

# 1 Лабораторная работа №1 – Компрессор динамического диапазона

*Цель:* Реализация в системе Matlab компрессоров динамического диапазона для достижения различных звуковых эффектов.

## 1.1 Теоретические сведения

### 1.1.1 Динамическая обработка звуковых сигналов

Динамическая обработка предназначена для сокращения динамического диапазона звуковых сигналов. Звуковые сигналы в радиовещании, телевидении и звукозаписи всегда подвергаются такой обработке, независимо аналоговые они или цифровые. Это связано с тем, что часто динамический диапазон природных звуков, звуков музыки и речи значительно шире динамического диапазона электроакустических трактов современной аппаратуры. Если не использовать динамическую обработку и передавать полностью натуральный динамический диапазон оркестра (70...80 дБ), то чтобы не мешали шумы помещения, уровень сигнала пианиссимо (тихое звучание) должен быть не ниже 40 дБ SPL (*SPL – sound pressure level*). Фортиссимо (громкое звучание) при этом будет достигаться при уровне звукового давления 120 дБ SPL, что близко к болевому порогу. **Динамическая обработка** звуковых сигналов позволяет слушать радио и телевизионные передачи со звуковым давлением не выше 100 дБ SPL в обычных жилых помещениях, в которых обычно уровень шума не менее 40...50 SPL и, даже в автомобилях, где уровень шума значительно больше. Без такой обработки при пиковом звуковом давлении не выше 80 дБ SPL (как в театре) звукозаписи можно было бы слушать только в полной тишине, закрыв окна и двери, иначе любой посторонний шум заглушит тихие места фонограммы.

Динамическая обработка звуковых сигналов производится с помощью *ограничителей (лимитеров), максимайзеров, компрессоров, экспандеров и гейтов*. Это все пороговые устройства, в которых при достижении сигнала установленного уровня их коэффициент передачи меняется скачком. Таким способом можно как сжать, так и расширить динамический диапазон сигнала.

## 1.2 Динамический диапазон сигналов и трактов

Человек слышит звук в чрезвычайно широком диапазоне звуковых давлений. Этот диапазон простирается от абсолютного порога слышимости до болевого порога 140 дБ SPL относительно нулевого уровня, за который принято давление 0,00002 Па (рис. 1.1).

Зона риска на этом рисунке обозначает область звуковых давлений, которые при длительном воздействии могут привести к полной потере слуха. Болевой порог для тональных звуков зависит от частоты, для звуков с произвольным спектром за болевой порог принят уровень давления 120 дБ SPL.

