

## Обработка цифрового звука

### Сведение

**Сведéние** или **микширование** (*mixing, смешение*) — стадия создания из отдельных записанных треков конечной записи, следующий после звукозаписи этап создания фонограммы, заключающийся в отборе и редактировании (иногда реставрации) исходных записанных треков, объединении их в единый проект и обработке эффектами. Редактирование часто выделяется в самостоятельный этап работы. Стадия сведения исторически возникла в 60-е годы XX века с появлением многодорожечных магнитофонов, позволивших записывать 8 каналов одновременно. Первоначально, сведение не являлось отдельной стадией процесса создания фонограмм, а было составным элементом звукозаписи, так как технические возможности сохранять весь проект отсутствовали, и сведение осуществлялось по мере записи нового материала без возможности возврата к более ранним стадиям. С увеличением числа каналов стала возможной работа со всем проектом, что выделило сведение в самостоятельный этап работы. В 90-е годы XX века значительное распространение получили цифровые технологии сведения. В 1994 году с выпуском программно-аппаратного комплекса Pro Tools III, стала возможной обработка эффектами реального времени, позволившая подбирать параметры обработки во время прослушивания. Cubase VST в 1996 году положил начало сведению в том виде, в котором оно наиболее распространено по настоящее время.

Сведение в проектах электронной музыки — этап следующий после её создания. Этап звукозаписи при работе над электронным проектом чаще всего отсутствует. Граница между созданием и сведением электронной музыки размыта, проект попадает на сведение уже частично сведенным, так как многие виртуальные синтезаторы уже имеют обработки.

В результате сведения многоканальный проект выводится в монофоническую, стереофоническую или многоканальную фонограмму, которая, обычно, получает свой окончательный вид в процессе именуемом мастеринг.

Сведение — не чисто технический процесс соединения различных треков в единое целое, это скорее творческая деятельность, от которой зависят особенности звучания результата. Цель сведения различается в зависимости от концепции проекта.

Критерием оценки в проектах, ориентированных на реалистическое отображение процесса исполнения музыки, служит протокол OIRT (International Radio and Television Organisation):

- Пространственное впечатление записи, которое включает в себя такие составляющие, как: передача ощущения объёма помещения, в котором располагаются звуковые источники, естественность передаваемого пространства, реверберационные отражения, планы звуковых источников, учёт традиций звукового решения пространственного образа в музыке различных стилевых направлений, отсутствие дефектов вызванных наложением нескольких звуковых пространств (многопространственность).
- Прозрачность фонограммы, которая определяется текстовой разборчивостью, различимостью звучания отдельных инструментов или групп инструментов, ясностью передаваемого пространства.
- Музыкальный баланс фонограммы создаваемый логичными соотношениями громкостей между частями произведения, соотношениями громкостей голосов, инструментальных групп и отдельно взятых инструментов.
- Тембр фонограммы, как целостного произведения, удобство восприятия тембрового звучания, естественность отражения тембральной окраски инструментов и выгодность подачи тембров.
- Стерефоничность фонограммы, как целостной звуковой панорамы, характеризующейся симметричностью положения прямых сигналов и отражений, равномерностью и естественностью расположений звуковых источников, учетом традиций решения панорамирования в музыке различных стилей.
- Технические замечания по качеству звукового образа, дефектов, проявляющихся в нелинейных искажениях, неверной передаче частотной характеристики, резонансах, различных видах помех и шумов.
- Характеристика исполнения, которая заключается не только в отсутствии исполнительского брака, заключающегося в неверных нотах, ритмических ошибках, интонационных ошибках, недостаточной ансамблевой сыгранности, но и в качестве использования выразительных средств, таких как темп и его агогические отклонения, пропорциональность динамических оттенков динамического плана произведения в целом и градаций динамики на уровне интонации.
- Аранжировка произведения представленного в виде фонограммы оценивается в случае исполнения переложений музыкальных произведений для других составов исполнителей.
- Динамический диапазон фонограммы в протоколе отражает не только соотношение полезного сигнала и шумов, соотношение уровней звуков между пиками и самыми тихими фрагментами, но и соответствие динамики условиям, в которых фонограмма будет прослушиваться, традиционным для конкретных стилей музыки представлениям о решении динамического плана, естественность и логичность в передаче динамических оттенков, акцентов и кульминаций.

