

Лекция 5-6. Звуковой сигнал как случайный процесс

План лекции

- 1. Статистические характеристики звуковых сигналов (ЗС).**
- 2. Законы распределения уровней ЗС.**
- 3. Спектры, спектральная мощность ЗС.**
- 4. Динамический диапазон и средняя мощность ЗС.**

Звуковое колебание
(давление p , сила
звука I , звуковая
мощность P)

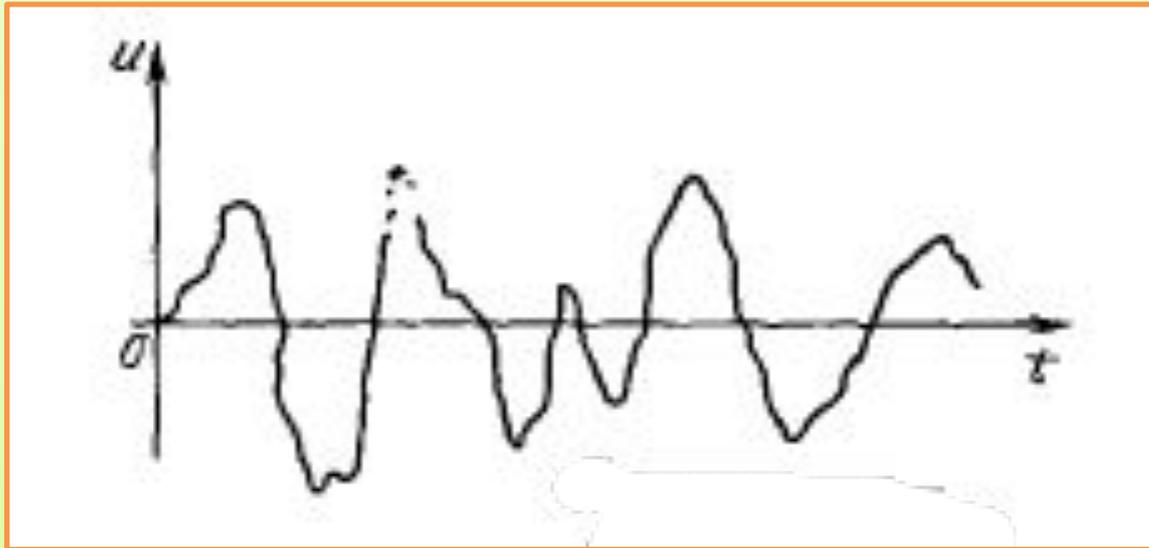
Цепи системы звукового вещания

Электрический
сигнал
(напряжение, ток,
электрическая
мощность)

1. Звуковой сигнал (ЗС) представляет собой случайный процесс, а его акустические или электрические величины изменяются со временем непрерывно.

Его графическое представление изображается совокупностью **реализаций** (отрезков) случайных функций.

Пусть каждая реализация есть изменяющееся во времени t напряжение u (или давление p) за определенный интервал $T_{ин}$ – это будет кривая изменения мгновенного значения ЗС во времени:



- Уровень ЗС – характеристика сигнала в определенный момент времени и представляет собой выпрямленное и усредненное за некоторый предшествующий промежуток времени t_1 напряжение $\overline{u(t_1)}$, отнесенное к некоторой условной величине U_0

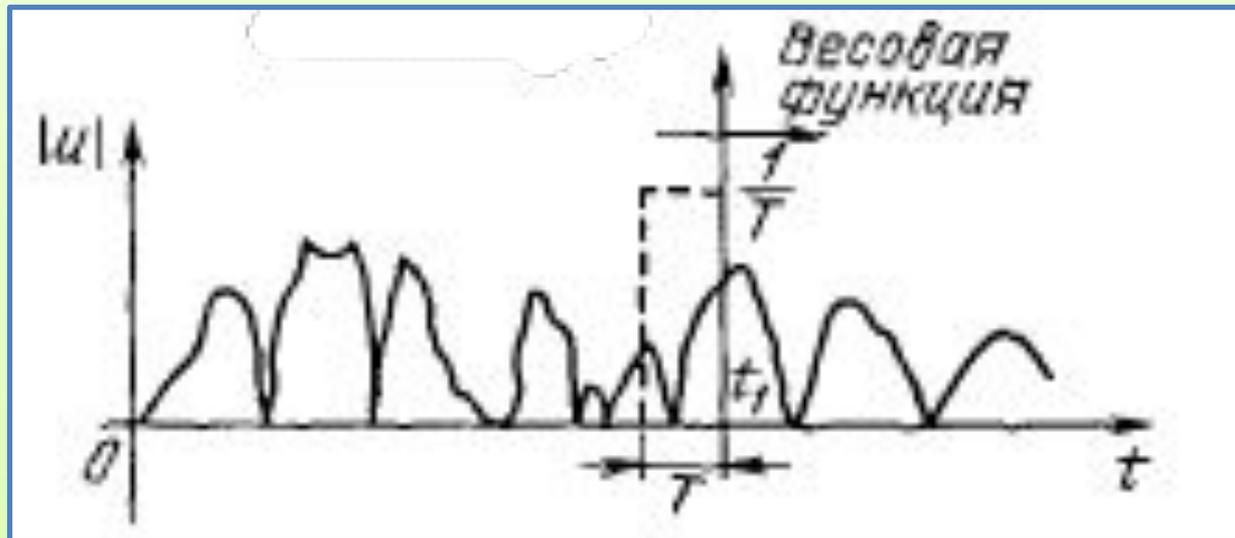
$$L_3(t_1) = 20 \lg \left(\frac{\overline{u(t_1)}}{U_0} \right)$$

