18.03 Difference Equations and Z-Transforms

Jeremy Orloff

Difference equations are analogous to 18.03, but without calculus. On the last page is a summary listing the main ideas and giving the familiar 18.03 analog. The two line summary is:

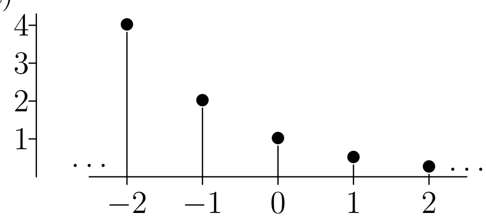
1. In 18.03 the answer is e^{at}, and for difference equations the answer is a^n.

2. The differential operator D has both algebraic and analytic analogs in difference equations.

Sequences $x[n]$ (also called signals or discrete functions)

Examples:

a)

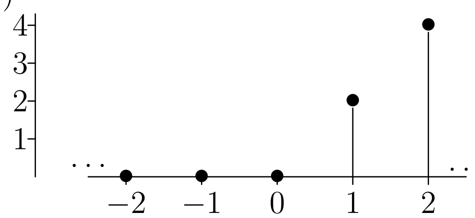


$x[n] = (1/2)^n$

$\Rightarrow \; \ldots\, x[-2] = 4,\; x[-1] = 2,$

$x[0] = 1,\; x[1] = 1/2,\; \ldots$

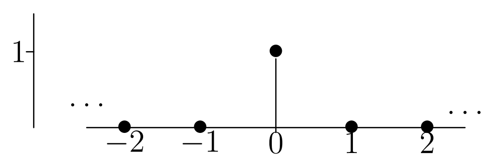
b)



$$x[n] = \begin{cases} 0 & \text{if } n < 0 \\ 2n & \text{if } n \geq 0 \end{cases}$$

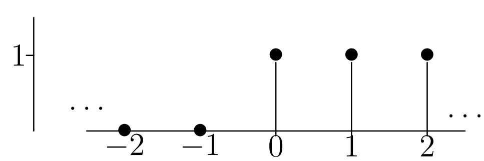
c) **Unit sample:**

$$\delta[n] = \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{otherwise} \end{cases}$$



d) **Unit step:**

$$u[n] = \begin{cases} 0 & \text{if } n < 0 \\ 1 & \text{if } n \geq 0 \end{cases}$$



Difference Equations

Example 1: $y[n] - y[n-1] = x[n]$. (x = input, y = output)

Example 2: $y[n] + 8y[n-1] + 7y[n-2] = x[n]$.

Example 3: $y[n] + 8y[n-1] + 7y[n-2] = x[n] - x[n-1]$.

Example 4: $y[n] - ny[n-1] = x[n]$

Examples 1-3 are constant coefficient equations, i.e. linear time invariant (LTI). Example 4 is not constant coefficient. We will focus on constant coefficient equations.

(continued)

In practice it's easy to compute as many terms of the output as you want: *the difference equation is the algorithm.*

Example: For the difference equation $y[n] - \frac{1}{2}y[n-1] = u[n]$ find $y[n]$ for $n \geq 0$. Assume rest IC $y[-1] = 0$.

(Here $u[n]$ is the unit step function.)

answer: Rewrite the equation as $y[n] = u[n] + \frac{1}{2}y[n-1]$.

Make a table:

n	-1	0	1	2	3	4	...
$u[n]$	0	1	1	1	1	1	...
$y[n]$	0	1	3/2	7/4	15/8	31/16	...

We have already seen difference equations with Euler's formula. For example the IVP $y' = ky$; $y(0) = 1$ becomes the difference equation

$y_{n+1} = y_n + khy_n = (1+kh)y_n \;\Leftrightarrow\; y_{n+1} - (1+kh)y_n = 0.$

Here instead of $y[n]$ we wrote y_n

Z-transform (analog of Laplace transform)

Let $x[n]$ be a sequence. Its z-transform is $X(z) = \sum_n x[n]z^{-n}$.

(analogous to $F(s) = \int_0^\infty f(t)\mathrm{e}^{-st}\,dt$; $\mathrm{e}^{-s} \leftrightarrow z^{-1}$.)

When it's useful we will denote the z-transform of x by $\mathcal{Z}x$ (similar to using $\mathcal{L}x$ for Laplace).

Example 1: z-transform of $\delta[n]$ is 1.

Example 2: z-transform of $u[n]$ is $U(z) = \sum_{n=0}^{\infty} z^{-n} = 1 + z^{-1} + z^{-2} + \ldots = \frac{1}{1-z^{-1}}$.

Example 3: If $x[n] = 0$ for $n < 0$ and $x[n] = a^n$ for $n \geq 0$ then $X(z) = \sum_{n=0}^{\infty} a^n z^{-n} = \frac{1}{1-az^{-1}}$.

