



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

Математика для Звукорежиссеров

Анатолий Ананьев



Μαθηματικά
для звукорежиссеров

Анатолий Ананьев



ВЫПУСК № 2

Киев 2014

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.

А64 Математика для звукорежиссера. Выпуск 2. 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.

© А.Б. Ананьев, 2014

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

Содержание

Введение : Спектры звуковых сигналов	5
1 Комплексные числа	6
2 Комплексная экспонента	9
3 Спектральное представление сигналов	13
4 Как работает спектр	18
5 Что мы пропустили	20
6 Заключение	22

Введение : Спектры звуковых сигналов

Мы говорим слово "спектр" настолько часто, что просто неприлично не понимать в точности, что именно имеется в виду.

Для специалиста в обработке звуковых сигналов спектр - первый и главный инструмент, с которым он вторгается в структуру звука. Спектры для сигналов, частотные характеристики для электроакустической аппаратуры. Люди, которые пытаются действовать "на слух" и плохо разбираются в этих "документах на звук", вряд ли могут считаться профессионалами.

Я спрашиваю себя - когда возникает сложный, а иногда и хитрый, вопрос, связанный со свойствами слышимого звука, должны ли люди бежать с просьбой разобраться в этом вопросе к звукорежиссеру, или наоборот - звукорежиссер должен бежать с такой просьбой к людям?

Спросите и вы себя.

Тех, кто будет пользоваться этой книгой, прошу уважать мои авторские права. Разрешаю копировать материал только для личного пользования, в случае использования фрагментов документа прошу указывать источник и автора.

1 Комплексные числа

Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk.

Так написал немецкий математик Леопольд Кронекер (1823-1891). В переводе это значит "Цифры (числа) созданы любовью Бога, всё остальное — дело рук человека".

1. Математика удивительная область деятельности, очень близкая к музыкальной. Различие между ними состоит, вероятно, в том, что математики зачастую не в состоянии правильно спеть мелодию, а музыканты ничего за пределами кассы не могут правильно сосчитать. Все остальное - сходство.

Главное сходство в том, что, не заимствуя почти ничего из окружающей природы, математики и музыканты одной лишь силой воображения создают новые миры, подчиненные внутренней логике. И эти миры, если они доступны другим людям, начинают играть в жизни этих "других" очень большую роль.

Я не знаю, что такое "Бог", но наверно это должно быть что-то похожее на математиков и музыкантов.

Само по себе появление целых чисел кажется мне чудом. Ну, посудите сами, в процессе развития человечества человек развился до такого состояния, что почувствовал возможность отделить понятие количества от количества конкретных предметов - "три стула, пять конфет". Мы говорим просто "три" и уже понимаем, что мы сказали.

С этого места начинается математика.

2. То, что Кронекер назвал просто числами (элементами счета), сегодня называют натуральными числами $\{1, 2, 3, \dots\}$. Они позволяли выполнять операцию сложения чисел из этого ряда.

Развитием этого ряда стали целые числа $\{-3, -2, -1, 0, 1, 2, 3, \dots\}$, которые позволили, кроме того, выполнять операции вычитания и умножения любых чисел из этого ряда.

Рациональные числа представимы в виде M/N , где M - целое число, а N - натуральное. Рациональные чаще оказываются дробными числами. Множество рациональных чисел позволяет выполнять все алгебраические операции над ними.

Множество действительных (или по-другому называемых вещественных) чисел дополнительно к рациональным включает в себя т.н. иррациональные числа, которые не могут быть представлены в виде отношения M/N , а значит в любом цифровом представлении имеют бесконечное количество цифр. Таким, например, является значение диагонали квадрата с единичными сторонами $\sqrt{2} \approx 1.41421356\dots$

Математические задачи в значительной мере связаны с решением уравнений. Линейные уравнения благополучно решаются в пределах множества рациональных чисел (например $2x + 3 = 0$). Но квадратные уравнения в школьной программе иногда объявляют "не имеющими решения".

Таким уравнением, к примеру, является

$$x^2 - 2x + 10 = 0 \quad \text{ф.1-1}$$

решение которого имеет вид

$$x_{1,2} = 1 \pm \sqrt{1-10} \quad \text{ф.1-2}$$

или, иначе $x_{1,2} = 1 \pm 3 \cdot \sqrt{-1} \quad \text{ф.1-3}$

Это и есть решение "уравнения, не имеющего решения". Уравнение ф.1-1 не имеет решения среди действительных чисел. Поэтому, чтобы легализовать решения типа ф.1-3, ввели составной формат "комплексных чисел", содержащий $\sqrt{-1}$, и для корня ввели отдельный заменяющий его символ - иногда (чаще у математиков) это символ i , инженеры чаще используют символ j . Таким образом решение ф.1-3 может быть записано в виде $x_{1,2} = 1 \pm 3j$.

Форму вида $a + j \cdot b$, ф.1-4

где a и b - вещественные (то есть, возможно и отрицательные) числа, $a + j = \sqrt{-1}$,

называют комплексным числом.

Часть числа вида "а" называют его действительной частью, а вида "j·b", которую можно также записывать в виде "b·j" - мнимой частью.