Данная концепция используется при сведении большей части академических жанров, значительной части фольклорной и джазовой музыки. При этом, часто, наилучшие результаты могут быть получены, если условия записи были близки к идеальным, а запись выполняется на стереопару микрофонов и вообще не требует сведения.

При сведении поп-, рок-, электронных и хип-хоп проектов чаще всего фонограмма рассматривается не как документальная фиксация создания музыкантами произведения искусства, а как самостоятельный вид искусства, выражающий свой собственный художественный образ. Для оценки сведения в рамках данной концепции не существует единого протокола.

## Мастеринг

**Мастеринг звукозаписей** — процесс подготовки и переноса записанной и сведенной фонограммы на какой-либо носитель для последующего размножения.

При подготовке к мастерингу полученные после записи исходники редактируются; при необходимости также осуществляется домонтаж/подмонтаж частей произведения. При необходимости могут применяться другие приёмы работы со стереофонограммой: шумоподавление, полосовая компрессия, расширение/сужение стереобазы и т. д. Таким образом и получают материал для создания каждой аудиокomпозиции.

Затем, когда все в отдельности композиции (мастер-треки) готовы, их собирают в нужном порядке, добавляют между ними переходы (напр. добавляются натуральные шумы и т. д.) либо паузы и получают финализированный продукт (аудиоальбом, фонограмму фильма, микстейп и т. д.)

Исходный материал обрабатывается путём эквализации и компрессии (лимитирования). Для **каждого вида издания (носителя)** обычно применяются собственные настройки мастеринга в зависимости от типа носителя, аудиопотребителя и воспроизводящего звук прибора (особенностей его звучания).

**Виртуальный цифровой мастеринг** (ремастеринг) — альтернативный к классическому аппаратному мастеринг звуковых фонограмм, в котором вся редакция осуществляется в виртуальной цифровой среде, как правило, внутри компьютерной системы с помощью звуковых редакторов или специального программно-аппаратного комплекса, в котором звук не проходит дополнительных аналогово-цифровых преобразований, оставаясь в цифровом формате. Получил широкое распространение с конца 1990-х годов.

Последовательность мастеринга:

**Удаление артефактов.** Щелчки, шумы, фон, неправильно поставленная дикция (шипящие и свистящие), западание некоторых частот — все это нам нужно ликвидировать.

Шумоподавление - процесс для удаления шумов готовых фонограмм. Программы или модули (как встроенные, так и внешние), осуществляющие такой тип шумоподавления, называются *денойзерами* (denoiser).

Главным функциональным элементом шумоподавления является *фильтр*, изменяющий полосу пропускания в зависимости от спектра обрабатываемого сигнала. Управляющая этим фильтром электронная схема анализирует входной сигнал, а затем на основе этого анализа меняет параметры фильтра таким образом, чтобы обеспечить максимально благоприятные условия для передачи полезных составляющих сигнала и одновременно максимально ослабить его шумовые компоненты.

Если коснуться практики, то на сегодняшний день большая часть работ по очистке звука сводится к избавлению сэмпла от высокочастотного шума (hiss). Существует два метода. Первый из них - фильтрация. Для начала нужно отфильтровать все верхние частоты, которые забиты именно шумом. А второй метод заключается в том, чтобы программа сканировала образец шума, а затем по алгоритму вычла бы этот шум из волны.