Convolution

Start with two sequences $x[n]$ and $y[n]$ their convolution is defined as

$$(x * y)[n] = \sum_k x[k]\,y[n-k].$$

This arises in the following way. $X(z) = \sum_k x[k]z^{-k}$, $Y(z) = \sum_m y[m]z^{-m}$.

$$\Rightarrow X(z)Y(z) = \left(\sum_k x[k]z^{-k}\right)\left(\sum_m y[m]z^{-m}\right) = \sum_k \sum_m x[k]y[m]z^{-k-m} = \sum_n \left(\sum_k x[k]y[n-k]\right) z^{-n}$$

i.e., $X(z)Y(z) = \mathcal{Z}(x * y)$.

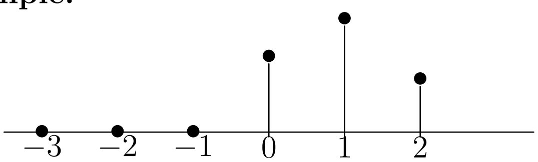
(continued)

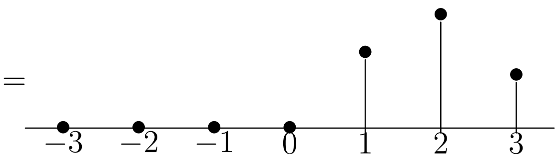
Delay Operator (algebraic analog of D)

Delay operator: $(\mathcal{R}y)[n] = y[n-1]$. (Also called *right shift operator.*)

So, $(\mathcal{R}^2 y)[n] = y[n-2]$ etc. Algebraically we work with $\mathcal{R}$ in difference equations and $\mathcal{Z}$-transforms in much the same way we work with D in differential equations and Laplace transforms.

Example:

$y =$ 

$\mathcal{R}y =$ 

Theorem: (Z-transform of $\mathcal{R}$) $\mathcal{Z}(\mathcal{R}x) = z^{-1}X$.

Proof: $\mathcal{Z}(\mathcal{R}x) = \sum_n (\mathcal{R}x)[n]z^{-n} = \sum_n x[n-1]z^{-n}$

$$= z^{-1}\sum_n x[n-1]z^{-(n-1)} = z^{-1}\sum_m x[m]z^{-(m)} = z^{-1}X(z). \quad \text{QED}$$

General constant coefficient difference equations and the z-transform.

General form: $P(\mathcal{R})y = Q(\mathcal{R})x$. z-transform $P(z^{-1})Y = Q(z^{-1})X$.

x is called the input and y is the output or response.

Example: Example 3 above has $(1 + 8\mathcal{R} + 7\mathcal{R}^2)y = (1-\mathcal{R})x$

Using the formula for the z-transform of $\mathcal{R}$ we get $(1+8z^{-1}+7z^{-2})Y = (1-z^{-1})X$.

System (or transfer) function

Theorem: The difference equation $P(\mathcal{R})y = Q(\mathcal{R})x$ with initial conditions $x[n] = 0$ and $y[n] = 0$ for $n < 0$ (called *rest initial conditions*) satisfies $Y = \dfrac{Q(z^{-1})}{P(z^{-1})}X$.

We call $\dfrac{Q(z^{-1})}{P(z^{-1})}$ the *system function.* Often, we will denote it $H(z)$.

Thus, with rest IC, $Y(z) = X(z)H(z)$.

(The need for rest IC is explained in the odds and ends section later in these notes.)

Unit sample response

By definition, the unit sample response satisfies the equation

$P(\mathcal{R})h = Q(\mathcal{R})\delta$ with rest IC. It's easy to see that $\mathcal{Z}(h) = H$ is the system function.

Theorem: The equation $P(\mathcal{R})y = Q(\mathcal{R})x$ with rest IC has solution $y = x * h$, where h is the unit sample response of the system.

Proof: From above we know $Y(z) = X(z)H(z) = \mathcal{Z}(x * h)$. QED

(continued)

Example: Solve $y[n] - ay[n-1] = \delta[n]$, with rest IC.

Z-transform: $(1 - az^{-1})Y = 1 \;\Rightarrow\; Y = \dfrac{1}{1 - az^{-1}} = 1 + az^{-1} + (az^{-1})^2 + \ldots$.

$$\Rightarrow\; y[n] = \begin{cases} 0 & \text{for } n < 0 \\ a^n & \text{for } n \geq 0 \end{cases}$$

Poles, stability and homogeneous equations

The system $P(\mathcal{R})y = Q(\mathcal{R})x$ has system function $H(z) = \frac{Q(z^{-1})}{P(z^{-1})}$.

So the poles of $H(z)$ are exactly the roots of $P(z^{-1})$.

(We need to assume P and Q have no common factors.)

As in differential equations these poles give us the solutions to the *corresponding homogeneous equation*, i.e., $P(\mathcal{R})y = 0$.

Example: Solve the homogeneous equation $P(\mathcal{R})y = y[n] + 8y[n-1] + 7y[n-2] = 0$.

Trial solution: $y[n] = a^n$.

Substitution: $a^n + 8a^{n-1} + 7a^{n-2} = 0 \;\Rightarrow\; a^{n-2}(a^2 + 8a + 7) = 0$.

Characteristic equation: $a^2 + 8a + 7 = 0$.

Roots: $a = -7, -1$.

