

Слышимость временного расплывания и несогласованности во времени акустических сигналов

M.N. Kunchur¹

Department of Physics and Astronomy
University of South Carolina, Columbia, SC 29208
(Dated: 19.07.2007 [received]; 29.08.2007 [published])

Technical Acoustics, <http://www.ejta.org>, 2007, 17.

Copyright **Technical Acoustics** [ISSN 1819-2408] (2007) EEA.

Перевод С. В. Гапоненко

Рассогласованность по времени между динамическими головками в акустической системе и временное расплывание сигнала в компонентах и кабелях долгое время рассматривались как возможная причина деградации качества звуковоспроизведения. Отмечалось также, что слушатели предпочитают более высокие частоты дискретизации (напр., 96 кГц), чем частота 44,1 кГц цифровых компакт-дисков, даже хотя частота Найквиста 22 кГц в последнем случае превышает номинальный высокочастотный предел слышимости одиночных тонов $f_{\max} = 18$ кГц. Эти качественные и хорошо известные наблюдения указывают на возможность того, что человеческий слух, возможно, обладает чувствительностью к временным ошибкам τ , которые короче, чем обратная угловая частота $[2\pi f_{\max}]^{-1} = 9$ мкс, что делает необходимым расширять полосу частот аудиоаппаратуры до величин, больших f_{\max} , чтобы обеспечить качество. Слепые тесты настоящей работы дают количественное подтверждение этого путем анализа различимости временного рассогласования сигналов пространственно смещенных громкоговорителей. Эксперимент показал, что порог смещения равен $d = 2$ мм, что соответствует различимости времени задержки примерно 6 мкс.

Введение

Идеальная звуковоспроизводящая цепочка будет воспроизводить точную копию исходного акустического сигнала, принятого записывающим микрофоном. В этом случае конечный воспроизведенный акустический сигнал будет иметь такую же форму во времени и такой же Фурье спектр, как и исходный сигнал. По практическим соображениям цепь воспроизведения не может иметь ни бесконечно малое время отклика, ни бесконечно широкую полосу. Однако эти свойства не являются необходимыми или желательными, поскольку ухо имеет собственные ограничения и по времени, и по частоте.

Обычно время отклика τ и высокочастотная граница f_{\max} идут «рука об руку». Для линейных систем существует обратное соотношение между τ и $\omega_{\max} (=2\pi f_{\max})$ – граничной круговой частотой. Однако механизм слуха сложен и нелинеен. Более того, извлечение спектральной и временной информации осуществляется с вовлечением совершенно различных цепочек рефлекторной дуги (нейронных цепей) с различными типами нейронов (обладающих различными скоростями отклика) и различными топологиями нейронных цепей. В результате временное разрешение не связано напрямую с наивысшей слышимой частотой и предельное время τ может быть намного короче, чем $1/\omega_{\max}$.

В звуковоспроизводящей системе также имеются сложности отклика (такие, как следствие диэлектрической релаксации, механических колебаний в кабелях, реверберации в корпусах акустических систем и другие процессы, приводящие к запасанию и медленному высвобождению энергии), не позволяющие считать ее линейной системой, что также разрушает простую связь между τ и $1/\omega_{\max}$. Вследствие этого адекватный частотный отклик не должен гарантировать, что система будет обладать звуковой прозрачностью (т.е. не будет создавать слышимого ухудшения сигнала). В аудио-сообществе

¹ Electronic address: kunchur@sc.edu ; URL: <http://www.physics.sc.edu/kunchur>

признается, что размывание по времени является ключевым фактором нарушения прозрачности [1,2] и что временное рассогласование может вызывать слышимые ошибки в спектральном отклике [3]. Кроме этих научных работ, популярная литература и реклама, адресованные энтузиастам hi-end-аудио, изобилуют заявлениями о важности «временного согласования» динамиков акустической системы (по этой причине некоторые модели динамиков имеют наклонные панели), о «когерентности» малых динамиков относительно больших (отсюда следует предпочтение некоторыми слушателями компактных мониторов по сравнению с большими дипольными панелями), о «временной когерентности» кабелей (т.е. предотвратить дисперсию, чтобы все частоты приходили одновременно), и множество иных эффектов, относящихся к расплыванию во времени. Как отмечалось в литературе [4], существуют известные заявления слушателей о замечаемом улучшении воспроизведения при частотах дискретизации выше частоты дискретизации компакт-диска (CD) 44,1 кГц даже при том, что слушатели не могут слышать чистые тона выше частоты Найквиста 22 кГц. Такие тонкие эффекты могут быть замаскированы во многих массовых коммерческих аудиосистемах и в аудиометрической аппаратуре, которая используется в психоакустических исследованиях, вследствие ограниченного разрешения системы – узкое место, возникающее вследствие ограничений аппаратуры, а не слуха.

В литературе описаны несколько предыдущих попыток по исследованию величины минимальной временной особенности τ в источнике, различимой на слух. Пороговые значения τ , определенные в этих работах не были неожиданными в том смысле, что они всегда были медленнее временной шкалы, ожидаемой исходя из частотного ограничения уха, т.е. $\tau > 1/\omega_{max}$. Более того, типы временных характеристик, изученных в этих работах – такие, как отрезки тишины или повторяемые пульсации – не имеют прямого отношения к звуковоспроизведению, так как такие искажения не могут естественным образом возникнуть в аудиотракте.

В настоящей работе сообщается о детектировании временной задержки короче, чем любая из ранее опубликованных, и которая имеет порог τ короче $1/\omega_{max}$. Более того, метод и тип временной особенности – расхождение между длинами пространственных путей двух динамических головок акустической системы – это искажение, которое реально может иметь место в реальном аудио-оборудовании, поскольку большинство акустических систем содержит несколько динамических головок и даже те, что не содержат несколько головок, будут иметь временное расплывание вследствие конечных размеров динамической головки. Такие расхождения времени прихода могут иметь даже большее значение в многоканальных системах окружающего звучания. Настоящая работа обеспечивает научную основу известным заявлениям аудиофилов о том, что для высокого качества воспроизведения требуется временной отклик в микросекундном диапазоне, и дает надежный количественный стандарт для оценивания влияния временных задержек и расплывания в цепочке компонентов аудиосистемы на ухудшение качества воспроизведения.

1. Исходные данные

Основная цель высококачественного звуковоспроизведения состоит в воспроизведении звука с точностью, достаточной для того, чтобы ошибки во всех областях (частотной, временной и т.д.) были ниже порогов детектируемости. Отвлекаясь от стереовоспроизведения и аспектов пространственной локализации, заметим, что моно-звуки могут быть восприняты как различные вследствие (а) различных частотных компонентов, (б) различных уровней компонентов, (в) различной длительности и временного нарастания компонентов, и (г) различных фаз отдельных компонентов. Первые два различия часто характеризуют общим термином «спектральные», хотя, строго говоря, следует говорить о различиях в амплитудном спектре и спектре интенсивностей, так как все изменения, включая те, что связаны с фазой и временем, могут быть описаны через изменения в комплексном Фурье-спектре. За исключением простых линейных систем, взаимосвязь частоты, амплитуды, времени и фазы не прямая. В аудиосистемах, например, разделительный фильтр может внести частотно-зависимую разность фаз без физической задержки одной полосы частот относительно другой, однако влияние неравных расстояний слушатель –

сабвуфер и слушатель – сателлитная акустическая система лучше всего описывается с помощью задержки между двумя полосами частот.