Комплексные числа участвуют в алгебраических операциях по общим правилам. При этом действительные и мнимые части выражений всегда разделены символом j , который, однако, тоже подчиняется алгебраическим правилам с учетом того, что все-таки $j = \sqrt{-1}$:

$$j^0 = 1; j^1 = j; j^2 = -1; j^3 = -j; j^4 = 1; j^5 = j; \text{ и т.д.} \\ \frac{1}{j} = j^{-1} = \frac{j}{j \cdot j} = -j; \quad \text{ф.1-5}$$

3. Математические новшества зачастую находят понимание и объяснение с общежитейской точки зрения, и от этого становятся менее абстрактными. Отрицательные числа воплотились в бухгалтерских "дебетовых" и "кредитовых" суммах. Комплексные числа можно объяснить тем, что мы хотим по каким-то причинам иметь доступ к парам действительных чисел, выполняя

алгебраические действия над парами в целом, но ни в коем случае не выполняя их "внутри" пары над ее элементами.

Заметим, что математика оперирует еще более сложными числовыми структурами, например, т.н. кватернионами, содержащими четыре взаимно независимых вещественных компонента.

2 Комплексная экспонента

1. Вообще-то слово "экспонента" значит "показатель степени". Exponens - с лат. "показывающий", это число, которое пишут возле другого числа ("основания") справа сверху, и оно показывает, сколько раз нужно взять множителем это основание.

Но отдельно еще существует понятие "экспоненциальная функция". Это такая показательная функция, у которой основанием служит число e , названное по имени замечательного математика Леонарда Эйлера. В разговорной манере кое-что сместилось - в экспоненциальной функции e^x формально можно сказать, что x является экспонентой, но так, разумеется, не говорят, особенно учитывая, что в этой функции x уже не обязан быть целым числом, а является значением функции, обратной к показательной, а именно - логарифмической (см. вып. 1). Говорят "экспонента", имея теперь в виду экспоненциальную функцию e^x в целом, которую часто записывают в виде $\exp(x)$.

2. Число $e \approx 2,71828$ - замечательное число. Его происхождение таково.

Представим себе, что некий банк начисляет 100% годовых. Таким образом в конце года мы получим удвоенную вложенную сумму.

Но теперь допустим, что этот добрый банк позволяет кроме того начислять сложные проценты ежемесячно. То есть в конце каждого месяца банк начисляет $1/12$ от годовой суммы, но разрешает на снимать проценты, а накапливать их вместе с основной суммой. Через 1 месяц, таким образом, вклад составит в относительных единицах величину $(1 + 1/12)$. Через два месяца

это будет $\left(1 + \frac{1}{12}\right) \cdot \left(1 + \frac{1}{12}\right) = \left(1 + \frac{1}{12}\right)^2$. Через год $\left(1 + \frac{1}{12}\right)^{12} \approx 2,6130$

и банки это хорошо понимают, поэтому не разрешают накопление процентов при ежемесячном начислении.

Теперь допустим, что начисление соответствующего процента на вклад происходит ежедневно. В этом случае, по аналогии, в конце года сумма будет

составлять $\left(1 + \frac{1}{365}\right)^{365} \approx 2,7146$. Как видим, прирост по сравнению с предыдущим вариантом совсем незначительный.

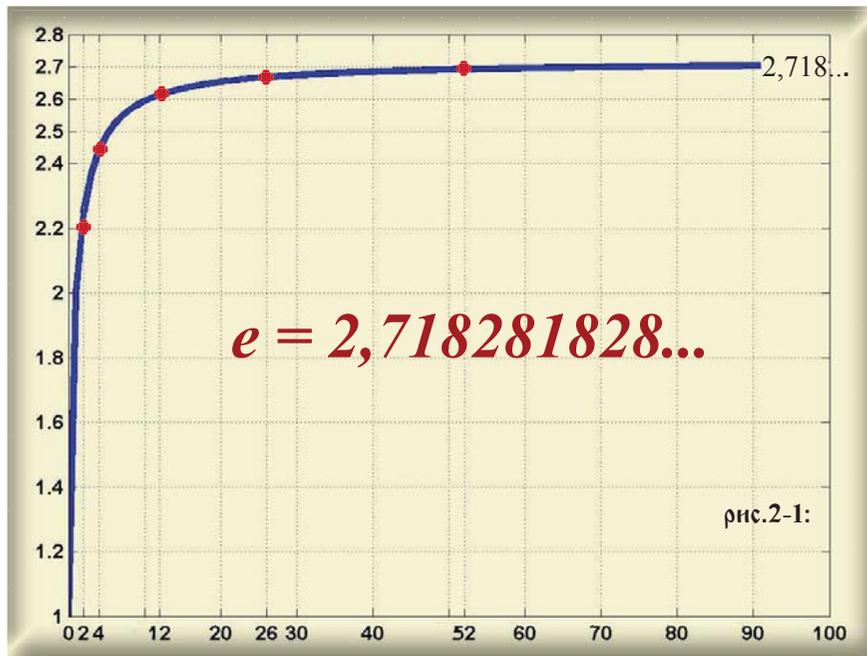
При почасовом накоплении $\left(1 + \frac{1}{8760}\right)^{8760} \approx 2,7181$. При поминутном

$\left(1 + \frac{1}{525600}\right)^{525600} \approx 2,7183$.

Итог в общем виде получается как предел последовательности и имеет значение

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = 2.718281828459045\dots \text{(число } e) \quad \text{ф.2-1}$$

На рисунке 2-1 показан характер приближения последовательности к пределу при увеличении n .



Красными точками отмечены состояния для приведенного ранее весьма условного примера: при полугодовом начислении (2), при поквартальном (4), помесечном (12), каждые две недели (26), еженедельном (52). Как видно, последовательность сходится к пределу довольно быстро.

