



Ананьев
Анатолий Борисович,
кандидат технических наук,
доцент кафедры звукорежиссуры КНУКиИ,
выпускник кафедры акустики КПИ 1966 г.

Анатолий Ананьев
Акустика Музыка и Речи для Звукорежиссеров

Анатолий Ананьев



Акустика

Музыка и Речи для звукорежиссеров

Киев 2014

Начало создания этой книги пришлось на тревожное осеннее время 2013 года, когда произошли события, повергшие в шок киевлян, затем всю страну, а еще чуть позже - весь мир. Лишившись покоя, с несколькими почти круглосуточно одновременно работающими телеканалами на своих мониторах, я старался погасить нервное напряжение, насколько возможно углубившись в написание книги. Работа не избавляла от тревоги и горечи, а только направляла силы в те немногие сферы, где от меня могла быть польза.

Мы многое увидели и узнали за это тяжелое время. Возможно мы стали мудрее и сильнее. Мы с надеждой смотрим в будущее страны, в которой живет такой народ.

Пусть каждый, кто откроет эту книгу, вспомнит добрым словом всех тех, кто заплатил очень дорогую цену за то, чтобы мы могли смотреть в будущее страны с надеждой.

НАРОДОВІ УКРАЇНИ ПРИСВЯЧУЮ
НАРОДУ УКРАИНЫ ПОСВЯЩАЮ
DEDICATED TO THE PEOPLE OF UKRAINE

22 февраля 2014

Содержание

Введение	7
1 Область слухового восприятия	8
Речь и вокал	8
Музыкальные инструменты	10
Области слухового восприятия	13
Музыка в областях	16
Несколько слов об уровнях	21
2 Спектры звуковых сигналов	24
Кто такой спектр	24
Интерлюдия	26
Фазирование колонок	26
Фазовая характеристика эквалайзера	27
Изъятие вокала из песни	27
Как спектр работает	32
Критические полосы спектра	33
Несинусоидальные представления	36
3 Дискретные сигналы и их спектры.	38
Дискретизация сигналов	38
Еще один вид децибелов	41
Особенности спектров дискретных сигналов	42
Элайзинг	44
Ресемплинг	46
Теорема Котельникова	47
Оверсемплинг	47
4 Спектрограммы	49
5 Музыкальный строй. Темперация.	53
Интервалы, порожденные обертонами	53
Квинтовый круг	55
Равномерные темперации	56
К вопросу о точности интонирования	59
6 Музыкальные звуки	60
Сигналы и шумы	60
Структура музыкальных звуков	61
Шкала частот в изображении музыкального звука	63
Особенности изображения музыкальных звуков	64
Вибрато музыкального звука	65

7 Музыкальные инструменты	70
8 Певческий голос	77
Основные вокальные характеристики певческого голоса	77
Вибрато	77
Форманты	80
Примеры спектральных характеристик голоса певцов	83
9 Режиссура речевой деятельности	86
Речевые форманты	86
Звонкие речевые согласные	89
Интонационное поведение в речевой деятельности	90
Логические ударения	91
Интонационные шаблоны в речи	92
Паразитное заполнение пауз	94
Послесловие	97

**Київський Національний Університет Культури і Мистецтв
Інститут Кіно і Телебачення
Кафедра звукорежисури**

Посібник орієнтований на студентів та спеціалістів, що навчаються або підвищують свій фаховий рівень зі звукорежисури. Може бути корисний всім, хто бажає набути базових знань з акустики музичних та мовних сигналів.

Ананьев А.Б.

A64 Акустика Музыки и Речи для звукорежиссера. Учебное пособие.
К.: 2014. -

Пособие ориентировано на студентов и специалистов, которые обучаются или повышают свой профессиональный уровень в звукорежиссуре. Может быть полезно всем, кто желает приобрести базовые знания по акустике музыкальных и речевых сигналов.

20.05.14 - правка на стр.54

© А.Б. Ананьев, 2014

© А.Б. Ананьев, иллюстрации, оригинал-макет, 2014

Введение

Эта книга является дополнением и продолжением ранее изданной "Акустики для звукорежиссеров". Материал, приведенный здесь, частично повторяет несколько глав из моих более ранних книг "Акустика для звукорежиссеров" и "Элементы музыкальной акустики". Но здесь содержание этих глав существенно переработано в соответствии с целями именно этой книги и, в частности, адаптировано к звукорежиссерской аудитории, как правило далекой от математических методов обработки сигналов.

Стиль изложения здесь рассчитан на читателя, который достаточно хорошо усвоил содержание "Акустики для звукорежиссеров" и не нуждается в повторении и разъяснении сведений, которые изложены в упомянутой книге. Так, например, здесь я не разъясняю смысл и характер кривых равной громкости, не упоминаю основополагающей роли синусоидальных колебаний в аудиоинженерии, не касаюсь базовых психофизиологических свойств слуха. Некоторые разъяснения возникают лишь в связи с основной задачей этой книги - разобраться (насколько это возможно) в специфике музыкальных и речевых сигналов с точки зрения применения их в звукорежиссерской деятельности.

В необходимых местах я ссылаюсь на упомянутую книгу в виде, например [Ак, 37], что означает "Акустика для звукорежиссеров" стр 37.

Независимо от роли, которую в свое время сыграли мои студенты в нашей исследовательской деятельности, я с удовольствием приношу дань благодарности тем из них, кого вспоминал при подготовке этой книги:

Анна Путилова

Елена Чепурная

Марина Кирилук

Ольга Тарасова

Михаил Авраменко

Ирина Джола

Антон Кочетков

Юлия Бакута

Екатерина Твердохлеб

и, как всегда, Екатерина Ананьева.

Александра Нечай - вновь бесценный редактор моих трудов.

Тех, кто будет пользоваться этой книгой, прошу уважать мои авторские права. Разрешаю безвозмездно копировать книгу только для личного пользования, и только целиком, без изъятий. В случае цитирования отдельных фрагментов книги прошу указывать источник и автора.

Огромная благодарность неизвестному мне автору замечательного фото из интернета на обложке.

1 Область слухового восприятия

Речь и вокал

1. В литературе широко распространены иллюстративные изображения области слухового восприятия человека [Ак, 83]. Мы также начинаем изложение материала с этой темы. И первой своей задачей поставим разделить общий частотный диапазон воспринимаемых звуков на характерные поддиапазоны в соответствии с ролью, которую они играют в звуковых сигналах, а также со слуховым впечатлением, которое производят части различных звуковых сигналов, расположенные в этих поддиапазонах. Прежде всего обратимся к звукам человеческого голоса - исторически основного подвластного человеку источнику сигналов.

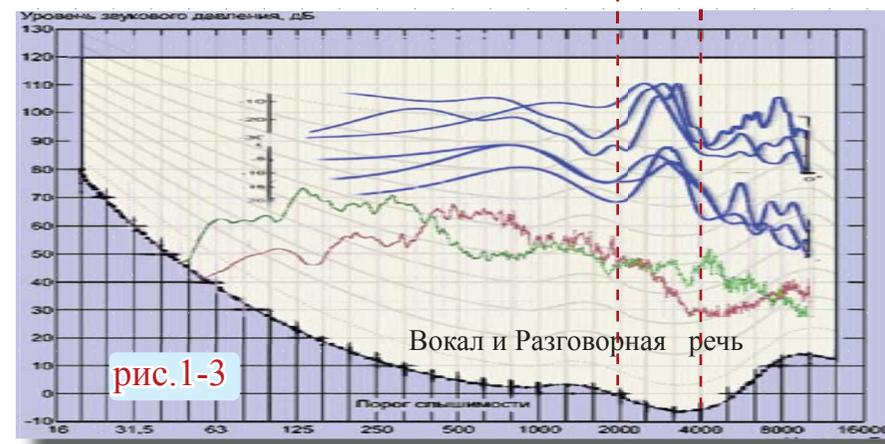
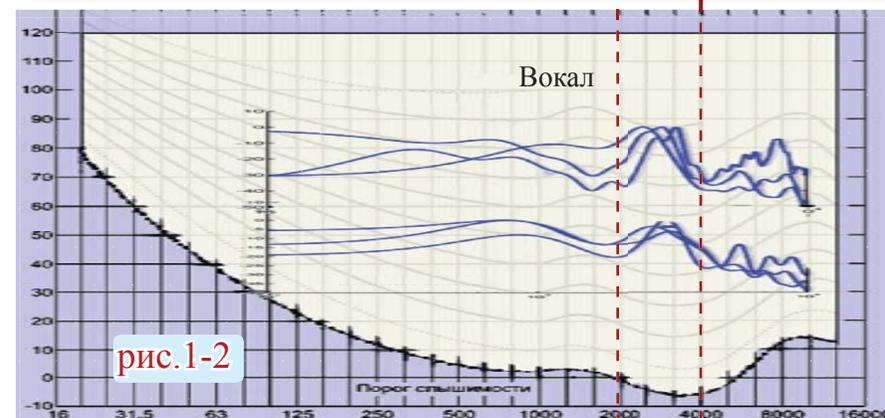
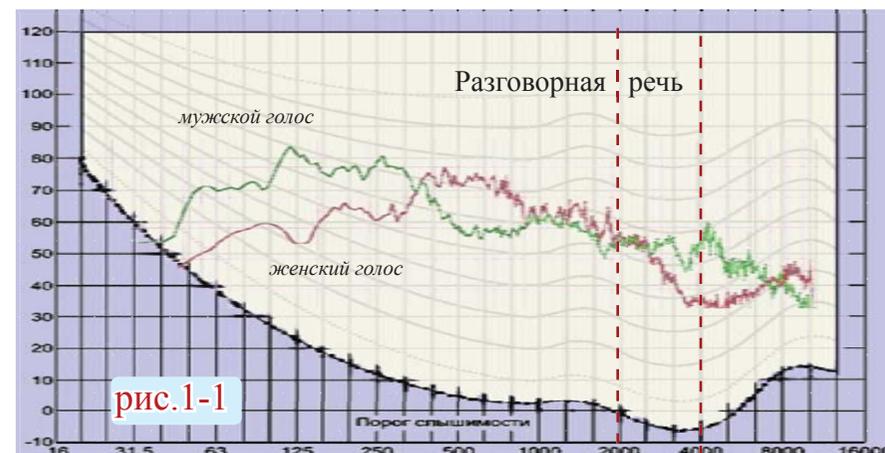
На рисунках 1-1..1-3 приведена нижняя граница области слухового восприятия - порог слышимости, относительно которого размещены спектральные характеристики человеческих голосов, и показан ход кривых равной громкости.

В частности, на рис.1-1 зеленой линией показан спектр отрезка речевого сигнала относительно низкочастотного мужского голоса (низкий баритональный тембр), и там же красной линией показан спектр отрезка речи женского голоса, тембр которого близок к представлениям о вокальном типе меццо-сопрано.

На рис.1-2 показаны сглаженные спектры гласных звуков, спетых профессиональными вокалистами, как мужчинами, так и женщинами. Характер обработки при сглаживании (выделении огибающей спектров) здесь таков, что он не дает верных значений спектральной функции для различных частот, но достаточно хорошо указывает характер распределения энергии звуковых сигналов по частотному диапазону.

На рис.1-3 все указанные графики совмещены, и это позволяет сформулировать основные существенные различия между спектрами сигналов разговорной речи и "вокальной речи" в терминологии В.П.Морозова.

2. Сразу отметим, что сравнение приведенных здесь данных не является абсолютно корректным, потому что источником для анализа спектров разговорной речи явились целые произнесенные фразы, а для вокального варианта выбраны лишь отдельные спетые гласные звуки. Причин такого выбора несколько, но главной можно назвать то, что с точки зрения звукорежиссера качество передачи разговорной речи определяется ее разборчивостью, что требует внимания к частотному диапазону всех артикулируемых звуков. При передаче же вокального материала существенным является сохранение



тембровых качеств голоса, которые в пении проявляются преимущественно при исполнении длительных гласных звуков.

3. Из приведенных рисунков сразу очевидна радикальная разница между поведением спектров речи и вокала, которая состоит в наличии ярко выраженного энергетического подъема у вокалистов в области примерно от 2 до 4 кГц. Этот подъем называется верхней певческой формантой (ВПФ), и она удивительным образом совпадает у разных певцов - как мужчин, так и женщин.

Подробности, связанные с ВПФ, мы обсудим в дальнейшем, а здесь отметим, что исследователи певческих голосов обращают внимание на своеобразное звучание этой форманты, что позволяет опытному исследователю ощущать ее наличие или отсутствие в голосе вокалиста. Эти обстоятельства позволяют сразу выделить область 2..4 кГц в качестве характерного поддиапазона области слухового восприятия.

Музыкальные инструменты

4. На рисунках 1-4..1-9 приведены спектральные картины, характерные для различных музыкальных инструментов. Источником звуков является распространенная в интернете "Энциклопедия музыкальных инструментов", в которой представлены так называемые "музыкальные катушки" - в них есть возможность хроматически воспроизвести диапазон звуков, доступных для исполнения на выбранном инструменте.

Кумулятивный (накопительный) спектр такой хроматической гаммы показывает частотный диапазон, который, в принципе, может занимать исполнение пьесы на инструменте, с учетом, разумеется, всей обертоновой структуры звуков этого инструмента [Ак,156].

Представленные спектры носят иллюстративный характер, их размах на картинках по вертикали не соответствует децибельной шкале области слухового восприятия, но для всех графиков при фактических измерениях спектров этот размах был установлен равным 60 dB. Графики спектров расположены на рисунках так, чтобы их правый край очевидно указывал на значение частоты, при котором, таким образом, спектральный уровень инструмента ниже максимального значения в этом графике на 60 dB.

Нетрудно видеть, что для всех выбранных оркестровых музыкальных инструментов эта граничная частота находится вблизи значения 8 кГц.

Таким образом, еще один маркер, определяющий поддиапазоны области слухового восприятия, может быть установлен в точке 8 кГц. Заметим, что поддиапазоны 2..4 кГц и 4..8 кГц находятся в октавном соотношении.

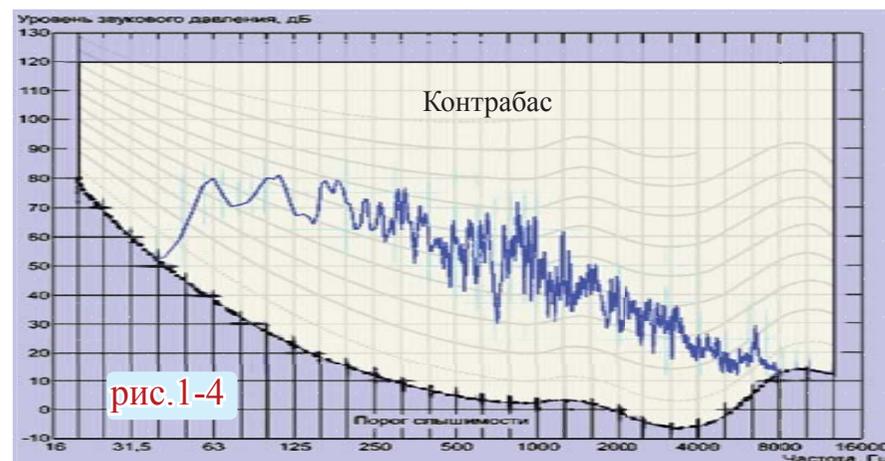


рис.1-4

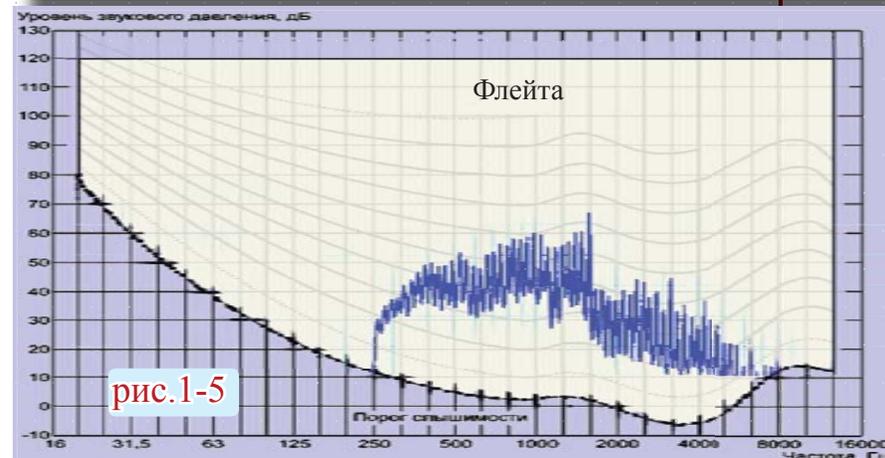


рис.1-5

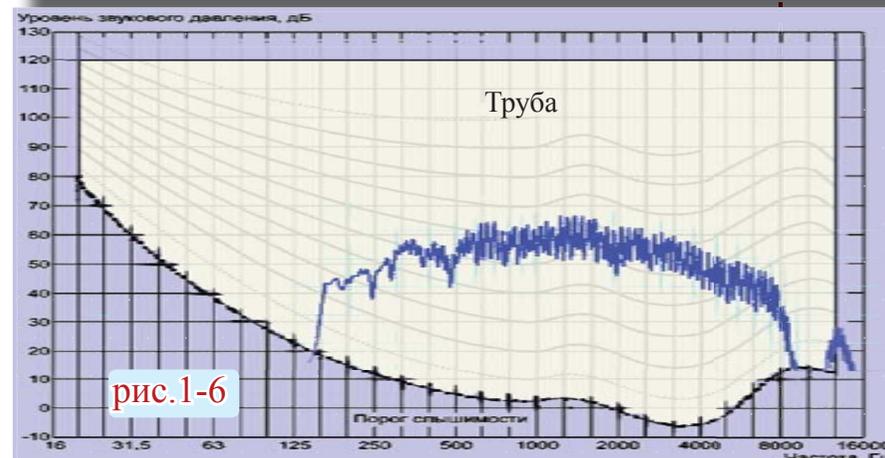
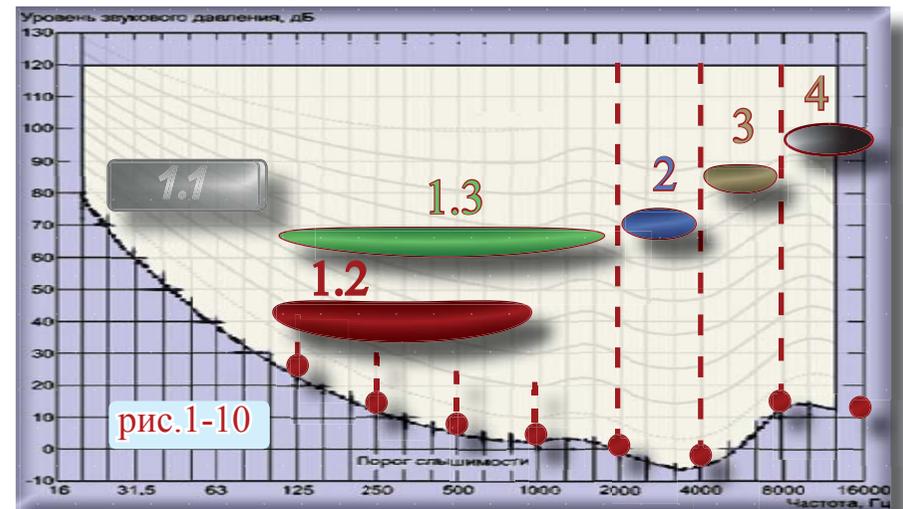
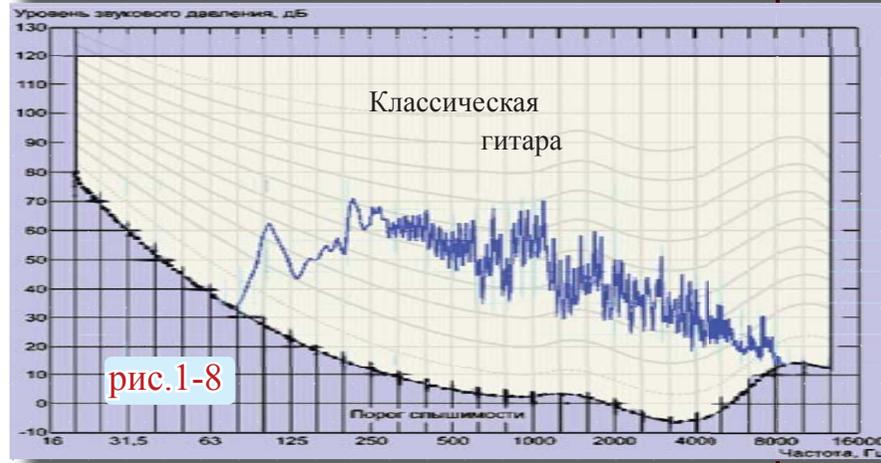
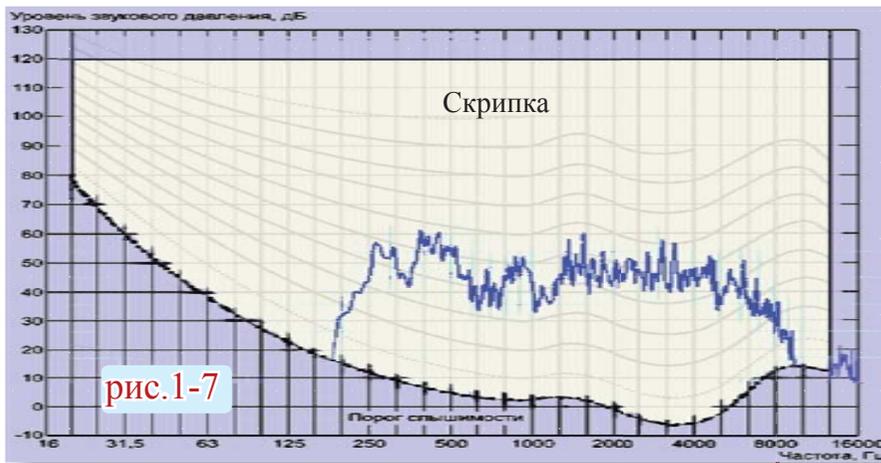


рис.1-6



Итоговое разделение области может иметь вид, показанный на рис.1-10.

Такое разделение не является бесспорным. Более того, каждый практикующий звукорежиссер в зависимости от характера типичных для него творческих задач сам определяет для себя характерные частотные области, в которых он манипулирует звуковым материалом. Предложенное здесь условное разделение области слухового восприятия на поддиапазоны служит целям этой книги и объясняется следующим образом.

Области слухового восприятия

5. Всю область слышимых частот мы мыслим как совокупность 4 функционально основных частотных диапазонов, внутри которых можно заметить и дополнительные (даже перекрывающиеся) поддиапазоны:

1. 20 Гц - 2 кГц,
 1. 20 Гц - 100 Гц,
 2. 100 Гц - 1 кГц,
 3. 100 Гц - 2 кГц
2. 2 - 4 кГц,
3. 4 - 8 кГц,
4. 8 - 20 кГц.

Обсуждение глобальных границ области (20 Гц и 20 кГц) состоялось ранее [Ак, 84] и к нему мы не возвращаемся.

6. Область самых низких частот, примерно до 100 Гц, принадлежит преимущественно шумам. Это могут быть шумы окружающей природной

среды, и шумы, входящие в состав монтируемого звукового (речевого или музыкального материала). Диапазон шумов, конечно, в принципе простирается и выше 100 Гц, но здесь мы подчеркиваем, что ниже этой частоты мы не ожидаем сколько-нибудь значимых интонированных звуков, принадлежащих музыкальному или речевому материалу. Ну, разве что контрабас сыграет в нижнем регистре самую низкую ноту. Поэтому можно несколько преувеличенно считать, что эта область безраздельно принадлежит только шумам, даже если они по своей природе захватывают, кроме того, и области интонированных звуков, смешиваясь там с ними.

7. Область 100 Гц - 2 кГц, это принципиально очень важная область. В ней располагается основная часть мелодико-гармонического содержания музыки, здесь же находятся форманты речевых гласных. У певческих голосов эта область принадлежит так называемой нижней певческой форманте (НПФ), в которой сконцентрированы обертоны, определяющий тембровую индивидуальность голоса.

Разумеется, положение обертоновой структуры звуков зависит от типа голоса или инструмента, частота их основного тона может лежать примерно в пределах от 100 Гц (середина большой октавы) до 1000 Гц (конец второй октавы), куда достигает профессиональное колоратурное сопрано в высокой тесситуре.

Мы здесь вновь говорим о голосах, но надеюсь, что вы давно заметили - инструментальное изложение мелодий примерно следует расположению голосов поющих людей. Слушатели рефлекторно молча "подпевают внутри себя" мелодии и попытка вынести мелодию за пределы возможного голосового "сопровождения" вполне может привести к неудаче музыкального произведения. Позже мы увидим, что изложение мелодии вне указанных выше пределов приводит также к принципиальным дефектам в восприятии музыкальной интонации вследствие свойств слуха [Ак,114].

8. В любом случае, если говорить о певческих голосах, примерно к 2 кГц заканчивается область основного тембрового резонансного подъема и в области 2 кГц - 4 кГц начинается область энергетического подъема в спектре голосов квалифицированных вокалистов, названная ранее верхней певческой формантой (ВПФ). Эта область обладает характерным специфическим звучанием, поэтому есть смысл обращать внимание на характер спектральной функции в этом поддиапазоне также и для музыкальных инструментов (у которых "верхней форманты" нет).

9. Всю область частот свыше 4 кГц можно назвать "областью (или зоной) призвуков". Это не значит, что компонентами звука, находящимися в ней, можно пренебрегать, или непременно следует их подавлять. Просто каждая

из указанных здесь частотных областей формирует весьма характерный по тембровой окраске вклад в общий тембр звука, будь то шумовой сигнал или звук музыки. Высокочастотные компоненты, лежащие выше 4 кГц в различных сочетаниях могут добавлять неприятно звучащие свистящие и шипящие (сами по себе) элементы. Поскольку обычно эти добавки энергетически незначительны, они лишь дополняют общий тембровый "портрет" звука, но, назвав их призвуками, мы тем самым указываем их функциональную и эстетическую роль в этом портрете.

10. В зоне призвуков метка 8 кГц является характерной. Во первых, как уже было отмечено, вблизи этой частоты заканчивается основная энергетически значимая часть спектра музыкальных инструментов. Во вторых (как будет показано дальше) именно эта частота ограничивает звучащую часть свистящих звуков "з". "ц". "с". если они отчетливо выражены в речевом материале.

Еще заметим, что черные шеллачные патефонные (грамофонные) пластинки, которые в относительно давние времена предшествовали нынешним виниловым носителям, содержали записи, верхняя частота которых, в силу технологии изготовления этих пластинок, составляла примерно 8 кГц. Такие музыкальные материалы, безусловно, были лишены качества, присущего современным записям, однако люди слушали и коллекционировали их. Часть уникальных записей таким образом дошла до нашего времени.

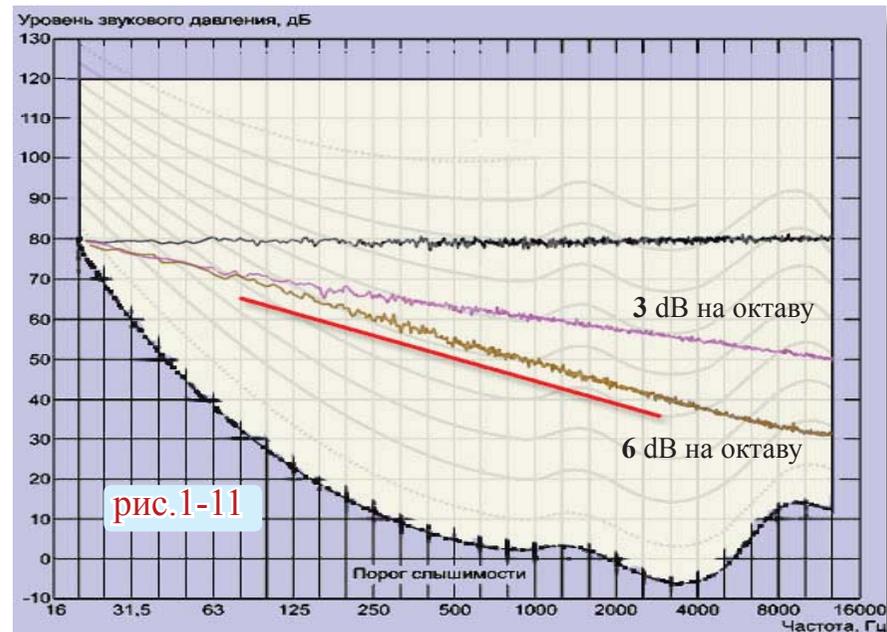
11. Части звукового материала, расположенные в области выше 8 кГц, содержит призвуки, доступные не всем слушателям. Во всяком случае, это относится к компонентам звука с частотами выше 14 кГц. Заметим, что спектр записей на свежечупленных виниловых пластинках, как говорят, простирается до 17-18 кГц, однако после нескольких проигрываний, вследствие "проработки" дорожек иглой он снижается до 13-14 кГц. Именно в области выше 13-14 кГц происходит утрата звучащих высокочастотных компонент с возрастом.

Говорят, что, начиная примерно с 30 лет, утрата (то есть подъем порога слышимости в этой области) составляет примерно по 10 dB на каждые 10 лет жизни. Конечно, характер области слышимости таков, что можно добиться восприятия плохо слышимого звука за счет существенного увеличения его интенсивности (а значит и громкости). Но мы исключаем такие нелегальные манипуляции звуком и, следуя мнению, в частности, американских авторов, считаем, что нормальным восприятием звука в домашних условиях следует считать уровни интенсивности вблизи слушателя порядка 60-70 dB.