Аналогично, слуховой аппарат воспринимает различие в фазе и времени по-разному, и ошибки в фазе и времени имеют неэквивалентные последствия. В линейной цепи задержка Δt синусоидального сигнала связана с его фазовым сдвигом θ через соотношение $\theta = 2\pi f \Delta t$. В слуховом процессе звуковой сигнал раскладывается на отдельные частотные каналы набором чувствительных волосковых ячеек внутреннего уха, настроенных на различные характеристические частоты и определенным образом упорядоченные пространственно вдоль базилярной мембраны в улитке. Эта частотно-пространственно разделенная информация переносится волокнами слухового нерва к ядру улитки. Нервные импульсы вдоль волокон слухового нерва отслеживают фазу своих соответствующих акустических сигналов для частот вплоть до 4 кГц. Однако слуховой механизм в значительной степени не сохраняет фазовую информацию между компонентами разных частот, что приводит к знаменитому закону Ома (второму) [5,6], благодаря чему ухо не сильно чувствительно к фазовым сдвигам для сильно разнесенных частот (несмотря на большие различия в форме волнового процесса). Этот факт полезен для проектирования разделительных фильтров, где разность фаз НЧ и ВЧ выходов важна в той степени, в какой это влияет на амплитудный отклик [3]. Выше 4 кГц волокна слухового нерва откликаются приблизительно равномерно (одинаково) на длительность тона без синхронизации между запускающими образами и фазой акустического сигнала.

Хотя фазовая когерентность для различных частот не очень важна, временная когерентность – совершенно иное дело. Слуховая система очень чувствительна к синхронизации нарастающих фронтов различных частот: как хорошо известно, тембр инструментов становится неоднозначным, если удалить участки нарастания и затухания нот [7]. В ядре улитки быстро откликающиеся ячейки-осьминоги действуют как синхронные ключи «И» (каждая примерно с 60 входами) для совмещения (объединения) сигналов волокон слухового нерва от различных частотных каналов [8,9]. Эти ячейки четко откликаются на хорошо обозначенную во времени многочастотную «активность» на переднем фронте звукового импульса и поэтому их выходной сигнал есть мера скорости нарастания исходного звукового сигнала [10]. Если начальная временная неопределенность в сигнале волокон слухового нерва представлена в виде Гауссовой функции плотности вероятности $f(t) = f_0 e^{-(t/t_0)^2}$, значение $t_0 \sim 125 \mu s$ можно принять как грубую оценку начального временного расплывания, поскольку волокна слухового нерва теряют синхронизацию фазы с акустическим исходным сигналом вблизи 4 кГц и откликаются только на положительный полупериод [11,12]. Вероятность того, что N сигналов придут одновременно, чтобы возбудить осьминожную ячейку, пропорциональна произведению вероятностей. Это дает функцию вероятности на выходе $f'(t) \propto (f_0)^N e^{-(t(t_0/\sqrt{N}))^2}$ с уменьшенным временным расплыванием t_0/\sqrt{N} . Кроме начального фактора объединения $N \sim 60$ для каждой осьминожной ячейки, выходной импульс ячейки претерпевает дополнительные слияния (объединения) на более высоких нейронных уровнях (напр., в сферических ветвистых ячейках в пространственном лемнискусе (от *лат.* лента?)). Не ясно, насколько эти слияния увеличивают N и улучшают временное разрешение. Однако ясно, что максимальный фактор слияния не может превышать общее количество волосковых элементов внутреннего уха, которое равно примерно 4000. Таким образом, для импульсных исходных сигналов временная резкость (детальность, точность) слуховой системы может быть оценена как 2-16 мкс, если принять t_0/\sqrt{N} с $N = 60-4000$. Заметьте, что эта величина τ имеет очень мало общего с высокочастотной границей слышимости f_{\max} . Как было описано ранее, слуховой сигнал происходит от волосковых ячеек, расположенных так, что ближайшие ко входу улитки (базальный край) воспринимают самые высокие частоты. С возрастом высокочастотные волосковые ячейки постепенно гибнут и соответственно понижается f_{\max} . Так как τ зависит только от N и прямо не зависит от f_{\max} , значение τ не так чувствительно к возрасту. Напр., 50%-ное падение f_{\max} от 18 до 9 кГц соответствует потере 1 из 10 слышимых октав, т.е. уменьшению N от примерно 4000 до примерно 3600. Это ухудшит τ лишь на 5% (т.е. на корень квадратный из 4000/3600). Это может объяснять, почему пожилые слушатели, испытывающие трудности при различении консонансных звуков, тем не менее различают

предельно малые несовершенства (преимущественно во временной области) в высококачественных системах звуковоспроизведения.

Другая очевидная «асимметрия» между частотным и временным восприятием возникает при детектировании джиттера (дрожание импульсов по временной шкале). Было показано, что на слух можно заметить временной джиттер величиной всего 0,1 мкс для тонов, состоящих из пакетов высокочастотных импульсов [13]. Можно наивно подумать, что ухо может слышать частоты в мегагерцовом диапазоне. Однако пакет точно периодических импульсов содержит только гармоники фундаментальной частоты, и джиттер вносит дополнительную частоту в этот спектр, давая тем самым признак для различения [14].

Эффект от ограничения полосы частот в звуковоспроизводящей цепочке имеет еще одно следствие. Независимо от значения верхней частоты среза или степени рассогласования всегда существует некоторое ослабление частот в полосе слышимости. Повлияет ли это заметно на тембр или нет -- зависит от порога слуховой системы по различению изменения интенсивности (его называют минимальной заметной разностью). Этот порог чувствительности опять-таки вовсе не связан сильно с f_{max} .

Наконец, даже в случае стабильных тонов, частоты выше f_{max} могут повлиять на восприятие, если их уровень высок. Хотя наружное и внутреннее ухо стремятся отфильтровать ультразвуковые частоты, часть этой энергии тем не менее достигает внутреннего уха, особенно вследствие проводимости костей [15—19]. Эта ультразвуковая энергия может затем стимулировать каналы слышимых частот либо напрямую, либо через генерацию слышимых продуктов нелинейностей слухового процесса [20,21]. Такой ультразвук, даже если он не слышим сам по себе, может, как предполагается, изменить восприятие тембра при наложении на слышимые гармоники [22,23].

Таким образом, частоты выше f_{max} , хотя обычно и не относятся к слышимым, если представлены в форме чистых тонов, могут тем не менее влиять на тембр, а их удаление может понизить качество воспроизведения. Как следствие, минимальное время отклика аудиотракта, который ухудшает прозрачность, может быть намного меньше, чем $1/\omega_{max}$.

Было поставлено несколько экспериментов с целью определения минимальной временной особенности, которая может различима. Некоторые из этих экспериментов содержат попытку разделить роль нейрофизиологических механизмов, которые изоспектральны по амплитуде (полагаясь полностью на временную организацию и фазу) по отношению к спектральным механизмам (полагаясь на различия в амплитудном спектре). Если рассматриваются аудиоприменения, например, определение минимальной цифровой скорости дискретизации для достижения прозрачности, значение имеет минимальная временная ошибка, которая может быть замечена, не важно при этом, какой признак использует наше ухо, чтобы сообщить о различии. Многие из предыдущих экспериментов исследовали слышимость временных «щелей» в синусоидах [24-26] и шуме [27-30] (шум брался изоспектральный) и дали значения τ в интервале 0.2—10 мс. В другом эксперименте [31] «щель» Δt вводилась внутрь импульса и слушатели старались отличить один импульс шириной 20 мкс от пары импульсов по 10 мкс с задержкой $dt = \Delta t + 10$ мкс. Порог детектирования оказался равным $dt \sim 20$ мкс и было показано, что различение имеет спектральную природу (узкие импульсы имеют очень широкий спектр и разделение пары порождает заметные спектральные изменения в широком интервале частот). Изоспектральные варианты этих экспериментов [32,33], в которых пара неодинаковых импульсов сопоставлялась с их обращенной во времени копией, дали значения около 200 микросекунд. В недавнем эксперименте использовался [34] повторяющийся дрожащий шум. Исходный сигнал состоял из повторяющихся добавлений копий шумового сигнала, которые постепенно задерживались таким образом, чтобы получалось периодическое дрожание. Порог оказался более 12,5 микросекунд. Заметим, что временные искажения (отрезки тишины и последовательного дрожания) не относятся к тем, что естественно возникают в аудиотракте, и поэтому не имеют прямой связи со звуковоспроизведением. Более того, пороговые значения τ , которые наблюдались, довольно велики и превышают значение $1/\omega_{max}$.