3. Показательная функция с числом e в основании обладает рядом замечательных свойств. Одно из них важно для нас сейчас.

В вычислительной математике активно используется аппарат разложения функций в степенные ряды. Подробности таких разложений не входят в излагаемый здесь материал. Мы воспользуемся готовыми формулами представлений в виде ряда функции e^x , $\cos(x)$ и $\sin(x)$. Итак:

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \dots \quad \text{ф.2-2}$$

Рассмотрим это разложение для мнимого аргумента jx .

С учетом формул ф.1-5 получим

$$e^{jx} = 1 + \frac{jx}{1!} + \frac{(jx)^2}{2!} + \frac{(jx)^3}{3!} + \frac{(jx)^4}{4!} + \frac{(jx)^5}{5!} + \dots = \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots\right) + j \left(\frac{x}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots\right) \quad \text{ф.2-3}$$

А теперь обратим внимание на то, что разложение в ряд для тригонометрических функций $\cos(x)$ и $\sin(x)$ имеет вид, соответственно:

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \quad \text{ф.2-4}$$

$$\sin(x) = \frac{x}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \quad \text{ф.2-5}$$

откуда очевидно следует (с заменой x на α из уважения к тому факту, что \sin и \cos тригонометрические, т.е.угловые функции):

$$e^{j\alpha} = \cos \alpha + j \cdot \sin \alpha \quad \text{ф.2-6}$$

Это и есть знаменитая и очень важная формула Эйлера, которая связывает показательную экспоненциальную функцию мнимого аргумента с тригонометрическими функциями.

4. Теперь вернемся к формам представления комплексных чисел.

Классическая форма записи комплексного числа $a + j \cdot b$ позволяет сразу опознать его, но не всегда является наиболее удобной. Она называется алгебраической формой записи. Существует также тригонометрическая форма записи того же самого числа, а именно

$$R \cdot e^{j\varphi}, \text{ где } R = \sqrt{a^2 + b^2}, \varphi = \arctg \frac{b}{a} \quad \text{ф.2-7}$$

Нетрудно видеть, что в такой интерпретации в соответствии ф.2-6

$$R \cdot e^{j\varphi} = R \cdot \cos \varphi + j \cdot R \cdot \sin \varphi$$

и, следовательно

$$a = R \cdot \cos \varphi, \quad b = R \cdot \sin \varphi \quad \text{ф.2-8}$$

Разумеется, что алгебраическая и тригонометрическая формы совершенно эквивалентны друг другу, и выбор одной из них - вопрос исключительно удобства пользования.

Итак, благодаря формуле Эйлера, мы имеем равенство

$$a + j \cdot b = R \cdot e^{j\varphi} \quad \text{ф.2-9}$$

где величины a , b , R , φ связаны между собой формулами ф.2-7 и ф.2-8.

Для нас тригонометрическая форма представления не просто удобнее, а, по сути, является единственной формой, в которой мы можем понимать результаты, выраженные с помощью комплексных чисел.

Величина R называется модулем или амплитудой числа, а величина φ - фазой (фазовым углом).

Мы вновь, тем самым, привязываемся к модели данных в виде некоего

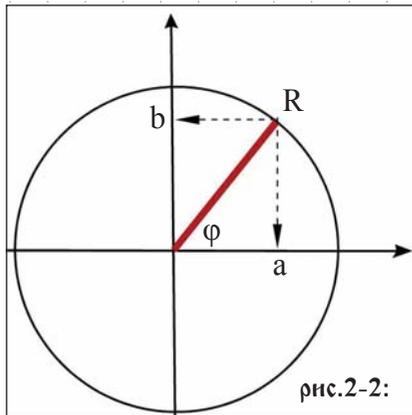


рис.2-2:

вектора длины R , находящегося в декартовой (прямоугольной) оси координат под углом φ к горизонтальной оси (косинусов). При этом вертикальная ось, это ось синусов, так что b/a - это тангенс угла наклона вектора. Но эта модель должна уже стать привычной для специалиста, имеющего дело с сигналами, и во многих случаях она помогает сформировать зрительные образы для выполняемых преобразований сигналов.

Заметим, что

$$R \cdot e^{-j\varphi} = R \cdot \cos \varphi - j \cdot R \cdot \sin \varphi \quad \text{ф.2-10}$$

поскольку $\cos(\varphi)$ четная функция (т.е. $\cos(-\varphi) = \cos(\varphi)$), а $\sin(\varphi)$ нечетная функция (т.е. $\sin(-\varphi) = -\sin(\varphi)$).

3 Спектральное представление сигналов

1. Теперь мы готовы к тому, чтобы обсудить спектральную форму представления сигнала.

Классические формулы спектральной теоремы (теоремы Фурье, разложения Фурье) в экспоненциальной форме имеют вид:

$$s_T(t) = \sum_{k=-\infty}^{\infty} \dot{c}_k e^{jk\frac{2\pi}{T}t} \quad \text{ф.3-1}$$

$$\dot{c}_k = \frac{1}{T} \int_{-T/2}^{T/2} s_T(t) e^{-jk\frac{2\pi}{T}t} dt \quad \text{ф.3-2}$$

Разъясним каждый символ в этих формулах.

2. Греческий символ Σ означает суммирование последующих элементов, в которых "индекс суммирования" k принимает все возможные целочисленные значения, указанные в заголовке - в данном случае от минус бесконечности до плюс бесконечности. Символ $\int \{...\} dt$ - это "определенный интеграл", вычисляющий алгебраическую (с учетом знаков) площадь под выражением, заключенным в скобки $\{...\}$. Точка над символом \dot{c} означает, что c - комплексное число.