В расположении звукового материала в области призвуков нет каких-либо явных закономерностей. В некоторых случаях можно увидеть зоны энергетического подъема вблизи частоты 6 кГц, иногда вблизи 8 кГц, вблизи 11 кГц,...

Эти подъемы привносят в тембр звуков оттенки, о которых говорят "стеклянный", "металлический", "шипящий", "кислотный" и т. п. Иногда эти свойства звука составляют его отличительную черту, иногда от них стараются избавиться путем обработки сигнала.

Музыка в областях

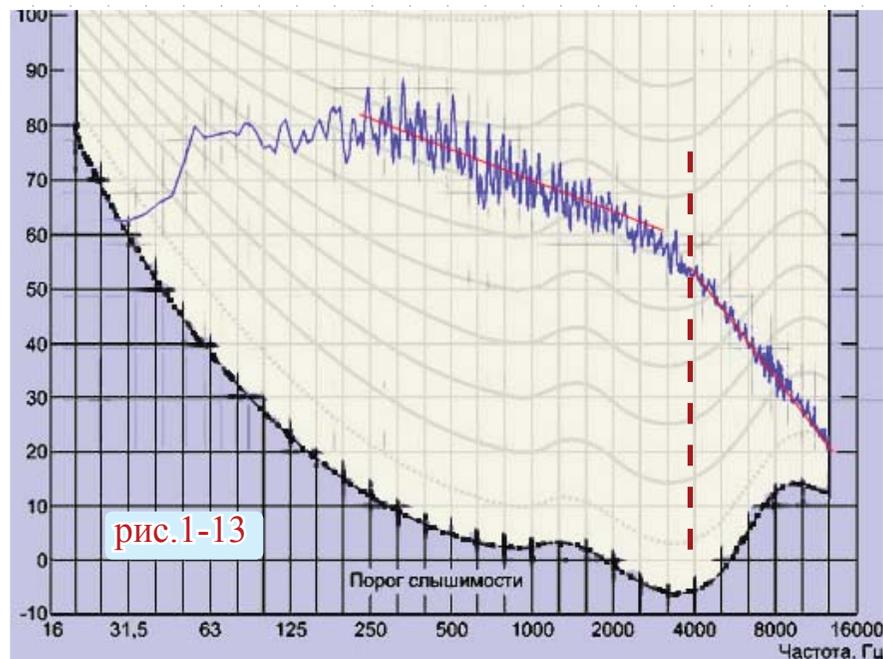
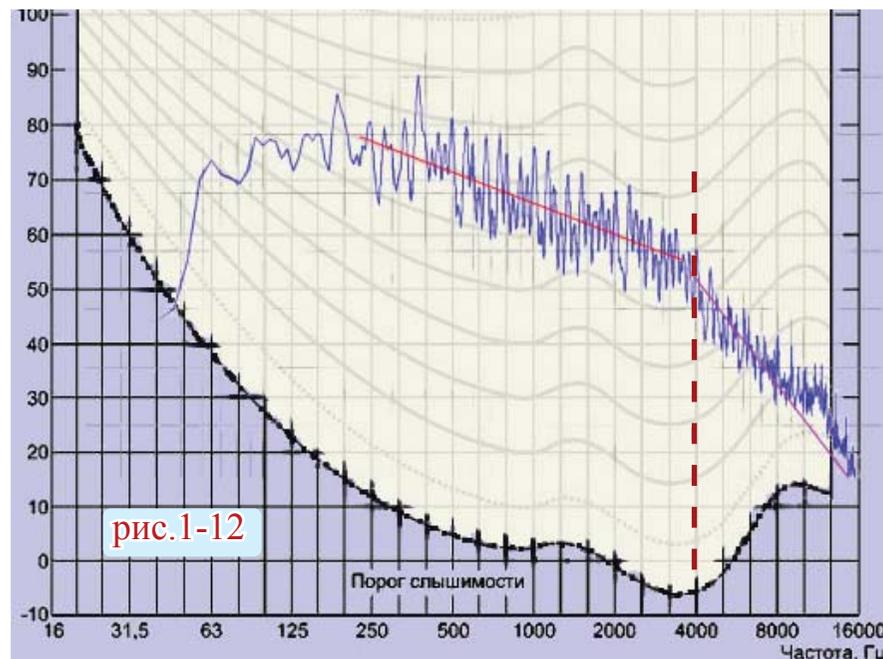


12. Для ориентации в характере спектров музыки и речи мы приводим на рисунке 1-11 спектральные функции так называемых: "белого" шума (горизонтальная черная линия), "розового" шума (фиолетовая линия), интенсивность которого спадает с показателем 3 dB на октаву и "коричневого" шума (коричневая линия), спадающего с показателем 6 dB на октаву.

Красная прямая линия сопровождает коричневый шум в качестве шаблона, он будет нанесен на спектральные характеристики образцов музыкальных и речевых фрагментов.

Первый из таких образцов приведен на рис.1-12 - это увертюра из второй сюиты Баха, исполняемая преимущественно струнными смычковыми инструментами камерного оркестра.

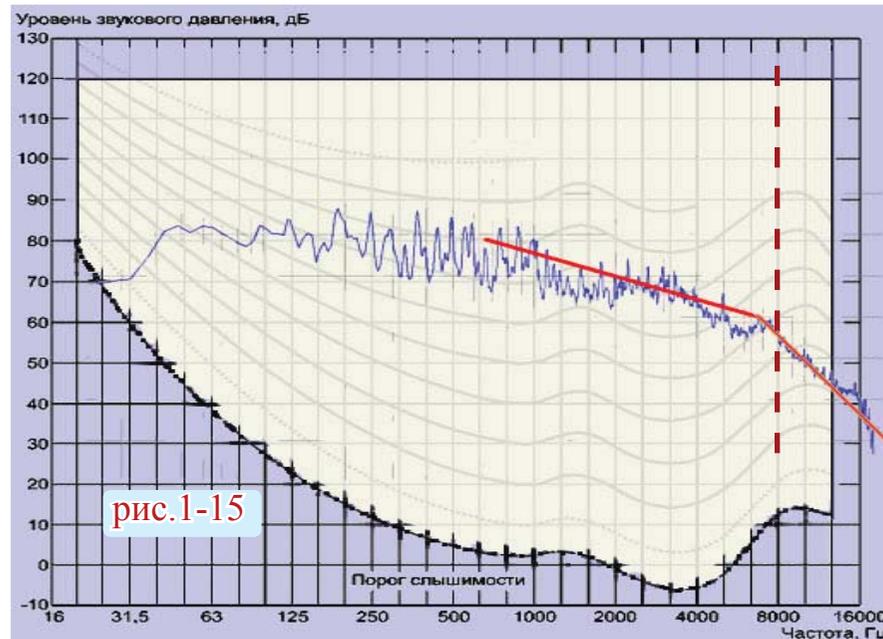
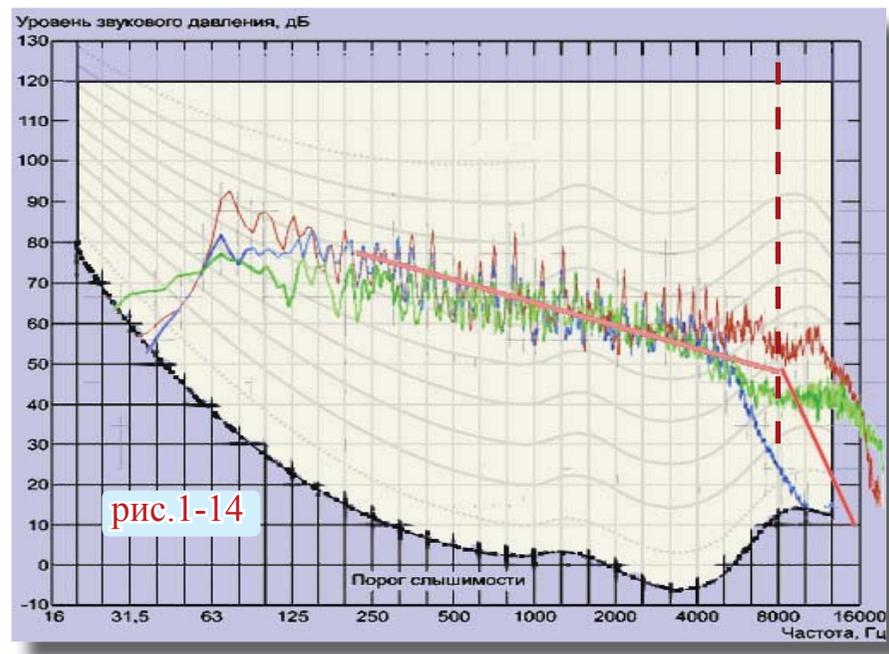
Второй образец (рис.1-13) принадлежит оркестровому фрагменту из второго концерта Рахманинова для фортепиано с оркестром, в котором выразительно участвуют и духовые инструменты симфонического оркестра.



В обоих рисунках отчетливо различаются сравнительно плоская часть спектра примерно до 500 Гц, фрагмент, убывающий со "скоростью" 6 dB на октаву в диапазоне примерно от 500 до 4 кГц, и более крутой участок спектра, убывающего с показателем крутизны примерно 20 dB на октаву, в области свыше 4 кГц.

На рис.1-14 сведены вместе спектры одного и того же фрагмента известной пьесы "In The Mood" в аранжировке Дж.Гарланда - Г.Миллера, исполненной биг-бэндом. В биг-бэнде отсутствуют струнные инструменты и весьма широко представлены медные духовые, обертоновый ряд которых простирается в область относительно высоких частот. В связи с этим общая спектральная картина выглядит несколько иначе, нежели рис.1-13. Синяя линия отражает усредненный спектр записи, сделанной в 1939 году в Бродвейской радиостудии на устаревшем (по нынешним меркам) оборудовании. Зеленой линией показан спектр записи, реализованной на виниловой пластинке в более поздние годы. Красная линия принадлежит спектру студийной записи в формате CD-Audio с оркестром Хуго Штрассера (Германия). Все три записи реализуют одну и ту же классическую аранжировку с идентичными составами оркестра. Поведение спектральных функций в области крайних высоких частот зависит, разумеется, от типа носителя записи, но при достижении верхней доступной границы спектр спадает далее с крутизной порядка 40 dB на октаву.

Контрастной по характеру является музыка группы Queen "Show Must Go On", спектр фрагмента которой показан на рис.1-15.



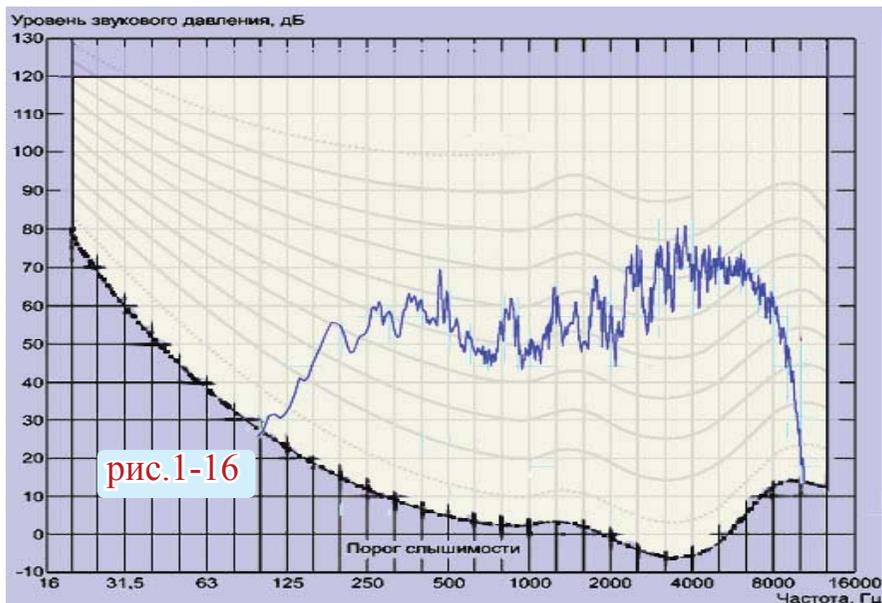
Наличие вокала и электронных инструментов вносит некоторую специфику в спектр, однако и здесь, начиная с 8 кГц, он спадает с крутизной порядка 20 dB на октаву.

Вообще, спектры ансамблей с синтезированными электроникой звуками могут заметно отличаться от приведенных здесь образцов. Зачастую целью такого синтеза является экзотичность, экстравагантность этих звуков, и она отражается в спектральных характеристиках. Специфичный стиль музыки даже при традиционных музыкальных инструментах может вызвать отклонения спектра от привычной картины. Так, например, очевидно, что "злоупотребление" в использовании ударных тарелок, (спектр звука тарелки Ride от Zinjan показан на рис.1-16) может привести к существенному подъему общего спектра вблизи 8 кГц.

13. Таким образом, в работе над музыкальным материалом можно высказать следующие пожелания.

Производя различные корректирующие манипуляции со звуком, следует очень бережно относиться к его части, расположенной в области до 2 кГц, бережно - к части, расположенной в области от 2 до 4 кГц, осторожно - к части, расположенной в области от 4 до 8 кГц, и внимательно - к части, расположенной выше 8 кГц.

Поясним это, отчасти повторив сказанное ранее.



В области ниже 2 кГц располагается основная часть, определяющая характер тембров как музыкального так и речевого материала. Изменения в этой области могут существенно деформировать тембровую картину звукового полотна.

Область 2 - 4 кГц формирует тембровую добавку к звуку со специфичным ярким "блестящим" колоритом. У певцов это верхняя певческая форманта, у других музыкальных инструментов - тоже выразительная компонента, уровень которой иногда есть смысл немного поднять. В этой области чувствительность слуха наиболее высока и микс от этого становится более звучным.

В диапазоне 4 - 8 кГц располагаются, так сказать, "деликатные" компоненты спектра звуков. У музыкальных инструментов верхние обертоны, расположенные здесь, подчеркивают тембровую разновидность инструмента, например бренд гитары или тип саксофона (альт/сопрано). В речевом материале вмешательство в эту область может привести к шепелявости звонких согласных.

Компоненты звука в области выше 8 кГц, как было сказано ранее, в большинстве случаев носят призвуковый характер. И вмешиваясь в эту область, следует внимательно следить, освобождаем ли мы музыкальные материал от призвука, или сам музыкальный материал носит характер "призвука", который нужно бережно сохранить, как это может быть при сольном звучании оркестрового треугольника, некоторых перкуссий и т. п. Речевой материал в этой области ничего ценного не содержит кроме шумов, которые могут быть беспощадно подавлены.

Несколько слов об уровнях

14. Теперь обсудим "вертикаль" области слухового восприятия, градуированную в децибелах. Будем поначалу ориентироваться на средние частоты этой области, где чувствительность слуха наиболее высока - это примерно диапазон от 500 Гц до 6 кГц.

Известно, что громкость 1 сон соответствует уровню интенсивности в 40 dB SPL [Ак, 107] для указанного диапазона частот. Говорят, что это - примерно громкость тихого шепота на расстоянии 1 м. Заметим, что уровень в 30 dB - это практически полная тишина для релаксации организма. Ничтожные шумы проникают при этом в помещение из внешнего окружения и вследствие работы различных систем обеспечения в зданиях. Для достижения уровня порядка 20 dB придется применять специальные звукопоглощающие средства, известные проектировщикам звуковых студий.

Итак воспроизведение музыкального материала с динамическим оттенком *pp*, видимо предполагает уровень интенсивности воспроизведения порядка 40 dB. Заметим, что оркестровое, фортепианное, виолончельное и скрипичное *pp* исполнители наверняка будут трактовать по разному, однако здесь речь идет о желательных уровнях при воспроизведении записей музыки для группового прослушивания. Оттенок *ff* даже в оркестровом *tutti* вряд ли следует позволять воспроизводить с уровнем, значительно превышающем 90 dB.

Таким образом, можно согласиться с известной в литературе условной шкалой музыкальных уровней:

<i>pp</i> - 40 dB	<i>p</i> - 50 dB
<i>mp</i> - 60 dB	<i>mf</i> - 70 dB
<i>f</i> - 80 dB	<i>ff</i> - 90 dB

Заметим, что могут встретиться еще экзотические указания *ppp* и *fff*, для которых оставим ориентиры 30..40 dB и 90..100 dB - с кратковременным "всплеском" эмоций такого рода можно согласиться.

Разумеется, вряд ли кто-то захочет систематически контролировать уровни звучания материала с такой точностью, в концертном зале это вообще бессмысленно. Но при студийной записи, а также в условиях домашнего воспроизведения эти величины можно иметь в виду. Заметим также, что шаг в 10 dB соответствует приблизительно удвоению слышимой громкости звука: музыканты могут соотносить с этим обстоятельством свои собственные трактовки указанных в нотах нюансов, соглашаясь с ним или отступая от него.

Разумеется, приведенные цифры весьма умозрительны и условны. Для одиночного фортепиано указывают, например, иногда соответствие *pp* - 35 dB.

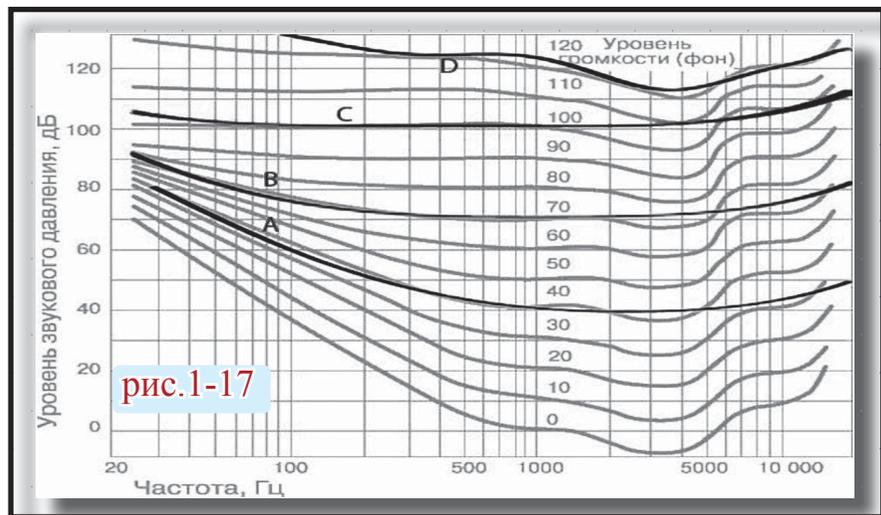
Но камерный оркестр такое пианиссимо не сыграет, кроме того в концертном зале уровень шума будет заведомо выше.

Фортиссимо большого оркестра может, конечно, развить в tutti уровень порядка 100 и больше децибел вблизи сцены, если выдержит дирижер. Но вряд ли можно рекомендовать такие уровни слушателям в зале.

Приведенные значения уровней обеспечивают динамический диапазон музыки 40..50 dB (в среднеквадратичных значениях) и эту величину можно считать правильной как для ее записи, так и для воспроизведения при прослушивании (разумеется, полный динамический диапазон записи с учетом необходимого превышения шумового "порога" и возможных пиковых выбросов звука должен превышать эту величину).

15. Вероятно, мы уже забыли, что речь здесь все время шла о частотном диапазоне с границами 500 Гц .. 6 кГц, именно поэтому децибелы для музыкальных динамических оттенков можно было приблизительно воспринимать как dB SPL [Ак,99]. Однако в музыкальной деятельности очень важен как раз диапазон от 100 до 500 Гц, в котором кривые равной громкости круто поднимаются вверх в сторону низких частот.

Вообще-то вычисление уровня интенсивности, само по себе, не предполагает сигнал в виде синусоидальных колебаний - звуковой сигнал любого вида можно возвести в квадрат, прологарифмировать с коэффициентом 20, затем результат немного усреднить (проинтегрировать) во времени, то есть сделать его инерционным, чтобы успевать следить за изменениями этого результата. Но, если такие операции совершить с сигналом, не применяя никакой



обработки, то полученная цифра укажет уровни звука, распространяющегося в пространстве. Для того, чтобы узнать уровни слышимого человеком сигнала, нужно звук пропустить через фильтр, частотная характеристика которого близка к частотной характеристике слуховой системы, а такая характеристика должна соответствовать кривым равной громкости, зеркально отраженным относительно горизонтальной линии.

Фильтры подобного рода присутствуют в профессиональных измерителях уровней (Sound Level Meter), называемых часто шумомерами. Поскольку кривые равной громкости непараллельны, при различных интенсивностях входных сигналов приходится применять несколько различающиеся фильтры. Наиболее часто встречаются так называемые фильтры типа А (для слабых сигналов) и фильтры типа С (для громких сигналов в пределах допустимых санитарных норм). Уровни звука, измеренные с применением таких фильтров, выражают, соответственно, в децибелах с обозначением dBA и dBC.

Таким образом, в случае измерений уровней реальных звуков для динамических музыкальных оттенков до *mf* следует пользоваться шкалой dBA, а свыше *mf* - dBC.

16. В приводимых материалах мы не скрываем приблизительность используемых данных и не стесняемся неточности получаемых результатов. Кривые равной громкости получены не с вашим личным участием и не характеризуют ваши конкретные свойства в данный момент в данной ситуации. Частотные характеристики фильтров приближенно соответствуют лишь некоторым усредненным таким кривым (см. перевернутые "под кривые равной громкости" характеристики фильтров А,В,С и D на рис.1-17, [Ак,112]).

Психофизиология слухового восприятия - это наука не точных формул, а приблизительных графиков. И иногда она отвергает более точные формулы (например закон Стивенса) в пользу более удобных графиков (закон Вебера-Фехнера и кривые равной громкости). Графики время от времени перепроверяют и уточняют. Это единственно возможный путь.

Вот почему я думаю, что отклонение интуитивно ощущаемого уровня звука от точного результата измерения на 5 dB вообще не является ошибкой. Но отклонение на 10 dB является неприятной ошибкой, а отклонение на 20 dB свидетельствует о профессиональной несостоятельности звукорежиссера.

Разумеется, последнее соображение не может относиться к уровням выше 90 ..95 dB, при которых разумная деятельность человека невозможна.

2 Спектры звуковых сигналов

Кто такой спектр

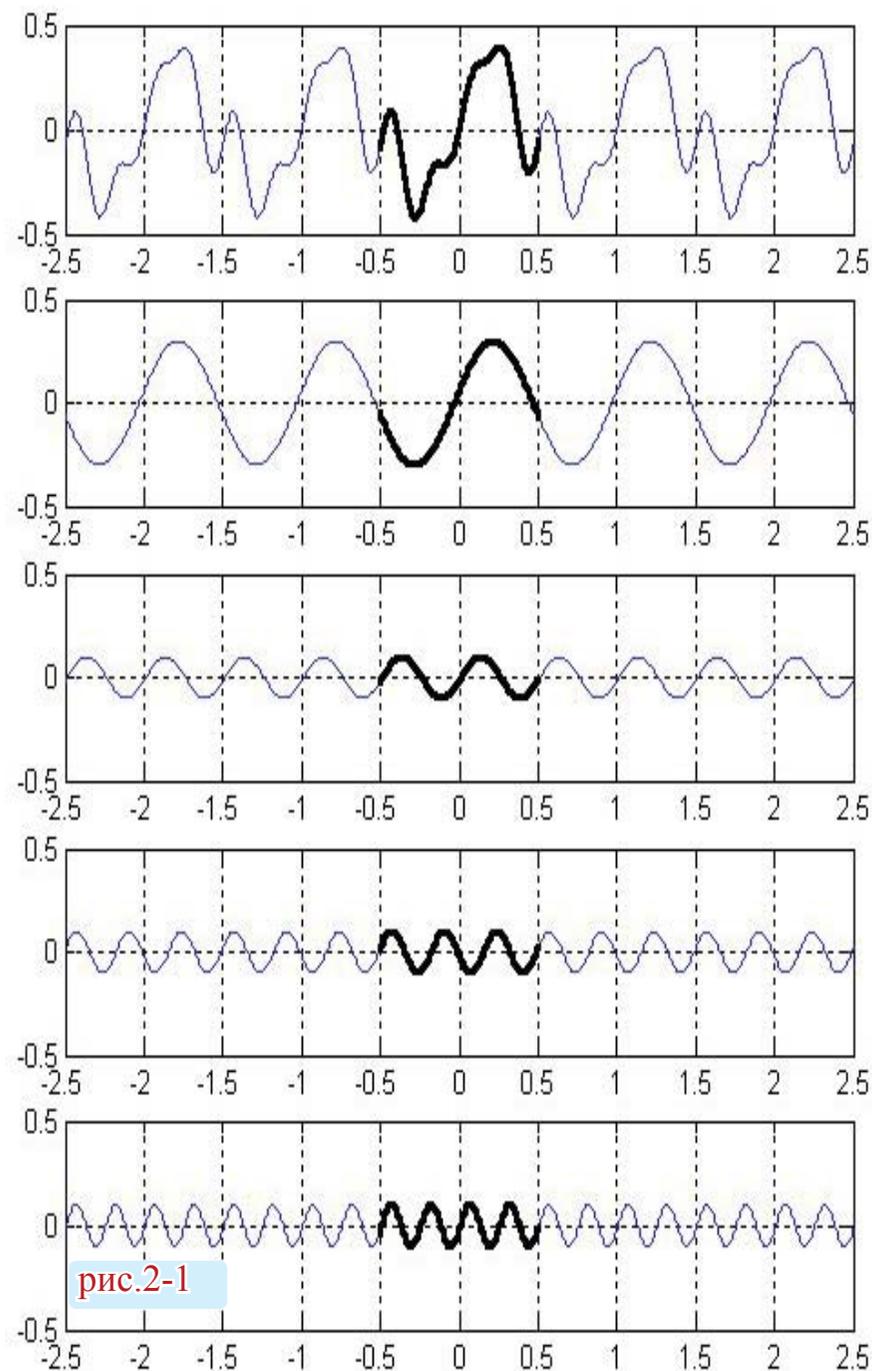
1. Рассуждения в предыдущей главе целиком основывались на спектральных характеристиках звуковых сигналов. Это неудивительно, поскольку при необходимости дать какую-либо объективную оценку сигналу первое возникающее желание квалифицированного аудиоспециалиста - увидеть спектр этого сигнала. Само собой разумеется, что до этого нужно проконтролировать аудиоредактором изображение этого сигнала для того, чтобы убедиться - сигнал не претерпел ограничений уровня, что привело бы к его нелинейным искажениям.

Средств объективного анализа звуковых сигналов не так уж много и среди универсальных инструментов спектральный анализ - первый и главный инструмент. Поэтому рассмотрим его подробно.

2. Основой спектрального анализа является теорема Фурье, в связи с которой употребляются также термины, "ряд Фурье", "преобразование Фурье" [Ак, 54]. Это несколько различающиеся, но идейно близкие преобразования сигналов. В иллюстративных материалах мы применяем формулы для наиболее современного варианта спектральных преобразований, а именно "Дискретное Преобразование Фурье" (ДПФ). Есть еще один термин - Быстрое Преобразование Фурье (БПФ или FFT - Fast Fourier Transform) - это чрезвычайно быстрый и красивый метод вычисления ДПФ, которым пользуются повсеместно при спектральных вычислениях.

3. На рис.2-1 показан общий принцип разложения в ряд Фурье периодического сигнала. Именно для периодических сигналов была изначально сформулирована теорема Фурье. На рисунке выделено разложение в ряд одного периода T такого сигнала (для простоты величина периода выбрана равной единице, что изображено интервалом от $-0,5$ до $+0,5$). При этом один из важнейших фактов состоит в том, что частоты синусоидальных сигналов, на которые может быть разложен исходный сигнал, predeterminedены сразу же и всегда равны величинам $1/T, 2/T, 3/T, 4/T, 5/T, \dots$

Так, например, для периодического сигнала произвольной формы, если его период составляет 200 мсек (0,2 сек), то набор составляющих содержит синусоиды с частотами 5, 10, 15, 20, 25, ... (Гц). А, если период равен 2 сек, то частоты синусоид в наборе равны 0,5, 1, 1,5, 2, 2,5, ... Гц. Это твердое правило. Что же касается количества синусоид в наборе, то оно зависит от вида исходного разлагаемого сигнала. На приведенном рисунке оказалось, что ис-



ходный сигнал составлен всего из 5 синусоид. В общем же случае их количество формально (математически) может быть бесконечным.

Мы еще вернемся к вопросу об этой бесконечности (она несколько не должна беспокоить), а сейчас обратимся к рисункам 2-2..2-11 на странице 30. На них изображен, прежде всего, прямоугольный импульс (рис 2-2), а внизу на рис.2-9,2-10 показаны спектры амплитуд и фаз в разложении Фурье этого импульса.

В этом месте у нас небольшая интерлюдия¹.