Настоящий эксперимент исследует точно разновидность временных искажений, которая проявляется в типичном аудиотракте: искажения, вызванные пространственным рассогласованием двух динамических головок. Возникающее небольшое расширение формы сигнала также тесно связано с временным размытием вследствие конечного времени релаксации и конечной полосы пропускания аудиокомпонентов. Полученный порог заметности $\tau \approx 6 \mu s$ не просто короче всех значений, опубликованных в литературе, но впервые оказывается ниже величины $1/\omega_{max}$. В отличие от всех вышеупомянутых экспериментов, задержки в нашем эксперименте вводятся механически, что исключает появление побочных нелинейных продуктов и переходных процессов, которые могут генерироваться при смешивании, задерживании, временном обрезании сигналов электронными или электрическими средствами. Полученный результат также проливает свет на возможные физиологические механизмы и их

связь с принятыми значениями минимальной заметной разности уровня звука, опубликованными в литературе.

2. Методика

2.1. Аппаратура

Конфигурация эксперимента показана на Рис. 1. Две динамические головки расположены вертикально одна над другой с лицевыми сторонами, параллельными одна другой. Верхняя головка установлена на рельс, который может перемещаться вперед и назад между точкой жесткой фиксации (выровненное положение) и регулируемой фиксацией с помощью микрометрического винта (для смещенного положения) заданного значения смещение d . Слушатель усаживался на расстоянии $D = 4,3$ м лицом к громкоговорителям, уши на высоте посередине между двумя динамиками. Динамики центрированы по отношению к обоим ушам, так что оба уха принимают одинаковый сигнал.

Комната имеет форму прямоугольного параллелепипеда высотой 2,7 м, ширина 3,6 м, длина 5,8 м. Ось динамик—слушатель лежит вдоль длинного измерения и в центре относительно боковых стен, эта ось находится на высоте 1,1 м над полом. Пол и стены комнаты были покрыты акустическим ковром, потолок закрыт акустическими обоями. Эти материалы имеют коэффициент поглощения $S > 0,7$ на частотах, важных в рассматриваемом случае (>7 кГц). Кроме того, в определенных стратегических местах для подавления отражений были установлены плиты из стекловолкна толщиной 38 мм – 6 панелей общей площадью около 9 кв. м.

Громкоговорители – пара ленточных ВЧ-головок Aurum Cantus G2Si (Jinlang Audio Co. Ltd., Penglai City, P. R. of China) с частотным откликом 2–40 кГц, чувствительностью 96 дБ/Вт на расстоянии 1 м номинальным сопротивлением 6 Ом. Оба динамика подключены параллельно к одному источнику прямоугольных импульсов частотой 7 кГц. Этот источник состоял из аналогового генератора сигналов (модель 4001 производства Global Specialties Instruments, Cheshire, Connecticut) за которым был установлен широкополосный усилитель (с полосой частот по уровню 0–2,2 МГц). Рис. 2 показывает форму сигнала по напряжению на клеммах одного из динамиков, измеренных цифровым запоминающим 500 МГц-осциллографом LeCroy модель LT322, (LeCroy Corporation, Chestnut Ridge, New York), который оцифровывал сигнал с частотой выборок 200 MS/s (миллион выборок в сек) и вертикальным разрешением 12 бит. Этот же осциллограф использовался для других измерений спектра и формы сигналов. Обращаем внимание на хорошо управляемый отклик с ничтожным звоном и выбросами и временем нарастания-спада менее 2 мкс. Измеренный джиттер составил 68 нс ($< 0,05$ % периода).

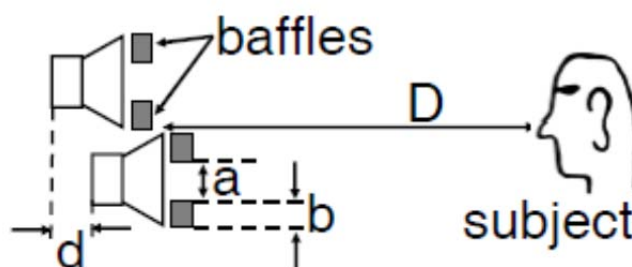


FIG. 1: Схема эксперимента. Расстояние динамик – слушатель $D=4.3$ м, апертурная длина $a=1,5$ см, расстояние между центрами динамиков $a + 2b = 9.9$ см. Смещение d изменяется. В слепых тестах слушатель пытается различить настроенный ($d=0$) и расстроенный варианты (d не равно 0) для интервала значений d 2–10 мм ($\tau \sim 6\text{--}30$ μs).

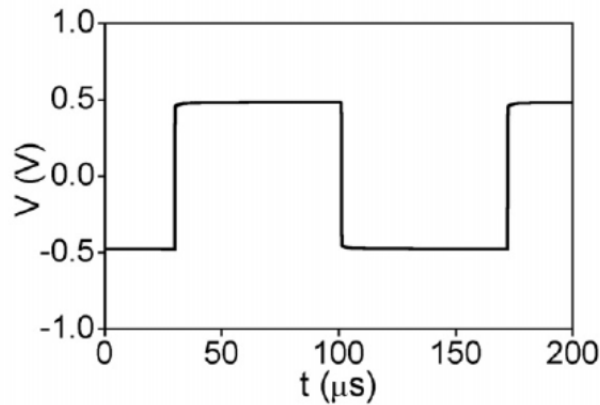


FIG. 2: Форма импульса напряжения на клеммах громкоговорителя, записанная с частотой выборки 200 MS/s (мегасэмпл в сек)

Прямоугольные импульсы частотой 7 кГц были выбраны потому, что они имеют только четные гармоники, которые (за исключением основной частоты) находятся ниже порогов слышимости для одиночных тонов. Однако выходной акустический сигнал динамика не будет совершенной репликой электрического входного импульса. Кроме того, может возникать изменение гармонических коэффициентов из-за непостоянства частотного отклика, могут появиться субгармоники и другие фальшивые ангармонические компоненты [35], когда громкоговоритель работает при высоких уровнях, особенно при неадекватном демпфировании. В настоящем эксперименте уровень накачки динамика был скромный (примерно $0,5 \text{ В}$ пиковое входное напряжение и 69 дБ SPL вместе расположения слушателя), а демпфирование -- эффективным (выходное сопротивление источника сигнала $< 40 \text{ мОм}$) для предотвращения негармонических искажений. Отсутствие негармонических искажений проверялось анализатором спектра для акустического выходного сигнала с использованием измерительного микрофона ACO Pacific (ACO Pacific, Inc., Belmont, California) model 7016 и предусилителя 4012 с усилением 40 дБ . Частотный отклик микрофона с усилителем был плоским по уровню $\pm 3 \text{ дБ}$ в полосе $4 \text{ Гц} - 120 \text{ кГц}$.

Спектр мощности акустического сигнала от динамика показан на рис. 3а. Часть б) показывает увеличенное изображение основного пика (f_1) с линейными осями, полная ширина на середине высоты FWHM равна 0.77 Гц (примерно 0.01% от f_1). Пики субгармоник от шума отличить невозможно. Абсолютный уровень звукового шума (включая окрестность субгармоники $3,5 \text{ кГц}$) был менее 0 дБ SPL (т.е. ниже пунктирной линии на рис. 3а). Выше f_1 также не было заметных ангармонических компонентов. Как ожидалось, в спектре преобладают нечетные гармоники ($7, 21, 35, 49 \text{ кГц}, \dots$), простирающиеся за границы слышимого диапазона. Поскольку сигнальная цепь полностью аналоговая, фальшивые частоты, которые могут появиться из-за ошибок дискретизации в цифровых цепях, отсутствовали.