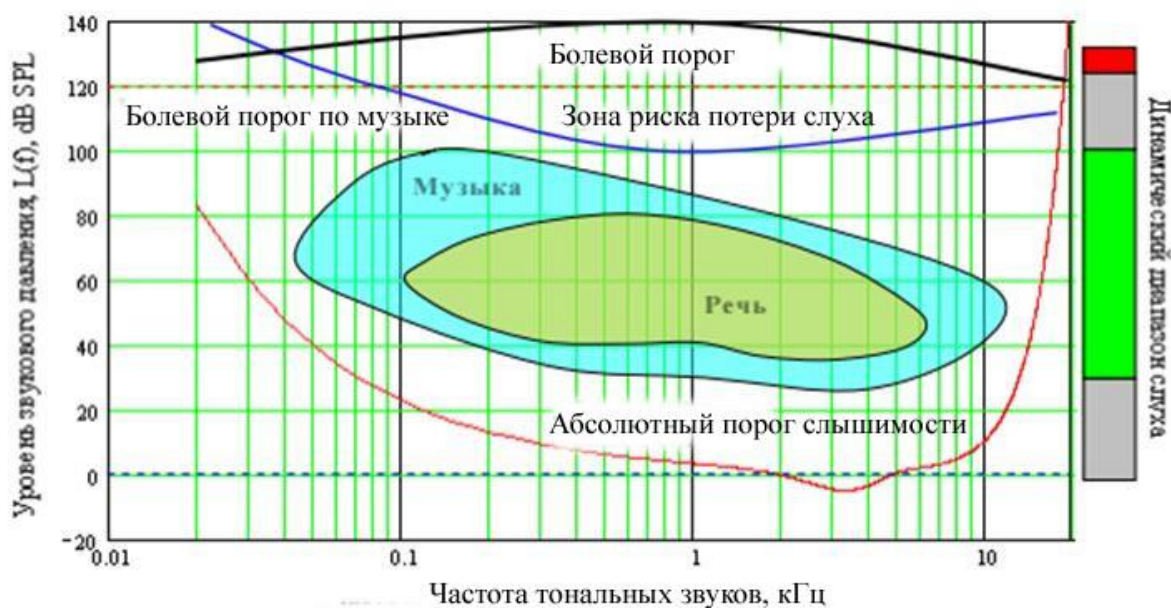


Рисунок 1.1 – Области слышимости слуха

В тишине чувствительность слуха человека повышается, а при наличии громких звуков – понижается, слух адаптируется к окружающей звуковой среде, поэтому динамический диапазон слуха не такой большой – около 70..80 дБ. Сверху он ограничен давлением 100 дБ SPL, а снизу шумом с уровнем -30...35 дБ SPL. Этот динамический диапазон может сдвигаться вверх и вниз до 20 дБ. Для комфортного восприятия музыки рекомендуется, чтобы звуковое давление не превышало 104 дБ SPL в домашних условиях и 112 дБ SPL, в специально оборудованных помещениях.

**Динамический диапазон музыки** определяется отношением в децибелах самого громкого звука (фортиссимо) и самого тихого звука (пианиссимо). **Динамический диапазон симфонической музыки** составляет 65...75 дБ, а на концертах рок-музыки он возрастает до 105 дБ, при этом

пики звуковых давлений могут достигать 122...130 дБ SPL. Динамический диапазон вокальных исполнителей не превышает 35...45 дБ

Шум помещения маскирует звук и этим уменьшает динамический диапазон. Влиянием шума можно полностью пренебречь, если его уровень на 10...20 дБ SPL ниже минимального уровня музыкальных звуков. В студиях звукозаписи уровень шума не превышает 20...30 дБ SPL, ночью в квартирах «тихих» домов этот уровень равен 40 дБ SPL, любой разговор увеличивает уровень шума до 60 дБ SPL. Именно поэтому тихая музыка часто тонет в шумах помещения прослушивания и невольно возникает желание увеличить громкость.

### 1.3 Громкость звука и динамический диапазон

С динамическим диапазоном звукового сигнала тесно связана такая важная характеристика фонограммы (записи звука) как громкость звучания. Громкость звука является субъективной характеристикой. Она наиболее сильно зависит от двух факторов - звукового давления и спектрального состава звука. В качестве одной из оценок громкости используется понятие **уровня громкости**, измеряемого в фонах. В соответствии со стандартом ISO на частоте 1 кГц уровень громкости в фонах полностью определяется уровнем звукового давления в децибелах (например, 40 дБ SPL – 40 фон).

Связь между уровнем громкости и звуковым давлением на других частотах определяется экспериментальными графиками, называемыми кривыми равной громкости или изофонами (рис. 1.2).

Из приведенных графиков кривых равной громкости следует, что для натурального звучания запись должна воспроизводиться при том же уровне громкости, при котором она была осуществлена. Когда запись, сделанная на высоком уровне громкости, воспроизводится на более низких уровнях, низкие и высокие частоты как бы пропадают, частотный баланс нарушается. Если запись сделана на низком уровне громкости, а воспроизводится на высоком - пение и речь кажутся бубнящими, так как слушатель воспринимает в них низкие частоты значительно более громкими, чем при прослушивании естественного источника на более «тихих» уровнях.