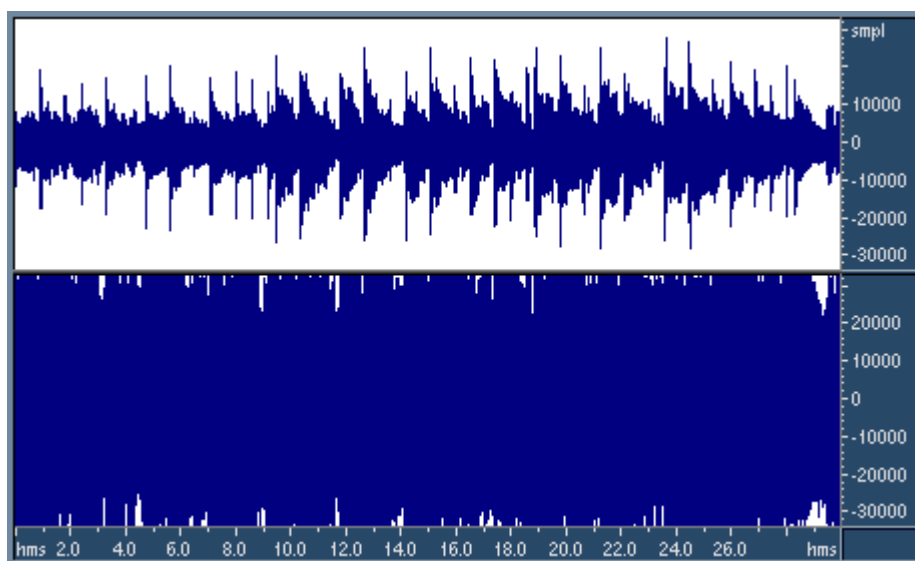
Большинство выпускаемых денойзеров в качестве самого фильтра применяют только фильтр низких частот, который обрезает все высокочастотные помехи, лежащие за пределами полосы частот с полезным сигналом.

У денойзеров помимо достоинств есть также крупный и неустранимый недостаток: вследствие поочередного исчезновения из звукового сигнала отдельных узких полос, его составляющих, сигнал приобретает характерный «металлический» призывок. К сожалению, его невозможно устранить, хотя этот звук можно сделать достаточно малозаметным путем тщательного подбора параметров работы самого денойзера.

**Эквализация (Соблюдение частотного баланса).** Эквализация фонограммы при мастеринге должна, как правило, носить больше косметический характер. Если исходный материал имеет некоторые артефакты, то его лучше пересвести заново. К таким артефактам чаще всего относятся «инструментальная каша» в диапазоне от 500 Гц до 3 кГц, переизбыток реверберации, противофаза и т.д. Частотная картина композиции должна быть единым целым, где понятно расположение инструментов, слышна их отчетливость и прозрачность.

### *Динамическая обработка.*

*Максимайзер (цифровой пиковый лимитер)* - это прибор динамической обработки, повышающий уровень звукового сигнала. Громкость звука зависит не только от уровня звука (или звукового давления), но и от его спектрально-временного состава. Если частотный баланс фонограммы уже определен и менять его нежелательно, то для повышения громкости фонограммы нужно повышать уровень сигнала. Зачем повышать громкость? На это есть две причины. Первая из них заключается в том, что громкая музыка чаще всего кажется "красивее", чем тихая, и больше привлекает внимание. Поэтому большинство продюсеров всеми силами стремятся повысить уровень фонограммы при мастеринге: ведь от этого может зависеть ее коммерческий успех. Вторая причина повышения громкости - желание наиболее полно использовать динамический диапазон носителя аудиозаписи, будь то компакт-диск или аналоговая лента. Также важно максимально использовать динамический диапазон воспроизводящего устройства, чтобы запись не тонула в шумах. При записи звука носители обычно ограничивают величиной перегрузки *пиковый* уровень сигнала, а не его среднеквадратичную мощность (это несколько упрощенная, но близкая к реальности модель для большинства аналоговых и цифровых носителей звука). Отношение пикового уровня фонограммы к ее среднеквадратичному уровню (RMS) называется *пик-фактором* (crest-factor, кросс-фактор). Прямоугольная волна (меандр) имеет единичный пик-фактор 0 дБ. Пик-фактор синусоиды равен 3 дБ. Фонограммы с широкой динамикой или резкими пиками обладают высоким пик-фактором (20 дБ и более), а сильно компрессированные фонограммы - низким пик-фактором (10...15 дБ). Ясно, что при ограниченной пиковой мощности, фонограмма с меньшим пик-фактором может достичь более высокой громкости. С целью уменьшения пик-фактора фонограммы ее пропускают через устройства динамической обработки (рис. 1). Рассмотрим их принципы работы.