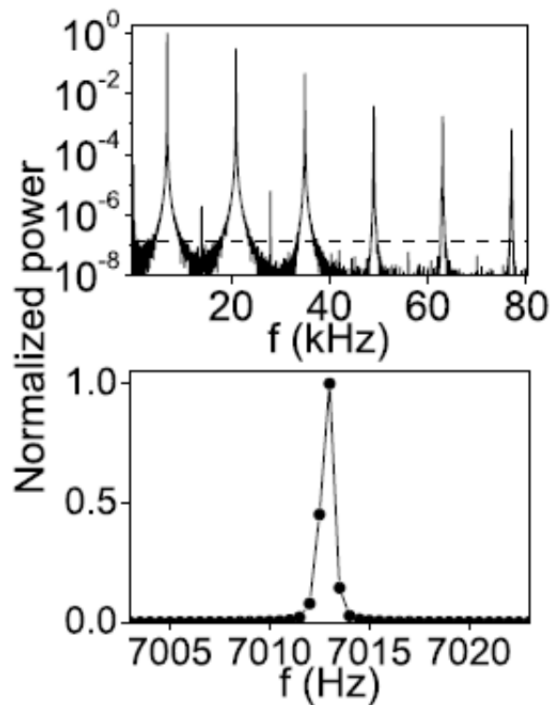


FIG. 3: Спектр мощности неусредненного акустического выходного сигнала одного громкоговорителя на расстоянии 0,7 м. Коэффициенты нормированы по отношению к основному максимуму. (a) Log-линейный график в окне 20 Hz–80 kHz с шагом 20 Hz при частоте выборок 2MS/s. Горизонтальная пунктирная линия соответствует абсолютному уровню 0 dB SPL. (b) Линейно-линейный график области вблизи главного максимума с шагом 0,5 Hz при частоте выборок 50 kS/s.

2.2. Акустические сигналы в месте расположения слушателя

Два сигнала достигают ушей слушателя с задержкой

$$\tau = d/c, \quad (1)$$

где c – скорость звука. Слушатель находится на расстоянии $D = 4,3$ м лицом к динамикам, его уши расположены на уровне посередине между двумя динамиками. Расположение динамиков и слушателя должно быть таким, чтобы линия, соединяющая среднюю точку между ушами слушателя и среднюю точку между двумя динамиками, была перпендикулярна плоскости, задаваемой передними поверхностями динамиков при нулевом смещении. Это проверялось лазерным лучом, затем слушатель тщательно подстраивал положение головы на слух. С этого момента слушатель поддерживал голову в фиксированном положении. Уровень звука при нулевом смещении составлял 69 dB SPL в месте нахождения слушателя.

Ур-ие (1) было бы точным и полным, если бы динамики были точечными источниками, а отражения в комнате отсутствовали. Однако звук, порождаемый протяженным источником, претерпевает временное расплывание δt (так как волны от разных точек излучающей поверхности достигают ушей в разные моменты времени), что уменьшает точность эксперимента. Звук генерируется в каждом динамике вертикальной алюминиевой лентой длиной 6 см. Как показано на рис. 1, применялись поглощающие экраны с целью ограничения апертуры (и эффективной длины ленты) до значения $a = 1,5$ см при $b = 4,2$ см, таким образом временное расплывание уменьшалось до

$$\delta t \simeq (a^2 + 2ab)/2cD \simeq 0.5 \mu s, \quad (2)$$

что весьма мало в сравнении с задержками, применявшимися в эксперименте. Расстояние по вертикали между центрами двух динамиков $a + 2b = 9,9$ см приводило к угловому разделению между этими источниками в месте расположения слушателя величиной $1,3^\circ$.

Отражения в комнате могут также ухудшать определение времени задержки. Конечно, были приняты меры по уменьшению отраженной энергии, но даже лучшая безэховая камера не имеет полного поглощения. Поэтому необходимо количественно оценить влияние отраженной энергии на эксперимент. Сигнал в положении слушателя будет состоять из прямого излучения от обоих динамиков плюс сумма отражений. Каждая траектория отраженного звука для одного динамика будет связана с соответствующей (подходящей) траекторией отраженного звука для другого динамика и поэтому все отражения можно рассматривать в виде таких подобранных пар. Обозначая каждую такой пару буквой n ($n = 0$ для пары прямых сигналов), можно записать, что относительное смещение динамиков d внесет разницу длин траекторий d_n между траекториями в паре. Из геометрического рассмотрения ясно, что $d_n < d$ всегда.

Действительно, для первых отражений от боковых стен $d_n \approx dD/\sqrt{D^2 + w^2}$ где w – это ширина комнаты поперек оси динамик—слушатель, для первых отражений от пола и потолочного покрытия

$$d_n \approx dD/\sqrt{D^2 + (2h + b + a/2)^2},$$

где h – это вертикальное расстояние от оси динамик—слушатель до пола или потолка. Эти d_n соответствуют задержкам $\tau_n = d_n/c \leq d/c$. Прямой сигнал и отражение от от задней стены (за слушателем) имеют $d_n = d$ и точно соответствуют первичной задержке величиной d/c . Отражения более высоких порядков имеют задержки, уменьшающиеся в прогрессии.

Пара отраженных сигналов кроме более короткой пошаговой задержки между сигналами двух динамиков может также иметь дополнительную начальную разность геометрических длин путей l_n между членами пары. Для путей, возникающих при одном отражении от пола или потолка

$$l_n \approx [(a + 2b)D/2h] \left[\sqrt{\frac{D^2 + 4h^2}{D^2}} - \sqrt{\frac{D^2}{D^2 + 4h^2}} \right] \quad (3)$$

а для путей с одним отражением от каждой из стен

$$l_n = 0. \quad (4)$$

Каждая пара путей-отражений имеет среднюю длину до слушателя D_n , которая растет с увеличением порядка отражения, а для прямых сигналов $D_0 = D$.

С такими определениями результирующий общий сигнал в точке расположения слушателя от обоих динамиков, просуммированный по всем отражениям, может быть записан как (для каждой гармоники)

$$A \cos(2\pi f[t + l'/c]) + A \cos(2\pi ft) = \sum_n A_n \cos(2\pi f[t + (D_n + l_n)/c]) + \sum_n A_n \cos(2\pi f[t + D_n]), \quad (5)$$

когда динамики выровнены, и записывается в виде

$$A \cos(2\pi f[t + \{l' + d'\}/c]) + A \cos(2\pi ft) = \sum_n A_n \cos(2\pi f[t + \{d_n + D_n + l_n\}/c]) + \sum_n A_n \cos(2\pi f[t + D_n]), \quad (6)$$

когда имеется смещение.² Первые члены с обеих сторон в обоих уравнениях соответствуют сигналу от верхнего динамика, а вторые члены – сигналу от нижнего динамика. l' – это эффективное начальное смещение пути, а d' ($< d$) – эффективное смещение, усредненное по всем отражениям.

Амплитуды A_n быстро спадают (относительно A_0 прямых сигналов) для путей, испытывающих многократные отражения. Даже отражения первого порядка сильно ослаблены по отношению к прямым

² Последние слагаемые в двух последних ф-лах должны иметь вид $[t + D_n/c]$ (прим. перев.).

звукам вследствие трех факторов: (1) поглощение отражающей поверхностью, (2) более длинный путь D_n и, следовательно, спадание по закону $1/(D_n)^2$, (3) узкая угловая характеристика ленточного динамика, излучающего преимущественно в прямом направлении. Это приводит к тому, что отношения (A_n^2/A_0^2) составляют 2% для отражений от пола, потолка и боковых стен, и примерно 20% для отражения от дальней стены. Хотя отражение от дальней стены может давать вклад в стоячие волны, оно не изменяет ни начального сдвига, ни эффективного смещения, т.е. для этого отражения $l_n = 0$, $d_n = d_0$.

Изменение пути l' уменьшает начальную, до смещения, интенсивность звука на коэффициент $\cos^2(\pi f l' / c)$. Смещение динамика ослабляет этот уровень на величину

$$\Delta L_p = 10 \log \left[\frac{\cos^2 \{ \pi f (d + l') / c \}}{\cos^2 \{ \pi f l' / c \}} \right]. \quad (7)$$

Если голова слушателя соответствующим образом центрирована, $l_0 = 0$ для прямых сигналов, а если голова смещена на δy (по оценкам < 3 см), то $l_0 = \delta y (a + 2b) / D$. Другие вклады в l' идут от отражений от пола и потолка. На основании ур-ий 2 и 5 их оценка составляет $l' < 1,2$ мм. Из ур-ия 7 теперь можно получить верхнюю оценку ослаблений. Эти оценки даны в таблице 1 для основной частоты.