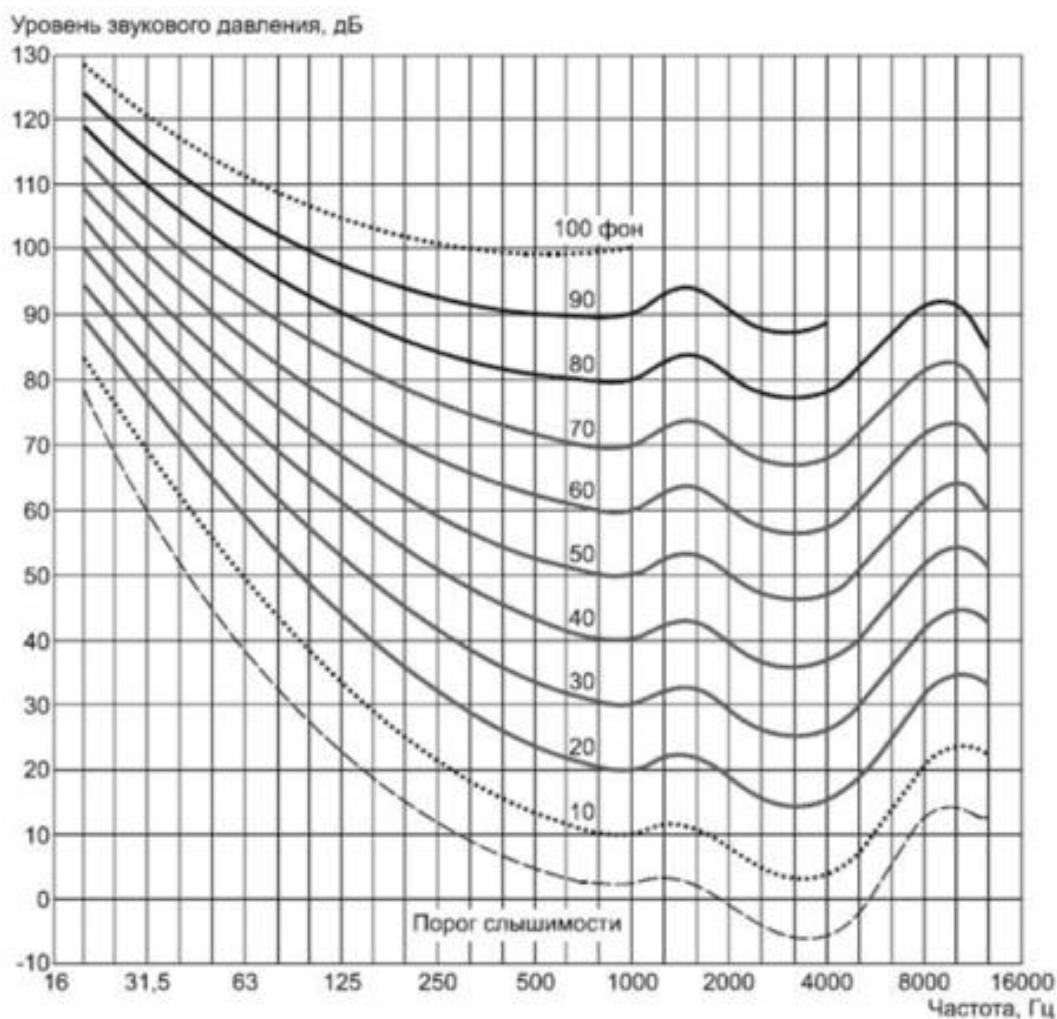


Рисунок 1.2 – Кривые равной громкости

#### 1.4 Принцип динамической обработки

**Динамическая обработка** звуковых сигналов производится для сужения или расширения динамического диапазона. Прибор или программный модуль, выполняющий эту функцию, называется **динамическим процессором (DP)**. Динамическая обработка применяется только для нестационарных сигналов с подъемами и провалами огибающей звукового сигнала. Уровень огибающей определяется с помощью пикового детектора, при этом одновременно вычисляется среднеквадратическое значение сигнала с временем интегрирования примерно равным постоянной времени слуха – 50...60 мс.

