

- [107] L. H. Doherty and S. A. Stone, "Forward scatter from rain," *IEEE Trans. Antennas Propagat.*, vol. AP-8, pp. 414-418, July 1960.
- [108] K. L. S. Gumm and T. W. R. East, "The microwave properties of precipitation particles," *Quart. J. Roy. Meteor. Soc.*, vol. 80, pp. 522-545, Oct. 1954.
- [109] D. Gibble, "Effects of rain on transmission performance of a satellite communication system," in *IEEE Int. Conv. Rec.*, pt. 6, p. 52, 1964.
- [110] J. Rutz, "More on wet radomes," *IEEE Trans. Antennas Propagat.*, vol. AP-13, p. 823, Sept. 1965.
- [111] L. Anderson, "Measurements of 20 GHz transmission through a wet radome," *IEEE Trans. Antennas Propagat.*, vol. AP-23, Sept. 1975.
- [112] B. G. Evans, A. J. Fryatt, J. Read, and P. T. Thompson, "Investigation of the effects of precipitation on parabolic antennas employing linear orthogonal polarization at 11 GHz," *Electron. Lett.*, vol. 7, pp. 375-377, July 1971.
- [113] A. B. Crawford, D. C. Hogg, and L. E. Hunt, "A horn-reflector antenna for space communication," *Bell Syst. Tech. J.*, vol. 40, pp. 1095-1116, July 1961.
- [114] C. Dragone and D. C. Hogg, "The radiation pattern and impedance of offset and asymmetrical near-field Cassegrainian and Gregorian antennas," *IEEE Trans. Antennas Propagat.*, vol. AP-22, pp. 472-475, May 1974.
- [115] R. H. Turtin, "A multibeam satellite antenna for the 20 and 30 GHz radio bands," *Bell Syst. Tech. J.*, July-Aug. 1975.
- [116] C. Dragone, "An improved antenna for microwave radio systems consisting of two cylindrical reflectors and a corrugated horn," *Bell Syst. Tech. J.*, vol. 53, pp. 1351-1379, Sept. 1974.
- [117] K. Funakawa and Y. Otsu, "Characteristics of the slant path rain attenuation at 35 GHz obtained by solar radiation and atmospheric emission observations," in *I.U.C.R.M. Colloquium Proc.*, pp. III 9-1-III 9-5, Oct. 1973.
- [118] J. E. Ahluti and J. E. Hall, "Site diversity advantage for satellite communication at 11.6 GHz," *Electron. Lett.*, vol. 10, pp. 527-528, Dec. 12, 1974.
- [119] G. Brunsard, "Rain-induced cross polarization and raindrop canting," *Electron. Lett.*, vol. 10, pp. 411-412, Oct. 3, 1974.

Модели слуха¹⁾

УДК [007 : 57].001.57

Models of Hearing²⁾

ШРЕДЕР (Manfred R. Schroeder), член ИИЗР

Заказная статья

Люди издавна размышляли о природе слуха человека и животных. Некоторые догадки, дополненные экспериментом (и наоборот), привели нас в конце концов к пониманию механизма слуха — от тонких монопольных фазовых эффектов (запрещаемых акустическим законом Ома) до бинаурального стереофонического эффекта (приветствуемого почти каждым). Многие из этих недавно приобретенных знаний нашли свое формальное выражение (как и в других областях) в моделях: математических или физических моделях (или тех и других) «работы» уха или того, как оно могло бы работать, «если бы было создано авторами умозрительных моделей, а не в процессе эволюции». В настоящей статье после кратких сведений о строении человеческого уха представлены модели обоих видов. Автор не претендует на полноту изложения, он просто намеревался показать, как развиваются исследования в этой интересной и все еще не понятой до конца области слуховых ощущений.

1. ВВЕДЕНИЕ

А. От Лукреция до Тартини

Механизм слухового ощущения привлекал внимание ученых и служил предметом умозрительных размышлений человека с зарождения цивилизации. При попытках человека познать природу слуха важную роль играли модели. Еще в первом веке до нашей эры римский поэт и философ Лукреций постулировал модель слуха, согласно которой различные тоны воспринимаются небольшими крупинками пес-

ка во внутреннем ухе. Можно считать, что модель Лукреция как бы заложила основы «корпускулярной теории звука», предвосхитив весьма забавным образом фононы современной квантовой акустики. В своей книге [1] Лукреций так рассуждал о речи и слухе:

«Слышится, прежде всего, всякий звук или голос, как только...

В уши проникнув, своим они телом нам чувство затронут. Ибо и голос и звук непременно должны быть телесны. Если способны они приводить наши чувства в движение. Голос, к тому же, гортань нам часто скребет, и наружу Крик исходя, все пути горловые шершавит немало. Ибо, лишь в узкую щель, накопясь большого толку, Первоначала начнут голосов вырываться наружу. Выход, конечно, и уста, за днаемши ими, скребется.

...Также ты знаешь, каков причиняет телу убыток. Сколько и нервов берет, сколько сил у людей отнимает. Бесперывная речь, от сиянья зари восходящей. Произносимая вплоть до глубокого сумрака ночи. Если, к тому же, она произносится голосом громким. Значит, признать мы должны непременно, что голос телесен. Если от долгих речей убывает часть нашего тела.

...Дальше, когда из глубин вытесняются нашего тела Звуки, которые мы через рот испускаем наружу. Гибкий, искусный язык на слона разделяет их быстро. С помощью также и губ, принимающих должную форму.

...Часто бывает и так, что в народном собрании до слуха Всех долетает одно глашатай громкое слово: Голос единственный здесь дробится на много отдельных. Вдруг, потому что идет, по отдельным ушам разбегаясь. Звучную форму словам, направляемым в них, придавая. Та же часть голосов, что до самых ушей не доходит, Попусту мимо идет, и, расставшись в воздухе, гибнет. Часть же другая, в пути отскакивая от твердых предметов, Звук отдает и порой морочит подобием слова.

...Видеть пришлось мне места, где раз шесть или семь отдается

Изданный звук, где холмы от холмов отражают немедля Слово за словом, на них отечая и вторя друг другу. Эти места, по словам в соседстве живущего люда,

¹⁾ Получена 26 декабря 1974 г., в исправленном виде — 25 мая 1975 г.

²⁾ Ориг., стр. 1332—1350.

The author is with the Drittes Physikalisches Institut, Universität Göttingen, F.R. Germany, and Bell Laboratories, Murray Hill, N.J.

Служат пристанищем Нимф, козлоногих Сатиров и Фавнов.
Эти они, говорят, затевают веселье и ночью
Шумно играют везде, тишину и покой нарушая...
...Я продолжаю. Тебе удивляться нисколько не надо,
Что «клопы» преграды, глазам ничего не дающие видеть,
Звуки доходят до нас и касаются нашего слуха.
Часто, мы видим, идет разговор за затворенной дверью,
Или, действительно, там по извилистым голос отверстиям
Может свободно идти, где для образов нету прохода:
Ведь расшелятся они, если эти отверстия не прямы,
Как у стекла, где проплыть может призрак любой без
препятствий.

Кроме того, разлетается голос по всем направлениям,
Ибо один из других голоса возникают: лишь только
Голос раздался один, как дробится сейчас же на много
Так же, как искры огня рассыпаются в новые искры.
Так в потаенных местах голоса постоянно роятся,
И раздается их звук, пробуждая окрестности всюду...

Хотя некоторые идеи Лукреция представляются сейчас фантастическими, в его произведении можно распознать те понятия, которые сегодня, 2000 лет спустя, формируют основы нашего понимания звука: *энергия, резонанс, дифракция* и даже *принцип Гюйгенса* (связанный здесь с распространением огня).

Тем не менее прошло много столетий, прежде чем были получены дополнительные количественные данные о звуке и слухе. В XVIII в. Тартини [2], знаменитый итальянский скрипач, описал свои «terzi suoni» — «третьи» тоны, которые вырабатывает само ухо из одновременно воспроизводимых двух тонов. Эти тоны имеют очень важное значение при попытках смоделировать и понять функционирование внутреннего уха и, в частности, его *нелинейные* свойства. Вопреки бытовавшим долгое время взглядам ухо не является приемником с высокой степенью линейности даже по отношению к величине низкого и темпирования звука. Более того, «комбинационные тоны», наподобие обнаруженных Тартини, в действительности начинают восприниматься ухом вблизи порога слышимости, где механические смещения во внутреннем ухе составляют доли ангстрем (10^{-8} см), т. е. сопоставимы с диаметром атома водорода! Происхождение этих нелинейностей по-прежнему остается загадкой.

Однако тот факт, что они возникают на атомном уровне, заставляет нас обратиться к молекулярным процессам и, более того, к тем молекулярным процессам, которые тесно связаны с обменом веществ во внутреннем ухе. Как показали опыты на животных, нелинейные явления изменялись, когда прекращалось питание уха кровью.

Помимо Тартини, познанию слухового восприятия содействовали своими наблюдениями (либо предположениями) многие другие композиторы и музыкальные исполнители. Однако систематическое исследование слухового аппарата (с разработкой подробных моделей) началось лишь с середины XIX в.

В. От Ома и Зеебека к Гельмгольцу

Современные исследования слухового восприятия ведут свое начало от работ Ома, Зеебека и Г. Гельмгольца.

Ом [3] постулировал свой акустический «закон фаз», согласно которому воспринимаемый тембр звука зависит исключительно от его спектра *мощности* и не зависит от фазовых углов его частотных составляющих. Хотя этот закон, как установили современные исследования, и допускает исключения, тем не менее он представляет одно из фундаментальных поло-

жений «психоакустики» — раздела психофизиологии, посвященного изучению слуха путем прослушивания тестов. Наряду с физиологическими исследованиями и математическим моделированием психоакустические эксперименты с использованием сложнейших акустических сигналов играют главную роль в современных исследованиях слуха. Тщательно поставленные психоакустические эксперименты позволяют проникнуть (если можно так выразиться) через ухо к определенным центрам сознания в нашем мозгу. И хотя Ом и его современники не располагали столь совершенным оборудованием, какое имеется в настоящее время (в частности, ЦВМ для генерации высокоточных звуковых сигналов), Ом, несомненно, является отцом (или, может быть, дедом) психоакустического подхода к исследованию уха.

Зеебек обнаружил появление так называемой «периодической высоты», т. е. ощущение тембра звука без наличия физических компонент в акустическом стимуле на воспринимаемой частоте [4]. Данные, полученные Зеебеком, а затем Шоутоном [5] и его последователями об «остаточной высоте» [6] (названной так вследствие ощущения высоты даже после устранения соответствующей частотной составляющей в сигнале), до сих пор являются предметом исследований и для психоакустиков, и для физиологов, и для разработчиков моделей.

Гельмгольц не только интересовался ощущением тона [7], но выполнил также большой объем физиологических исследований по анатомии уха.

На основе своих наблюдений Гельмгольц выдвинул резонансную теорию функционирования внутреннего уха; хотя эта теория и не приемлема уже в буквальном смысле, она позволила уловить самую суть механизма работы внутреннего уха, а именно *частотную избирательность*.



Фиг. 9 Ухо человека. Звуковые волны, попадающие в ушную раковину через наружный слуховой проход, подводятся к барабанной перепонке, расположенной на входе среднего уха. В среднем ухе имеются три маленькие слуховые косточки, которые выполняют функцию преобразователя импеданса между воздушной средой наружного уха и жидкой средой внутреннего уха. При высоких интенсивностях звука подвижность косточек понижается вследствие непроизвольного сокращения прижатых к ним мышц («ак, тический рефлекс»), что предохраняет среднее ухо от повреждения из-за перегрузки. Во внутреннем ухе звуковые волны бегут вдоль тонкой «базиллярной мембраны», резонансная частота которой изменяется на протяжении ее длины от 20 кГц до почти 100 Нз, обеспечивая частотный анализ поступающего сигнала.

В наше время мы многим обязаны Бекешу [8], который открыл бегущие волны на базилярной мембране (БМ), имеющие чрезвычайно важное значение, и, будучи в начале своей деятельности инженером по телефонии, правильно оценил среду распространения этих бегущих волн как неоднородную линию передачи: высокие частоты распространяются по БМ только на короткие расстояния и затем быстро затухают, тогда как низкие частоты распространяются дальше по БМ (чем ниже частота, тем больше расстояние), прежде чем начнут затухать. Такой низкочастотный характер свойств в совокупности с локальным резонансом обуславливает частотную избирательность смещений БМ, впервые отмеченную самим Бекешом и значительно позже Джонстоном [9], Роде [10], Кольдфеллем [11], Вильсоном [12] и Хельфенштайном [13], которые использовали самые совершенные методы физики, например, на основе сверхчувствительного эффекта Мессбауэра — Доплера, лазерной интерферометрии и зондирования конденсаторным микрофоном.

II. АНАТОМИЯ И ОСНОВНЫЕ ХАРАКТЕРИСТИКИ УША

На фиг. 1 представлено схематическое изображение наружного, среднего и внутреннего уха, основанное на анатомических исследованиях Гельмгольца. Попадая в наружное ухо (ушную раковину), звуковые волны проходят через наружный слуховой проход к барабанной перепонке, расположенной на входе среднего уха. Среднее ухо содержит три мелкие слуховые косточки, которые передают звуковые колебания «овальному окну» на входе внутреннего уха. Внутреннее ухо, или улитка (названа по внешнему сходству с улиткой), заполнено жидкостью и разделено мембранами на несколько каналов (фиг. 2). Одна из этих мембран (а именно БМ) поддерживает кортиев орган — орган слухового ощущения. Кортиев орган содержит волосковые клетки, которые преобразуют смещения БМ относительно текториальной мембраны в нервные



Фиг. 2. Поперечное сечение улитки; видны наполненные жидкостью каналы и базилярная мембрана с волосковыми клетками. Смещение базилярной мембраны относительно текториальной воспринимается небольшими пучками волосков (на верхушках волосковых клеток), действующих наподобие звукоусилителя фонографа в процессе преобразования механических колебаний в электрические сигналы (нервные импульсы), возникающие в слуховом нерве.

импульсы. Существуют два типа волосковых клеток: внутренние и наружные; и хотя об «организации» рецепторов улитки (см. [9]) известно уже многое, причины существования этих двух видов клеток по-прежнему до конца не выяснены.

А. Ушная раковина и наружный слуховой проход

Несмотря на относительно простое строение (по сравнению со сложной структурой и наличием иннервации во внутреннем ухе), ушная раковина и наружный слуховой проход выполняют некоторые важные функции в процессе локализации звука, объяснить которые удалось совсем недавно.

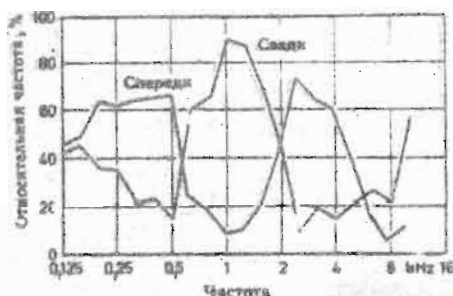
Долгое время оставалось непонятным, каким образом человек может определить местоположение источника звука в «медианной плоскости» (воображаемой вертикальной плоскости, проходящей через середину отрезка, соединяющего барабанные перепонки, и перпендикулярной к нему). Способность человека различать направления прихода звука в горизонтальной плоскости объяснялась разницей интенсивностей и фаз звуковых волн, приходящих к каждому уху [14]. Однако для звуковых источников в медианной плоскости (при симметричной форме головы) звуковые волны вблизи левого и правого ушей представляют собой одинаковые функции времени для всех углов возвышения звукового источника. Тем не менее для целого ряда сигналов человек легко может различить направления звука: вперед (0°), вверх (90°), назад (180°) и даже несколько промежуточных направлений.

Как это происходит? Одна довольно старая теория относит эту способность за счет небольших (непроизвольных) движений головы. Но способность к локализации в медианной плоскости сохраняется даже в том случае, когда голова слушающего четко зафиксирована.

Ответ, как было выяснено в течение нескольких последних лет, заключается в следующем [14]. При разных углах падения звука дифрагирует различным образом у головы и у ушных раковин. Поэтому звуковые волны, входящие в наружный слуховой проход, в результате дифракции фактически оказываются «профильтрованными». Для каждого направления на спектр приходящего звукового сигнала накладываются характерные пики и провалы. Если звук обладает достаточно широким спектром, как в случае речи, музыки и большинства шумов, эти пики и провалы в спектре акустического сигнала, воздействующего на барабанные перепонки, могут быть осознаны и использованы для определения угла падения. По всей видимости, в результате огромного опыта слушания люди научились ассоциировать различные спектральные характеристики с соответствующими вертикальными направлениями.

Если это предположение правильно, то широкополосный шум, спектр которого сформирован в соответствии с процессом дифракции для некоторого заданного угла падения, должен вызывать ощущение соответствующего субъективного направления независимо от действительного направления. Проверки субъективного восприятия таких профильтрованных сигналов весьма убедительно подтвердили эти предположения (фиг. 3).

Еще одна загадка, которую не удавалось разрешить в течение многих десятилетий, была наконец объяснена на основе рассмотрения спектральных характеристик.