Интерлюдия

4. Мы уже говорили [Ак,23], что синусоидальное колебание характеризуется всего тремя параметрами - частотой колебания (либо однозначно с ней связанным периодом, но лучше и привычнее все же частота), амплитудой и фазой. Скажем это иначе - дайте мне три этих числа и больше ничего не нужно - синусоидальное колебание полностью определено. Поскольку форма его известна и зависит только от этих трех чисел, то мы можем, говоря о синусоидальных колебаниях, не рисовать их а обсуждать для синусоид только эти тройки чисел. Из этих троек, как мы уже говорили, все значения частот синусоид, входящих в спектр сигнала, жестко связаны с его длительностью и только длительностью. Другого выбора для этих частот нет, поэтому обсуждать их здесь далее нет смысла, остаются только двойки чисел. А форма звукового сигнала (звукорежиссеры иногда говорят "волны") определяет, как раз, именно эти пары чисел для каждой синусоиды - ее амплитуду и фазу. Вот эти-то числа и представлены в графической форме на рисунках, названных "спектр амплитуд" и "спектр фаз".²

Фазирование колонок

5. С фазами у нас отношения сложные. Поскольку величина фазы просто фиксирует нулевую точку отсчета времени синусоидального колебания, то для одиночной синусоиды она вообще не представляет интереса. Интерес, и немалый, к фазам возникает, когда речь идет о нескольких взаимодействующих в пространстве колебаниях, то есть о пространственном звуковом поле. Вот тогда различие в их фазах оказывается важным и иногда даже критическим фактором.

1 Интерлюдия - вставной промежуточный фрагмент в общем музыкальном произведении

2 Меня упрекают иногда профессионалы-акустики за чрезмерное "разжевывание" материала. Но выхода другого я не вижу. Одной-двумя формулами можно, так сказать, "закрыть вопрос". Однако большинству звукорежиссеров формулы чужды и могут только оттолкнуть, а чрезвычайно важность материала побуждает меня донести его до понимания коллег, которые не специализировались в этих вопросах.

В качестве примера приведем фазирование колонок левого и правого каналов стереосистемы. Как правило, сигнальные кабели для этих колонок маркированы цветом. Но в случае, если такая маркировка отсутствует, нетрудно маркировать кабели самому. Нужно поставить обе колонки рядом и подать на них синусоидальный (или узкополосный) сигнал, длина волны которого заведомо больше габаритных размеров колонки. Кабеля при этом подсоединить произвольным образом, но затем поменять полярность кабельной пары на одной из колонок. В одном из двух случаев громкость звучания пары колонок будет отчетливо больше - это и есть правильное соответствие кабельных пар обеих колонок.

Дело в том, что при "неправильном" противофазном подключении динамики левого и правого каналов работают так, что когда, например, динамик левого канала движется вперед (создавая волну сжатия частиц поля), динамик правого канала идет назад (создавая волну разрежения частиц поля), и в значительной мере происходит "перекачивание" частиц звукового поля от динамика к динамику. Это называется акустическим коротким замыканием. Воздействие же обоих динамиков на звуковое поле при синхронном их движении вперед (на сжатие) и назад (на разрежение) является наиболее эффективным и это ощутимо на слух.

Разнесение в пространстве колонок левого и правого каналов, разумеется, снижает влияние их фазировки на "частотную характеристику" создаваемого звукового поля. Но правильная фазировка вообще исключает роль стереобазы в этой характеристике.

Фазовая характеристика эквалайзера

6. В случае единственного сигнала даже фазовые сдвиги синусоидальных составляющих "внутри" этого сигнала не влияют на его слуховое восприятие. В этом можно убедиться, если сложить сигнал из нескольких синусоид и манипулировать фазовыми сдвигами между ними - форма сигнала будет изменяться, но на слух различие не обнаружится. Это, в частности, означает, что пропуская одиночный звуковой сигнал через фильтр (например эквалайзер) нас интересуют амплитудные характеристики этого фильтра - при различных фазовых характеристиках слуховой результат фильтрации (эквалайзинга) будет один и тот же. Заметим, что это обстоятельство соответствует известным теориям слухового восприятия звуковых сигналов.

Изъятие вокала из песни

7. Еще одно интересное обстоятельство следует из упомянутого слухового безразличия к фазам (еще раз напомним - сигнала, а не группы сигналов

в пространственном звуковом поле, где происходит интерференционное сложение компонент сигналов).

Инвертировав звуковой сигнал (перевернув его "вверх ногами"), мы тем самым инвертируем все синусоидальные его составляющие - это можно трактовать как внесение в них только лишь дополнительного фазового сдвига без изменения амплитуд этих синусоид. В соответствии со сказанным ранее нет ничего удивительного в том, что инвертированный сигнал звучит в точности так же, как исходный. Именно на этом свойстве основано так называемое "изъятие вокала из песни". Одновременно заметим, что изменение полярности (противофаза) входного сигнального кабеля одиночной колонки не приводит к каким-либо изменениям слышимого сигнала и требует лишь внимания к возможным заземлениям аппаратуры во избежание электрического короткого замыкания на выходе усилителя.

Прежде всего отметим, что изъятие вокала или другой музыкальной партии из микса возможно лишь в том случае, если эта партия по какому-либо признаку может быть отделена от всего остального микса. Во всех иных случаях возможно лишь скрупулезное и частичное "точечное выковыривание" фрагментов вокала из общей смеси, что решением задачи не является. Высказанная мысль очевидна и лишь акцентирует внимание на том, что природу обмануть невозможно ни в каком случае.

Спектр и положение вокальной партии во времени сильно смешиваются с остальными партиями и почти никогда не оставляют в этом смысле надежды. Таким образом, из монофонического сигнала извлечь вокал невозможно, как и невозможно отделить коровье молоко от добавленной в него в дальнейшем воды. Реальный шанс на изъятие вокала может появиться в стереозаписи, если голос записан точно посередине стереобазы, а аккомпанирующие инструменты четко разделены между левым L и правым R каналом без перекрытия (часть целиком в левом, часть целиком в правом). В этом случае независимо от временного и спектрального перекрытия вокала V и сопровождающего аккомпанемента процедура изъятия вокала может быть реализована так.

1. Исходное состояние каналов выглядит как $L + V, R + V$;
2. Инвертируем один из каналов (напомню, что его звучание от этого не изменяется): $-(L + V)$;
3. Суммируем получившуюся пару каналов: $-L - V + R + V = R - L$;
4. Задание выполнено! Мы получили монотрек $R - L$, который содержит только аккомпанемент, звучащий адекватно.

Некоторые аудиоредакторы, автоматически выполняющие эту работу, возвращают "стерео"пару каналов, но нетрудно убедиться, что оба канала в этой паре идентичны.

Отличным примером исходного материала, который позволяет весьма чисто удалить вокал таким образом, является запись песни Flowers On The Wall в исполнении Нэнси Синатры (из альбома Boots 1966 г).

Здесь возникает несколько вопросов.

Мы получили монотрек. Нельзя ли получить стерео с изъятим вокалом? Можно ли не изъять, а выделить вокал? Если мы сумели его изъять, то, возможно, некоторыми манипуляциями можно ликвидировать аккомпанемент, оставив лишь вокал. Что делать, если вокал не находится в центре звуковой картины, или если звук каждого инструмента аккомпанемента распределен между левым и правым каналами?

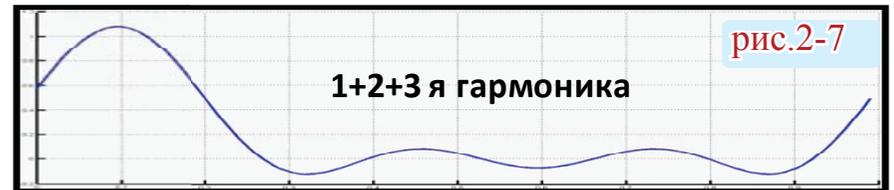
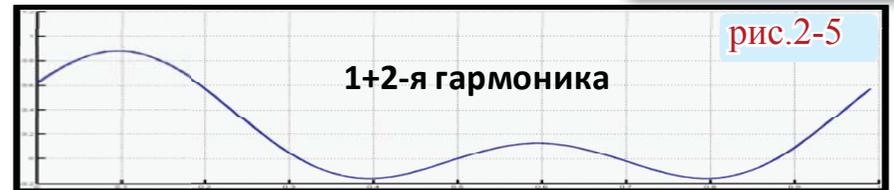
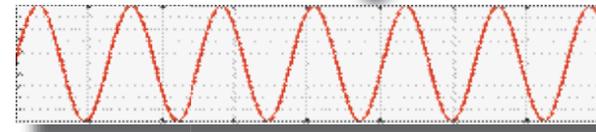
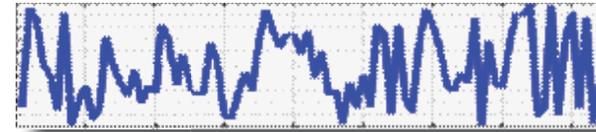
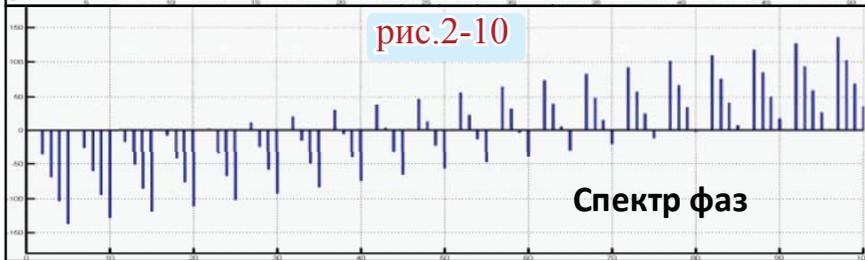
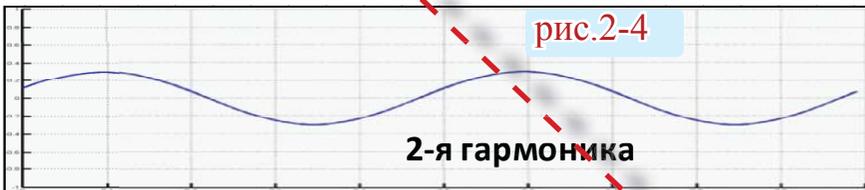
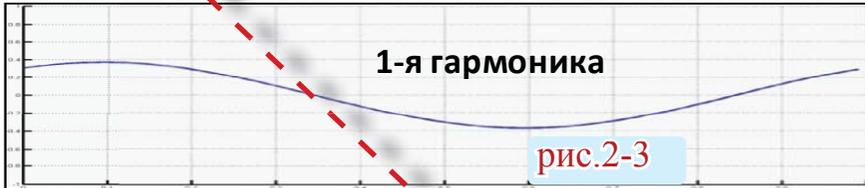
Начнем с конца. При этом все-таки изъятием вокала будем считать результат, в котором инструментальные партии сохранены в неизменном на слух состоянии.

1. Если вокал находится не в центре панорамы, значит вокальная партия записана в каналах с различным уровнем. Можно экспериментально уравнивать вокал в каналах, при этом инструментальный стереобаланс каналов будет нарушен и эти нарушения попадут в итоговый монотрек.
2. Если вокал находится в центре, но звуки инструментов частично попадают в оба канала ("левые инструменты" L представлены в каналах как "левый в левом" LL и "левый в правом" LR, и аналогично представлены "правые" инструменты как RR и RL), то формально можно записать сигналы каналов, как $LL + RL + V, RR + LR + V$, что приводит к инверсии - $(LL + RL + V)$ и монотреку $RR - RL - (LL - LR)$.

Нетрудно видеть, что инструментальные партии деформированы тем более, чем больше их перекрытие в каналах. Практика изъятия вокала показывает, что в большинстве случаев эти деформации оказываются неприемлемы.

3. Что касается не подавления а извлечения вокала в отдельный трек, а также возможности формирования полноценного стереоматериала без вокала, то мы предоставляем читателям самостоятельно разочароваться в таких возможностях.

Конец интервью



Как спектр работает

8. Я надеюсь, мы понимаем, почему в дальнейшем разговоре о спектральных свойствах сигнала исключим также и фазы, оставив для рассмотрения лишь спектр амплитуд. Это настолько естественно, что зачастую именно его называют просто спектром сигнала.³

Повторим обращение к рисунку 2-2, на котором изображен прямоугольный импульс а внизу на рис.2-9,2-10 показаны спектры амплитуд и фаз в разложении Фурье этого импульса.

Рассмотрим рисунок 2-9 под названием "Спектр амплитуд". На этом рисунке должны быть отображены амплитуды синусоидальных составляющих сигнала (по вертикальной оси) в зависимости от порядкового номера синусоид, который отображается на горизонтальной оси.

Но на горизонтальной оси картинке намного удобнее указывать не собственно номер синусоиды, а сразу ее частоту в герцах, поскольку для каждого номера k синусоидальной компоненты частота, как уже говорилось, равна $f_k = k \cdot f_1$, где $f_1 = 1/T$, а T - длительность сигнала. Спектр сейчас, это картинка в координатах амплитуда×частота.

Теперь посмотрим, как спектр "работает". На рис.2-3 показана 1-я гармоника⁴ с частотой $1/T$, амплитудой, взятой из спектра амплитуд, и начальной фазой, взятой из спектра фаз. Здесь мы не проигнорировали фазу, потому что собираемся не послушать синусоиду а складывать с ее участием прямоугольный импульс (см. примечание 3).

На рис.2-4 показана 2-я гармоника с частотой $2/T$, амплитудой, взятой из спектра амплитуд, и начальной фазой, взятой из спектра фаз, а справа. на рис.2-5 (стр.31) сумма этих первых двух гармоник.

Рис.2-6 показывает 3-ю гармонику с частотой $3/T$, а рис.2-7 справа. на стр.31 сумму первых трех гармоник. Видно, как "деликатно" начинает эта сумма формировать выпуклость в районе будущего импульса, одновременно подавляя все остальное пространство.

Сумма первых пяти гармоник (рис.2-8) делает это еще отчетливее, и это неудивительно - отсчитайте самые левые пять амплитуд в "спектре амплитуд", они энергетически весьма значимы.

3 Анализируя сигнал, мы интересуемся только спектром амплитуд. Но, синтезируя его, то есть складывая сигнал из синусоидальных компонент спектра, разумеется, необходимо учитывать фазовое состояние каждой компоненты.

4 Слово "гармоника" встречается в двух контекстах - либо это синоним слова "синусоида", либо элемент "гармонической последовательности" колебаний, частоты которых находятся в упомянутом соотношении $f_k = k \cdot f_1$.

Следующие за ними гармоники заметно слабее по амплитудам, но зато их очень много. И вот уже сумма 25 первых гармоник (рис.2-11) вполне отчетливо отрисовывает контур будущего импульса.

Можно поверить в то, что первые несколько сот гармоник дают результат, почти неотличимый от исходного импульса, который был подвергнут разложению на эти самые гармоники.

9. Именно таким образом, рассматривая картину спектра сигнала, мы видим области наибольших амплитуд синусоидальных составляющих, и понимаем, что именно они, в основном, формируют образ сигнала, а компоненты с маленькими амплитудами дополняют этот образ, придавая ему окончательный вид.

Современные алгоритмы Фурье-представления сигналов, а именно Дискретное Преобразование Фурье (ДПФ), вычисляемое методом Быстрого Преобразования Фурье (БПФ), разработанные для дискретизованных (цифровых) сигналов, устанавливают однозначное соответствие между количеством дискретных отсчетов исследуемого сигнала и количеством синусоид в спектральном его представлении, а именно: для реальных сигналов количество синусоид вдвое меньше числа дискретных отсчетов сигнала во времени. В теоретическом же недискретном разложении Фурье установлен факт, что по мере увеличения частоты в спектре сигнала, в принципе, начиная с некоторого места, амплитуды спектральных компонент начинают очень быстро убывать, так что с этими компонентами практически можно не считаться.

Интерес к спектральному представлению является неотъемлемой частью процесса обработки сигнала с помощью эквалайзера. Такая обработка исключительно "на слух" соответствует невысокому уровню квалификации аудиооператора, поэтому высококачественные программные и аппаратные средства обработки звуковых сигналов непременно отображают состояние спектра звука в процессе эквализации.

Критические полосы спектра

10. Здесь следует иметь в виду психофизиологический параметр, который называется "критическая полоса" (иногда в литературе подобное понятие называют "частотная группа"). Суть этого понятия состоит в том, что слух аккумулирует эффект от воздействия близлежащих между собой синусоидальных компонент спектра ("близлежащих" как раз в пределах этой "критической" полосы частот), - аккумулирует в виде увеличения результирующей громкости от этих компонент. Это значит, что, если некоторое количество синусоид изначально было распределено в относительно широкой полосе частот, то они производили определенное тембровое впечатление и облада-

ли определенной громкостью. Если теперь эти синусоиды начать сближать между собой по частоте так, что они уместятся в пределах критической полосы, то, конечно же, тембровое впечатление от них изменится. Но изменится и общая производимая ими громкость - она станет больше.

Всегда трудно разделить субъективные впечатления от различных параметров звукового сигнала: например, вы сохраняете общую интенсивность сигнала, но перераспределяете его спектр в сторону более низких частот - и вам кажется, что сигнал стал громче. Вы, при неизменной интенсивности, сравниваете стереосигнал и его моно-трансформацию - и вам вновь кажется, что стереовариант звучит громче. Способность осознанно разделять различные свойства (параметры) звукового сигнала относится к высоким достоинствам профессиональных звукорежиссеров. Из этого замечания понятно, что уловить наличие критических полос, экспериментально манипулируя звуковым содержимым в домашних условиях, весьма трудно. Но лабораторные статистические эксперименты были проведены и дают результаты, отображенные на рис.2-12.

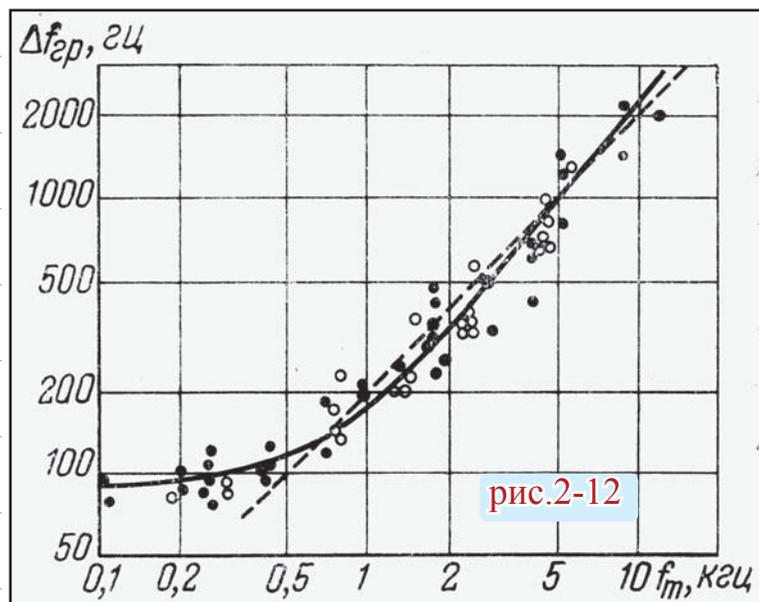


рис.2-12

На горизонтальной оси этого рисунка нанесены значения частоты, принадлежащие области слухового восприятия от 100 Гц до 20 кГц. Вертикальная ось указывает величину критической полосы, ассоциированной с точкой на горизонтальной оси частот (обозначение $\Delta f_{гр}$, где гр - группа, заимствовано из литературного источника иллюстрации).

11. Заметим очень важное обстоятельство. У критических полос не существует фиксированных границ. В любой точке f_1 оси частот области слухового восприятия мы можем указать (критическую) полосу частот $(f_1 - \Delta f_{гр} \dots f_1 + \Delta f_{гр})$, в пределах которой компоненты сигнала совместно формируют ощущение громкости. Но то же самое можно сказать и о любой другой точке с частотой f_2 , которая находится рядом с f_1 . Таким образом получается, что критические полосы, так сказать, "перекрываются". Это непросто осознать, если не обратить внимание на рис.2-13, который упрощенно иллюстрирует работу периферийной слуховой системы. Каждый элемент этой системы (чувствительный волосок приемной базилярной мембраны внутреннего уха) в соответствии с упрощенной теорией слухового восприятия Г.Гельмгольца является биологическим резонатором, реагируя на частоты, близкие к его собственной резонансной частоте. Частотные характеристики таких резонаторов перекрываются и синусоида некоторой частоты, как это видно из рис.2-13, одновременно участвует в формировании отклика целого ряда приемных волосков, которые передают импульсы возбуждения по нейронной цепи в центры обработки коры головного мозга.

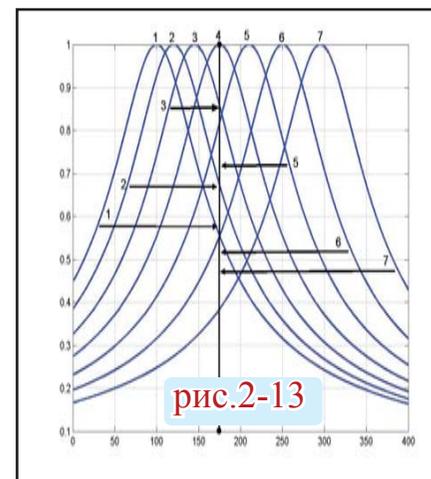


рис.2-13

12. Вообще-то, формальный подсчет результирующей громкости при воздействии звуковых сигналов является делом крайне неблагодарным, и, можно сказать, даже неосуществимым. Попытки дать методики такого подсчета не привели к ощутимым успехам. Пожалуй, успешным является лишь среднестатистическое увязывание субъективной громкости с уровнями громкости (через кривые равной громкости) и уровнями интенсивности, за что следует быть благодарными С.Стивенсу.

Что же касается критических полос, то наш труд не напрасен - знание этих величин позволяет грамотно и обоснованно выполнять некоторые операции по эквалайзингу звукового материала.

Позже мы повторно обратимся к понятию критических полос, а сейчас отметим, что выполняя тембровую коррекцию звукового материала в высоко-частотной области, которая привносит в звук оттенки типа "шипящий", "стеклянный", "скрипучий", "пронзительный" и т.п., мы, разумеется, подавляем отдельно расположенные энергетически ярко выраженные квазигармониче-

ские компоненты спектра. Но в случаях, когда подобные призвуки порождены не индивидуальным "всплеском", а широкой частотной областью, дающей ощутимый нежелательный вклад в итоговый тембр, подавлять эти призвуки следует, регулируя эквалайзером уровень сигнала в частотном интервале порядка критической полосы для этой области.

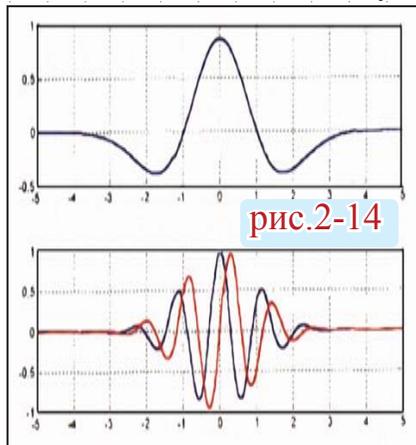
Из рис.2-12 видно, что ширина критической полосы на низких частотах до 500 Гц практически постоянна и составляет 100 Гц. Начиная с 500 Гц и выше ширина критической полосы составляет примерно 20% от точки, выбранной вами за ориентир. Так, например, желая воздействовать на громкость призвуков в частотной области с центром 12 кГц, следует манипулировать интервалом частот порядка 2,4 кГц.

Несинусоидальные представления

13. В заключение отметим, что существует, однако, формализованное разбиение всей частотной шкалы на "критические полосы" с фиксированными границами. Каждая такая "полоса" образует условную единицу, которую называют "барк" - всего в области слухового восприятия от 20 Гц до 16 кГц укладываются последовательно друг за другом 24 барка, границы между ними составляют (в герцах) величины: 100, 200, 300, 400, 510, 630, 770, 920, 1080, 1270, и т.д. Такая формальная техническая разбивка предпринята для сжатия аудиоданных в алгоритме MPEG и не соответствует, как уже было сказано, реальной психофизиологии формирования громкостных характеристик слуха.

Еще раз заметим, что критические полосы слуха будут упомянуты нами в дальнейшем при обсуждении проблем восприятия музыкальных звуков.

14. Вполне возможно, что у читателя уже возникал вопрос: "А что,



синусоида - действительно какая-то священная функция, единственная, способная сформировать любые мыслимые сигналы?"

Нет, это не так. Существует удивительно красивый раздел математики, которая сама по себе в целом представляется храмом, сравнимым по красоте, пожалуй, только с храмом музыки - эти два творения абстрактного разума, как я думаю, представляют собой вершину деятельности человеческого интеллекта. Этот раз-

дел называется Функциональный Анализ и он создает базовые основания для существования целого ряда "элементарных" типовых функций, из которых можно складывать произвольные сигналы.

Относительно недавно, например, были очень модны так называемые "вейвлеты", образцы которых показаны на рис.2-14. Произвольный сигнал можно сформировать из комбинации элементов такого рода, перемещая их влево\вправо вдоль образа сигнала и сжимая\растягивая в каждой точке расположения. (разумеется каждому состоянию элемента соответствует вычисляемый определенным образом коэффициент).

Но нетрудно сообразить, что таким элементарным составляющим соответствуют слуховые ощущения в виде либо острых сосредоточенных, либо размытых во времени щелчков, шлепков и т.п. Составить целостный слуховой образ сигнала из таких элементарных ощущений не представляется возможным. В то же время синусоидальные элементы в соответствии с акустическим законом Ома и слуховой теорией Гельмгольца как раз образуют удовлетворительную базу для понимания, исследования, документирования и операционной обработки именно звуковых сигналов.

Иные элементные базы, в частности вейвлеты, полиномиальные представления и прочие успешно применяются в задачах кодирования (сжатия) сигналов для эффективной их передачи по телекоммуникационным каналам связи.

3 Дискретные сигналы и их спектры.

Дискретизация сигналов

1. Практически все современные тракты записи звука в своих входных каскадах содержат т.н. "аналого-цифровой преобразователь (АЦП)". Непрерывный звуковой сигнал, например с выхода микрофона, поступая на АЦП, преобразуется в последовательность чисел, которые, как правило, аккумулируются в запоминающем устройстве, образуя в итоге "единицу хранения", называемую файлом.

Процесс преобразования непрерывного сигнала в числовую (иногда говорят "цифровую") последовательность характеризуется несколькими параметрами преобразования. Для сформированной числовой последовательности, которая отображает исходный сигнал, может быть вычислен спектр - именно это имелось в виду в предыдущей главе, когда упоминалось Дискретное Преобразование Фурье и Быстрое Преобразование Фурье. Такому спектру присущи свои важные особенности и параметры. Именно они и будут рассмотрены в настоящей главе.

Преобразование непрерывного (аналогового) сигнала в числовую (цифровую) последовательность удобно представить в виде двух шагов (рис.3-1).

2. Первый состоит в том, что сигнал "дискретизируют", то есть выбирают отдельные его значения через равные интервалы времени, которые на рис.3-1 названы "шаг дискретизации". Обозначим шаг дискретизации символом Δt , он измеряется в секундах или долях секунды (миллисекунды $1\text{мс}=0.001\text{с}$, микросекунды $1\text{мкс}=0.000001\text{с}$). Количество таких значений (чаще говорят "отсчетов", иногда "дискрет") за единицу времени составляет величину $f_d = 1/\Delta t$, и эту величину называют частотой дискретизации. Частоту дискретизации выражают в единицах "Герц" и она не обязана быть целым числом - если она, например, равна 10.2 Гц, то это всего лишь означает, что за 10 секунд будут выбраны 102 отсчета. Термин "Герц" относится не только к аудио- или видеоколебаниям. Герцы указывают количество любых событий, происходящих за 1 секунду, если эти события носят периодический (строго регулярный) характер. Если вы систематически постукиваете ногой о пол, то весьма вероятно, что делаете это с частотой 2 герца.

Так называемый "стандарт CD-Audio" указывает частоту $f_d=44100\text{ Гц}$, значение которой кроме свойств слуха (см. далее) было выбрано из соображений аудио и видео совместимости для основных видеостандартов. Распространены также значения $f_d=22050\text{ Гц}$ и $f_d=11025\text{ Гц}$, полученные из основного значения последовательным делением пополам. В так называемых DAT-устройствах

(Digital Audio Tape - цифровой "магнитофон") принято значение $f_d=48\text{ кГц}$, для которого существуют и кратные этой величине значения $f_d=96\text{ кГц}$ и $f_d=192\text{ кГц}$ (о них позже).

3. Итак, в результате дискретизации сигнала мы располагаем последовательностью чисел, представляющих отдельные систематически "выхваченные" значения этого сигнала. Но что представляют из себя сами такие числа? Невозможно поверить, например, что в числовой файл может попасть число 0.33333333... или что-то ему подобное, хотя математике известно, что практически все отсчеты сигнала должны иметь бесконечный "хвост" знаков в десятичном или любом ином изображении.

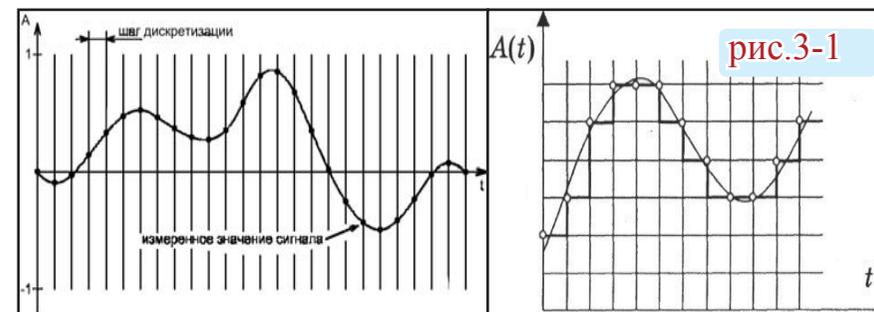
Структура хранения чисел в файле (не говоря уж о здравом смысле) подсказывает, что все "хвосты" в числах должны быть урезаны, начиная с некоторого заданного места.