Принцип цифровой **динамической обработки** основан на почти **мгновенном изменении коэффициента передачи** сигнала, когда уровень огибающей звукового сигнала становится выше (ниже) установленного порога. Порог может задаваться как пиковым значением уровня, так и среднеквадратическим.

Основными параметрами ДР являются: **амплитудная характеристика** (в виде графика или функции), порог срабатывания, время установления (attack)  $T_y$  и время восстановления  $T_b$  коэффициента передачи. Перелом в амплитудной характеристике называется точкой перегиба. Входной уровень, соответствующий перегибу, называется порогом, задается он в децибелах. Угол наклона графика передаточной характеристики выше (ниже) точки перегиба определяет коэффициент передачи подъемов (провалов) огибающей звукового сигнала.

В зависимости от вида передаточной характеристики динамического процессора, он может выполнять функции лимитера, компрессора, экспандера или ограничителя шума (гейта). Все эти функции могут быть реализованы в одном устройстве или с помощью одного программного модуля. На рис. 1.3 приведены типовые передаточные характеристики этих устройств.

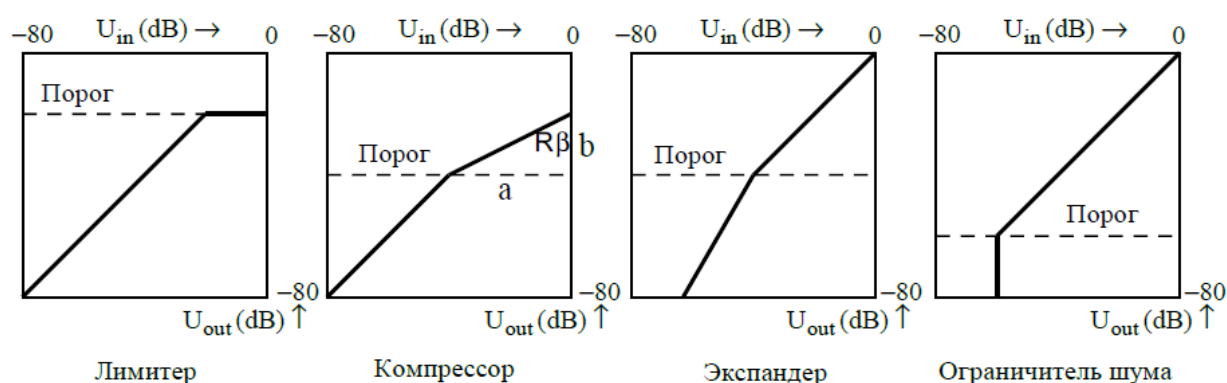


Рисунок 1.3 – Амплитудные (передаточные) функции компрессоров

Передаточная характеристика ДР однозначно определяет регулировочную характеристику, в соответствии с которой меняется коэффициент передачи регулирующего элемента ДР. Изгибы и изломы передаточной характеристики предопределяют дискретный характер изменения коэффициента передачи регулирующего элемента ДР, и они в значительной мере определяют форму амплитудной огибающей динамически обработанного звукового сигнала.

**Лимитер** ограничивает динамический диапазон сверху – пики и подъемы огибающей, значение которых превышают пороговый уровень, он ослабляет до уровня, равного установленному порогу. Компрессор также ограничивает динамический диапазон сверху – подъемы огибающей, величина которых превышает пороговый уровень, он ослабляет в соответствии с установленным коэффициентом передачи выше точки перегиба.

**Экспандер** расширяет динамический диапазон снизу – провалы огибающей, уровень которых ниже порогового он увеличивает путем уменьшения коэффициента передачи ниже точки перегиба. Гейт ограничивает динамический диапазон снизу – все сигналы в провалах огибающей, уровень которых ниже порогового он не пропускает.