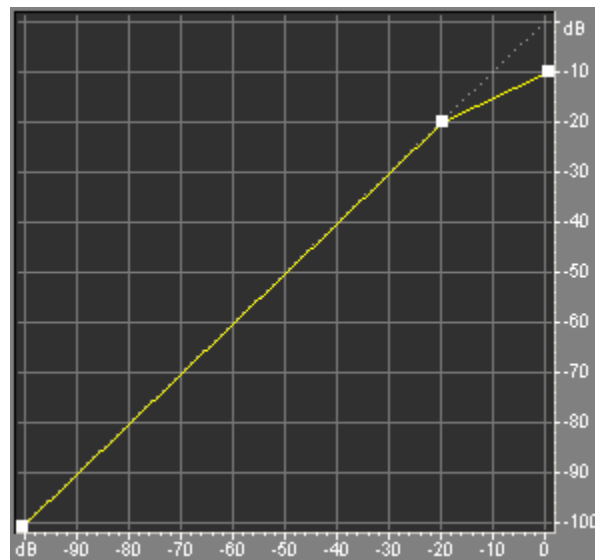


**Рис. 1.** Фонограмма до динамической обработки и после. Уменьшение пик-

фактора. Клиппинга здесь нет, и звук вполне приемлем для радио.

Основными устройствами для работы с уровнями фонограммы являются *устройства динамической обработки*. Принцип действия этих устройств заключается в анализе уровня входящего в них аудио-сигнала и изменении этого уровня по некоторому закону. Основными параметрами устройств динамической обработки являются передаточная характеристика и время атаки/восстановления.

*Передаточная характеристика* (не путать с амплитудно-частотной характеристикой) - это зависимость желаемого выходного уровня звука от входного уровня. В соответствии с передаточной характеристикой, устройство динамической обработки определяет тот коэффициент усиления, который нужно применить к входному сигналу в каждый момент времени. Пример передаточной характеристики показан на рис. 2. Такое устройство динамической обработки называется *компрессором*; оно пропускает без изменения звуки с амплитудой до -20 дБ и уменьшает амплитуду всех звуков выше -20 дБ. Таким образом, компрессор делает громкие звуки тише, сужая динамический диапазон фонограммы.

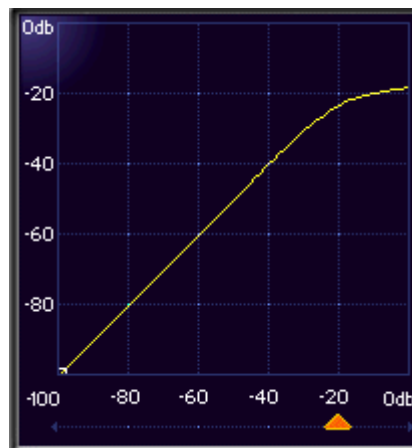


**Рис. 2.** Передаточная характеристика компрессора. Порог равен -20 дБ, степень компрессии 2:1.

Перелом в передаточной характеристике называется *точкой неергиба* (knee). Входной уровень, соответствующий колену, называется *порогом* (threshold). Угол наклона передаточной характеристики выше порога определяет *степень сжатия* (ratio, степень компрессии). Степень сжатия 2:1 означает, что при увеличении входного уровня на 2 дБ выше порога выходной уровень вырастет лишь на 1 дБ. Если степень сжатия равна единице, то уровень звука при прохождении через прибор не изменится. Если она стремится к бесконечности, то устройство будет ограничивать амплитуду выходного звука значением порога. Такие устройства называются *лимитерами*, они ограничивают динамический диапазон. Если же степень сжатия меньше единицы, например 1:1.5, то это значит, что когда входной

уровень превышает порог, устройство будет повышать выходной уровень по сравнению с входным. Такие устройства называются *экспандерами*, они расширяют динамический диапазон. Существуют и другие типы устройств динамической обработки: гейты, дакеры, левелеры и пр., со своими специфическими передаточными характеристиками и параметрами работы.