Обзор этой части символов носил поверхностный характер, поскольку наше внимание должно быть сосредоточено на иных обстоятельствах, которые мы сейчас будем рассматривать.

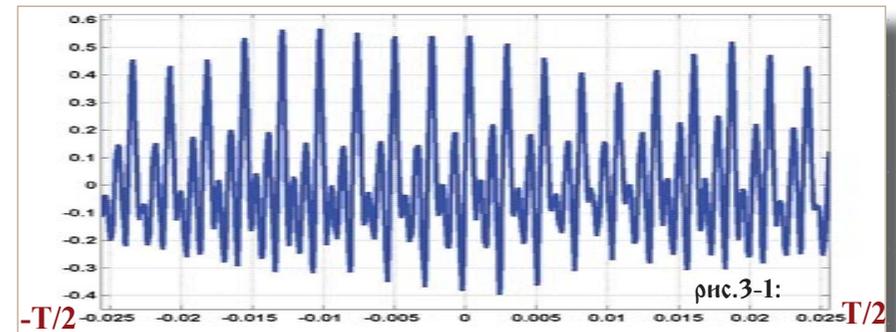


рис.3-1:

На рис.3-1 показан фрагмент записи звука, спетого знаменитой певицей Марией Каллас. Запись расположена на оси времени в соответствии с ф.3-2 так, что нулевая метка времени находится посередине фрагмента. Именно развертка сигнала во времени мыслится как первичный материал для после-

дующего спектрального анализа. Поэтому формулу ф.3-2 называют "прямым преобразованием Фурье" - в ней над сигналом выполняются математические операции, которые приводят к формированию спектра.

Формулу ф.3-1 называют "обратным преобразованием Фурье" - она из спектральных компонент синтезирует развертку сигнала во времени.

3. Рассмотрение начнем с обратного преобразования ф.3-1. В этой формуле видно, что сигнал формируется из суммы комплексных экспонент, которые, как теперь понятно, компактно изображают синусоидальные и косинусоидальные функции. Функции эти домножаются на комплексные коэффициенты \hat{c}_k , вычисленные при выполнении прямого преобразования.

Общая идеология, связывающая формулы преобразований Фурье такова.

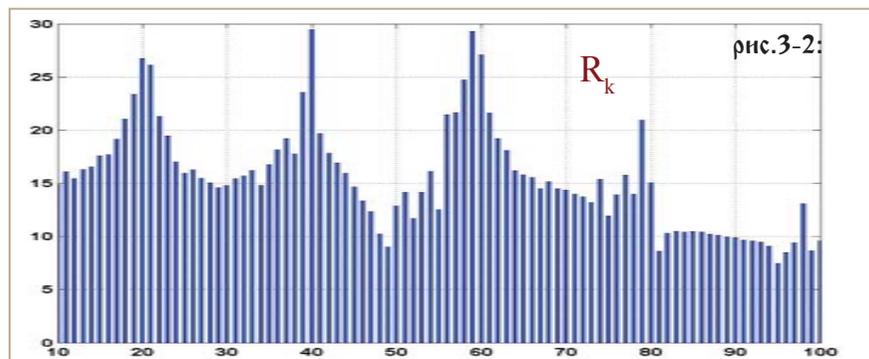
Если образ сигнала во времени вида рис.3-1 не требует дополнительного анализа или обработки, то нет необходимости обращаться к спектральным преобразованиям.

Однако в очень многих случаях именно анализ свойств звука и необходимость изменения этих свойств требуют получения его (звука) спектральных характеристик. В этом случае применение прямого преобразования ф.3-2 позволяет получить коэффициенты $\hat{c}_k = R_k \cdot e^{j\varphi_k}$, которые, вообще говоря, и представляют собой спектр сигнала.

Изменяя эти коэффициенты, мы можем непосредственно влиять на тембровые качества звука, поскольку, в соответствии с акустическим законом Георга Ома, элементарные слуховые ощущения порождаются именно колебаниями синусоидального типа, которыми управляют (см.ф.3-1) коэффициенты \hat{c}_k .

Отрегулировав значения \hat{c}_k нужным образом, мы выполняем обратное преобразование ф.3-1 и получаем измененный звуковой сигнал.

4. Проиллюстрируем сказанное.



На рис.3-2 показаны вычисленные в соответствии с ф.3-2 значения модулей R_k комплексных коэффициентов \hat{c}_k (см.ф.2-9) для небольших по-

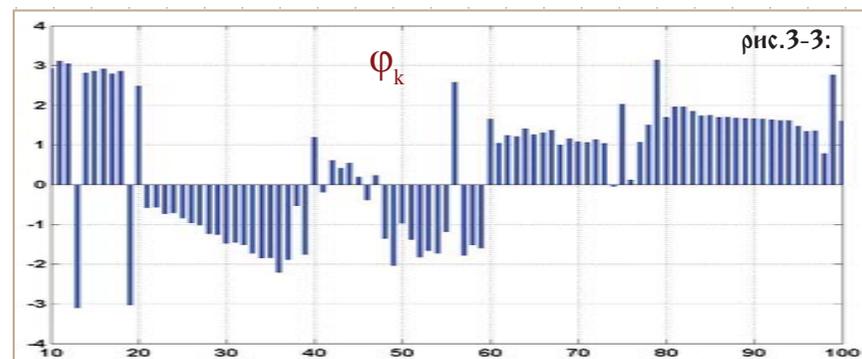
ложительных номеров k. Этот набор модулей, как правило представленный в графическом виде, называется амплитудным спектром сигнала (в данном случае сигнала на рис.3-1).

