THE CORE IDEAS IN OUR TEACHING

GILBERT STRANG

What will our students remember? One answer comes quickly but it is a counsel of despair: *nothing at all.* At the other extreme is an impossible hope that we all cherish: *everything we say.* Let me look for an intermediate answer, closer to reality, possibly by changing the question.

I have come to believe that each course has a central core. We may not see it ourselves, when we teach a new topic every day. For the calculus course, I won't even venture an answer: at least not here. My examples will be differential equations and linear algebra, because writing a textbook forced me to uncover (painfully slowly!) the underlying structure of the course.

May I begin with linear algebra. The ideas of a vector space and a basis for that space are central. We have a serious job to help students understand these words. The building blocks are "linear combinations" and "linear independence". We certainly need good examples, and good bases for them. I think it is here that the course becomes coherent—or it can scatter into unconnected examples of isolated ideas.

I will start with a matrix A. A more abstract person would start from a linear transformation. But we are aiming for a basis; we are choosing coordinates; they bring us to a matrix. There are four fundamental subspaces associated with that matrix:

1.	Its nullspace $\mathbf{N}(A)$ (the kernel)	dimension	$n-r$
2.	Its column space $\mathbf{C}(A)$ (the range)		r
3.	Its row space, which is $\mathbf{C}(A^T)$		r
4.	The nullspace $\mathbf{N}(A^T)$ of the transpose		$m-r$

These are the spaces that we want students to remember. I draw them as often as possible (two in $\mathbf{R}^n$ and two in $\mathbf{R}^m$). I count their basis vectors to find their dimension: the first big theorems in linear algebra. The rank r determines all dimensions. I propose multiple choices of A—the beauty of this subject is in the wonderful variety of matrices. And I connect the four subspaces to factorizations of A, which are really choices of bases that lie at the absolute center of pure and applied linear algebra.

The bases in U and Q and S and V become increasingly perfect.

$\mathbf{A = LU}$	Elimination gives an echelon basis for the row space
$\mathbf{A = QR}$	Gram-Schmidt gives an orthonormal basis for $\mathbf{C}(A)$
$\mathbf{A = S\Lambda S^{-1}}$	Eigenvectors give a basis in which A is diagonal
$\mathbf{A = U\Sigma V^T}$	Orthonormal bases in the columns of U and V.

We are constantly constructing bases for the fundamental subspaces. Elimination and Gram-Schmidt orthogonalization end after finitely many steps. Diagonalization by eigenvectors is deeper and better, but A must be square and nondefective. The Singular Value Decomposition produces perfect bases v_i and u_i for all four subspaces – orthonormal and also diagonalizing for every matrix A:

$$Av_i = \sigma_i u_i \ (i \le r) \qquad Av_i = 0 \text{ and } A^T u_i = 0 \ \ (i > r)$$

The success of the SVD comes from the spectral theorem for symmetric matrices: $A^T A$ has a full set of orthonormal eigenvectors v_i. Beautifully, the u_i turn out to be orthonormal eigenvectors of AA^T. This can be a highlight for the last day of a linear algebra course.

For an earlier day, one idea is to ask students to "read" a few matrices:

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

The rotation is familiar, the projection is almost too easy. The difference matrix is also the incidence matrix for a simple graph (three nodes in a line). Incidence matrices of a larger graph are terrific examples – all four subspaces have a meaning.

May I turn from subspaces to the basic course on differential equations. Part of this course is a collection of methods to solve separable equations, exact equations, logistic equations $y' = ay - by^2$, and more. We go forward to systems of equations, and test nonlinear equations for stability. But the coherent part (the central problem) is to solve **linear equations with constant coefficients**. How can we present their solutions?

I believe we have to answer this question. It is the ODE equivalent of solving $Ax = 0$ and $Ax = b$ and $Ax = \lambda x$. It certainly rests on the most important functions in this course: *exponentials* e^{st} *and* $e^{\lambda t}$. By working with exponentials, we (almost) turn the differential equation into algebra.

