

## **Двоичное кодирование звуковой информации. Глубина кодирования и частота дискретизации.**

Кодированию данного вида информации в учебной литературе пока уделяется не очень много внимания. По-видимому, это объясняется тем, что интенсивное использование компьютера для обработки звука началось позднее, нежели чисел, текстов и графики. Приятное исключение составляет учебник [1], который не просто традиционно излагает способ дискретизации звука, но и дополнительно демонстрирует простейшие приемы цифровой обработки звуковых файлов, а так же содержит ряд других сведений, имеющих отношение к данному вопросу. Еще одним источником информации, использованным при написании данных материалов, послужили несколько подробных популярных статей Д. Симаненкова, опубликованных в тематических выпусках журнала «Компьютера» [2].

Итак, рассмотрим интересный и все чаще используемый на практике процесс кодирования звуковой информации подробнее.

Из курса физики известно, что звук есть колебания среды. Чаще всего средой является воздух, но это совсем не обязательно. Например, звук прекрасно распространяется по поверхности земли: именно поэтому в приключенческих фильмах герои, стараясь услышать шум погони, прикладывают ухо к земле. Напротив, существует весьма эффективный школьный опыт, который показывает, что при откачивании воздуха мы перестаем слышать звук находящегося под герметичным колпаком звонка. Важно так же подчеркнуть, что существует определенный диапазон частот, к которому принадлежат звуковые волны: примерно от нескольких десятков герц до величины немного более 20 кГц<sup>1</sup>. Значения этих границ определяются возможностями человеческого слуха.

Благодаря роли звуковых сигналов в практической жизни человека процессы генерации и закономерности распространения звука изучены достаточно хорошо. Чаще всего звуковые колебания преобразуются в электрические, что легко осуществляется в помощью микрофона. Как правило, электрический сигнал от микрофона очень слаб и нуждается в усилении, что на современном уровне развития техники проблемы так же не представляет. Форму полученных колебаний (т.е. зависимость интенсивности сигнала от времени) можно наблюдать на экране обычного осциллографа; к сожалению, для получения наглядной устойчивой картины сигнал должен быть периодическим.

Важную роль в анализе звуковых (или полученных из них электрических) колебаний играет также спектральный анализ, т.е. нахождение распределения интенсивности различных частот в исходном сигнале. Математической основой такой процедуры служит разложение изучаемой функции в ряд по гармоническим функциям (синусам или косинусам) – так называемый «Фурьеанализ». Полученные в результате обработки спектры также обычно представляются графически в координатах «частота (абсцисса) – интенсивность (ордината)». Чтобы представить себе, как выглядит спектр звукового сигнала, достаточно взглянуть на информационный дисплей современного высококачественного аудиоконкомпекса. Спектральные характеристики сигналов и технических звуковых устройств имеют огромное теоретическое и практическое значение.

Наблюдать примерную форму сигналов и их спектральные характеристики на современном компьютере не только не представляет никакого труда, но и напротив, доступно каждому. Сделать это можно. Даже не используя специализированного

---

<sup>1</sup> Интересно сопоставить характерные звуковые частоты с тактовой частотой типового микропроцессора – различие составляет примерно **6 порядков**, что говорит об огромных возможностях компьютера в обработке звуковой информации.

программного обеспечения (примеры такого ПО можно посмотреть в [2]). Достаточно запустить стандартную мультимедийную программу Windows под названием *Звукозапись* и распространенный универсальный проигрыватель звуковых файлов *Winamp*. Первая программа будет показывать форму сигнала, а вторая – его спектр.

Звуковые сигналы в окружающем нас мире необычайно разнообразны. Для их записи с целью последующего воспроизведения необходимо как можно точнее сохранить форму кривой зависимости интенсивности звука от времени. При этом возникает одна очень важная и принципиальная трудность: звуковой сигнал *непрерывен*, а компьютер способен хранить в памяти только *дискретные* величины. Отсюда следует, что в процессе сохранения звуковой информации она должна быть «оцифрована», т.е. из аналоговой непрерывной формы переведена в цифровую дискретную. Данную функцию выполняет специальный блок, входящий в состав звуковой карты, который называется *аналого-цифровой преобразователь – АЦП*.

Каковы основные принципы работы АЦП?

