

**Гармоника** - это синусоидальное колебание:

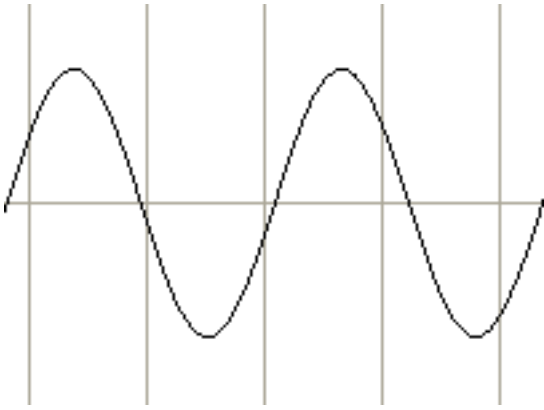
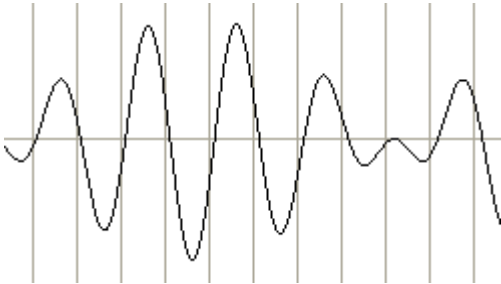
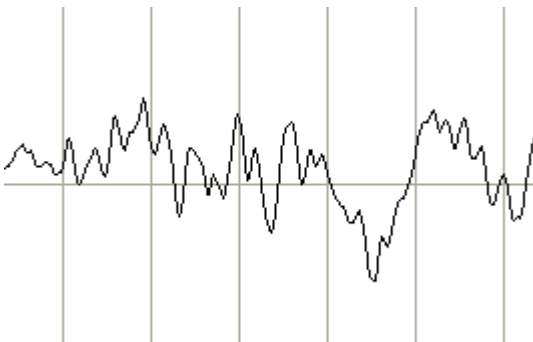


График показывает, как изменяется воздушное давление на барабанную перепонку слушателя: вверх и вниз по дуге, периодически. Воздух давит то сильнее, то слабее. Сила воздействия совсем невелика, и колебания происходят очень быстро: сотни и тысячи толчков каждую секунду. Такие периодические колебания мы воспринимаем как звук.

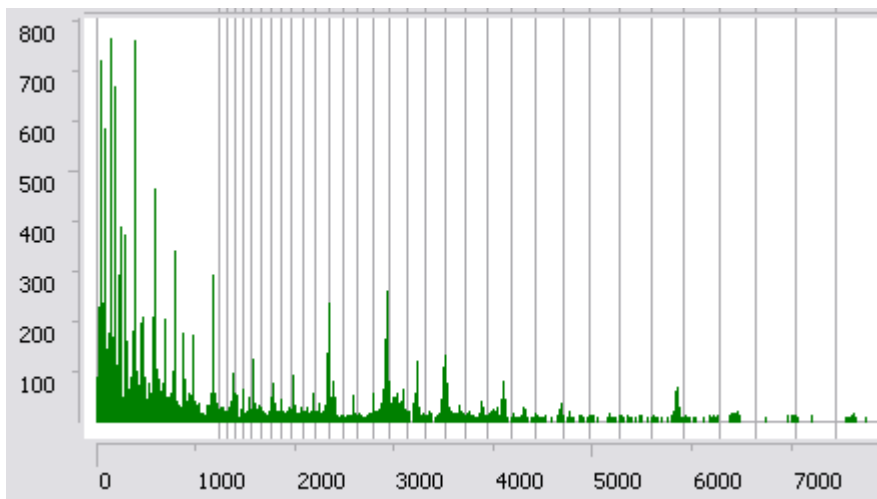
Сложение двух разных гармоник дает колебание более сложной формы:



Сумма трех гармоник - еще сложнее, а естественные, природные звуки и звуки музыкальных инструментов складываются из большого количества гармоник, и в сумме может получиться что-нибудь вроде этого:



В такой сумме может быть неопределенно много гармоник, но среди них заметный (на слух) вклад вносят лишь некоторые, самые сильные (самые громкие). Поэтому при анализе разумно ограничивать число гармоник сотнями, может быть, тысячами. Набор гармоник, выделенных из звукового колебания - это **спектр**. Выглядит он примерно так (по горизонтали - частота, по вертикали - амплитуда гармоники):



Каждая гармоника характеризуется тремя параметрами: амплитудой, частотой и фазой.

