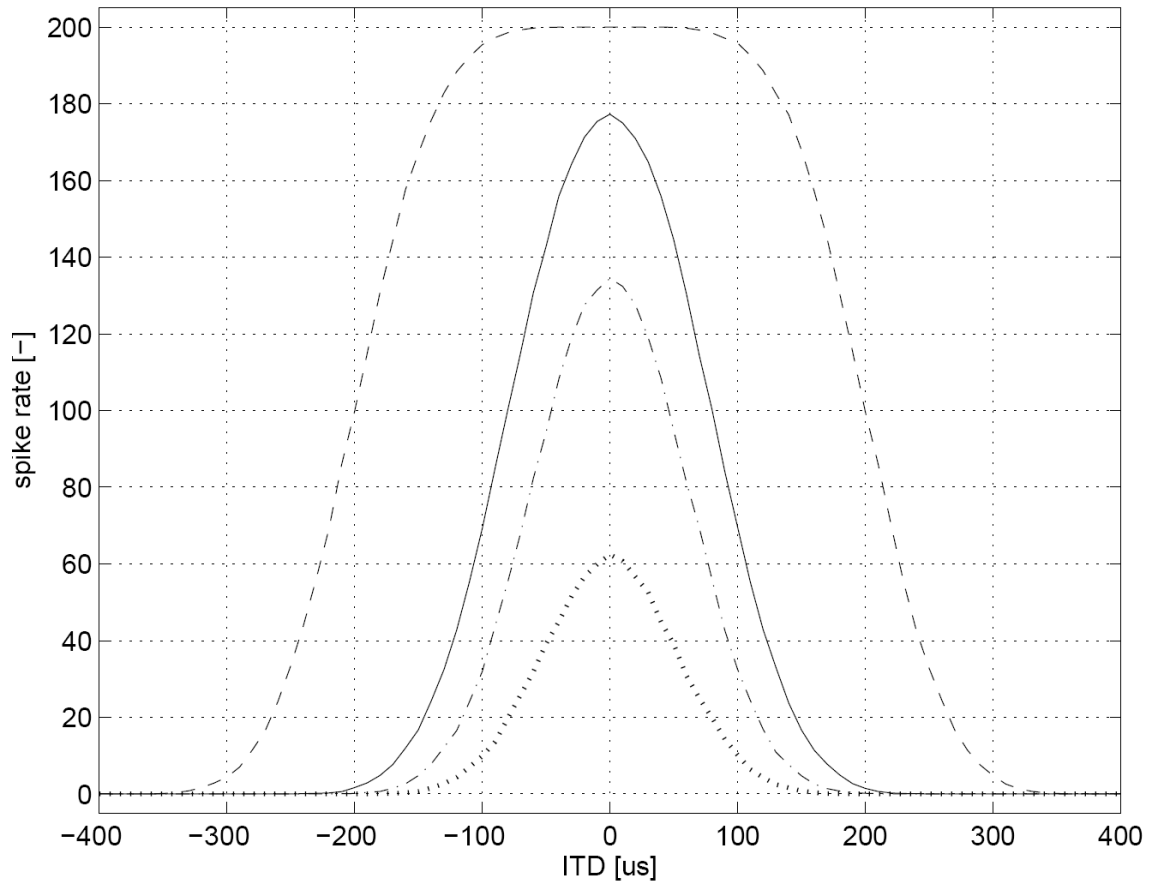


Figure 4: **Effect of the size of the time window.** The width of the time windows for CDs sets the range of ITDs detectable by the circuit. Four widths are compared here: $\Delta w = 200 \mu\text{s}$ (--), $80 \mu\text{s}$ (-), $50 \mu\text{s}$ (-.) and $20 \mu\text{s}$ (...). Plotted output activity is dependent on ITD, as in Figure 3. Only the excitatory part is active in this plot. From an engineering point of view, an optimal curve uses the full dynamic range of available output activities, $\Delta w = 50 \mu\text{s}$.

The size of the time window Δ_w is the key parameter of coincidence detection. Figure 4 shows the model output function for $\Delta_w = 200, 80, 50$ and $20 \mu\text{s}$. In Figure 4 only the excitatory curves are shown, since the inhibitory curves in the full model are analogous to those in Figures 3 and 5. The size of this window in the real neuronal circuit is set by the ionic currents giving rise to the generation of postsynaptic potentials (Svirskis *et al.* 2003, Szalisznyo 2006). Oertel *et al.* (2000) give the upper estimate of the window size for neurons in the auditory pathway specialized in coincidence detection. They give values in the range lesser than 300-1000 μs , with the smaller values for *in vitro* and larger values for *in vivo* preparations.

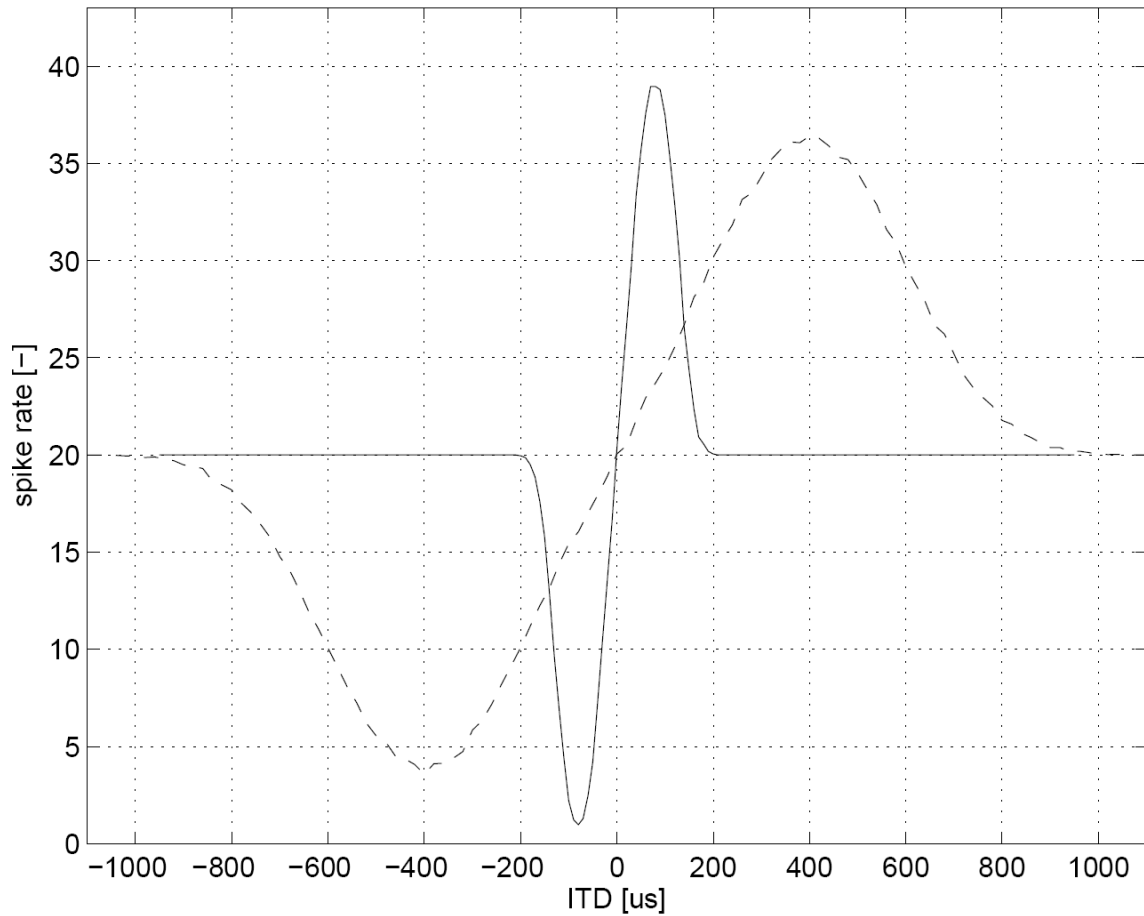


Figure 5: **Effect of the timing jitter magnitude.** Two magnitudes of timing jitter Δ_{JA} and Δ_{JB} are shown here, $100 \mu\text{s}$ as a solid line (-) and $400 \mu\text{s}$ as a broken line (--). They also set the slope of the model output curve and thus the range of ITDs detected. Together with the previous Figure 4, this shows the robustness of the model with respect to variation of parameters. Figures 4 and 5 illustrate the range of parameter values accessible to the model.

In general, the magnitude of timing jitter (denoted Δ_{JA} and Δ_{JB}) in a neuronal nucleus is dependent on the degree of neuronal convergence in the previous stages of processing. In spike trains propagating to a higher order nucleus, one of three jitter changes may occur: the wave of spikes can sharpen (the jitter decreases), the jitter may not change, or the spike volley can become blurred (the jitter increases). All three variants can be obtained with the perfect integrator neuronal model in the regime of coincidence detection, for different sets of parameters (Reed *et al.* 2002). All three variants were shown to exist in experimental recordings (Gerstner *et al.* 1996, Joris 1996, Marsalek *et al.* 1996). We do not discuss these differences further here. The output activity shown in Figure 5 is dependent on ITD for two different magnitudes of timing jitter. Figure 5 shows two functions of the full model only (with both excitation and inhibition) for two values of timing jitter, 100 μ s and 400 μ s. Oertel *et al.* (2000) give jitter values of 200 μ s for *in vivo* and values of 20-40 μ s for *in vitro* preparations.