Эта картинка весьма неважного качества, но иного при таком коротком сигнале (длительность $T \approx 50$ мсек) и нельзя было ожидать. Связь между различными параметрами спектрального анализа мы обсудим позже.

Однако даже по такой картинке можно сделать некоторые суждения о природе и параметрах звука, не видя его развертки и не услышав его:

это музыкальный звук с обертоновой структурой. Хорошо виден основной тон и первые два обертона, последующие два проявляются слабо ввиду малой их энергии. Наиболее сильные обертоны возвышаются над своим окружением на 12..15 dB и создадут отчетливое ощущение интонации. Частота исполняемой ноты соответствует примерно 20-й гармонике спектрального разложения.

5. На рис. 3-3 показаны вычисленные в соответствии с ф.3-2 величины фазовых углов φ_k комплексных коэффициентов \hat{c}_k для тех же значений номеров k. Этот набор, как правило представленный в графическом виде, называется фазовым спектром сигнала. Величины углов на рисунке представлены в радианах и принимают значения в диапазоне $\{-\pi .. +\pi\}$.



Эта картинка, в отличие от предыдущей, малоинформативна. Нельзя сказать, что фазовые соотношения в сигнале не играют роли, но роль эта далеко не так очевидна.

Поэтому аудиоредакторы по запросам спектра звукового сигнала всегда возвращают картину именно амплитудного спектра, усовершенствованную так, как это сейчас сделаем и мы.

6. Давайте повнимательнее посмотрим на аргумент комплексной экспоненты в формуле суммирования ф.3-1:

Параметр k нам известен - это индекс суммирования. Сейчас будем говорить о половине суммы для значений $k > 0$.

$$e^{j \cdot k \cdot \frac{2\pi}{T} \cdot t} = \cos(k \cdot 2\pi \cdot \frac{1}{T} \cdot t) + j \cdot \sin(k \cdot 2\pi \cdot \frac{1}{T} \cdot t) \quad \text{Ф.3-3}$$

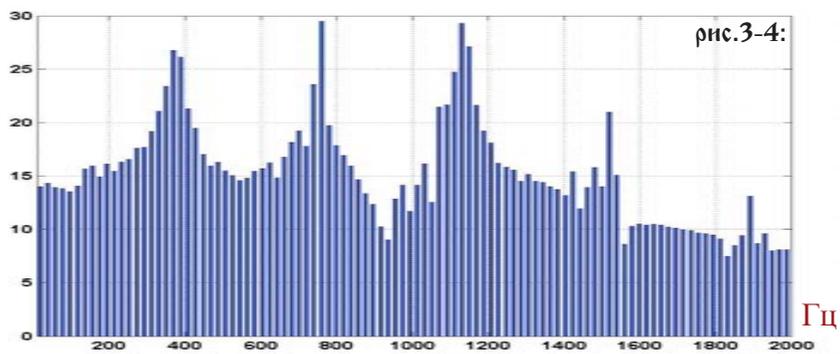
При $k = 1$ тригонометрические функции имеют аргумент $2\pi \cdot f \cdot t$, где $f = \frac{1}{T}$, а значит T при $k = 1$ является периодом этих функций и f - наименьшая частота синусоидальных (косинусоидальных) составляющих спектра, обозначим ее f_1 .

Относительно терминологии заметим, что мы называем все семейство колебаний подобной формы "гармоническими" колебаниями, имея в виду, что начальная фаза у них может быть произвольной.

Однако есть еще термин "гармоники", который подразумевает ряд колебаний синусоидального типа, частоты которых образуют арифметическую прогрессию с начальной частотой f_1 и шагом также равным f_1 , то есть

$$f_1, 2 \cdot f_1, 3 \cdot f_1, 4 \cdot f_1, 5 \cdot f_1, 6 \cdot f_1, 7 \cdot f_1 \dots$$

Спектр сигнала, таким образом, составлен из гармоник, амплитуды которых могут быть отображены на рисунке вида 3-2, но на горизонтальной оси вместо номеров гармоник k следует указать соответствующие значения частоты $k \cdot f_1$.



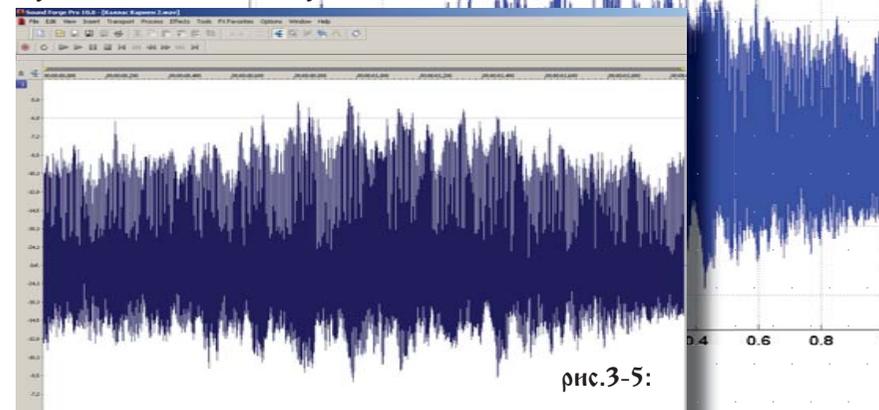
В результате мы имеем спектральную картину, близкую к тем, которые предоставляют нам аудиоредакторы, и можем определить, что исполняется нота с частотой примерно 380 Гц (это чуть выше, чем фа# первой октавы).