Степень сжатия (компрессии) определяется отношением коэффициентов передачи ниже и выше точки перегиба передаточной функции. Поэтому она может задаваться как отношение приращений входного и выходного сигналов в децибелах, например, 18 : 6.

Гейт - это экспандер с коэффициентом передачи равным нулю. Этот DP понижает коэффициент усиления ниже порога вплоть до нуля. Жесткое удаление шума вызывает неприятное ощущение, как будто вам заткнули уши, поэтому для этих целей он используется редко, хотя и разрабатывался для понижения шума в паузах речи. В основном он применяется для динамической обработки звука ударных инструментов.

В работе DP очень важную роль играют его временные параметры: время установления  $T_y$ , часто называемое временем срабатывания, и время восстановления  $T_v$ , в значительной мере определяющие длительность переходных процессов. Эти параметры измеряются в единицах времени, тогда как скорость установления и восстановления измеряется в дБ/с. Последнее означает, что время срабатывания и время восстановления не являются величинами постоянными, а зависят от уровня звукового сигнала. Эта зависимость проявляется в том, что чем больше уровень сигнала, тем больше время установления и меньше время восстановления.

Выбор значений этих параметров зависит от режима работы динамического процессора (компрессор, экспандер, лимитер или гейт). Существенное значение имеет характер музыкального материала. Иногда эти параметры выбираются исходя из необходимости решения каких либо творческих задач.

Чем меньше время установления и восстановления, тем громче получается результирующий звук, но тем более изломанной получается амплитудная огибающая выходного сигнала DP, и тем больше возникает интермодуляционных искажений. Ясно, что чем больше время установления и восстановления, тем медленнее будет меняться во времени амплитудная огибающая, тем более гладкой она будет. Однако, при большом времени установления и восстановления начинает проявляться эффект "проваливания

громкости" (pumping). Провалы громкости возникают вокруг кратковременных пиков в сигнале. Таким образом, выбор времени установления/восстановления - это компромисс между интермодуляционными искажениями и эффектом проваливания громкости

**Время установления**  $T_y$  определяет время между моментом нарастания или спада огибающей входного сигнала до порогового значения (снизу или сверху) и временем, через которое устанавливается заданный уровень выходного сигнала динамического процессора.

Чем меньше время установления - тем на более коротких пиках уровня срабатывает DP. При малом времени установления можно практически исключить возникновение пиков сигнала при скачкообразном увеличении его уровня. Однако при этом звучание может стать недостаточно акцентированным. Так же важно помнить, что малое время установления может вызвать искажения, особенно на низких частотах. Низкочастотные сигналы имеют большой период, и компрессия, действующая посреди периода, звучит очень неестественно.

При большом времени установления компрессор будет пропускать короткие атаки звукового сигнала, превышающие порог, т.к. не будет успевать на них реагировать. Это может быть нежелательным, например, для лимитеров. В сигнале на выходе компрессора будут присутствовать пики.

Для лимитеров чаще всего выбирают время установления в пределах 0,5 - 1,0 мс, для речевых компрессоров время установления должно составлять 1... 2 мс, в обычных компрессорах время установления - порядка 10...100 мс.

**Время восстановления**  $T_v$  определяет время между моментом спада пика огибающей входного сигнала до порогового уровня (сверху или снизу) и временем, через которое на выходе процессора восстановится исходный уровень звукового сигнала. При малом времени восстановления обеспечивается более существенное сглаживание. Если время восстановления слишком велико, то компрессор дольше находится в активном состоянии и воздействует на динамический диапазон даже тогда, когда это нежелательно. Это дает заметный на слух эффект пульсации звука, так как компрессия не приводит к сглаживанию сигнала.

Обычно рекомендуется для инструментальной музыки в качестве грубого приближения и отправной точки для более тонкой настройки выбирать время восстановления около 500 мс. Это соответствует промежутку между

двумя тактами при темпе 120 четвертей в минуту. Для речевых компрессоров время восстановления - 300 мс, в обычных компрессорах время восстановления - порядка 100...1000 мс.