A related question was raised in review: How many ipsi- and how many contra-lateral synaptic connections (and of what proportion of excitatory and inhibitory connections) would a real MSO neuron need in order to achieve the accuracy observed physiologically in these neurons? Probabilistic spiking transmits information only with a given probability (reliability). In a previous investigation we calculated the time to spike required by the coincidence detection circuit for two given probabilities of achieving the decisive spike (reliabilities), $p_1 = 50\%$ and $p_2 = 95\%$ (Marsalek and Lansky, 2005). These probabilities were calculated based on the number, K , of unitary events. The calculation procedure, known as a Bernoulli process, is described in the 2005 paper, however, the values of K are not tabulated in the paper, since they are only intermediate results of the calculation. For the whole range of sound frequencies relevant for the human MSO, the number of events, K , attained values from 1 up to 50, higher values for higher frequencies. However, this number of unitary events can be achieved either by waiting for N sound periods, or by parallel processing by several neurons. In the latter case, the MSO nucleus must consist of at least M copies of the circuit, but may even have N tonotopic channels, such that $MN > K$. The numbers K , M , N represent the minimal

values required for a functioning circuit. Numbers of neurons and connections in the MSO are probably higher, since most nuclei show a redundancy in the number of their neurons and connections.

4 Discussion

Binaural sound localization is achieved with remarkable precision throughout the animal kingdom. The timing precision of individual spikes in most neurons in the auditory pathways of various species is lower than animals' behavioral assessment of the ITD. Given that this precision is important for survival and easily demonstrated, the means by which it is achieved by the neural circuits forms a fascinating and as yet unresolved question. In this paper we have presented a stochastic model employing both excitatory and inhibitory synaptic inputs to address this question in the neural circuit of the mammalian medial superior olive.

In recent years it has been demonstrated in experiments on gerbils (Brand *et al.* 2002) and also on other mammals (Grothe 2003) that synaptic inhibition plays a critical role in the sound localization circuit. The original theory of delay lines array by Jeffress applies in birds (Carr and Konishi 1988). It is not clear whether it also applies in mammals. Recordings in cats do not show the same ITD tuning curve slopes as those in rodent recordings (Joris and Yin 2007, Yin and Chan 1990). Harper and McAlpine (2004) present a theoretical explanation of these differences, including the discussion of human data. Psychophysical experiments studying the circuit in humans using subjective response might resolve this question in the future. These experiments range from normal hearing (Middlebrooks and Green 1991), to electrical hearing sense in cochlear implantees (Laback and Majdak 2008). Of course it is possible that the mammalian circuit uses some entirely different mechanism to those presented here and in previous investigations.

Another observation in the abovementioned experiments shows that the maximal response does not occur at the best ITD, but that the best ITD occurs where the slope of the response curve is maximal. As early as the 1970s and 1980s, some authors discussed

the possibility that the maximal slope of the response of a coincidence detector measuring the ITD between spike trains may relay the information (Goldberg and Brown 1969, Phillips and Brugge 1985). The possibilities to test this proposal numerically were limited at that time. Here we exploit the advancement of computational tools in a numerical study of the problem.

Most neurons do not achieve high spike timing precision, but in neurons in the periphery, the information about timing must be somehow encoded and preserved before reaching the first binaural neurons. Such timing precision is enabled by cellular and sub-cellular mechanisms and is best studied using deterministic equations, specifically differential equations. One can compare the work of Svirskis *et al.* (2003), which studies the state space of the nonlinear dynamics of ion currents in the MSO neurons, with that of (Szaliszno 2006), in which the state space of the nonlinear ion currents in the LSO is studied. The authors of these two papers also performed experimental recordings. In both the circuits of the MSO and the LSO, the nonlinearities are necessary for the proper function of coincidence detectors. Other models (including this work) are based on stochastic spiking, or cast the neural circuit in terms of a logical circuit. See for example a review of Colburn (1996), or (Marsalek 2000). Experimental investigations can also be supported by conceptual models (Brand *et al.* 2002). It is argued frequently that the ITD tuning curve slope offset is due to the optimized wiring of the circuit (Grothe 2003, Harper and McAlpine 2004).

Most theories which include inhibitory mechanisms claim that inhibition is precisely timed and brief. The models assume high precision of spike times in all neurons in the circuit. Not all spikes are so precise, as is shown in experiments of Batra *et al.* (1997) and Joris (1996). The spatial organization of the neurons themselves is important. Agmon-Snir *et al.* (1998) and Zhou *et al.* (2005) suggest models based on a spatial organization of neurons in the MSO.

The parameter space of the binaural model is limited at higher frequencies, where ambiguities arise due to the relatively short sound wavelength in comparison with the ITD. In the mammalian brainstem, the circuit of the MSO processes the low frequencies

and another twin circuit, the LSO, processes the high frequencies. The two circuits are developed in parallel in human, cat, dog and other experimental species although in some mammals only one of them is present (Grothe 2003). The relative importance of the ITD and the IID within these respective sound frequency ranges in human was elucidated by means of confusing these two cues in psychophysical experiments (Wightman and Kistler 1992). Breebaart *et al.* (2001a and 2001b) brought a wider perspective towards the possible mechanisms used in binaural processing by using the excitatory-inhibitory mechanism where previous models dealt only with the excitatory-excitatory interaction. Their model uses the canonical structure of a grid of parallel tonotopic pathways together with delay lines and covers both the low frequency (ITD) and high frequency (IID) ranges with similar mechanisms. The model of Breebaart *et al.* (2001) comprises of building blocks of the signal processing circuits. Colburn (1996) and his numerous collaborators developed a series of binaural circuit models of varying complexity. They stress the importance of the coincidence detector within the circuit. Stern and Trahiotis (1996) review existing models and their own circuit implementations are close to delay line concepts.

Our model differs from those mentioned above in that it employs randomness in the spike arrival time between synapses. Random delay and spike timing jitter might seem like an impediment, especially in models with precisely timed inhibition. The statistical properties of spike trains average out errors in individual spikes and enable the neural computation of azimuth at the same time.

When two spikes from opposite sides arrive at the first binaural neuron, their coincidence must be detected with higher precision. Jeffress (1948) was first to notice this. All subsequent sound localization models have to assume that these particular neurons are coincidence detectors. Neurons as coincidence detectors have been used frequently in sound localization models (Gerstner *et al.* 1996, Marsalek 2000, Zhou *et al.* 2005). Let us also mention the classical model describing neural coincidence detectors in visual circuits in invertebrates, starting with the Reichardt model (Srinivasan and Bernard 1976, Zanker *et al.* 1999).

The first binaural neuron must use coincidence detection to extract information from interaural sound timing disparity. This neuron must indeed be precise as a coincidence detector. Such neuron outputs spike only if left and right inputs coincide within a short time window, in the time range of microseconds. In our model, stochastic properties make use of timing jitter, which at first glance is merely signal deteriorating noise.

From the point of view of implementation of the model, it does not matter exactly where the net ITD is placed in the two (left and right side) converging branches of the neural pathway. If an additional delay is added in both sides, it cancels out. The net delay between the two sides only matters when compared with the ITD, as written in Equation (1).

In a follow-up to this paper (Drapal and Marsalek 2010), we show that our numerical model can be connected to cochlear implant software emulator to give similar results to those observed by psychophysical methods in implanted volunteers (Laback and Majdak, 2008). As the possibilities to experiment with binaurally implanted and bimodal hearing subjects expand, it is possible that some psychophysical experiment will determine, whether delay line, stochastic delay, or both mechanisms are used in human hearing, even before definitive electrophysiological experiments on mammals are performed (Joris and Yin 2007).

In this paper we present a model for a neural algorithm performed by a circuit in the MSO. This is part of a general quest to capture the multitude of neural algorithms serving specific purposes. Let us give two closing examples of these algorithms. One is the case of spatial maps in the auditory brainstem of birds (Peña 2003). Peña (2003) shows how the brainstem circuit implements spike rate multiplication in order to calculate the location of a sound source. An analogous case from the 1980s concerns the neocortical visual circuit in the higher order visual areas which calculates the location of the illusory contour (von der Heydt *et al.* 1984). These authors demonstrate that a higher neocortical visual projection area can respond to a virtual object, an illusory contour, which is a result of a neural algorithm, as if it were a real solid object in a visual scene. Both these models were first proposed as hypotheses of a specific neural computation. The existence of

neural circuits implementing the respective neural computations was subsequently confirmed by a targeted experimental recording.

Abbreviations and symbols

CD – coincidence detection, **ECD** – excitatory coincidence detection, **ICD** – inhibitory coincidence detection, **EE** – excitatory–excitatory interaction, **EI** – excitatory–inhibitory interaction, **IID** – interaural intensity difference, **IPD** – interaural phase difference, **ISD** – interaural spectral difference, **ITD** – interaural time difference, **LSO** – lateral superior olive, **MSO** – medial superior olive, **PDF** – probability density function, Δ_j – time jitter (delay random variable), Δ_w – coincidence detection window, T – sound period and Δ_{max} – maximum delay.

The **beta density** is a probability density function written in a standard form as:

$B_{ab}(x) = x^{b-1}(1-x)^{a-1} B(a,b)^{-1}$, for $x \in [0,1]$, and $B_{ab}(x) = 0$ otherwise, where the parameters $a, b > 0$ and $B(a,b)$ is the (Euler) beta function.

Conflict of Interests

The authors declare no conflict of interests.

Acknowledgments

Thanks to Michal Andrlík, Nick Dorrell, Abigail Klopper, Lubomir Kostal, Benjamin Lindner, Jose Louis Peña, Daniel Šuta and to four anonymous referees. The research was supported by two grants from the Czech Ministry of Education (MSMT), Research Initiatives no. 6840770012 and no. 0021620806.