Two solutions: $y_1[n] = (-7)^n$, $y_2[n] = (-1)^n$.

General solution: $y = c_1y_1 + c_2y_2$. (This follows from the linearity of $P(\mathcal{R})$.)

(Below we will discuss the existence and uniqueness theorem that guarantees this gives all possible solutions.)

Note, $\dfrac{1}{P(z^{-1})} = \dfrac{1}{1 + 8z^{-1} + 7z^{-2}} = \dfrac{z^2}{z^2 + 8z + 7}$. So the roots of the characteristic equation are the same as the zeros of the denominator which are the poles of the system function.

Stability

As in 18.03, we say the system $P(\mathcal{R})y = Q(\mathcal{R})x$ is stable if the homogeneous solution $y_h[n] \to 0$ as $n \to \infty$ no matter what the initial conditions.

That is, the initial conditions don't affect the long-term behavior of the system.

Theorem: The system $P(\mathcal{R})y = Q(\mathcal{R})x$ is stable if and only if all the poles of the system function have magnitude < 1. (We assume P and Q have no common factors.)

Proof: As in the example just above, the general homogenous solution is a linear combination of sequences of the form $y_j[n] = a_j^n$, where a_j is a pole of the system function. This goes to 0 if and only if $|a_j| < 1$.

(continued)

Graphically the system is stable if all the poles are inside the unit circle. (Compare this with differential equations where the homogeneous solution is built from functions of the form $y_j(t) = e^{at}$, so we need a in the left half-plane.)

Example:
Difference equations:

Stable discrete system Unstable discrete system

Differential equations:

Stable continuous system Unstable continuous system

Odds and ends

Causality: Causality means the future doesn't affect the past. Our systems $P(\mathcal{R})y = Q(\mathcal{R})x$ are causal because $y[n]$ depends only on $y[k]$ and $x[k]$ for $k \leq n$.

Linearity: $P(\mathcal{R})$ is linear, i.e. $P(\mathcal{R})(c_1y_1 + c_2y_2) = c_1P(\mathcal{R})y_1 + c_2P(\mathcal{R})y_2$.
Proof: Immediate from the definition of $\mathcal{R}$ and $P(\mathcal{R})$.

Using linearity we see that the general solution to $P(\mathcal{R})y = Q(\mathcal{R})x$ is given by $y = y_p + y_h$, where y_p is any particular solution and y_h is the general homogeneous solution.

Example: Find the general solution to $y[n] - \dfrac{1}{2}y[n-1] = u[n]$.

Homogeneous solution: Characteristic equation: $a - \dfrac{1}{2} = 0 \;\Rightarrow\; a = \dfrac{1}{2} \;\Rightarrow\; y_h[n] = c\left(\dfrac{1}{2}\right)^n$.

Particular solution: Use rest IC, so $y[n] = 0$ for $n < 0$.
We could find $y[n]$ for $n \geq 0$ directly, instead we'll find it using the z-transform.

$$(1 - \frac{1}{2}z^{-1})Y = \frac{1}{1 - z^{-1}} \;\Rightarrow\; Y = \frac{1}{(1 - \frac{1}{2}z^{-1})(1 - z^{-1})} = \frac{A}{1 - \frac{1}{2}z^{-1}} + \frac{B}{1 - z^{-1}}.$$

Using coverup we find $A = -1$, $B = 2 \;\Rightarrow\; y_p[n] = \begin{cases} 0 & \text{for } n < 0 \\ 2 - (1/2)^n & \text{for } n \geq 0 \end{cases}$

Note, for $n \geq 0$ we can write the answer in the form $y_p[n] = 1 + \dfrac{1}{2} + \dfrac{1}{4} + \cdots + \left(\dfrac{1}{2}\right)^n$.

Finally, the general solution is $y = y_p + y_h$.

(continued)

Transient: If a system is stable then $y_h[n] \to 0$ for all initial conditions. In this case we call y_h the transient.

Exponential Input Theorem: A solution to $P(\mathcal{R})y = a^n$ is $y[n] = \dfrac{a^n}{P(a^{-1})}$.

Proof: $\mathcal{R}a^n = a^{n-1} = a^{-1}a^n \;\Rightarrow\; P(\mathcal{R})a^n = P(a^{-1})a^n$.

(see below for the extended version of this theorem)

Discretizing DEs: In previous sections we have seen the algebraic analogy between the differential operator D and the right shift operator $\mathcal{R}$. That is, the algebra is similar when dealing with either of these operators. But, for numerical methods we need more than this. To discretize a differential equation we need to replace D by something that is analytically analogous. You have already seen this when we learned Euler's method. Here we will discretize DEs two different ways: using forward differences and using backward differences. We will relate them to numerical methods you have already learned. We will also discuss issues of stability.

Let $y_c(t)$ be a function of the continuous variable t.
We can *discretize* y_c by picking a *stepsize* h and a start time t_0 and letting

$$y[n] = y_c(t_0 + nh).$$

Here are two ways to discretize the differential operator D (using the same stepsize h).

