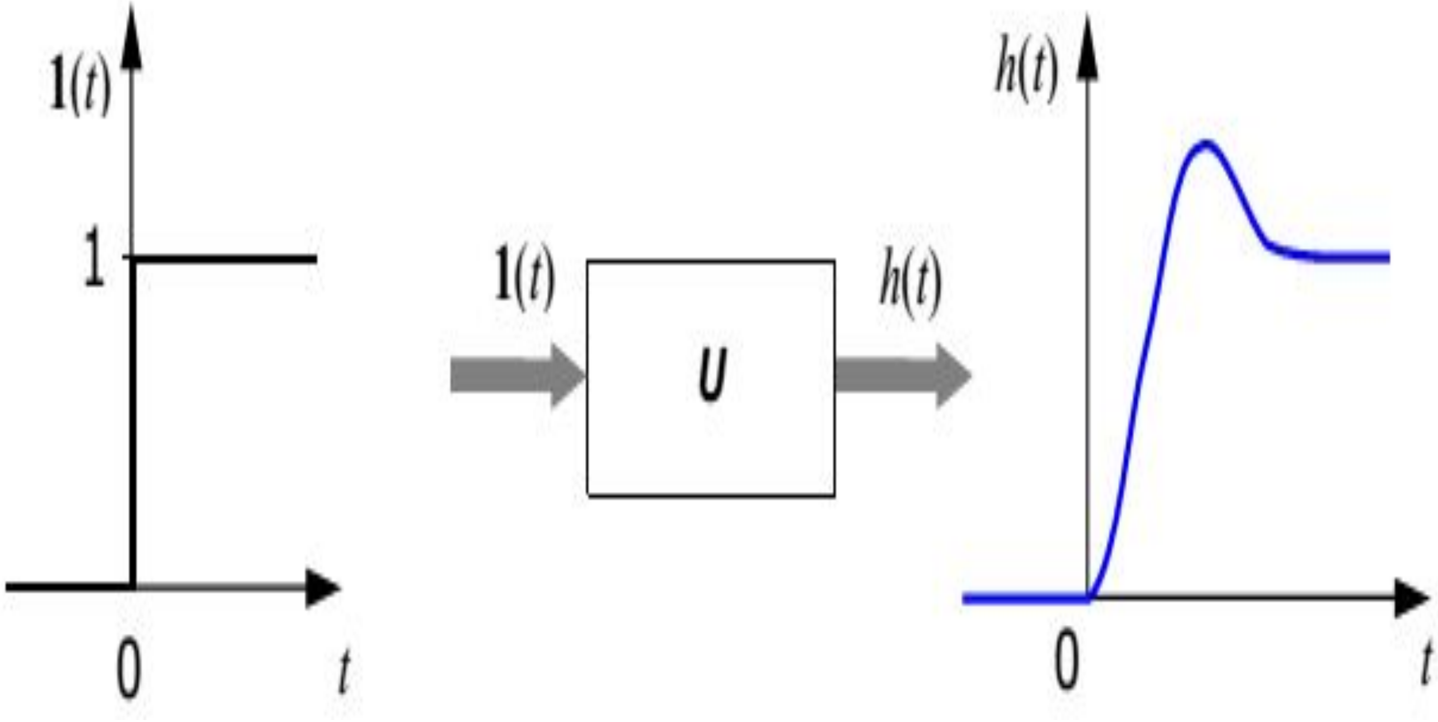
Переходная функция

Один из методов построения моделей «вход-выход» – определение реакции объекта на некоторый стандартный сигнал. Один из простейших сигналов – так называемый «единичный скачок» («единичный ступенчатый сигнал»), то есть мгновенное изменение входного сигнала с 0 до 1 в момент $t = 0$. Формально этот сигнал определяется так:

$$1(t) = \begin{cases} 0, t < 0 \\ 1, t \geq 0 \end{cases}$$

Реакция объекта на единичный скачок называется **переходной функцией** и обозначается $h(t)$:



Пусть модель объекта задана дифференциальным уравнением первого порядка:

$$T\frac{dy(t)}{dt}+y(t)=k\cdot x(t)\,, \tag{16}$$

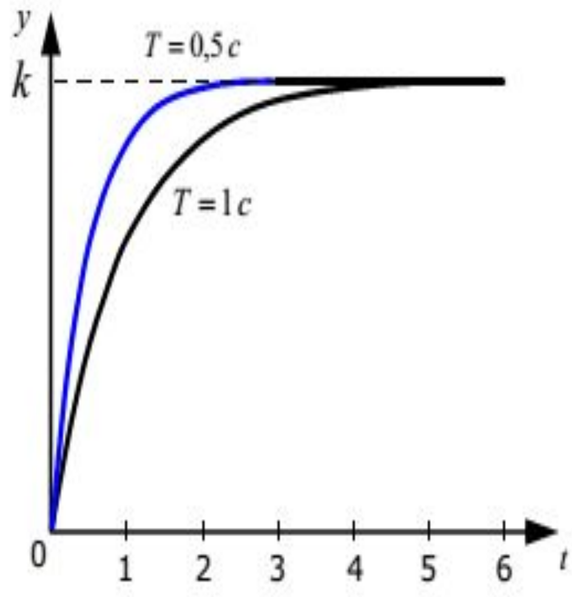
где k – безразмерный коэффициент, а T – некоторая постоянная, которая имеет размерность времени (измеряется в *секундах*). Найдем переходную характеристику этого звена. Решая уравнение (16) при $x(t)=1$ ($t>0$), получаем

$$y(t)=k+C_1\cdot\exp\left(-\frac{t}{T}\right),$$

где постоянная C_1 должна определяться из начальных условий. Поскольку нас интересует переходная характеристика, начальные условия считаем нулевыми, то есть $y(0)=0$, что дает $C_1=-k$ и поэтому

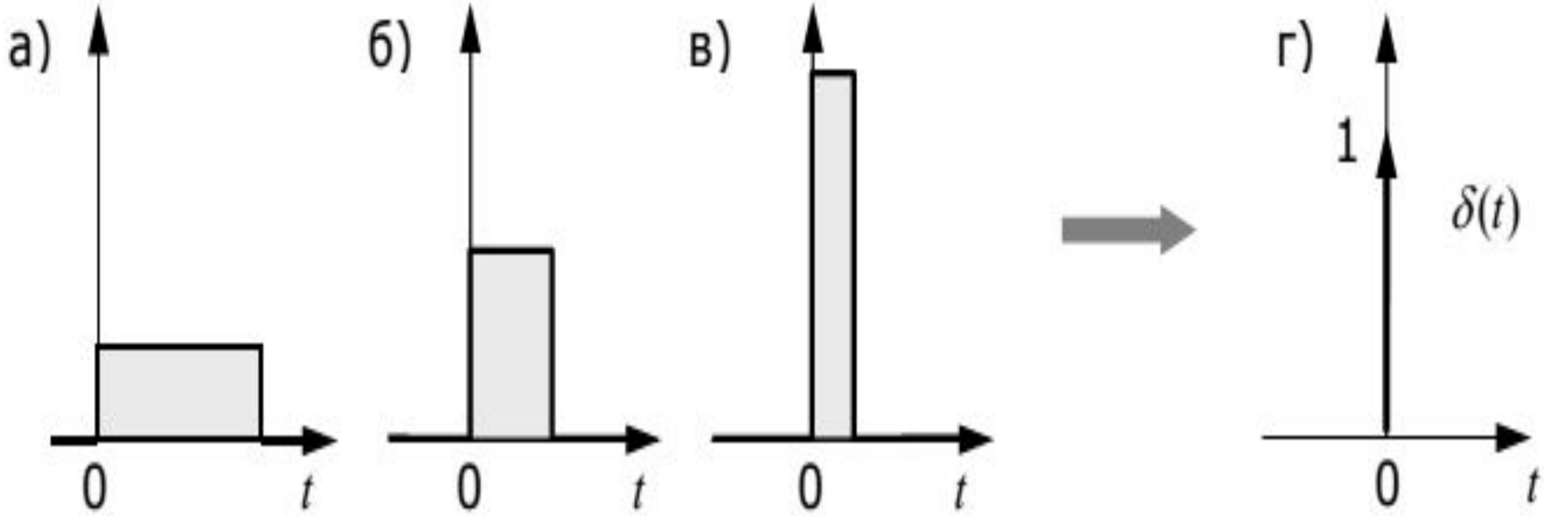
$$h(t)=y(t)=k\left[1-\exp\left(-\frac{t}{T}\right)\right]. \tag{17}$$

На рисунке показаны переходные характеристики (17) при различных значениях параметра T, который называется **постоянной времени** звена:



Импульсная характеристика

В качестве тестового сигнала можно, в принципе, использовать любой сигнал. Например, можно изучать реакцию системы на прямоугольный импульс. Вопрос в том, чтобы определить некоторый стандартный вид этого импульса. На рисунках а)-в) показаны три импульса, имеющих одинаковые площади. Для простоты будем считать, что эта площадь равна единице.