References

- AGMON-SNIR H, CARR CE, RINZEL, J: The role of dendrites in auditory coincidence detection. *Nature* **393**: 268–272, 1998.
- BATRA R, KUWADA S, FITZPATRICK D: Sensitivity to interaural temporal disparities of low-and high-frequency neurons in the superior olivary complex. II. Coincidence detection. *J Neurophysiol* **78**: 1237–1247, 1997.
- BECKIUS GE, BATRA R, OLIVER DL: Axons from anteroventral cochlear nucleus that terminate in medial superior olive of cat: observations related to delay lines. *J Neurosci* **19**: 3146–3161, 1999.
- BRAND A, BEHREND O, MARQUARDT T, McALPINE D, GROTHE B: Precise inhibition is essential for microsecond interaural time difference coding. *Nature* **417**: 543–547, 2002.
- BREEBAART J, VAN DE PAR S, KOHLRAUSCH A: Binaural processing model based on contralateral inhibition. I. Model structure. *J Acoust Soc Am* **110**: 1074–1088, 2001a.
- BREEBAART J, VAN DE PAR S, KOHLRAUSCH A: Binaural processing model based on contralateral inhibition. II. Dependence on spectral parameters. *J Acoust Soc Am* **110**: 1089–1103, 2001b.
- CARR CE, KONISHI M: Axonal delay lines for time measurement in the owl's brainstem. *Proc Natl Acad Sci USA* **85**: 8311–8315, 1988.
- COLBURN HS: Binaural models. In: *Auditory Computation*. HL HAWKINS, TA McMULLEN, AN POPPER, RR FAY (eds), Springer Verlag, New York, 1996, pp 332-400.
- DRAPAL M, MARSALEK P: Stochastic model shows how cochlear implants process azimuth in real auditory space. *Chinese J Physiol* **53**: In press, 2010.
- GERSTNER W, KEMPTER R, VAN HEMMEN JL, WAGNER H: A neuronal learning rule for sub-millisecond temporal coding. *Nature* **383**: 76–78, 1996.

- GOLDBERG JM, BROWN PB: Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: Some physiological mechanisms of sound localization. *J Neurophysiol* **32**: 613–636, 1969.
- GROTHER B: New roles for synaptic inhibition in sound localization. *Nat Rev Neurosci* **4**: 540–50, 2003.
- HARPER NS, MCALPINE D: Optimal neural population coding of an auditory spatial cue. *Nature* **430**: 682–686, 2004.
- JEFFRESS LA: A place theory of sound localization *J Comp Physiol Psychol* **41**: 35–39, 1948.
- JORIS P, YIN TCT: A matter of time: internal delays in binaural processing. *Trends Neurosci* **30**: 70–78, 2007.
- JORIS PX, SMITH PH, YIN TC: Coincidence detection in the auditory system: 50 years after Jeffress. *Neuron* **21**: 1235–1238, 1998.
- JORIS PX: Envelope coding in the lateral superior olive. II. Characteristic delays and comparison with the responses in the medial superior olive. *J Neurophysiol* **76**: 2137–2156, 1996.
- KOCH C: *The Quest for Consciousness: A Neurobiological Approach*. Roberts and Co, Englewood, 2004.
- LABACK B, MAJDAK P: Binaural jitter improves interaural time–difference sensitivity of cochlear implantees at high pulse rates. *Proc Natl Acad Sci USA* **105**: 814–817, 2008.
- MARSALEK P: Coincidence detection in the Hodgkin-Huxley equations. *Biosystems* **58**: 83–91, 2000.
- MARSALEK P, DRAPAL M: Mechanisms of coincidence detection in the auditory brainstem: Examples. In: *Mathematical Modeling of Biological Systems*, A DEUTSCH, R BRAVO DE LA PARRA, R DE BOER, O DIEKMANN, P JAGERS, E KISDI, M KRETZSCHMAR, P LANSKY, H METZ (eds), vol. 2. Birkhaeuser, Boston, 2008, pp 255–264.

- MARSALEK P, KOCH C, MAUNSELL J: On the relationship between synaptic input and spike output jitter in individual neurons. *Proc Natl Acad Sci USA* **94**: 735-740, 1997.
- MARSALEK P, KOFRANEK J: Sound localization at high frequencies and across the frequency range. *Neurocomputing* **58–60**: 999–1006, 2004.
- MARSALEK P, KOFRANEK J: Spike encoding mechanisms in the sound localization pathway. *Biosystems* **79**: 191–198, 2005.
- MARSALEK P, LANSKY P: Proposed mechanisms for coincidence detection in the auditory brainstem. *Biol Cybern* **92**: 445–451, 2005.
- McALPINE D, GROTHE B: Sound localization and delay lines – do mammals fit the model? *Trends Neurosci* **26**: 347–350, 2003.
- McALPINE D, JIANG D, PALMER A: A neural code for low-frequency sound localization in mammals. *Nat Neurosci* **4**: 396–401, 2001.
- MIDDLEBROOKS J, GREEN D: Sound localization by human listeners. *Ann Rev Psychol* **42**: 135–159, 1991.
- OERTEL D, BAL R, GARDNER SM, SMITH PH, JORIS PX: Detection of synchrony in the activity of auditory nerve fibers by octopus cells of the mammalian cochlear nucleus. *Proc Natl Acad Sci USA* **97**: 11773–11779, 2000.
- PEÑA JL: Binaural processing in the synthesis of auditory spatial receptive fields. *Biol Cybern* **89**: 371–377, 2003.
- PHILLIPS DP, BRUGGE JF: Progress in neurophysiology of sound localization. *Ann Rev Psychol* **36**: 245–274, 1985.
- REED MC, BLUM JJ, MITCHELL CC: Precision of neural timing: Effects of convergence and time-windowing. *J Comput Neurosci* **14**: 35–47, 2002.
- SRINIVASAN MV, BERNARD GD: A proposed mechanism for multiplication of neural signals. *Biol Cybern* **21**: 227–236, 1976.
- STERN RM, TRAHOTIS C: Models of binaural perception. In: *Binaural and Spatial Hearing in Real and Virtual Environments*. R GILKEY, TR ANDERSON (eds), Lawrence Erlbaum Associates, New York, 1996, pp 499-531.

- SVIRSKIS G, DODLA R, RINZEL J: Subthreshold outward currents enhance temporal integration in auditory neurons. *Biol Cybern* **89**: 333–340, 2003.
- SZALISZNYO K: Role of hyperpolarization-activated conductances in the lateral superior olive: a modeling study. *J Comput Neurosci* **20**: 137–152, 2006.
- VON DER HEYDT R, PETERHANS E, BAUMGARTNER G: Illusory contours and cortical neuron responses. *Science* **224**: 1260–1262, 1984.
- WIGHTMAN FL, KISTLER DJ: The dominant role of low-frequency interaural time differences in sound localization. *J Acoust Soc Am* **91**: 1648–1661, 1992.
- YIN T, CHAN J: Interaural time sensitivity in medial superior olive of cat. *J Neurophysiol* **64**: 465–488, 1990.
- YOUNG ED: Cochlear nucleus. In *Synaptic Organization of the Brain, 4th edition*, GM SHEPHERD (ed) Oxford University Press, New York, 1998, pp. 121–157.
- ZANKER JM, SRINIVASAN MV, EGELHAAF M: Speed tuning in elementary motion detectors of the correlation type. *Biol Cybern* **80**: 109–116, 1999.
- ZHOU Y, CARNEY L, COLBURN H: A model for interaural time difference sensitivity in the medial superior olive: interaction of excitatory and inhibitory synaptic inputs, channel dynamics, and cellular morphology. *J Neurosci* **25**: 3046–3058, 2005.

A.3 Neural coding in sound localization pathway, roles of timing jitter and random delay

Marsalek, P. and Drapal, M. (2009). Neural coding in sound localization pathway, roles of timing jitter and random delay. *Neuronal Coding, 8th International Workshop*, pages 105–109.

NEURAL CODING IN SOUND LOCALIZATION PATHWAY, ROLES OF TIMING JITTER AND RANDOM DELAY

Petr Marsalek^{1,2} and Marek Drapal^{1,2}

(1) Faculty of Biomedical Engineering, Czech Technical University
Nam. Sitna 3105, CZ-272 01 Kladno, Czech Republic
marsalek; drapal@fbmi.cvut.cz

(2) Charles University Prague, First Medical Faculty
Department of Pathological Physiology,
U nemocnice 5, CZ-128 53, Praha 2, Czech Republic
Petr.Marsalek; Marek.Drapal@lf1.cuni.cz

ABSTRACT

We present a model of sound localization in mammals. This model is consistent with new experimental findings of how the ITD (interaural time delay) is processed in mammals. In the paper we describe in more detail several key components of our model: neurons acting as coincidence detectors, spike timing jitter, random delays and other parameters, which can be identified with the biophysical properties of auditory neurons. We model the low frequency part of the sound localization circuit. Under the low frequency part we understand that part of animal's hearing range, in which the ITD cue is used. The disparity of the inputs between the two ears is called interaural and the neurons, which have access to the spikes originating from both ears are called binaural. Our model describes the first binaural neurons found when progressing from periphery. There are analogous binaural neurons in birds (neurons of nucleus laminaris) and in mammals (neurons of superior olive), yet they are embedded in different neural circuits [2]. The level of description used in our paper is therefore the level of the neural circuit for the ITD processing.

Once the function of the two circuits is understood sufficiently in experiment, the theoretical description of two differing neural circuits will give us an explanation how the extremely short time differences in the range of tens of microseconds are used to indicate the azimuth of the incoming sound. Enough experimental literature is now available for both these circuits. We can take the experimental description of these neural circuits and use them as test bed for the theories how sound is encoded into spike trains in lower processing stages of the auditory pathway.