Forward difference: $(\Delta_f\, y)[n] = \dfrac{y[n+1] - y[n]}{h}$, i.e. $\Delta_f = \dfrac{\mathcal{L} - \mathcal{I}}{h}$, where $\mathcal{L}$ is the *left shift operator.*

Backward difference: $(\Delta_b\, y)[n] = \dfrac{y[n] - y[n-1]}{h}$, i.e. $\Delta_b = \dfrac{I - \mathcal{R}}{h}$.

Example: Use forward and backward differences to discretize the first order constant coefficient IVP.

(C) $$y_c' + ky_y = f_c(t), \quad y_c(0) = 1.$$

Pick a stepsize $h > 0$ and discretize $f_c(t)$ by $f[n] = f_c(nh)$.
Using forward differences replace D by Δ_f. Then $y_c(nh) \approx y[n]$, where

$$(\Delta_f\, y)[n] + k\, y[n] = f[n], \quad y[0] = 1.$$

$$\Rightarrow\; \frac{y[n+1] - y[n]}{h} + k\, y[n] = f[n] \;\Leftrightarrow$$

(F) $$y[n+1] - (1 - hk)\, y[n] = hf[n].$$

Using backward differences replace D by Δ_b. Then $y_c(nh) \approx y[n]$, where

$$(\Delta_b\, y)[n] + k\, y[n] = f[n], \quad y[0] = 1.$$

(continued)

$$\Rightarrow \frac{y[n]-y[n-1]}{h} + k\,y[n] = f[n] \;\Leftrightarrow\; (1+hk)\,y[n] - y[n-1] = hf[n] \;\Leftrightarrow$$

$$(B) \qquad y[n] - \frac{1}{1+hk}\,y[n-1] = \frac{h}{1+hk}\,f[n].$$

Example: (continued) Examine the stability of the original continuous system (C) and its discretized versions (F), (B).

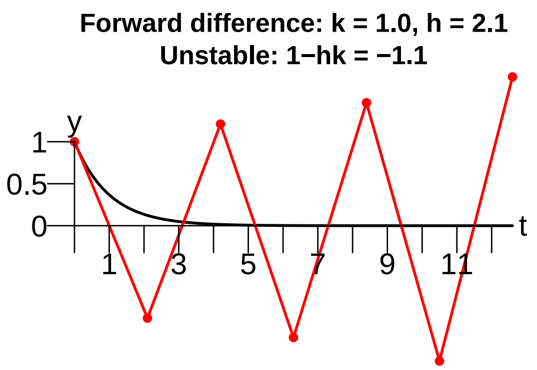
The continuous system (C) has system function $\dfrac{1}{s+k}$. which has a pole at $s = -k$.
The homogeneous equation (the DE with $f_c(t) = 0$) has solution $y_c = Ce^{-kt}$.
Looking at either the poles or the homogeneous solution we conclude the system is stable for $k > 0$.

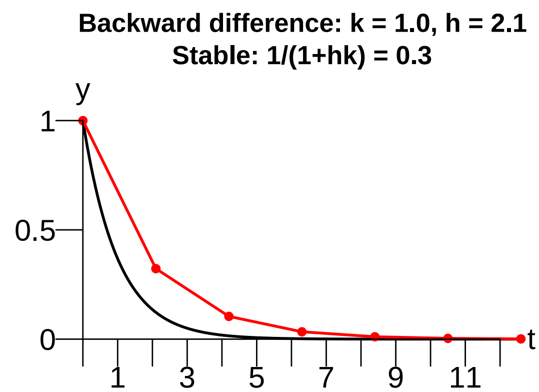
The discrete system (F) has system function $\dfrac{1}{1-(1-hk)z^{-1}}$.
The system has a pole at $z = 1 - hk \Rightarrow$ it's stable if $|1 - hk| < 1$.
If $k > 0$ then we need $0 < h < 2/k$ for stability. Thus, even when (C) is stable, (F) is only stable for small h.

The discrete system (B) has system function $\dfrac{1}{1-z^{-1}/(1+hk)}$.
The system has a pole at $z = 1/(1+hk) \Rightarrow$ it's stable if $1/|1+hk| < 1$.
If $k > 0$ then (B) is stable for all $h > 0$. Thus, when (C) is stable so is (B) for any h.

We can see the above results graphically by discretizing the homogeneous equations using various values of h. In all of the plots below, the black exponential curve is the continuous solution $y_c(t) = e^{-kt}$. Notice how, when discretizing a stable DE, forward differences change quality as h decreases, while the backward differences do not.