Фиг. 2. Локализация звуковых источников в «медианной плоскости» (спереди, сверху, сзади) определяется спектром сигнала. При отсутствии каких-либо других указаний направления звука слушатель воспринимает звуковые волны, содержащие основную часть энергии в области 1 кГц, локализованными сзади. Звуковая энергия с частотами ниже 500 Гц и в области 3 кГц воспринимается локализованной «спереди». Источники звука, содержащие большую часть энергии в области 8 кГц, распознаются локализованными «сверху». За счет широкого спектра фильтруется благодаря дифракция у головы и ушных раковин, в результате чего подчеркиваются определенные диапазоны частот в зависимости от угла падения звуковых волн. Таким образом, слушатель может выделять различные направления в медианной плоскости (для которой сигналы у обоих ушей одинаковы и, следовательно, отсутствуют указания о времени или интенсивности) [14].

Эта загадка состоит в следующем. Почему, используя головные телефоны, мы слышим звук таким образом, как будто он возникает *внутри* нашей головы? Почему так трудно или даже невозможно осуществить внешнюю локализацию («вынести наружу») источника звука, как это мы обычно делаем, слушая первичные источники или громкоговорители?

Согласно гипотезе, которая одно время внушала некоторое доверие, дело заключается в том, что при слушании через головные телефоны источники звука следуют за движением головы и что, по-видимому, такие нестационарные источники вызывают звуковые образы внутри головы слушателя. Другая «теория» связывала внутреннюю локализацию источников звука с механическим давлением на ушники на голову. Зауэр [15] (из фирмы CBS Laboratories) удалось опровергнуть обе теории: в его остроумном эксперименте были использованы головные телефоны, которые не оказывали давления на голову и сигналы в которых автоматически изменялись при движении головы слушателя так, чтобы соответствовать модификациям звукового восприятия, ощущаемым человеком при слушании в свободном звуковом поле. И тем не менее многие слушатели (включая автора) оказались неспособными связать звуковые образы с внешним миром.

Неудачные попытки внешней локализации звуковых образов при слушании через головные телефоны сейчас объясняются следующими обстоятельствами. При применении головных телефонов в наружном звуковом проходе между барабанной перепонкой и мембраной головного телефона возникают стоячие волны. Эти волны создают эффект фильтрации, результат которой несколько отличается от спектральных пиков и провалов, обусловленных дифракцией у головы слушателя в условиях свободного звукового поля. Поэтому у слушающего через головные телефоны не могут возникнуть ассоциации с каким-либо внешним источником и, следовательно, звуковые источники свя-

зываются у него с единственным оставшимся местом — внутри собственной головы.

Если эта довольно смелая теория в основном верна, то *инверсная* фильтрация (устраняющая эффект стоячих волн в слуховом проходе) в сочетании с частотной характеристикой, соответствующей слушанию в свободном звуковом поле, должна обеспечить внешнюю локализацию источника звука. Именно это и продемонстрировал Лоуз [16] и другие исследователи.

Недавно Уоллес [17] сконструировал головные телефоны, в которых перед слуховым проходом размещался звукопоглощающий материал, в результате чего стоячие волны между мембраной головного телефона и барабанной перепонкой удалось свести к минимуму. С этими головными телефонами слушатели оказались способными «вынести наружу» источники звука, хотя, по-видимому, не столь успешно, как с электрическими фильтрами, которые не только устраняют эффект стоячих волн, но и заменяют фильтрующее действие дифракции у головы человека.

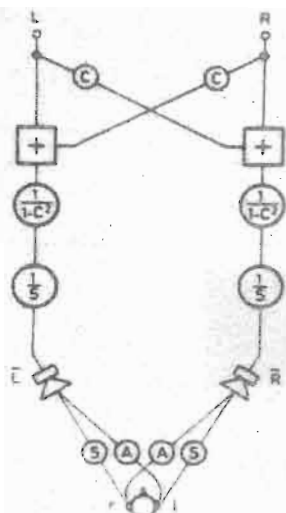
В. Возможность получения стереоэффекта высшего качества

Еще в 1962 г. Этол и автор [18] продемонстрировали способ применения фильтрации звуковых сигналов для получения эффекта правильной локализации. В обычной двухканальной стереофонической системе местоположение воспринимаемых источников звука обычно ограничено пространством между двумя громкоговорителями. Физические источники звука, расположенные сзади, сверху и по бокам (за пределами отрезка, соединяющего два громкоговорителя), не могут быть четко воспроизведены. И все же, поскольку мы имеем только два уха (и поскольку движения головы не имеют решающего значения в процессе локализации), достаточно двух громкоговорителей, чтобы выдать полное восприятие акустического пространства. Это обеспечивается при условии, что звуковые волны, исходящие от двух громкоговорителей, «специально обработаны» так, чтобы произвести на барабанные перепонки слушателя звуковое давление, не отличное от давления, которое создавалось бы в свободном звуковом поле заданными источниками (включая источники, расположенные сзади, сверху и далеко по бокам).

Воплощение этой идеи требует предварительного измерения комплексных передаточных функций (т. е. амплитуды и фазы как функций частоты) между громкоговорителем в безэховой камере и правой и левой барабанными перепонками слушателя.

Предположим, что перед слушателем размещены два громкоговорителя, один, скажем, на 30° правее, а другой — на 30° левее. Обозначим передаточную функцию от громкоговорителя до барабанной перепонки на той же самой стороне через $S(f)$, а передаточную функцию до барабанной перепонки на другой стороне — через $A(f)$.

Если бы $A(f)$ равнялась нулю, т. е. громкоговорители не оказывали бы «перекрестного» воздействия на «далекие» уши, то задача приведения специального звукового сигнала к каждой барабанной перепонке решалась бы весьма просто. Сигналы, подводимые к громкоговорителям, следовало бы профильтровать инверсно по отношению к передаточной функции «той же самой стороны» $S^{-1}(f)$. Однако вследствие дифракции звука вокруг головы человека *существует пере-*



Фиг. 4. Как передать сигнал от входа «R» (вверху) посредством громкоговорителей только к правому уху? Буквы в кружках обозначают передаточные функции. $C = -A/S$, где S — передаточная функция от громкоговорителя до уха на той же стороне, а A — передаточная функция до уха на другой стороне. Если в качестве входных сигналов «R» и «L» использовать записи с головы марионетки, оборудованной на месте «барабанных перепонки» высококачественными микрофонами, то можно получить весьма достоверную стереофоническую передачу с явно выраженными и соответствующим образом расположенными источниками звука для всех возможных направлений, включая направления сверху и снизу.

крестное воздействие на ухо «другой стороны». Такое перекрестное воздействие должно быть исключено.

Этого можно достигнуть с помощью другого громкоговорителя, излучающего соответствующим образом профильтрованный сигнал для компенсации перекрестного воздействия. Конечно, этот компенсирующий сигнал также создает «перекрестные» помехи в том ухе, для которого он не предназначен, что в свою очередь должно быть скомпенсировано новым компенсирующим сигналом, подводимым к первому громкоговорителю — и так далее до бесконечности.

Полное решение такого рода задачи многократной фильтрации и подавления иллюстрируется фиг. 4, где $C(f) = -A(f)S^{-1}(f)$ — передаточная функция фильтра в цепи компенсации перекрестного воздействия. Сквозная передаточная функция от правого входа (R) к правому уху (r) имеет вид

$$R_r(f) = (1 - C^2)^{-1} S^{-1} S + C(1 - C^2)^{-1} S^{-1} A,$$

при этом если $C = -AS^{-1}$, то $R_r = 1$, что и требуется.

Сквозная передаточная функция от правого входа к левому уху (l) также может быть легко получена с помощью фиг. 4:

$$R_l(f) = (1 - C^2)^{-1} S^{-1} A + C(1 - C^2)^{-1} S^{-1} S = 0,$$

что и требуется.

Практическая реализация изображенной на фиг. 4 схемы фильтрации продемонстрировала поразительные результаты. Хотя единственными источниками звука были два громкоговорителя, реально воспринимаемые звуковые образы могли быть локализованы

далеко в стороне и даже *сзади* слушатель... Фактически воспринимался даже угол *возвышения* источника звука (людьми с правильной формой головы!). Поскольку рассматриваемая система в целом линейна, одновременно может быть воспроизведено большое количество источников звука с их эхо-сигналами без взаимной интерференции при условии, что слушатель находится в нужном положении между громкоговорителями и не поворачивает свою голову более, чем на $\pm 10^\circ$ относительно фронтального направления. Иллюзия пространства настолько убедительна, что слушатель способен «заглянуть» для поисков невидимых источников звука. Однако в тот момент, когда он поддается этому искушению, иллюзия реальности исчезает, переходя зачастую в ощущение источников «внутри головы».

Способ воспроизведения звука, иллюстрируемый фиг. 4, открывает совершенно новые возможности в изучении акустики концертных залов [19]. Ранее при сравнении двух залов нужно было основываться на суждениях по прослушиванию музыкальных отрывков, воспроизводимых в различное время, часто различными оркестрами, под управлением различных дирижеров. Даже если все остальные моменты были одинаковы, тот факт, что между двумя исполнениями музыкального произведения проходит несколько дней, недель, а то и месяцев, делает эту весьма тонкую оценку качества звучания если не невозможной, то чрезвычайно ненадежной.

С появлением нового способа воспроизведения оказалось возможным немедленное сравнение существенных черт идентичных музыкальных программ. Автор никогда не забудет того момента, когда он впервые «пересадил» себя из кресла Берлинской филармонии в кресло зала Венского музыкального общества, слушая симфонию Моцарта «Юпитер» в исполнении Британского оркестра. На основании прежних посещений у него сложилось мнение (но не было полной уверенности), что акустика этих двух залов неодинакова; теперь все различия стали совершенно явными.

С. Среднее ухо

Людям, страдающим от различных недугов, среднее ухо может быть лучше всего знакомо в связи с тяжелыми инфекционными заболеваниями, которым оно подвержено. Однако назначение среднего уха не в том, чтобы доставлять нам одни неприятности. Наши уши в процессе эволюции достигли такой степени совершенства, при которой маловероятны излишества в их строении. Поэтому не удивительно, что некоторые важнейшие слуховые функции осуществляются именно в пространстве между барабанной перепонкой и овальным окном — (входом к внутреннему уху, см. фиг. 1).

Наиболее важной из них является *согласование акустических импедансов* между воздушной средой в наружном ухе и жидкой средой во внутреннем ухе. Характеристический импеданс воздуха равен $414 \text{ кг} \times \text{м}^{-2} \cdot \text{с}^{-1}$, а воды — $1\,480\,000 \text{ кг} \cdot \text{м}^{-2} \cdot \text{с}^{-1}$, или почти в 3600 раз больше. Однако несмотря на то, что внутреннее ухо заполнено жидкостью, похожей на воду, его входной импеданс значительно меньше, поскольку эта жидкость может «выпучивать» мембрану над круглым окном. И все же на пути от наружного к внутреннему уху импеданс должен быть преобразован на довольно значительную величину. В основном

это преобразование зависит от отношения площади барабанной перепонки к площадке у основания стремечка. Некоторую роль играет и эффект рычага при смещении слуховых косточек. В результате преобразование импеданса составляет примерно 1 : 20 и передача акустической мощности через среднее ухо улучшается более чем в пять раз.

Другая важная функция среднего уха заключается в «регуляции усиления», осуществляемой посредством так называемого *акустического рефлекса* [20], который предохраняет чувствительное ухо от перегрузки и возможного повреждения.

При высоких уровнях звука коэффициент передачи среднего уха уменьшается и в амплитудной характеристике уха начинает преобладать квадратичный член. Если в наружное ухо одновременно поступают два или более первичных тона, то в результате этой нелинейности появляются *комбинационные тоны*. Так, если в одно ухо поступают два тона большой амплитуды с частотами f_1 и f_2 , то можно услышать «разностный тон» с частотой $f_2 - f_1$. Его амплитуда растет пропорционально произведению амплитуд первичных тонов, как и следовало ожидать для квадратичного закона нелинейности.

Иногда можно услышать также суммарный тон (с частотой $f_1 + f_2$), однако, как правило, его интенсивность весьма мала вследствие низкочастотного характера передаточной функции среднего уха, которая ослабляет частоты выше ~ 1 kHz.

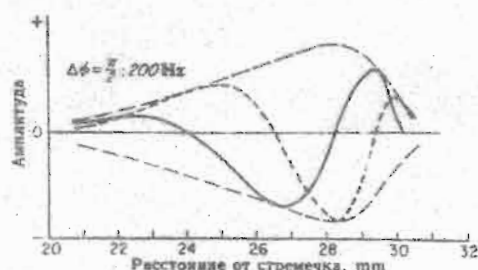
Особенно впечатляющее доказательство искажений в среднем ухе можно получить с помощью двух широкополосных шумовых сигналов, сформированных путем частотной модуляции по случайному закону двух несущих с фиксированными частотами, отличающимися на Δf . Эти шумовые сигналы подводятся к уху через два отдельных преобразователя (чтобы исключить влияние возможных искажений в преобразователе).

При низких уровнях звукового давления ощущается суммарный сигнал (линейная сумма подводимых сигналов) в виде другого широкополосного шума без заметных периодичностей. Однако при достаточно высоких уровнях прослушивается тон с частотой Δf . Поскольку этот тон отсутствует в стимулах, он должен «формироваться» самим ухом. Более того, этот тон не является «субъективным» (в том смысле, что он не создается в высших нервных центрах), а физически присутствует в среднем и внутреннем ухе. В самом деле, его можно подавить тональным сигналом частоты Δf с соответствующими амплитудой и фазой, который подводится к наружному уху.

Д. Внутреннее ухо

«Улитка» (медицинский термин «*cochlea*» соответствует латинскому названию улитки со спиральной раковиной) во внутреннем ухе является частью нашего органа слуха, обладающей частотно-избирательными свойствами. Поперечное сечение улитки представлено на фиг. 2, на которой видны три наполненных жидкостью канала. Два из этих каналов разделены базилярной мембраной, которая поддерживает *кортиева орган*, где волосковые клетки осуществляют преобразование механических колебаний в нервные импульсы.

Впервые на частотно-избирательные свойства БМ обратил внимание Гельмгольц, который представил ее для наглядности в виде ряда настроенных струн (как в пианино), резонирующих на различных частотах.



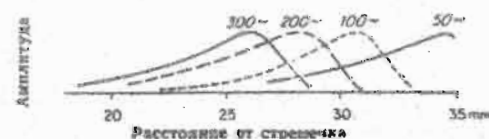
Фиг. 2. Бегущие волны на базилярной мембране для 200-Гц тона, согласно [8].

Однако когда Бекеш действительно попытался с помощью микроскопа проследить за вибрациями БМ при воздействии акустических стимулов, он «видел» бегущие волны, распространяющиеся (с уменьшающейся скоростью) от стремечка у основания улитки до геликотремы у апекса, находящегося на расстоянии 35 мм от стремечка. На фиг. 5 такие бегущие волны с частотой 200 Гц показаны для двух моментов времени, разнесенных на 1,25 мс, или на четверть периода. Отклонение БМ (в сильно увеличенном масштабе) для первого момента показано сплошной линией. Отклонение БМ спустя четверть периода показано короткими штрихами. За это время максимум отрицательного отклонения БМ сместился примерно от 27 до 28,5 мм, что соответствует фазовой скорости в 1,2 м/с или 1/300 скорости звука в воздухе (а по отношению к скорости звука в воде это составляет еще меньшую долю).

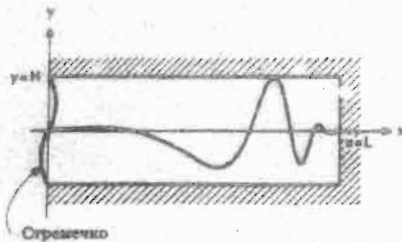
Из фиг. 5 легко установить, что на больших расстояниях от стремечка фазовая скорость более низкая, тогда как вблизи стремечка она значительно выше. Но фазовая и групповая скорости на БМ являются функциями не только расстояния, но также и частоты. На языке электротехники БМ представляют собой неоднородную (дисперсионную) линию передачи.

Две кривые, изображенные на фиг. 5 длинными штрихами, описывают эволюцию положительных и отрицательных пиков волны по мере их перемещения вдоль БМ. Эта «огibaющая» для частоты 200 Гц достигает своего наибольшего значения вблизи 28 мм. Как можно видеть из фиг. 6, на которой показаны огibaющие для четырех различных частот, чем ниже частота волны, тем дальше она распространяется по БМ, прежде чем начнет затухать. Например, волна с частотой 300 Гц до начала затухания распространяется приблизительно до 25 мм, тогда как волна с частотой 100 Гц достигает своего максимума вблизи 30 мм.

Таким образом, каждая частота звукового диапазона имеет свое собственное «место» на БМ, где она будет вызывать колебания максимальной величины. На основании этих наблюдений были разработаны так называемые «теории места» для интерпретации



Фиг. 3. Огибающие бегущих волн на базилярной мембране. Чем ниже частота, тем дальше распространяется волна, прежде чем она достигнет «зона» и начнет затухать.