Аналогично определяется уровень $L_3(t_2)$ в момент времени t_2 и т.д.

Что такое среднее значение сигнала? Обычно усреднение мгновенного значения проводится с постоянным весовым коэффициентом, т.е.

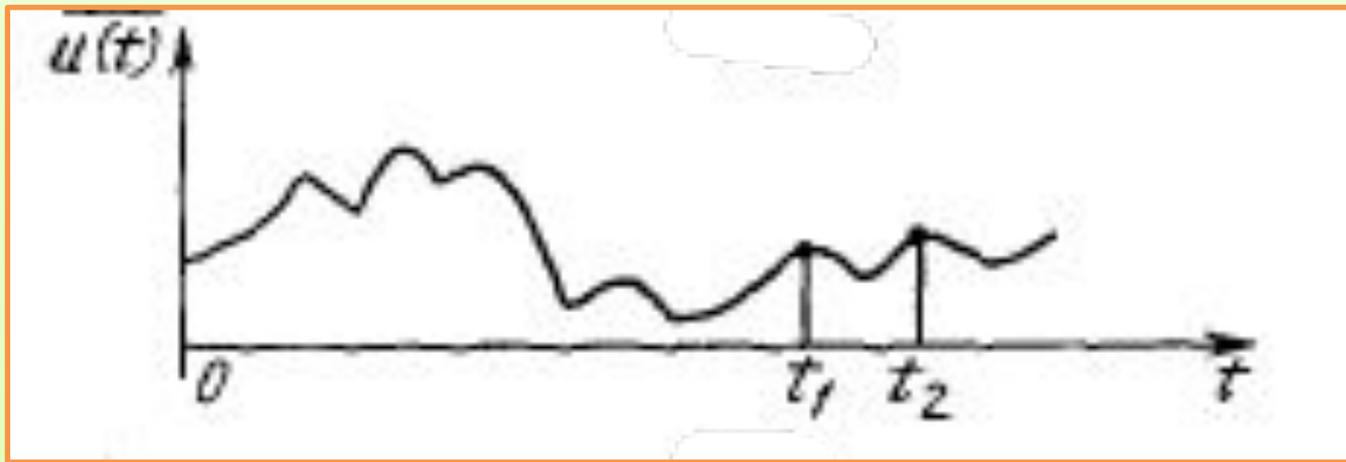
$$\overline{u(t_1)} = \frac{1}{T} \int_{t_1-T}^{t_1} |u(t)| dt$$

• Таким образом, это выражение дает среднее за время T значение модуля $u(t)$, причем всем выпрямленным значениям $u(t)$ приписывается один и тот же вес dt/T в интервале $t_1 - T \dots t_1$ и нулю вне этого интервала (T – интервал усреднения):



Аналогично и для других моментов времени t_2 , t_3 и т.д.

•Т.к. ЗС – нестационарный случайный процесс, то полученные таким образом для разных моментов времени t_1, t_2, \dots, t_n значения $\overline{u(t_1)}, \overline{u(t_2)}, \dots, \overline{u(t_n)}$ будут в общем случае разными и дадут зависимость изменения во времени среднего значения выпрямленного напряжения:



При $T \rightarrow 0$ зависимость средних не отличается от временных зависимостей мгновенных значений. При больших T средние значения тем меньше меняются во времени, чем больше T .

• Если существует минимальный интервал усреднения $T = T_0$, при котором среднее значение выпрямленного напряжения не меняется со временем, т.е. при $T \geq T_0$:

$$\overline{u(t_1)} = \overline{u(t_2)} = \dots = \overline{u(t_n)} \quad (1)$$

Такой сигнал называется стационарным, а время T_0 - интервал стационарности. Обычно это значение около 2...3 мин, но для речевых сигналов меньше, чем для музыкальных, т.к. слух человека не может усреднять сигнал за такой большой интервал времени. Слуховое ощущение в каждый момент времени (например, t_1) определяется не только мгновенным значением сигнала в этот момент, но и более ранними его значениями, но не сильно удаленными по времени от данного момента.