**Амплитуда** - это размах колебаний. Чем размах больше, тем сильнее воздействие на барабанную перепонку, и тем громче звук, который мы слышим. Амплитуда может измеряться несколькими способами:

- В абсолютных величинах. Если формат звука 8-битный, то это величина от 0 до 255. Если формат 16-битный, то это величина от -32767 до +32767. Для 32-битных параметров в данном редакторе нормальные величины также заключены между -32767 и +32767, но допускается выход за эти границы.
- В процентах - от максимальной величины (если не оговорено другое): от -100% до +100%.
- В виде коэффициента, в долях максимальной величины (если не оговорено другое): от -1 до +1.
- В децибелах. "**Децибелы** (сокращенно дБ) - это специальная единица, которая показывает усиление или ослабление звука по сравнению с каким-то "опорным" уровнем. Формулы для преобразования коэффициента (k) в децибелы (D) и обратно:

$$D = 10 \log_{10} k$$

$$k = 10^{D/10}$$

Из этих единиц первые три более наглядны, когда смотришь график колебаний. А децибелы ближе к тому, как это воспринимается на слух. Если принять за 0 дБ максимальную громкость в обычных компьютерных наушниках, то -30 дБ - это примерно такой уровень громкости, который едва можно расслышать при регуляторе громкости, выведенном на максимум. При равномерном изменении амплитуды от 0 до -30 дБ, субъективно (на слух) громкость тоже изменяется равномерно, хотя на самом деле - по экспоненте. Это особенность физиологии слуха (закон Фехнера).

**Частота** колебаний - второй параметр гармоники. Частота показывает, сколько толчков давления воздуха происходит за одну секунду. Как обычно, частота измеряется в Герцах (Гц) или килогерцах (кГц). кГц = 1000 Гц.

Человек может слышать гармоники с частотой от 20 Гц до 20 кГц (могут быть небольшие индивидуальные отличия). Этот диапазон называется "звуковым". Все, что ниже - это "инфразвук", все, что выше - "ультразвук". Ультразвук и инфразвук человек не слышит по крайней мере как звуки. Если инфразвук или ультразвук имеет очень большую мощность (громкость), то он может оказывать физическое воздействие на организм. К сожалению, трудно найти достоверные материалы об эффектах такого воздействия, а легенд на эту тему ходит очень много. Кроме того, инфразвуковые или ультразвуковые колебания могут вызывать вторичные колебания каких-нибудь предметов уже в звуковом диапазоне (какое-нибудь дребезжание), и тогда эти вторичные колебания будут слышны.

Большие частоты воспринимаются как "высокие", "тонкие" звуки. Выше 10 кГц - писк, свист. Маленькие частоты воспринимаются как "низкие", "басовые" звуки, рокот. Обычные частоты гармоник мужского голоса - ниже, женского - выше, детского - еще выше.

Кроме частоты, могут применяться другие единицы измерения, которые однозначно связаны с частотой - это период колебаний и высота.

**Период** колебаний (T) - это просто величина, обратная к частоте (f):

$$T = 1 / f$$

**Высота** связана с музыкальной записью, нотами. Высота может быть вычислена по формулам:

$$f = 440.0 * 2^{(12 V + N - 57) / 12}$$

- где f - частота, V - номер **октавы** от 0 до 9, N - номер **ноты** от 0 до 11. Можно выполнить и обратное преобразование из частоты в номер ноты и октавы:

$$12 V + N = 12 * \log_2(f / 440.0) + 57$$

Понятно, что для произвольно заданной частоты f может не найтись целых чисел V, N. Тогда говорят об **отклонении** частоты от ближайшей ноты.