Фиг. 7. Двумерная гидромеханическая модель улитки, исследованная Лессером и Беркли [21]: наполненный жидкостью жесткостенный «ящик» с гибкой мембраной.

восприятия высоты тона, согласно которым высота чистого тона определяется положением максимума колебаний БМ.

Серьезным препятствием для обоснования таких теорий оказалась относительно низкая разрешающая способность БМ по частоте (низкая добротность), которую обнаружил Бекеш. Психоакустическими исследованиями установлено, что минимально заметная разница в частоте между двумя тонами, последовательно воспроизводимыми перед слушателем, составляет менее 3 Hz на частоте 1000 Hz! Приверженцы модельных представлений попытались объяснить этот примечательный факт, предположив, что после грубой механической фильтрации возбуждений происходят сложнейшие процессы дискриминации нервных импульсов. Мы не будем повторять их запутанные (и, вероятно, ошибочные) рассуждения, а лишь заметим, что, согласно данным более поздних работ, механическая фильтрация в БМ обладает значительно большей частотной избирательностью, чем это следует из работ Бекеша. Причина, из-за которой Бекеш не смог обнаружить такое высокое разрешение по частоте, заключается в том, что он работал с мертвыми животными и для того, чтобы увидеть в свой микроскоп смещения БМ, должен был использовать звуковые сигналы с большими амплитудами. Сейчас благодаря работам Роде [10] мы знаем, что частотная избирательность существенно снижается при больших уровнях акустического возбуждения, а также спустя несколько минут после прекращения обмена веществ.

III. ДВУМЕРНАЯ МОДЕЛЬ БАЗИЛЯРНОЙ МЕМБРАНЫ

Каким образом можно понять природу бегущих волн и резонансов на БМ? На фиг. 7 показана чрезвычайно упрощенная двумерная модель улитки по Лессеру и Беркли [21], в которой фигурируют только гибкая БМ (размеченная в плоскости листа), два прилегающих канала с жидкостью и твердый «футляр». При каждом значении x координаты положения вдоль мембраны БМ характеризуется импедансом $Z(x)$, зависящим от массы $m(x)$, сопротивления $r(x)$ и жесткости $s(x)$:

$$Z(x) = j\omega m(x) + r(x) + s(x)/j\omega, \quad (1)$$

где ω — круговая частота. Полагая жидкость несжимаемой, для «потенциала скоростей» $\phi(x, y)$ в жидкой среде получаем следующее дифференциальное уравнение (уравнение Лапласа):

$$\Delta \phi = 0. \quad (2)$$

Путем дифференцирования из потенциала скоростей могут быть получены составляющие мгновенной скорости u_x и u_y , а также давление p в жидкости:

$$u_x = \frac{\partial \phi}{\partial x}, \quad u_y = \frac{\partial \phi}{\partial y} \quad (3)$$

и

$$p = -\rho \frac{\partial \phi}{\partial t} = -j\omega \rho \phi, \quad (4)$$

где ρ — плотность жидкости.

Граничные условия для ϕ у недеформируемых стенок модели записываются в виде

$$\frac{\partial \phi}{\partial n} = u_n = 0, \quad (5)$$

где u_n — составляющая скорости, нормальная к этой стенке.

В случае гибкой стенки, соответствующей овальному окну, получаем

$$\frac{\partial \phi}{\partial x} = u_x(y), \quad (6)$$

где $u_x(y)$ — (заданная) скорость стремечка.

Наконец, граничные условия на поверхности БМ определяются из уравнения движения («скорость равна давлению, деленному на импеданс»):

$$\frac{\partial \phi}{\partial y} = \frac{2\rho}{Z(x)} j\omega \phi. \quad (7)$$

Лессер и Беркли [21] получили решение этих уравнений в виде пространственного ряда Фурье для $\phi(x, 0)$. Коэффициенты Фурье находятся (путем обращения матрицы высокого ранга) из системы совместных линейных уравнений, коэффициенты которых в свою очередь получаются в результате обобщенных интегральных преобразований Фурье функции импеданса $Z(x)$.

Весьма обнадеживает тот факт, что несколько решений, полученных до сего времени, описывают ожидаемое поведение бегущих волн при наличии резонанса. Однако излишне творить, что при использованном математическом подходе невозможно получить сколько-нибудь правильной физической картины того, что же в действительности происходит в рассматриваемой модели (и соответственно в БМ).

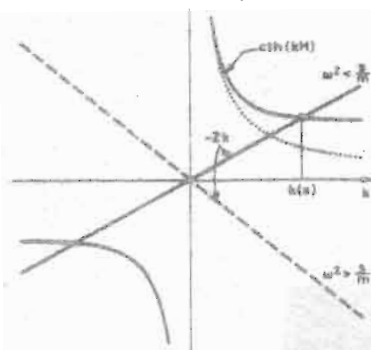
Можно лучше понять физику процессов, а также установить разумные границы применимости модели (это весьма существенно), если бегущую волну считать решением и исследовать, как она движется.

Легко убедиться в том, что (2) удовлетворяется, а граничное условие при $y = \pm H$ приблизительно выполняется, если

$$\phi(x, y) = \text{ch} \{(H - y)k(x)\} e^{-j\omega \phi(x)} \quad (8)$$

при условии

$$\left| \frac{\partial k}{\partial x} \right| \ll |k|^2, \quad (9)$$



Фиг. 8. Графическое решение для постоянной распространения $k(x)$ в модели Лессера — Беркли.

которое в свою очередь означает, что импеданс $Z(x)$ не может изменяться слишком резко вдоль x . Более определенно в количественном отношении это требование может быть записано в виде условия

$$\left| \frac{d \ln(Z)}{dx} \right| \ll |k|, \quad (10)$$

которое, по-видимому, выполняется (по крайней мере приблизительно) в случае реальных импедансных функций.

Более детальный анализ показывает, что предположение о распространении по БМ бегущих волн (с пренебрежимыми отражениями) оправдывается, если расстояние d , на котором резонансная частота БМ изменяется в ϵ раз, всюду велико по сравнению с длиной волны, деленной на $\pi\sqrt{\delta}$ (δ — коэффициент потерь, сопутствующих смещениям БМ), — условие, в точности выполняемое в организме. (Знала ли это природа, когда она создавала улитку? Или эволюция была в затруднении: что делать со *стоячими* волнами?) С другой стороны, δ не может быть слишком велико, поскольку диапазон чувствительности БМ, обладающей ограниченной длиной, должен перекрывать частотный спектр, простирающийся на много октав.

Граничные условия на БМ приводят к следующему уравнению для «постоянной распространения» $k(x)$, выраженному в неявной форме:

$$-\frac{Z(x)}{2\rho i\omega} k(x) = \text{cth}(Hk(x)). \quad (11)$$

Временно предположив, что потери пренебрежимо малы ($r=0$ в (1)), получаем

$$\frac{1}{2\rho} \left(\frac{s}{\omega^2} - m \right) k = \text{cth}(Hk). \quad (12)$$

В случае $\omega^2 < s/m$ это уравнение имеет решение с вещественным k , которое может быть найдено графически, как показано на фиг. 8.

Однако при $\omega^2 > s/m$ решение с вещественным значением постоянной распространения k отсутствует. Другими словами, бегущие волны выше частоты среза $\omega_c = \sqrt{s/m}$ не реализуются на БМ. Частота среза ω_c монотонно уменьшается с увеличением x , так как, согласно измерениям Бекеша, по мере увеличения расстояния x вдоль мембраны резко падает ее жест-

кость $s(x)$. (Напротив, масса, связанная со смещениями БМ, изменяется сравнительно мало.)

При $\omega^2 > s/m$ постоянная k становится чисто мнимой величиной. Если ввести постоянную затухания

$$\alpha = jk, \quad (13)$$

то (12) преобразуется в уравнение

$$\frac{1}{2\rho} \left(m - \frac{s}{\omega^2} \right) \alpha = \text{ctg}(H\alpha), \quad (14)$$

которое фактически имеет множество решений, указывающих на наличие целого семейства затухающих волн. Обращаясь снова к (8) и полагая $k = -j\alpha$, легко установить, что эти решения соответствуют *стоячим* волнам в направлении y , затухающим в направлении x . Имеется один «основной» тип колебаний (соответствующий резонансу в направлении y на половине длины волны), характеризуемый наименьшей величиной α , который будет испытывать наименьшее затухание и поэтому сохранится дольше всех типов справа от «характеристического положения» x_c , определяемого соотношением $s(x_c) = \omega^2 m(x_c)$.

При конечных (но небольших) потерях ($r(x) \neq 0$) постоянная распространения k будет, разумеется, комплексной величиной, непрерывно изменяющейся от почти действительной до почти чисто мнимой по мере прохождения волны через характеристический участок.

Амплитуда скорости смещения БМ (та же, что и амплитуда скорости жидкости вблизи обеих поверхностей БМ) определяется (в используемом грубом приближении) выражением

$$|v(x)| = \frac{2\omega\rho}{|Z(x)|} |\text{ch}(Hk(x))|. \quad (15)$$

Поскольку в формуле для $v(x)$ присутствует «резонансный знаменатель»

$$|Z(x)| = \left[\left(\omega m - \frac{s}{\omega} \right)^2 + r^2 \right]^{1/2}, \quad (16)$$

эта скорость будет высокой вблизи характеристического («резонансного») участка, где $|Z(x_c)|$ сводится к r .

Таким образом, можно прийти к заключению, что модель Лессера — Беркли в точности воспроизводит данные наблюдений Бекеша (бегущие волны, резонанс и затухание). Однако действительная ценность модели зависит от того, насколько она пригодна для количественного описания реальных процессов. До настоящего времени попытки количественно проверить модель не предпринимались — особенно в отношении варианта с независимо определенными (а не специально подобранными) параметрами $m(x)$, $r(x)$, $s(x)$. Поэтому вопрос, можно ли БМ количественно промоделировать «мембраной», характеризующейся переменным по месту импедансом, остается открытым. Учитывая сложность анатомии улитки и самой БМ, логично предположить, что потребуются более сложные модели. Действительно, были предложены и проанализированы модели [23], в которых БМ рассматривается как *пластинка* (а не мембрана).

С другой стороны, для приложений, требующих меньшей (а не большей) строгости, работа с моделью

может быть облегчена путем дальнейшего ее упрощения, особенно при отработке модели с помощью ЦВМ — одного из основных инструментов для изучения поведения модели в широком диапазоне изменения различных условий. В настоящее время такое моделирование затрудняется требуемым объемом вычислений, характерным для двумерных структур типа модели Лессера — Беркли. В связи с этим возникает вопрос, существуют ли одномерные приближенные модели БМ. В следующем разделе дан положительный ответ, но с некоторыми оговорками.

IV. ДЛИННО- И КОРОТКОВОЛНОВЫЕ МОДЕЛИ

А. Модель на основе электрической аналогии

Предположим, что длина бегущей волны на БМ в рассмотренной выше модели велика по сравнению с шириной канала улитки H ; тогда гиперболический котангенс в (11) может быть приближенно выражен формулой

$$\text{cth}(Hk(x)) \approx 1/Hk(x) \quad (17)$$

и, следовательно,

$$k^2(x) = \frac{-2j\omega\rho}{Z(x)H} \quad (18)$$

При тех же предположениях скорость смещения БМ окажется равной [см. (7) и (8)]

$$v(x) = \frac{2j\omega\rho}{Z(x)} \quad (19)$$

Специалистам, и особенно специалистам в области электротехники, должно быть очевидно, что (18) и (19) представляют собой уравнения некоторой (неоднородной) линии передачи с продольной индуктивностью на единицу длины

$$L_1(x) = 2\rho H \quad (20)$$

и поперечной полной проводимостью на единицу длины

$$Y(x) = \frac{1}{Z(x)}$$

Если эту механическую линию передачи представить ее электрическим аналогом, то скорость $v(x)$ будет соответствовать поперечному току на единицу длины, а разность давлений с обеих сторон БМ — напряжению на линии [22].

Одна секция длиной Δx такой линии передачи в виде электрического аналога с импедансом

$$Z(x) = j\omega m(x) + r(x) + s(x)/j\omega, \quad (21)$$

или

$$Z(x) = j\omega L + R + 1/j\omega C,$$

иллюстрируется фиг. 9.

Сразу становятся очевидными основные черты ее характеристик передачи и затухания: при небольших значениях x (соответствующих участку вблизи стре-



Фиг. 9. Схема небольшой секции базиллярной мембраны, составленная на основе электрических аналогий (одно-волновая модель). На низких частотах ($\omega \ll 1/\sqrt{LC}$) данная секция по существу выполняет функции линии задержки. На высоких частотах ($\omega \gg 1/\sqrt{LC}$) она представляет собой (индуктивный) аттенуатор. При $\omega = 1/\sqrt{LC}$ поперечный ток (соответствующий скорости смещения базиллярной мембраны) имеет резонансный пик. Различные частоты резонируют на различных участках, расположенных вдоль мембраны.

мечка) $Z(x) \approx 1/j\omega C$. Таким образом, линия передачи ведет себя по существу как линия задержки (с задержкой $\sqrt{L_1 C}$ на единицу длины). В другом крайнем случае, когда $s(x) \ll \omega^2 m(x)$, $Z(x) \approx j\omega L$ и линия действует по существу как индуктивный аттенуатор. На характеристическом участке, где $Z(x) = R$, поперечный ток (т. е. скорость смещения БМ) обладает ярко выраженным максимумом.

Полагая массу m и сопротивление r постоянными, а жесткость $s(x)$ изменяющейся по экспоненциальному закону, можно аналитически проинтегрировать (18) и получить приближенное выражение [22, 24] (пренебрегая концевыми эффектами) для напряжения (разности давлений) в точке

$$p(x, \omega) = p(0, \omega) \omega_m^{-1/2} [\omega_c^2 - (\omega - j\omega_c \delta/2)^2]^{1/4} \exp[-jK \arcsin(\omega/\omega_c - j\delta/2)], \quad (22)$$

где $\omega_c = \omega_c(x) = \sqrt{s/m}$, $\omega_m = \omega_c(0)$, а $\delta = \omega_c r/s$ — коэффициент потерь. Безразмерная постоянная K определяется формулой

$$K = \frac{2\rho H d}{m}, \quad (23)$$

а d — расстояние, на котором частота среза ω_c изменяется в e раз.

Для того чтобы получить скорость смещения БМ в виде функции положения x и частоты ω , давление $p(x, \omega)$ следует умножить на локальную поперечную полную проводимость $Y(x)$. Поскольку на характеристической частоте $Y(x)$ принимает весьма большое значение, можно заключить, что при $\omega = \omega_c(x)$ скорость смещения БМ имеет резонансный характер.

В. Коротковолновые модели

Другой одномерный приближенный вариант модели Лессера — Беркли основан на допущении, что длина волны на БМ повсюду мала по сравнению с шириной

H канала улитки. При таком предположении [25, 26] и при $\operatorname{Re}\{k\} > 0$ выражение (11) можно переписать в виде следующего чрезвычайно простого соотношения для постоянного распространения:

$$k(x) = -\frac{2i\omega\rho}{Z(x)}. \quad (24)$$

К сожалению, хотя длина волны на БМ велика по сравнению с H в основной (базальной) области, сделанное предположение приводит к не имеющему физического смысла решению за пределами резонансного участка [26]. Таким образом, подобные коротковолновые модели оказываются непригодными.

V. НЕЛИНЕЙНЫЕ МОДЕЛИ МЕХАНИЗМА ВНУТРЕННЕГО УХА

A. Комбинационные тоны

В слуховом восприятии проявляется целый ряд нелинейных эффектов, возникновение которых в значительной степени обусловлено нелинейностью механизма работы внутреннего уха. Среди этих нелинейных эффектов наиболее (и раньше других) известны некоторые комбинационные тоны, в частности комбинационный тон частоты $2f_1 - f_2$, иногда называемый (нижним) «кубическим дифференциальным тоном». По существу это один из «третьих тонов» (terzi suono Тартини), который возникает в случае, когда в наружное ухо поступают два первичных тона с частотами f_1 и f_2 ($f_2 > f_1$). Тот факт, что комбинационный тон (КТ) возникает во внутреннем ухе (а не в среднем ухе или в результате преобразований сигналов в нервной системе) неопровержимо вытекает из психоакустических и нейрофизиологических экспериментов.

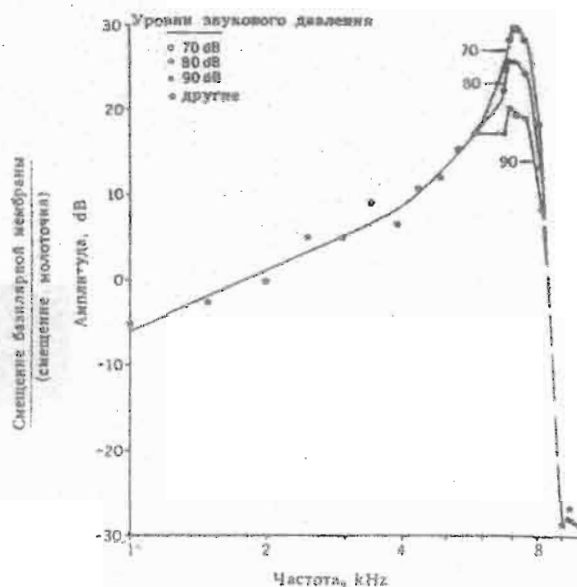
1. КТ с частотой $2f_1 - f_2$ становится слышимым при уровнях звука, близких к порогу слухового ощущения, когда среднее ухо еще обладает высоколинейными характеристиками.