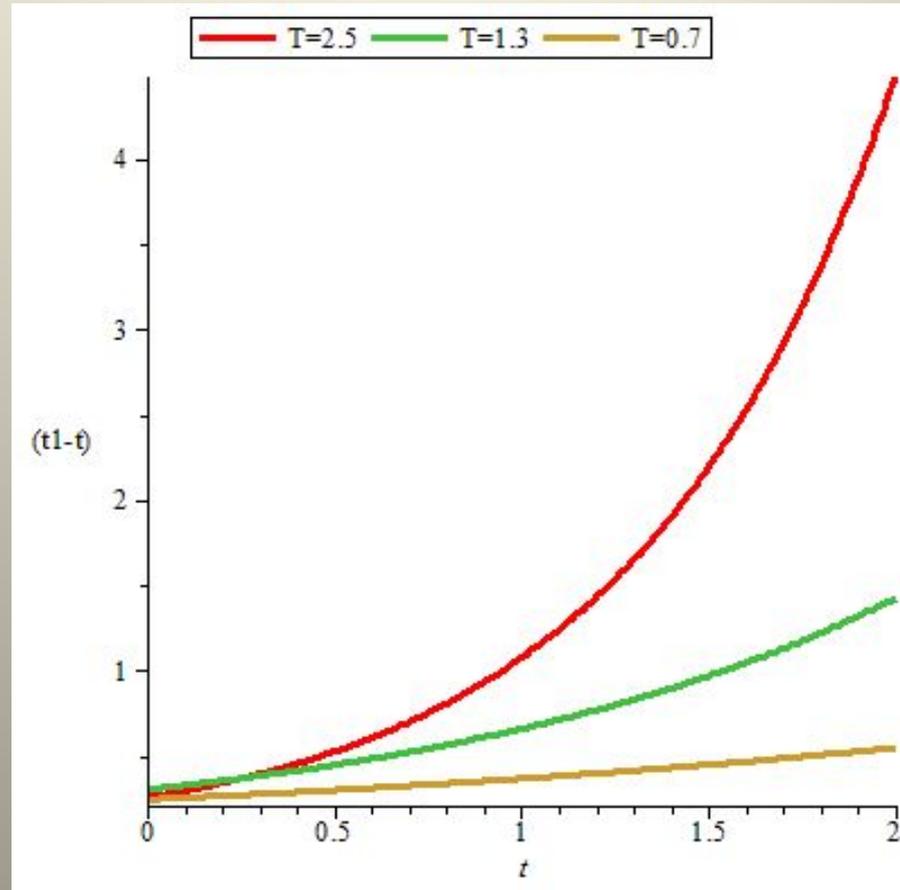
• Значит, при моделировании ЗС и определении его уровня усреднение его мгновенных значений надо выполнять с переменным множителем веса. Такой функцией является

$$\lambda(t_1 - t) = \frac{1}{T} \exp\left(-\frac{t_1 - t}{T}\right) \text{ при } t \leq t_1 \quad (2)$$

где $\lambda(t_1 - t) = 0$ при $t > t_1$; T – длительность предшествующего интервала времени (временная «память» системы).

На следующем рисунке показано, как зависит весовая функция от длительности предшествующего интервала T .

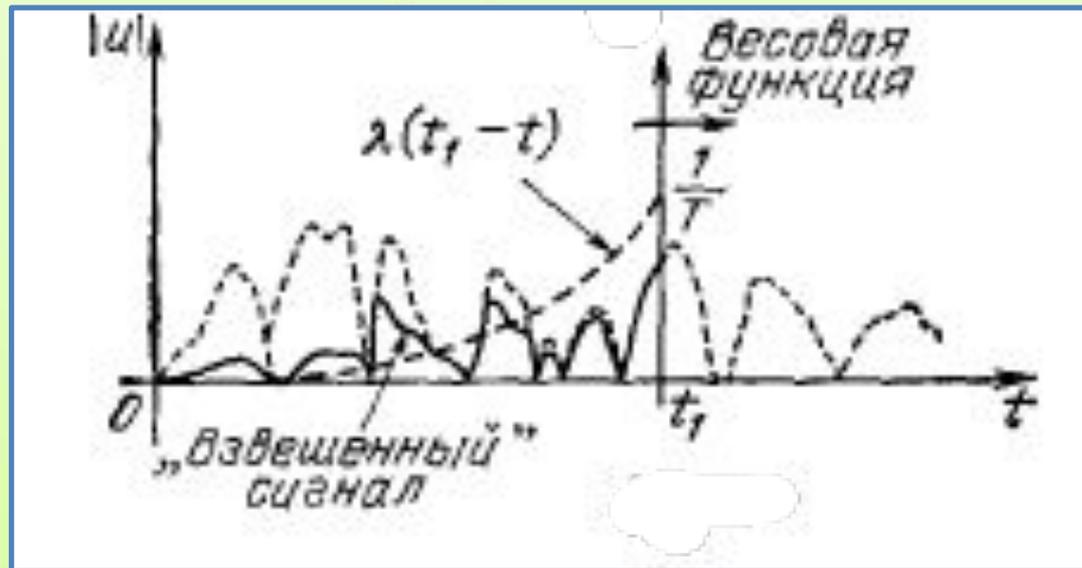
- Временная зависимость весовой функции $\lambda(t_1 - t)$ при различных интервалах T