Together with and even before the blossom of experimental literature on this topic, theories of how these circuits possibly function were presented and compared to the experimental data. The first theory suggested that the ITD's are processed by a delay line. This was proposed by Jeffress in 1948 [4]. Only several decades later it was demonstrated by experiments on barn owl that a variant of delay line exist in a neural circuit of barn owl and also of other birds [1]. The neural circuit of birds converges on the binaural nucleus laminaris.

Even though several experimentalists attempted to find a structure similar to avian delay line in mammals, the results were not conclusive [5]. We joined the theoretical effort and described necessary and sufficient properties of binaural neurons to make them able to use the information contained in the ITD [7]. New experimental findings started to indicate that a mammalian sound localization circuit can be wired differently than the avian [10]. McAlpine with colleagues also presented several purely theoretical works, amongst them [3], to explain some properties of the ITD coding by neural population. Some of the coding properties are shared by both avian (Figure 1) and mammalian (Figure 2) circuits. Together with Lansky we were able to describe

All Figures: **General legend to all the three figures**

The three figures show activities in percentage of maximal activity (on the y-axis) of three types of binaural neurons of three respective neural circuits (avian in Figure 1, mammalian in Figure 2 and model of mammalian in Figure 3) in response to different values of the ITD in microseconds (on the x-axis). There are three bands for the ITD tuning shown on the figures by three parts of the curves with decreasing heaviness. The thick solid part shows the interval of the ITDs corresponding to the azimuth and based on the head size. The medium dotted part shows the interval which is out of the "straight" ITDs, yet it can still be used for the azimuth calculation. The thin interrupted parts correspond to the continuation of the ITD for the next phase differences due to the continuity of the phase information.

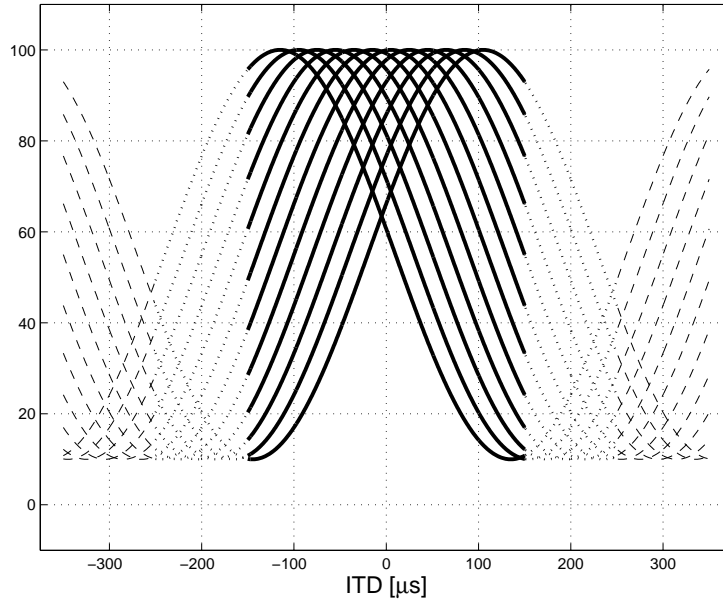


Figure 1: **Schematic responses of binaural neurons in the delay line (laminar nucleus)**

For axes and common details see the general legend above. This figure shows set of twelve representative curves corresponding to set of twelve neurons of the delay line [4] tuned by their maximal response to a particular ITDs, $20 \mu\text{s}$ apart. This tuning corresponds to the position of the neuron within the delay line.

neural operations of the coincidence detection circuits in terms of random spike arrivals and their probability densities [8]. We followed up with more estimates, in particular studying the standard deviation of spike timing, sometimes called spike timing jitter [9]. We also study alternative ways to measure spike timing variability [6].

The implementation details of the model in (Figure 3) will be presented in the paper. Briefly, the latencies are simulated together with the stochastic arrival of spike at the first binaural neuron. We used several modifications of the model. We are now in a process of rigorously showing that some of the model variants are computationally equivalent. We use both inhibitory and excitatory inputs to the binaural neuron. This is consistent with the mammalian literature. Our model makes possible to describe neural coding in lower stages of auditory processing and therefore the model can be used for studying both natural and electronic sensation, the latter evoked by cochlear implant stimulation.

We are capable of recalculating the input to the auditory nerve by a cochlear implant emulator (we use a published and patent free realistic software emulation of cochlear implant). In last decade, many deaf patients in Europe have been implanted in both cochleas by a cochlear

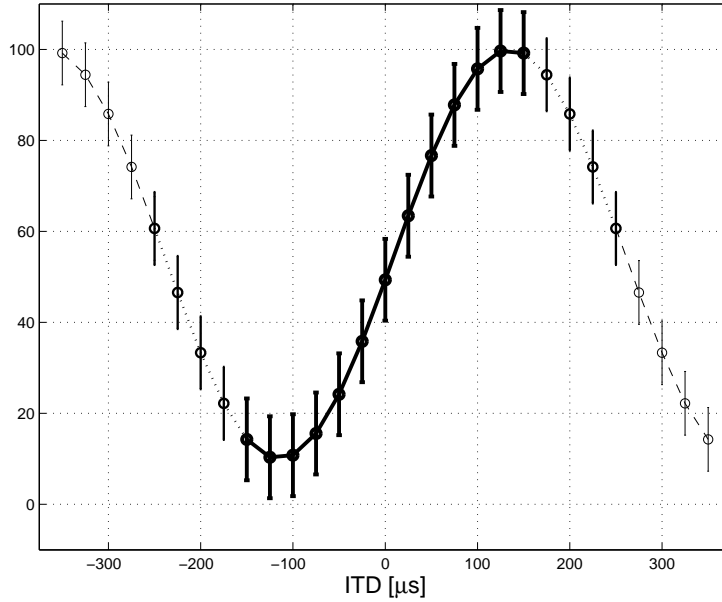


Figure 2: **Schematic responses of binaural neurons in the mammalian circuit (MSO)** For axes and common details see the general legend above. The information about the azimuth is extracted from at least two (maybe there are more than two) tuning curves of broadly tuned delay channels [2]. Only one of the two channels is shown here. Error bars illustrate the stochastic dimension of the tuning: the actual response varies from trial to trial and the exact azimuth is extracted from pooling of the signal from a small set of neurons.

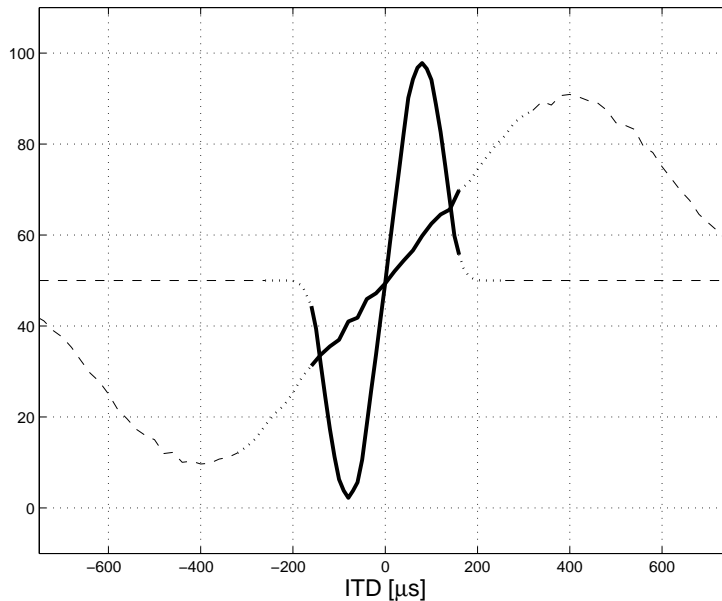


Figure 3: **Responses of model neurons with two sets of parameters (stochastic model)** For axes and common details see the general legend above. This figure shows the effect of varying the stochastic delay and timing jitter in a model of MSO neurons. The higher activity versus ITD slope has timing jitter $100 \mu\text{s}$ and delay $100 \mu\text{s}$ and the lower slope has the values of jitter $600 \mu\text{s}$ and of the delay $400 \mu\text{s}$ [6, 9]. This illustrates the robustness of the model with respect to variation of parameters. Maximal responses of the model were normalized to be compared to previous figures. Note that the ITD range is doubled compared to the previous two figures to show the different spans of the two activity versus ITD slopes.

implant of identical manufacturer. Existing technology enables synchronized stimulation of both cochleas in these patients. Cochlear implants can switch between several coding strategies, some useful for speech understanding and some other useful for open space hearing. It has been shown previously in experiments that there exist a trade-off between sound localization and speech comprehension in these patients. Another finding demonstrates that specific types of spike timing jitter are beneficial in these patients. In our work we demonstrate theoretically, why different types of jitter are useful. Our improvements of our sound localization model may lead to the formulation of novel neural coding strategies for cochlear implants.

Keywords: Coincidence detection, directional hearing, interaural time delay, stochastic neuronal model.