2. Амплитуда КТ очень сильно зависит от разности частот первичных тонов, уменьшаясь со скоростью 100 дБ на октаву, что напоминает частотно-избирательные свойства БМ.

3. Нейрофизиологическими экспериментами [27] установлено, что в первичных слуховых нейронах, характеристическая частота которых близка к 3 кГц, практически отсутствует нервная активность (свыше спонтанного возбуждения), когда ухо животного стимулируется сигналами с частотами либо $f_1 = 4,0$ кГц, либо $f_2 = 5,0$ кГц, однако при одновременном воздействии обоих тонов появляется сильная активность, изменяющаяся синхронно с частотой $2f_1 - f_2$.

4. Согласно данным как психоакустических, так и нейрофизиологических экспериментов, КТ с частотой $2f_1 - f_2$ можно подавить, если воздействовать на наружное ухо третьим синусоидальным колебанием с частотой $2f_1 - f_2$ и с соответствующими амплитудой и фазой.

Фактически КТ во всех возможных ситуациях проявляет себя как первичный тон, волна возбуждения от которого бежит по БМ в направлении к вершине улитки. По всей видимости, сама БМ (а также те органы, которые принимают участие в ее движениях) представляет наиболее вероятное место возникновения КТ.



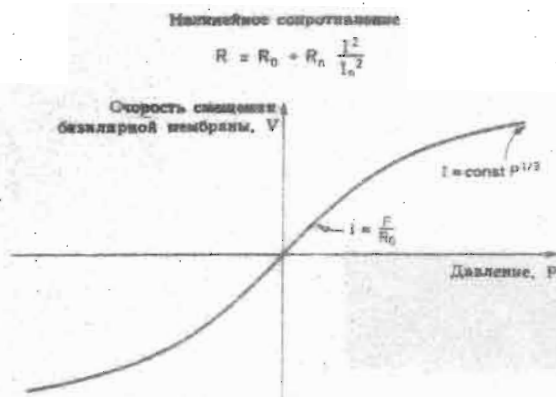
Фиг. 10. Относительная амплитуда колебаний базилярной мембраны в зависимости от частоты для трех уровней звукового давления (SPL). SPL измерены относительно порога слухового ощущения с номинальным значением $0,0002 \text{ дин/см}^2$. Наличие вблизи резонанса различных относительных амплитуд свидетельствует о нелинейном характере смещений базилярной мембраны [10]. Эта нелинейность является причиной некоторых нелинейных эффектов слуха, таких, например, как определенные «комбинационные тоны», которые «создает» само ухо из тонов, входящих извне. На существование комбинационных тонов впервые обратил внимание в XVIII в. итальянский скрипач Тартини [2].

Ключ к пониманию природы рассматриваемой нелинейности дают результаты измерений смещений БМ относительно смещений молоточка (фиг. 10); в этих экспериментах, проведенных Роде, использовался эффект Мессбауэра. По мере увеличения уровня звукового давления от 70 до 80 и 90 дБ относительная характеристика вблизи резонанса становится более плоской (а максимумы кривых несколько смещаются в сторону низких частот). Такие закономерности вполне объяснимы в рамках рассмотренной выше длинноволновой модели (см. фиг. 9), если сопротивление R , соответствующее потерям при смещениях БМ, увеличивается с возрастанием скорости смещения БМ.

B. Нелинейные потери

Оказалось, что на основе механизма нелинейных потерь, впервые предложенного Кимом и др. [28] (хотя и в рамках физически менее обоснованной модели), можно объяснить, по крайней мере качественно, а иногда и количественно, все известные нелинейности работы внутреннего уха. Чтобы учесть нелинейные эффекты, Хаббард и Гейслер [29] модифицировали длинноволновую модель, основанную на электрической аналогии, положив сопротивление R (см. фиг. 9) зависящим от протекающего через него тока.

Автор [30] исследовал, к каким результатам приводит предположение о квадратичном возрастании со-



Фиг. 11. Некоторые нелинейные эффекты во внутреннем ухе можно объяснить, сделав предположение о нелинейном характере потерь. В соответствии со схемой, составленной на основе электрических аналогий (фиг. 9), сопротивление R может зависеть от протекающего через него тока I , согласно приведенной выше формулы. В этом случае зависимость между скоростью смещения базиллярной мембраны и давлением становится нелинейной, приобретая, как здесь показано, характер амплитудного сжатия [30].

противления R с увеличением тока I

$$R = R_0 + \beta I^2, \quad (25)$$

где β — положительная постоянная.

При резонансе емкость C и индуктивность L , пропорциональные соответственно податливости и массе единицы длины БМ, обладают противоположными по знаку и равными по абсолютной величине импедансами, так что результирующий импеданс на характеристической частоте равен просто R . При напряжении возбуждения U (соответствующем перепаду давлений в жидкости, действующему на мембрану) ток («скорость смещения мембраны») равен

$$I = \frac{U}{R} = \frac{U}{R_0 + \beta I^2}. \quad (26)$$

В данном случае ток I связан с напряжением U кубическим уравнением. Его решение в графической форме (показанное на фиг. 11) представляет функцию амплитудного сжатия, которая при больших значениях напряжения описывается законом кубического корня

$$I = \gamma U^{1/3}, \quad (27)$$

где $\gamma = \beta^{-1/3}$.

Воздействие многотоновых входных сигналов на нелинейную модель БМ (включавшую 175 секций, подобных изображенной на фиг. 9, с зависящим от тока сопротивлением) было имитировано на ЦВМ [31] для изучения природы КТ. Однако в данном случае представляется более поучительным вывести некоторые простые аналитические соотношения без «маскирующего» эффекта сложного процесса моделирования на ЭВМ.

С. Двухтоновое подавление

Рассмотрим сначала нелинейное взаимодействие двух тонов, проявляющееся в электрической активности акустического нерва и называемое «двухтоновым подавлением». Если в наружное ухо поступает «чистый

тон» (синусоидальной формы) достаточной интенсивности, то в нервных волокнах, обладающих «характеристическими частотами», равными (или близкими) частоте акустического сигнала, возникают серии коротких импульсов, синхронных с акустическими колебаниями (стимулами). Если к первому тону добавить второй с несколько большей или меньшей частотой, то электрическая активность нервных волокон, обусловленная первым тоном, может быть подавлена при условии, что второй тон обладает достаточной громкостью — он должен быть обычно громче первого. Такое двухтоновое подавление было тщательно изучено Заком [32] и другими исследователями.

Причина двухтонового подавления была приписана, по крайней мере предположительно, эффекту нейронного торможения, называемому «латеральным торможением». Латеральное торможение: хорошо подтверждается в зрении и широко известно благодаря одному из его перцептуальных проявлений, а именно так называемых «полос Маха» [33]. Оно связывает нейроны, действующие друг на друга, посредством торможения (в противовес возбуждению), т. е. возбуждение тормозящего нейрона уменьшает вероятность возбуждения заторможенного нейрона.

Однако для объяснения двухтонового подавления, по-видимому, не требуется привлекать столь экзотических явлений, как латеральное торможение. В этом можно еще больше убедиться, исследуя структуру нервных волокон, которая способна увеличить латеральное торможение без какой бы то ни было связи с периферической системой слуха.

Чтобы упростить обсуждение, мы предположим, что нелинейность, вызывающая эффект двухтонового подавления, по существу локализована лишь в одном месте на БМ — вблизи «характеристического» участка БМ для подавляющего тона с частотой f_1 . Сигнал возбуждения при двухтоновом подавлении имеет следующую форму:

$$U = A \cos(2\pi f_1 t) + B \cos(2\pi f_2 t), \quad (28)$$

где A и B — амплитуды подавляемого и подавляющего тона соответственно. (Поскольку f_1 и f_2 , как правило, — несоизмеримые частоты, учет фазовых углов не обязателен).

При больших амплитудах скорость смещения базиллярной мембраны вблизи «характеристического» участка для f_1 определяется выражением

$$I = \gamma [A \cos(2\pi f_1 t) + B \cos(2\pi f_2 t)]^{1/3}. \quad (29)$$

Если пренебречь временными интервалами в окрестностях нулевых значений $\cos(2\pi f_1 t)$, то выражение (29) для случая $B \gg A$ можно разложить в ряд

$$I = \gamma B^{1/3} \cos^{1/3}(2\pi f_2 t) + \frac{1}{3} \gamma B^{-2/3} \cos^{-2/3}(2\pi f_2 t) \times \\ \times A \cos(2\pi f_1 t) + \dots \quad (30)$$

Вычисляя компоненту ряда Фурье I_1 для этого выражения на частоте f_1 , получаем

$$I_1 \propto A B^{-2/3}. \quad (31)$$

Таким образом, в рамках этих грубых приближений составляющая скорости смещения БМ на частоте f_1 уменьшается с увеличением B . Действительно, если составляющая f_1 сама не возбуждает исследуемый нерв,

то его результирующее возбуждение представляется величиной I_1 , которая, согласно (31), уменьшается с увеличением B . Это именно то явление, которое обнаруживается в экспериментах по двухтоновому подавлению, когда сама составляющая f_2 не возбуждает исследуемые нервные волокна.

Если A увеличивается, то для поддержания амплитуды смещения I_1 на постоянном уровне необходимо увеличивать B пропорционально $A^{2/3}$. Заксом получены данные [32, фиг. 3] о том, что при увеличении A на 5 dB для поддержания скорости возбуждения нерва на постоянном уровне требуется увеличить B на 7—9 dB, что согласуется с полученными выше теоретическими результатами (в соответствии с которыми B нужно увеличить на 7,5 dB, чтобы компенсировать эффект от увеличения A на 5 dB). По другим данным Закса [32, фиг. 6, c], при увеличении A на 10 dB постоянство скорости возбуждения обеспечивается увеличением B на 12—16 dB при теоретической величине 15 dB.

Закс также установил, что отношение «у. регулируемой» скорости возбуждения (скорость возбуждения минус спонтанная скорость) при двухтоновой стимуляции к управляемой скорости возбуждения при одном подавляемом тоне зависит только от отношения B/A .

Предположим, что управляемая скорость возбуждения r первичного слухового волокна является степенной функцией скорости смещения БМ:

$$r(I_1) \sim I_1^n. \quad (32)$$

Использование этого соотношения вполне оправдано, если скорость возбуждения далека от насыщения, и показатель степени n выбран близким к 2. Тогда при $B \ll A$, комбинируя (32) и (31), для случая двухтоновой стимуляции получаем

$$r_{AB} \sim A^n B^{-2n/3}, \quad (33)$$

а в случае единственной составляющей на частоте f_1 :

$$r_A \sim A^{n/3}. \quad (34)$$

Из (33) и (34) найдем отношение

$$\frac{r_{AB}}{r_A} = \left(\frac{B}{A}\right)^{-2n/3}, \quad (35)$$

которое действительно является функцией A/B . При $n=2$ показатель степени в (35) равен $-1,33$, что хорошо совпадает с одним из результатов, полученных Заксом для f_1 и f_2 в пределах одной критической полосы (показатель степени равен $-1,34$, см. фиг. 4, b в работе [32]).

Однако при большем разномес частот абсолютная величина показателя степени становится, как правило, меньше, получаемой из (35) — эффект, лежащий за пределами возможностей проведенного здесь упрощенного анализа с учетом нелинейности только на одном участке БМ. Чтобы получить дополнительное подтверждение нашей гипотезы о том, что двухтоновое подавление является результатом не нервной (тормозящей) деятельности, а механических нелинейностей, нужно провести значительно более подробный анализ с моделированием на ЭВМ.

Д. Комбинационный тон $2f_1 - f_2$

Если при двухтоновой стимуляции наружного уха [см. (28)] частоты f_2 и f_1 весьма близки ($f_2 > f_1$), а нелинейность, проявляемая БМ, соответствует упрощенной формуле (27), то в случае $A \cos(2\pi f_1 t) \gg B \cos(2\pi f_2 t)$ можно получить следующее биномиальное разложение для I :

$$I = 7A^{1/3} \cos^{1/3}(2\pi f_1 t) + \frac{1}{3}A^{-2/3} \cos^{-2/3}(2\pi f_1 t) \times B \cos(2\pi f_2 t). \quad (36)$$

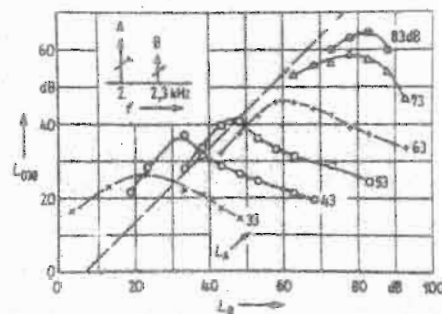
Анализ Фурье [34] показывает, что амплитуда скорости смещения БМ на частоте $2f_1 - f_2$ равна

$$I(2f_1 - f_2) \sim A^{-2/3} B. \quad (37)$$

Напротив, при $B \gg A$ анализ Фурье дает значение амплитуды КТ в виде

$$I(2f_1 - f_2) \sim A^2 B^{-5/3}. \quad (38)$$

Таким образом, если амплитуду низкочастотного тона A поддерживать постоянной, а амплитуду высокочастотного тона B увеличивать, начиная со значений $B \ll A$, то амплитуда комбинационного тона сначала будет расти пропорционально B . Как только B станет велико по сравнению с A , амплитуда КТ начнет уменьшаться по закону $B^{-5/3}$. Следовательно, где-то вблизи $L=A$ должно находиться максимальное значение амплитуды комбинационного тона. Именно такое достаточно любопытное немонотонное поведение и было обнаружено Цвикером [35] в ходе психоакустических экспериментов (фиг. 12).

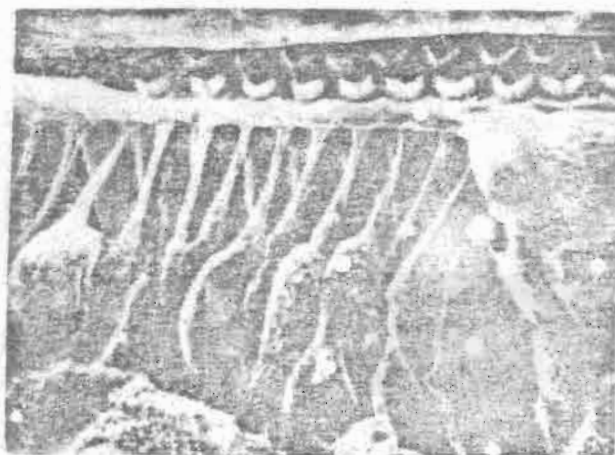


Фиг. 12. Уровень L_{AB} комбинационного тона $2f_1 - f_2$ (1,7 кГц в данном случае) начинает понижаться по мере увеличения верхнего первичного тона ($f_2 = 2,3$ кГц) выше определенного уровня. Такое специфически «немонотонное» поведение легко объясняется с помощью нелинейной характеристики, показанной на фиг. 11 [34].

Аналогично, если величина B поддерживается постоянной, а A увеличивается, то амплитуда КТ сигнала (пока $A \ll B$) возрастает пропорционально A^2 , а в дальнейшем (при $A \gg B$) уменьшается по закону $A^{-5/3}$. Следовательно, и в данном случае амплитуда КТ должна достигать максимального значения приблизительно при $A=B$. Такая закономерность действительно наблюдается экспериментально.

Наконец, если обе первичные амплитуды изменяются пропорционально друг другу

$$A \sim B,$$



Фиг. 13. Волосковые клетки, преобразующие во внутреннем ухе механические колебания в электрические импульсы. На верхушках клеток видны пучки волосков (отделенные на этой иллюстрации от текториальной мембраны). Изображение получено Спундлингом с помощью растрового электронного микроскопа (см. также [g]).

то рассматриваемый здесь случай нелинейности должен привести к асимптотической зависимости амплитуды КТ от кубического корня из первичной амплитуды

$$I(2f_1 - f_2) \sim A^{1/3}, \quad (39)$$

что находится в полном соответствии с данными Цвикера [36] и Гольдштейна [37].

Немонотонное поведение амплитуды КТ (когда одна из первичных амплитуд поддерживается постоянной) и ее зависимость по закону кубического корня (когда обе первичные амплитуды изменяются пропорционально друг другу) долгое время ставили акустиков в тупик, поскольку нормальная кубическая нелинейность, характеризующаяся соотношением вход-выход

$$y \sim x^3,$$

как можно показать, приводит к выражению для амплитуды КТ на частоте $2f_1 - f_2$

$$I(2f_1 - f_2) \sim A^2 B,$$

которое совершенно не соответствует экспериментальным данным.