• При таком способе усреднения для момента времени t_1 среднее значение сигнала

$$\overline{u(t_1)} = \int_{-\infty}^{t_1} \lambda(t_1 - t) |u(t)| dt$$

Взвешивание выпрямленных мгновенных значений реального ЗС показано на рисунке:



• Среднее выпрямленное значение напряжения, усредненное за промежуток времени с весовой функцией $\lambda(t_1 - t)$ и выраженное в децибелах, называется **динамическим уровнем ЗС**

$$L_3(e) = 20 \lg \left(\frac{\overline{u(t)}}{U_0} \right)$$

Иначе – изменяющиеся во времени уровни ЗС называются динамическими.

Вывод: значение уровня зависит не только от мгновенных значений сигнала, но и веса и от длительности T , т.е. временных параметров прибора, которым производилось измерение.

Что дает знание статистических свойств ЗС?

1. Понимание процессов, протекающих как в отдельных устройствах, так и во всей системе звукового вещания.
2. Правильная трактовка свойств и параметров звуковещательной аппаратуры.
3. Создание испытательных сигналов, максимально имитирующих реальные ЗС.
4. Разработка оптимальных систем и устройств, в наилучшей степени учитывающих как свойства ЗС, так и особенности их слухового восприятия.

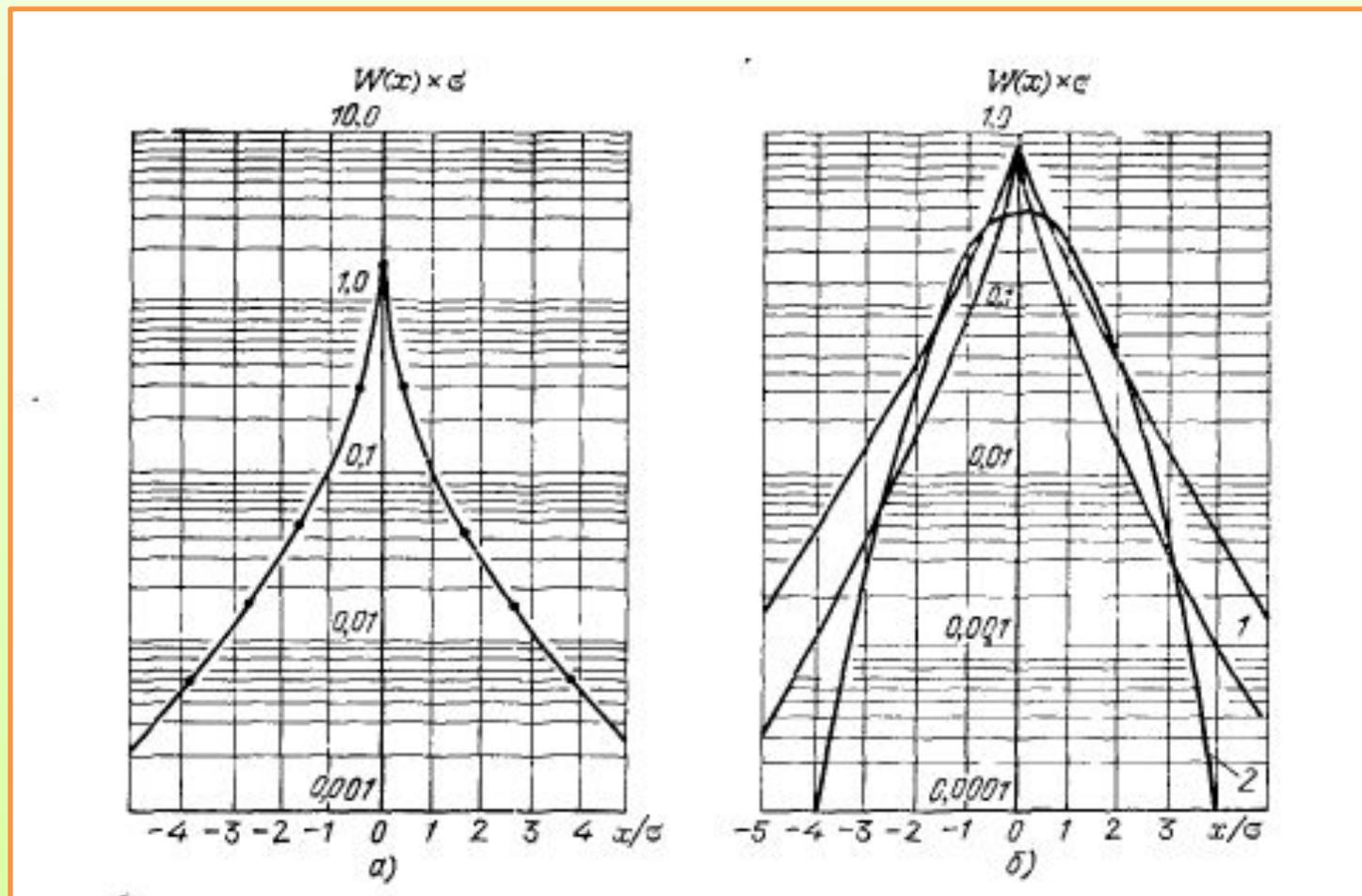
2. Статистические свойства ЗС характеризуются прежде всего законами распределения мгновенных значений и уровней звуковых сигналов (напряжений) во времени; длительностью пауз и длительностей непрерывного существования разных уровней.