Точность, с которой можно определить эту частоту, неудовлетворительна. При длительности сигнала ~50 мсек, наименьшая частота в спектре и шаг по частоте (его называют "разрешающей способностью спектра") равны $1/0.05 \text{ сек} \approx 20 \text{ Гц}$. То есть гармоники, отрисованные в спектре, имеют частоты

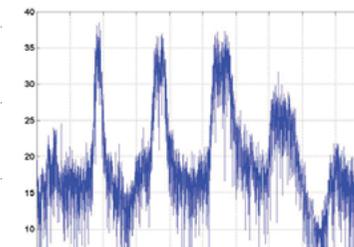
$$20, 40, 60, 80, 100, 120, \dots, 320, 340, 360, 380, 400, 420, 440, \dots \text{ (Гц)}$$

И никакие другие частоты в спектре этого сигнала мы отыскать не можем.

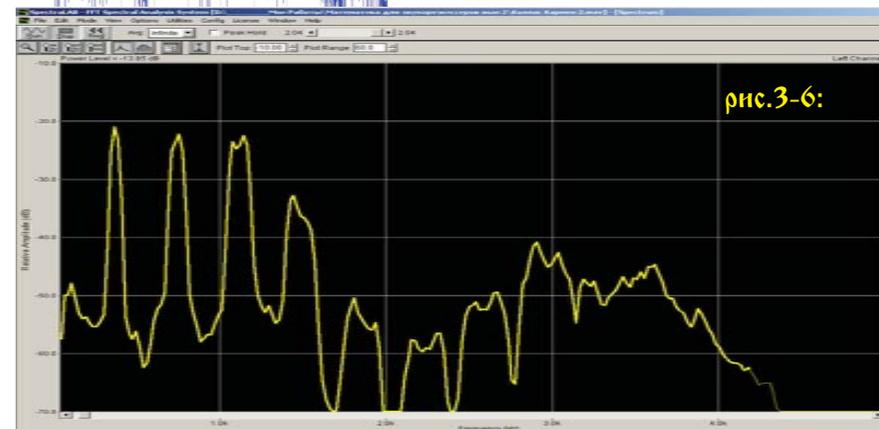
Для более качественного измерения спектра необходимо обеспечить большую длительность исследуемого сигнала.



На рисунках 3-5 показаны более длительные (примерно 2 сек) фрагменты исполнения этой же ноты Марией Каллас. Нижний слой получен в замечательном вычислительном пакете MatLab, верхний принадлежит программе Sound Forge.

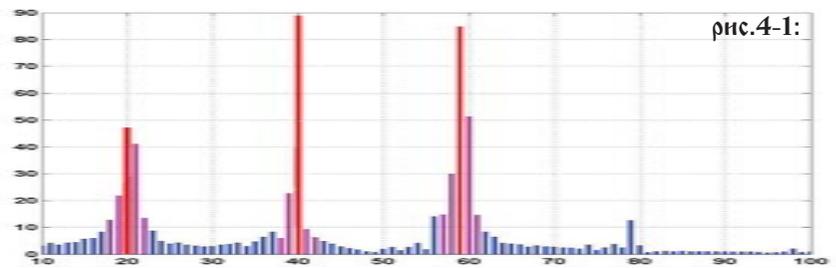


На нижнем слое рисунка 3-6 показан вычисленный с помощью MatLab спектр 2-секундного фрагмента в интервале до 2 кГц, как и на рис.3-4. На верхнем слое приведен спектр этого фрагмента в более широком частотном диапазоне, полученный с помощью программы SpectraLab.



При длительности $T \approx 2$ секунды наименьшая частота и разрешение спектра составляют величины примерно $f_1 = 1/(2 \text{ сек}) = 0,5 \text{ Гц}$, что является абсолютно удовлетворительными характеристиками анализа в данном случае.

4 Как работает спектр



На рис.4-1 вновь показан спектр фрагмента рис.3-1, аналогичный спектру рис.3-2. Но теперь вертикальная ось градуирована не в децибелах, а в натуральных относительных уровнях. Дело в том, что мы собираемся не обсуждать слуховые аспекты этого звука, а проводить математические манипуляции, и нам нужно видеть метрические соотношения между гармониками.