## 1.5 Практические аспекты реализации компрессора

### 1.5.1 Схема компрессора

На рис. 1.4 показана общая схема компрессора (динамического процессора). Сначала измеряется текущий уровень сигнала, а затем вычисляются требуемые коэффициенты усиления для данного уровня  $g(n)$  исходя из передаточной характеристики компрессора.

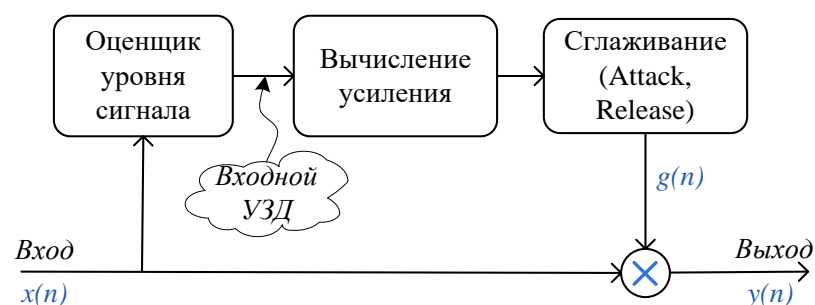


Рисунок 1.4 – Схема компрессора

Этап сглаживания полученных коэффициентов усиления не является обязательным, иногда временные константы времени установки (attack) и восстановления (release) «встраиваются» в оценщик уровня сигнала.

Сигнал в компрессоре может обрабатываться в блочном режиме, либо в режиме отсчет-за-отсчетом. Первый больше характерен для программной реализации, а второй лучше подходит для аппаратной реализации.

### 1.5.2 Детекторы (оценщики) уровня сигнала

Оценка уровня сигнала может производиться как по **абсолютному** значению, так и по **среднеквадратическому** значению (англ. RMS – root mean square). В первом случае усредняются **абсолютные** значения сигнала –  $|x(n)|$ , в во втором – квадраты мгновенных значений сигнала –  $x^2(n)$ .

Чаще всего в детекторах используют формулу **экспоненциального усреднения**. Для детектора **абсолютного** уровня она принимает вид:

$$p(n) = \begin{cases} \alpha p(n-1) + (1-\alpha)|x(n)|, & \text{если } |x(n)| \geq p(n-1) \\ \beta p(n-1) + (1-\beta)|x(n)|, & \text{если } |x(n)| < p(n-1) \end{cases} \quad (1)$$

где  $\alpha$  – параметр, зависящий от времени установки,  $\beta$  – параметр, зависящий от времени восстановления.



Для детектора *среднеквадратичного* уровня используют выражения:

$$p(n) = \begin{cases} \alpha p(n-1) + (1-\alpha)x^2(n), & \text{если } x^2(n) \geq p(n-1) \\ \beta p(n-1) + (1-\beta)x^2(n), & \text{если } x^2(n) < p(n-1). \end{cases} \quad (2)$$

Следует помнить, что в случае оценки уровня по *абсолютному* значению для перевода в децибелы следует пользоваться формулой:

$$P = 20 \log_{10} p(n),$$

а в случае *среднеквадратического* детектора:

$$P = 10 \log_{10} p(n).$$

Существует также **непосредственный способ** оценки уровня сигнала, основанный на вычислении мощности сигнала в скользящем окне:

$$P(x[n]) = 20 \log_{10} \sqrt{\frac{1}{m} \sum_{i=0}^{m-1} x^2[n-i]} = 10 \log_{10} \left( \frac{1}{m} \sum_{i=0}^{m-1} x^2[n-i] \right), \quad (3)$$

где  $x[n]$  – входной сигнал, а  $m$  – длина временного окна (в отсчетах), используемого для вычисления мощности сигнала. Временное окно должно быть короче, чем период стационарности речевого сигнала (5-20 мс), но длиннее, чем период одного речевого колебания (0,3-3 мс для 0,3-3 кГц). Например, если выбрать длину окна равную 4 мс, она составляет 64 отсчета при частоте дискретизации 16 кГц.