Таким образом, то, что Цвикер назвал «необычным поведением амплитуды» КТ частоты $2f_1 - f_2$, естественно вытекает из простого предположения о квадратичном характере нелинейности сопротивления R (фиг. 9), представляющего собой потери в БМ.

VI. ПРЕОБРАЗОВАНИЕ МЕХАНИЧЕСКИХ КОЛЕБАНИЙ В НЕРВНОЕ ВОЗБУЖДЕНИЕ

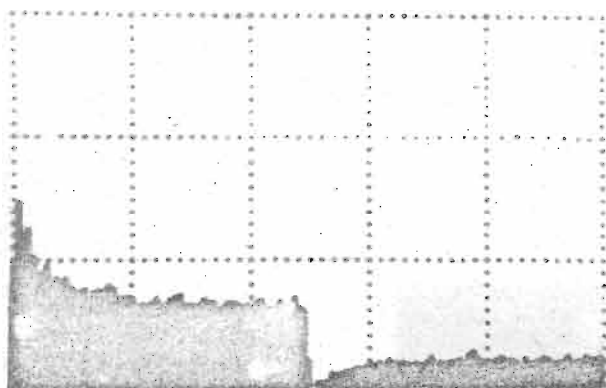
А. Некоторые сведения о первом возбуждении во внутреннем ухе

При попадании в наружное ухо акустической синусоидальной волны БМ приходит в движение, которое (приблизительно) описывается синусоидальной вре-



Фиг. 14. Три окончания слухового нерва (NE) у основания волосковой клетки. Изображение получено с помощью растрового электронного микроскопа. Изгиб волосков на верхушке клетки изменяет электрическую проводимость некоторой части мембраны у волосковой клетки, модулируя таким образом ток, вызванный разностью электрических потенциалов между клеткой и ее окружением. Перекрывающийся, по-видимому, способствующий освобождению «пузырька», содержащего небольшое количество химического передаточного вещества, которое перемещается к участку внутри волосковой клетки, расположенному напротив нервных окончаний [38]. Когда таких пузырьков накопится достаточно большое количество, в нервном окончании «возбуждается» электрический импульс, который через несколько промежуточных стадий обработки попадает в мозг.

менной функцией. Это движение передается волосковым клеткам, «сядающим» на БМ, через «реснички» — небольшие пучки волосков, растущих на каждой волосковой клетке. Считается, что эти волоски, по крайней мере те из них, которые расположены на наружных волосковых клетках, прикреплены к текториальной мембране (см. фиг. 13). Смещения БМ относительно текториальной мембраны приводят к изгибу (или к деформации либо смещению) этих небольших волосков, которые таким образом выполняют функцию, в какой-то степени похожую на работу иглы звукоснимателя в процессе преобразования механических вибраций в электрические сигналы. Однако если в случае иглы звукоснимателя и процесса электромеханического преобразования картина достаточно ясна, то подробности процесса преобразования в ухе до конца еще не изучены. Предполагается [38], что движения ресничек изменяют электрическую проводимость биологической мембраны у поверхности волосковых клеток, модулируя таким образом электрический ток, протекающий через волосковую клетку. Источниками этого тока, как установлено во многих работах, явля-



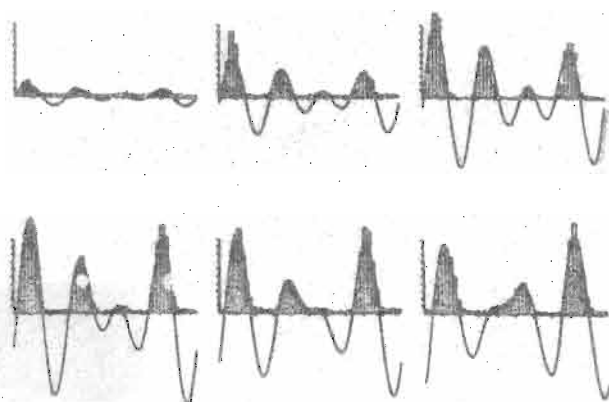
Фиг. 15. Нервная «адаптация». Частота нервных импульсов (вертикальная шкала) быстро возрастает в первый момент стимулирования, а затем медленно снижается со временем (горизонтальная шкала). После прекращения стимула (в центральной части фигуры) частота импульсов мгновенно падает почти до нуля, а затем медленно «восстанавливается» до нормальной «частоты спонтанного возбуждения» (обычно в пределах от 50 до 100 нервных импульсов в секунду). Отмеченные явления адаптации и восстановления довольно просто объясняются процессами истощения и регенерации химического передаточного вещества в волосковой клетке [39]. (См. [h].)

ются разности потенциалов между различными каналами (лестницами) внутреннего уха.

Полагают также, что этот переменный ток в свою очередь способствует выделению в волосковой клетке синаптических «пузырьков». Эти пузырьки, небольшие образования, содержащие химическое «передаточное» вещество, смещаются к внутренней поверхности волосковой клетки напротив окончания афферентного нервного волокна (фиг. 14), которое они и возбуждают. Результаты такого рода нервного возбуждения могут быть записаны с помощью микроэлектродов, крошечные кончики которых (диаметром менее одной тысячной миллиметра!) вводятся в тело нервного волокна.

Длительность электрических импульсов (также называемых спайками) в слуховом нерве составляет примерно 0,5 мс, а средняя частота повторения превышает 100 импульсов в секунду даже в отсутствие акустического возбуждения, т. е. «спонтанно». Типичная величина частоты спонтанного возбуждения составляет около 50 импульсов в секунду. При воздействии достаточно сильного установившегося акустического сигнала частота возбуждения может достигнуть 150 и более импульсов в секунду.

В начале воздействия сильного акустического сигнала частота возбуждения может достигнуть 1000 импульсов в секунду. Однако более высокие частоты уже невозможны вследствие «рефрактерного периода» (или «мертвого времени») нерва, наступающего после каждого возбуждения. Это мертвое время, в течение которого нервное волокно восстанавливает исходные свойства своей оболочки, имеет для слухового нерва величину порядка 1 мс. После начального периода возбуждающего сигнала высокая первоначальная частота возбуждения быстро спадает до так называемой *адаптированной частоты* (фиг. 15). Такая «адаптация» слухового нерва свойственна и другим нервным структурам.



Фиг. 16. «Периодические гистограммы», демонстрирующие вероятность нервного возбуждения (вертикальные полоски), и наложенные друг на друга акустические сигналы. Эта фигура показывает, что характеристика преобразования механических колебаний в нервные импульсы, осуществляемого в ухе, сходна с линейным однополупериодным выпрямлением [52].

Адаптация является одним из наиболее распространенных проявлений нервной деятельности, которое должно быть принято во внимание при построении любой сколько-нибудь серьезной модели. В модели, которая будет рассмотрена в настоящем разделе, адаптация осуществляется за счет механизма «истощения», т. е. механизма, когда для возбуждения присоединенного нерва расходуется содержимое квантов (пузырьков) химического вещества волосковой клетки.

Помимо адаптации и рефрактерности, важнейшими проявлениями слуховой нервной активности являются разновидности *линейности* (односторонней) и *автоматической регулировки усиления*. Линейность здесь означает, что вероятность возбуждения слухового нерва, определенная за много циклов периодического воздействия, точно (почти без искажений) повторяет «положительные» участки стимулирующего сигнала. Вероятность возбуждения отрицательными участками сигнала оказывается ниже уровня спонтанной частоты (фиг. 16). Таким образом, процесс преобразования, по крайней мере для ограниченного диапазона изменения амплитуд, оказывается весьма схожим с процессом *линейного однополупериодного выпрямления*.

Автоматическая регулировка усиления, присущая процессу преобразования, состоит в том, что в пределах определенного диапазона изменения амплитуд вероятность возбуждения почти не зависит от амплитуды сигнала (фиг. 17). Особо примечателен тот факт, что механизм регулировки усиления практически не искажает *форму* сигнала. Поэтому регулировка усиления в слуховых нейронах осуществляется по



Фиг. 17. Периодические гистограммы для двух уровней синусоидального акустического сигнала, отличающихся на 35 дБ. Вероятности нервного возбуждения только слегка различаются по форме. Общий результат (на 1000 периодов) от 124 возр-с лишь до 140. (См. [h].)

принципу, отличному от применяемого в мгновенном компрессоре, не говоря уже об амплитудном ограничителе. Ее действие напоминает работу регулятора громкости с временем отклика, приблизительно равным 20 мс.

В. Вероятностная модель

Хотя рассмотренный процесс преобразования носит довольно сложный характер тем не менее его можно описать с помощью математической модели [39], основанной на трех простых предположениях.

1. «Кванты» (электрохимического) реагента генерируются в волосковой клетке с некоторой постоянной средней скоростью (r квантов в секунду).

2. Кванты исчезают и тем самым возбуждают соответствующее афферентное нервное волокно с вероятностью в единицу времени, пропорциональной их количеству $n(t)$ и функции «проницаемости» $p(t)$, зависящей от входного сигнала.

3. Кроме того, кванты исчезают независимо от стимуляции, не вызывая при этом возбуждения нервного волокна, с вероятностью в единицу времени, равной $g - n(t)$, где g — некоторая постоянная.

Для функции проницаемости можно использовать следующую формулу «мягкого однополупериодного выпрямителя»:

$$p(t) = \frac{1}{2} p_0 \{x(t) + [x^2(t) + 1]^{1/2}\}. \quad (40)$$

Вероятность возбуждения в единицу времени $f(t)$ определяется (см. предположение 2) выражением

$$f(t) = n(t)p(t).$$

При достаточно высокой основной частоте периодического сигнала возбуждения число квантов за период сигнала будет оставаться относительно постоянным. Поэтому вероятность возбуждения в единицу времени может быть аппроксимирована выражением

$$f(t) \approx \bar{n}p(t), \quad (41)$$

т. е. можно считать, что она приблизительно пропорциональна $p(t)$.

При малых сигналах возбуждения сама величина $p(t)$ — p_0 пропорциональна сигналу, а при больших сигналах $p(t)$ пропорциональна сигналу на выходе однополупериодного выпрямителя; все это согласуется с данными нейрофизиологических измерений.

В установившемся режиме число генерируемых квантов в среднем равно числу исчезающих квантов, т. е.

$$r = \overline{n(t) \cdot p(t)} + \overline{n(t)g}, \quad (42)$$

где r — скорость генерирования квантов (см. предположение 1), $n(t)$ — количество квантов, а g — вероятность исчезновения кванта, приходящаяся на один квант в единицу времени (см. предположение 3).

Первый член правой части выражения (42) представляет собой среднюю вероятность \bar{f} возбуждения нерва. Поэтому в стационарных условиях

$$\bar{f} = r - \bar{n}g. \quad (43)$$

Из (41) для средней вероятности возбуждения получаем:

$$\bar{f} = \bar{n}p.$$

Отсюда с учетом (43) среднее число квантов \bar{n} определяется следующим приближенным выражением:

$$\bar{n} \approx r/(\bar{p} + g), \quad (44)$$

т. е. при больших сигналах возбуждения ($\bar{p} \gg g$) величина \bar{n} обратно пропорциональна амплитуде сигнала. Следствием этого является асимптотическая независимость вероятности возбуждения от амплитуды сигнала:

$$f(t) \approx rp(t)/(\bar{p} + g) \approx rp(t)/\bar{p}. \quad (45)$$

Такая «нормировка» вероятности возбуждения является одной из важнейших характеристик, наблюдаемых на периодических гистограммах; она выведена здесь непосредственно из предположений, на которых основана обсуждаемая модель.

С. Модель в виде RC-цепи

Интересно отметить, что электрическим аналогом поведения рассматриваемой модели в среднем может служить простая схема с RC-цепью, показанная на фиг. 18; она содержит генератор тока с параллельно включенными постоянной проводимостью g , емкостью C и переменной проводимостью $p(t)$.

Ток источника r соответствует скорости генерирования квантов, ток $f(t)$, протекающий через переменную проводимость, — вероятности возбуждения нерва, заряд, накопленный на конденсаторе C , — среднему числу квантов реагента в волосковой клетке.

Поскольку схема на фиг. 18 содержит четыре элемента (параметра), а модель только три, то один из этих параметров, например емкость C , может быть выбран произвольно. Остальные три должны удовлетворять следующим соотношениям:

$$\begin{aligned} r &= f_{\max}, \\ p_0 &= C f_0 f_{\max} T, \end{aligned} \quad (46)$$

и

$$g = C(f_{\max} - f_0)f_{\max}T.$$

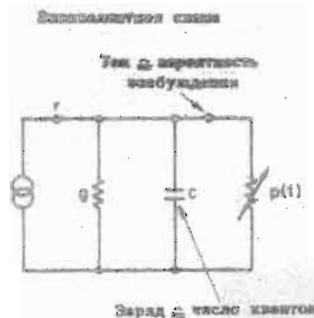
Многие свойства рассматриваемой модели, включая те, которые уже обсуждались, можно воспроизвести с помощью схемы, показанной на фиг. 18. Так, например, истощение квантов соответствует разряду конденсатора C , когда под воздействием акустического сигнала увеличивается проводимость $p(t)$ (см. (40)). Аналогичным образом, восстановление скорости возбуждения до спонтанной, наступающее после сильного акустического воздействия, соответствует увеличению тока через проводимость p по мере перезарядки конденсатора C .

Ожидаемое количество квантов удовлетворяет простому дифференциальному уравнению

$$dn/dt = r - n(t)[p(t) + g], \quad (47)$$

решение которого имеет вид

$$n(t) = \left[r \int_0^t p(\tau) d\tau + n(0) \right] / f(t), \quad (48)$$



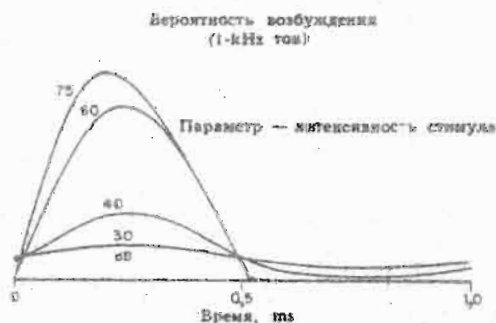
Фиг. 18. Схема, построенная на основе электрических аналогов для модели [39] процесса преобразования механических колебаний в нервные импульсы, осуществляемого волосковой клеткой. Проводимость $r(t)$ изменяется под действием стимула по закону однопериодного выпрямления. Высокая средняя проводимость разряжает конденсатор C , ограничивая тем самым ток, представляющий собой вероятность возбуждения нерва, присоединенного к волосковой клетке. Полагают, что такой процесс «опорожнения» конденсатора объясняет нервную «адаптацию», которую иллюстрирует фиг. 15.

где

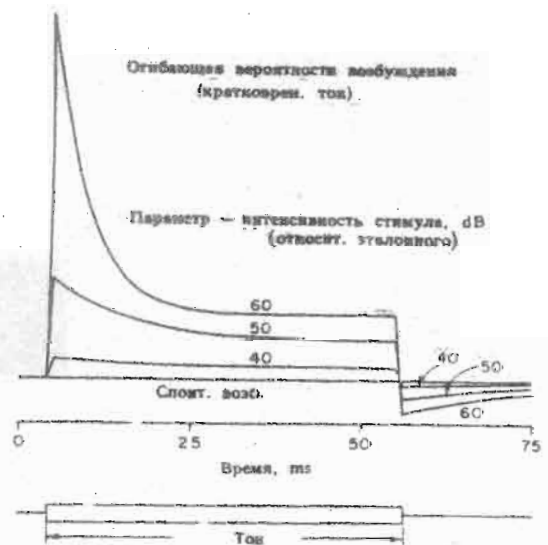
$$I(t) = \exp \left[gt + \int_0^t p(\theta) d\theta \right]. \quad (49)$$

На фиг. 19 показаны расчетные кривые вероятностей возбуждения при воздействии синусоидального сигнала с различными амплитудами. Они соответствуют периодическим гистограммам, полученным в результате нейрофизиологических исследований, демонстрирующих хорошо известный переход от «линейной» реакции на сигнал к нормированному отклику типа однопериодного выпрямления.

На фиг. 20 показан ряд кривых усредненных вероятностей возбуждения при воздействии коротких импульсов тона с различными амплитудами. На этих кривых, соответствующих «постстимульным» гистограммам, можно наблюдать обычно возникающий начальный выброс, последующую адаптацию, подавление скорости возбуждения (сразу после пре-



Фиг. 19. Вероятности возбуждения модели [39] для синусоидального акустического сигнала при четырех входных уровнях. Вероятность возбуждения линейно связана с мгновенными значениями сигнала при уровнях не выше 30 dB. При уровнях выше 60 dB вероятность возбуждения соответствует сигналу после однопериодного выпрямления и при дальнейшем увеличении уровня сигнала изменяется незначительно. Результаты, полученные с помощью этой модели, соответствуют периодическим гистограммам, показанным на фиг. 17.