Для понимания многих аспектов работы со звуком знание музыкальных нот необязательно. Достаточно, понимания, что октава и нота - это два числа, два номера, которые по формуле дают частоту. В музыке каждая нота и октава имеют собственное уникальное название, но при вычислениях удобнее работать с номерами.

Обратите внимание, что и здесь равномерное повышение номера ноты вызывает экспоненциальный рост частоты. Это также связано с биологией слуха. При равномерном изменении ноты высота повышается приблизительно равномерно, если судить субъективно, на слух.

При увеличении номера ноты на 1 частота увеличивается в  $Q = (2^{1/12})$  раз. Такое увеличение частоты в Q раз называется "повышением на **полутон**", а уменьшение в Q раз - "понижением на полутон". Изменение частоты в x раз соответствует изменению высоты на  $\log_{Qx}$  полутонов.

Таким образом, полутон - это относительная единица изменения частоты, т.е. она показывает не величину частоты в герцах, а соотношение между двумя частотами. 12-кратное изменение частоты в Q раз дает в сумме 2-кратное изменение частоты. Изменение частоты в 2 раза называют изменением на одну **октаву**. Обратите внимание, что словом "октава" называют и номер для вычисления частот, и относительную единицу измерения частоты.

И есть еще две единицы измерения: **тон** (2 полутона) и **цент** (1/100 полутона).

Единица	Величина
Октава	2
Тон	$2^{1/6}$
Полутон	$2^{1/12}$
Цент	$2^{1/1200}$

**Фаза** - это смещение колебаний во времени (третий и последний параметр гармоника).

Фаза может измеряться в градусах или радианах. В зависимости от фазы смещается нулевой отсчет на графике. Для задания гармоника достаточно указать фазу от  $-180$  до  $+180$  градусов, поскольку при значениях, выходящих за этот диапазон, график гармоника повторяется.

Также нередко "фазой" называют разного рода задержки по времени, и тогда ее измеряют в секундах или миллисекундах. Такого рода задержки по-разному воспринимаются на слух, но здесь уже восприятие примерно соответствует времени, а не логарифмам или степеням. То есть, для восприятия фазы закон Фехнера не работает.

В целом график одной гармоника можно выразить формулой:

$$p(t) = A \sin(2 \pi f t + \varphi)$$

- где  $t$  - время,  $p$  - давление воздуха,  $f$  - частота,  $\varphi$  - фаза,  $A$  - амплитуда.

## Обертон

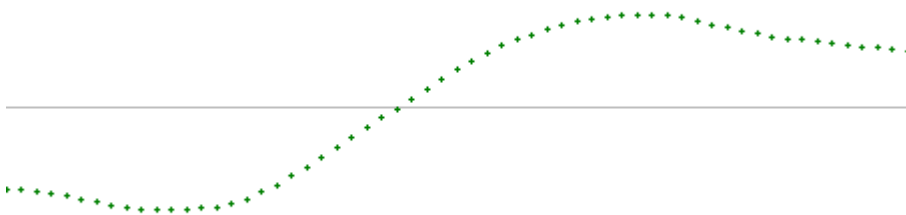
Несколько гармоник с кратными частотами  $f, 2f, 3f, 4f, \dots$  субъективно воспринимаются как единый звук, как звучание одного инструмента. Тогда первая гармоника ( $f$ ) - называется основной (или базовой) гармоникой, а остальные - "обертонами" ( $2f$  - первый обертон,  $3f$  - второй обертон и так далее). Иногда основную гармонику называют нулевым обертоном.

## Тембр

Тембр - это набор гармоник из записи звучания какого-нибудь инструмента. Если тембр состоит из одной гармоника, то звук кажется чистым, но "бедным". Такой тембр имеют гудки в телефонной трубке. Если есть много гармоник, звук более "сочный". Если все гармоники близки по частоте к основной частоте и обертонам (то есть, как бы сгруппированы вокруг них), то такой звук воспринимается, как "нота" определенной высоты, и из таких звуков может образоваться музыка, мелодия. Если гармоники разбросаны хаотично, то такой тембр характерен для ударных инструментов, шумов, шорохов. Из таких тембров нельзя создать полноценную мелодию, можно только задать ее ритмический рисунок.