В выражении (3) рассчитывается *среднеквадратическое* значение уровня. *Абсолютный* уровень сигнала для непосредственного способа оценивается по формуле:

$$P(x[n]) = 10 \log_{10} \left( \frac{1}{m} \sum_{i=0}^{m-1} |x[n-i]| \right). \quad (4)$$

### 1.5.3 Определение временных констант<sup>1</sup>

Реакция на единичную ступеньку аналоговой системы (первого порядка) имеет вид:

$$g(t) = 1 - \exp(-t/\tau), \quad (5)$$

где  $\tau$  – временная постоянная. В результате дискретизации (5) получаем

$$g(nT_s) = 1 - \exp\left(-\frac{nT_s}{\tau}\right) = 1 - z_{\infty}^n, \quad z_{\infty} = \exp(-T_s/\tau) \quad (6)$$

Применяя z-преобразование получаем:

<sup>1</sup> Zolzer U. Digital audio signal processing, p. 230

$$G(z) = \frac{z}{z-1} - \frac{1}{1-z_{\infty}z^{-1}} = \frac{1-z_{\infty}}{(z-1)(1-z_{\infty}z^{-1})}. \quad (7)$$

Если определять время установки, как  $t_a = t_{90} - t_{10}$ , т.е. такое время за которое уровень сигнала повышается с 10% до 90%, то получим

$$0.1 = 1 - \exp\left(-\frac{t_{10}}{\tau}\right) \rightarrow t_{10} = 0.1\tau$$

$$0.9 = 1 - \exp\left(-\frac{t_{90}}{\tau}\right) \rightarrow t_{90} = 0.9\tau$$

Связь между временем установки  $t_a$  и временной постоянной  $\tau$  можно установить следующим образом:

$$\begin{aligned} \frac{0.9}{0.1} &= \exp\left(\frac{t_{90} - t_{10}}{\tau}\right) \\ \ln\left(\frac{0.9}{0.1}\right) &= (t_{90} - t_{10})/\tau \\ t_a = t_{90} - t_{10} &= 2.2\tau \end{aligned}$$

Отсюда получаем выражение для полюса:

$$z_{\infty} = \exp(-2.2T_s/t_a). \quad (8)$$

Система для реализации данного отклика на ступенчатый сигнал можно получить, используя связь между передаточной функцией системы и z-образом реакции на единичную ступеньку:

$$H(z) = \frac{z-1}{z} G(z). \quad (9)$$

Используя (7) и (9) передаточную функцию может быть записана как:

$$H(z) = \frac{(1-z_{\infty})z^{-1}}{1-z_{\infty}z^{-1}} \quad (10)$$

с полюсом  $z_{\infty} = \exp(-2.2T_s/t_a)$ , определяющим время установки, восстановления или усреднения.

Таким образом, для уравнения

$$p_{RMS}^2(n) = \alpha p_{RMS}^2(n) + (1-\alpha)x^2(n) \quad (11)$$

мы имеем передаточную функцию

$$H(z) = \frac{1-\alpha}{1-\alpha z^{-1}}, \quad (12)$$

откуда

$$\alpha = \exp(-2.2T_s/t_a). \quad (13)$$

#### 1.5.4 Блочная (фреймовая) обработка сигнала в компрессоре

При блочном режиме на вход компрессора поступает входной сигнал  $x^{(m)}[r]$  поступает в режиме фрейм за фреймом. Задача компрессора состоит в расчете линейной функции усиления  $g[r]$  для формирования выходного сигнала:

$$y^{(m)}[r] = x^{(m)}[r]g[r], \quad r = 1, \dots, R.$$

Как правило на функцию  $g[r]$  накладывают ограничение линейности. В этом случае достаточно определить значения  $g[1]$  и  $g[R]$ , а остальные значения находятся путем линейной интерполяции. Ниже приведен псевдокод работы компрессора. В результате работы алгоритма определения значения  $g[1]$  и  $g[R]$ , которые затем интерполируются для получения  $g[r]$ .