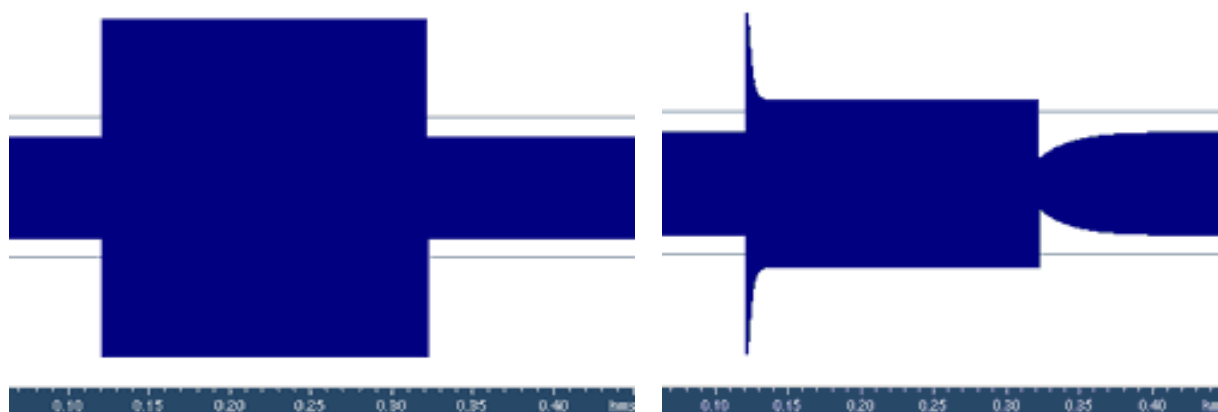
Иногда передаточная характеристика сглаживается, чтобы в ней не было острых углов (рис. 3). Этот режим называется *soft knee* или *soft threshold* (мягкий порог). Компрессор с мягким порогом начинает немного уменьшать уровень сигнала еще до достижения им величины порога.



**Рис. 3.** Мягкий порог.

Работу устройства динамической обработки можно описать следующей схемой. Устройство следит за входным уровнем и в соответствии с ним регулирует выходной уровень, т.е. *применяет к входному сигналу некоторую амплитудную огибающую (коэффициент усиления), меняющуюся со временем*. Для хорошего звучания результирующего сигнала нужно соблюсти несколько условий. Самое главное из них таково: амплитудная огибающая должна быть гладкой, без разрывов и, по возможности, без изломов. Действительно, если амплитудная огибающая имеет разрывы, то выходной звук тоже будет иметь разрывы в форме волны, слышимые как щелчки и треск. Изломы в амплитудной огибающей также будут приводить к искажениям выходного сигнала.

Для сглаживания амплитудной огибающей в устройствах динамической обработки имеются два параметра: *время атаки* (attack, время срабатывания) и *время восстановления* (release). Они определяют скорость реакции устройства на изменения входного уровня. Время атаки показывает, за какое время устройство реагирует на превышение порога (атаку), а время восстановления показывает, за какое время устройство реагирует на возвращение входного уровня обратно под порог.



**Рис. 4.** Звук до и после обработки компрессором.

Пусть на вход компрессора подается сначала слабый сигнал, не превышающий порога, а затем - атака, превышающая порог (рис. 4). В соответствии с передаточной характеристикой, компрессор должен пропустить слабый сигнал без изменения, а уровень громкого сигнала (атаки) - ослабить. Время атаки указывает, за какое время компрессор изменит свой коэффициент усиления от единичного до результирующего, предписанного передаточной характеристикой. Если вслед за громким сигналом входной уровень снова опускается ниже порога, то компрессор переходит в стадию восстановления и снова увеличивает свой коэффициент усиления до единичного. Время, за которое коэффициент усиления вернется к единичному значению, и будет временем восстановления.