Импульсная характеристика

Что будет, если мы будем уменьшать ширину импульса, сохраняя его площадь? Очевидно, что высота импульса будет расти и в пределе (когда ширина стремится к нулю) станет бесконечной. Таким образом, мы получили еще один классический тестовый сигнал – *единичный импульс* или *дельта-функцию Дирака* $\delta(t)$. Это идеальный (невозможный в реальной жизни) сигнал, который равен нулю во всех точках, кроме $t = 0$, где он уходит к бесконечность, причем его площадь (интеграл по всей оси времени) равен единице:

$$\delta(t) = \begin{cases} \infty, & t = 0 \\ 0, & t \neq 0 \end{cases}, \qquad \int_{-\infty}^{\infty} \delta(t)\, dt = 1.$$

Импульсная характеристика

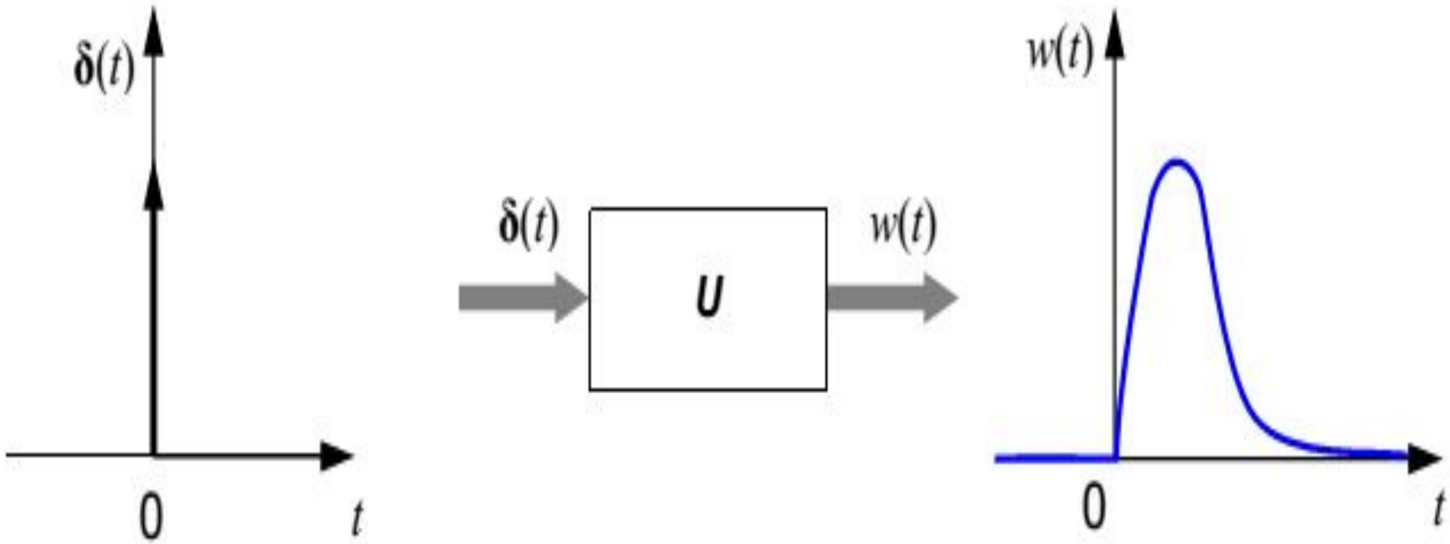
Поскольку бесконечный импульс невозможно нарисовать, на графике он изображается стрелкой, высота которой равна единице (см. рисунок г).