Второй шаг преобразования аналогового сигнала в цифровую форму как раз состоит в том, что истинное значение каждого отсчета заменяется на наиболее близкое к нему значение из наперед заданной числовой шкалы, соответствующей принятому стандарту так же, как, например, мы в своей практике заменяем истинную температуру своего тела величиной с точностью до первого десятичного знака: 36.6; 37.2; и т.п.

Такую операцию называют квантованием (построение объекта с помощью дискретного набора элементов). Заметим, что дискретизацию можно назвать квантованием по времени, и раньше иногда так и говорили.

Операция квантования отображена на рис.3-1, где можно усмотреть замену значений гладкой кривой значениями, принадлежащими точкам, которые разбивают непрерывную вертикальную ось на шкалу с дискретными значениями.

Итак, операция превращения реальных (не ограниченных по точности) значений отсчетов сигнала в числа, принадлежащие шкале с ограниченным фиксированным набором значений называется квантованием.



Поскольку этим числам предстоит храниться в файле компьютерного типа, то и шкала для них выбирается, исходя из компьютерных форматов. Компьютерные форматы ориентированы на двоичные системы представления чисел. Наименьшая "единица хранения" числа в них называется байт и содержит 8 двоичных позиций (бит), в которых могут храниться целые числа из десятичного диапазона 0..255 (беззнаковые) или -128..+127 (числа со знаком)¹.

В стандарте CD-Audio размер числовой шкалы составляет два байта, то есть 16 двоичных позиций (бит), в которых могут храниться целые числа без знака из десятичного диапазона 0..65535, или со знаком из десятичного диапазона -32768..+32767. Именно в двухбайтовом формате (иногда называемом "машинным словом") целых чисел со знаком хранятся в большинстве случаев данные звуковых сигналов в компьютерных файлах.

Двухбайтовый формат хранения наиболее популярен сегодня, хотя существуют и более крупные форматы. Попутно заметим, что для алгоритмической обработки сигналов их целочисленные значения переводятся в формат с дробной частью, позволяющий относительно большое количество десятичных знаков "после запятой". А после выполнения процессором нужных вычислений с высокой точностью результаты с округлением вновь переводятся в целочисленный формат хранения.

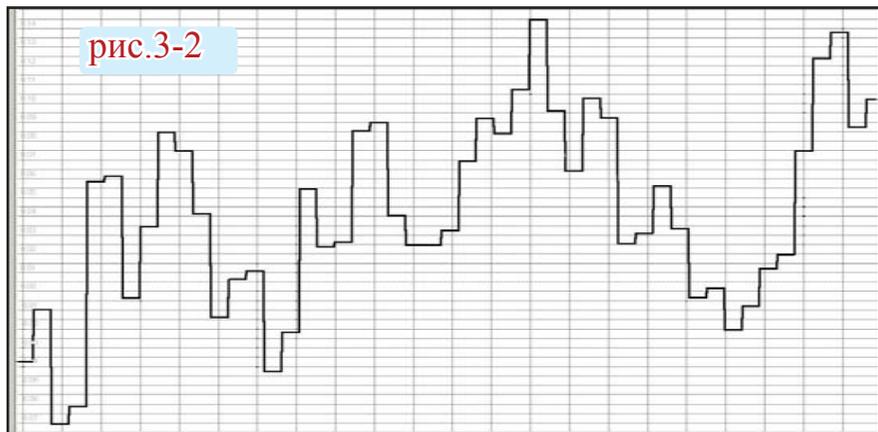


рис.3-2

4. На рис.3-2 показан в большом увеличении фрагмент звукового файла в окне одного из аудиоредакторов. На этом рисунке отдельные равноудаленные значения отсчетов объединены в ступенчатую структуру исключительно

¹ С деталями компьютерных форматов чисел можно ознакомиться, в частности, по популярным материалам, размещенным в интернете.

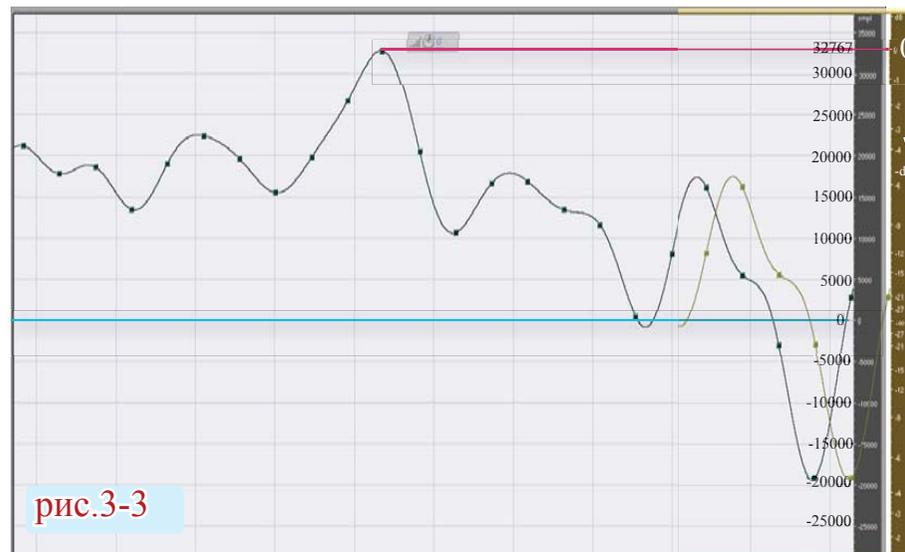


рис.3-3

но для визуального удобства зрителя. Рис.3-3 показывает другой фрагмент звукового файла в ином очевидном зрительном формате. На этом рисунке важна шкала чисел, которые указывают уровни для точек, изображающих дискретные отсчеты. Обратите внимание на максимальное значение 32767 - лучшие аудиоредакторы позволяют ручное изменение значений отдельных отсчетов файла путем перетаскивания мышью, в данном случае, точек строго вертикально вверх/вниз. Однако превысить величину 32767 при таком редактировании не удастся - это верхний доступный предел при двухбайтовом (в данном случае) хранении файла.

Еще один вид децибелов

5. Именно здесь уместно отметить смысл суффиксного обозначения децибел dBFS (dB Full Scale - децибелы в полной шкале). Это обозначение появилось исключительно благодаря внедрению цифрового представления сигналов, в котором количество байт, отведенных для хранения чисел, определяет максимально возможное хранимое число.

Мы уже говорили, что порог для отсчета уровней в децибелах выбирают произвольно в зависимости от характера измерения этих уровней. Для dB SPL в качестве порогового избрали значение 10^{-12} Вт/м² в районе минимально ощущаемой интенсивности области слухового восприятия. А для цифрового представления записанных сигналов порогом решили сделать максимально возможное в данной системе записи число, и такие децибелы обозначают как dBFS - в двухбайтовой системе dBFS отсчитывается относительно уже упо-

мянутой величины 32 767. Справа на рис.3-3 показана шкала в децибелах для этого же фрагмента звукового файла в том же аудиоредакторе. Видно, что 0 dB соответствует сигнальному отсчету, величина которого равна 32767 (когда величина порога очевидна из контекста, можно, например, вместо dBFS указывать просто dB). Понятно, что все иные возможные значения отсчетов меньше этой пороговой величины, а значит аргумент логарифмической функции меньше единицы. Поэтому все значения dBFS отрицательны, в чем нетрудно убедиться, работая с любым аудиоредактором (нашу картинку нужно увеличить, чтобы увидеть значения на шкале децибел).

6. Итоговое положение состоит в том, что приведенные нами два логических шага невозможны один без другого и обязаны составлять единое одновременно реализуемое действие. Невозможно выделить дискретные отсчеты сигнала, не придав им значение, пригодное для запоминания в цифровом (компьютерном) устройстве. А также невозможно квантовать некоторое значение реального сигнала, не указав его точное положение во времени. То есть дискретизация и квантование физически неразрывны настолько, что в литературе иногда употребляют лишь одно из этих слов, имея в виду при этом всю процедуру целиком. В последнее время предпочтение отдают слову "дискретизация" сигнала. В качестве характеристики такого процесса указывают сразу пару параметров - частоту дискретизации в герцах и формат хранения чисел в битах, например, для стандарта CD-Audio: 44100 Гц 16 бит.

Особенности спектров дискретных сигналов

7. Как было уже указано, дискретизация сигнала приводит к существенным особенностям его спектральных свойств. Обратимся к ним, но прежде заметим, что математический спектр аналогового сигнала для практика-звукорежиссера тоже обладает непривычными особенностями.

Дело в том, что спектральная функция образуется в результате именно математических интегральных преобразований исходного сигнала, связанных с именем Фурье, а они таковы, что картина спектральной функции оказывается распространенной в обе стороны от нуля по оси частот (см.рис.3-4). Напомним - на картинах мы изображаем только амплитудно-спектральную функцию. Справа от нуля герц располагается спектральная картина привычного вида, а слева от нуля располагается вычисленная по Фурье картина при "отрицательных" значениях частоты, и мы подчеркнули эти странные значения, зеркально отразив числа по оси частот в левой части рисунка.

Слова "отрицательные частоты" встречаются при обсуждении математического спектра сигналов и необходимо внести ясность в их употребление.

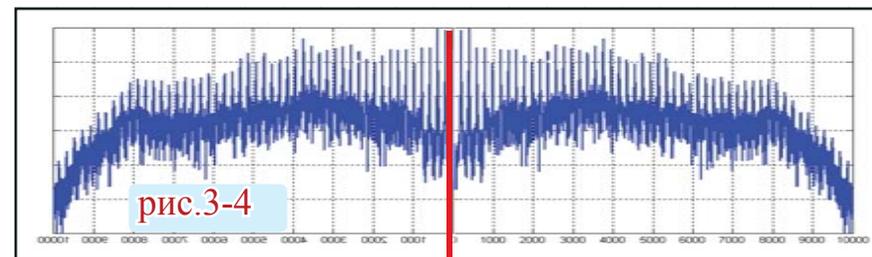


рис.3-4

Разумеется, никаких реальных отрицательных частот в природе не существует. Пять герц имеют в виду повторение события 5 раз в секунду, а величина -5 Гц не поддается осмыслению. Появление "отрицательных" частот связано с тем, что математические формулы для спектральной функции содержат определение гармонического ряда частот $f_k = k \cdot f_1$ как для положительных значений коэффициента $k = 1, 2, 3, \dots$, так и для отрицательных $k = -1, -2, -3, \dots$. Именно эти отрицательные коэффициенты, умноженные на положительное значение $f_1 = 1/T$ порождают значения, которые отображают в отрицательной части математической оси частот и называют отрицательными частотами.

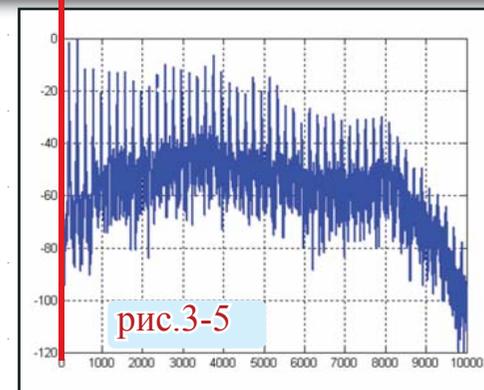


рис.3-5

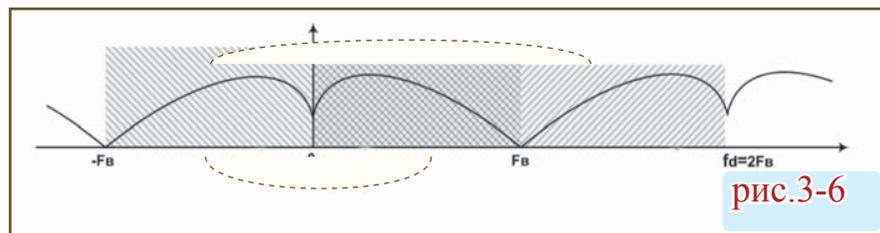
8. Для реально существующих сигналов² вычисленный по Фурье спектр (см.рис.3-4) обладает такими свойствами: амплитудный спектр в (левой) области отрицательных частот является точным зеркальным отображением спектра в правой части относительно вертикальной оси, проходящей через 0 Гц. Фазовый спектр (мы условились не изображать его на рисунках) слева тоже симметричен фазовому спектру справа, но все значения "левых" фаз при этом противоположны по знаку "правым" фазам.

Такие свойства позволяют обсуждать распределение энергии сигнала по частотам, ориентируясь только на спектральную картину в правой области ("положительных", рис.3-5) частот, этому существует математическое доказательство. Инженерные манипуляции аналоговым сигналом, разумеется, тоже реализуют для физически существующих параметров. Однако при любых

² в отличие от так называемых "аналитических сигналов" которые являются обобщенными математическими моделями

формальных математических операциях следует включать в рассмотрение именно строгое определение амплитудно-частотной спектральной функции на всей оси частот с их "отрицательными" и положительными значениями.

9. Дискретизация сигнала вносит специфику, прежде всего, вследствие строгого математического представления спектра. Сама математическая модель дискретизации сигнала несколько отличается от представленных ранее популярных иллюстраций и включает в себя так называемые δ -функции (дельта-функции), которые придают процессу дискретизации математическую строгость. Еще раз заметим, что результаты математических расчетов и преобразований находят хорошее подтверждение в практике их применения, в противном случае они были бы отвергнуты.



На рис.3-6 показан математический спектр (амплитудный) исходного аналогового сигнала, по форме напоминающий рис.3-4. Здесь он заштрихован, условно говоря "слева направо" - левый верх\правый низ. Дискретизация такого сигнала приводит к неожиданному результату - спектр дискретизованного сигнала не просто повторяет форму спектра исходного аналогового сигнала, а неограниченно повторяет (клонировует) исходную спектральную функцию в сторону положительных частот до бесконечности и в сторону отрицательных частот до минус бесконечности. Фрагменты таких клонов видны на рис.3-6, причем необходимая для наших пояснений часть ближайшего "справа" клона заштрихована "справа налево". Дуга внизу рисунка показывает две части исходного математического спектра, дуга вверху указывает, две совершенно идентичные половинки исходного спектра и его "ближайшего правого" клона в спектре дискретизованного сигнала³.

Элайзинг

10. Таким образом, спектр сигнала при его дискретизации становится периодическим. Доказано, что величина периода, то есть расстояние по оси

³ Если необходимо, можно математический спектр сигнала составить из двух половинок - правой половины исходного спектра (смешанная штриховка) и левой половины упомянутого правого клона (штриховка влево).

частот между идентичными значениями периодического спектра в точности равно частоте дискретизации fd . Это приводит к важным следствиям.

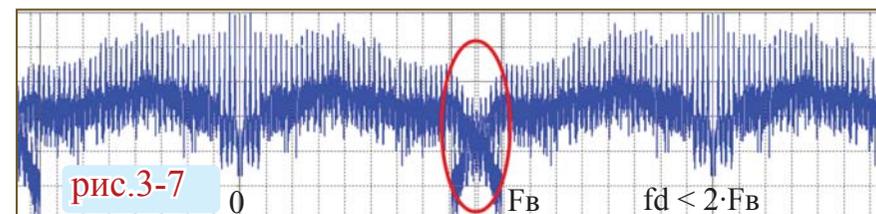
11. В случае, если физический спектр реального аналогового звукового сигнала целиком сосредоточен в интервале частот от 0 Гц до $Fв$ Гц, математический его спектр заключен в границах $-Fв .. Fв$ (Гц), как это показано на рис.3-6. Нетрудно видеть, что выбрав частоту дискретизации по правилу $fd = 2 \cdot Fв$, мы избегаем наложения копий спектров друг на друга - в этом случае спектральные копии точно стыкуются между собой. В случае $fd \geq 2 \cdot Fв$ тем более - между копиями образуется свободное пространство.

Итак, мы получили первое важное правило:

- если энергия сигнала ограничена верхней частотой $Fв$, то частота дискретизации такого сигнала должна составлять величину $fd = 2 \cdot Fв$.

Обсудим это правило.

12. Допустим, что мы выбрали частоту дискретизации сигнала больше необходимой величины: $fd > 2 \cdot Fв$. В этом случае размер звукового файла увеличивается пропорционально частоте fd . Это невыгодно, поэтому, даже создавая запасной промежуток между спектральными копиями, увеличивать частоту fd следует незначительно.



13. Выбрав частоту дискретизации меньше рекомендуемой величины ($fd < 2 \cdot Fв$), мы создаем ситуацию, при которой произойдет "перехлест" частей математического спектра. Рис.3-7 показывает такую ситуацию. Но физический спектр связан со своей математической моделью и дискретизованный сигнал, поступающий в файл для записи, обработки и последующего воспроизведения, просуммирует все синусоидальные компоненты в частотном диапазоне от 0 Гц до $Fв$ Гц, таким образом деформируя исходную звуковую картину в высокочастотной части спектра.

Такое нежелательное наложение спектров называется элайзинг (иногда говорят алиасинг - aliasing).

14. Очевидно, что для правильного выбора частоты дискретизации нужно быть уверенным, что за пределами верхней частоты $Fв = fd/2$ отсутствуют

какие-либо компоненты исходного дискретизируемого сигнала. С точки зрения психофизиологии слухового восприятия для верхней границы области слухового восприятия в 20 кГц частота дискретизации должна составлять 40 кГц, и стандартная величина 44100 Гц удовлетворяет этому условию. Но даже верхнюю частоту 20 кГц в сигнале еще нужно обеспечить, потому что реальное пространственное звуковое поле непременно содержит природные и промышленные звуки в значительно более широком слышимом для человека на аппаратурно регистрируемом диапазоне частот.

Это значит, что любое записывающее дискретное устройство обязано содержать входной низкочастотный фильтр сигнала, верхняя граница которого в два раза меньше установленной в устройстве частоты дискретизации.

Ресемплинг

15. Иногда возникает необходимость в так называемом ресемплинге (resampling, передискретизация). Слово "передискретизация" имеет также английский эквивалент oversampling, о котором мы скажем позже.

Ресемплинг это переформатирование файла с изменением частоты дискретизации f_d . Повышать частоту дискретизации в подавляющем большинстве случаев не имеет смысла - файл уже хранит взятые ранее отсчеты, и новых отсчетов взять неоткуда, можно лишь выполнять математические интерполяционные действия.

Понижение же частоты дискретизации бывает полезным достаточно часто - например, нет смысла хранить материал диалогового интервью в файле, ранее записанном с частотой 44100 Гц. В этом случае есть смысл снизить частоту дискретизации до 22050 Гц и аудиоредакторы позволяют сделать это, просто выбрасывая из файла все четные или нечетные отсчеты, размер файла таким образом уменьшается вдвое. Однако, если при этом не снизить до ~11 кГц верхнюю частоту F_v ранее отфильтрованного до ~22 кГц сигнала, то деформация звука за счет элайзинга может оказаться очень значительной - ведь записи интервью могли сопутствовать внешние шумы, лежащие в диапазоне 11 .. 22 кГц.

Все известные мне средства ресемплинга перед выполнением собственно этой операции автоматически выполняют предварительную дополнительную фильтрацию сигнала с новой верхней частотой $F_v = f_d/2$ для новой частоты дискретизации.

Иногда ресемплинг выполняется более сложным образом. Например при изменении частоты дискретизации с 48 кГц на 44100 Гц необходимо не выбрасывать отсчеты, а рассчитывать новые значения для других моментов времени на основе имеющихся. Это возможно применением так называемой

интерполяции: на рис.3-3 видно, что характер волны в значительной мере сохранится, если точки отсчетов соединить прямыми линиями. В этом простейшем случае на основе каждой пары точек нетрудно рассчитать новые значения, находящиеся между этой парой. Более точные расчеты используют более сложные методы, например так называемую "сплайн-интерполяцию".

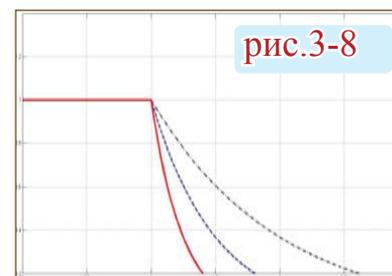
Теорема Котельникова

16. У дискретных отсчетов сигнала, взятых с частотой $f_d = 2 \cdot F_v$, есть замечательное свойство, сформулированное в теореме Котельникова⁴. Популярным языком можно сказать, что в соответствии с этой теоремой для сигнала, вся энергия которого сосредоточена в полосе от 0 Гц до F_v Гц, вся информация сосредоточена в дискретных отсчетах, взятых с частотой $f_d = 2 \cdot F_v$. Другими словами, все значения непрерывного сигнала, которые лежат между этими отсчетами, являются информационным балластом и имеют право быть отброшенными, потому что могут быть абсолютно точно восстановлены математическим способом.

С этим нетрудно согласиться, если вновь вернуться к рис.2-8. Обратите внимание насколько плавно изменяется суммарный сигнал при сложении первых низкочастотных компонент спектра. И только значительно более высокочастотные синусоиды формируют быстрые изменения верхушки и крутой фронт результирующего импульса. Запретив частоты выше F_v , мы блокируем, так сказать, "непредсказуемо быстрые" изменения сигнала, делая его в определенной мере "предсказуемым" и легко математически восстанавливаемым с помощью известных интерполяционных формул.

Таким образом, частота дискретизации звукового сигнала $f_d = 2 \cdot F_v$ обоснованно может быть признана оптимальной частотой для этой цели (дискретизации) в случае, когда за пределами F_v в спектре сигнала отсутствует энергия.

Оверсемплинг



17. Теперь очевидна важность качественной фильтрации сигнала фильтром нижних частот с верхней границей F_v . Лучше всего было бы резко отсечь в спектре входного сигнала все компоненты с частотами выше F_v , но реализация таких фильтров вызывает большие сомнения. Чем выше требова-

⁴ сформулированной в 1933 г. акад. В.А.Котельниковым, в западной литературе ее называют теоремой Шеннона, а частоту $f_d = 2 \cdot F_v$ частотой Найквиста.

ния к крутизне среза фильтра в области частоты F_v , тем более сложным и громоздким он становится. Инженерная мысль нашла достаточно качественное и коммерчески оправданное решение, которое связано с передискретизацией типа оверсемплинг (oversampling). Упрощенно эта мысль проиллюстрирована рис.3-8. Давайте отфильтруем сигнал, условно говоря, "простым и дешевым" фильтром с верхней частотой и частотой дискретизации, соответственно, в четыре раза больших номинальных значений F_v и f_d . Затем результат отфильтруем еще раз таким же простым фильтром, только с частотами не в четыре, а в два раза больше номинальных. И наконец в третий раз отфильтруем второй результат таким же относительно простым фильтром но теперь уже окончательно с номинальными значениями частот. Нетрудно видеть, что таким путем можно добиться, в целом, относительно высокого качества фильтрации не слишком дорогой ценой. Например, в точке, где нежелательная компонента входного процесса при первой фильтрации подавляется на уровне -20 dB, после третьей фильтрации она будет подавлена на уровне - 60 dB.

Эту совокупность операции и называют oversampling⁵, который, например для аппаратуры DAT с частотой дискретизации 48 кГц начинается со стартовой частоты $48 \cdot 2 \cdot 2 = 192$ кГц, известной многим читателям нормативной документацией⁶.

⁵ Иногда термину oversampling в массовых популярных материалах дают несколько иное толкование, но мы будем придерживаться указанного здесь варианта.

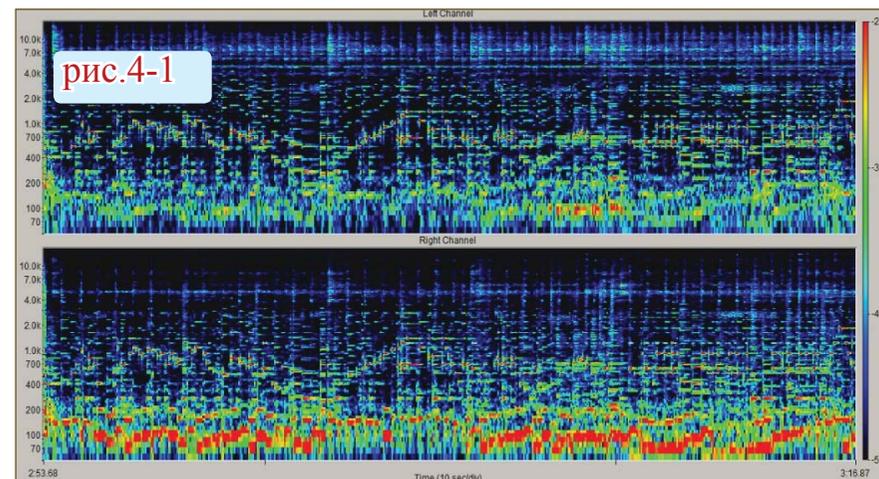
⁶ Рекламные упоминание столь высоких частот дискретизации в ряде случаев провоцируют загадочные рассуждения о каких-то мистических свойствах слухового восприятия, известных, вероятно, только менеджерам по продаже звуковой аппаратуры.

4 Спектрограммы

1. Нужно сказать, что спектр в его рассмотренной форме не является всеобъемлющим инструментом, способным ответить на все интересующие вопросы о структуре звукового сигнала. Для вычисления спектра необходимо выделить некоторый отрезок сигнала длительностью T секунд и быть при этом убежденным, что на протяжении этих T секунд спектральный состав сигнала не изменяется. При этом подробность описания спектра (по сути, точность отрисовки его графика, говорят еще "разрешающая способность" спектра) зависит от величины T - расстояние по частоте между ближайшими точками спектра равно $1/T$ Гц. Чем точнее мы хотим видеть форму распределения энергии по частоте, тем большее время звук должен оставаться совершенно неизменным. Часто ли возникает такая возможность?

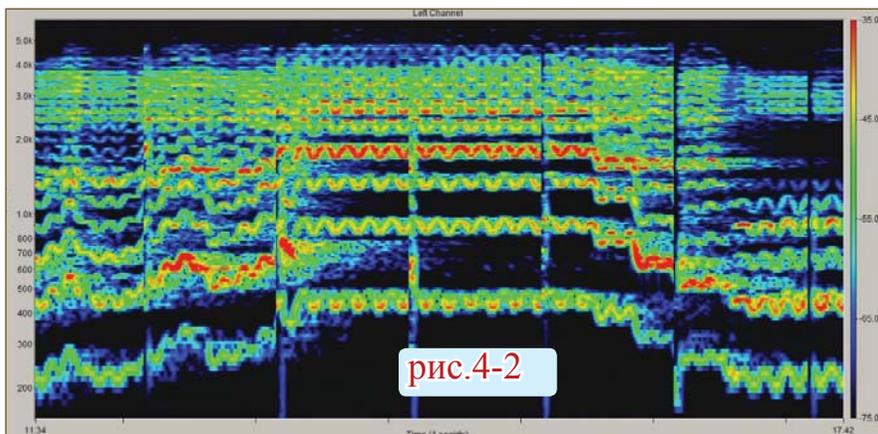
При академических исследованиях звуков можно попросить исполнителя удерживать постоянство звука в течение нескольких секунд. В реальной аудиопрактике такое постоянство - редкость. Звуковое полотно изменяется очень динамично и его изменения на протяжении нескольких секунд не позволяют ответить на вопрос - спектр чего именно мы вычисляем.

Спектральные наблюдения за динамикой тембровых и интонационных изменений звука реализуются с помощью спектрограмм, примеры которых приведены на рис.4-1,4-2.



На первом из рисунков отображен фрагмент спектрограммы стереозаписи ансамблевой пьесы "Take The A Train" с Оскаром Питерсоном. Видеодемонстрация документирует то, что не всегда отчетливо фиксирует

ухо: партия контрабаса практически целиком сосредоточена в одном из каналов, фортепиано расположено строго в центре панорамы. За пределами этого рисунка содержатся детали, которые позволяют судить о других подробностях тембрового баланса в каналах. Здесь можно было бы видеть даже небольшое рассогласование во времени отдельных инструментов, если оно имело место.



Следующий рисунок (рис.4-2) - это фрагмент арии, исполняемой Лучиано Паваротти. Картина не только дает возможность видеть распределение гармоник в голосе певца, включая верхнюю певческую форманту (красная область посередине рисунка, о певческих формантах см. в главе 8), но и судить о его интонационной стабильности, о характере ведения вибрато (волнистые треки обертонов, о вибрато см. в главе 6) - вибрато хорошо заметно даже в коротких промежуточных нотах.

Работа со спектрограммами требует определенных навыков. По своей сути они представляют непрерывную серию относительно коротких спектральных измерений, отображенных в трехмерном пространстве, в котором на плоскости рисунка расположены оси времени и частоты, а спектральные амплитуды сравнительно грубо отображаются цветом в соответствии со шкалой, показанной здесь узкой полоской справа.

2. На практике спектрограммы разворачиваются (движутся) на экране в реальном времени в сопровождении собственно звукового сигнала. Как правило, для достижения нужного качества картины приходится подбирать параметры спектрального анализа методом проб и ошибок, и количество ошибок, разумеется, резко снижается с приобретением опыта.

Основные параметры спектрограммы таковы.

- Размер T фрейма (кадра во времени) для вычисления "мгновенного" спектра. Этот размер должен обеспечивать компромисс - не быть слишком малым, чтобы разрешение по частоте в спектре позволяло видеть детали распределения энергии по частоте, и не быть слишком большим, чтобы не захватывать существенные изменения тембровой картины во времени.