Определения времени атаки и восстановления могут различаться у разных производителей. В некоторых устройствах под временем восстановления понимается не полное время возвращения коэффициента усиления, а время его возвращения, скажем, до половины обратного пути. Часто коэффициент усиления возвращается к исходному значению по экспоненте, и в этом случае лишь второе определение имеет смысл. В некоторых устройствах время атаки задается скоростью изменения коэффициента усиления (дБ/сек) или, наоборот, временем изменения коэффициента усиления на 6 дБ.

Время атаки и время восстановления измеряются в миллисекундах и могут меняться в широких пределах для различных приборов динамической обработки и в зависимости от конкретных задач. Так, например, в компрессорах обычное время атаки - порядка 10...100 мс, а типичное время восстановления - порядка 100...1000 мс. Ясно, что чем больше время атаки и восстановления, тем медленнее будет меняться во времени амплитудная огибающая, тем более гладкой она будет. Однако при большом времени атаки компрессор будет пропускать короткие атаки, превышающие порог, т.к. не будет успевать на них реагировать. Это может быть нежелательным, например, для лимитеров.

**Создание звуковой картины.** В определенных случаях требуется добавление



реверберации, расширение стереобазы. При этом вы должны понимать, что внедрение трехмерного звука и реверберации может повлиять в худшую сторону на параметр RMS. Поэтому нужно выбирать между абсолютной прослушиваемостью и эстетическим фактором.

На самом деле многие звукорежиссеры уже не раз отказывались от реверберации в пользу более громкого звучания фонограммы, но тогда очень часто терялся этот самый эстетический фактор. Как вариант — использовать системы многополосной реверберации, где вы можете добавить эффект только для определенного частотного диапазона. Картина получается более аккуратной. В принципе по правилам психоакустики достаточно только обозначить присутствие эффекта, и слушатель будет его ощущать подсознательно. Чтобы получить более яркое звучание, часто необходимо использовать такое устройство, как эксайтер, который позволяет насытить сигнал дополнительными гармониками и улучшить параметры стереозвучания.

## **Аудиоредакторы**

### **Sony Sound Forge**

Программа Sound Forge долгое время являлась флагманским продуктом компании Sonic Foundry, а в мае 2003 года успешно перекочевала под другую торговую марку — Sony Pictures Digital. К радости огромной армии поклонников этого звукового редактора, подразделение Sony продолжило развитие Sound Forge. На сегодняшний день это отличный и удобный цифровой аудиоредактор, в состав которого входит набор утилит, предназначенных для работы со звуком. При помощи данной программы можно эффективно обрабатывать звуковые композиции, накладывать множество эффектов, осуществлять точную и быструю запись и редактирование сэмплов, кодировать данные, конвертировать звуковые файлы в различные форматы и т.д.

Это одна из самых популярных и действительно полезных программ, предназначенных для обработки аудио на профессиональном уровне. С ее помощью можно менять аудиосигнал до невозможности или же редактировать записанную партию любого музыкального инструмента. Sony Sound Forge успешно объединяет в себе практически полный набор современных звуковых эффектов и мощные средства редактирования звуков для последующего их использования в сэмплере.

Отличительные особенности Sound Forge таковы:

- функциональность и эргономичность интерфейса;
- большое количество возможностей редактирования и обработки эффектами;
- поддержка большинства новых технологических стандартов, включая аудио с характеристиками 32/64 бита с плавающей точкой и частотой дискретизации от 2

до 192 кГц;  
постоянная поддержка со стороны разработчика.