Иногда определяют дельта-функцию как производную от единичного ступенчатого сигнала $\mathbf{1}(t)$. Действительно, эта производная равна нулю при всех значениях t, кроме нуля, где она обращается в бесконечность.

Реакция системы на единичный импульс (дельта-функцию) называется **импульсной характеристикой** и обозначается $w(t)$:

Импульсная характеристика

Реакция системы на единичный импульс (дельта-функцию) называется **импульсной характеристикой** и обозначается $w(t)$:



Импульсная характеристика

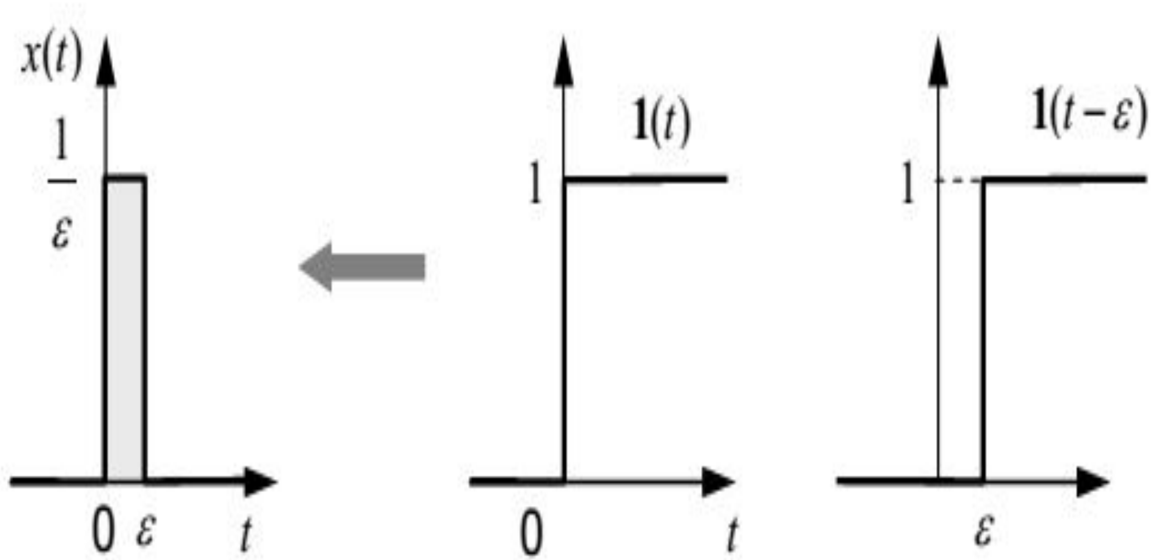
Рассматривая дельта-функцию как предельный случай прямоугольного сигнала единичной площади, можно найти связь между переходной функцией и импульсной характеристикой.

Пусть ширина прямоугольного импульса равна ε, а высота – $1/\varepsilon$. Такой импульс можно представить в виде разности двух ступенчатых сигналов

$$x(t) = \frac{1}{\varepsilon}[\mathbf{1}(t) - \mathbf{1}(t-\varepsilon)],$$

где $\mathbf{1}(t-\varepsilon)$ – это единичный ступенчатый сигнал, который приходит в момент $t=\varepsilon$, то есть, смещен по времени на ε (см. рисунок далее).

Импульсная характеристика



Импульсная характеристика

Так как для линейных систем справедлив принцип суперпозиции, сигнал на выходе будет равен разности реакций системы на входы $\mathbf{1}(t)$ и $\mathbf{1}(t-\varepsilon)$, умноженной на коэффициент $1/\varepsilon$. Учитывая, что реакция на сигнал $\mathbf{1}(t)$ – это переходная функция $h(t)$, получаем

$$y(t) = \frac{1}{\varepsilon}[h(t) - h(t-\varepsilon)].$$

Переходя к пределу при $\varepsilon \to 0$

$$w(t) = \lim_{\varepsilon \to 0} \frac{h(t) - h(t-\varepsilon)}{\varepsilon} = \frac{dh(t)}{dt}$$

переходная функция –

$$h(t) = \int_{0}^{t} w(\tau)\, d\tau \, .$$

Весовая функция

Дифференцируя переходную характеристику (17) звена первого порядка, получаем соответствующую импульсную характеристику:

$$w(t) = \frac{d}{dt}\left(k\left[1-\exp\left(-\frac{t}{T}\right)\right]\right) = \frac{k}{T}\exp\left(-\frac{t}{T}\right).$$