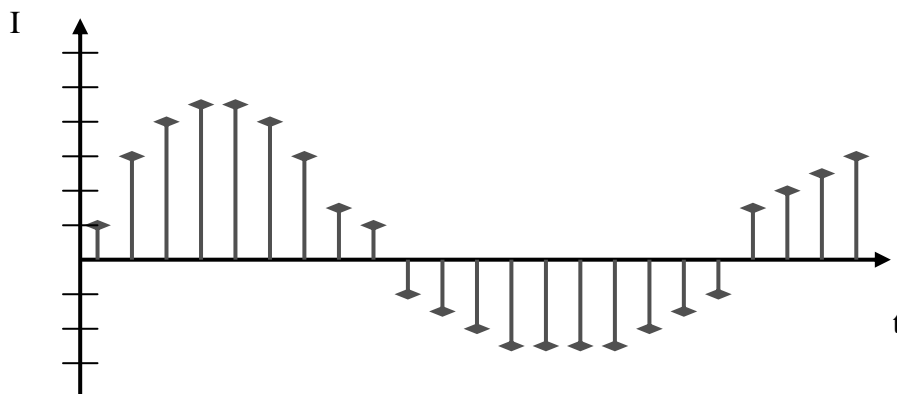
*Во-первых, он производит дискретизацию записываемого звукового сигнала по времени.* Это означает, что измерение уровня интенсивности звука ведется не непрерывно, а, напротив, в определенные фиксированные моменты времени (удобнее, разумеется, через равные временные промежутки). Частоту, характеризующую периодичность измерения звукового сигнала, принято называть *частотой дискретизации*. Вопрос о ее выборе далеко не праздный, и ответ в значительной степени зависит от спектра сохраняемого сигнала: существует специальная теорема Найквиста, согласно которой частота «оцифровки» звука должна как минимум в 2 раза превышать максимальную частоту, входящую в состав спектра сигнала. Считается, что редкий человек слышит звук частотой более 20 000 Гц (20 кГц). Поэтому для высококачественного воспроизведения звука верхнюю границу обычно с некоторым запасом принимают равной 22 кГц. Отсюда немедленно следует, что частота звукозаписи в таких случаях должна быть не ниже 44 кГц<sup>2</sup>. Названная частота используется, в частности, при записи музыкальных компакт-дисков. Однако часто такое высокое качество не требуется, и частоту дискретизации можно значительно снизить. Например, при записи речи вполне достаточно частоты дискретизации 8 кГц. Заметим, что результат при этом получается хоть и не блестящий, но для восприятия речи вполне удовлетворительный<sup>3</sup> - вспомните, как вы слышите голоса своих друзей по телефону.

*Во-вторых, АЦП производит дискретизацию амплитуды звукового сигнала.* Это следует понимать так, что при измерении имеется «сетка» стандартных уровней (например, 256 или 65536 – это количество характеризует *глубину кодирования*), и текущий уровень измеряемого сигнала округляется до ближайшего из них. Напрашивается линейная зависимость между величиной входного сигнала и номером уровня. Иными словами, если громкость возрастает в 2 раза, то интуитивно ожидается, что и соответствующее ему число возрастает вдвое. В простейших случаях так и делается, но, как показывает более детальное изучение, это не самое лучшее решение. Проблема в том, что в широком диапазоне громкости звука человеческое ухо не является линейным. Например, при очень громких звуках (когда «уши закладывает») увеличение или уменьшение интенсивности звука почти не дает эффекта, в то время как при восприятии шепота очень незначительное падение уровня может привести к полной потере разборчивости. Поэтому при записи цифрового звука, особенно при 8-битном кодировании, часто используются различные неравномерные распределения уровней громкости, в основе которых лежит логарифмический закон (*μ-law*, *A-law* и другие; впечатляющие формулы для них можно посмотреть в дополнительной книге [1]).

<sup>2</sup> Обычно используется значение 44032 Гц, которое делится нацело на 256 [1].

<sup>3</sup> Известно, что высокие частоты в основном влияют на «окраску» (тембр) человеческого голоса.

Итак, в ходе оцифровки звука мы получаем поток целых чисел, представляющих собой стандартные амплитуды сигналов через равные промежутки времени. На *рис.1* приведенное выше описание процесса дискретизации проиллюстрировано графически:



*Рис. 1*

На рис. 1 представлен процесс «оцифровки» зависимости интенсивности звукового сигнала  $I$  от времени  $t$ . Отчетливо видна дискретизация по времени (равномерные отсчеты на горизонтальной оси) и по интенсивности сигнала (требуемое при этом округление схематически изображено «изломами» горизонтальных линий разметки). Подчеркнем, что на рисунке степень дискретизации для наглядности сознательно утрирована: реально различие между соседними уровнями дискретизации по обеим осям значительно меньше и, следовательно, форма сигнала передается гораздо точнее.