## Дискретизация

Для обработки на компьютере звуковые колебания подвергаются **дискретизации**. Это значит, что функция, представляющая звуковую волну, строится по точкам, которые идут с некоторым шагом:



Шаги очень маленькие - тысячи раз в секунду. Частота этих шагов называется **частотой дискретизации**. Для единообразия принято несколько стандартных частот дискретизации: 8000, 11025, 22050, 44100, 48000 и 96000 Гц. Из них наиболее распространенная - 44100.

## Преобразование Фурье

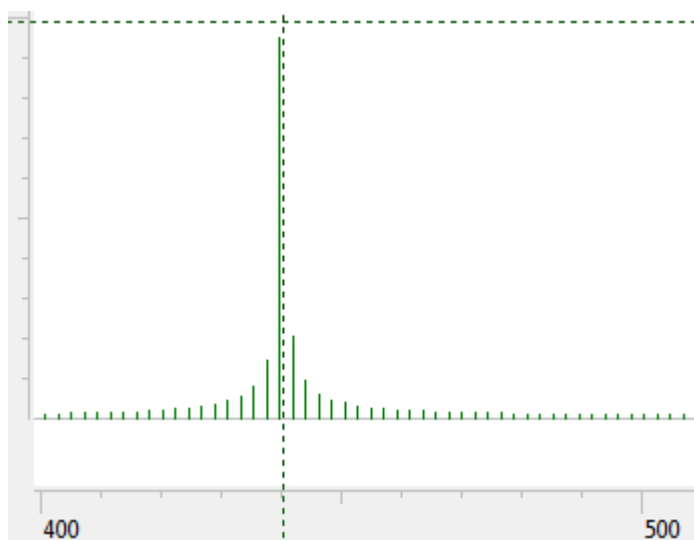
Для того, чтобы получить спектр звукового колебания, существуют разные методы, и самый распространенный среди них - **преобразование Фурье**. В компьютерной технике для него применяется конкретный алгоритм, именуемый "быстрым дискретным преобразованием Фурье", сокращенно **БПФ** или ДПФ (Fast Fourier Transformation = FFT).

Представьте себе, что мы подвергли дискретизации некоторый фрагмент звуковой волны. Пусть частота дискретизации равна  $S$ , и в результате дискретизации получено  $N$  точек функции, представляющей волну. Теперь представьте себе набор частот, идущих с равным интервалом:

$$0, S/N, 2S/N, 3S/N \dots kS/N \dots (N-1)S/N$$

$k$  - номер частоты БПФ.

Теперь можно легко представить себе, что делает преобразование Фурье. Оно рассчитывает "отклик" каждой из этих частот. Если в спектре исходного сигнала есть частота  $kS/N$ , то частота с номером  $k$  "откликнется", а остальные останутся нулевыми. Но так бывает редко. Обычно частоты в спектре исходного сигнала попадают в интервал между соседними частотами БПФ. Например, если частота попадает в интервал между  $kS/N$  и  $(k+1)S/N$ , то "отклик" будет в основном в частотах с номерами  $k$ ,  $k+1$  и (значительно меньше) рядом с ними:



На рисунке выше вертикальными зелеными линиями показана величина "отклика" частот БПФ, а пунктиром - частота в исходном сигнале. Таким образом БПФ позволяет, если и не вычислить спектр исходного сигнала в точности, то хотя бы получить о нем более-менее наглядное представление.

Частоту в исходном сигнале называют **истинной частотой**. Эффект, при котором истинная частота представлена в спектре БПФ группой близких частот, - **эффектом размазывания** или рассеяния. Дополнительные ухищрения помогают его уменьшить, а иногда и вовсе устранить.