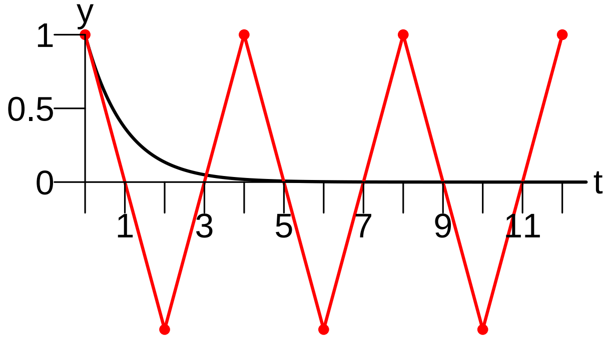
$$
\begin{array}{llll}
(CH) & y_c' + ky = 0; & y_c(0) = 1 & \Rightarrow\; y_c(t) = e^{-kt}. \\
(FH) & y[n+1] - (1-hk)\,y[n] = 0; & y[0] = 1 & \Rightarrow\; y[n] = (1-hk)^n \\
(BH) & y[n] - \dfrac{1}{1+hk}\,y[n-1] = 0; & y[0] = 1 & \Rightarrow\; y[n] = (1+hk)^{-n}
\end{array}
$$



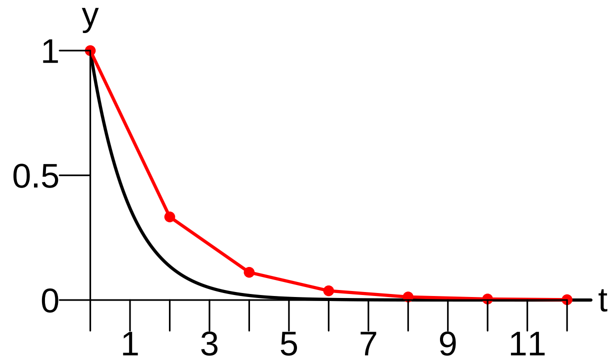


(continued)

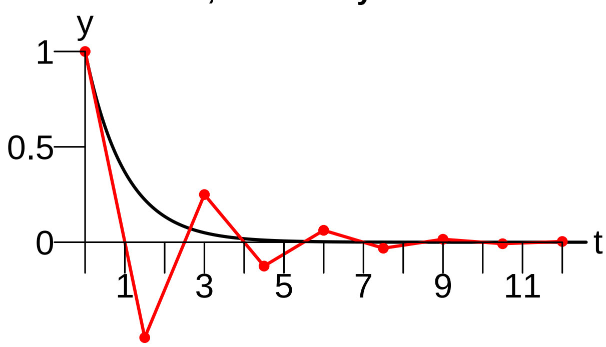
Forward difference: k = 1.0, h = 2.0
Unstable: 1−hk = −1.0



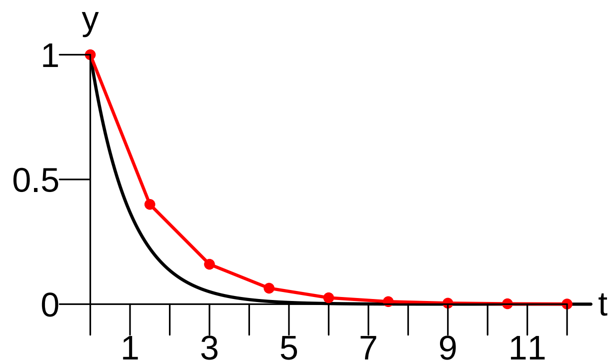
Backward difference: k = 1.0, h = 2.0
Stable: 1/(1+hk) = 0.3



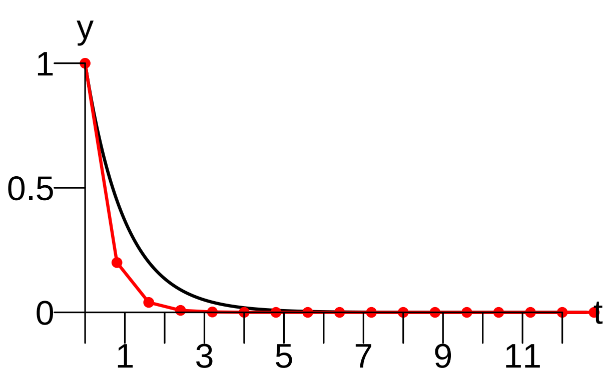
Forward difference: k = 1.0, h = 1.5
Stable, oscillatory: 1−hk = −0.5



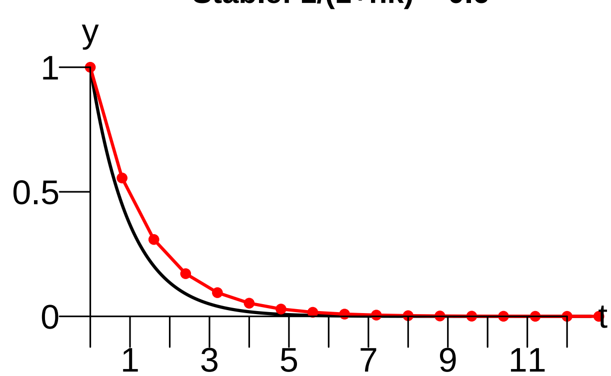
Backward difference: k = 1.0, h = 1.5
Stable: 1/(1+hk) = 0.4