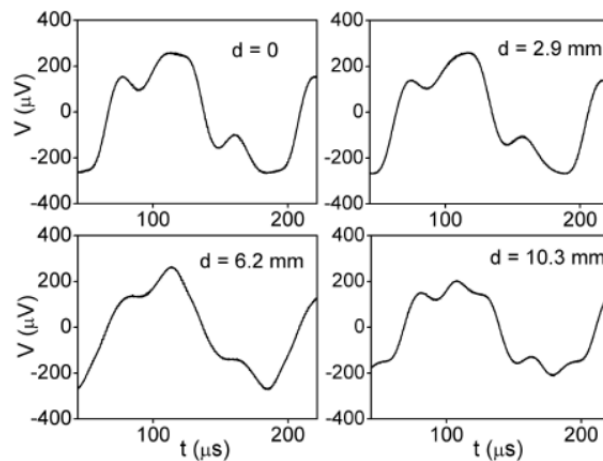


FIG. 4: Форма акустических сигналов, измеренных на месте слушателя для 4 вариантов расположения динамиков. График $d = 0$ соответствует контрольному сигналу (который равен удвоенному сигналу одного динамика). Регистрация с частотой выборки 20 MS/s.

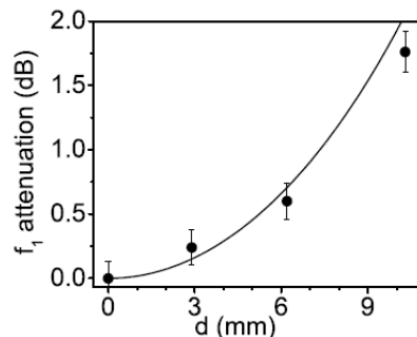


Fig. 5. Уменьшение мощности основной частоты 7 кГц как функция смещения динамика. Точки – измеренное уменьшение акустического сигнала в месте расположения слушателя. Сплошная линия – расчет по формуле 7.

Табл. 1. Теоретические верхние оценки ослабления сигнала на основной частоте 7 кГц для различного смещения

d (mm)	2.0	2.3	2.9	3.9	6.2	10.3
$-\Delta L_p$ (dB)	0.16	0.19	0.27	0.44	0.98	2.6

Предыдущее обсуждение и расчеты были основаны на геометрии эксперимента и элементарной теории сигналов. С другой стороны, можно просто измерить действительный конечный акустический сигнал в точке расположения слушателя с помощью микрофона и анализировать и оценивать изменения. На рис. 4 показаны формы акустических сигналов, измеренные в точке нахождения слушателя для четырех смещений. Кривая $d=0$ соответствует ур-ию 5, другие кривые соответствуют ур-ию 6. Формы сигналов (измеренные описанным ранее оборудованием: измерительным микрофоном АСО и предусилителем, за которым следовал осциллограф LeCroy) получены усреднением по 16000 записей с частотой выборок 20 мегасэмплов в секунду с вертикальным разрешением 12 бит.

Принимая во внимание 7 кГц-периодичность сигналов и отсутствие негармонических компонентов (что подтверждено усредненным спектром на рис. 3 и его предыдущим анализом), импульсы на рис. 4 можно представить в виде рядов Фурье и полностью описать коэффициентами C_n и фазами θ_n в разложении $V(t) = \sum C_n \cos(2\pi f_n t + \theta_n)$, где $f_n = n \times 7$ кГц. Эти коэффициенты даны в табл.2 (нормированные на коэффициент первой гармоники контрольного сигнала $C_1(0)$). Столбцы для гармоник, где все величины C находятся ниже уровня шума (0,005) были исключены (это имеет место для всех четных гармоник кроме $n = 2$). Фаза каждой гармоники задана относительно фазы основной гармоники для того же значения d (т.е. по отношению к $\theta_n(d)$), абсолютная фаза и разность фаз для различных значений d , конечно, не важны. Кроме шумовой погрешности, значения ненулевых коэффициентов изменяются незначительно в зависимости от положения микрофона из-за малых локальных изменений интенсивности, вызванных слабыми парциальными стоячими волнами. Эта чувствительность к положению приводит к неопределенности около $\pm 0,015$ для ненулевых коэффициентов, что эквивалентно ошибке в измеренных ослаблениях в диапазоне от $\pm 0,13$ до $\pm 0,16$ дБ.

Для идеального прямоугольного сигнала, гармонические коэффициенты имеют значения $C_n = 1/n$ для нечетных n и $C_n = 0$ для четных n . Из таблицы 2 видно, что четные C_n действительно ничтожно малы и что нечетные C_n следуют $1/n$ хорошо до $n = 3$. Выше присутствует высокочастотное затухание отклика динамика и коэффициенты ниже, чем теоретические значения $1/n$. Общий уровень звука при $d = 0$ равен 69 дБ SPL. Большая часть этой мощности (88%) сконцентрирована в 7 кГц-колебаниях основной частоты. Уровни всех гармоник кроме 7 кГц для всех d находятся ниже порогов их слышимости [35—37]. Последние два нижних столбца в табл.2 дают ослабление общей среднеквадратичной мощности $(C_1^2(d) + C_2^2(d) + C_3^2(d) + \dots)$ и основной гармоники (только слышимые Фурье-компоненты) относительно их значений для $d = 0$. Эти измеренные значения основной компоненты показаны на рис. 5 и, как видно, форма импульсов более-менее соответствует рассчитанным верхним оценкам в табл. 1.

В сравнении с выше описанными акустическими сигналами в воздухе, сигналы на барабанной перепонке будут обеднены высокими гармониками из-за фильтрации ушным каналом. Так, например, отношение третьей гармоники к первой в барабанной перепонке будет ниже чем измеренные отношения $C_3(d)/C_1(d)$ в Табл. 2. Однако, изменение каждой амплитуды Фурье в барабанной перепонке будет точно такое, как соответствующее измеренное значение $C_n(d)/C_n(0)$ в табл. 2. Таким образом, ослабление на основной частоте в барабанной перепонке будет точно таким, как измеренные значения, записанные в последнем нижнем столбце таблицы, а общее среднеквадратичное ослабление в барабанной перепонке будет чуть ниже, чем дано в предпоследнем нижнем столбце этой таблицы.

2.3. Процедура

В эксперименте испытуемые сидели перед двумя близко расположенными динамиками (рис.1). Для контрольной ситуации динамики выравнивались и находились на одинаковом расстоянии от ушей испытуемого, для измерений динамики смещались один относительно другого на расстояние d . Форма акустического сигнала возле ушей становилась все болл смазанной (Рис. 4 и ур-ие 6), гармоники все более ослаблялись (ур-ие 7) по мере увеличения d . Контрольный звук ($d = 0$) воспринимался как более резкий и яркий? Чем звук при смещении ($d \neq 0$), пока d не становилось таким малым, что разницы не было. Цель состояла в нахождении порогового значения d , которое едва различимо на слух. В слепых тестах испытуемый пытается определить, соответствует ли неизвестный звук контрольному расположению или смещению с различным значением d . Оказалось, что испытуемым обычно требуется слушать звук в течение нескольких секунд, чтобы сформировать устойчивое представление о тембре, быстрое переключение «туда—назад» затрудняет различение.