References

- [1] Carr, C. E., Konishi, M. (1988). Axonal delay lines for time measurement in the owl's brainstem. *Proc Natl Acad Sci USA*, **85**, 8311–5.
- [2] Grothe, B. (2003). New roles for synaptic inhibition in sound localization. *Nat Rev Neurosci.*, **4**, 540–550.
- [3] Harper, N. S., McAlpine, D. (2004). Optimal neural population coding of an auditory spatial cue. *Nature*, **430**, 682–686.
- [4] Jeffress, L. A. (1948). A place theory of sound localization. *J. Comp. Physiol. Psychol.*, **41**, 35–39.
- [5] Joris, P., Smith, P., Yin, T. (1998). Coincidence detection in the auditory system: 50 years after Jeffress. *Neuron*, **21**, 1235–1238.
- [6] Kostal L., Marsalek, P. (2009) Analysis of synaptic input and output spike-timing precision by information-theoretic methods. *Neural Coding 2009 Abstract Book*, this issue.
- [7] Marsalek, P. (2000) Coincidence detection in the Hodgkin-Huxley equations. *Biosystems*, **58**, 83–91.
- [8] Marsalek, P., Lansky, P. (2005) Proposed mechanisms for coincidence detection in the auditory brainstem. *Biol Cybern.*, **92**, 445–451.
- [9] Marsalek P., Drapal M. (2007) Mechanisms of Coincidence Detection in the Auditory Brainstem: Examples. in: *Mathematical Modeling of Biological Systems*, vol. II. A. Deutsch, R. Bravo de la Parra, R. de Boer, O. Diekmann, P. Jagers, E. Kisdi, M. Kretzschmar, P. Lansky and H. Metz (eds). Birkhaeuser, Boston, 255-264.
- [10] McAlpine, D., Jiang, D., Palmer, A. R. (2001). A neural code for low-frequency sound localization in mammals. *Nat. Neuroscience*, **4(4)**, 396–401.

A.4 Stochastic model shows how cochlear implants process azimuth in real auditory space

Drapal, M. and Marsalek, P. (2010). Stochastic model shows how cochlear implants process azimuth in real auditory space. *The Chinese Journal of Physiology*, 53(6), pages 439–446. **IF (2008) 0.698.**

Stochastic Model Shows How Cochlear Implants Process Azimuth in Real Auditory Space

Marek Drapal¹ and Petr Marsalek^{1, 2, 3, 4}

¹Charles University of Prague, Department of Pathological Physiology
U nemocnice 5, CZ-128 53, Praha 2, Czech Republic

²Czech Technical University of Prague, Faculty of Biomedical Engineering
Nam. Sitna 3105, CZ-272 01, Kladno, Czech Republic

³Max Planck Institute for the Physics of Complex Systems
Nöthnitzer Str. 38, D-011 87, Dresden, Germany
and

⁴Czech Academy of Sciences, Institute of Physiology
Videnska 1083, CZ-142 20, Praha 4, Czech Republic

Abstract

Interaural time difference (ITD) is a major cue for sound azimuth localization at lower sound frequencies. We review two theories of how the sound localization neural circuit works. One of them proposes labeling of sound direction in the array of delay lines by maximal response of the tuning curve (Jeffress model). The other proposes detection of the direction by calculating the maximum slope of tuning curves. We formulate a simple hypothesis from this that stochastic neural response infers sound direction from this maximum slope, which supports the second theory. We calculate the output spike time density used in the readout of sound direction analytically. We show that the numerical implementation of the model yields results similar to those observed in experiments in mammals. We then go one step further and show that our model also gives similar results when a detailed implementation of the cochlear implant processor and simulation of implant to auditory nerve transduction are used, instead of the simplified model of auditory nerve input. Our results are useful in explaining some recent puzzling observations on the binaural cochlear implantees.

Key Words: coincidence detection, directional hearing, interaural time delay, stochastic neuronal model

Introduction

In previous studies we developed a stochastic model of the sound localization circuit (21-23). The work presented here aims to demonstrate that both simplified simulation of natural sound input and detailed simulation of cochlear implants (CI) yield a similar output in our binaural model. The goal of future work is an explanation of the tradeoff between complex sound recognition and localization observed in bilaterally implanted patients, as well as possible

improvement of the CI stimulation protocols.

We present a model of sound localization in mammals. This model is consistent with new experimental findings of how the ITD (interaural time delay) is processed in mammals. In this paper we describe in more detail several key components of our model: neurons acting as coincidence detectors, spike timing jitter, random delays and other parameters, which can be identified with the biophysical properties of auditory neurons. We model the low frequency part of the sound localization circuit. Under the low

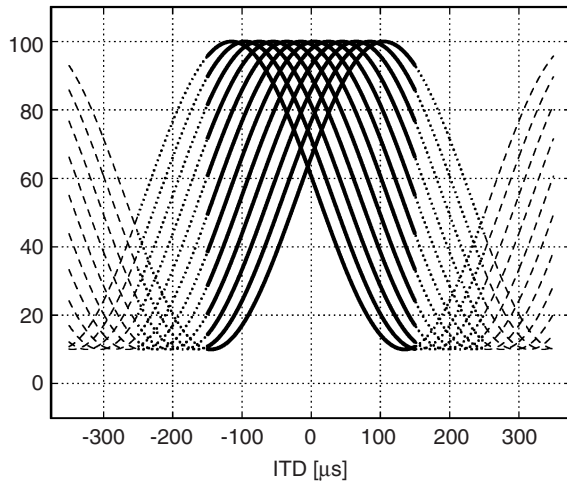


Fig. 1. Schematic responses of binaural neurons in the delay line (laminar nucleus). This and the following Figure show activities in the percentage of maximal activity (on the y-axis) of two respective types of binaural neurons of two respective neural circuits (avian in this Figure and mammalian in the next) in response to different values of the ITD in microseconds (on the x-axis). There are three bands for the ITD tuning corresponding to different line weights. The thick solid lines show the interval of the ITDs corresponding to the azimuth and based on the head size. The dotted lines show the interval outside the ITDs, which can be also used for the azimuth calculation. The dashed lines correspond to the continuation of the ITD for the next phase differences due to the continuity of the phase information. This Figure shows a set of twelve representative curves corresponding to the set of twelve neurons of the delay line (13) tuned by their maximal response to particular ITDs, 20 μ s apart. This tuning corresponds to the position of the neuron within the delay line.

frequency part we understand that part of the animal's hearing range, in which the ITD cue is used. The disparity of the inputs between the two ears is called interaural and the neurons, which have access to the spikes originating from both ears are called binaural. Our model describes the first binaural neurons found when progressing from periphery. There are analogous binaural neurons in birds (neurons of nucleus laminaris) and in mammals (neurons of superior olive), yet they are embedded in different neural circuits (7, 15). The level of description used in our paper is therefore the level of the neural circuit for the ITD processing.

Both the avian and mammalian circuits are subject to continuous experimentation (10). The experiments are demanding because stimuli have to have high time precision in the microsecond range and because the neural circuits are in intimate proximity of centers of animal's vital functions. Experimental literature is now available with detailed discussion

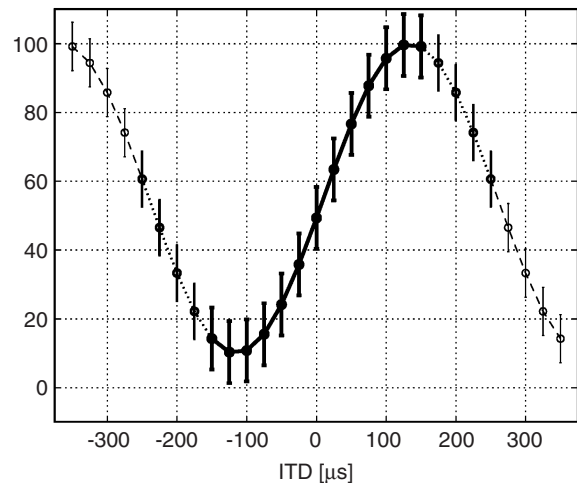


Fig. 2. Schematic responses of binaural neurons in the mammalian circuit (MSO). For axes and details see the caption for the previous Figure. The information about the azimuth is extracted from at least two (or possibly more) tuning curves of broadly tuned delay channels (10). Only one of the two channels is shown here. Error bars illustrate the stochastic dimension of the tuning: the actual response varies from trial to trial and the exact azimuth is extracted by pooling the signal from a small set of neurons.

of both these circuits (3, 7). We can take the experimental description of these neural circuits and use them as a testing ground for the theories conveying how sound is encoded into spike trains in lower processing stages of the auditory pathway.

Theories of how these circuits might function have long been presented and compared to experimental data. The first theory suggested that the ITD's are processed by a delay line. This was proposed by Jeffress in 1948 (13). Only several decades later it was demonstrated by experiments on barn owls that a variant of delay line exists in a neural circuit of the barn owl and also that line was found in other birds (7, 10). The neural circuit of birds converges on the binaural nucleus laminaris.

Even though several experimentalists have attempted to find a structure similar to avian delay line in mammals, the results are not conclusive (15). We described necessary and sufficient properties of binaural neurons to make them able to use the information contained in the ITD (21). New experimental findings indicate that a mammalian sound localization circuit may be wired differently than the avian (10). McAlpine and colleagues (24) also presented several purely theoretical works, amongst them reference No. 11, in which they explain some properties of the ITD coding by a neural population. Some of the coding properties are shared by both avian (Fig. 1) and mammalian (Fig. 2) circuits. We described neural

operations of the coincidence detection circuits in terms of random spike arrivals and their probability densities (23). We followed up with more estimates, in particular studying the standard deviation of spike timing, sometimes called spike timing jitter (17, 22).