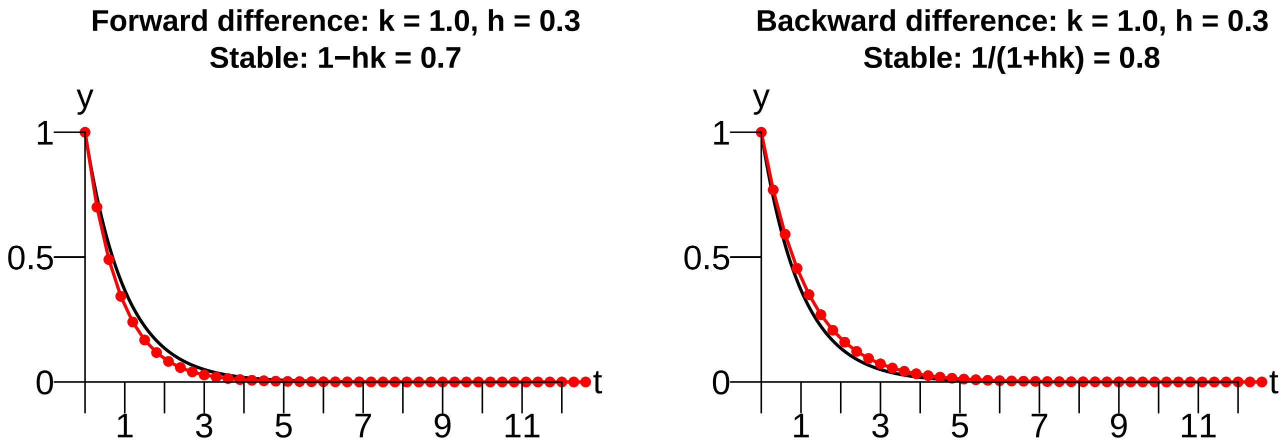
Forward difference: k = 1.0, h = 0.8
Stable: 1−hk = 0.2



Backward difference: k = 1.0, h = 0.8
Stable: 1/(1+hk) = 0.6



(continued)



Relation to numerical methods for solving differential equations

In the examples above we can write $y_c' = F(y_c, t) = -ky_c + f(t)$. The discretizations are specific cases of the following methods.

Forward differences $\leftrightarrow$ *Euler's method:*

$$t[n] = nh \quad \text{and} \quad y[n+1] = y[n] + h\,F(y[n], t[n]).$$

Backward differences $\leftrightarrow$ *Backward Euler:*

$$t[n] = nh \quad \text{and} \quad y[n+1] = y[n] + h\,F(y[n+1], t[n+1]).$$

The Backward Euler is also called *implicit Euler* because it only gives an implicit formula for $y[n+1]$.

In general, for a constant coefficient linear DE Euler's method corresponds to discretizing D using Δ_f and implicit Euler corresponds to using Δ_b. Backward Euler maintains stability for any stepsize, while (forward) Euler only maintains stabiliy for small stepsize. This is why, despite the extra computation the implicit techniques are often preferable.

Likewise, there are forward and backward Runge-Kutta methods and the discussion of stability holds for them also.

Analogies with differential equations

1. Algebraic analogy: Most often in 18.03 our system was $P(D)y = f$, but sometimes we had $P(D)y = Q(D)f$. For example, for LRC circuits we have differential equations in q (the charge on the capacitor) and i (the current in the circuit). The input voltage E is handled slightly differently in each equation.

i. $Lq'' + Rq' + \frac{1}{C}q = E$. Input $= E$ is used as is.

ii. $Li'' + Ri' + \frac{1}{C}i = E'$. Input $= E$ is modified.

(continued)

In (i) we have $P(D)q = E$. In (ii) we have the more general $P(D)i = Q(D)E$, which has system function $\dfrac{Q(s)}{P(s)}$. This is algebraically analogous to the difference equation $P(\mathcal{R})y = Q(\mathcal{R})x$ with system function $\dfrac{Q(z^{-1})}{P(z^{-1})}$.

2. Analytic analogy: This is covered in the section 'Discretizing DEs' above. There we saw two ways to discretize D, $D \to \Delta_b$ and $D \to \Delta_f$.

Theorem: (Existence and uniqueness) If $P(\mathcal{R})$ has degree m then the IVP

$$P(\mathcal{R})y = 0; \quad y[0] = b_0,\, y[1] = b_1,\, \ldots,\, y[m-1] = b_{n-1}$$

has a unique solution.

Proof: This is clear. Simply solve for $y[n]$ recursively as we did in the first order example. We show a degree two example.

Example: Solve $y[n] + a_1 y[n-1] + a_2 y[n-2] = 0$, $y[0] = b_1$, $y[1] = b_1$.
General equation: $y[n] = -a_1 y[n-1] - a_2 y[n-2]$.
$y[0] = b_0$, $y[1] = b_1 \;\Rightarrow\; y[2] = -a_1 b_0 - a_2 b_2 \;\Rightarrow\; y[3] = -a_1 y[2] - a_2 y[1]$, etc.
We see that $y[n]$ is uniquely determined.