Для дискретного сигнала величина $T = N \cdot \Delta t = N / f_d$, где N - количество дискретных отсчетов, размещенных во фрейме. Особенность вычисления Быстрого Преобразования Фурье (БПФ, англ. FFT) состоит в том, что N должно представлять собой степень двойки из ряда 2, 4, 8, ..., 512, 1024, 2048, 4096, 8192, 16384, ... Для частоты дискретизации 44100 Гц это соответствует длительностям фрейма: 45, 90, 180 мксек ... 11, 22, 46, 93, 186, 372 мсек ...

3. Очевидно, что первые из приведенных длительностей фрейма совершенно не подходят для анализа звукового сигнала. Однако, если ориентироваться, например на исполнение музыкальных пьес, то фрейм приблизительно в 100 мсек ($N = 4096$) соответствует 10 нотам в секунду (быстрый свинговый джаз), приблизительно 200 мсек ($N = 8192$) - 5 нот в секунду (подвижная пьеса) ... Таким образом, для частоты $f_d = 44100$ Гц вполне можно выбрать размер фрейма (в отсчетах), равным 4096 или 8192. Зачастую так и делают. Уменьшив в два раза величину частоты до $f_d = 22050$ Гц, тех же параметров можно достичь, соответственно, при ($N = 2048$ и $N = 4096$).

А что же изменится при таком уменьшении частоты дискретизации? Во-первых, уменьшится в два раза размер звукового файла, и уменьшится в два раза верхняя частота сигнала $F_v = f_d / 2$ - с 22050 Гц до 11025 Гц (с автоматической перефильтрацией), что тоже может быть удобно при отрисовке спектрограммы!

- Однако заметим, что столь подробное отображение процесса вибрато на рис.4-2 невозможно при указанных размерах фрейма без дополнительных условий. Частота вибрато обычно близка к 5 Гц, и для того, чтобы увидеть колебание с такой частотой, нужно сформировать не менее десяти точек на период колебания, что соответствует длительности фрейма не более $1 / (5 \cdot 10) = 20$ мсек. Разрешающая способность

1 Я "нагружаю" Вас этими техническими подробностями, потому что открыв аудиоредактор в режиме спектрограммы, Вы увидите эти числа среди параметров настройки и будет полезно понимать их значение.

спектра при этом составит величину более 50 Гц, что дает неприемлемо низкое качество спектральной картины²

Решением проблемы является реализация спектрограммы с перекрытием (англ. overlapping) фреймов. Для такой схемы выбор размера фрейма с $N = 8192$ (разрешающая способность спектра 5 Гц) и величина перекрытия фреймов в процентах порядка 90%, как это очевидно из рис.4-2, обеспечивает хорошее решение поставленной задачи.

4. Рассмотренные здесь способы обработки звуковых сигналов не являются исчерпывающими. Кроме них известны так называемые корреляционные методы обработки, с помощью которых можно ответить на некоторые дополнительные вопросы. Специалисты инженерного профиля могут также путем моделирования сигналов получать очень интересные и практически полезные результаты, расширяющие мастерство звукорежиссеров. Для нас важно, чтобы в случаях экспертной оценки свойств звукового сигнала звукорежиссер мог, по крайней мере, адекватным образом сформулировать свой вопрос аудиоспециалисту, обладающему специальными знаниями и программным обеспечением, и верно понимать полученный от него ответ и документальные результаты.

Но основные контрольные средства, с которыми работают как искусственные специалисты, так и звукорежиссеры без специальной подготовки - это спектральные средства. Развертка сигнала во времени в аудиоредакторе позволяет судить о степени его компрессии, спектр стабильных в тембровом отношении участков звука позволяет судить о распределении энергии сигнала по оси частот, а спектрограмма позволяет, в определенной мере, делать выводы о развертывании звуковой картины во времени.

Мы надеемся, что изложенный здесь материал приблизит звукорежиссера к более сознательному применению этих средств.

² Это хороший термин - способность, разрешающая видеть отдельно стоящие спектральные компоненты - между ними ничего сигнального анализ не показывает. Шаг в 50 Гц в этом смысле для спектра звукового сигнала неудовлетворителен.

5 Музыкальный строй. Темперация.

Интервалы, порожденные обертонами

1. Среди возможных первоисточников музыкальных инструментов вероятным прообразом струнных инструментов считают охотничий лук. Звук спущенной тетивы, зависящий от размеров лука, материала тетивы и ее натяжения, вполне мог породить интонационные эксперименты, которые привели к организации специального звучащего устройства, свободного от охотничьих задач.

Во всяком случае можно счесть обоснованным выбор звучащей струны в качестве базового элемента для описания элементарных свойств музыкального звука.

Исследование природы такого звука показывает наличие в его составе гармонического ряда колебаний, весьма близких к синусоидальным (тем более синусоидальных, чем строже выполняются технические и технологические требования к струне).

Для основного тона с частотой f_0 гармонический ряд содержит составляющие с частотами $2 \cdot f_0, 3 \cdot f_0, 4 \cdot f_0, 5 \cdot f_0, 6 \cdot f_0, 7 \cdot f_0, \dots$

Последовательные интервалы между гармониками соответствуют отношению частот: $2, 3/2, 4/3, 5/4, 6/5, 7/6, \dots$

Природа человеческого восприятия устроена так, что автономные звуки с некоторыми из этих соотношений частот воспринимаются слухом, как элементы интонационного звукоряда, и образуют между собой музыкальные интервалы:

2 - октава, т.е. $2 \cdot f_0$;

$3/2$ - квинта (ч.5) через октаву, приведенная к первой октаве $1.5 \cdot f_0$;

$4/3$ - кварта (ч.5) перед второй октавой, в первой октаве $1.333 \dots \cdot f_0$;

$5/4$ - большая терция (б.3), в первой октаве $1.25 \cdot f_0$;

$6/5$ - малая терция (м.3), в первой октаве $1.2 \cdot f_0$;

$7/6$ - ?, в первой октаве $1.17 \cdot f_0$;

$8/7$ - ?, в первой октаве $1.143 \cdot f_0$;

$9/8$ - большая секунда (б.2), в первой октаве $1.125 \cdot f_0$;

Интервалы между последующими гармониками сигнала, постепенно сужаются от большой секунды к малой и далее, образуя все более диссонирующие созвучия. Группа интервалов со знаком вопроса воспринимается музыкально воспитанным слушателем, как интервалы промежуточные между м.3 и б.2

Заметим, что по преданиям лира Орфея в наиболее архаичном варианте содержала четыре струны, которые давали от "начального" тона интервалы кварты, квинты и октавы. Это соответствовало речитативному стилю художественной речи, близкому к торжественной разговорной манере, которая и сегодня в значительной мере опирается на интервалы кварты и квинты.

Указанные выше интервалы дают возможность построить более широкую интонационную шкалу - обращение терции дает сексту, обращение секунды - септиму. Звукоряд, построенный на интервалах между обертонами музыкального звука, называют натуральным звукорядом. Его возможности весьма ограничены как при использовании струнного инструмента, так и духового типа, например, фанфар.

2. Развитие инструментальной музыкальной базы, в частности, появление клавишных инструментов с фиксированной нотной шкалой - органов, клавесинов и т.п., остро поставило вопрос о предварительной настройке этих инструментов, что породило термин "темперация", то есть построение интервальной шкалы, способной обеспечить выполнение музыкальных художественных задач, с помощью расчетов.

Обилие и сложность рассуждений при таких расчетах не позволяют привести их обзор. Однако проблему можно ощутить путем относительно простого рассмотрения, которое мы сейчас выполним.

Очевидно, что наиболее совершенным в интонационном смысле является интервал октава. Звуки, существующие в таком соотношении, сливаются воедино практически идеально. Следующим "заслуживающим абсолютного

Т1	квинтовый круг		в первой октаве	
	квинтовые шаги	в десятичном виде	приведенные	ноты
1	1	1.0	1	до
2	3/2	1.5	1.5	соль
3	9/4	2.25	1.125	ре
4	27/8	3.375	1.6875	ля
5	81/16	5.0625	1.265625	ми
6	243/32	7.59375	1.8984375	си
7	729/64	11.390625	1.423828125	фа#
8	2187/128	17.0859375	1.06787109375	до#
9	6561/256	25.62890625	1.60180664062	соль#
10	19683/512	38.443359375	1.20135498046	ре#
11	59049/1024	57.6650390625	1.8020324707	ля#
12	177147/2048	86.4975585937	1.35152435302	ми#(фа)
13	531441/4096	129.74633789	1.01364326476	си#(до)

"доверия" является интервал квинта - его можно получить, зажав жестко закрепленную струну на 2/3 от основания. Давайте на этом и остановимся.

Квинтовый круг

3. Если последовательно шаг за шагом строить квинты от стартовой ноты, назовем ее "До", то получим так называемый "квинтовый круг" (иногда говорят квинтовая спираль, имея в виду, что Си# не дает точного октавного соотношения 2:1 с До):

До-Соль-Ре-Ля-Ми-Си-Фа#-До#-Соль#-Ре#-Ля#-Ми#(Фа)-Си#(До).

В таблице Т1 приведены расчетные данные относительных частот для квинтового круга Эти частоты, образованные для каждого k-го шага по формуле $(3/2)^{(k-1)}$ (колонки 2,3), далее мы приводим к первой октаве, а именно - значение частоты делим на 2 (то есть заменяем октавным эквивалентом в более низком регистре) до тех пор, пока это значение не окажется в промежутке от 1 до 2 (условно "первая" октава, колонка 4).

Названия нот (колонка 5) здесь тоже условны, причем музыканты понимают, почему, например, вместо Фа указана нота с названием Ми#, а для звукорежиссера (при всем уважении) это неважно. Важно, что циклический проход квинтами по оси частот на 13 шаге приводит почти (с небольшим смещением) к стартовой ноте До (напомним, приведенной к первой октаве)

4. Однако квинтовыми шагами можно смещаться не только в сторону повышения частоты, но и в сторону ее понижения. Квинтовый круг в этом случае имеет вид

Здесь повторяется история с двойным названием нот, которое мы соблюдаем, но не разъясняем, поскольку это относится к музыкально-теоретическим сведениям.

До-Фа-Сиб-Миб-Ляб-Реб-Сольб-Доб(Си)-Фаб(Ми)-Сибб(Ля)-Мибб(Ре)-Лябб(Соль)-Ребб(До)

Общие результаты расчетов, систематизированные в виде хроматической гаммы, приведены в таблице Т2.

Из таблицы видно, что квинтовые шаги в разных направлениях оси частот приводят к различающимся значениям октавного интервала, который в обоих случаях не составляет точно удвоенной стартовой частоты. Кроме того, ни одна из двенадцати интонируемых таким образом ступеней не совпадает попарно по частоте.

Темперация, построенная на основе квинтового круга, называется Пифагорейской и, очевидно, что она не является удовлетворительной. Музыкальное мелодическое движение вверх по частоте не вполне совмещается с движением вниз из-за расхождений в частоте интонируемых позиций с

до 1 = 1.000	до 1 = 1.000
до# = 1.068	реб = 1.053
ре = 1.125	миbb(ре) = 1.110
ре# = 1.201	миb = 1.185
ми = 1.266	фаb(ми) = 1.249
ми#(фа) = 1.352	фа = 1.333
фа# = 1.424	сольb = 1.405
соль = 1.500	ляbb(соль) = 1.480
соль# = 1.602	ляb = 1.580
ля = 1.688	сиbb(ля) = 1.665
ля# = 1.802	сиb = 1.778
си = 1.898	доб(си) = 1.873
си#(до) = 2.014	ребb(до) = 1.973

одинаковым хроматическим номером. Это препятствует однозначному построению аккордовых последовательностей и модуляции тональности в процессе исполнения музыкальных произведений.

Такие обстоятельства породили иные методы поиска интервальных соотношений, способных удовлетворить профессиональный музыкальный слух. Итогом длительных поисков явилась компромиссная равномерно-

темперированная шкала, которая базируется на двух постулатах, постулат октав - шкала вместе с каждой частотой f_1 должна содержать частоты $2 \cdot f_1$ и $f_1/2$, а также постулат транспонирования - одноименные интервалы, сыгранные от любой ноты, должны быть интонационно идентичны.

Равномерные темперации

5. В принципе, равномерно-темперированная шкала не обязательно должна быть 12-ступенной. Реализация указанных постулатов возможна при любом количестве ступеней в октаве, и нетрудно доказать, что для m -ступенной равномерной шкалы каждый хроматический k -й шаг должен приводить к замене частоты f_k на $f_k \cdot \sqrt[m]{2}$

В соответствии с законом Вебера-Фехнера приращение частоты тона, необходимое для ощущения некоторого заданного интервала "Г" пропорционально не только приращению частоты, но и начальной частоте f_0 "опорного" тона, от которого интервал отсчитывается. Равномерно-темперированная шкала соответствует этому принципу, границы такого соответствия мы обсудим позже.

Таким образом ощущение любого музыкального интервала "Г" формируется в зависимости от соотношения (f_1 / f_0) . Поскольку октава соответствует удвоению частоты, то основание логарифмической (по Веберу-Фехнеру) шкалы высоты тона удобно выбрать равным двум.

Что касается выбора величины m , то европейская музыкальная практика приучила нас к 12-ступенной октаве, хотя в некоторых культурах известны

17 и 24-ступенные темперации. Ряд теоретических работ предлагал 29, 72 и даже 84-ступенные равномерные темперации. Выбор количества ступеней связан с тем, чтобы обеспечить в такой темперации хорошее приближение к первым интервалам натурального звукоряда, которые наиболее естественно воспринимаются слухом, прежде всего к квинте. Очевидно, что чем на большее число ступеней будет разбита октава, тем больше шансов, что какая-то из ступеней окажется близка к чистой квинте.

4. Наши собственные расчеты на эту тему приводят к выводу, что наиболее качественную квинту дают всего несколько равномерных темпераций, из которых призовые места занимают 53-ступенная, 41-ступенная и 12-ступенная (84-ступенная не рассматривалась).



рис.5-1

На рис.5-1 показана реконструкция так называемого органа Бозанкетта с 53 равномерно распределенными ступенями в октаве, клавиатура здесь содержит примерно полторы октавы. Существует свидетельство, что в принципе играть на таком инструменте возможно, однако сколько-нибудь серьезное исполнительство в данном случае вызывает большие сомнения.

Очевидно, что при сохранении относительно высокого качества интервалов наиболее выгодным из перечисленных с технической точки зрения

является 12-ступенный строй. Практически соответствуя по количеству ступеней в октаве квинтовому кругу, он обеспечивает высокую степень приближения к основным интервалам, интонируемым по гармоникам музыкального звука - квинте и большой терции. Заметим, что именно 12 ступеней, а не 11, или 13; или иное небольшое количество равномерно расположенных в октаве ступеней обеспечивают наиболее относительно высокое качество соблюдения интервалов. Это доказано нашими исследованиями, из которых следует, что именно 12-ступенная равномерная темперация обеспечивает наибольшее приближение "равномерной" квинты к чистой квинте с соотношением $3/2$, проигрывая в этом отношении только 53-ступенной темперации, технических возможностей которой мы уже коснулись. Для 12-ступенной шкалы хроматический шаг по частота равен $\sqrt[12]{2} = 1.05946 \approx 1.06$.

В равномерно-темперированной шкале, естественно, отсутствуют расхождения между нотами, находящимися на одинаковых хроматических позициях, например, до# и реb - в этом случае ноту с относительной частотой, равной 1.06 можно назвать любым из этих двух названий в зависимости от избранной тональности.

Такое совпадение частоты для разноименных нот называется энгармонизмом, а сами ноты с разными названиями называются энгармонически равными.

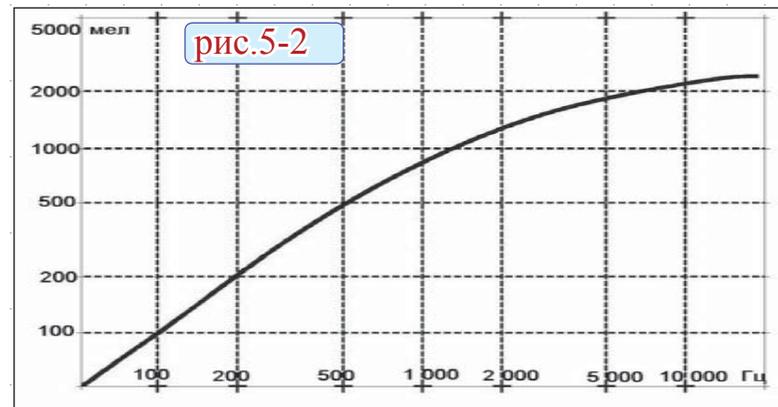
5. В таблице Т3 приведен весь хроматический ряд нот с названиями и значениями частоты от "правого" квинтового круга (в сторону более высоких частот). Справа в колонке 3 приведены значения относительных частот в пределах октавы для равномерно-темперированной шкалы, еще правее - отклонения этих частот от "пифагорейских" значений. Как видно из таблицы, все ступени гаммы имеют некоторые отклонения от этих значений, причем для квинты и терции эти отклонения - наименьшие. Сами по себе эти отклонения, вероятно, указывают на несовершенство равномерной шкалы не большее, чем отсутствие энгармонизма в пифагорейской, но особо чувствительные к слуховым несовершенствам музыкальные специалисты иногда высказывают неудовлетворение качеством созвучий в равномерной шкале.

Нужно, однако, признать, что развитие музыкальной мысли, опирающееся на модуляции, политональность, сложные гармонические последовательности не имеет иной реальной альтернативы кроме системы равномерной темперации для инструментов с фиксированной настройкой.

нога	Квинтовый круг	12-ступенная шкала	отклонение
до 1	1	1	
до#	1.068	1.05946309435930 \approx 1.059	-0.009
ре	1.125	1.12246204830937 \approx 1.122	-0.003
ре#	1.201	1.18920711500272 \approx 1.189	-0.012
ми	1.266	1.25992104989487 \approx 1260	-0.006
фа	1.352	1.33483985417003 \approx 1.335	-0.017
фа#	1.424	1.41421356237310 \approx 1.414	-0.010
соль	1.5	1.49830707687668 \approx 1.498	-0.002
соль#	1.602	1.58740105196820 \approx 1.587	-0.015
ля	1.688	1.68179283050743 \approx 1.682	-0.006
ля#	1.802	1.78179743628068 \approx 1.782	-0.020
си	1.898	1.88774862536339 \approx 1.888	-0.010
до 2	1.014	2.00000000000000	

К вопросу о точности интонирования

6. Касаясь точности соблюдения интонационного музыкального строя, нельзя обойти психофизиологическое свойство слуха, в соответствии с которым ощущение интервала октавы при удвоении частоты простого тона не строго соблюдается во всем слышимом диапазоне частот [Ак, 114].



На рис.5-2 показана зависимость интонационной интервальной единицы "мел" от частоты. Можно сказать, что мел - это единица высоты тона, но понимать эту единицу удобнее так, что именно удвоение числа мел соответствует субъективно воспринимаемому ощущению октавы. Видно, что до 500 Гц удвоение частоты тона в герцах соответствует удвоению количества мел. Выше границы в 500 Гц (вторая фортепианная октава) ощущение октавного интервала при удвоении частоты в Гц незначительно деформируется, а в третьей фортепианной октаве для среднестатистического слушателя с музыкальным слухом удвоения частоты уже недостаточно для полноценного ощущения октавного интервала. Полагают, что в еще более высоком регистре, где, собственно, трудно оценивать музыкальные интервалы вообще, эти интервалы, выраженные в герцах, существенно "расширяются".

Музыкальная практика и со стороны композиторов и со стороны музыкантов интуитивно считается с этими обстоятельствами, не нагружая высокие регистры мелодико-гармоническими функциями, а реализуя в них высокотехничные но не особо требовательные к точности интонирования исполнительские задачи.

Разумеется, для высоких обертонов звука такая зависимость (рис.5-2) повышает степень их диссонансности, но можно согласиться с тем, что для консонансных созвучий чистота интервалов играет большую роль, а для подобного рода диссонансов в высоком регистре чистота несущественна.

6 Музыкальные звуки

Сигналы и шумы

1. Простейшее разделение звуков на виды позволяет видеть в звуках "сигналы" и "шумы". Обычный принцип их разделения состоит в том, что сигналам приписывают наличие некоторой "информации", а шумы препятствуют получению и обработке этой информации и в таком контексте являются синонимом слова "помехи".

Подход звукорежиссера не отрицает терминов "сигнал" и "шум", однако смещает акценты в этих понятиях, помещая слово "шум" в два различных контекста. Тот шум, который "помеха", относится к паразитным звукам, порожденным аппаратурой записи и внешними нежелательными источниками. Звуки же, которые часто тоже называют шумами в бытовой лексике, могут относиться к окружающей человека внешней среде (транспортные, производственные, природные), и иногда (в особенности в кино и телевидении) представляют творческий интерес для звукорежиссера, являются предметом его целенаправленных поисков, и в таком случае играют художественную роль "сигналов".

Упомянутые шумы не поддаются классификации. Можно, конечно разделить их на стуки, шорохи, шипения, свисты и т.д. В библиотеках шумов для звукорежиссеров зачастую так и делают, но такое разделение является волюнтаристским, произвольным и почти не позволяет ввести какие-либо классификационные параметры для описания этих звуков.

Ориентируясь на слуховые способности человека, представляется разумным в интересах звукорежиссуры различать между собой две иные категории, называя их "интонированные звуки" и "шумы", то есть неинтонированные. Интонированные, это те звуки, которым можно приписать некоторую интонационную высоту тона. Другими словами, если человек с развитым музыкальным слухом может спеть или взять на музыкальном инструменте ноту, эквивалентную по высоте тона слышимому звуку произвольной природы, то этот произвольный звук интонируемый (или интонированный). Все остальные звуки, с которыми интонационных ассоциаций не возникает, это шумы, и им можно давать произвольные словесные описания. Еще раз - если для произвольного звука музыкальным слухом может быть опознана некоторая музыкальная "высота тона", то он (для этого слуха) интонированный.

Естественно, что интонируемыми звуками, в принципе, можно исполнить мелодию. Наличие интонационного свойства позволяет использовать нотную запись для воспроизведения таких звуков, и это является качеством, принципиально отличающим их от шумов.

Структура музыкальных звуков

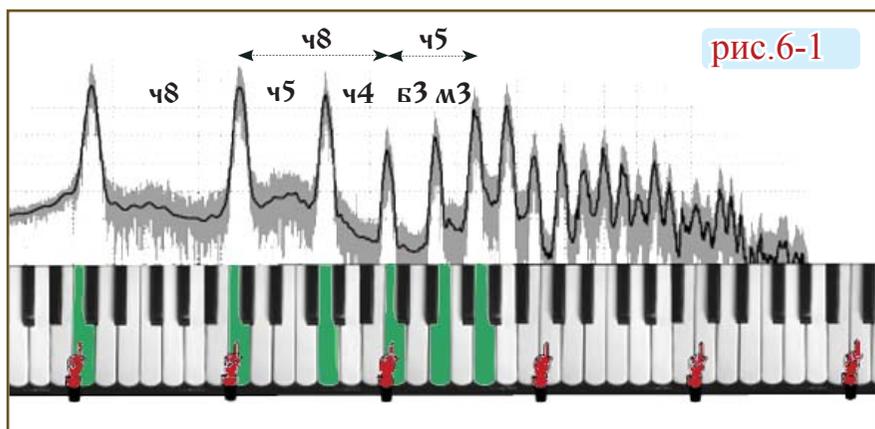
2. Среди интонированных звуков мы особо выделяем музыкальные звуки. Их характерной особенностью является обертоновая структура. Музыкальный звук содержит "основной тон" - квазисинусоидальный сигнал с наименьшей в этой структуре частотой f_0 , которая создает интонационную основу звучания¹. Структура музыкального звука содержит ряд "обертонов" - квазисинусоидальных сигналов, частоты которых образуют гармонический ряд $2f_0, 3f_0, 4f_0, 5f_0$, и т.д.

Слово "квазисинусоидальный" означает следующее. Музыкальный звук может быть образован и строго синусоидальными сигналами, организованными в указанную структуру. Но такой звук с эмоциональной точки зрения окажется выхолощенным, "мертвым", и может иметь ценность только в экспериментальных исследованиях. Реальные музыкальные звуки содержат элементы обертоновой структуры, по форме близкие к синусоидальным колебаниям, но подверженные некоторой нестабильности и в частоте и по амплитуде - именно "подвижность" параметров колебаний при сохранении их интонационной стабильности создает благоприятный психологический эффект. Заметим, что небольшие периодические вариации амплитуды и частоты звука (далее мы назовем их "вibrато") не аннулируют ощущения интонационной стабильности, то есть соответствия звука одной и той же ноте.

Количество обертонов может варьироваться в довольно широких пределах - от нескольких единиц до нескольких десятков. Их количество и соотношение между их амплитудами определяют качество звука, называемое тембром. Здесь заметим, что встречаются высказывания, в которых под словом "тембр" подразумевают еще и наличие дополнительных свойств звука таких как, атака, вibrато, реверберация, биения, влияние кривых равной громкости и даже расстояние до источника звука. В таких высказываниях смешаны совершенно разнородные обстоятельства, которые, с одной стороны, влияют на слуховое впечатление от звука, а с другой стороны - имеют самостоятельные названия и регулируются самостоятельными параметрами. Мы будем понимать под словом "тембр" исключительно психофизиологическое свойство, порожденное спектральной характеристикой конкретного интонационно стабильного звука, предоставленного для записи и измерений.

Роль обертонов музыкального звука в его тембровом ощущении может быть проиллюстрирована рис.6-1. Мы уже говорили, что если условно назвать основной тон "До1", первые пять обертонов соответствуют нотам "до2", "соль2", "до3", "ми3", "соль3", где цифра указывает номер октавы, отсчитываемой от "До1".

1 Физически основной тон с этой частотой может в структуре и вовсе отсутствовать, но оставшаяся часть структуры музыкального звука (см. далее) создает субъективное ощущение его наличия. В этом случае говорят о кажущемся (виртуальном) или фантомном основном тоне.



Иногда обертоны называют гармониками² из-за указанного выше соотношения их частот. Номера гармоник соответствуют числовому множителю при основной частоте; первая, вторая, третья и т.д. Нумерация обертонов иная - основной тон так и называют, а далее следуют: первый обертон (вторая гармоника), второй обертон (третья гармоника) и т.д.

Очевидно что первые шесть гармоник музыкального звука образуют мажорное трезвучие в очень широком расположении. Последующие несколько обертонов дают в звучание ноты тембровый вклад в виде все более деформированных ("сужающихся") малых терций вплоть до 9..10 гармоники, начиная с которых в тембре участвуют диссонансные секунды. Гармоники, существенно далеко отстоящие от начальной частоты (15-20-е и далее) дополняют тембр микротональными созвучиями так, что тембр приобретает оттенки от пронзительного до остросвистящего, скрипящего и т.п. Все дело здесь в энергетических соотношениях между различными участками спектра и, разумеется, в частотном диапазоне (регистре), в котором располагается нота.

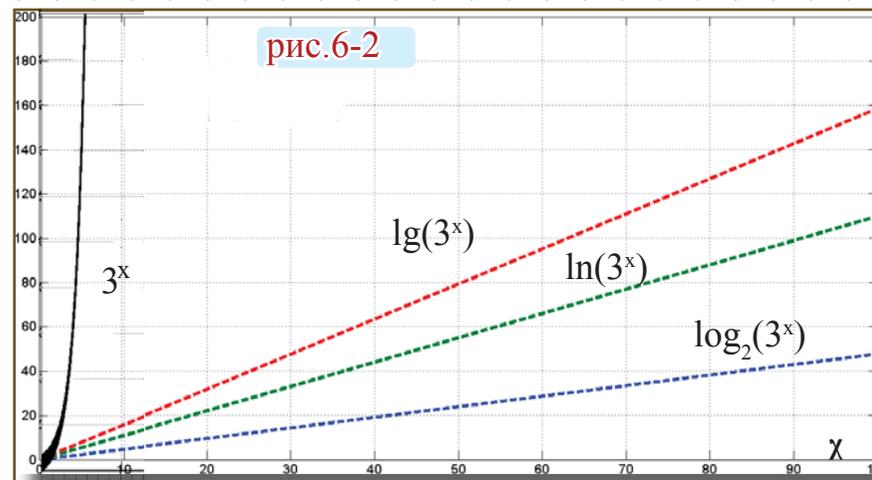
Так, например, для баритонального диапазона звук со спектром, показанным на рис.6-1 будет восприниматься более благоприятно, чем в высоком теноровом диапазоне, где даже обертоны, принадлежащие к первой десятке, попадают в область "писков" и "свистов", создавая особо неприятные созвучия..

Понимание роли обертонов в формировании слуховых ощущений безусловно помогает проводить осознанный эквалайзинг звукового материала для достижения поставленных художественных целей.

² Сходное название встречается и в связи с тем, что одиночные колебания синусоидального типа с произвольной начальной фазой иногда называют гармоническими колебаниями.

Шкала частот в изображении музыкального звука

3. Теперь понятно как опознать индивидуальный музыкальный звук на спектральной картине - для этого достаточно увидеть в спектре систематическую "зубчатую" структуру. "Зубья" в такой структуре могут быть некачественными - различаться по форме, быть сплюснутыми и т.д. Но расположение этих энергетических выступов подчинено определенной системе и ее мы сейчас обсудим.



На рис.6-2 показан график степенной функции 3^x а также графики десятичного, натурального и двоичного логарифмов от этой функции. Основание степенной функции выбрано произвольно так, чтобы оно не совпадало с основаниями логарифмических функций.