### **Adobe Audition**

Эта программа имеет достаточно длинную историю. Когда-то она носила название Cool Edit и была довольно упрощенным вариантом нынешней программы. По существу, первая версия Audition являла собой клон последней версии Cool Edit Pro (на которой и закончилась история этого редактора) На сегодняшний день Adobe Audition — профессиональный инструмент для микширования и обработки аудиоматериалов. Программа предоставляет средства для микширования, редактирования, создания мастер-копий и обработки аудиоэффектов:

- редактирование отдельных аудиофайлов;
- смешивание одновременно до 128 треков;
- создание петель и применение более 45 цифровых встроенных эффектов;
- производство высококачественной аудиопродукции: Adobe Audition поддерживает файлы с разрядностью до 32 бит и частотой дискретизации до 192 кГц;
- поддерживается запись на любые носители (включая магнитную пленку, диски CD, DVD или DVD- Audio). Вся внутренняя обработка выполняется с 32-битной разрядностью, что позволяет достичь высокого качества выходного материала;
- удобный интерфейс;
- встроенные режимы просмотра (многодорожечный и для редактирования);
- спецэффекты реального времени;
- средства анализа;
- функция восстановления звука и поддержка видеоряда;
- поддержка виртуальных инструментов VSTi;
- современные средства микширования;
- средства для адаптивного устранения шума и инструменты для поэтапной коррекции, которые помогают восстановить старые записи.

### **Audacity**

Audacity — это небольшой кроссплатформенный бесплатный аудиоредактор, работающий под Windows, Linux/ Unix, Mac OS 9 и Mac OS X. Он создан группой энтузиастов на C/C++ и распространяется под лицензией GPL.

Вполне приличное количество встроенных средств обработки звука: усиление (Amplify), усиление басов (Bass Boost), эхо, вау-вау (Wah-wah), все виды фейдов, удаление шума и т.д. Кроме того, версии под Windows и Mac OS знают VST-

плагины. Linux-версия Audacity поддерживает и вовсе несметное количество различных плагинов — уже больше сотни.

В Audacity есть стандартный набор встроенных эффектов и фильтров: шумоподавление, усиление/ослабление сигнала, фэйзер, эхо, FFT-фильтр и др. В комплект Audacity также входит бесплатный VST-ревербератор Freeverb2. Кроме того, вы можете создавать собственные плагины на основе скриптового языка, поддерживаемого Audacity. Другими словами, у вас есть возможность не только написать математическую формулу для обработки звукового сигнала, но и сформировать под это дело графический интерфейс. Такие плагины просто появляются в меню Effect, как и обычные эффекты и плагины. Дополнительные плагины можно скачать с сайта редактора.

К сильным сторонам Audacity относятся широкие возможности по обработке сигнала и способность программы узнавать любые форматы аудио — как сжатые (MP3, OGG), так и не сжатые (WAV, AIF).

Благодаря оригинальной системе хранения аудиоданных Audacity очень быстро открывает и записывает свои проект-файлы. Пока работа над аудиосигналом производится непосредственно в самом пакете, сигнал хранится на диске в виде небольших файлов собственного формата Audacity, каждый из которых представляет собой отдельный фрагмент аудио. В Sound Forge и Audition обработка производится непосредственно над цельным wav-файлом, из-за чего запись и открытие файла происходят достаточно медленно. Собственный формат Audacity не совместим ни с какими другими программами, да это и не нужно, учитывая его сугубо служебные функции. Audacity позволяет экспортировать аудио в любой распространенный формат, кроме MP3. Алгоритм кодирования в MP3 защищен копирайтом, так что кодек придется скачивать отдельно. На сайте проекта Audacity есть нужные ссылки.

Инструментов для редактирования в Audacity не очень много. Это обычные утилиты для нарезки звука, масштабирования и перемещения. Есть еще инструмент для редактирования огибающих громкости.

Не обладая всеядностью некоторых звуковых редакторов, Audacity тем не менее поддерживает чтение и запись таких форматов, как .wav, .aiff, .au, OggVorbis и .mp3. Запись в последний — только при наличии внешнего энкодера. Например, LAME.