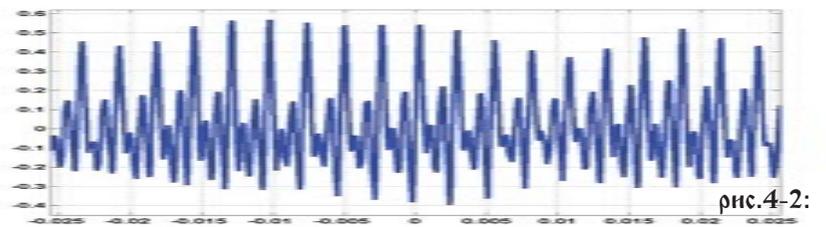
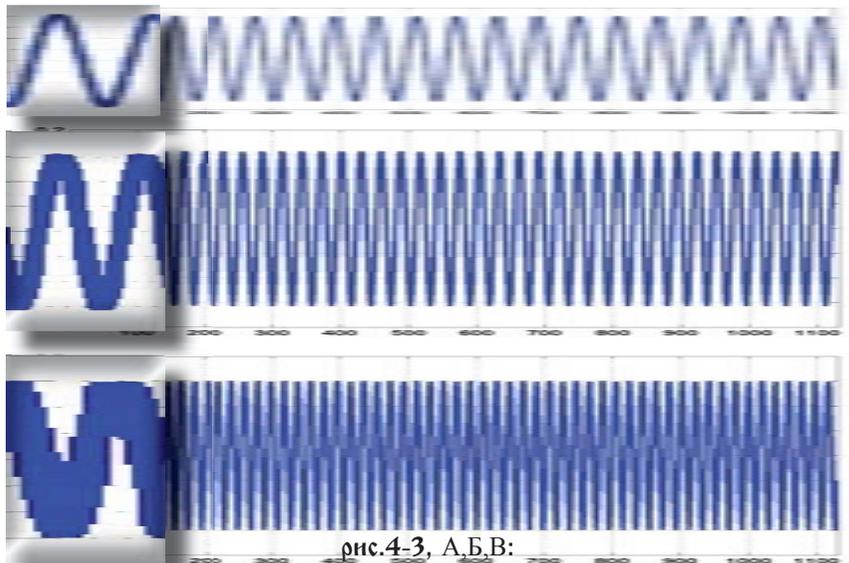


Рис.4-2 повторяет рис.3-1, а ниже располагаются образы наиболее крупных гармоник этого звука, отмеченных на рис.4-1 ярким красным цветом.

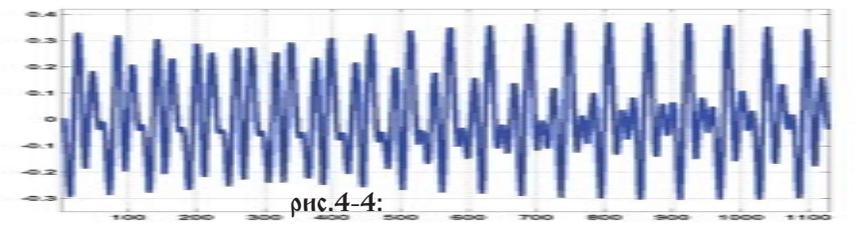


В начальной части рисунков 4-3 показана стартовая часть этих гармоник, где нетрудно видеть, что все они имеют различные начальные фазовые углы φ_k .

Грубость контуров стартовой части объясняется высокой степенью увеличения этой части на рисунках.

Амплитуды гармоник соответствуют значениям R_k в спектре рис.4-1.

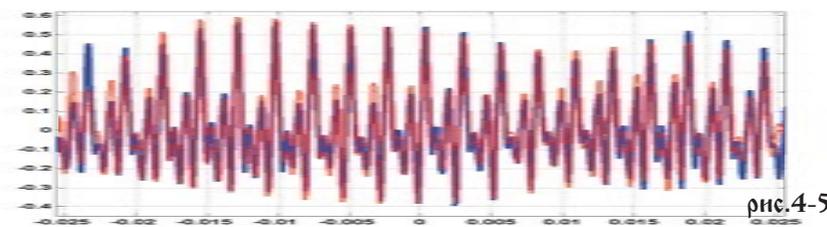
Именно таким образом элементы комплексного спектра \hat{c}_k участвуют в формировании образа звука рис.4-2 на основе 3-х максимальных гармоник с номерами 20, 40, 59 и частотами $20 \cdot f_1$, $40 \cdot f_1$ и $59 \cdot f_1$, где $f_1 = 1/T \approx 20$ Гц.



На рис.4-4 показан результат суммирования этих трех гармоник с учетом их спектральных параметров \hat{c}_k . Безусловно, можно увидеть сходство в поведении этой суммы и исходного звука на рис.4-2, на слух это сходство тоже будет заметно. Однако, разница здесь тоже достаточно ощутима.

Поэтому выберем большее количество гармоник, а именно, по 5 гармоник спектра с центрами в указанных ранее максимумах. Выбранные 15 гармоник на рис.4-1 окрашены в красно-малиновые тона.

Суммирование этой группы гармоник приводит к результату, который на рис.4-5 красными линиями наложен на копию исходного образа рис.4-2.



Здесь очевидна высокая степень совпадения звуковых картин и можно не сомневаться, что слуховые впечатления также будут идентичными

Общее количество спектральных компонент при частоте дискретизации сигнала 22050 Гц составляет $11025 / \sim 20 \approx 550$ элементов. Оказывается, что в данном случае достаточно пятнадцати наиболее энергетически значимых компонент, чтобы почти восстановить форму сигнала. Остальные более слабые компоненты лишь дополняют образ до полного совпадения с исходным.

5 Что мы пропустили

Мы прошли коротким маршрутом по пути, связывающему представления звукового сигнала во времени и в спектральной области. Однако по дороге некоторые вещи обошли молчанием, а некоторые даже не упомянули. Здесь мы кратко охарактеризуем то, что сочли не главным в этом материале.

1. В формуле 3-1 индекс суммирования "пробегаёт" значения от минус бесконечности до плюс бесконечности. Мы рассматривали только положительные k , поскольку доказано, что для реальных сигналов (а не особых математических моделей) значения $R_{-k} = R_k$. То есть в целом $R(k)$ - четная функция. Для того, чтобы увидеть полную картину, нужно рисунок 3-2 зеркально отразить влево относительно вертикальной оси.

Поэтому аудиопрограммы и мы в этой книге не показываем "отрицательной" половины спектра - она неинформативна.