Model

The integrate and fire neuronal model, also called the perfect integrator, is one of the simplest neural models. This model captures the integration property of a neuron, especially for time scales shorter than the neuronal time constant. The time constant can be measured electro-physiologically and especially in auditory neurons it can be assumed to be large compared to the rate of synaptic inputs. With a given integration time window and complex input, such a neuronal model possesses rich output behavior, which can be attributed to hypothetical neural computations (9, 31). In agreement with our notation in reference No. 17 the neural computations of the perfect integrator can be summarized as:

$$\left. \begin{array}{l} \text{1-st input : } f_{\text{in}}(t) \\ \vdots \\ \text{n-th input : } f_{\text{in}}(t) \end{array} \right\} \rightarrow \text{first } (k) \text{ spikes} \rightarrow f_{\text{out}}(t), \quad [1]$$

where f_{in} are input spike timing densities, f_{out} is the output density, k is the number of spikes needed to fire the action potential and n is the total number of inputs. Here we consider only small amounts of input spikes, $k, n = 2, 3, 4$ and a short time integration window Δ_{W} . Due to the short time window, such a neuron acts as a coincidence detector. We will see that the neural computation performed by this model is quite complex.

The mammalian brainstem neural circuitry is shown in Fig. 3. Since there is chain of delays on both sides at each synapse, we can generally suppose that we have n ipsilateral delays in the ipsilateral (A) branch of the pathway, $\Delta_{A1}, \Delta_{A2}, \dots, \Delta_{An}$ and m contralateral delays in the contralateral (B) branch of the pathway, $\Delta_{B1}, \Delta_{B2}, \dots, \Delta_{Bm}$. The total delay difference between those two pathways, Δ_{AB} , is equal to

$$\Delta_{\text{AB}} = \Delta_{\text{AJ}} - \Delta_{\text{BJ}} = \sum_{i=1}^n \Delta_{\text{Ai}} - \sum_{j=1}^m \Delta_{\text{Bj}}. \quad [2]$$

We assume that these random delays with (timing) jitter (subscript J) on sides A and B (left and right) are mutually independent and identically distributed non-negative random variables. We denote them as $\Delta_{\text{AJ}}, \Delta_{\text{BJ}}$ with a maximum value of Δ_{max} . The following constraints have to be imposed on them, with the third reflecting the property that the coincidence detection (time) window Δ_{W} should be shorter

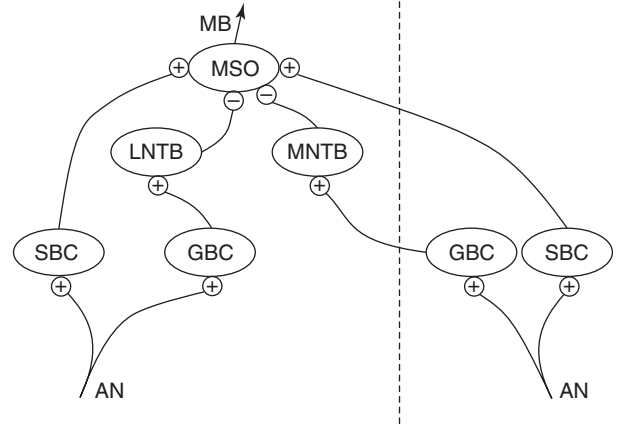


Fig. 3. Mammalian brainstem wiring. This scheme shows synaptic connections in the mammalian brainstem converging onto the nucleus of the medial superior olive (MSO) and passing the spike trains further to the midbrain (MB). Vertical line denotes the dividing mid line between the left and right brainstem sides. Only one half of the circuit converging to the left MSO is shown. The neurons of the processing stages from bottom to top are: neurons of the ganglion spirale, also known as auditory nerve (AN) neurons, neurons of the cochlear nuclei, namely globular bushy cells (GBC) and spherical bushy cells (SBC). Input of the former is inverted through the medial (MNTB) and lateral nuclei of trapezoid body, while input of the latter connects directly to the MSO. The polarity of synapses is denoted by the plus (+) sign for excitatory and minus (-) sign for inhibitory.

than or equal to the maximum delay, and that the sound period T should be greater than or equal to the maximum delay:

$$\begin{aligned} 0 \leq \Delta_{\text{JA}} \leq \Delta_{\text{max}}, \quad 0 \leq \Delta_{\text{JB}} \leq \Delta_{\text{max}}, \\ 0 \leq \Delta_{\text{W}} \leq \Delta_{\text{max}} \leq T. \end{aligned} \quad [3]$$

These are the theoretical conditions which have to be met to solve the analytical calculation. Using numerical simulations we verify that the model is robust in parameter variations within these conditions. Some of the extrapolated values mimic the phase ambiguity observed in the auditory experiments.

Let us substitute to Equation [1] $n = 2$ and for the spike timing probability density f_{in} the beta distribution. Such a spike timing is observed in the auditory pathway (18). The beta distribution has a probability density function written as:

$$f_{\text{in}}(t) = \frac{1}{B(a, b)} H(t) H(1-t) t^{a-1} (1-t)^{b-1}, \quad [4]$$

with parameters $a, b \geq 1$, normalized by $B(a, b)$ to give unity integral. $B(x, y)$ and $\Gamma(x)$ are respectively (Euler) beta and gamma functions and can be cal-

culated as $B(x, y) = \Gamma(x)\Gamma(y)/\Gamma(x + y)$ and $\Gamma(x) = \int_0^{\infty} t^{x-1} \exp(-t) dt$ (1). $H(x)$ is the (Heaviside) step function: $H(x) = \begin{cases} 0, & \text{for } x < 0, \\ 1, & \text{for } x \geq 0. \end{cases}$ The mean, standard deviation and variation coefficient, of the beta distribution expressed in parameters a and b are:

$$\mu = \frac{a}{a+b}, \quad \sigma = \sqrt{\frac{ab}{(a+b)^2(a+b+1)}},$$

$$C_V = \sqrt{\frac{b}{a(a+b+1)}}, \quad [5]$$

respectively. For $a = b = 1$ the beta distribution in Equation [4] is the uniform distribution and for $a = 1$ and $b = 2$ it is the triangular distribution.

Results

Analytical Calculations

For two independent random variables Δ_{AJ} and Δ_{BJ} we can obtain a new random variable defined as a difference in Equation [2]. The convolution formula for the sum of the two random variables Δ_{AJ} and $-\Delta_{BJ}$ with distributions f_{inA} and f_{inB} prescribes that distribution f_{out} of this new variable Δ_{AB} is written as convolution integral

$$f_{out} = f_{inA} * f_{inB}. \quad [6]$$

For the evaluation of this convolution formula we can use the property of the Laplace integral transform (30) that it replaces the convolution operation (*) with multiplication (·).

$$L\{f_{inA} * f_{inB}\} = L\{f_{inA}\} \cdot L\{f_{inB}\}. \quad [7]$$

We obtain the sought distribution f_{out} by means of the inverse Laplace transform:

$$f_{out} = L^{-1}\{L\{f_{inA}\} \cdot L\{f_{inB}\}\}. \quad [8]$$

Let us substitute f_{inA} and f_{inB} in this Equation with the beta distribution [4], $f_{inA} = f_{in}(t)$ and $f_{inB} = -f_{in}(1-t)$. We can use the properties of the Laplace transform that it is a linear operator and can have as an argument the unit step (Heaviside) function. Even though the two unit steps define the function in three pieces, the symbolic manipulation software we used, the MATLAB(TM) symbolic toolbox, calculates the output function seamlessly. One of the simplest non-trivial examples is to set $a = 2$ and $b = 4$ in formula [4]. In this example, f_{in} is the 4th degree polynomial, $\sum_0^4 c_i t^i$, such that the output of the convolution will be the 9th degree polynomial:

$$f_{out}(t) = \begin{cases} -c_9 t^9 - c_8 t^8 + c_7 t^7 - c_6 t^6 + c_5 t^5 - c_4 t^4 + c_3 t^3 - c_2 t^2 + c_1 t^1 + c_0, & 0 \leq t \leq 1, \\ +c_9 t^9 - c_8 t^8 - c_7 t^7 - c_6 t^6 - c_5 t^5 - c_4 t^4 - c_3 t^3 - c_2 t^2 - c_1 t^1 + c_0, & -1 \leq t \leq 0, \\ 0, & \text{otherwise,} \end{cases} \quad [9]$$

where the absolute values of the polynomial coefficients are $c_9 = 0.6349$, $c_7 = 2.8571$, $c_4 = 20$, $c_3 = 33.33$, $c_2 = 17.1429$, $c_0 = 1.5873$, $c_8 = c_6 = c_5 = c_1 = 0$ and signs are as denoted. The signs of the coefficients confirm that this polynomial is an even function.

In this way we arrived at the formula for the excitatory coincidence detection (23). When we add inhibition, from Equation [9] we obtain the inhibitory coincidence detection function as a new density f_{out} obtained by the difference of the excitatory f_{exc} and the inhibitory f_{inh} inputs,

$$f_{out}(t) = f_{exc}(t) - f_{inh}(t - T_\phi) + C. \quad [10]$$

T_ϕ corresponds to a phase shift, and C is a baseline firing rate to get nonnegative firing. The two ITD tuning curves with and without inhibition in Fig. 7 can be compared to experiment in Ref. No. 3. To

highlight the differences between the distributions discussed here, Fig. 5 shows f_{in} of Equation [4] and f_{out} of Equation [9], compared to normal density $N(0, \sigma)$ with the standard deviation σ set equal to the standard deviation of f_{out} . In order to see the differences in their tails, these functions are shown in both semi-logarithmic and linear scales.