**Вход:**  $x^{(m)}[r]$

$\alpha$  – параметр экспоненциального усреднения, зависящий от времени установки;

$\beta$  – параметр экспоненциального усреднения, зависящий от времени восстановления.

**Инициализация:**  $p = 2^{-15}$

$P_{start} = 10\log_{10}(p)$

**for**  $r = 1, 2, \dots, R$

$p_{inst} = (x^{(m)}[r])^2$

**if**  $(p_{inst} > p)$  **then**

$p = \alpha p + (1 - \alpha)p_{inst}$

**else**

$p = \beta p + (1 - \beta)p_{inst}$

**end if**

**end for;**

$P_{end} = 10\log_{10}(p)$  // входное УЗД

$P_{start}^{out} = \text{IO\_Func}(P_{start})$  // выходное УЗД (начало фрейма)

$P_{end}^{out} = \text{IO\_Func}(P_{end})$  // выходное УЗД (конец фрейма)

$G_{start} = P_{start}^{out} - P_{start}$  // усиление в дБ (начало фрейма)

$G_{end} = P_{end}^{out} - P_{end}$  // усиление в дБ (конец фрейма)

$g[1] = 10^{G_{start}/20}$  // линейный коэффициент усиления (начало фрейма)

$$g[R] = 10^{G_{end}/20}, \quad // \text{ линейный коэффициент усиления (конец фрейма)}$$

В приведенной программе переменная  $p$  хранит текущее (среднее) значение уровня мощности сигнала. Первому кадру сигнала ( $m=0$ )  $p$  присваивается значение  $2^{-15}$ , в последующем значение  $p$  сохраняется от фрейма к фрейму. Таким образом,  $P_{start}$  для  $m$ -го фрейма равно значению  $P_{end}$  ( $m-1$ )-го фрейма. Текущее значение мощности сигнала вычисляется путем экспоненциального усреднения. Параметр экспоненциального усреднения выбирается в зависимости от того, что происходит – нарастание или спад сигнала. Через  $IO\_Func()$  в программе обозначена функция амплитудная характеристика компрессора.

## 2 Задание

Лабораторная работа включает реализацию динамического процессора (компрессора) в среде Matlab заданного типа в соответствии с вариантом (таблица 1).

Таблица 1 – Варианты умножителей

Вариант	Функция динамической обработки	Детектор уровня	Режим обработки
1	Ограничитель (лимитер)	Абсолютный, экспоненциальный, см. (1)	Блочный
2	Компрессор	Среднеквадратичный, экспоненциальный, см. формулу (2)	Отсчет-за-отсчетом
3	Экспандер	Абсолютный, непосредственный, см. формулу (4)	Отсчет-за-отсчетом
4	Ограничитель шума	Среднеквадратичный, непосредственный, см. формулу (3)	Блочный
5	Ограничитель шума	Абсолютный, экспоненциальный, см. (1)	Блочный
6	Экспандер	Абсолютный, непосредственный, см. формулу (4)	Отсчет-за-отсчетом
7	Компрессор	Среднеквадратичный, экспоненциальный,	Блочный

		см. формулу (2)	
8	Ограничитель (лимитер)	Среднеквадратичный, непосредственный, см. формулу (3)	Отсчет-за- отсчетом

### 2.1 Порядок выполнения работы

1. Изучить теоретические сведения по теме лабораторной работы.
2. Получить у преподавателя задание для выполнения практической части работы.
3. Реализовать в Matlab динамический процессор в соответствии с вариантом.
4. Показать результат работы устройства преподавателю.
5. Оформить и защитить отчет по лабораторной работе.

### 2.2 Содержание отчета

1. Цель работы.
2. Краткие теоретические сведения.
3. Matlab-описание динамического процессора с комментариями.
4. Графики, подтверждающие правильность работы алгоритма.
5. Выводы по работе.