При этом нужно правильно понимать слова "отрицательные частоты" которые нередко встречаются в популярных текстах о спектрах. Разумеется никаких отрицательных частот не существует, значения частоты на спектральной оси образованы величиной $1/T$ сек. Эта положительная частота, умножаясь на отрицательные коэффициенты k из формулы ф.3-2, образует отрицательные величины, которые условно называют "отрицательными частотами" для того, чтобы обозначить принадлежность таких частот к левой (от вертикальной оси) части графика, если эту часть по каким-то причинам обсуждают.

Разумеется, при вычислении обратного преобразования Фурье применяют полный спектр, с положительными и "отрицательными" частотами, соблюдая верность формуле ф.3-1.

2. Формула ф.3-1 содержит бесконечную сумму. Такое представление является формальным и соответствует случаю, при котором анализируемый сигнал содержит мгновенный скачок уровня. Для формирования такого скачка "нулевой длительности" нужно гармоническое колебание бесконечной частоты. Разумеется, что реальные сигналы с которыми имеет дело звукорежиссер являются значительно более плавными и вычисленные для них спектры содержат вполне конечное число гармоник.

3. Мы не обращали внимания на фазовый спектр, а лишь воспользовались его элементами, когда в соответствии с ф.3-1 суммировали некоторые спектральные компоненты для получения картины нового сигнала.

Распределение энергии сигнала по частотному диапазону видно из расположения на оси частот компонентов амплитудного спектра, но и фазовый спектр также в ряде случаев играет важную и весьма деликатную роль. Эти случаи находятся за пределами настоящей работы и заслуживают отдельного рассмотрения.

4. В своем исторически первоначальном виде теорема Фурье была изложена не в формате комплексных экспонент, а непосредственно на базе тригонометрических функций для периодических сигналов.

В современном понимании неограниченное продолжение полученного в ф.3-2 спектрального набора гармоник по оси времени автоматически периодизирует исходный отрезок анализируемого сигнала, а экспоненциальное представление входит в красивую теорию из функционального анализа. И мы, не вникая в математические подробности, оперируем результатами современных подходов к спектральной теореме.

5. Повсеместно применяемые цифровые технологии, то есть работа с дискретизованными сигналами в виде файлов, породили несколько иные алгоритмы спектрального анализа. Они называются Дискретное Преобразование Фурье (ДПФ) - прямое ф.5-1 и обратное ф.5-2.

$$\dot{c}_k = \sum_{m=0}^{N-1} s_m e^{-jkm \frac{2\pi}{N}}, \quad k = 0..N-1 \quad \text{рис.5-1:}$$

$$s_m = \frac{1}{N} \sum_{k=0}^{N-1} \dot{c}_k e^{jkm \frac{2\pi}{N}}, \quad m = 0..N-1 \quad \text{рис.5-2:}$$

Разница между этими формулами и ф.3-1,3-2 состоит в том, что здесь не видны параметры времени. Мы берем N дискретных отсчетов сигнала и получаем N комплексных отсчетов спектральных коэффициентов. Комплексные экспоненты подтверждают, что и здесь участниками анализа и синтеза являются тригонометрические функции, правда тоже в дискретизованном виде.

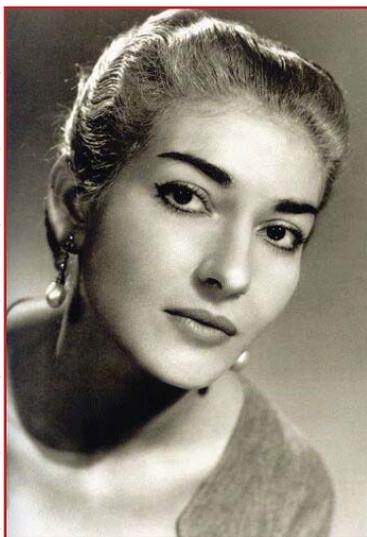
Здесь также явно не видны параметры дискретизации, а правильная дискретизация играет огромную роль в обработке звукового сигнала. Длительность сигнала T по аналогии с формулами ф.3-1,3-2 равна N/f_d , где f_d - частота дискретизации..

Вообще дискретизованные сигналы обладают свойствами, которые тоже заслуживают отдельного обсуждения. Вероятно, такое обсуждение состоится позже. Здесь же отметим, что ДПФ является родственным классическому алгоритму Фурье, но не следует непосредственно из него. Тем не менее, пара ДПФ ф.5-1,5-2 является взаимно-однозначными преобразованиями.

Существует знаменитая вычислительная реализация спектрального метода в виде ДПФ, она называется Быстрое Преобразование Фурье (БПФ, Fast Fourier Transform - FFT). Это очень красивый и очень быстрый вычислительные алгоритм, повсеместно употребляемый в программном и аппаратном обеспечении, о чем в большинстве случаев пользователи даже не подозревают. Все, что сделано и еще будет сделано нами, реализовано с помощью БПФ.

6 Заключение

Упомянув в тексте замечательную оперную вокалистку Марию Каллас, невозможно не познакомить вас с ней хотя бы внешне.



Мы ознакомились со спектральным представлением звуковых сигналов. Это знакомство не охватило многие практические аспекты работы со спектрами. Для тех кому рассказанное здесь показалось интересным и полезным, наш путь не окончен. Это не очень простой путь, но ни в какой серьезной области серьезной деятельности я простых путей не видел.

И могу только повторить:

те, кто хочет иметь право считать себя специалистом по звуку, могут вновь убедиться, что они работают в области знаний, сложность которой во всех отношениях заслуживает уважения.