Convolution formula as a result of linear time invariance

Consider the equation $P(\mathcal{R})y[n] = Q(\mathcal{R})x[n]$ with rest IC. Let h be the unit sample response. We will rederive the formula $y = x * h$ using linearity and time invariance.
Let $y[n] = (x*h)[n] = \sum_k x[k]h[n-k]$.
The sequence $h[n]$ is the solution to the equation $P(\mathcal{R})y[n] = Q(\mathcal{R})\delta[n]$.
Time invariance means that $h[n-k]$ is a solution to $P(\mathcal{R})y[n] = Q(\mathcal{R})\delta[n-k]$.
We can write $x[n] = \sum_k x[k]\delta[n-k]$, so by linearity we have

$$\begin{aligned} P(\mathcal{R})y[n] &= P(\mathcal{R})\sum_k x[k]h[n-k] = \sum_k x[k]P(\mathcal{R})h[n-k] = \sum_k x[k]Q(\mathcal{R})\delta[n-k] \\ &= Q(\mathcal{R})\sum_k x[k]\delta[n-k] = Q(\mathcal{R})x[n] \end{aligned}$$

We have shown that $y = x * h$ is a solution.

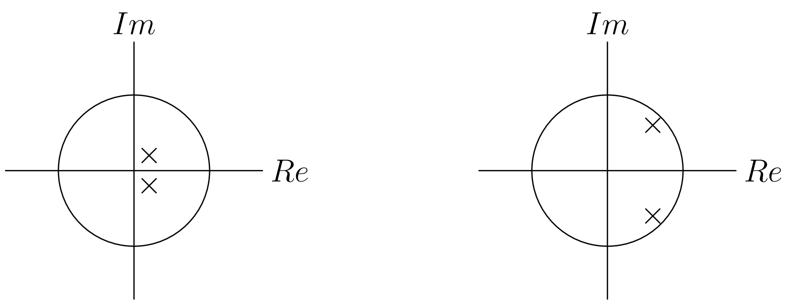
(continued)

Growth and decay rates

If a is a complex number then if $|a| < 1$ the rate that a^n *decays* to 0 depends on $|a|$, the closer to 1 the slower a^n decays. Likewise if $|a| > 1$ the rate that a^n *grows* depends on $|a|$.

If a_1, a_2, $\ldots$, a_m are complex numbers then the growth or decay rate of the linear combination $y[n] = \sum c_j a_j^n$ is given by the biggest value of $|a_j|$. If all $|a_j| < 1$ then it is a decay rate and the bigger the rate (the closer to 1) the slower $y[n]$ decays.

Example: Both systems are stable. System A has a faster decay rate than system B, i.e., the transient disappears faster for system A than for system B.



System A System B

The need for rest IC: If $y[n]$ is the solution to $P(\mathcal{R})y = Q(\mathcal{R})\delta$ then we needed rest IC to write $Y(z) = \dfrac{Q(z^{-1})}{P(z^{-1})}$. We'll explain this using a simple example.

Consider $P(\mathcal{R})y[n] = y[n] - y[n-1] = \delta[n]$.

Particular solution with rest IC: $y_p[n] = \begin{cases} 0 & \text{for } n < 0 \\ 1 & \text{for } n \geq 0 \end{cases}$.

Homogeneous solution: $y_h[n] = c$.

General solution: $y[n] = y_p[n] + y_h[n] = \begin{cases} c & \text{for } n < 0 \\ 1 + c & \text{for } n \geq 0 \end{cases}$.

Since $P(\mathcal{R})(y_p + y_h) = \delta$ the z-transform gives $P(z^{-1})(Y_p + Y_h) = 1$. The reason we can't simply divide by $P(z^{-1})$ is because $P(z^{-1})Y_h(z) = (1 - z^{-1}) \displaystyle\sum_{n=-\infty}^{\infty} cz^{-n} = 0$.

Algebraically we say that $P(z^{-1})$ and $Y_h(z)$ are *zero divisors*, that is, they are non-zero but when multiplied together they give 0. Just like dividing by 0, we have to be careful doing division with zero divisors. By demanding rest IC we only consider z-transforms of sequences y with $y[n] = 0$ for $n < 0$. It is easy to see that this set (called a ring) does not have zero divisors.

(continued)

Extended Exponential Input Theorem

Shift Law: Let $v[n]$ be a sequence and $w[n] = a^n v[n]$ then $(1-a\mathcal{R})w = a^n(1-\mathcal{R})v$. (Proof is easy algebra).

Example: Solve $(1 - a\mathcal{R})x = a^n$.

The exponential input theorem fails because $P(a^{-1}) = 1 - aa^{-1} = 0$.

Try $x = a^n v$. Substituting x into the equation and using the shift law gives

$(1 - a\mathcal{R})x = a^n \;\Leftrightarrow\; (1 - aR)(a^n v) = a^n(1 - \mathcal{R})v = a^n \;\Leftrightarrow\; (1 - \mathcal{R})v = 1$

$\Rightarrow\; v[n] = n \;\Rightarrow\; x[n] = na^n$.

(If we want we can add the homogeneous solution $x_h[n] = Ca^n$ to x.)

Example: Solve $(1 - a\mathcal{R})^2 x = a^n$.