Start with the simplest right hand sides $f(t) = 0$ and e^{st}.

$$Ay'' + By' + Cy = 0 \qquad Ay'' + By' + Cy = e^{st}$$

The key idea is to expect solutions $y = Ge^{st}$:

$$G(As^2 + Bs + C)e^{st} = 0 \qquad G(As^2 + Bs + C)e^{st} = e^{st}.$$

On the left, two values of s are allowed: the roots s_1 and s_2 of $As^2 + Bs + C = 0$. On the right, any s is allowed (and the possibilities $s_1 = s_2$ and $s = s_1$ and $s = s_1 = s_2$ need special attention). Normally we have

$$y_n = y_{\text{nullspace}} = c_1 e^{s_1 t} + c_2 e^{s_2 t}$$

$$y_p = y_{\text{particular}} = G(s)e^{st} = \frac{1}{As^2 + Bs + C}\, e^{st}.$$

Those two parts of $y(t)$ connect linear differential equations to linear algebra. The complete solution combines all y_n with one y_p. Linearity is in control and the consequence is $y = y_n + y_p$.

I apologize for asking you to read what you know so well. The simplicity of $y = Ge^{st}$ has to be recognized and remembered. This is where calculus meets algebra. G is the prime example of an undetermined coefficient (determined by the equation). An elementary course could continue as far as $f(t) = e^{i\omega t}$ and $\cos \omega t$ and $\sin \omega t$ and stop. The serious question is to *solve the differential equation for all* $f(t)$.

I see two instructive ways to reach $y(t)$. Both begin with special right hand sides, and combine the solutions. The combination has to be an integral and not just a finite sum: calculus is needed now. Here are the good options:

1. Combine exponentials e^{st} with weights $F(s)$ to get $f(t)$. By linearity, the solution $y(t)$ will combine the exponentials $F(s)G(s)e^{st}$.

2. Combine impulses $\delta(t - s)$ with weights $f(s)$ to get $f(t)$. By linearity, the solution $y(t)$ will combine the impulse responses $f(s)g(t - s)$.

Where e^{st} is localized at frequency s, the delta function $\delta(t - s)$ is completely localized at time s.

Method **1** uses the Laplace transform. The transform of $f(t)$ gives the right weights $F(s)$:

$$F(s) = \text{ transform of } f(t) \quad y(t) = \text{ inverse transform of } F(s)G(s)e^{st}.$$

For each s, that solution $F(s)\,G(s)\,e^{st}$ is easy. The hard part is the *inverse* Laplace transform, to combine those solutions into $y(t)$.

Realistically, we know a very limited number of transform pairs. Method **1** almost limits us to the same short list as before: f can combine $e^{(a+i\omega)t}, \cos\omega t, \sin\omega t, t$, and their products. This is a space of functions whose derivatives stay in the space. You can guess that I am advocating Method **2**. It begins with an impulse $\delta(t)$:

(1) $$Ag'' + Bg' + Cg = \delta(t) \quad \text{with } g(0) = 0 \text{ and } g'(0) = 0.$$

Introducing that delta function is a good thing! We are finding the *fundamental solution* $g(t)$—the Green's function, the growth factor, the impulse response. This is a high point in the course. And it is easy to do, because *this same* $g(t)$ *also solves the homogeneous equation*:

(2) $$Ag'' + Bg' + Cg = 0 \quad \text{with } g(0) = 0 \text{ and } g'(0) = 1/A.$$

The solution must have the form $g(t) = c_1 e^{s_1 t} + c_2 e^{s_2 t}$. The two initial conditions give c_1 and c_2 and a neat formula for $g(t)$:

(3) $$\mathbf{g(t)} = \frac{\mathbf{e^{s_1 t} - e^{s_2 t}}}{\mathbf{A(s_1 - s_2)}} \qquad \left(\text{or } g(t) = \frac{te^{s_1 t}}{A} \text{ when } s_1 = s_2\right).$$

Then the original equation, with any right side $f(t)$, is solved by

(4) $$y_{\text{particular}}(t) = \int_0^t g(t-s)\, f(s)\, ds.$$

Discussion In coming quickly to the formula for $y(t)$, I have left multiple loose ends. Let me go backwards more slowly, as we would certainly do in a classroom. Methods **1** and **2** are closely connected. The Laplace transform of $\delta(t)$ is 1. Then equation (1) transforms to

(5) $$(As^2 + Bs + C)\; G(s) = 1.$$

The transfer function $G(s) = 1/(As^2 + Bs + C)$ *is the Laplace transform of the impulse response* $g(t)$. These functions can be written in terms of A, B, C or s_1 or s_2. A lot of effort has gone into choosing good parameters! The damping ratio $B/\sqrt{4AC}$ and the natural frequency $\sqrt{C/A}$ are two of the best.