Другое название импульсной характеристики – *весовая функция*. Это название связано с тем, что для произвольного входного сигнала $x(t)$ выход системы $y(t)$ при нулевых начальных условиях вычисляется как интеграл

$$y(t) = \int_{-\infty}^{t} x(\tau)\, w(t-\tau)\, d\tau = \int_{0}^{\infty} x(t-\tau)\, w(\tau)\, d\tau\ .$$

Здесь функция $w(t)$ как бы «взвешивает» входной сигнал $x(t)$ в подынтегральном выражении.

Заметим, что импульсная характеристика дает неполную информацию об объекте, поскольку не учитывает ненулевые начальные условия.

Передаточная функция

Пусть модель объекта задана линейным дифференциальным уравнением второго порядка, связывающим вход $x(t)$ и выход $y(t)$:

$$b_2 \frac{d^2 y(t)}{dt^2} + b_1 \frac{dy(t)}{dt} + b_0 y(t) = a_1 \frac{dx(t)}{dt} + a_0 x(t) \qquad (18)$$

где a_i $(i = 0{,}1)$ и b_i $(i = 0{,}1{,}2)$ – постоянные.

Введем *оператор дифференцирования* $p = \frac{d}{dt}$, который действует на сигнал $x(t)$ по правилу $p\,x(t) = \frac{dx(t)}{dt}$. Обратите внимание, что запись $p\,x(t)$ обозначает *не умножение* оператора p на $x(t)$, а *действие* этого оператора, то есть дифференцирование $x(t)$.

Теперь запишем производные сигналов $x(t)$ и $y(t)$ по времени в операторной форме

$$\dot{y}(t) = \frac{dy(t)}{dt} = py(t), \quad \ddot{y}(t) = \frac{d^2 y(t)}{dt^2} = p^2 y(t), \quad \dot{x}(t) = \frac{dx(t)}{dt} = px(t).$$

Подставляя эти выражения в (18), получим

$$b_2 p^2 y(t) + b_1 py(t) + b_0 y(t) = a_1 px(t) + a_0 x(t). \qquad (19)$$

Передаточная функция

Подставляя эти выражения в (18), получим

$$b_2 p^2 y(t) + b_1 p y(t) + b_0 y(t) = a_1 p x(t) + a_0 x(t)\,. \tag{19}$$

Можно формально вынести за скобки $y(t)$ в левой части равенства (19) и $x(t)$ в правой части:

$$(b_2 p^2 + b_1 p + b_0)\; y(t) = (a_1 p + a_0)\; x(t)\,. \tag{20}$$

Левая часть (20) означает, что оператор $b_2 p^2 + b_1 p + b_0$ действует на сигнал $y(t)$, а в правой части оператор $a_1 p + a_0$ действует на сигнал $x(t)$. «Разделив» (условно, конечно) обе части (20) на оператор $b_2 p^2 + b_1 p + b_0$, связь выхода и входа можно записать в виде

$$y(t) = \frac{a_1 p + a_0}{b_2 p^2 + b_1 p + b_0} x(t) = W(p)\, x(t)\,, \tag{21}$$

где запись $W(p)\, x(t)$ означает не умножение, а действие сложного оператора

$$W(p) = \frac{a_1 p + a_0}{b_2 p^2 + b_1 p + b_0}\,. \tag{22}$$

на сигнал $x(t)$. Иначе говоря, формула $y(t) = W(p)\, x(t)$ – это не что иное, как символическая запись уравнения (18), которую удобно использовать.

Передаточная функция

Часто передаточной функцией называют функцию $W(\lambda)$, которая получается из (22) в результате замены оператора p на некоторую независимую переменную λ. Эта фукнция представляет собой отношение двух полиномов (многочленов) от λ.

Передаточная функция $W(\lambda)$ называется **правильной**, если степень ее числителя *не больше*, чем степень знаменателя; **строго правильной**, если степень числителя *меньше* степени знаменателя; **неправильной**, если степень числителя *больше*, чем степень знаменателя. Например, функция $\dfrac{1}{\lambda+1}$ – строго правильная и одновременно правильная; $\dfrac{\lambda}{\lambda+1}$ – правильная, но не строго правильная (иногда такие функции называют *биправильными*), а $\dfrac{\lambda^2+\lambda+1}{\lambda+1}$ – неправильная.