Again, try the solution $x = a^n v$. Substituting gives

$(1 - a\mathcal{R})^2 x = a^n \;\Leftrightarrow\; a^n(1 - \mathcal{R})^2 v = a^n \;\Leftrightarrow\; (1 - \mathcal{R})^2 v = 1 \;\Rightarrow\; (1 - \mathcal{R})v = n$

$$\Rightarrow\; v[n] = \frac{n(n+1)}{2} \;\Rightarrow\; x[n] = \frac{n(n+1)}{2}a^n.$$

Define $S_k[n]$ as $S_0[n] = 1$, and $S_k[n] = \dfrac{n(n+1)\cdots(n+k-1)}{k!}$ for $k \geq 1$.

Note: $S_k[0] = 0$ and $S_k[1] = 1$.

Lemma: $(1 - \mathcal{R})S_k = S_{k-1}$.

Proof: $(1 - \mathcal{R})(S_k)[n] = S_k[n] - S_k[n-1] = \dfrac{n(n+1)\cdots(n+k-1) - (n-1)n\cdots(n+k-2)}{k!}$

$$= \frac{[(n+k-1) - (n-1)]n(n+1)\cdots(n+k-2)}{k!} = \frac{n(n+1)\cdots(n+k-2)}{(k-1)!} = S_{k-1}[n].$$

Theorem (extended exponential input)

Suppose $Q(s)$ is a polynomial and $Q(a^{-1}) \neq 0$ then the difference equation

$Q(\mathcal{R})(1 - a\mathcal{R})^k x = a^n$ has solution $x[n] = \dfrac{a^n}{Q(a^{-1})}S_k[n]$.

Proof: By the shift law and the lemma $(1 - a\mathcal{R})^k x = \dfrac{a^n}{Q(a^{-1})}(1 - a\mathcal{R})^k S_k = \dfrac{a^n}{Q(a^{-1})}S_0 = \dfrac{a^n}{Q(a^{-1})}$.

Therefore $Q(\mathcal{R})(1 - a\mathcal{R})^k x = Q(\mathcal{R})\dfrac{a^n}{Q(a^{-1})} = a^n$.

One-sided sequences

We state without proof some 'one-sided' results.

1. $(1 - \mathcal{R})x = u[n]$ has solution $x[n] = (n+1)u[n]$.
2. $(1 - \mathcal{R})^k x = u[n]$ has solution $x[n] = S_k[n+1]u[n]$.
3. $(1 - a\mathcal{R})^k x = a^n u[n]$ has solution $x[n] = a^n S_k[n+1]u[n]$.

(continued)

Summary

(correspondence between difference and differential equations)

Difference equations	*Differential equations*
Sequences: $x[n]$	Functions: $x(t)$
z-transform: $\mathcal{Z}(x) = X(z) = \sum_n x[n]z^{-n}$	Laplace transform: $\mathcal{L}(x) = X(s) = \int_0^\infty x(t)\mathrm{e}^{-st}\,dt$
Convolution: $(x * y)[n] = \sum_k x[k]y[n-k]$	$(x * y)(t) = \int_0^t x(u)y(t-u)\,du$
$\mathcal{Z}(x * y)(z) = X(z)Y(z)$	$\mathcal{L}(x * y) = X(s)Y(s)$

Operators
$$\left.\begin{array}{rcl} \mathcal{R} &=& \text{delay} = \text{right shift} \\ \Delta_b &=& \text{backward difference} \\ \Delta_f &=& \text{forward difference} \end{array}\right\} \qquad D = \text{derivative}$$

$$\left.\begin{array}{rcl} (\mathcal{R}x)[n] &=& x[n-1] \\ (\Delta_b x)[n] &=& \dfrac{x[n]-x[n-1]}{h} \\ (\Delta_f x)[n] &=& \dfrac{x[n+1]-x[n]}{h} \end{array}\right\} \qquad Dx = x'$$

(The correspondence $\mathcal{R} \leftrightarrow D$ is an algebraic one and Δ_f, $\Delta_b \leftrightarrow D$ is an analytic one.)

Causal LTI system: $P(\mathcal{R})y = Q(\mathcal{R})x$ — $P(D)y = Q(D)x$ (18.03 usually: $P(D)y = x$)

(x = input, y = response, assume P and Q have no common factors.)

Difference equations	*Differential equations*
System function: $H(z) = \dfrac{Q(z^{-1})}{P(z^{-1})}$	$H(s) = \dfrac{Q(s)}{P(s)}$
Unit sample resp.: $P(\mathcal{R})h = Q(\mathcal{R})\delta$, rest IC	Unit impulse resp.: $P(D)h = Q(D)\delta$, rest IC
$\mathcal{Z}(h) = H$	$\mathcal{L}(h) = H$
$P(\mathcal{R})y = Q(\mathcal{R})x$; rest IC $\Rightarrow Y = XH,\ y = x * h$	$P(D)y = Q(D)x$; rest IC $\Rightarrow Y = XH,\ y = x * h$
Stability: poles of H inside unit circle	poles of H in left half-plane
Decay rate of transient determined by pole with greatest magnitude	decay rate determined by right most pole (greatest real part)

(In 6.01 they also write $\dfrac{Q(\mathcal{R})}{P(\mathcal{R})}$ for the system function.)