We must also explain why equations (1) and (2) have the same solution $g(t)$. Mechanically, this comes from partial fractions:

$$\frac{1}{As^2+Bs+C} = \frac{1}{A(s-s_1)(s-s_2)} = \frac{1}{A(s_1-s_2)}\left(\frac{1}{s-s_1} - \frac{1}{s-s_2}\right).$$

The inverse Laplace transform confirms that $e^{s_1 t}$ and $e^{s_2 t}$ go into $g(t)$.

Here is a truly "mechanical" explanation of (1) = (2). A bat hits a ball at $t = 0$. The velocity jumps instantly to $g'(0) = 1/A$. This comes from integrating $Ag'' + Bg' + Cg = \delta(t)$ from $t = 0$ to $t = h$. The left side produces the jump in Ag' and the integral of $\delta(t)$ is 1. The other terms disappear as $h \to 0$, leaving $Ag'(0) = 1$.

In working with $\delta(t)$, some faith is needed. It is worth developing and it is not misplaced. A delta function is an extremely useful model. So is its integral the step function, which turns on a switch at $t = 0$. By linearity, the step response is the integral of $g(t)$.

Finally, let me connect Method **1** directly to Method **2**. In the first method, the Laplace transform of $y(t)$ is $F(s)\,G(s)$. In the second method, $y(t)$ is the *convolution* of $f(t)$ with $g(t)$. The connection is the Convolution Rule: The transform of a convolution $f(t) * g(t)$ is a multiplication $F(s)\,G(s)$.

In the language of signal processing, any constant coefficient linear equation can be solved in the "s-domain" or the "t-domain." The poles s_1, s_2 of the transfer function $G(s) = 1/(As^2 + Bs + C)$ control the behavior of $y(t)$: oscillation, decay, or instability. The whole course develops out of the quadratic formula for those roots s_1 and s_2.

Note The actual course would start with **first order equations:**

$$y' - ay = 0 \qquad y' - ay = e^{st}$$

The null solutions are $y_n = ce^{at}$. The particular solution is $y_p = e^{st}/(s-a)$. The transfer function is $G(s) = 1/(s-a)$. The fundamental solution (impulse response, growth factor, Green's function) solves

$$\begin{array}{lll} g' - ag = \delta(t) & \text{with} & g(0) = 0 \\ g' - ag = 0 & \text{with} & g(0) = 1 \end{array}$$

This function is simply $\mathbf{g = e^{at}}$. At this early point it doesn't need all those names! We recognize it as 1/(integrating factor). Its Laplace transform is $G(s) = 1/(s-a)$. For systems $y' = Ay$ we have the matrix exponential $\mathbf{g} = \mathbf{e}^{\mathbf{A}\mathbf{t}}$. The solution $y_n + y_p$ for any right hand side $f(t)$ and initial condition $y(0)$ is

$$(6) \qquad y(t) = y(0)e^{at} + \int_0^t e^{a(t-s)}\, f(s)\, ds.$$

The input $f(s)$ at time s grows in the remaining time $t-s$ by the factor $e^{a(t-s)}$. The solution $y(t)$ (the integral) combines all of these outputs $e^{a(t-s)}\, f(s)$.

That single paragraph translates into weeks of teaching, even without $\delta(t)$. Perhaps first order equations with constant coefficients might be the one topic that is understood and remembered ? I don't like to think so, because a teacher has to remain an optimist.

I plan to prepare video lectures going at a normal pace, and linked to `http://math.mit.edu/dela`. That website has much more about differential equations and linear algebra and a new textbook for those courses.