Фиг. 20. Результат моделирования, иллюстрирующий черную адаптацию при возбуждении кратковременным тоном. Внизу показана продолжительность тона (50 мс). (Ср. с фиг. 15.)

кращения стимулирующего сигнала) до величины, меньшей спонтанной скорости в установившемся режиме, и медленное возвращение (в данном случае с постоянной времени $\tau = 20$ мс) к спонтанной скорости возбуждения после окончания импульса.

Другие важные данные нейрофизиологических экспериментов связаны с интервальными гистограммами, отражающими вероятность появления интервалов различной длительности между следующими друг за другом импульсами. Математический анализ дискретно-временного варианта рассматриваемой модели, проведенный Логаном и Шеппом [40] при отсутствии входного сигнала (которая в этом случае формирует марковский процесс), показывает, что распределение интервалов подчиняется закону геометрической прогрессии. Этот результат хорошо совпадает с данными по измерению нервных импульсов [41], за исключением эффектов рефракторности.

На этом мы закончим рассмотрение моделей БМ и преобразований энергии в нервной системе. В следующем разделе мы вернемся к одной из первоначальных психоакустических проблем — вопросу о том, чувствительно ли наше ухо только к амплитудному спектру акустических стимуляторов или оно все же может — в нарушение акустического закона Ома — выделять различные фазовые комбинации.

VIII. МОНАУРАЛЬНАЯ ФАЗОВАЯ ЧУВСТВИТЕЛЬНОСТЬ

A. Могут ли наши уши распознать фазовые сдвиги или люди невосприимчивы к фазе звука?

Одна из наиболее волнующих проблем теории слухового восприятия — та, которая ставила в тупик создателей моделей, психоакустиков и совсем недавно начала занимать энтузиастов высококачественной записи и любителей — касается способности или неспособности уха человека воспринимать фазу сигнала. Еще в середине прошлого века Ом сформулиро-

вал свой знаменитый закон, согласно которому «слуховое восприятие зависит только от амплитудного спектра звука и не зависит от фазовых углов различных частотных составляющих этого спектра».

Для того чтобы избежать очевидных нарушений акустического закона Ома, мы вынуждены несколько уточнить его формулировку. Так, в приведенной выше цитате перед выражением «амплитудный спектр» нужно добавить «кратковременный». Если этого не сделать, то легко придумать, например, такой пример, противоречащий закону фаз:

1) выделим речевой сигнал длительностью 100 с;
2) определим его фурье-преобразование (предполагая, что сигнал периодически повторяется каждые 100 с), в результате получим частотные составляющие, расположенные через каждые 1/100 Hz;

3) проведем рандомизацию фазовых углов, т. е. выберем каждый фазовый угол из равномерного распределения между 0 и 2π rad;

4) произведем обратное фурье-преобразование.

В результате формируется сигнал, который в течение промежутков времени короче 100 с достаточно для практики точно выглядит и звучит подобно гауссовскому шуму (со спектром мощности, равным спектру мощности исходного речевого сигнала). Таким образом, операция над фазовыми углами привела в данном случае не только к изменению тембра акустического сигнала, но по сути дела к полному превращению сигнала из (разборчивой) речи в (случайный) шум.

Если бы фурье-преобразование проводилось в интервале не 100 с, а 50 ms, что примерно соответствует времени анализа речи при слуховом восприятии, то рандомизация не изменила бы тембра звука по крайней мере при качественной оценке его на слух.

Таким образом, говоря о «фазовой глухоте» уха, мы должны помнить, что это относится к кратковременным спектрам.

Чтобы объяснить то удивительное свойство нечувствительности нашего слуха к фазе и, следовательно, к форме сигнала, Ом и Гельмгольц предложили следующую модель: 1) в ухе существует набор «полосовых фильтров», перекрывающих весь диапазон звуковых частот, и 2) соответствующие органы в ухе измеряют амплитуду на выходе каждого фильтра, и в мозг передается только эта информация.

Считалось, что полосовые фильтры реализованы в БМ в виде «настроенных струн». Естественно, что волосковым клеткам отводилась роль устройств для измерения амплитуды.

В конце прошлого и первой половине нашего столетия велась острая дискуссия о том, насколько справедлив акустический закон Ома. Как оказалось, большая часть опубликованных данных о невыполнении этого закона была получена на недостаточно хорошей аппаратуре, генерировавшей продукты искажений, взаимодействие которых (с полезным сигналом и друг с другом) в самом деле зависело от относительных фазовых углов. (Вспомним, что многие из ранних работ по восприятию фаз были выполнены до того, как появились электронные фильтры и усилители!)

Тем не менее важная доля истинных фазовых эффектов не была опровергнута, — по-видимому, лучше всего известен демонстрирующий фазовые эффекты «АМ — ЧМ»-эксперимент [42]. Акустическим возбудителем в этом эксперименте служит АМ-сигнал несущей частоты (скажем, 2000 Hz), частота модуляции

равна, например, 100 Hz. Такой сигнал состоит из трех составляющих с частотами 1900, 2000 и 2100 Hz.

Если теперь фазу одной из боковых частот — 1900 Hz либо 2100 Hz — изменить на 180° , то АМ-сигнал превратится в «квази-ЧМ-сигнал» (КЧМ-сигнал) («квази», потому что частотная модуляция сопровождается небольшой амплитудной модуляцией: с удвоенной модулирующей частотой).

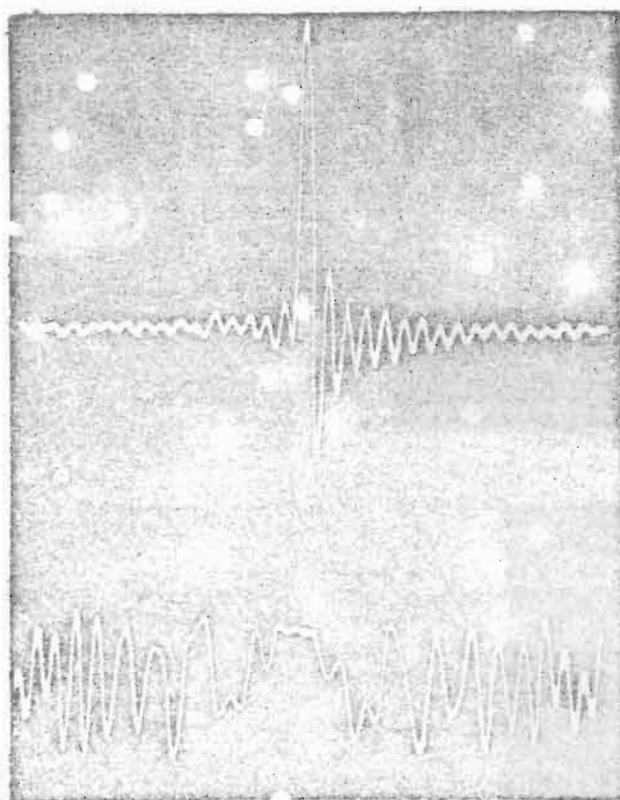
В процессе поочередного прослушивания АМ- и КЧМ-сигналов тембр их звучания воспринимается существенно различно. (Жаль, что Ом и Гельмгольц не довелись слышать эти сигналы. Это поколебало бы их уверенность при формулировке закона фаз. Они были бы вынуждены отказаться от своих упрощенных идей — и от этого пострадали бы их последователи, чьи работы основывались на модели слуха, включающей набор настроенных фильтров. Пессимистичное обобщение, с которым, например, был вынужден работать Гельмгольц, иногда помогает формулировать великие идеи — даже если эти идеи в конце концов оказываются лишь приблизительными описаниями. Осмелился бы Ньютон сформулировать свои законы движения и тяготения, если интерпретация данных Кеплера была бы осложнена учетом релятивистских эффектов?)

Результаты АМ — ЧМ-эксперимента возродили интерес к способности уха воспринимать форму сигнала (а не только его амплитудный спектр). Автор настоящей статьи, например, заинтересовался возможным влиянием формы сигнала на качество синтезированной речи. В какой степени невысокое качество синтезированной вокодером речи [43] обусловлено влиянием формы сигнала? В соответствии с концепциями Гельмгольца вокодеры воспроизводят кратковременный амплитудный спектр речевых сигналов безотносительно к фазовым углам анализируемого речевого сигнала. В частности, речевые сигналы синтезируются из квазипериодических последовательностей импульсов, все гармоники которых имеют одинаковые (нулевые) фазовые углы. Если эти фазовые углы рандомизировать, то получится сигнал со значительно меньшим пик-фактором (определяемым как частное от деления размаха сигнала на его среднеквадратичное значение). Результат такой рандомизации фазовых углов показан на фиг. 21. Фактически ЭВМ создала 4096 случайных комбинаций фаз, и на фиг. 21 представлен сигнал с наименьшим пик-фактором.

При прослушивании двух сигналов, изображенных на фиг. 21, обнаруживается неожиданно большая разница в качестве их звучания — и это несмотря на идентичность их амплитудных спектров (31 гармоника равной амплитуды от 100 до 3100 Hz). Синтезированные речевые сигналы, полученные из этих двух сигналов «возбуждения», также звучали различным образом (с меньшим «гудением» в случае сигнала с низким пик-фактором).

Сначала эту разницу в качестве звучания объяснили большим различием в пик-факторе. Однако последующие эксперименты с сигналами с такими же низкими пик-факторами показали, что каждый из этих сигналов звучит по-разному, за исключением случая, когда их фазовые углы связаны «линейным» законом [44, 45], т. е. когда два набора фазовых углов ϕ_k и ψ_k связаны соотношением

$$\psi_k = \phi_k + \alpha + \beta k, \quad (50)$$



Фиг. 21. Один период сигнала, составленного из 31 гармоник с равными амплитудами. Вверху: все гармоники имеют примерно нулевую фазу. Внизу: тот же самый амплитудный спектр, но фазовые углы 0 либо π гдѣ выбирались таким образом, чтобы минимизировать «пик-фактор» (оба сигнала представлены в одинаковом масштабе). При основной частоте ниже 200 Hz [44] эти два сигнала имели совершенно различное звучание, что противоречит акустическому закону Ома [3], который утверждает, что ухо «глухо к фазе». Автору удалось сыграть простые мелодии путем изменения фазовых углов отдельных гармоник изображенного вверху сигнала, сохраняя их амплитуды неизменными («фазовый орган»).

где α и β — произвольные постоянные, а k — номер гармоники.

В выражении (50) член βk вызывает простую задержку на $(\beta/2\pi f_1)s$, где f_1 — основная частота.

Более важную роль играет член α , поскольку он приводит к изменению формы сигнала, хотя, как уже было указано, качество звучания при этом остается неизменным.

Какая же характеристика сигнала $s(t)$ инвариантна при изменении фазы каждой гармоники на одну и ту же величину?

Ответ: огибающая сигнала $e(t)$, определяемая соотношением

$$e^2(t) = s^2(t) + \hat{s}^2(t), \quad (51)$$

где $\hat{s}(t)$ — так называемое преобразование Гильберта от $s(t)$:

$$\hat{s}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(\tau)}{t - \tau} d\tau. \quad (52)$$

С помощью преобразования Гильберта может быть определен аналитический сигнал $\sigma(t)$

$$\sigma(t) = s(t) - j\hat{s}(t), \quad (53)$$

фурье-образ которого для всех отрицательных частот стремится к нулю.

Аналитический сигнал может быть записан иначе с использованием определенной выше огибающей $e(t)$ и фазовой функции $\gamma(t)$:

$$\sigma(t) = e(t) \exp[j\gamma(t)], \quad (54)$$

где

$$\gamma(t) = \arctg[-s(t)/\hat{s}(t)]. \quad (55)$$

Поскольку в сигнале $\sigma(t)$ отсутствуют отрицательные частоты, сдвиг фазы на α эквивалентен умножению $\sigma(t)$ на $\exp(j\alpha)$:

$$\sigma_\alpha(t) = e(t) \exp[j\gamma(t) + j\alpha]. \quad (56)$$

Соответствующий сдвинутый по фазе реальный сигнал определяется соотношением

$$s_\alpha(t) = e(t) \cos[\gamma(t) + \alpha], \quad (57)$$

откуда видно, что его огибающая $e(t)$ не зависит от фазового сдвига.

Другими словами, преобразование фазы вида $\psi_k = \phi_k + \alpha + \beta k$, которое не изменяет качества звучания сигнала, не изменяет и его огибающую (вносится лишь не ощущаемая на слух и, на моноaurальном восприятии задержка).

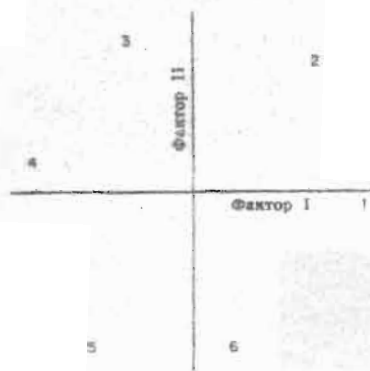
В таком случае самое естественное — это предположить, что ухо действует как детектор огибающей волн на БМ, которые оно «воспринимает». Именно такая точка зрения и была впервые выдвинута автором [44] и де Буром [45] и получила широкое признание среди акустиков.

К сожалению, гипотезы слухового восприятия огибающей и правило линейного преобразования фазы вступают в противоречие с данными некоторых экспериментов по слуховому восприятию сигналов, содержащих только две частотные составляющие. Для таких сигналов любое преобразование фазы является линейным преобразованием вида (50), а огибающая определяется исключительно амплитудным спектром. Поэтому любые преобразования таких сигналов не будут ощутимы на слух. Однако Джеффрис и др. [46] показали, что это не так, хотя обнаруженные ими фазовые эффекты и были весьма тонкими.

В. Проверка гипотезы огибающей

В 1971 г. Холлом и автором [47] был проведен эксперимент с шестью сигналами, имевшими две частотные составляющие: $f_1 = 200$ Hz и $f_2 = 400$ Hz; у составляющей с $f_1 = 200$ Hz фаза всегда была равна 0 ($\phi_1 = 0$), а составляющая с $f_2 = 400$ Hz могла обладать 6 различными фазовыми углами: $\phi_2 = 0, \pi/3, 2\pi/3, \pi, 4\pi/3$ и $5\pi/3$.

Эти сигналы были предложены слушателям в процессе так называемых «триадных» сравнительных испытаний.



Фиг. 22. Анализ психологических факторов при вынесении оценок несхожести стимулов, составленных из двух синусоидальных колебаний с частотами 200 и 400 Гц и различными фазовыми углами верхней гармоники (0, 60, 120, 180, 240, 300°). Точками 1—6 представлены шесть различных сигналов. Их расстояния друг от друга связаны монотонной зависимостью с их субъективными различиями, определенными тремя тренированными слушателями [47]. «Пространство восприятия» (субъективное), показанное на этой фигуре, тесно связано с пространством сигналов (физическим): верхние гармоники всех шести сигналов имеют одинаковые амплитуды, но различаются по фазе с шагом в 60°. Это типичный «случай из учебника», который представляет взаимосвязь между физическим стимулом и психологической реакцией так, как она истолковывается «многомерным психологическим шкалированием». (Дополнительную информацию по вопросам «психофизики» см. в [1].)

В любой момент времени слушатель мог прослушать три из шести сигналов и должен был сказать, какие два из этих трех сигналов наиболее близки по звучанию и какие два — наименее близки. Полученные таким образом оценки схожести были использованы для построения «пространства подобия», где исследуемые шесть сигналов были представлены точками таким образом, что евклидовы расстояния между ними были монотонно связаны со степенью их несхожести.

Пространство подобия, полученное в результате этих экспериментов, показано на фиг. 22; шесть сигнальных точек разместились почти точно на окружности, а их угловые расстояния оказались близко соответствующими разностям фаз сигналов.

Таким образом, при двухтоновом сигнале не только воспринимается разность фаз, но и сами результаты дают некоторую закономерную картину. Кроме того, они характеризуются хорошей воспроизводимостью и мало меняются при смене слушателей.

С «гипотезой огибавшей» приходится распрощаться. Оказывается, ухо все-таки ощущает разницу в форме сигнала по крайней мере при определенных условиях — даже в том случае, когда их огибавшие одинаковы.

В экспериментах с двухтоновыми сигналами при отношении частот $f_1 : f_2 = 2 : 3$ Холл [48] наблюдал монауральные фазовые эффекты, причину которых можно было однозначно усмотреть во взаимодействии «разностного тона» $f_2 - f_1$ и комбинационного тона $2f_1 - f_2$; частоты каждого из этих тонов равны $f_1/2$, а фазы отсчитываются на $\Delta\phi = 2\phi_2 - 3\phi_1$. В этом случае амплитуда результирующей составляющей, рассматриваемая в функции ϕ_2 , должна иметь период, равный π , — именно так, как и установил Холл.

Эти (и другие) экспериментальные данные указывают, что некоторые монауральные фазовые эффекты или даже большая их часть являются результатом зависящей от фазы интерференции продуктов искажений во внутреннем (и среднем) ухе [49]. Эта гипотеза нашла дальнейшее подтверждение в опытах Баннена [53], который показал, что область существования монауральных фазовых эффектов в значительной степени совпадает с областью существования комбинационных тонов.