Нулями передаточной функции называются корни ее числителя, а **полюсами** – корни знаменателя. Например, функция $W(\lambda)=\dfrac{\lambda-1}{\lambda^2+3\lambda+2}$ имеет нуль в точке $\lambda=1$ и два полюса в точках $\lambda=-1$ и $\lambda=-2$.

Передаточная функция и преобразование лапласа

Рассмотрим снова уравнение (18):

$$b_2 \frac{d^2 y(t)}{dt^2} + b_1 \frac{dy(t)}{dt} + b_0 y(t) = a_1 \frac{dx(t)}{dt} + a_0 x(t) \qquad (29)$$

Применим к левой и правой частям преобразование Лапласа, считая, что все начальные условия нулевые. Получается *уравнение в изображениях*, связывающее преобразования Лапласа входа $X(s)$ и выхода $Y(s)$:

$$b_2 \cdot s^2 Y(s) + b_1 \cdot sY(s) + b_0 \cdot Y(s) = a_1 \cdot sX(s) + a_0 \cdot X(s)$$

Можно вынести за скобки $Y(s)$ в левой части и $X(s)$ в правой части:

$$(b_2 s^2 + b_1 s + b_0) \cdot Y(s) = (a_1 s + a_0) \cdot X(s).$$

Разделив обе части этого равенства на $b_2 s^2 + b_1 s + b_0$, получаем

$$Y(s) = \frac{a_1 s + a_0}{b_2 s^2 + b_1 s + b_0} \cdot X(s) = W(s) \cdot X(s), \quad \text{где } W(s) = \frac{a_1 s + a_0}{b_2 s^2 + b_1 s + b_0}. \qquad (30)$$

Передаточная функция и преобразование лапласа

Сравнение (22) и (30) показывает, что $W(s)$ – это *передаточная функция* объекта, записанная в виде функции от комплексной переменной s, а не от оператора дифференцирования p, как в (22).

Таким образом, при нулевых начальных условиях *изображение выхода линейного объекта вычисляется как произведение его передаточной функции на изображение входного сигнала*.

Из (30) следует и другой важный вывод: *передаточная функция равна отношению изображений по Лапласу выхода и входа при нулевых начальных условиях*.

пример

Рассмотрим пример использования преобразования Лапласа для вычисления выхода системы при известном входном сигнале. Пусть объект управления описывается уравнением первого порядка (16):

$$T\frac{dy(t)}{dt} + y(t) = k \cdot x(t) \tag{31}$$

и на его вход поступает единичный ступенчатый сигнал $x(t) = \mathbf{1}(t)$. Требуется найти сигнал выхода $y(t)$, который в данном случае представляет собой переходную характеристику.

пример

Решим эту задачу с помощью передаточных функций и изображений сигналов по Лапласу. Чтобы найти изображение выхода по формуле (30), нужно знать изображение входного сигнала $X(s)$ и передаточную функцию звена $W(s)$. Изображение входа находим по табличным данным (см. (25)), а передаточную функцию – из (31), повторяя приведенные выше рассуждения:

$$X(s) = \frac{1}{s}, \quad W(s) = \frac{k}{Ts+1}.$$

Теперь находим изображение выхода

$$Y(s) = \frac{1}{s} \cdot \frac{k}{Ts+1} = \frac{k}{s} - \frac{kT}{Ts+1}.$$

и представляем его в виде суммы элементарных дробей:

$$Y(s) = \frac{k}{s} - \frac{k}{s+1/T}.$$

Используя принцип суперпозиции для изображений (27), вычисляем оригинал – сигнал выхода:

$$y(t) = k \cdot \mathcal{L}^{-1}\left\{\frac{1}{s}\right\} - k \cdot \mathcal{L}^{-1}\left\{\frac{1}{s+1/T}\right\}.$$

пример

Обратные преобразования Лапласа находим по таблице

$$y(t) = k - k \cdot \exp\left(-\frac{t}{T}\right) \quad \text{при } t > 0,$$

что совпадает с (17). Таким способом можно вычислять реакцию системы на известный входной сигнал без прямого решения дифференциального уравнения.