Заметим, что степенную функцию можно понимать как обобщение геометрической прогрессии, в частности прогрессии, образованной последовательностью частот октавных интервалов: $f_0, f_0 \cdot 2, f_0 \cdot 4, f_0 \cdot 8, f_0 \cdot 16, \dots, f_0 \cdot 2^k \dots$

Поэтому понятно, что, если частоты, соответствующие октавам, изображать шкалой в натуральном числовом выражении, то метки частоты для последовательных нот будут все более отдаляться друг от друга. Если же эти частоты изображать в логарифмической шкале, то (поскольку логарифм от степенной функции при любом основании дает линейную зависимость, как это видно на рис.6-2) упомянутые степенные (октавные) метки будут располагаться на равном расстоянии друг от друга, то есть на графике октавы будут равноразмерны (как и на клавиатуре фортепиано).

На рис.6-3,6-4 показаны два варианта отображения спектра звука певческого голоса: с линейной шкалой частоты (рис.6-3) и с логарифмической шкалой частоты (рис.6-4). Обсудим эти изображения.

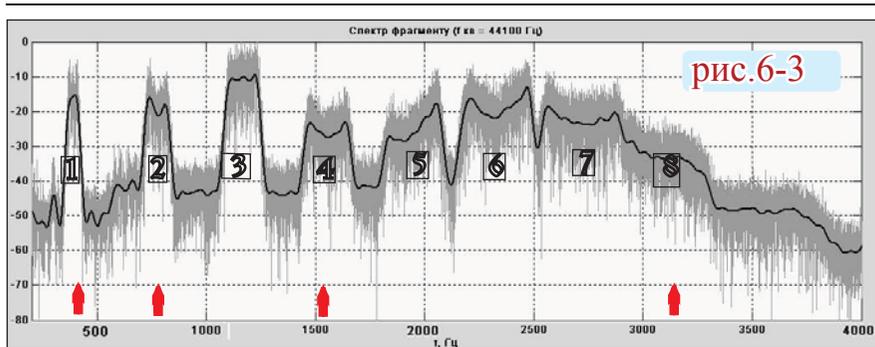


рис.6-3

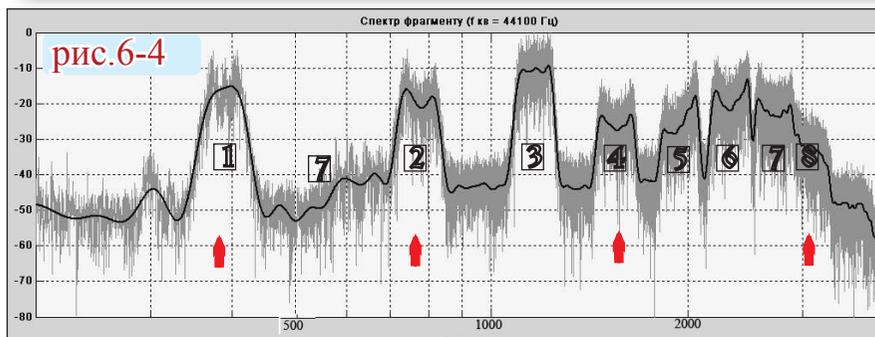


рис.6-4

Особенности изображения музыкальных звуков

4. На рис.6-3 видно как арифметическая прогрессия обертоновой структуры музыкального звука $f_0, f_0 \cdot 2, f_0 \cdot 3, f_0 \cdot 4, f_0 \cdot 5, \dots, f_0 \cdot k \dots$ соотносится с геометрической прогрессией октавных интервалов $f_0, f_0 \cdot 2, f_0 \cdot 4, f_0 \cdot 8, f_0 \cdot 16, \dots, f_0 \cdot 2^k \dots$ при линейной шкале частот. Мы повторяем сказанное выше: октавы (красные стрелки) расширяются на линейной шкале, а обертоны равноудалены друг от друга.

На рис.6-4 показана взаимосвязь тех же видов прогрессии при логарифмической шкале, изображение которой становится "линейным" для степенных функций и геометрических прогрессий и, разумеется, "нелинейным" для арифметических. На этом рисунке отчетливо видно как внутрь второй по счету октавы попадает один обертон, внутрь третьей - уже три обертона, и для последующих октав количество "внутренних" обертонов будет расти.

Понимание этих закономерностей позволяет по виду структуры музыкального звука на спектральной картине сразу определить характер шкалы оси частот даже без числовых меток на ней.

Остается вопрос - какую из этих двух шкал следует рекомендовать как более результативную. Здесь нет однозначного ответа, он зависит от цели, которую вы ставите, изображая спектр музыкального звука. Если вы собираетесь оценить степень гармоничности обертонов (в смысле соответствия их

частот арифметической прогрессии, то есть равноудаленность друг от друга), симметрию их формы и прочие детали такого рода, то линейная шкала предпочтительнее. Если же хотите посмотреть расположение их в области слухового восприятия по аналогии с материалами гл. 1, то для этой цели более подходит логарифмическая шкала.

5. Следует иметь в виду, что, в отличие от горизонтальной оси частот на спектральной картине, определить характер вертикальной шкалы при отсутствии разметки невозможно. То есть без соответствующих надписей мы не можем судить, отображены ли уровни обертонов в натуральном (процентном соотношении), или в нелинейно "сжатом" - логарифмическом (dBFS).

Вибрато музыкального звука

5. Вибрато - это частотная и/или амплитудная модуляция звука³. Это слово уже встречалось в нашем тексте. Музыкальные исполнители хорошо понимают значение этого художественного приема и способы его выполнения.

На рис.6-5 показано глубокое амплитудное вибрато из пьесы "Маленький цветок" (Petite Fleur, 1952 в исполнении Сиднея Беше (кларнет)). Отдельно существующее амплитудное вибрато - редкое явление. Как правило оно (с небольшим размахом) сопровождается частотное вибрато, которому исполнители отдают предпочтение.

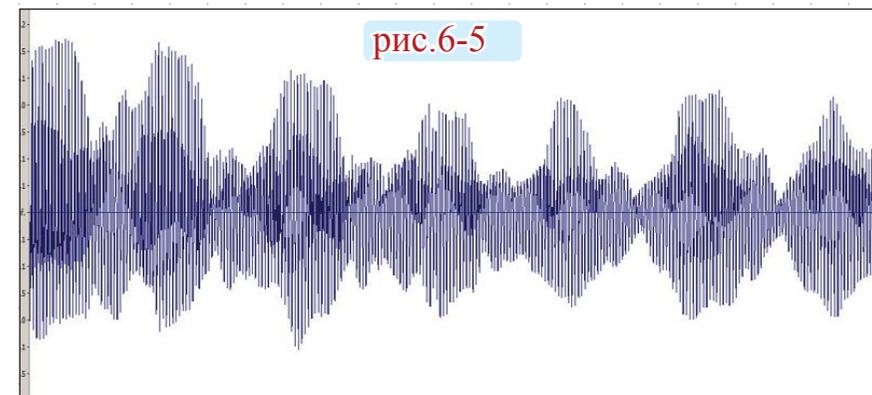
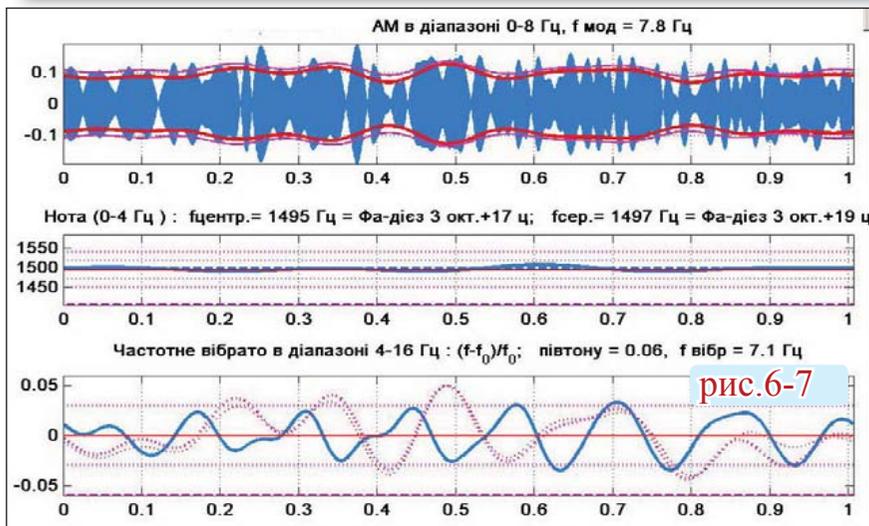
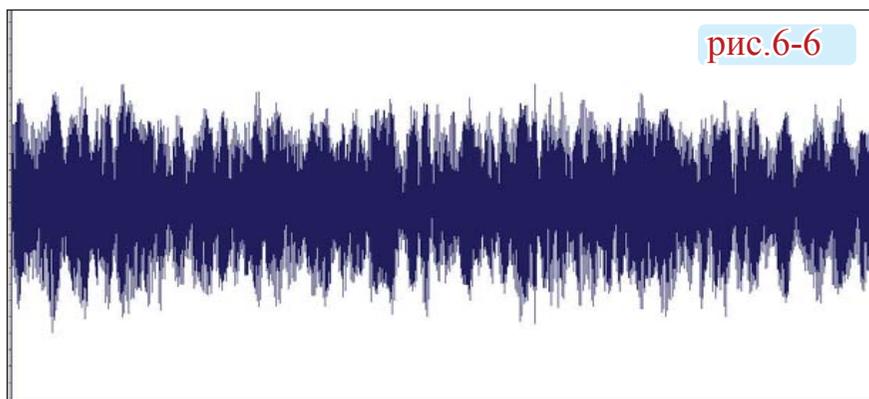


рис.6-5

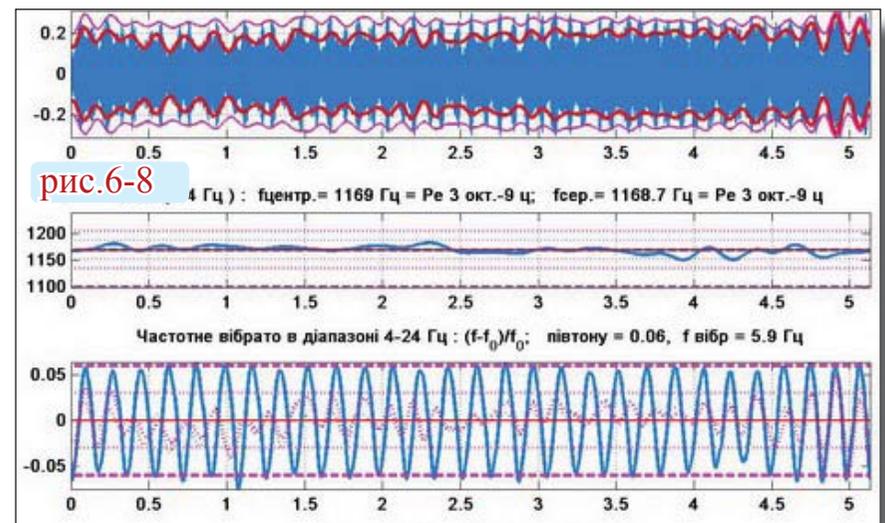
На рис.6-6 приведена запись ноты, исполненной французской певицей Мирей Матье - звуку присуще частотное вибрато, как это обычно бывает у певцов, однако частотная модуляция на рис.6-6 не видна. Ее можно видеть лишь после частотной демодуляции такого сигнала. Пример результата такой демодуляции показан на рис.6-7. Здесь развернуты во времени три продукта демодуляции. Верхняя лента показывает один из обертонов звука голоса и на его фоне красной линией колебания огибающей ("амплитуды") сигнала,

³ Подробное объяснение явления модуляции содержится в [Ак,25].



что порождает практически незаметные изменения громкости. Средняя лента отображает поведение средней частоты звука, из нее следует, что Мирей Матье безупречно выдерживает интонационную стабильность исполняемой ноты. Нижняя лента содержит функцию модуляции частоты, то есть отклонения частоты от среднего значения, приведенного выше - это и есть частотное вибрато. Размах этого отклонения не превышает полутона, а характер изменения во времени отдаленно напоминает синусоидальный.

Полезно сравнить приведенные данные с материалами рис.6-8, которые были получены при исследовании голоса баритона Дмитрия Хворостовского. В отличие от предыдущего звука этот длится в течение пяти секунд. Амплитудные колебания незначительны и это свидетельствует о профессиональном умении длительно удерживать постоянный поток воздуха при пении. Интонационная



стабильность звука (средняя лента) при этом вполне удовлетворительна. Что же касается частотного вибрато, то его следует признать феноменальным. На протяжении пяти секунд певец выдерживает практически синусоидальную манипуляцию частотой с абсолютно равным размахом частоты почти в целый тон. Столь глубокое и безукоризненное по форме вибрато свидетельствует об исключительных природных данных голосового аппарата певца и хорошем его физиологическом состоянии. Заметим, что высокое качество параметров исполняемых звуков, разумеется, приобретается по мере постановки голоса в процессе обучения и певческой практики и впоследствии неизбежно утрачивается с возрастом.

6. Явление вибрато мы продемонстрировали на примере певческих голосов и это неудивительно, поскольку вибрато является одним из основных эффектов, украшающих пение. Однако можно видеть применение вибрато и в инструментальной музыке. И, если для тростевых инструментов вибрато более доступно в амплитудном варианте (рис.6-5), то исполители на смычковых инструментах пальцевой вибрацией незначительно систематически изменяют длину колеблющейся части струны и, тем самым, модулирует частоту издаваемого тона.

Принято считать, что наиболее благоприятно воспринимается вибрато с частотой порядка 4.5 .. 5.5 Гц, однако эти значения следует считать усредненными показателями. Известно, например, что французские шансонье (исполнители популярных жанров) тяготеют к завышенному вибрато с частотой порядка 7 Гц (рис.6-7). Немецкие вокальные исполнители старинной музыки иногда демонстрируют заниженную частоту вибрато - порядка 4 Гц. Такие, казалось бы, небольшие отклонения отчетливо регистрируются слухом, создавая весьма специфические музыкальные ассоциации. Для относительно высокой частоты вибрато (порядка 8 Гц) существуют отдельные наименования

ния - иногда такой вариант исполнения называют "тремоло", иногда "барашек". Такое вибрато чаще всего встречается у неопытных певцов с неподготовленным голосовым аппаратом.

Кроме частоты характеристикой вибрато является также его глубина - так называют размах частоты в процессе ее модуляции. Этот параметр у квалифицированного певца обычно составляет величину между музыкальным полутоном и тоном и зависит от степени эмоциональной напряженности при исполнении ноты. Во всяком случае, при глубине порядка тона и более вибрато начинает доминировать среди других свойств исполняемой ноты, воздействуя на слушателя.

7. Наличие частотного вибрато возможно распознать на спектральных картинах звука. Мы уже видели вибрато Паваротти на спектрограммах в виде регулярных во времени отклонений частоты от центрального значения. Но спектральные картины интонационно стабильного звука тоже хорошо указывают на наличие в звуке вибрато (мы уже отмечали, что наличие вибрато, в принципе, не нарушает свойства интонационной стабильности, поскольку и в этом случае можно указать исполняемую ноту).

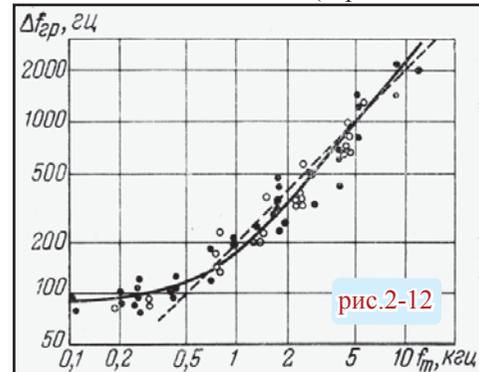
Частотное вибрато (частотная модуляция) расширяет спектр сигнала [Ак,68], не изменяя его энергию. Это значит, что при линейной шкале оси частот образы обертонов за счет вибрато должны расширяться, но при этом становиться ниже по уровню так, чтобы площадь рисунка, отведенная обертоном, сохранялась: чем глубже вибрато - тем обертон становится одновременно шире и ниже.⁴ Далее заметим, что, например, общее вибрато с глубиной в полутоном отражается в каждом обертоне полутоновым смещением частоты. То есть при сдвиге частоты основного тона до величины $f_0 + \Delta f$ обертоны сдвинутся до величин $2(f_0 + \Delta f)$, $3(f_0 + \Delta f)$, $4(f_0 + \Delta f)$, и т.д. На рисунке каждый обертон будет расширяться тем более, чем больше его номер (см. рис.6-3). И наоборот, в случае логарифмической шкалы для реальных значений частоты обертонов их ширина имеет вид $\log(f_0 + \Delta f)$, $\log(f_0 + \Delta f) + \log 2$, $\log(f_0 + \Delta f) + \log 3$, $\log(f_0 + \Delta f) + \log 4$, и т.д., что совершенно не может быть заметно на картине, то есть ширина всех обертонов оказывается одинаковой (см. рис.6-4).

Повторно обратившись к рис.6-3 и рис.6-4, с помощью приведенных объяснений мы можем окончательно осознать признаки, по которым следует распознавать структуру музыкального звука в различных частотных шкалах.

8 В заключение вернемся к вопросу параграфа "Критические полосы спектра" главы 2 (стр.33) и повторим рис.2-12. Из этого рисунка видно, что полоса, в пределах которой гармонический сигнал самостоятельно формиру-

⁴ При необходимости получить стандартный спектральный портрет звука следует попросить музыканта исполнить "прямой" звук без вибрато.

ет ощущение громкости, составляет примерно 100 Гц в интервале до 500 Гц. А это значит, что даже весьма низкочастотный музыкальный звук, например, с основным тоном 100Гц (середина Большой октавы на фортепиано) с обертонами в 200 Гц, 300 Гц, 400 Гц и т.д. (разделенными интервалом в 100 Гц, формирует тембр в нижней форманте, комбинируя независимые громкости каждого из обертонов. При этом высокие 10..15-е обертоны взаимодействуют по громкости уже отчасти совместно, но тембровую роль этих диссонансирующих обертонов мы уже указывали.



Для тона с частотой, например 250 Гц (вблизи до 1 октавы) нетрудно видеть, что обертоны, принадлежащие нижней певческой форманте (т.е. ниже 2 кГц, подробнее о ней см. главу 8), в силу своей удаленности друг от друга на 250 Гц тоже самостоятельны в формировании общей громкости, хотя и заходят в более высокочастотную зону.

Для тона с частотой 600 Гц (начало второй октавы) вновь видно, что и здесь при удалении друг от друга на эту величину обертоны НПФ находятся вне критической полосы соседних обертонов и тоже независимы.

Таким образом, нижняя певческая форманта (как и нижние обертоны любого музыкального звука) играет еще и ту роль, что формирует базовую часть тембра за счет независимого сложения эффектов громкости, производимых каждым из ее обертонов. В то же время высокие обертоны звука в большей или меньшей мере взаимодействуют между собой в пределах общих критических полос. Это, в частности, может быть учтено при синтезе или эквалайзинге музыкального звука, в котором для нижних обертонов конкретной ноты можно управлять тембром за счет регулировки уровня отдельного обертона, а для высших - регулировкой уровня совокупности обертонов в пределах критической полосы, что уже было отмечено ранее в главе 2.

7 Музыкальные инструменты

1. Практически каждая книга, касающаяся музыкальной акустики, содержит материал об акустических свойствах музыкальных инструментов. В большинстве случаев такой материал рассказывает о физическом устройстве этих инструментов и волновых процессах в них. Раскрытие физики звукообразования вместе с простыми формулами, поясняющими эту физику, безусловно является полезным для целого ряда специалистов, профессия которых непосредственно связана с музыкальными источниками звуков.

Некоторые источники звука основаны на собственных колебаниях осциллятора (генератора звука), однократно возбужденного внешним усилием - струны гитары, арфы и т.п. Некоторые - на автоколебаниях¹ генератора - струны скрипки, трости в язычковых инструментах. Некоторые на вынужденных колебаниях резонирующего столба воздуха - флейты. Конструкции музыкальных инструментов содержат важные составные компоненты, которые играют существенную роль в звукообразовании - это резонаторы, настраиваемые воздуховоды, рупорные элементы на концах некоторых инструментов и т.п.

Мы, однако, полагаем, что звукорежиссеру более интересны свойства собственно звуков, издаваемых музыкальными инструментами, и в этой книге остановимся на этих свойствах.

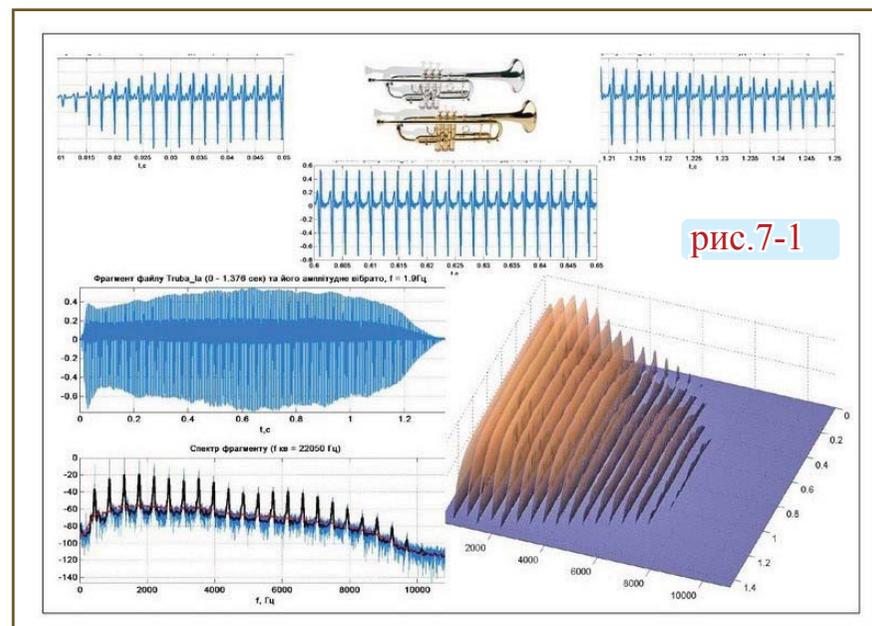
На рис.7-1..7-8 показаны комплекты визуальных характеристик звуков музыкальных инструментов в следующем порядке:

- Сверху располагаются фрагменты временной развертки записи звуков, а именно - начало звука, его середина и завершение.
- Ниже слева располагается полный временной образ звука, разумеется звук должен быть интонационно стабильным (одна нота), допускается лишь вибрато (модуляция частоты и/или амплитуды сигнала).
- Еще ниже находится спектральная картина (амплитудный спектр) сигнала, рассмотренный в гл.2.
- Внизу справа располагается трехмерная спектральная картина звука. На плоскости из глубины на зрителя ориентирована ось времени, а почти горизонтально идущая ось слева направо отражает шкалу частот так, что более низкие частоты располагаются слева. Эта картина сходна со спектрограммой, но амплитуды кратковременно измеренных спектров отражены не просто цветом, а так сказать "поставлены в полный рост".

¹ незатухающие колебания за счет энергии постоянного внешнего воздействия.

Приведенные здесь визуальные образцы-карты параметров музыкальных звуков являются иллюстративным материалом для нескольких характерных групп инструментов, и позволяют обсудить особенности их звучания, не привлекая непосредственного слухового материала. Разумеется, любой музыкант без всяких иллюстраций внутренне представляет себе звучание выбранных здесь инструментов. Но визуальные характеристики позволяют выполнять детальный объективный анализ, а также тренируют специалиста для экспертной оценки звуков, которые недоступны для слухового восприятия, по их характеристикам..

Оркестровая труба

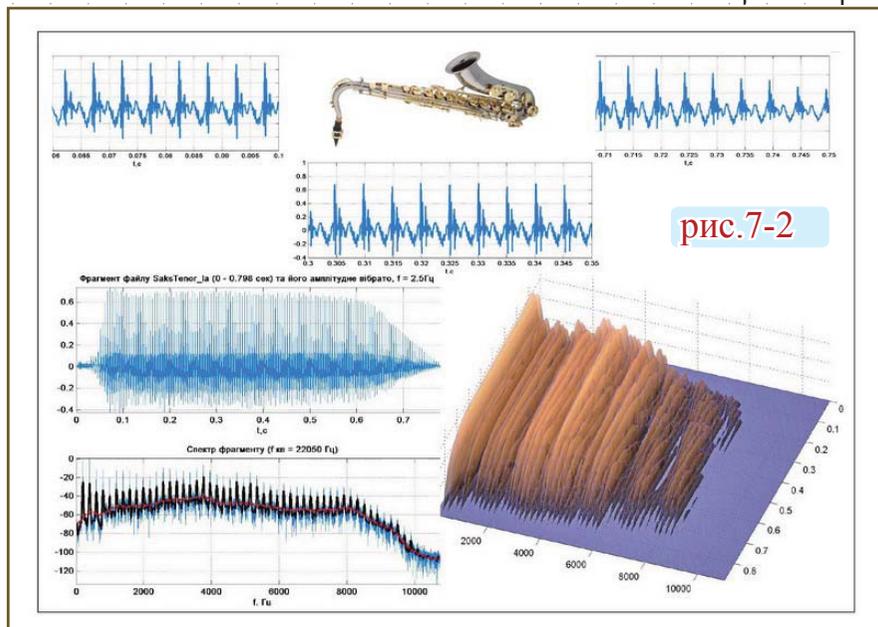


2. Оркестровая труба. Звук этого инструмента обладает весьма отчетливо выраженной периодичностью, высокой интонационной стабильностью и тембровой устойчивостью. Тембр трубы практически неизменен на протяжении всего звучания от атаки до финального спада, что видно как форме временных фрагментов, так и по трехмерной спектральной картине. Лишь самые высокие спектральные компоненты (обертоны) звука возникают несколько позже начала звука и исчезают перед его окончанием. Это объясняется тем, что в амбушюрных (мундштучных) инструментах начало и конец звука всегда несколько приглушены. Количество обертонов звука трубы достаточно велико - до десятка первых гармоник формируют выразительный тембр инструмента и еще примерно десяток более высоких и энергетически значимых обертонов придают громкому звуку пронзительный характер. Все гармоники от-

четливо выражены с чистыми межгармоническими промежутками. Наиболее значимые, отрисованные тонкими синими линиями вплоть до 0 dBFS (более толстые линии прорисовывают общую огибающую спектра), имеют уровень порядка 60dB.

Кроме сольного исполнения оркестровые трубы востребованы в оркестровых проектах. "Прорезающие" в tutti любую аккордовую ткань трубы доминируют благодаря своей интонационной и тембровой стабильности при весьма мощном звуке.

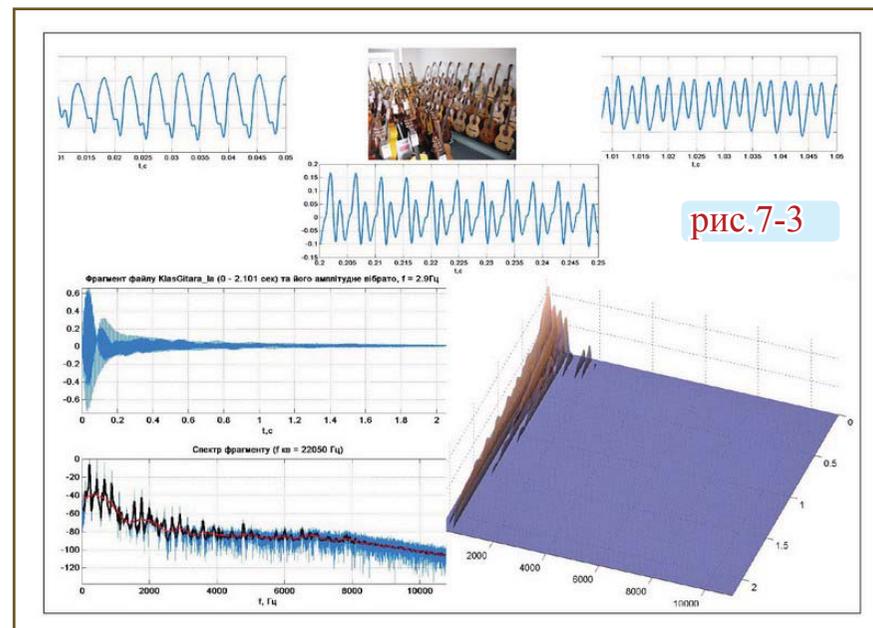
Тенор-саксофон



3. Тенор-саксофон. Даже внутреннее извилистое заполнение промежутков сигнала между максимумами периодов во временных фрагментах свидетельствует о наличии выразительных высокочастотных компонент в спектре. В самом деле, спектр сигнала весьма богат обертонами - в энергетически значимой части спектра их около сорока. Наиболее высокочастотные из них располагаются в зоне шипящих и свистящих призывков, которые заметно влияют на тембр звука. Трехмерная картина отчетливо выявляет поведение обертонов звука во времени: тембр звука удерживается стабильным, плотным в широком диапазоне частот, и это следует иметь в виду при построении аккордов с другими инструментами, играющими партию выше тенора. Так, например, одиночный альт-саксофон в дуэте с тенором будет играть роль подголоска

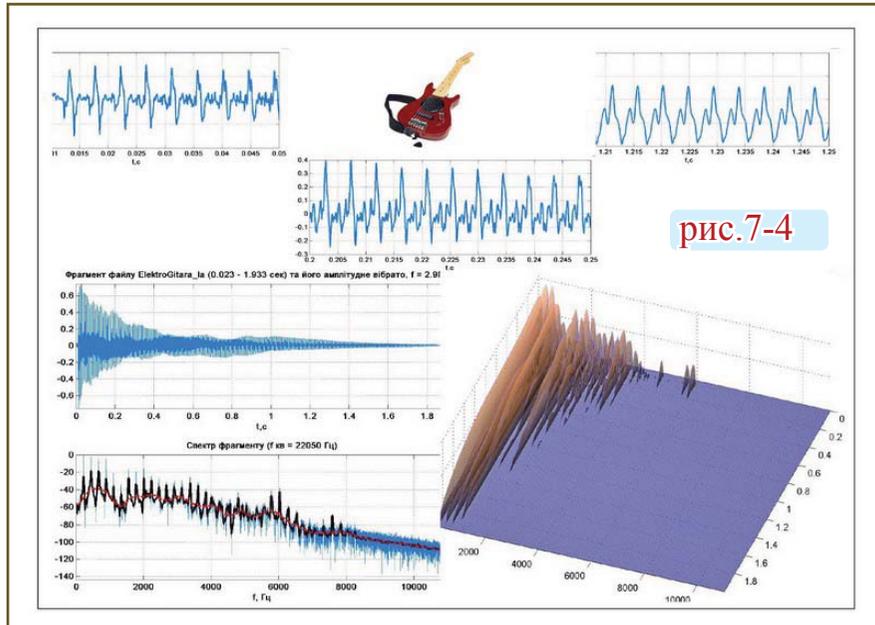
вследствие эффекта маскировки. При желании выделить функцию алта его следует либо разнести с тенором в пространстве, либо усилить функцию за счет присоединения к альту еще одного инструмента в унисон.