С. Новая модель монаурального восприятия фаз: внутренний спектр

Изложенное позволяет сформулировать новую модель монаурального восприятия фаз, основанную на представлении о «внутреннем спектре». Эта модель возрождает первоначальную гипотезу восприятия качества звучания «только за счет амплитудного спектра» с существенной оговоркой, что спектр сигнала, поступающего в наружное ухо, видоизменяется под действием нелинейных искажений, возникающих в среднем и внутреннем ухе. Спектр, полученный в результате такой модификации, называется внутренним спектром; именно он представляет собой ту информацию, которая преобразуется в волосковыми клетками в нервные импульсы и подвергается дальнейшей обработке в слуховых нервных путях вплоть до акустического отдела коры головного мозга.

До какой степени внутренний спектр инвариантен к линейным преобразованиям фазы внешнего акустического сигнала? Наиболее заметными являются те искажения, которые обусловлены только двумя соседними частотными составляющими. (Это следует из того, что искажения во внутреннем ухе резко понижаются с увеличением разности частот, а искажения в среднем ухе, которые не обнаруживают подобной чувствительности к разности частот, весьма незначительны при уровнях звука, обычных для слухового восприятия.)

Среди продуктов искажений, обусловленных взаимным влиянием соседних частотных составляющих, наиболее слышными являются те, чьи «индексы» отличаются на ± 1 . Так, для двух первичных составляющих f_1 и f_2 следует принимать во внимание следующие продукты искажений:

$$\begin{aligned} 2f_1 - f_2, \\ 3f_1 - 2f_2, \\ 4f_1 - 3f_2 \text{ и т.д.} \end{aligned}$$

или в общем виде $F_m = mf_1 - (m-1)f_2$. Фазовые углы ψ_m этих составляющих зависят от фазовых углов первичных составляющих ϕ_1 и ϕ_2 следующим образом:

$$\psi_m = m\phi_1 - (m-1)\phi_2. \quad (58)$$

Линейное преобразование фазы внешнего сигнала

$$\phi'_1 = \phi_1 + \alpha + \beta f_1$$

приводит к появлению у рассматриваемой составляющей нового фазового угла ψ' :

$$\psi' = \psi + \alpha + \beta[mf_1 - (m-1)f_2]. \quad (59)$$

В квадратных скобках здесь записана частота составляющей F_m . Таким образом,

$$\psi' = \psi + \alpha + \beta F_m, \quad (60)$$

т. е. фаза продукта искажений подвергается тому же самому линейному преобразованию, что и фаза внешнего акустического сигнала. Следовательно, если внешний сигнал содержит составляющую с частотой F_m , то его фаза изменится точно на ту величину, на какую изменится фаза продукта искажений, возникающих внутри уха. Таким образом, результирующая амплитуда сигнала с частотой F_m (внутренний спектр) остается неизменной при любом линейном преобразовании фаз! Инвариантными являются как *огбающая* внешнего сигнала, так и *внутренний спектр*, и поэтому больше нельзя утверждать, что при линейном преобразовании фаз отсутствуют наблюдаемые моноaurальные фазовые эффекты объясняется лишь «детектированием *огбающей*» в ухе. Модель внутреннего спектра приводит точно к такому же результату!

Д. Фазовый орган

Модель внутреннего спектра может объяснить результаты весьма любопытного эксперимента по моноaurальному восприятию фаз, проведенного автором в 1958 г. в фирме Bell Telephone Laboratories. В качестве одного из сигналов возбуждения в этих экспериментах была выбрана импульсная последовательность с частотой повторения 100 импульсов в секунду, обработанная ФНЧ с частотой среза 3100 Hz. Ее спектр состоял из 31 гармоники с равными амплитудами и нулевыми фазами. Затем, когда фаза одной из гармоник была изменена на π , стал слышен тон наряду с создаваемым импульсами гудением, высота которого соответствовала частоте той гармоники, фаза которой была изменена.

Появление этого тона вначале было истолковано по аналогии с тем, что происходит с формой сигнала: сигнал измененной формы получали из исходной импульсной последовательности путем *вычитания* синусоиды соответствующей частоты и удвоения амплитуды инвертированной гармоники. Эту синусоиду можно было наблюдать на осциллографе, и автор был настолько наивен, что удовлетворился этим объяснением. Действительно, если фазы изменялись в соответствии с некоторой простой мелодией (например, «Mary had a little lamb...»), то эта мелодия воспринималась вполне отчетливо, хотя на протяжении всего времени звучания амплитудный спектр оставался неизменным (плоским). Были приняты, конечно, все меры, чтобы исключить какие бы то ни было переходные явления. Иначе говоря, после каждого «фазового тона» следовала достаточно длинная пауза, чтобы устранить влияние сигналов на БМ и явлений последствия в процессе адаптации нервной системы.

Справедливость модели восприятия фаз, основанной на изменении формы сигнала, получила еще одно подтверждение благодаря тому, что фазовый тон не прослушивался при воспроизведении через громкоговоритель в реверберационной камере, которая приводит к *рандомизации* фазовых углов, а также когда исходный фазовый спектр обладал случайным характером. В этих случаях в сигнале не появлялось четко различимой синусоидальной составляющей после изменения его фазы. Отсюда «само собой» следовало, что

отчетливого π на нельзя услышать — по крайней мере так тогда считалось.

Существовало, правда, одно обстоятельство, которое нельзя было объяснить в рамках модели, основанной на изменении формы сигнала: волосковые клетки никогда не «воспринимают» форму сигнала в целом так, как она видна экспериментатору на экране осциллографа. Волосковые клетки воспринимают лишь небольшую узкополосную часть сигнала, в которой не может появиться «дополнительная» синусоида при изменении фазового угла.

В модели восприятия фаз, основанной на понятии внутреннего спектра, не возникает такого рода проблем, связанных с восприятием формы сигнала, и хотя для объяснения фазовой чувствительности весьма соблазнительно привлечь представление о форме сигнала, сигнал никогда не может достичь слуховой нервной системы в непреобразованном виде. Внутренний спектр импульсной последовательности по существу является плоским. (Продукты нелинейных искажений с сопоставимыми амплитудами «добавляются» к плоскому спектру внешнего сигнала. А поскольку нелинейность имеет «компрессирующий» характер, то это добавление фактически приводит к *уменьшению* амплитуд внешнего спектра.)

Если фазу одной гармоники изменить на π , то искажения фактически будут приводить к увеличению амплитуды внутреннего спектра на частоте этой гармоники. (На более низких частотах появятся дополнительные изменения.) Иначе говоря, внутренний спектр будет теперь не плоским, а «выпуклым», что при достаточной большой величине «горба» окажется ощутимым на слух.

Вопрос о том, в какой мере эта модель способна объяснить различные моноaurальные фазовые эффекты, в настоящее время изучается коллективами ученых в различных частях земного шара. Однако уже сейчас ясно, что все моноaurальные фазовые эффекты невозможно объяснить только через внутренний спектр. Ниже рассмотрены два примера — один очевидный, а другой — не столь очевидный.

Е. Что нельзя объяснить через понятие внутреннего спектра

Возьмем сигнал, состоящий из множества гармоник с равными амплитудами и фазовыми углами 0, 0, π , π , 0, 0, π , π , 0, 0... Один период такого сигнала формируется из одного положительного импульса, за которым следуют три равноотстоящих отрицательных импульса (с амплитудами, равными одной трети амплитуды положительного импульса). При основной частоте 1 Hz слышно четыре щелчка в секунду. Если теперь все фазы установить равными нулю, то три отрицательных импульса в каждом периоде повторения исчезнут и останется только положительный импульс. При этом вместо четырех щелчков будет восприниматься только один щелчок в секунду. Такая перемена не может быть отнесена за счет изменений во внутреннем спектре, но совершенно естественно может быть увязана с изменением формы сигнала.

Имеются и другие примеры такого рода, где полученные результаты легко объясняются изменениями формы сигнала, однако они не представляют особого интереса, поскольку рассматриваемые эффекты проявляются только при достаточно низких скоростях повторения, требующих «долговременного» анализа

(по-видимому, в высших мозговых центрах, включая открытое сознание, как в приведенном выше примере с подсчетом числа импульсов). Рассмотренный выше случай отрезка речевого сигнала длительностью 100 с, фазовые углы которого были рандомизированы и который превратился поэтому в шумовой сигнал, является еще одним примером, выпадающим из области кратковременного слухового анализа и не дающим возможности сделать какие-либо выводы по восприятию фаз.

Существуют и другие экспериментальные данные, не требующие, однако, длительного анализа, которые вступают в противоречие с моделями внутреннего спектра (и отбрасываются). Это вынуждает нас подойти с других позиций к моделям изменения формы сигнала.

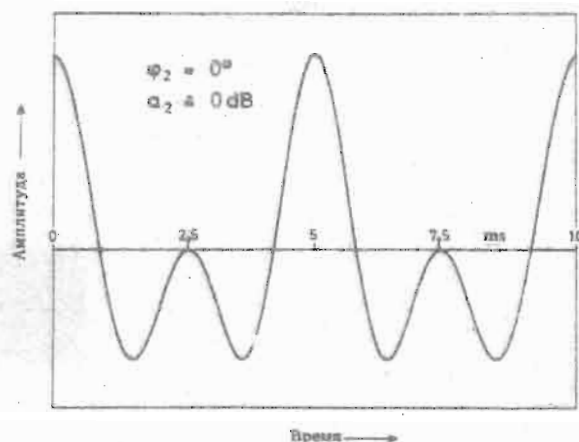
Г. Пороги маскировки и фаза

Тэрхардт и Фастл [50] поставили весьма остроумный эксперимент, в результате которого был получен, по-видимому, один из наиболее сильных монауральных фазовых эффектов. Интересно отметить, что для стимуляции они использовали в точности такой же акустический сигнал, с которым проводили эксперименты Холл и автор [44] в описанных выше триадных испытаниях на подобию: 200- и 400-Гц тоны с различными относительными фазовыми углами $s(t) = a_1 \cos(2\pi f_1 t) + a_2 \cos(2\pi f_2 t - \phi_2)$. Но Тэрхардт и Фастл ставили перед слушателями более определенные вопросы: вместо того, чтобы спрашивать, какие из предлагаемых стимулов звучат более похоже, а какие менее похоже, они спрашивали, удается ли услышать 400-Гц тон в присутствии 200-Гц тона. Изменяя амплитуду 400-Гц тона до тех пор, пока он не будет слышан, они измеряли так называемые «пороги маскировки», выше которых 200-Гц тон уже не «маскирует» 400-Гц тон.

Основной результат эксперимента Тэрхардта — Фастла представлен на фиг. 23: при изменении фазы 400-Гц тона с $\phi_2 = 0^\circ$ на $\phi_2 = 180^\circ$ этот тон можно услышать только при увеличении его амплитуды более



Фиг. 23. Сдвиг «порога маскировки» 400-Гц тона по мере изменения его фазового угла ϕ_2 от 0 до 360° . Маскирующим является 200-Гц тон. При $\phi_2 = 180^\circ$ для удержания 400-Гц тона на едва слышимом уровне «то амплитуду необходимо увеличить на 12 дБ. Еще один монауральный фазовый эффект? (Результаты измерений взяты из [50]. О результатах моделирования см. в [51].)



Фиг. 24. Сигнал, используемый для расчета интервальной гистограммы, состоящий из 200- и 400-Гц тонов с равными амплитудами при фазе 400-Гц тона $\phi_2 = 0^\circ$. При $\phi_2 = 180^\circ$ вместо двух небольших минимумов, как показано на фигуре, в сигнале появятся один большой минимум.

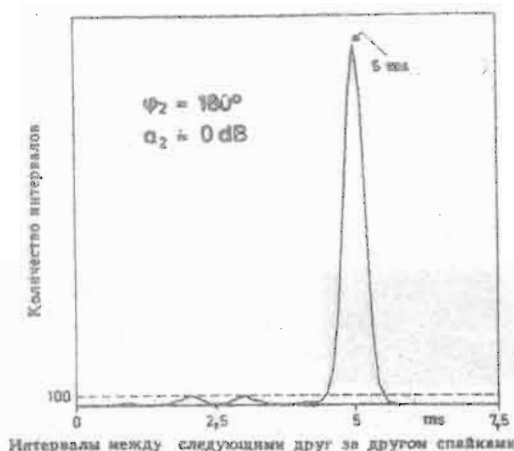
чем на 12 дБ. Иначе говоря, «порог маскировки» 400-Гц тона («маскируемого») жестко связан с его фазовым углом относительно 200-Гц тона («маскирующего»).

Как это можно объяснить? Очевидно, изменение относительной фазы приводит к резкому изменению формы сигнала. В самом деле, при $\phi_2 = 180^\circ$ форма меняется на обратную (умножается на минус единицу). Так, вместо одного положительного пика на основном периоде сигнала длительностью 5 мс при $\phi_2 = 0^\circ$ (см. фиг. 24), при $\phi_2 = 180^\circ$ появятся два положительных пика, отстоящих друг от друга на 2,5 мс.

Теперь предположим, что сигнал с заданной фазой 400-Гц тона попадает на БМ в неизменном виде и волосковые клетки «просматривают» его отрицательную часть. При $\phi_2 = 0^\circ$ они будут «видеть» (отрицательный) пик через каждые 2,5 мс и сигнализировать о «наличии 400-Гц тона». При $\phi_2 = 180^\circ$ до тех пор, пока не будет существенно повышена амплитуда 400-Гц тона, они «видят» (отрицательные) пики с интервалом 5 мс.

Насколько нужно увеличить эту амплитуду? По данным психоакустических измерений, на 12 дБ. Можно ли разработать некоторую количественную модель и что будет из нее следовать?

Предположим, что нервные импульсы, генерируемые волосковыми клетками, обрабатываются таким образом, что мозгу сообщается информация о расстояниях во времени между следующими друг за другом импульсами. Вполне допустимо, что такие механизмы (для измерения временных интервалов) существуют где-то вдоль слуховых путей к мозгу. В самом деле, поскольку нервные импульсы имеют примерно одинаковые амплитуды, то переносимая ими информация заключена именно во временном законе их следования. Такая временная информация является, конечно, не абсолютной (для синхронизации нервных импульсов люди не носят в своей голове кварцевые часы!), а относительной — по отношению к другим импульсам в тех же самых либо «соседних» нервных волокнах. (Это наиболее отчетливо проявляется при бинауральном восприятии звука, когда локализация в горизонтальной плоскости в основном определяется временем следования нервных импульсов.)



Фиг. 25. Интервальная гистограмма, полученная при моделировании преобразователя механических колебаний в нервные импульсы; стимул — двухтоновый сигнал (см. фиг. 24), $\phi_2 = 180^\circ$.

Предположим, что слуховой процессор будет вычислять «гистограмму» (т. е. распределение частота — появления) временных интервалов между следующими друг за другом импульсами в одной из ветвей нервных волокон. Какой вид будет иметь такая гистограмма? В рамках простой пороговой модели нервного возбуждения для рассмотренного выше сигнала, состоящего из 200- и 400-Гц тонов, эта гистограмма будет выглядеть так, как показано на фиг. 25. Относительная фаза ϕ_2 равна 180° . Для интервалов импульсов $\Delta t = 5$ ms (200-Гц тон) кривая имеет выраженный максимум и два небольших максимума в окрестностях $\Delta t = 2,5$ ms (400-Гц тон).

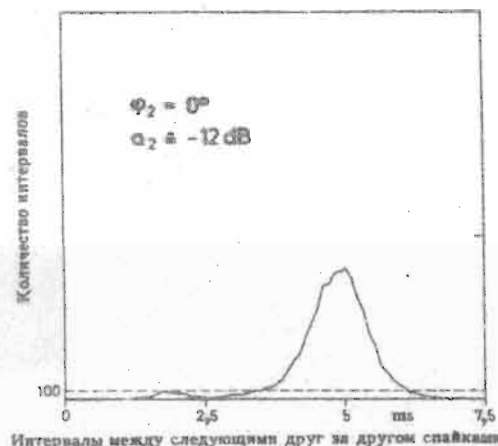
Предположим, что при $\Delta t = 2,5$ ms малые максимумы таковы, что они соответствуют порогу их восприятия слуховым анализатором и могут быть отождествлены с наличием 400-Гц тона.

Если фазовый угол изменить теперь на $\phi_2 = 0^\circ$, то величина максимума в окрестностях 2,5 ms резко возрастет. Чтобы эти максимумы гистограммы оставить на едва заметном («пороговом») уровне, амплитуда 400-Гц тона должна быть понижена — читатель легко может догадаться сам — на 12 dB! Соответствующая гистограмма ($\phi_2 = 0^\circ$, $a_2 = -12$ dB) показана на фиг. 26.

Такой же пороговый «эксперимент» был повторен путем машинного моделирования для всех фазовых углов ϕ_2 между 0 и 315° с приращением 45° . И для каждого ϕ_2 определялась соответствующая амплитуда a_2 400-Гц тона, при которой максимумы на интервальной гистограмме в окрестностях 2,5 ms имели пороговую величину. Полученные результаты показаны крестиками на фиг. 23 вместе с психоакустическими результатами пороговой маскировки, полученными Тэрхардтом и Фастлом. Как видно, совпадение очень хорошее.