Временная организация испытаний выглядела следующим образом. Для каждого d испытуемый слушал контрольный и «смещенный» звуки несколько раз, чтобы почувствовать тембр каждого из них. Затем звуки проигрывались в последовательности: «смещенный», неизвестный. Контрольный. Длительность смещенного звука в начале последовательности ограничивалась 20 сек., а длительности неизвестного и контрольного звуков были ограничены 10 сек. Механический процесс перемещения динамика из смещенного положения в основное занимает около 1 сек и обеспечивает мягкое и непрерывное изменение от «смещенного» к «несмещенному» сигналу без переходных процессов или скачков в сигнале, которые иногда происходят при электрическом переключении. Испытуемый(ая) оценивает идентичность неизвестного сигнала, сравнивая этот сигнал с его (ее) недавней памятью о контрольном и смещенном сигналах, затем проводится следующее испытание для той же установки d . Для каждого испытания неизвестный звук выбирается либо смещенным, либо контрольным (в среднем с равной вероятностью) в зависимости от последовательности случайных чисел, сгенерированную компьютером. Один из примеров такой последовательности {0, 0, 1, 1, 0, 0, 0, 1, 0, 1}. Когда все 10 испытаний для одного испытуемого закончены, в зависимости от доступности испытуемого, либо проводится новая серия испытаний с тем же испытуемым для следующего более низкого значения d , либо серия проводится для другого испытуемого для следующего более низкого значения d , для которого он(она) еще не проверялись. Всего было сделано 50 слепых испытаний на пяти испытуемых для каждой установки d . Тесты проводились для 6 значений d (2,0, 2,3, 2,9, 3,9, 6,2, 10,3 мм) в сравнении с контрольной величиной $d = 0$. В целом эксперимент состоял из 300 слепых тестов.

Табл. 2. Состав гармоник акустических сигналов. Коэффициенты $C_n(d)$ выражены в долях относительно $C_n(d = 0)$. Фазы θ_n в радианах выражены относительно θ_1 для такого же значения d . Две последних столбца внизу дают ослабление мощности в дБ общего среднеквадратичного уровня и первой гармоники ($C_1(d)$ относительно контрольных значений без смещения ($d = 0$)). Уровень шума и ошибка для коэффициентов равны 0.005 и 0.015 соответственно.

d (mm)	$f_1=7$ kHz		$f_2=14$ kHz		$f_3=21$ kHz		$f_5=35$ kHz	
	C_1	θ_1	C_2	θ_2	C_3	θ_3	C_5	θ_5
0	1.000	0.00	0.006	-3.81	0.332	-1.29	0.145	-0.84
2.9	0.973	0.00	0.006	-4.09	0.329	-1.54	0.104	-1.66
6.2	0.933	0.00	0.005	-4.49	0.217	-1.50	0.038	-5.35
10.3	0.816	0.00	0.004	-5.48	0.156	-0.67	0.131	-0.11
d (mm)	$f_7=49$ kHz		$f_9=63$ kHz		$f_{11}=77$ kHz		Attenuation	
	C_7	θ_7	C_9	θ_9	C_{11}	θ_{11}	rms	C_1
0	0.018	-1.94	0.007	-3.02	0.002	-3.79	0	0
2.9	0.012	-3.09	0.003	-4.50	0.001	-0.10	0.26	0.24
6.2	0.013	-0.90	0.008	-2.62	0.002	-4.13	0.90	0.60
10.3	0.007	-1.98	0.005	-1.16	0.002	-2.98	2.03	1.76

2.4. Слушатели

В исследовании участвовали 5 слушателей в возрасте 24—40 лет. У них не было повреждений слуха или неврологических болезней. Все участники были добровольцами и не получали оплаты. Совет университета Южной Каролины рассмотрел и согласовал предложение об этом исследовании и необходимую форму согласия участников.

3. РЕЗУЛЬТАТЫ

Результаты эксперимента показаны в таблице 3. Все участники на 100% в слепых тестах показали чувствительность вплоть до $d = 2,9$ мм. Показатели 50 из 50 соответствуют значениям хи-квадрат $\chi^2 = 50$, намного выше критического значения 3,84. (Количественная мера различимости в психофизических тестах – анализ величины хи-квадрат, определяемой как $\chi^2 = (C - T/2)^2 / (T/2) + (I - T/2)^2 / (T/2)$, где T – общее число испытаний, C – количество верных суждений, I – количество неверных суждений. Для испытаний с одной степенью свободы, каковым является настоящий эксперимент, в котором оценивают выровнен динамик или нет, критическое значение хи-квадрат составляет 3,84. Таким образом, например, счет 8 из 10 в тесте будет рассматриваться как статистически незначимый даже хотя доля правильных ответов равна 80%, так как в этом случае хи-квадрат равно 1,8, что меньше 3,84.) Самое короткое смещение, которое было четко зафиксировано составляло 2,3 мм, что соответствует задержке менее 6,7 мкс. Для этого случая, учитывая всех участников, было набрано 82% верных ответов, хи-квадрат = 20,48, а индекс различимости из теории сигналов составил $d' = 1,84$ с критерием $s = 0,97$. Для смещения $d = 2,0$ мм не получено надежного различения с контрольным сигналом (52% верных ответов, хи-квадрат = 0,08, $d' = 0,14$, $s = 0,18$).

Уравнение (1) связывает смещения с первичными задержками между двумя сигналами, а рис. 6 дает графическое представление всех результатов слепых тестов с учетом этих задержек. Часть (b) показывает на основе хи-квадрат-анализа, что граница между различимыми и неразличимыми задержками лежит вблизи 6 мкс.

Табл. 3. Результаты слепых тестов. Каждый ряд соответствует отдельному участнику, расположенному в порядке увеличения возраста. Цифры соответствуют числу правильных ответов (из 10) каждого участника для указанного смещения d в данном столбце.

10.3 mm	6.2 mm	3.9 mm	2.9 mm	2.3 mm	2.0 mm
10	10	10	10	9	10
10	10	10	10	8	4
10	10	10	10	10	4
10	10	10	10	9	4
10	10	10	10	5	4

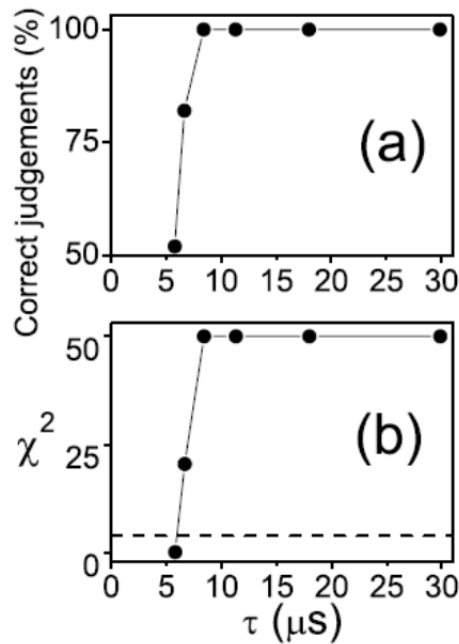


Рис. 6. Сводка результатов как функция времени задержки (усреднение по всем испытуемым). Каждая точка состоит из 50 слепых тестов. (а) Доля правильных суждений. (б) Значения хи-квадрат.ю Пунктир, соответствующий критическому значению 3,84, пересекает кривую данных около 6 мкс.

4. ЗАКЛЮЧЕНИЕ

Тракт звуковоспроизведения содержит много ступеней, которые могут вносить ошибки по времени, которые могут ухудшить качество звука:

- ограничение по ширине полосы частот (даже при абсолютно линейном поведении) приводит к конечному времени релаксации $\tau \sim 1/\omega_{max}$;
- применение цифровых носителей ограничивает самый короткий разрешаемый временной интервал величиной, равной примерно половине интервала выборки (который для компакт-дисков будет 11 мкс);
- пространственные размеры динамических головок акустических систем (или расстояний между головками в системах с несколькими головками) вносят временное смазывание и задержки.

В последнем случае уравнение 2 показывает, что дипольный громкоговоритель с единственной электростатической панелью высотой 1,5 метра при нахождении слушателя на расстоянии $D = 5$ м (ухо на середине высоты громкоговорителя) будет иметь временную размытость $a^2/2cD = 0,65$ мс (миллисек!). Это означает, что даже если весь остальной тракт будет иметь неограниченную полосу, входной сигнал в форме дельта-функции (короткий импульс) расплывется во временном окне протяженностью 650 мкс в месте расположения слушателя. Поэтому громкоговоритель, имеющий большой угловой размер при наблюдении с места слушателя с необходимостью будет иметь ухудшенное качество, что, возможно, объясняет, почему малые громкоговорители имеют тенденцию к субъективно более чистому и более когерентному звуку, хотя им может не хватать низкочастотного отклика.