Классическая гитара



4. Классическая гитара. Затухающую форму развертки во времени может иметь звук либо щипкового, либо ударного инструмента. Струнную щипковую природу этого инструмента раскрывают фрагменты развертки, которые существенно различаются на разных стадиях звучания. Атака звука почти синусоидальна, а развитие его вплоть до самого конца носит более сложный характер. Это обстоятельство вместе с амплитудным всплеском вскоре после атаки наводит на мысль, что при энергичном щипке в колебательном процессе кроме струны принял участие и гриф, как это бывает в так называемых связанных колебательных системах.

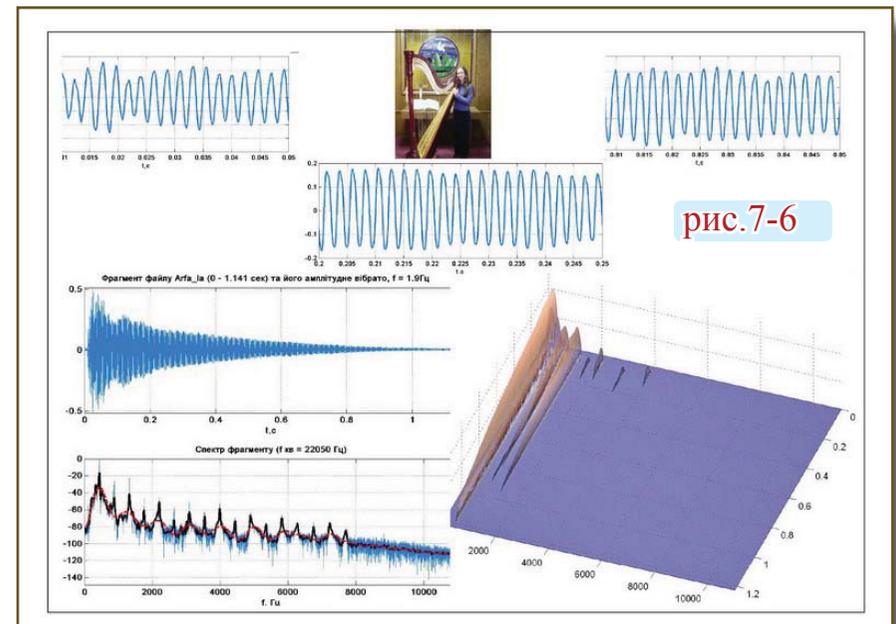
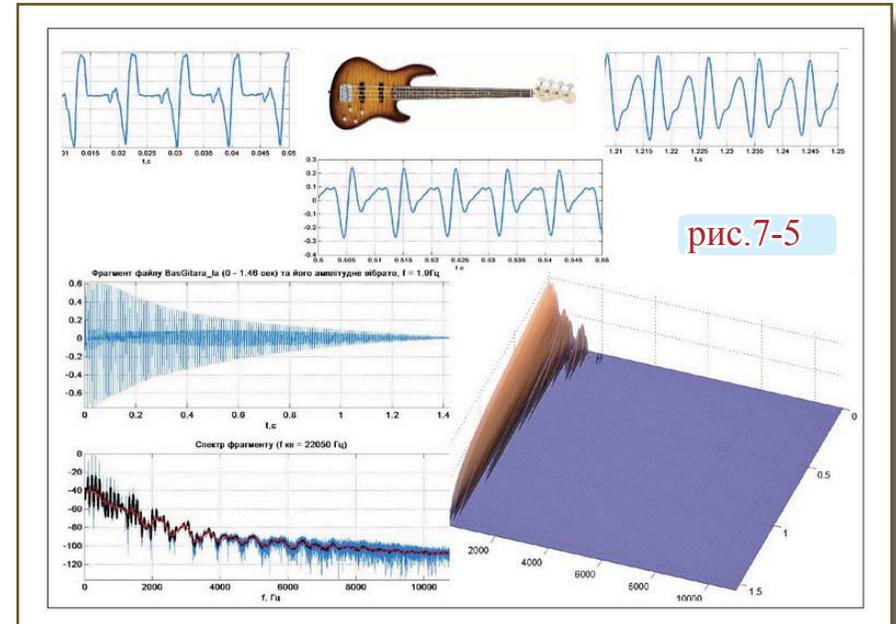
Спектр звука довольно беден. Тембр его преимущественно определяется первыми четырьмя гармониками, которые создают консонансное ощущение - очевидно, что здесь идет речь об игре на нейлоновых струнах без медиатора. Характер временных фрагментов в середине и конце звука, суди по их примерно одинаковой "гладкости", указывают, что несколько изменяются лишь амплитудно-фазовые соотношения между гармониками, что позволяет считать тембр звука по завершении атаки близким к стабильному.

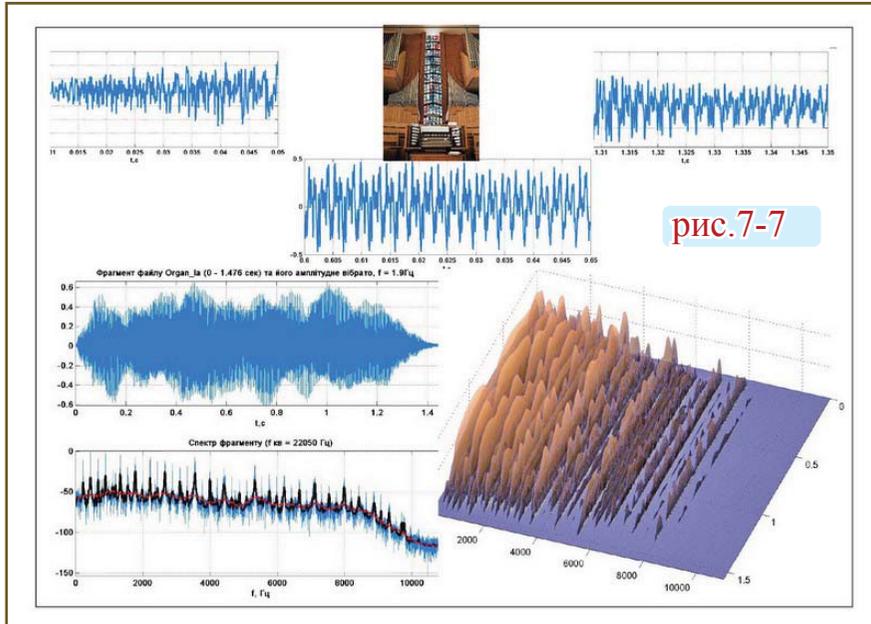


5. Электргитара. В приведенном примере очевидно, что спектр ее сигнала обогащен высокочастотными составляющими за счет электронных компонентов инструмента. Это видно и на фрагментах временной развертки сигнала, да и собственно спектральная картина демонстрирует обилие обертонов звука, другое дело, что эти обертоны активно присутствуют примерно в течение половины длительности звука, и затем постепенно исчезают, начиная с высоких обертонов. Показанный здесь звук в начальной фазе должен иметь диссонансный характер, однако в дальнейшем он содержит только "нижние" обертоны, что хорошо видно на трехмерной спектральной картине.

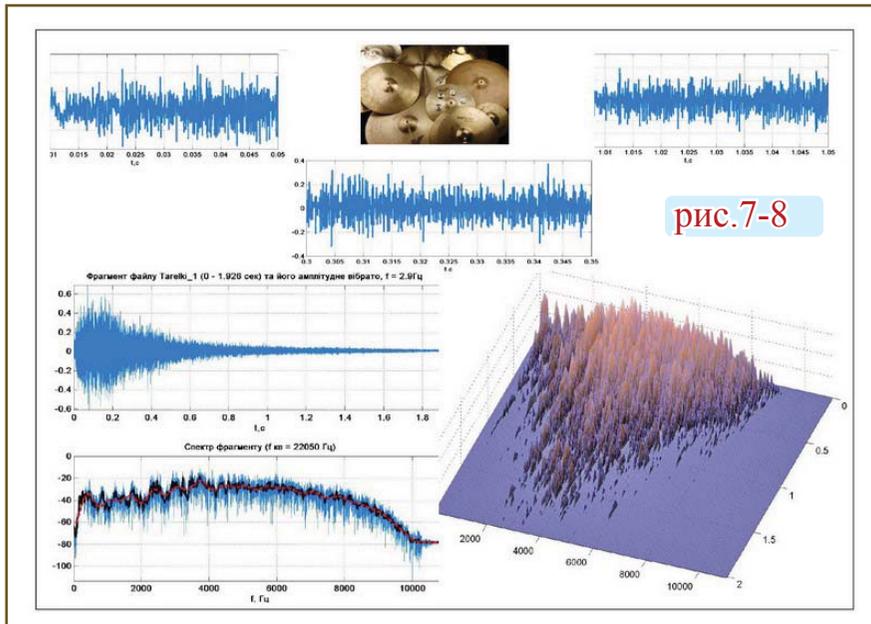
Можно полагать, таким образом, что подвижная игра аккордами на таком инструменте приведет к ощущению "грязи" в гармониях, в то же время соло гитары при исполнении кантильных партий может производить благоприятное впечатление.

6. Разумеется, подобного рода материалы полезны, в основном, для экспертной оценки качества звука, соотнося их с реальными слуховыми ощущениями. Далее мы приводим аналогичные иллюстративные материалы еще для нескольких музыкальных инструментов, предлагая читателю самостоятельно провести анализ по аналогии с предыдущими.





Тарелка Ride



8 ПЕВЧЕСКИЙ ГОЛОС

Основные вокальные характеристики певческого голоса

1. Работа звукорежиссера с певцом может носить тем более напряженный характер, чем ниже квалификация певца. Слово "квалификация" в бытовом контексте может означать и опыт его сценической деятельности, и умение работать с залом, и подбор репертуара - все составляющие, которые обеспечивают артисту успех. И нужно сказать, что такая квалификация вполне может создать исполнителю в массовых популярных жанрах авторитет, который компенсирует недостатки собственно вокальных данных, составляющих "основные средства производства" профессионально работающего певца.

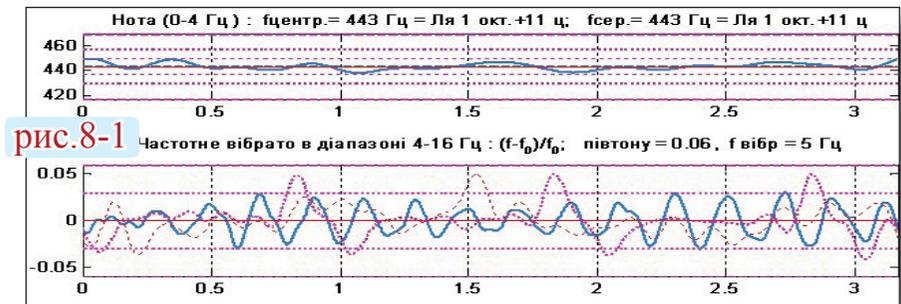
Одной из первых научных работ, посвященных высококачественным певческим голосам, явилась работа В.Бартоломью, помещенная в журнале Акустического общества Америки еще в 1934 г. В этой работе автор указал четыре основных свойства, которые должны быть присущи хорошо поставленному певческому голосу: 1 - вибрато, 2 - достаточная громкость при исполнении, 3 - нижняя певческая форманта, 4 - верхняя певческая форманта.

Рассмотрим эти характеристики, исключив из них лишь громкость (интенсивность) исполнения. Эта характеристика имеет значение в оперном исполнительстве, однако, например, обладатель качественного но негромкого голоса может реализовать себя в концертной деятельности, где микрофон давно стал неотъемлемым атрибутом сцены.

Вибрато

2. О вибрато нами было сказано уже немало. Этот элемент является общим и для певческого голоса, и для музыкальных инструментов, способных его воспроизвести.

На рис.8-1..8-3 приведены развертки во времени функций изменения частоты при тестовом исполнении фиксированной ноты ряда студентов-певцов му-



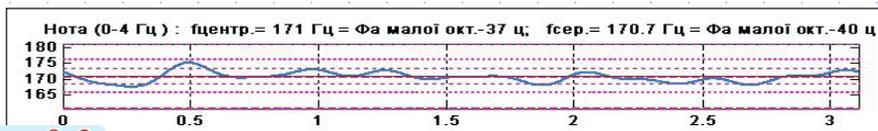


рис.8-2 Частотне вібрато в діапазоні 4-16 Гц : $(f-f_0)/f_0$; півтону = 0.06, f вібр = 4.9 Гц

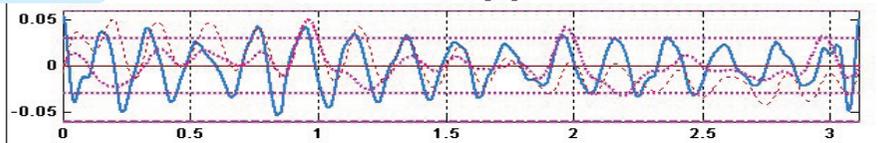
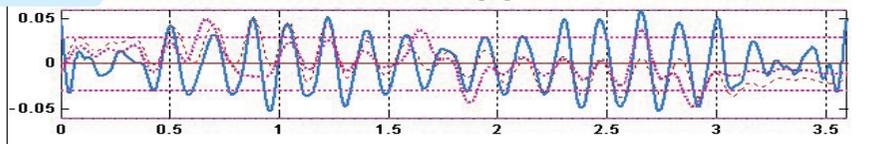


рис.8-3 Частотне вібрато в діапазоні 4-16 Гц : $(f-f_0)/f_0$; півтону = 0.06, f вібр = 5.6 Гц



зыкального учебного заведения¹. в каждом примере верхняя лента показывает медленную часть изменений (в диапазоне от 0 до 4 Гц), которая отражает способность певца удерживать общую интонационную стабильность звука. Нижняя лента каждого рисунка (для диапазона частот от 4 до 16 Гц) как раз иллюстрирует возможность певца исполнить вибрато, что и являлось заданием теста. Среди приведенных лишь пример рис.8-2 можно считать явлением вибрато, остальные варианты оставляют слуховое впечатление неуверенно исполненной ноты.

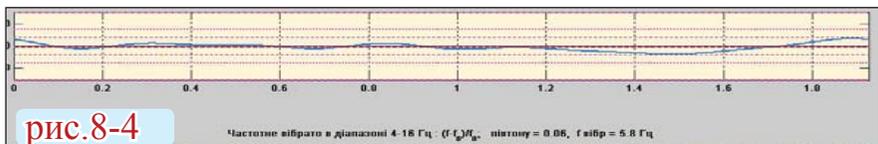


рис.8-4 Частотне вібрато в діапазоні 4-16 Гц : $(f-f_0)/f_0$; півтону = 0.06, f вібр = 5.8 Гц

¹ На верхних лентах каждого рисунка указана средняя удерживаемая частота для анализируемого обертона. А на нижней - отклонение от этой частоты в относительных долях: так ± 0.05 означают 5% отклонения. Заметим, что ± 0.06 составляют как раз отклонение в \pm полутон.

На рис.8-4 показаны аналогичные результаты для неопытной певицы-аматора. Очевидно, что исполнительница не может выполнить задание, хотя уверенно интонирует звук, обладая хорошим музыкальным слухом.

3. Далее приведем образцы вибрато от вокалистов, общепризнанных великими, без указания их имен. Относительно последнего образка заметим, что певец удерживал показанное здесь вибрато в процессе снижения дина-

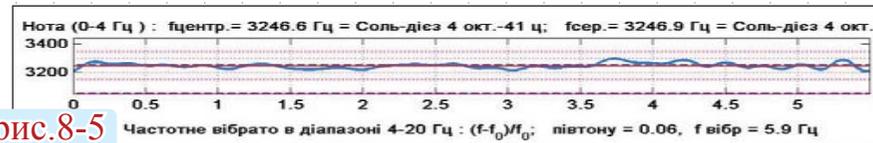


рис.8-5 Частотне вібрато в діапазоні 4-20 Гц : $(f-f_0)/f_0$; півтону = 0.06, f вібр = 5.9 Гц

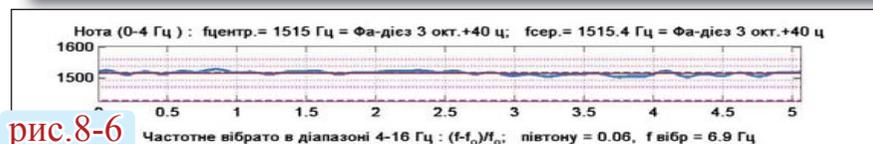


рис.8-6 Частотне вібрато в діапазоні 4-16 Гц : $(f-f_0)/f_0$; півтону = 0.06, f вібр = 6.9 Гц

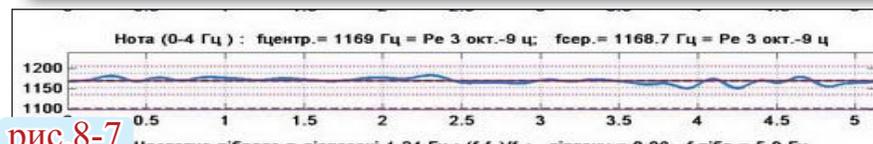


рис.8-7 Частотне вібрато в діапазоні 4-24 Гц : $(f-f_0)/f_0$; півтону = 0.06, f вібр = 5.9 Гц

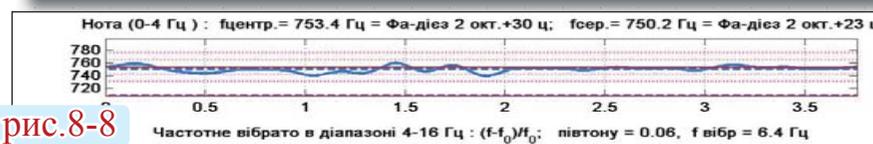


рис.8-8 Частотне вібрато в діапазоні 4-16 Гц : $(f-f_0)/f_0$; півтону = 0.06, f вібр = 6.4 Гц

мического уровня к пианиссимо, лишь несколько снизив размах модуляции. Практикующие певцы знают, что реализовать вибрато "в полный голос" намного легче, чем делать это в диапазоне пиано-пианиссимо.

Разница в приведенных выше (пункты 2 и 3) примерах очевидна и внимательное отношение звукорежиссера к характеристике вибрато исполнителя позволяет в определенной мере судить об исполнительском уровне. Свойство хорошего вибрато состоит в том, что оно не воспринимается сознанием, как отдельный элемент, а сливается с воспроизводимым звуком, придавая ему живость.

Форманты

4. Спектр профессионально исполненного вокального звука выдержанной интонационно стабильной ноты в области частот ниже 4 кГц, если его сгладить плавной кривой по верхушкам обертонов, в большинстве случаев выглядит как два куполообразных энергетических подъема. Левый, более низкочастотный из них называют нижней певческой формантой (НПФ), верхний же, расположенный в области 3 кГц - верхней певческой формантой² (ВПФ).

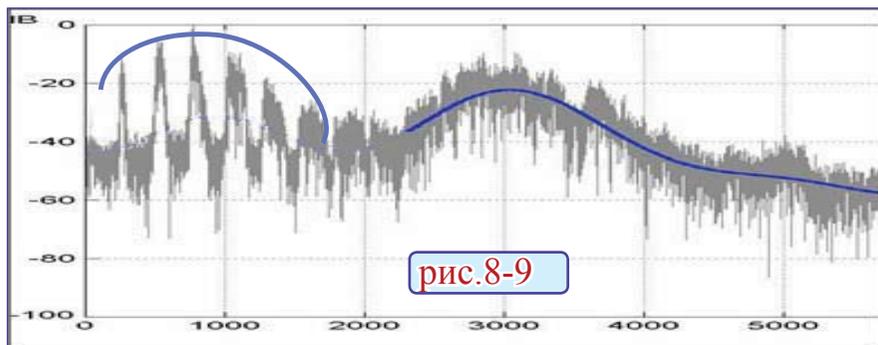


рис.8-9

Рис.8-9 удачно иллюстрирует наличие НПФ и ВПФ в спектре исполненной мужским голосом ноты. В данном случае НПФ состоит из шести обертонов и располагается в интервале до 2 кГц. Видно, что звуку сопутствует вибрато, которое расширяет полосы обертонов так, что для частот свыше 2 кГц они становятся неразделимы. Примерно с 2400 Гц и до 3700 Гц размещается ВПФ, а выше ее правой границы формально находится место для "паразитных" вокальных призвуков, которые в исследуемом звуке отсутствуют.

Именно обертоны в нижней форманте определяют тембровую основу голоса певца, делая его узнаваемым. Разумеется, большую роль в узнаваемости

² Формантой называют область энергетического подъема в спектре звука голоса. Наличие формант обусловлено характером резонаторных полостей голосового тракта.

играют и другие факторы - атака звука, манера дыхания, стиль вибрато и пр. Именно эти последние элементы украшают звук, окончательно формируя певческую индивидуальность. Сама по себе нижняя форманта певца звучит глуховато и "плоско". Но в рамках спектрального рассмотрения НПФ относится к элементам, звукорежиссерское вмешательство в которые недопустимо.

5. На этом же рисунке отчетливо выделена и верхняя певческая форманта (ВПФ). Она всегда расположена в области вблизи частоты 3 кГц. Относительно стабильное положение формантных областей на оси частот обусловлено тем, что они формируются внутренними резонансными полостями голосового аппарата певца, объем которых не очень различается даже среди худощавых и полных людей. Полнота тела определяется свойствами жировой и мускульной составляющих тела. А для голосовых качеств важны упругость и отсутствие каких-либо паразитных образований во внутренних формирующих звук элементах, включая голосовые связки, гортань, ротовую и носовую полости. Поэтому положение НПФ иногда связывают с "серединой" частотой 500 .. 600 Гц, а ВПФ, как уже было хорошо видно на рис.8-9, с частотой 3 кГц, хотя истинное ее положение, разумеется, немного зависит от типа голоса - детский, женский, мужской.

Верхняя форманта размещена в области наибольшей чувствительности слуха человека, сама по себе в отдельности производит впечатление, как это говорят, птичьего пения, безусловно украшает профессиональный вокал, отличая его от "домашего", иногда говорят "разговорного" пения, присущего так называемой "авторской песне".

Важно то, что при постоянном положении формантной резонаторной области в нее могут попадать обертоны с разными номерами в зависимости от конкретной тесситуры певца в процессе исполнения произведения. На рис.8-10 показан спектр звука исполненного певицей (колоратурное сопрано) в высокой тесситуре. В отличие от предыдущего рисунка к нижней форманте можно отнести лишь первые две или три гармоники. Третья гар-

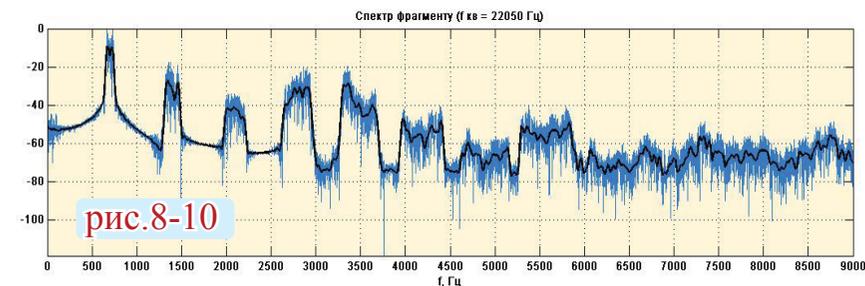


рис.8-10

моника (второй обертоном) соединяет форманты, и в верхнюю форманту, таким образом, условно попадают обертоны со второго по пятый. Энергетически значимый седьмой обертоном в районе 5.5 кГц и последующий сравнительно высокий уровень высокочастотных призвуков, видимо, обязаны своим происхождением почтенному возрасту певицы, который обязан был сказаться на состоянии ее вокального горлового аппарата.

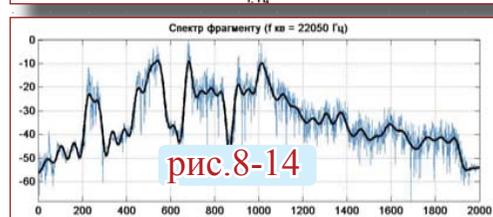
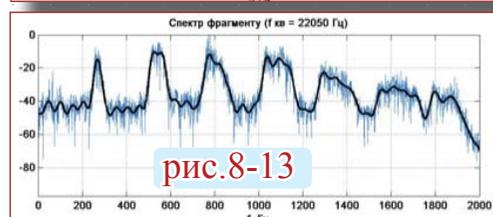
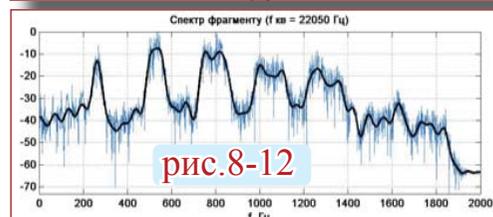
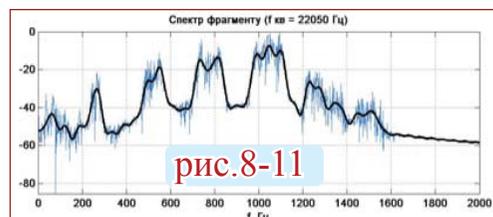
Заметим, что исполненная нота очевидно обладает существенно выраженным частотным вибрато.

6. На рис.8-11..8-14 приведены НПФ четырех певцов, исполняющих одну и ту же ноту в арии Ленского из оперы Чайковского. Характер форманты на нижнем рисунке 8-14 объясняется экстравагантной манерой пения певца, который тоже, в меру способностей, исполнил арию.

Несмотря на индивидуальность "портретов" нижней форманты видна схожесть положения ее максимума, которая, в частности, выражается в том, что основной тон звука занижен по уровню в сравнении с максимальными обертонами.

Здесь мы вспомним, что наличие частотной модуляции расширяет спектральную картину обертона и делает ее ниже по уровню. Поэтому в отсутствие вибрато разница между уровнями основного тона и максимума форманты окажется еще заметнее.

Заметим также, что соотношение между уровнем основного тона и обертонов на рис.8-10 объясняется именно тем, что здесь основной тон для высокой ноты колоратурного сопрано попадает как раз на область максимальных значений нижней форманты (примерно 500 .. 700 Гц).



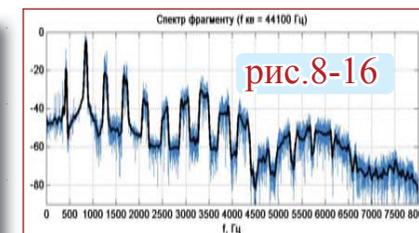
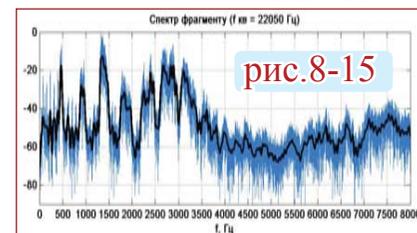
Примеры спектральных характеристик голоса певцов

На рис.8-15..8-22 приведена серия спектральных картин, на которых видны формантные области и вышележащие призвуковые области в звуках, исполненных рядом певцов преимущественно прошлого времени. Прокомментируем эти материалы, сделав предварительное замечание относительно имен исполнителей..

Я понимаю, что некрасиво интриговать, намекая на человека и не называя его. Но еще более некрасиво без особой необходимости публично хвалить одних людей и критиковать других не за поступки а за отпущенные природой данные. Здесь я, не называя имен исполнителей, выбрал тем самым менее некрасивое поведение и приношу за него извинения.

Рис.8-15 отображает спектр звука знаменитого в прошлом тенора, ВПФ которого сравнима по уровню с нижней формантой - у прославленных вокалистов зачастую так и бывает - располагаясь в области максимальной чувствительности слуха, эта область не только придает "колористический" блеск голосу, но и обеспечивает ему распространение на большие расстояния в зале на фоне аккомпанирующего оркестра, что певцы называют "полетностью" голоса.