VIII. ЗАКЛЮЧЕНИЕ

Мы охватили исключительно широкую область как в историческом плане (от Лукреция до 1974 г.), так и по существу дела: от дифракции у головы человека и механики внутреннего уха до преобразования и обработки сигналов в нервной системе, от физических и



Фиг. 26. То же самое, что на фиг. 25, но $\phi_2 = 0^\circ$. Амплитуда a_2 400-Гц тона была понижена на 12 dB с целью получения сходной гистограммы, подсчитанной вблизи 2,5 ms, которая отражала бы одинаковую степень восприятия 400-Гц тона в соответствии с моделью интервальной гистограммы, полученной автором для смещений порогов маскировки. Результаты для других фазовых углов ϕ_2 показаны крестиками на фиг. 23. Совпадение с экспериментальными данными хорошее.

физиологических измерений до психоакустических исследований слухового восприятия.

Многие из представлений, описанных выше, лишь прикл. женны и неполны, а многие очень интересные вопросы и вовсе опущены. Среди них теории восприятия высоты звука Ликлайдера [54] и Гольдштейна [55], модели бинаурального взаимодействия Дарлаха [56] и Колбурна [57], наблюдения Тэрхардта [58] о тембре и резкости звучания музыкальных аккордов, измерения потенциалов улитки [d] Даллоса и монументальная работа Стивенса [59] по психоакустическому шкалированию громкости, а также исследования психоакустических свойств звука [60]. Тем не менее автор надеется, что ему удалось по крайней мере выразить своего рода «ощущение» и, возможно, понимание того, в каком направлении ведутся исследования слухового восприятия и на каком этапе мы сегодня находимся в понимании одного из наших важнейших чувств: слухового ощущения.

Манфред Р. Шрёдер родился в Германии. Имеет степень доктора физических наук.

После переезда в США он стал штатным сотрудником фирмы Bell Telephone Laboratories, Марри-Хилл, шт. Нью-Джерси. В 1958 г. был назначен руководителем отделения акустических исследований; в 1963 г. — директором научно-исследовательской лаборатории акустики и речи, а в 1964 г. — директором научно-исследовательской лаборатории акустики, речи и механики, приняв ответственность за всю программу исследований в области акустики и ультразвука. В Bell Telephone Laboratories он занимался фундаментальными исследованиями в области архитектурной акустики, электроакустики, распространения звука под водой, речи и слуха. Он имеет около 40 соответствующих патентов США. В 1969 г. Шрёдер был назначен профессором физики и директором Третьего физического института Геттингенского университета. В 1972 г. исполнял обязанности заведующего кафедрой физики. Он принимал активное участие в многочисленных работах, выполняемых для правительства Соединенных Штатов. Входил в состав консультативной группы Федеральной комиссии связи «Национальный комитет по стереофоническому радиовещанию», которая выработала стандарты на стереофоническое радиовещание в США. Шрёдер — член Объединенного комитета Вооруженных сил и Национального исследовательского совета США по иссле-

дованиям в области слуха и биомеханики. В 1966 г. он был членом Совета по новым техническим и прикладным наукам Национальной консультативной комиссии при президенте по вопросам охраны здоровья людских ресурсов. В 1963 г. был приглашен Правительством СССР для консультации по акустике Кремлевского Дворца съездов. Он также являлся помощником редактора *Journal of the Acoustical Society of America*.

РЕКОМЕНДАЦИЯ ДЛЯ ДАЛЬНЕЙШЕГО ЧТЕНИЯ

- [a] W. A. van Bergeljk, J. R. Pierce, and E. E. David, Jr., *Waves and the Ear*. Garden City, New York: Doubleday, 1960.
- [b] H. Fletcher, *Speech and Hearing in Communication*. Princeton, N.J.: Van Nostrand, 1958.
- [c] E. Zwicker and F. Feldtkeller, *Das Ohr als Nachrichtenempfänger*. Leipzig, Germany: S. Hirzel Verlag, 1967.
- [d] P. Dallos, *The Auditory Periphery*. New York: Academic Press, 1973.
- [e] S. S. Stevens, F. Warshawsky, and the editors of L. E. *Sound and Hearing*. New York: Time Inc., Book Division, 1965.
- [f] W. R. Loewensin, Ed., *Principles of Receptor Physiology*. Berlin, Germany: Springer, 1971.
- [g] H. Spoendlin, *The Organization of the Cochlear Receptor*. White Plains, New York: S. Karger, 1966.
- [h] N. Kiang, *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve*. Cambridge, Mass.: M.I.T. Press, 1965.
- [i] G. V. Gersuni, Ed., *Sensory Processes at the Neuronal and Behavioral Levels*. New York: Academic Press, 1971.
- [j] D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics*. New York, Wiley, 1966.
- [k] В качестве вводного материала дополнительно можно воспользоваться цитированной выше литературой [7, 8, 14, 43].

ТРУДЫ ПОСЛЕДНИХ СИМПОЗИУМОВ

- [1] R. Plomp and G. F. Smoorenburg Eds., *Frequency Analysis and Periodicity Detection in Hearing*. Leiden, The Netherlands: Sijthoff, 1970.
- [2] M. B. Sachs Ed., *Physiology of the Auditory System*. Baltimore, Maryland: National Educational Consultants, 1971.
- [3] B. L. Cardozo, Ed., *Hearing Theory*. Eindhoven, The Netherlands: Institute for Perception Research, 1972.
- [4] A. R. Møller, Ed., *Basic Mechanisms in Hearing*. New York: Academic Press, 1973.
- [5] E. Zwicker and E. Terhardt, Eds., *Facts and Models in Hearing*. Heidelberg, Germany: Springer, 1974.

ЛИТЕРАТУРА

- [1] Лукреция, «О природе вещей». Книга четвертая. Изд-во АН СССР, 1945.
- [2] G. Tartini, *Trattato di Musica Seconda la Vera Scienza dell'Armonia*. Padua, Italy, 1754.
- [3] G. S. Ohm, *Ann. Phys. (Leipzig)*, vol. LIX, p. 513; vol. LXII, p. 1.
- [4] A. Sebeck, "Über die Definition des Tones," *Ann. Phys. Chem.*, vol. 63, pp. 353-368, 1844.
- [5] J. F. Schouten, "The perception of pitch," *Philips Tech. Rev.*, vol. 5, pp. 286-294, 1940.
- [6] J. F. Schouten, R. J. Ritsma, and B. L. Cardozo, "Pitch of the midline," *J. Acoust. Soc. Amer.*, vol. 34, pp. 1418-1424, 1962.
- [7] H. Helmholtz, *On the Sensation of Tone*. New York: Dover, 1954.
- [8] G. von Békésy, *Experiments in Hearing*. New York: McGraw-Hill, 1960.
- [9] B. M. Johnstone, K. T. Taylor, and A. J. Boyle, "Mechanics of the guinea pig cochlea," *J. Acoust. Soc. Amer.*, vol. 47, pp. 504-509, 1970.
- [10] W. S. Rhode, "Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique," *J. Acoust. Soc. Amer.*, vol. 49, pp. 1218-1231, 1971.
- [11] L. Kohlöffel, "A study of basilar membrane vibrations I. Punctum-detection: A new method for the analysis of microvibrations with laser light," *Acustica*, vol. 27, pp. 49-65, 1972.
- [12] J. P. Wilson and J. R. Johnstone, "Capacitive probe measures of basilar membrane vibration," in *Hearing Theory*, B. L. Cardozo, Ed. Eindhoven, The Netherlands: Institute for Perception Research, 1972, pp. 172-181.
- [13] W. Helfenstein, "Messung der Basilarmembranschwingungen," dissertation, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, 1974.
- [14] J. Blauert, *Räumliches Hören*. Leipzig, Germany: S. Hirzel Verlag, 1974.
- [15] Personal demonstration at CRS Laboratories, Stamford Conn., ca. 1960.
- [16] P. Laws, "Zum Problem des Entfernungshörens und der Im-Kopf-Lokalisierung von Hörereignissen," dissertation, Tech. Hochschule, Aachen, Germany, 1972.
- [17] Personal demonstration at Bell Laboratories, Murray Hill, N. J., 1974.
- [18] M. R. Schroeder and B. S. Atal, "Computer simulation of sound transmission in rooms," *IEEE Conv. Rec.*, pt. 7, pp. 150-155, 1963. (See also: M. R. Schroeder, "Digital simulation of sound transmission in reverberant spaces," *J. Acoust. Soc. Amer.*, vol. 47, pp. 424-431, 1970.)
- [19] M. R. Schroeder, "Computer models for concert hall acoustics," *Amer. J. Phys.*, vol. 41, pp. 461-471, 1973.
- [20] A. R. Møller, "Acoustic reflex in man," *J. Acoust. Soc. Amer.*, vol. 34, pp. 1524-1534, 1962.
- [21] M. B. Lesser and D. A. Berkeley, "Fluid mechanics of the cochlea, part 1," *J. Fluid Mech.*, vol. 51, pt. 3, pp. 497-512, 1972.
- [22] M. R. Schjeder, "An integrable model for the basilar membrane," *J. Acoust. Soc. Amer.*, vol. 53, pp. 429-434, 1973. (See also: J. Zwislöck, "Theory of the acoustical action of the cochlea," *J. Acoust. Soc. Amer.*, vol. 22, pp. 78-784, 1950; and L. C. Peterson and B. P. Bogert, "A dynamic theory of the cochlea," *J. Acoust. Soc. Amer.*, vol. 22, pp. 369-381, 1950.)
- [23] C. R. Steele, "Behavior of the basilar membrane with pure tone excitation," *J. Acoust. Soc. Amer.*, vol. 55, pp. 148-162, 1974.
- [24] G. Zweig, R. Lipetz, and J. R. Pierce, "The cochlear compromise," submitted to *J. Acoust. Soc. Amer.*
- [25] O. F. Ranke, "Hydromechanik der Schneckenflüssigkeit," *Zeitschrift f. Biologie*, vol. 103, pp. 409-434, 1950.
- [26] W. M. Siebert, "Ranke revisited—a simple short-wave cochlear model," *J. Acoust. Soc. Amer.*, vol. 56, pp. 594-600, 1974.
- [27] J. L. Goldstein and N. Y. S. Kiang, "Neural correlates of the aural combination tone $2f_1 - f_2$," *Proc. IEEE*, vol. 56, pp. 981-992, June 1968.
- [28] D. O. Kim, C. E. Molnar, and R. R. Pfeiffer, "A system of nonlinear differential equations modeling basilar membrane motion," *J. Acoust. Soc. Amer.*, vol. 54, pp. 1517-1529, 1973.
- [29] A. E. Hubbard and C. D. Geisler, "A hybrid-computer model for the cochlear partition," *J. Acoust. Soc. Amer.*, vol. 51, pp. 1895-1903, 1972.
- [30] M. R. Schroeder, "Recent advances in hearing research—nonlinear mechanics and neural transduction," in *Proc. 8th Int. Congr. Acoustics*, ("Invited Lectures"), pp. 47-61, 1974.
- [31] J. L. Hall, "Two-tone distortion products in a nonlinear model of the basilar membrane," submitted to *J. Acoust. Soc. Amer.*
- [32] M. B. Sachs, "Stimulus-response relation for auditory-nerve fibers: Two-tone stimuli," *J. Acoust. Soc. Amer.*, vol. 45, pp. 1025-1036, 1969.
- [33] F. Ratcliff, *Mech. Bands*. San Francisco, Calif.: Holden-Day, 1965.
- [34] M. R. Schroeder, "The amplitude behavior of the cubic difference tone," submitted to *J. Acoust. Soc. Amer.*
- [35] E. Zwicker, "Der kubische Differenzton und die Entstehung der Gehör," *Acustica*, vol. 20, pp. 206-209, 1968.
- [36] —, "D - ungewöhnliche Amplitudengang der nichtlinearen Verzerrungen des Ohrs," *Acustica*, vol. 5, pp. 67-74, 1955.
- [37] J. L. Goldstein, "Auditory nonlinearity," *J. Acoust. Soc. Amer.*, vol. 41, pp. 676-689, 1967.
- [38] A. Flock, "The physiology of the individual hair cell and their synapses," in *Basic Mechanisms in Hearing*, A. R. Møller, Ed. New York: Academic Press, 1973, pp. 273-306.
- [39] M. R. Schroeder and J. L. Hall, "Model for mechanical to neural transduction in the auditory receptor," *J. Acoust. Soc. Amer.*, vol. 55, pp. 1055-1060, 1974.
- [40] B. F. Logan and L. A. Shepp, "A birth and death model of neuron firing," to be published.
- [41] Мольнар, Пфайффер, «Интерпретация конфигураций спонтанной спайковой активности нейронов кохлеарного ядра», *ТИИЭР*, т. 56, № 6, стр. 110-123 (июнь 1968).
- [42] R. C. Matney and R. L. Miller, "Phase effects in monaural perception," *J. Acoust. Soc. Amer.*, vol. 19, pp. 780-797, 1947.
- [43] Шрёдер, «Вокодеры: анализ и синтез речи (Обзор прикладных исследований речевых сигналов за последние 30 лет)», *ТИИЭР*, т. 54, № 5, стр. 5-22 (май 1966).
- [44] —, "New results concerning monaural phase sensitivity," *J. Acoust. Soc. Amer.*, vol. 31, p. 1579 (A), 1959. (More details on this work can be found in J. R. Pierce, "Some work on hearing," *Amer. Scientist*, vol. 48, pp. 40-45, 1960.)
- [45] E. de Boer, "A note on phase distortion and hearing," *Acustica*, vol. 11, pp. 182-184 (L), 1961.
- [46] J. H. Craig and L. A. Jeffress, "Effects of phase on the quality of a two-component tone," *J. Acoust. Soc. Amer.*, vol. 34, pp. 1752-1760, 1962.
- [47] J. L. Hall and M. R. Schroeder, "Monaural phase effects for two-tone signals," *J. Acoust. Soc. Amer.*, vol. 51, pp. 1882-1894, 1972.
- [48] J. L. Hall, "Monaural phase effect: Cancellation and reinforcement of distortion products $f_3 - f_1$, and $2f_1 - f_2$," *J. Acoust. Soc. Amer.*, vol. 51, pp. 1872-1881, 1972.

- [49] T. Bunn, J. M. Fester, F. A. Bilsen, and G. van Brink, "Phase effects in a three-component signal," *J. Acoust. Soc. Amer.*, vol. 55, pp. 297-303, 1974.
- [50] E. Terhardt and H. Fast, "Zum Einfluss von Störtönen und Störgeräuschen auf die Tonhöhe von Sinustönen," *Acustica*, vol. 25, pp. 53-61, 1971.
- [51] M. R. Schroeder, "Neural models in auditory perception," in *Proc. 20th Int. Congr. Psychology* (Symp. Mathematical Models of Perception), Tokyo, 1972.
- [52] J. E. Rose, J. E. Hind, D. J. Anderson, and J. F. Brugge, "Some effects of stimulus intensity on response of auditory nerve fibers in the squirrel monkey," *J. Acoust. Soc. Amer.*, vol. 54, pp. 685-699, 1971.
- [53] T. Bunn, "Monaural phase perception and auditory nonlinearity," in *Proc. 8th Int. Congr. Acoustics*, vol. 1, p. 152, 1974.
- [54] J. Licklider, "Auditory frequency analysis," in *Information Theory*, E. C. Cherry, Ed. London, England: Butterworth, 1956, p. 253.
- [55] J. L. Goldstein, "An optimum processor theory for the central formation of the pitch of complex tones," *J. Acoust. Soc. Amer.*, vol. 54, pp. 1496-1516, 1973.
- [56] P. J. Metz, G. von Bismarck, and N. I. Durlach, "Further results on binaural unmasking and the EC mode. II. Noise bandwidth and interaural phase," *J. Acoust. Soc. Amer.*, vol. 43, pp. 1085-1091, 1968.
- [57] H. S. Colburn, "Theory of binaural interaction based on auditory nerve data. I. General strategy and preliminary results on interaural discrimination," *J. Acoust. Soc. Amer.*, vol. 54, pp. 1458-1470, 1973.
- [58] E. Terhardt, "Pitch, consonance, and harmony," *J. Acoust. Soc. Amer.*, vol. 55, pp. 1061-1069, 1974. (See also: E. Terhardt, "Zur Tonhöhenwahrnehmung von Klängen. II. Ein Funktionsschema," *Acustica*, vol. 26, pp. 187-199, 1972.)
- [59] S. S. Stevens, "Sensory power functions and neural events," in *Principles of Receptor Physiology*, W. R. Loswstein, Ed. Berlin, Germany: Springer, 1971, pp. 226-242.
- [60] J. J. Zwillocki, "In search of physiological correlates of psychoacoustic characteristics," in *Basic Mechanics in Hearing*, A. R. Møller, Ed. New York: Academic Press, 1973, pp. 787-806.