Настоящая работа дает наилучшее количественное оценивание (6 мкс) того, насколько такие временные ошибки могут иметь слышимое различие. Множество предыдущих экспериментов (изложенных

в вводной части), посвященных этому вопросу, использовали оборудование, временной отклик которого мог оказаться главным ограничением. Большинство исследований использовали довольно грубый цифровой синтез для источника сигнала, использовали усиление с недостаточной собственной скоростью отклика и электроакустические преобразователи с ограниченной полосой, работающие с неадекватным демпфированием. В настоящей работе использовалась аналоговый тракт, в котором прямоугольный сигнал, направленный в преобразователь (включая отклик генератора сигналов и усилителя), имел времена нарастания/спада примерно в 100 раз быстрее, чем цифровой синтез с частотой выборок 48 кГц. Очень низкое выходное сопротивление 40 мОм обеспечивает исключительное демпфирование, и, следовательно, хорошо контролируемую форму, как показано на рис. 2. Преобразователи в данной работе имеют намного более широкую полосу (спектр показан на рис. 3) в сравнении с типичными преобразователями, используемыми в аудиометрии (напр., наушники TDH-39). Таким образом, за счет исключения некоторых узких мест оборудования стало возможным продемонстрировать более короткие пороги для различения временных ошибок, чем полученные ранее (как отмечалось во Введении, временная разница, определенная в работе [31], соответствует задержке между импульсами около 20 мкс, что намного больше, чем порог времени, полученный в настоящей работе). Этот новый более низкий порог должен учитываться в конструировании аудиокомпонентов для достижения высочайшей прозрачности.

Хотя настоящая демонстрация различимости на микросекундной шкале использовала простые (прямоугольные) широкополосные сигналы, реальные музыкальные сигналы также содержат компоненты в этом спектрально-временном интервале. Измерения спектров различных музыкальных инструментов показывают, что спектры простираются в ультразвуковую область [38] и даже за пределы 100 кГц [39]. Было продемонстрировано, что некоторые инструменты (ксилофон, труба, малый барабан, цимбалы) имеют очень крутой передний фронт, так что их полный уровень, превышающий 120 дБ SPL, достигается быстрее, чем за 10 мкс [2,38]. Кроме ультразвукового спектрального содержания и микросекундных длительностей фронтов, третьим аспектом музыкальных звуков, требующим быстрого временного разрешения, является реверберация. Импульсный звук производит каскад отражений, их частота прихода к слушателю растет пропорционально квадрату времени. Темп прихода этих отражений $dN/dt \approx 4\pi c^3 t^2 / V$, где V – объем комнаты, достигает одного за 5 мкс через 1 сек для комнаты объемом 2500 куб.м [2]. Поэтому точность воспроизведения в микросекундном диапазоне необходима для сохранения исходной акустической реверберации. Настоящий экспериментальный результат, таким образом, дает реальную основу известным заявлениям аудиофилов о чувствительности к очень коротким ошибкам во времени (таким как недостаточность традиционно используемым частотам дискретизации в цифровом аудио), как указывалось во введении.

Хотя нейрофизиологическая основа наблюдающейся высокой временной разрешающей способности не является главным аспектом для звуковоспроизведения, настоящий результат тем не менее проливает свет на этот вопрос. В начальном моменте любого слухового ощущения участвует возбуждение внутренних волосяных рецепторов улитки. Три фактора могут изменить восприятие звука. Первый – это изменение частоты (частот) источника, которое изменяет характеристические частоты и места вдоль базилярной мембраны, в которых находятся максимально возбужденные внутренние волосяные рецепторы. Второй фактор – это изменение громкости, которое изменяет степень возбуждения и ширину возбуждаемых внутренних волосяных рецепторов (кривая настройки каждого внутреннего волосяного рецептора и соответствующего нервного волокна имеет ширину около трети октавы). Эти два «спектральных» фактора изменяют тонопическую картину возбуждения. Третий фактор, влияющий на восприятие звука – это изменение временного порядка того, как возбуждаются различные внутренние волосяные рецепторы и нервные волокна. В этом случае усредненные по времени тонопические «образы» (паттерны, картины) возбуждения не будут изменяться и, по-видимому, некоторый процесс сравнения времени или измерения времени имеет место на нейронной стадии за пределами улитки. В настоящем эксперименте суммарный составной сигнал от двух динамиков содержит идентичные частоты от каждого сигнала, так как при суммировании не привлекаются никакие электронные устройства или преобразователи, генерирующие нелинейный или негармонические побочные продукты. Как велики различия уровня сигнала и какова их возможная роль в процессе различения? Частоты (Таблица 3), представленные в порядке

резкого убывания интенсивности – 7, 14, 21, 35 кГц (высокие гармоники имеют интенсивности <1% относительно фундаментальной компоненты). Как обсуждалось в разделе 2.2, только компонента с частотой 7 кГц имеет интенсивность выше порога слышимости. Изменения уровня всех компонентов (индивидуально или коллективно) находятся ниже порогов заметности на слух. Для минимального заметного смещения $d = 2,3$ мм, имеем $\Delta L_p \approx -0.2$ дБ, (уменьшение интенсивности 5%) для среднеквадратичного уровня и уровня основной 7-кГц-компоненты (табл. 2 и 3, а также рис. 5). Пороги различения на слух уровней (для частоты, равной или превышающей 7 кГц и $L_p = 69$ дБ) известны из данных Jesteadt и др. [40] и составляют 0,7 дБ (15%-ное уменьшение интенсивности). Даже предел трехкратной стандартной ошибки для этого параметра равен 0,5 дБ (11%-ное уменьшение интенсивности). Поэтому изменения уровней в эксперименте (<0,2 дБ) оказываются ниже порога детектируемости и различение должно включать нечто большее, чем простые спектрально-амплитудные признаки сигнала.

Анализируя восприятие фазы, стоит сопоставить настоящий результат с результатом Plomp и Steeneken [41], которые исследовали различимость двух сложных низкочастотных звуков, которые имеют одинаковый амплитудный спектр и различаются только фазой. В их эксперименте один сигнал состоял только из синусоидальных (или только из косинусоидальных) членов в ряде Фурье, а другой имел чередующиеся синусы и косинусы (т.е. фазовый сдвиг 90 градусов между соседними гармониками). Для фундаментальных частот 292 и 584 Гц эти авторы пришли к заключению, что манипуляции с фазой не просто слышимы, но что различимость на слух была эквивалентна изменению уровней на 2 и 0,7 дБ/октаву, соответственно. Хотя эти результаты нельзя просто экстраполировать на 7 кГц, в настоящем эксперименте фазовый сдвиг в 15% между фундаментальной частотой и следующей явно выраженной (третьей) гармоникой, вызванный смещением динамика, явно эквивалентен разности уровней 0,7 дБ, так как это соответствует пределу заметности. Надо отметить, что в эксперименте [41] характеристические разности времен между сигналами были велики (более 100 мкс) вследствие их низких частот.

5. ДОБАВЛЕНИЕ ПОСЛЕ ВЕРСТКИ СТАТЬИ

В близком по смыслу эксперименте [42], сигналы размывались во времени путем пропускания через НЧ-фильтр (чтобы имитировать эффект ограничения полосы частот в аудио компонентах) вместо пространственного смещения динамиков. В этом эксперименте сигнал подавался испытуемым не с помощью громкоговорителей, а через наушники с частотным диапазоном шире звукового. Пороговая константа времени, которая была различима, имела величину 5 мкс, что сопоставимо с настоящей работой.