На рис.8-16 представлен спектр замечательной эстрадной певицы с неповторимо нежной атакой звука. Ее ВПФ заметна, хорошо заметен также и призвуковый подъем в области 6 кГц, однако тембровое ощущение от ее голоса формируется, кроме этого, отчетливым рядом чистых, "незамутненных" гармоник, которые даже при наличии вибрато указывают на крепкость и стабильность голосового формирующего аппарата.



Спектр звука на рис.8-17 принадлежит знаменитому в свое время совсем юному певцу, прославленному на весь мир. Мощное вибрато и мощная верхняя форманта наверняка способствовали успеху его вокала. Заметим, что ВПФ несколько смещена в сторону высоких частот - для детского организма небольших физических размеров, как мы уже указывали, это естественно.

На рис.8-18 показан спектр звука выдающейся эстрадной певицы, Обертоновый ряд отчетливый, но, видимо, резонаторные полости подверглись

влиянию времени (мы имеем в виду почтенный возраст исполнительницы), о чем свидетельствует и видимое увеличение зоны высокочастотных призвуков.

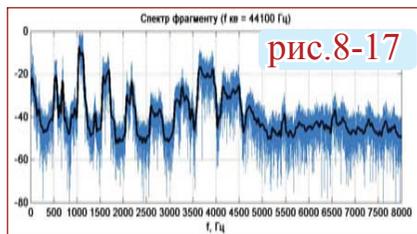


рис.8-17

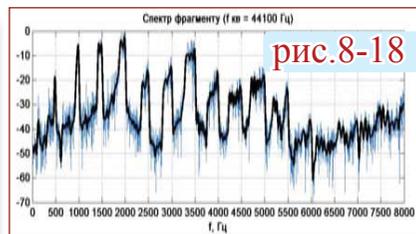


рис.8-18

Рис.8-19 содержит спектр звука весьма молодой (в свое время) участницы певческого соревнования. Она лидировала в этом соревновании, но спектральная картина в данном случае выглядит неубедительно. Нижняя форманта хороша, однако верхняя практически отсутствует. Звук и в самом деле воспринимается как глуховатый, но важно понимать, что по одной ноте, неудачно взятой или специально входящей в художественный замысел, ни в коем случае нельзя судить об исполнителе в целом.

Следующий рис.8-20 относится к молодой и неопытной исполнительнице, не имеющей вокальной подготовки. Спектр ее звука отличает обилие качественных обертонов, охватывающих и весьма расширенную область верхней форманты. Столь качественная форма обертонов здесь объясняется (это было подтверждено исследованием) полным отсутствием в звуке вибрато, которое, как мы указывали, расширяет форму обертонов при линейной шкале оси частот. Спектральный материал, несмотря на неканонический вид ВПФ и видимый подъем в высокочастотной зоне, подтверждает слуховое впечатление - голос певицы обладает "плотным" насыщенным тембром (меццо, близкое к контральто) и заслуживает систематической профессионализации под руководством опытного наставника.

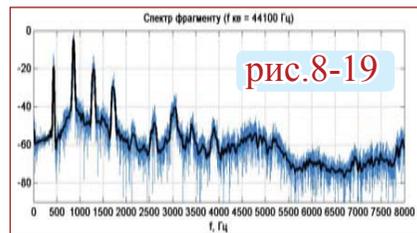


рис.8-19

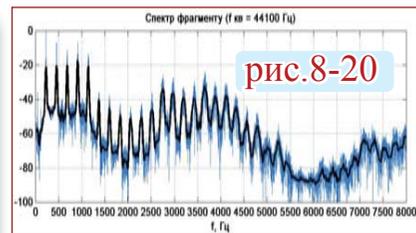


рис.8-20

Такого рода выводы являются одним из аргументов в пользу объективного анализа певческих голосов программно-инструментальными методами.

7. Последние два рисунка (рис.8-21,8-22) посвящены экзотическим вариантам пения. В верхней части рис.8-21 показана развертка во времени

звука от американского джазового вокалиста. Это редкий случай глубокого вокального амплитудного вибрато (см. также рис.6-5), которое настолько же непривычно воспринимать на слух, насколько, вероятно, и тяжело исполнять. Неудивительно, что при столь напряженной манере пения спектральная характеристика сильно отличается от характеристик классических вокалистов, хотя наличие верхней форманты свидетельствует о природных данных, присущих певцу.

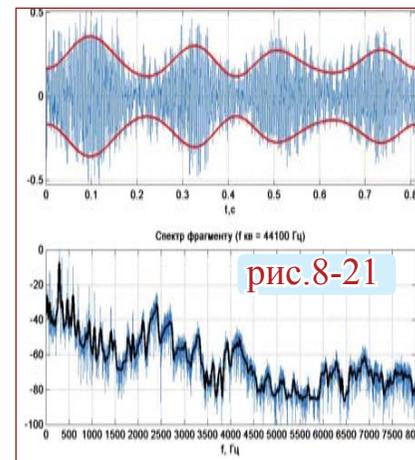


рис.8-21

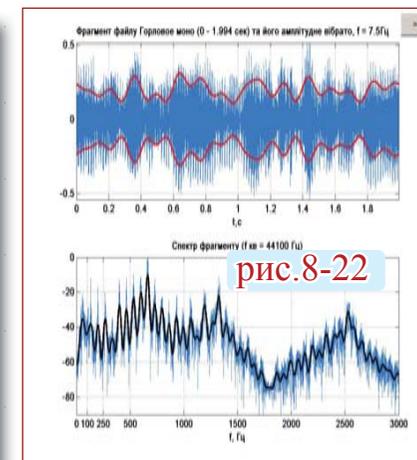


рис.8-22

На рис.8-22 отображены временная развертка и спектр звука при так называемом "горловом" пении. Оно представляет собой сложный тип звукоизвлечения, производящий впечатление очень низкочастотного интонированного хрипа, сквозь который прослушиваются более высокочастотные обертоны. Разумеется, ничего общего с классическим вокальным исполнительством такой вариант не имеет, но и ему соответствует обертоновая структура с энергетическими подъемами, которые и в такой манере указывают на наличие резонаторов в голосовом тракте.

9 Режиссура речевой деятельности

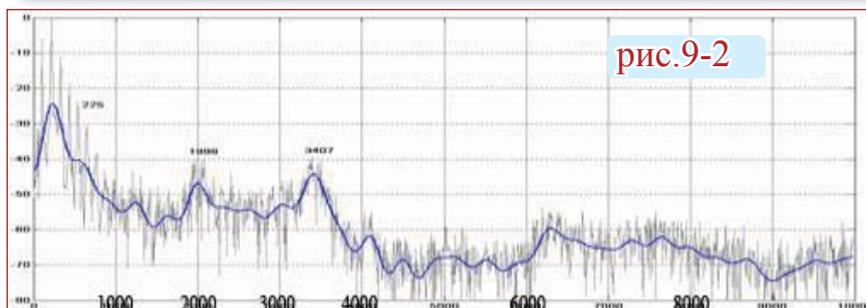
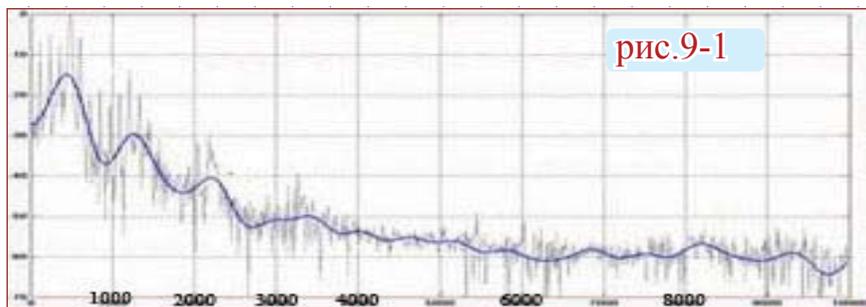
Речевые форманты

1 Мы не включили в название главы слово "акустика", хотя начнем разговор именно с акустических свойств элементов речи. Дело в том, что основной смысл речевой деятельности это обмен информацией. Важна ли эта информация, или она является лишь фоном для события, речь источника воспринимается слушателем, который извлекает смысл из иногда плохо организованного и плохо произнесенного текста.

Поэтому при соблюдении некоторых простых правил, относящихся к акустике, заботы звукорежиссера носят преимущественно редакторский характер, если у него есть такие полномочия.

2. Речевым гласным также присуще наличие формантных областей, но их характер и положение на оси частот различно для различных гласных звуков и, разумеется, для различных особенностей произношения, которое нередко затрудняет разборчивость речи.

На рис.9-1,9-2 приведены речевые спектра ударных гласных "А" и "И" из интервью известного российского певца. Мы выбрали речь певца, полагая, что его артикуляция наименее подвержена дефектам.



В "таблице 3" приведены среднестатистические значения центров формантных областей для классической русской речи.

Нетрудно видеть, что энергетические подъемы на рис.9-1,9-2 близки к указанным в таблице значениям.

Таблица 3. Канонические значения формантных частот гласных звуков, герц.

звук	F_1	F_2	F_3
А	600	1200	2400
Э	450	1700	2500
О	450	800	2500
У	350	750	2200
Ы	300	1800	2450
И	300	2200	3000
Е	400	1850	2600
Я	450	1500	2450

Ниже приведена аналогичного типа таблица для украинской неторопливой речи (колонки 1), речи в среднем темпе (колонки 2) и скороговорки (колонки 3)¹. Мы не приводим литературных ссылок, поскольку предметом этой книги является не исследование собственно фонетических особенностей речи, а указания звукорежиссеру частотных областей, особенно чувствительных для артикуляции гласных звуков.

Центральные частоты формант голосных звуков украинської мови

Звуки	F_1			F_2			F_3			F_4		
	1	2	3	1	2	3	1	2	3	1	2	3
[a]	670	670	630	1300	1300	1300	2200	2200	2290	3540	3540	3540
[e]	430	430	450	1550	1570	1600	2300	2300	2300	3560	3550	3590
[u]	360	380	330	2300	2300	2300	2700	2700	2750	3500	3500	3450
[i]	280	270	250	2320	2310	2350	2970	2970	2970	3600	3600	3600
[o]	450	450	420	750	740	730	2410	2390	2380	3460	3480	3450
[y]	370	370	370	630	630	640	2480	2480	2470	3550	3550	3500

В отличие от пения в спектрах речевых гласных усматривают четыре формантные области (F_1, F_2, F_3, F_4), из которых (как это видно в таблице 3) в основном

¹ Здесь имеется в виду разговорная манера носителей литературных русского и украинского языков.

считаются лишь с тремя первыми (F_1, F_2, F_3) но иногда и среди них в литературе ограничиваются лишь двумя первыми.

Комментируя приведенные таблицы, заметим, что центральные частоты формантных областей в обоих языках близки между собой. Однако носители литературного украинского языка, как правило, хорошо владеют и русским языком, что, по видимому, сближает артикуляционную манеру речи. В этом смысле, вероятно, было бы любопытно сопоставить соответствующие материалы, относящиеся к различным географическим регионам как в России так и в Украине.

Роль формант такова, что в хорошо артикулированном звуке можно с помощью эквалайзера, подавив одни энергетические подъемы, и увеличив в соответствии с таблицей другие области, приблизиться к иному гласному звуку. Разумеется, он может звучать не лучшим образом, потому что голосовой аппарат данного конкретного "пациента" не обязательно строго подчиняется приведенной здесь таблице, но адекватность нового гласного звука будет заметна.

Еще раз напомним что такое форманта. Формантой называют как область энергетического подъема в спектре звуков, обусловленную физической природой резонаторов, входящих в тракт формирования звука, так и те гармоники (обертоны) звука, которые попадают в эту область резонансного усиления. Мы по контексту всегда можем понять о чем идет речь. И если возникает желание применить слово "форманта" к области стабильного подъема уровня в спектре звука не голоса, а некоторого музыкального инструмента, что обусловлено его формой, то это не вызывает каких-либо возражений.

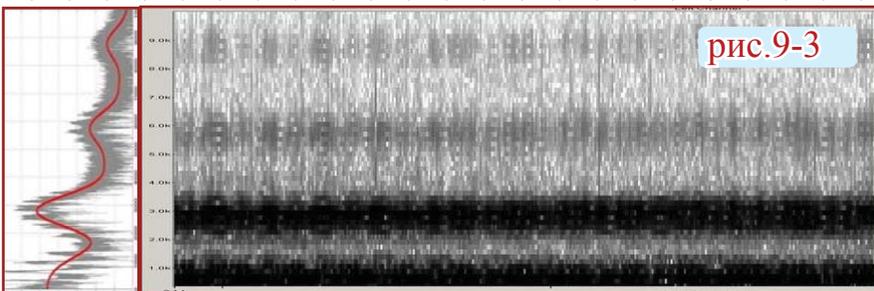


рис.9-3

Рисунок 9-3 иллюстрирует понятие форманты. Он организован так, что ось частот в левой спектральной картинке направлена вверх - так же как и в спектрограмме в правой части этого рисунка. В этой спектрограмме ось времени ориентирована горизонтально и сильно сжата. Видно, таким образом, что развивающийся во времени сигнал, гармоники и призвуки которого изменяют свои частотные параметры, попадая в зону формант, всегда усиливаются (почернение в одноцветной спектрограмме). Положение темных зон

формантного усиления неизменно, а их "содержание" меняется по мере звучания сигнала.

Обратив внимание на положение речевых формант в таблице "Таблица3", можно согласиться с мнением, что при профессиональном пении элементы голосового тракта занимают такое положение, что речевые форманты F_1 и F_2 примыкают друг к другу, образуя нижнюю певческую форманту, а верхняя певческая форманта, видимо, образуется на базе речевой форманты F_3 . Косвенным свидетельством этого служит тот факт, что словесное наполнение певческих арий зачастую трудно разобрать - гласные звуки в ариях по артикуляции приближаются друг к другу и опора в распознавании слов арии ложится в значительной мере на согласные звуки, в которых существенную роль играют звонкие согласные "З, Ц, С".

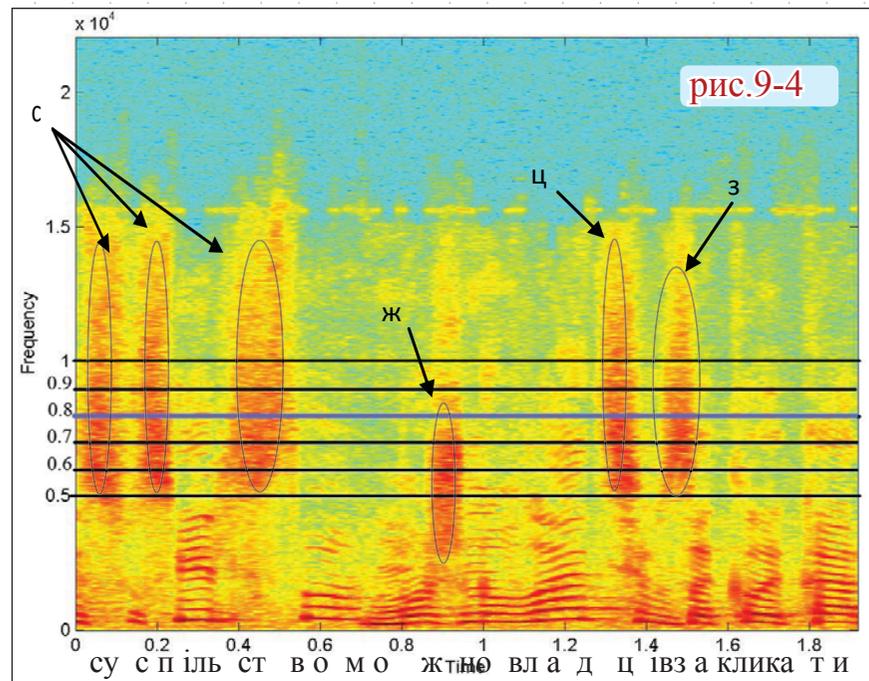
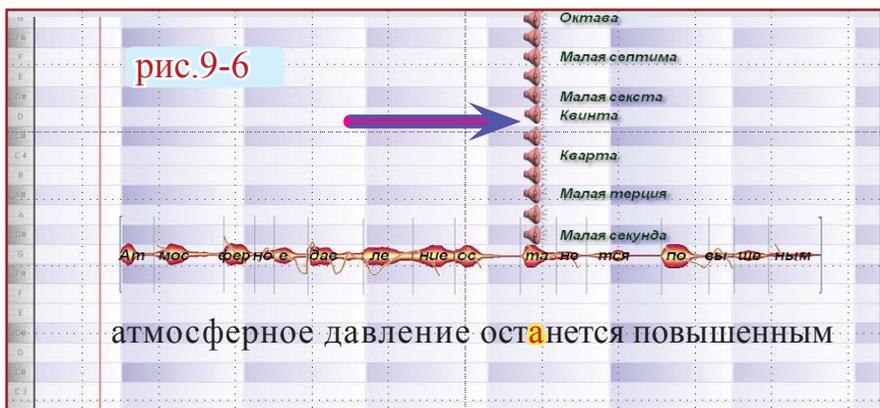


рис.9-4

Звонкие речевые согласные

2. На рис.9-4 приведена спектрограмма речевой фразы "супільство жовтої влади кліцтво" - фрагмента телевизионного новостного сообщения. Фраза хорошо артикулирована и на спектрограмме отчетливо видны следы как согласных звуков так и гласных, представленных обертоновой "гребенкой" (напомним- на спектрограммах ось частот сориентирована вер-

ге порядка квинты. В этом случае внимание слушателя сосредоточивается именно на факте стабильного состояния атмосферного давления.



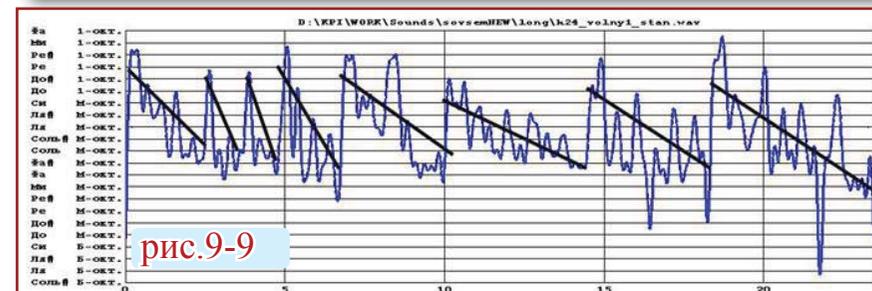
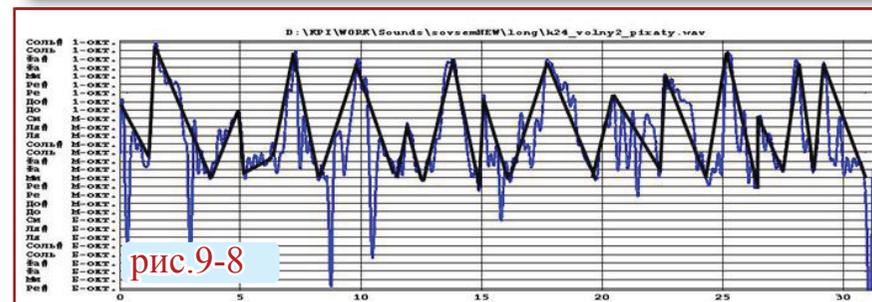
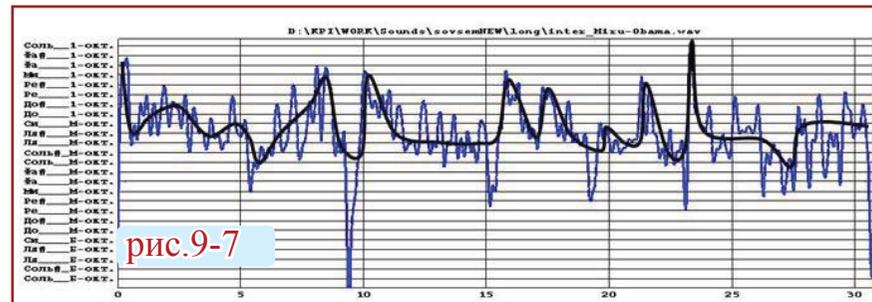
Заметим, что выбор слов во фразах для расстановки логических ударений при чтении лектором/диктором текста "с листа" требует существенных навыков. Ряд ведущих дикторов и журналистов обладают такими способностями, однако молодые медиаспециалисты зачастую оказываются неподготовленными к интеллектуальной подаче текста слушателям.

Мы полагаем, что звукорежиссер кроме непосредственных технических обязанностей несет еще и определенную моральную ответственность за качество производимого с его участием аудиоматериала в целом. И, заметив несоответствие произносимого текста логике содержания этого текста, может предложить принять участие в его подготовке к записи, помогая акцентировать наиболее значимые слова квинтовыми интонационными движениями.

Интонационные шаблоны в речи

7. На рис.9-7..9-9 показаны интонационные развертки (в координатах: время - высота тона в нотных обозначениях) трех примерно 30-секундных информационных сообщений от трех различных дикторов (так мы назовем любые источники речевого аудиоматериала, в том числе лекторов, теле- и радиожурналистов, политических комментаторов и т.п.). Быстроменяющиеся кривые показывают реальное их интонационное речевое поведение, а черные контуры спрямляют реальную картину, указывая тенденции этого поведения. Первая развертка принадлежит студийному диктору высокой квалификации, две последующие - выездным телерепортерам.

Нетрудно увидеть, что квалифицированный диктор разнообразит интонационную подачу материала. Весьма условно можно сказать, что профес-



сионал "поет" произносимый текст. В то же время менее подготовленные дикторы подвержены шаблонам интонационного произношения, очевидным из систематического расположения интонационных линий. Зачастую эти шаблоны порождены бытовыми разговорными манерами, малозаметными в обиходе. Но при передаче информации для массового слушателя следование таким шаблонам нарушает желательную расстановку логических фразовых ударений и существенно снижает качество восприятия содержательной части сообщений.

Приведенные примеры далеко не исчерпывают варианты интонационных шаблонов, которые также не должны ускользать от внимания звукорежиссера, взаимодействующего с дикторским персоналом

Паразитное заполнение пауз



8. Рисунок 9-10 представляет развертку во времени монолога длительностью 45 секунд широко известного телеведущего.

Зелеными полосами на развертке показаны фрагменты произносимого текста, заполненные паразитными частицами "Эээ...", "Ааа...", "Ммм..." и т.п. Эти полосы не являются паузами в собственном смысле слова, пустые паузы на этой развертке не отмечены - здесь же мы видим именно подачу паразитных голосовых сигналов, которые совершенно ничего не сообщают слушателю, но на психологическом уровне удерживают его внимание.

Этот дефект речи распространен повсеместно. Редкий лектор, тележурналист, политический или научный деятель способен связно донести излагаемый материал, особенно в тех случаях, когда оказывается перед микрофоном в прямом эфире, что иногда приводит людей к легкой форме нервного стресса. Сознание необратимости произнесенного текста вызывает напряжение в подборе подходящих слов, и ощущение необходимости чем-то заполнить время такого подбора.

Как раз телеведущий, о котором здесь шла речь, вряд ли нуждается в поиске слов для своего выступления. Скорее всего, обилие паразитных заполнителей у него является рефлекторным поведением, сложившимся от бесконтрольности благодаря его постоянному профессиональному успеху. Такие причины являются еще одним важным фактором - привычка и бесконтрольность делают этот дефект речевых технологий повсеместно допустимым.

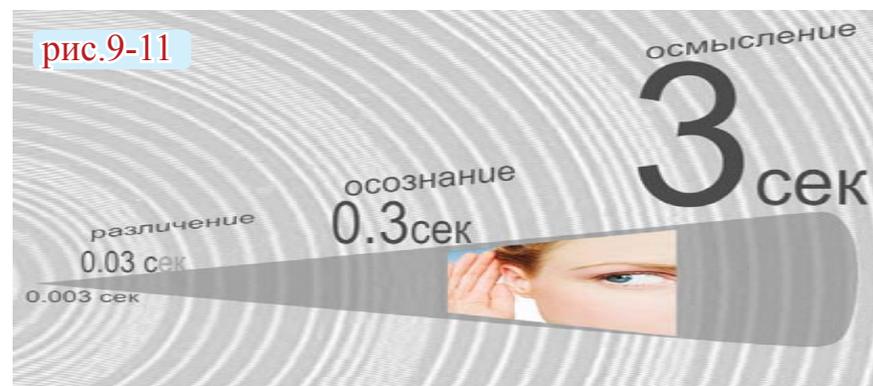
Нетрудно увидеть, что заполнители в зеленых полосах на рис.9-10 обладают весьма значительной энергией и иногда длительностью, и на слух они по громкости ничем не отличаются от осмысленной части текста. Для слушателя, заинтересованно старающегося осознать смысл произносимого, такие сбои являются раздражающим обстоятельством и существенно снижают общее впечатление о звуковой части общего медиапродукта.

9. Важно знать, что указанный в этом параграфе дефект речи может быть ликвидирован. Это не теоретическая декларация. Это проверено реальным достоверно известным опытом.

Для того, чтобы приобрести такой навык требуются два условия.

Первое желательное условие невозможно выполнить быстро. Речь идет о свободном владении "словарем" родного языка с пониманием оттенков употребляемых слов, их синонимов и синонимических выражений. Такое владение достигается только в результате внимательного чтения бумажных книг на протяжении длительного времени. Наградой становится автоматический выбор слов для выражения своей мысли без запинок и каких-либо популярных заменителей - "типа", "как бы" и т.п.

Второе условие связано с пониманием закономерностей восприятия произносимого речевого материала.



На рис.9-11 указаны условные границы времени, которые характеризуют процесс слухового восприятия человека: 3 мсек, 30 мсек, 300 мсек (0.3 сек) и 3 сек. Эти границы весьма условны и называют лишь примерный порядок величин, на которые следует ориентироваться.

В пределах интервалов длительностью до 3 мсек слух не успевает отреагировать на содержание физического звукового стимула, поэтому, в частности, время срабатывания аудиокомпрессоров стараются выбрать существенно меньше этой величины.

В пределах до 30 мсек слух улавливает степень сложности звукового сигнала, но не в состоянии разобраться в этой сложности. Приблизительно в таких пределах смесь многократных "первых" отражений звуковых волн от поверхностей внутри закрытого помещения складывается с исходным "прямым" сигналом и создает психофизиологический эффект его усиления и уплотнения.

Звуки, воспринимаемые на протяжении примерно 300 мсек обрабатываются сознанием так, что слушатель вполне отдает себе отчет в том, что именно он слышит - он, например, может не только сформулировать был ли это отрезок шума или интонированный звук, но и (для людей с абсолютным слухом) назвать звучащую ноту.

Для звукового сигнала длящегося примерно 3 секунды можно уже говорить об обработке информации. Содержание такого сигнала должно осмысливаться сознанием, то есть могут формироваться первичные смысловые ассоциации, возникать аналогии и противопоставления. При таких коротких отрезках времени, разумеется, речь не идет о каких-либо размышлениях, но базовые эмоциональные и логические реакции уже возникают. Осмысление требует небольшого но самостоятельного перерыва в восприятии, свободного от слуховой нагрузки. Эти данные были получены зарубежными исследователями на основе усредненного анализа оптимального времени звучания строк поэтических текстов, но именно поэтический материал с его нетривиальным метафорическим изложением позволяет судить о длительности порции слухового материала, после которого следует дать сознанию кратковременную паузу для обработки услышанного.

10. Указанная величина в 3 секунды важна в контексте затронутой в этом параграфе темы. Разумеется длительность речевых фрагментов как и длительность пауз между ними зависят от сложности излагаемого материала и подготовленности слушателя. В бытовой скороговорке, когда люди торопятся говорить, перебивая друг друга, эти факторы малозначимы. Но в ответственном мероприятии человек, которого мы здесь условно называем диктором, должен строго следить за порциями своей речи размером в несколько секунд (например от 2 до 5 в зависимости от сложности и важности текста), и делать паузы (например от полусекунды до 2 секунд), причем в паузах он должен совершенно замолкать, давая возможность слушателю автоматически осмыслить сказанное.

Рецепт состоит, таким образом, в том, чтобы сознательно волевыми усилиями подчинить себя алгоритму "говорю - молчу - говорю - молчу - говорю - молчу -...").

И привыкнуть к этому не только возможно, но и возможно за относительно короткое время, сохранив эту привычку на всю оставшуюся жизнь.

Послесловие

К счастью мне все меньше приходится убеждать молодых коллег в полезности освоения акустических закономерностей для их аудиодейтельности. Ранние (в прошлом) стадии пренебрежения к академическим знаниям и попытки заменить их простейшими умениями, как мне кажется, сегодня отстают. Студенты начинают видеть, что роскошное компьютерное и аппаратное обеспечение, несмотря на обещания менеджеров по продажам, само по себе не только не гарантирует, но даже и не позволит добиться заметных профессиональных успехов без понимания базовых научных и инженерных основ, заложенных в современную звукорежиссуру.

Именно осознание того, что все богатство орудий твоего труда подчиняется не тебе, а законам, формулам, графикам, понимание которых дает тебе возможность управлять этими орудиями по своему усмотрению, может превратить простого "оператора передней панели" в свободного творца, называемого звукорежиссером, результаты работы которого в большей мере зависят от его таланта, чем от кнопки на передней панели прибора.

*Смотреть в будущее страны с надеждой -
это смотреть в ваше будущее, дорогие читатели.*