Numerical Calculations

In this section we briefly present implementation details of the model in Fig. 4. The latencies are simulated together with the stochastic arrival of spike at the first binaural neuron. We used several modifications of the model. We are now in a process of rigorously showing that some of the model variants are computationally equivalent. We use both inhibitory and excitatory inputs to the binaural neuron. This is

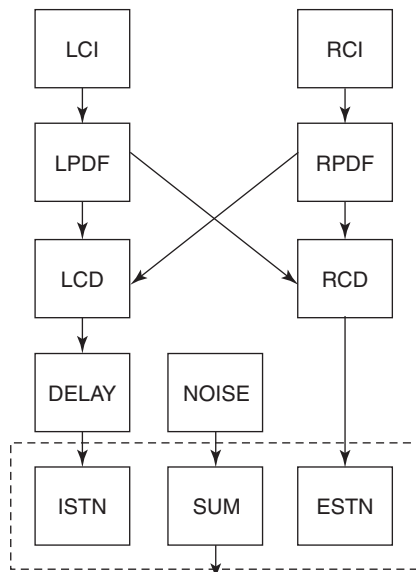


Fig. 4. Model circuitry. L and R CI are left and right cochlear natural or cochlear implant inputs. L and R PDF are analogously left and right probability density functions of spike generation times. L and R CD are coincidence detector neurons. This is an analogue to the processing at the first binaural neuron. DELAY is the axonal conduction delay and is shown only for the inhibitory branch. I and ESTN are inhibitory and excitatory synaptic transmissions, which are summed as the SUM of synaptic inputs together with NOISE. It can be shown that this circuitry is equivalent to the simplified wiring of the mammalian sound localization circuit, as described for example in reference No. 10.

consistent with the mammalian literature (10, 11, 24). Our model makes it possible to describe neural coding in lower stages of auditory processing and therefore the model can be used for studying both natural and electronic sensation, the latter evoked by cochlear implant stimulation (20).

We can recalculate the input to the auditory nerve by a cochlear implant emulator. We use a published and publicly available realistic software emulation of cochlear implant adapted from an open source in the MATLAB(TM) script language by Bruce and several other authors (4-6, 34). We then convert the CI signal into auditory nerve activity using a method implemented by these authors (6).

In the last decade, many deaf patients in Europe have received implants in both cochleas from the same manufacturer. Existing technology enables synchronized stimulation of both cochleas in these patients. Cochlear implants can switch between several coding strategies, some useful for speech understanding and some other useful for open space hearing. In the Nucleus (R) 24 cochlear implant system common coding strategies are Spectral Peak (SPEAK), Continuous

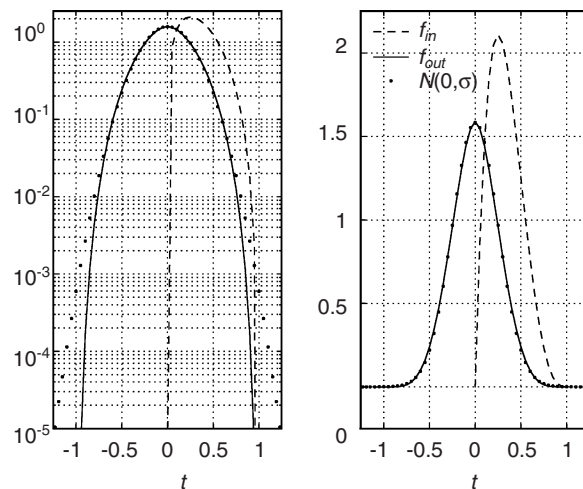


Fig. 5. Input and output probability density functions. The semi-logarithmic plot used in the left panel shows the divergence of the two tails of the normal and beta densities, as the tail of the normal density ($N(0, \sigma)$, dots) exceeds the range below the argument of -1 and above that of 1 . Semi-logarithmic plotting on the y-axis of the densities was used in the past, because the normal density is parabolic in these coordinates. In the right panel, the densities are shown on a linear scale for a more conventional comparison. Note that dots of the normal density virtually fall on the solid curve of the beta density in the linear panel. Standard deviation of the normal density σ was chosen to be equal to standard deviation of the output density f_{out} . The dashed line shows the input density f_{in} and the solid line shows the output density f_{out} on the whole interval, see Equation [9].

Interleaved Sampling (CIS) and Advanced Combination Encoder (ACE) (8). In our example we choose the SPEAK strategy (8, 25). Figure 6 shows successive steps of sound processing in the emulator. Only the channel with the highest energy out of more than twenty frequency channels of the CI model is shown.

It has been shown previously in experiments that there exists a trade-off between sound localization and speech comprehension in these patients (8). Another finding demonstrates that specific types of spike timing jitter are beneficial in these patients (19). In our work we demonstrate theoretically, why different types of jitter are useful. The improvements of our sound localization model may lead to the formulation of novel neural coding strategies for cochlear implants. Figure 7 shows data obtained in the CI emulator together with our model. On the x-axis is the interaural time delay between left and right inputs.

Discussion

Our results do not exclude the possibility that

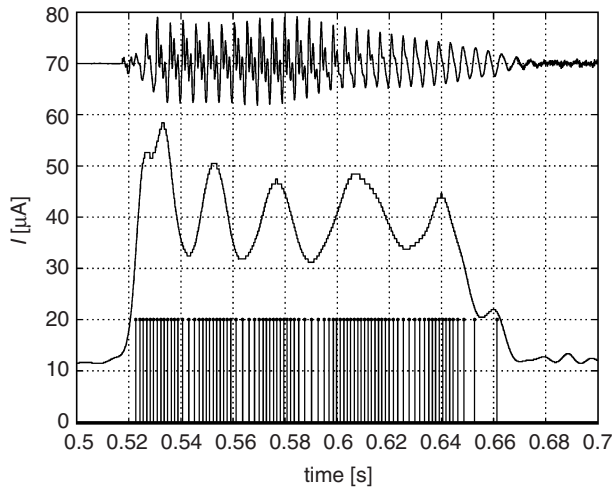


Fig. 6. Output example. Spikes in the auditory nerve are elicited by the current from the cochlear implant. The x-axis shows a time span of 200 ms. The upper trace shows complex speech sound in auditory space. This is a Czech syllable pronounced by a Czech native speaker. The units on the y-axis (μA) are related to the middle trace, which shows the stimulation current evoked by the sound. The bottom trace is a spike train elicited in a channel of cochlear implant with the highest energy content.

arrays of delay lines in the auditory brainstem neural circuit exist in mammals. We are only offering an alternative solution to the well known problem of time precision in sound localization. Our alternative solution is based on the concept of only one delay line. It is possible that the real neural circuit lies between these two extreme alternatives. More specifically, the circuit of the MSO might use a few channels tuned to specific ITDs. The continuum of the ITDs would be then represented in analogy to the continuum of colors, which is represented in the activity of three channels corresponding to just three cone pigments. ITD can be detected for both low frequency sounds and also for the envelopes of the amplitude modulated sounds (14).

All neural circuits extracting precise timing information must employ coincidence detectors. Output of these coincidence detectors can be passed with even higher timing precision to the next order neurons in auditory pathway (28). Precise spike timing is preserved even in higher relays of the auditory pathway, up to the thalamus and cortex, (27, 32), yet the function of such precisely timed spikes in the cortex is not known.

An important point for the interpretation of experimental data is the following. When looking at the experimental data, one cannot distinguish which of the two similar densities (with and without the tails, see Fig. 5) govern the data. We exemplified this

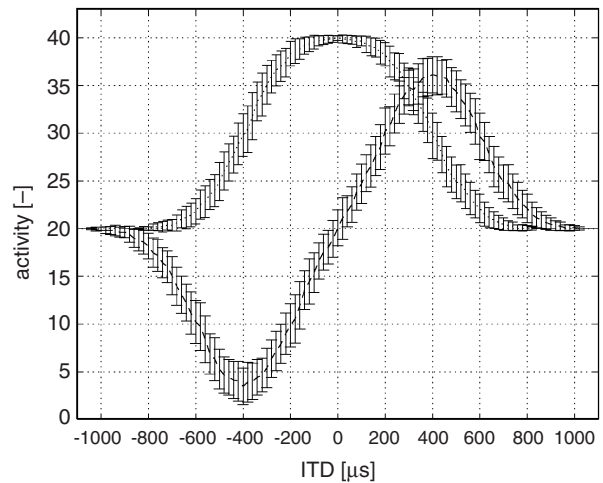


Fig. 7. Output of binaural model neuron. The model neuron is connected to the whole CI simulation package. ITD is on the x-axis in microseconds. Neuronal output activity is plotted on the y-axis. The dashed biphasic curve is the output of the full model with both excitation and inhibition. The dotted monophasic curve is the model without inhibition. Only one period of the output is shown on this picture for clarity, but the output is periodic, hence the x-axes could be easily converted to the interaural phase difference. The error bars are sample standard deviations obtained numerically in simulations. In both curves they are the result of 100 trials. In other words, these error bars do not show the level of the noise in the system. Instead, the noise is introduced into the system *via* the randomness of the random variables and the magnitude of the timing jitter.

in (22), where we demonstrated that the sample data of output density without the tails pass the Jarque-Bera (2) normality test with the confidence level p lower than any practically used value. Since publishing this finding (22), we have a more precise explanation as to the implications of the result of the Jarque-Bera test. Mathematically, the two densities differ because of their tails. Practically, when analyzing data or numerically simulated data samples, one data density cannot be distinguished from the other. It is important to realize that the output density is the density of timing jitter of the MSO neuron, and that the spike encoding this randomness is entering another neuron. The interpretation of data of spike timing density using the normal density is a common practice and it is not necessarily wrong. The difference we report here is subtle, but may matter for the proper description of the neural algorithm.

While the size of the CD window can be regulated in the neural circuit, the time jitter can also change based on the other parameters of the circuit, (28). Our results show that the proposed mechanism is robust to variation of these two parameters (time window and

time jitter).