Напомним: любая случайная величина X полностью описывается **функцией распределения $p(x)$** , которая есть вероятность события $X < x$, где x – некоторое текущее значение случайно величины. Функция $p(x) = p(X < x)$ – одномерная функция распределения случайной величины. Производная $W(x)$ от этой функции – одномерная **плотность вероятности** распределения случайной величины x , которая характеризует вероятность того, что случайная величина окажется распределенной в интервале от x до $x + \Delta x$.

- Для ЗС случайными величинами являются мгновенные значения напряжения u или звукового давления, уровни L_3 .

Законы распределения $p(x)$ или $W(x)$ в этом случае устойчивы, если длительность реализации больше интервала стационарности T_0 . На практике с достаточной точностью речевой сигнал рассматривается как квазистационарный случайный процесс при длительности анализа $T_{ин}$ больше 2...3 мин. А для музыкальных ЗС такой интервал не устанавливается, т.е. для получения устойчивых законов распределения музыкальных сигналов различных жанров следует существенно увеличить время анализа $T_{ин}$ - до нескольких часов. При коротком времени $T_{ин}$ нет устойчивости и анализ дает свойства только отдельных отрывков. Более того, анализируемые сигналы обрабатываются звукорежиссером и аппаратурой с определенными техническими характеристиками.

На рисунке 1 приведены опытные данные речевого (а) и музыкальных (б) ЗС:



Ось ординат – произведение плотности вероятности $W(x)$ и среднеквадратического отклонения σ (рассеяние значений случайной величины около ее математического ожидания); ось абсцисс – отношение мгновенных значений (например, u) к σ .

Речевой сигнал имеет устойчивое распределение плотности вероятности мгновенных значений, оно симметрично, одновершинное и имеет экспоненциальный характер

•Из рис. б видно, что разным типам звучаний (вокал, эстрадная музыка, симфоническая музыка и т.д.) соответствуют сходные по форме зависимости (область 1) и, кроме кривой 2, все они имеют экспоненциальный характер.

Кривая 2 характерна для хора с оркестром, джазовых композиций и хорошо аппроксимируется гауссовым законом распределения

$$W(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right).$$

Существуют обобщенные законы распределения музыкальных ЗС, основанные на экспериментах различных авторов