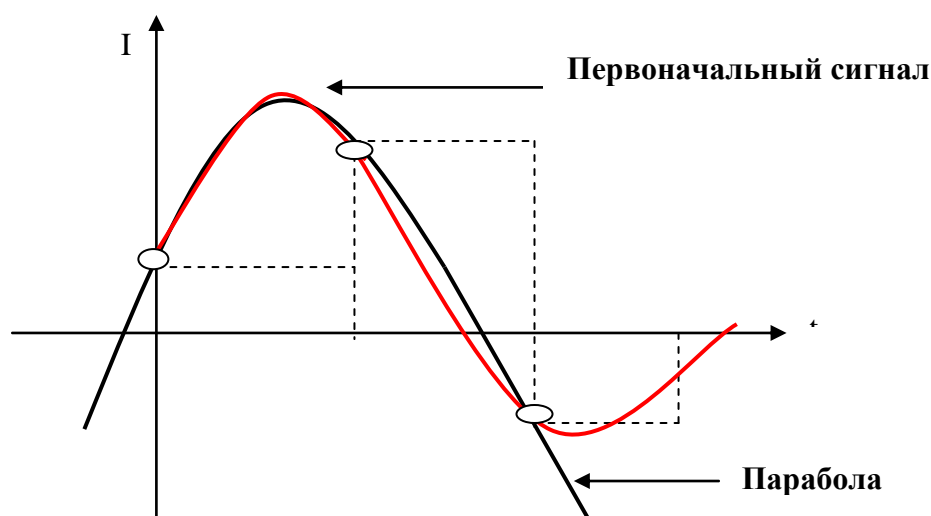
Мы рассмотрели лишь наиболее общие принципы записи цифрового звука. На практике для получения качественных звуковых файлов используется целый ряд дополнительных технических приемов. Их интересное и подробное описание дается в серии статей [2], на которую мы уже ссылались ранее.

Изложенный метод преобразования звуковой информации для хранения в памяти компьютера в очередной раз подтверждает уже неоднократно обсуждавшийся ранее тезис: любая информация для хранения в компьютере приводится к цифровой форме и затем переводится в двоичную систему. Теперь мы знаем, что и звуковая информация не является исключением из этого фундаментального правила.

Остается рассмотреть обратный процесс — воспроизведение записанного в компьютерный файл звука. Здесь имеет место преобразование в противоположном направлении — из дискретной цифровой формы представления сигнала в непрерывную аналоговую, поэтому вполне естественно соответствующий узел компьютерного устройства называется **ЦАП** — **цифро-аналоговый преобразователь**. Процесс реконструкции первоначального аналогового сигнала по имеющимся дискретным данным нетривиален, поскольку **никакой** информации о форме сигнала между соседними отсчетами не сохранилось. В разных звуковых картах для восстановления звукового сигнала могут использоваться различные способы. Наиболее наглядный и понятный из них состоит в том, что по имеющимся точкам рассчитывается степенная функция<sup>4</sup>, проходящая через заданные точки, которая и принимается в качестве формы аналогового сигнала. Ваших математических знаний вполне хватит, чтобы понять, как это делается. Возьмем, например, интерполяцию параболой  $I = at^2 + bt + c$  по трем заданным точкам. Подставив в эту формулу известные значения времени и приравняв их к сохраненным в файле значениям интенсивности звука  $I$ , получим три линейных уравнения с тремя

<sup>4</sup> Математики называют ее *полиномом*.

неизвестными  $a$ ,  $b$ ,  $c$ . Решить такую систему в состоянии любой добросовестный ученик. Качественный вид результата представлен на *рис. 2*.



*Рис. 2*

Видно, что на интерполируемом участке даже для параболы совпадение получается вполне удовлетворительное. Кроме того, технические возможности современных микросхем позволяют значительно увеличить степень полинома, а вместе с ней и точность реконструкции формы сигнала.

Подробно описанный нами метод восстановления является не единственным. К сожалению, другие, не менее эффективные, для своего понимания требуют глубоких специальных технических знаний [2].

Итак, мы подробно рассмотрели процессы преобразования естественных звуков к виду, пригодному для хранения в компьютере, и последующего их восстановления при воспроизведении. Разумеется, не следует требовать от учеников на экзамене большего. Тем не менее, назовем некоторые интересные вопросы, связанные с компьютерной обработкой звуковой информации, которые полезно знать любому грамотному пользователю. Это прежде всего сжатие (кто ни разу не использовал файлы MP3?), MIDI-запись музыки в виде необычайно компактных «нотных» команд для инструментов, форматы звуковых файлов и их особенности, возможности компьютеров в редактировании фонограмм (фильтрация, удаление помех и т.п.) и другие не менее важные темы.

### ***Основная литература***

1. Информационная культура: Кодирование информации. Информационные модели. 9 – 10-е классы. М.: Дрофа, 2000, 208 с.
2. Симаненков Д. Тема номера. Компьютерра № 30-31, с. 20-33; № 32, с. 20-28, 1998 (<http://www.computerra.ru/offline/1998/258> и <http://www.computerra.ru/offline/1998/260>).