Figures 1 and 2 hint at how various ranges of ITDs might be processed in mammals with various head sizes. To demonstrate this theoretical result in experiment might be difficult, but it is not impossible. Implanting two cochlear implants in both cochleas of a monkey and also in a guinea pig for comparison based on varied timing jitter is one suggestion for a demanding, yet realizable experiment.

It is also interesting to know that another type of the ITD jitter was already used in a variant of a human psychophysical experiment. A controlled neural jitter influences the sound localization performance, as reported in reference No. 19. This experiment is measurement of the sound localization performance of hearing impaired human volunteers. These implantees, who have cochlear implants on both sides, are capable of the ITD processing. In short, it is interesting to consider the working hypothesis of this paper, stating that the jittered binaural stimulation abolishes the adaptation at the binaural processing stage as opposed to the stimulation without time jitter. We can compare this hypothesis to what we know about the circuit. We would like to comment that the observation in reference No. 19 might also be explained by the active use of timing jitter by neurons, as proposed here. In the psychoacoustics of normal hearing subjects, this might be indirectly illustrated by careful manipulation of echoes and distractors, (16), but the interpretation of these experiments from the point of view of our theory is less straightforward.

Our model circuit does not reflect the detailed anatomical wiring of the real circuit because the real wiring is far too complex. In designing our model we started with the detailed circuit description. Any physiological description uses elementary neural operations: delay, summing of the postsynaptic potentials, coincidence detection, which can be understood as an operation of the “and” binary logical gate and inhibition, which can be understood as the “not” unary operation. Starting from the (complex) neural circuit, we can simplify it, when simplification of the series of operations is possible. Our model is the result of such a simplification. We can prove that the original and simplified circuits are equivalent from the computational perspective (22). It should be also noted that whenever (slow) software implementation would not make the real-time implementation possible immediately, most computations can be implemented using (much faster) digital signal processing hardware. We refer here to the hardware model of the interaural level difference processing in bats (12) as one out of many examples.

The neural circuits which use coincidences of two inputs and calculate the correlation between the

two neural signals have a long tradition following the Reichardt detector, a neural circuit for motion detection in fly vision (29, 33). In line with this tradition, early applications of neural correlations in sound localization circuits involved the possibility that not the peak, but the maximal slope of a response of the coincidence detector may relay the information, (26).

Figure 4 shows the neural operations as individual functional building blocks of a neural algorithm, which should not be mapped directly to neurons of the binaural pathway in Fig. 3. The spike timing jitter propagates through all the neurons of the ascending auditory pathway. In this paper we demonstrated, why a specific type of spike timing jitter is essential for azimuth processing in both normal hearing and hearing with cochlear implants.

Acknowledgments

This work was supported by the research projects MSM 0021620806 and MSM 6840770012 granted by the Ministry of Education, Youth and Sports of the Czech Republic. Thanks to Zbynek Bures, Abigail Klopfer, Paul Poon and Martin Vondrasek.

References

1. Abramowitz, M. and Stegun, I. Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables, New York: Dover, 1965.
2. Bera, A. and Jarque, C. Efficient tests for normality, heteroscedasticity and serial independence of regression residuals: Monte Carlo evidence. *Econ. Lett.* 7: 313-318, 1981.
3. Brand, A., Behrend, O., Marquardt, T., McAlpine, D. and Grothe, B. Precise inhibition is essential for microsecond interaural time difference coding. *Nature* 417: 543-547, 2002.
4. Bruce, I.C., Irlicht, L.S., White, M.W., O’Leary, S.J., Dynes, S., Javel, E. and Clark, G.M. A stochastic model of the electrically stimulated auditory nerve: Pulse-train response. *IEEE Trans Biomed. Eng.* 46: 630-637, 1999.
5. Bruce, I.C., White, M.W., Irlicht, L.S., O’Leary, S.J. and Clark, G.M. The effects of stochastic neural activity in a model predicting intensity perception with cochlear implants: Low-rate stimulation. *IEEE Trans Biomed. Eng.* 46: 1393-1404, 1999.
6. Bruce, I.C., White, M.W., Irlicht, L.S., O’Leary, S.J., Dynes, S., Javel, E. and Clark, G.M. A stochastic model of the electrically stimulated auditory nerve: Single-pulse response. *IEEE Trans Biomed. Eng.* 46: 617-629, 1999.
7. Carr, C.E. and Konishi, M. Axonal delay lines for time measurement in the owl’s brainstem. *Proc. Natl. Acad. Sci. U.S.A.* 85: 8311-8315, 1988.
8. Clark, G.M. Personal reflections on the multichannel cochlear implant and a view of the future. *J. Rehabil. Res. Devel.* 45: 651-694, 2008.
9. Gerstner, W., Kempter, R., van Hemmen, J.L. and Wagner, H. A neuronal learning rule for sub-millisecond temporal coding. *Nature* 383: 76-78, 1996.
10. Grothe, B. New roles for synaptic inhibition in sound localization. *Nat. Rev. Neurosci.* 4: 540-550, 2003.
11. Harper, N.S. and McAlpine, D. Optimal neural population coding of an auditory spatial cue. *Nature* 430: 682-686, 2004.
12. Horiuchi, T. and Hynna, K. Spike-based VLSI modeling of the

- ILD system in the echolocating bat. *Neural Networks* 14: 755-762, 2001.
13. Jeffress, L.A. A place theory of sound localization. *J. Comp. Physiol. Psychol.* 41: 35-39, 1948.
 14. Joris, P.X., Schreiner, C.E. and Rees, A. Neural processing of amplitude-modulated sounds. *Physiol. Rev.* 84: 541-577, 2004.
 15. Joris, P.X., Smith, P.H. and Yin, T.C.T. Coincidence detection in the Auditory system: 50 years after Jeffress. *Neuron* 21: 1235-1238, 1998.
 16. Kopco, N., Best, V. and Shinn-Cunningham, B.G. Sound localization with a preceding distractor. *J. Acoust. Soc. Am.* 121: 137-152, 2007.
 17. Kostal, L. and Marsalek, P. Neuronal jitter: Can we measure the spike timing dispersion differently? *Chinese J. Physiol.* 53(6): 454-464, 2010.
 18. Krishna, B.S. A unified mechanism for spontaneous-rate and first-spike timing in the auditory nerve. *J. Comput. Neurosci.* 13: 71-91, 2002.
 19. Laback, B. and Majdak, P. Binaural jitter improves interaural time-difference sensitivity of cochlear implantees at high pulse rates. *Proc. Natl. Acad. Sci. U.S.A.* 105: 814-817, 2008.
 20. Loizou, P.C. Introduction to cochlear implants. *IEEE Eng. Med. Biol. Mag.* 18: 32-42, 1999.
 21. Marsalek, P. Coincidence detection in the Hodgkin-Huxley equations. *Biosystems* 58: 83-91, 2000.
 22. Marsalek, P. and Drapal, M. Mechanisms of Coincidence Detection in the Auditory Brainstem: Examples, in: *Mathematical Modeling of Biological Systems*, edited by A. Deutsch, R. Bravo de la Parra, R. de Boer, O. Diekmann, P. Jagers, E. Kisdi, M. Kretzschmar, P. Lansky and H. Metz, volume 2, 255-264, Birkhaeuser, 2008.
 23. Marsalek, P. and Lansky, P. Proposed mechanisms for coincidence detection in the auditory brainstem. *Biol. Cybern.* 92: 445-451, 2005.
 24. McAlpine, D., Jiang, D. and Palmer, A.R. A neural code for low frequency sound localization in mammals. *Nat. Neurosci.* 4: 396-401, 2001.
 25. Millar, I.B., Tong, Y.C. and Clark, G.M. Speech processing for cochlear implant prostheses. *J. Speech Lang Hear Res.* 27: 280-296, 1984.
 26. Phillips, D.P. and Brugge, J.F. Progress in neurophysiology of sound localization. *Annu. Rev. Psychol.* 36: 245-274, 1985.
 27. Poon, P.W.F. and Chiu, T.W. Similarities of FM and AM receptive space of single units at the auditory midbrain. *Biosystems* 58: 229-237, 2000.
 28. Reed, M.C., Blum, J.J. and Mitchell, C.C. Precision of neural timing: Effects of convergence and time-windowing. *J. Comput. Neurosci.* 14: 35-47, 2002.
 29. Reichardt, W. Nervous integration in the facet eye. *Biophys. J.* 2: 121-143, 1962.
 30. Rektorys, K. (editor) *Survey of Applicable Mathematics*, London: Iliffe Books, 1969.
 31. Vaillant, J. and Lansky, P. Multidimensional counting processes and evoked neuronal activity. *IMA. J. Math. Control Inf.* 17: 53-73, 2000.
 32. Villa, A.E. and Abeles, M. Evidence for spatiotemporal firing patterns within the auditory thalamus of the cat. *Brain Res.* 509: 325-327, 1990.
 33. Zanker, J.M., Srinivasan, M.V. and Egelhaaf, M. Speed tuning in elementary motion detectors of the correlation type. *Biol. Cybern.* 80: 109-116, 1999.
 34. Zhou, Y., Carney, L.H. and Colburn, H.S. A model for interaural time difference sensitivity in the medial superior olive: Interaction of excitatory and inhibitory synaptic inputs, channel dynamics, and cellular morphology. *J. Neurosci.* 25: 3046-3058, 2005.

B Seznam příloh

Součástí této práce je datový nosič (CD) se zdrojovými kódy modelů, elektronickou verzí disertace, autoreferátu, publikovaných článků a obrázků.