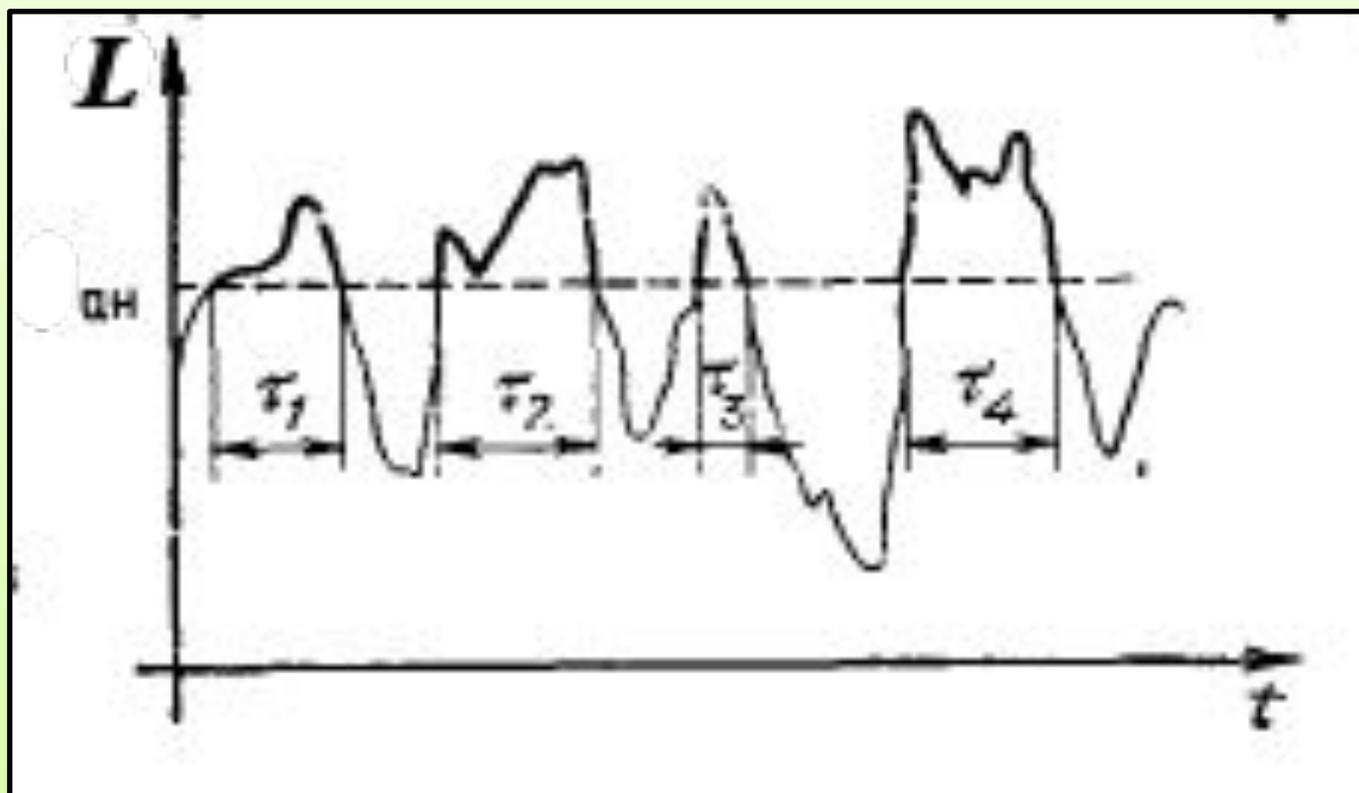
- Распределение уровне во времени реальных ЗС (речевых и музыкальных) зависит и от типа программы и длительности времени анализа, а также очень сильно от выбранной весовой функции. При достаточно малой длительности «памяти» T зависимости близки к рассмотренным выше, а увеличение T приближает закон распределения к гауссовскому:

$$p(L) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^L \exp\left(-\frac{(L-a)^2}{2\sigma^2}\right) dL \quad (3)$$

с параметрами $a = -24$ дБ, $\sigma = 10$ дБ.

При $T \rightarrow \infty$ вместо совокупности случайных величин получим только одно значение

- 3. Из можно узнать время τ , в течение которого достигается или превышает некоторое значение уровня анализа $L_{ан}$



• Но эти зависимости не могут сказать – из какого числа отрезков складывается общее время τ . Это важно практически, т.к. надо выбрать режим работы ламп (транзисторов) оконечного каскада усилителей системы звукового вещания большой мощности, при ручном или автоматическом регулировании уровней и т.п. Здесь также имеются экспериментальные исследования и соответствующие данные приводятся в литературе.

Отметим следующее – статистические зависимости плотности вероятности $W(\tau)$ длительности непрерывного существования (превышения) уровней $L > L_{ан}$ аппроксимируются так называемым гамма-распределением при $0 < \tau \leq 100$ мс

$$W(\tau) = \mu \frac{\tau}{\nu^2} \exp\left(-\frac{\tau}{\nu}\right)$$

• В этой формуле τ – длительность существования уровней (выбросов), с; ν – параметр, определяемый отношением $L_{ан}/L_{кв.макс}$ ($L_{кв.макс}$ – квазимаксимальный уровень ЗС, т.е. вероятность его превышения составляет 0,02); μ – параметр, зависящий от $L_{ан}$.

Пауза – это длительности выбросов, лежащих ниже квазимиимального уровня (вероятность его превышения 0,98) анализа ЗС. Для определения закона распределения пауз по длительности уровень анализа выбирается ниже номинального (максимально допустимый уровень, при котором нелинейные искажения устройства не превышают допустимого значения) на 40 дБ, т.к. при отклонении на 42...45 дБ от номинального энергия вещательного сигнала $\rightarrow 0$.

Длительность пауз ЗС колеблется в широких пределах, хотя вероятность появления пауз длительностью больше 2 с мала. Плотность вероятности распределения пауз по длительности наибольшая в интервале 50...150 мс и быстро убывает вне этого интервала. На речевые сигналы в среднем приходится 913 пауз в час, на музыкальные – 200. Длительность пауз с учетом перерыва между передачами составляет 8,6 % (первая общероссийская программа) и 5,1 % (программа «Маяк»).

Среднее время длительности пауз для активной речи существенно выше, чем для пассивной (чтение заранее подготовленного текста информации). Оно составляет (от общего времени передачи) 55 % для студентов и 47,5 % для общественных деятелей.