- [1] H. R. E. van Maanen, “Temporal decay: a useful tool for the characterization of resolution of audio systems?”, AES Preprint 3480 (C1-9), presented at the 94th convention of the Audio Engineering Society in Berlin (1993).
- [2] W. Woszczyk, “Physical and Perceptual Considerations for High-Resolution Audio”, Audio Engineering Society Convention Paper 5931 Presented at the 115th Convention 2003 October 10-13 New York, New York (2003).
- [3] N. Thiele, “Phase considerations in Loudspeaker Systems”, Audio Engineering Society Convention Paper 5307 Presented at the 110th Convention 2001 May 12-15 Amsterdam, The Netherlands (2001).
- [4] J. R. Stuart, “Coding for high-resolution audio systems”, J. Audio Eng. Soc., **52**, 117–144 (2004).
- [5] M. R. Schroeder, “Models of hearing”, Proc. of the IEEE, **63**, 1332 (1975).
- [6] W. M. Hartmann, “Signals, sound, and sensation (Modern Acoustics and Signal Processing)”, AIP Press (1996).
- [7] K. W. Berger, “Some factors in the recognition of timbre”, J. Acoust. Soc. Am. **36**, 1988 (1963).
- [8] D. Oertel, R. Bal, S. M. Gardner, P. H. Smith, and P.X.Joris, “Detection of synchrony in the activity of auditory nerve fibers by octopus cells of the mammalian cochlear nucleus”, Proc. Nat. Acad. Sci. **97**, 11773–11779 (2000).

- [9] N. L. Golding, D. Robertson, D. Oertel, “Recordings from slices indicate that octopus cells of the cochlear nucleus detect coincident firing of auditory nerve fibers with temporal precision”, *J. Neurosci.* **15**, 3138–3153 (1995).
- [10] M. J. Ferragamo and D. Oertel, “Shaping of synaptic responses and action potentials in octopus cells”, *Assoc. Res. Otolaryngol.* **21**, 96 (1998).
- [11] D. H. Johnson, “The response of single auditory-nerve fibers in the cat to single tones: synchrony and average discharge rate”, Ph.D. thesis, Department of Electrical Engineering, MIT, Cambridge, MA (1974).
- [12] S. A. Shamma, N. Shen, G. Preetham, “Stereoausis: Binaural processing without neural delays”, *J. Acoust. Soc. Am.* **86**, 989–1006 (1989).
- [13] I. Pollack, “Submicrosecond auditory jitter discrimination thresholds”, *J. Acoust. Soc. Am.* **45**, 1059–1059 (1969).
- [14] I. Pollack, “Spectral basis of auditory jitter discrimination”, *J. Acoust. Soc. Am.* **50**, 555 (1971).
- [15] B. H. Deatherage, L. A. Jeffress, and H. C. Blodgett, “A note on the audibility of intense ultrasound”, *J. Acoust. Soc. Am.* **26**, 582 (1954).
- [16] F. J. Corso, “Bone conduction thresholds for sonic and ultrasonic frequencies”, *J. Acoust. Soc. Am.* **35**, 1738–1743 (1963).
- [17] M. L. Lenhardt, R. Skellett, P. Wang, and A. M. Clarke, “Human ultrasonic speech perception”, *Science* **253**, 82–85 (1991).
- [18] M. L. Lenhardt, “Human ultrasonic hearing”, *Hearing Rev.* **5**, 50–52 (1998).
- [19] S. Fujioka et al., “Bone Conduction Hearing for Ultrasound”, *Trans. Tech. Com. Physio. Acoust. Soc. Japan*, H-97-4 (1997).
- [20] H. E. von Gierke, “Subharmonics generated in human and animal ears by intense sound”, *J. Acoust. Soc. Am.* **22**, 675 (1950).
- [21] K. Ashihara, K. Kurukata, T. Mizunami, and K. Matsushita, “Hearing threshold for pure tones above 20 kHz”, *Acoust. Sci. & Tech.* **27**, 12–19 (2006).
- [22] T. Oohashi, E. Nishina, N. Kawai, Y. Fuwamoto, and H. Imai, “High-frequency sound above the audible range affects brain electric activity and sound perception”, *J. Audio Eng. Soc. (Abstracts)* **39**, 1010 (1991).
- [23] S. Yoshikawa, S. Noge, M. Ohsu, S. Toyama, H. Yanagawa, T. Yamamoto, “Sound-quality evaluation of 96-kHz sampling digital audio”, *J. Audio Eng. Soc. (Abstracts)* **43**, 1095 (1995).
- [24] M. J. Shailer and B. C. J. Moore, “Gap Detection and the Auditory Filter: Phase Effects Using Sinusoidal Stimuli”, *J. Acoust. Soc. Am.* **81**, 1110–1117 (1987).
- [25] C. Formby, M. Gerber, L. Sherlock, and L. Magder, “Evidence for an across-frequency, between-channel process in asymptotic monaural temporal gap detection”, *J. Acoust. Soc. Am.* **103**, 3554–3560 (1998).
- [26] B. C. J. Moore, “An Introduction to the Psychology of Hearing”, 5th edition, Academic Press (2003).
- [27] R. Plomp “Rate of decay of auditory sensation”, *J. Acoust. Soc. Am.* **36**, 277–282 (1964).
- [28] M. J. Penner “Detection of temporal gaps in noise as a measure of the decay of auditory sensation”, *J. Acoust. Soc. Am.* **61**, 552–557 (1977).
- [29] D. A. Eddins, J. W. Hall, and J. H. Grose, “Detection of temporal gaps as a function of frequency region and absolute bandwidth”, *J. Acoust. Soc. Am.* **91**, 1069–1077 (1992).
- [30] D. P. Allen, T. M. Virag, and J. R. Ison, “Humans detect gaps in broadband noise according to effective gap duration without additional cues from abrupt envelope changes”, *J. Acoust. Soc. Am.* **112**, 2967–2974 (2002).
- [31] B. Leshowitz, “Measurement of the two-click threshold”, *J. Acoust. Soc. Am.* **49**, 462–466 (1971).
- [32] D. Ronken, “Monaural detection of a phase difference between clicks”, *J. Acoust. Soc. Am.* **47**, 1091–1099 (1970).
- [33] G. B. Henning, and H. Gaskell, “Monaural phase sensitivity with Ronken’s paradigm”, *J. Acoust. Soc. Am.* **70**, 1669–1673 (1981).

- [34] K. Krumbholz, R. D. Patterson, A. Bobbe, and H. Falstl, "Microsecond temporal resolution in monaural hearing without spectral cues?", *J. Acoust. Soc. Am.* **113**, 2790–2800 (2003).
- [35] K. Ashihara, and S. Kiryu, "Influence of expanded frequency band of signals on non-linear characteristics of loudspeakers", *Nippon Onkyo Gakkai Shi (J. Acoust. Soc. Jap.)* **56**, 549–555 (2000).
- [36] International Standards Organization minimum audible field (MAF) standard: ISO 389-7 (1996).
- [37] K. Kurukata, T. Mizunami, K. Matsushita, and K. Ashihara, "Statistical distribution of normal hearing thresholds under free-field listening conditions", *Acoust. Sci. & Tech.* **26**, 440–446 (2005).
- [38] P. Rogowski, A. Rakowski, and A. Jaroszewski, "Specific Hearing Loss in Young Percussion and Brass Wind Players due to Music Noise Exposures", *The 8th International Congress on Sound and Vibration*, Hong Kong, China, 2–6 July (2001).
- [39] J. Boyk, "There's life above 20 kilohertz! A survey of musical instrument spectra to 102.4kHz", <http://www.cco.caltech.edu/boyk/spectra/spectra.htm>. Copyright c_ 1992, 1997 James Boyk, Music Lab, California Institute of Technology.
- [40] W. Jesteadt, C. C. Wier, and D. M. Green, "Intensity discrimination as a function of frequency and sensation level", *J. Acoust. Soc. Am.* **61**, 169–177 (1977).
- [41] R. Plomp and H. J. M. Steeneken, "The effect of phase on the timbre of complex tones", *J. Acoust. Soc. Am.* **46**, 409–421 (1969).
- [42] M. N. Kunchur, "Temporal resolution of hearing probed by bandwidth restriction", *Acta Acustica united with Acustica* **94**, 594-603 (2008). (Preprint can be downloaded from <http://www.physics.sc.edu/kunchur/temporal.pdf>)