•3. Пусть $u(\xi)$ – изменение во времени мгновенного значения напряжения ЗС, тогда его текущая мощность

$$P(t) = \frac{1}{T} \int_{-\infty}^t \exp\left(-\frac{t-\xi}{T}\right) u^2(\xi) d\xi$$

где $\exp(\dots)$ - весовая функция. Очевидно, если $P(t)$ – случайная функция, то $W(P)$ есть плотность вероятности распределения текущей мощности во времени.

Существуют экспериментальные аппроксимации закона распределения текущей мощности для речевых и музыкальных сигналов, полученных по множеству реализаций (выборок) сигналов:

• для речевых сигналов

$$W(P) = 0,39 \exp\left(-\frac{P}{\bar{P}}\right) - 0,25 \exp\left(-\frac{5,7P}{\bar{P}}\right);$$

для музыкальных сигналов

$$W(P) = 0,24 \exp\left(-\frac{P}{\bar{P}}\right) - 0,17 \exp\left(-\frac{4,0P}{\bar{P}}\right)$$

\bar{P} - среднее значение текущей мощности и есть математическое ожидание случайной величины P . Для большинства жанров ЗС наиболее вероятные значения текущей мощности лежат на 3 дБ ниже долговременной средней мощности \bar{P} . Поэтому при установке регуляторов уровня ориентируются на значение \bar{P} .

- Уровень среднеминутной мощности определяется из выражения

$$L_{\text{ср.мин}} = 10 \lg(P_{\text{ср.мин}}/P) = 10 \lg(U_{\text{ср.мин}}^2/U_{\text{sin}}^2)$$

U_{sin} - эффективное значение синусоидально напряжения (частотой 1000 Гц), равное квазимаксимальному значению $L_{\text{кв.макс}}$; $U_{\text{ср.мин}}$ - среднее значение напряжения за время анализа (интервал наблюдения $T_{\text{ин}}$), равное 1 мин.

Для среднеминутной мощности также можно построить график закона распределения в зависимости от жанров ЗС. Он показывает с достаточной для практических применений гауссовский закон распределения случайно величины.

- Если ЗС пропустить через набор полосовых фильтров, охватывающих весь диапазон слышимых частот, то можно для него измерить плотности вероятности текущей мощности $W_i(P_i)$ как в выделенных узких полосах частот ΔF_i , так и в широкой полной ΔF полосе $W(P)$.

Для такого набора данных $W_1(P_1), W_2(P_2), \dots, W_n(P_n)$ и $W(P)$, можно определить значения текущей мощности P_1, P_2, \dots, P_n , вероятность превышения которых имеет заданное значение (обычно это 0,02). Тогда совокупность значений

$$L_i = 10 \lg(P_i/P)$$

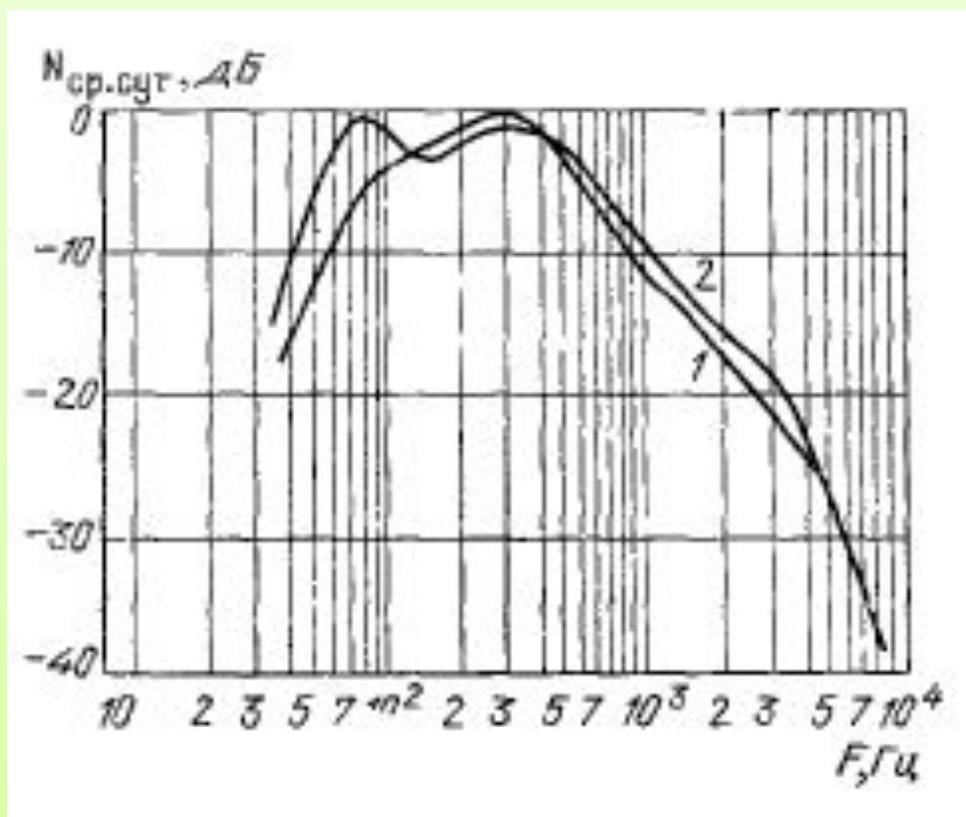
Это распределение максимальных уровней текущей мощности для совокупности выделенных полос частот.

- Если значения L_i отнести к ширине выделенной полосы частот ΔF_i и взять отношение

$$L_{fi} = 10 \lg \left(\frac{\Delta F P_i}{\Delta F_i P} \right)$$

то можно определить уровень спектральной плотности максимальной (средней) мощности или распределение соответствующих уровней (максимальных, средних) по частоте для исследуемого отрывка звукового сигнала.

На рисунке приведены среднестатистические суточные графики спектральной плотности мощности первой (1) и второй (2) общероссийских программ проводного радиовещания:



На графике уровень спектральной плотности мощности уменьшается на низких (ниже 70 Гц) и верхних (выше 3000 Гц) частотах, причем 0 дБ соответствует значению максимальной спектральной плотности мощности, достигаемой в сигнал этих программ. Формы зависимостей определяются в основном спектральной плотностью мощности сигнала того жанра, который в данной программе занимает наибольшее время и имеет максимальную мощность. Например, для первой программы определяющей является речь, а для второй оркестровая музыка и речь.

Отметим, что применение необычных способов инструментовки произведений современной музыки, а также электронных методов преобразования сигналов (синтезаторы) расширяет спектры музыкальных сигналов до 15000 ... 16000 Гц. Это предъявляет высокие технические требования к студийной аудиоаппаратуре

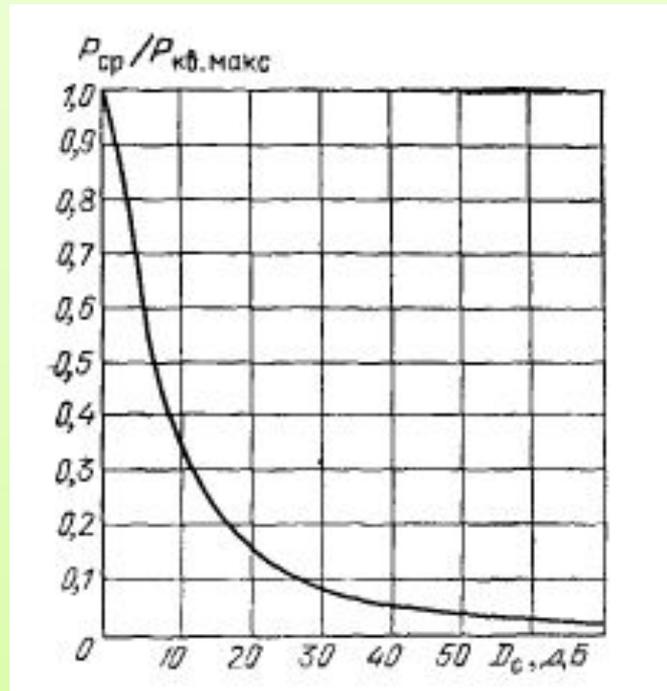
- 4. Общее определение динамического диапазона D уже приводилось как отношение квазимаксимального и квазимиимального уровней ЗС, а международный комитет по радиовещанию и телевидению вводит определение динамического диапазона как отношение максимального давления к минимальному, но не разъясняет их смысл. Такое определение D называется теоретическим. Но можно определить его и экспериментально из уровнеграммы как разность максимально и минимального уровней для достаточно большого интервала наблюдения

$$D_c = L_{\text{макс}} - L_{\text{мин}}$$

• Ясно, что это значение D_c зависит сильно от выбранного времени интегрирования измерителя уровней, т.е. чем больше это время, тем меньше будут регистрироваться кратковременные выбросы ЗС и тем меньшим окажется найденное значение D_c .

Очевидно – передача ЗС без искажений по каналу звукового вещания возможна при условии $D_c < D_k$, где D_k - динамический диапазон канала, зависящий от уровня шумов в тракте и от номинального напряжения. Так как для реальных ЗС это часто не выполняется, то необходима предварительная обработка ЗС.

- При распределении уровней ЗС по закону Гаусса можно определить отношение средней мощности сигнала звукового вещания к квазimaxимальной и оказывается, что это отношение зависит только от значения D_c :



• При реальных значениях $D_c = 20 \dots 40$ дБ в трактах звукового вещания отношение $P_{\text{ср}}/P_{\text{кв.макс}}$ меняется в пределах $0,053 \dots 0,155$ и остается меньше 1.

Малое отношение средней мощности к максимальной говорит о том, что режим номинальной мощности достигается лишь в течение незначительной части общего времени передачи. Это учитывают при разработке аппаратуры, особенно мощных вещательных усилителей, развивающих номинальную мощность в течение коротких (не более $10 \dots 20$ мс) промежутков времени.