

Substitutions in Differential Equations and Systems

Class Notes for Math 44 / WRS

March 27, 2007

Substitution in differential equations is much like substitution in ordinary integrals. Sometimes a well-chosen substitution allows us actually to solve an equation. Sometimes it allows us to simplify an equation before we resort to numerical or qualitative techniques. There are no general rules for finding good substitutions, but some examples may be helpful.

1 First-order equations

We begin with some examples involving first-order equations of the form $y' = f(y, t)$.

Example 1 Solve the equation $y' = -\frac{1}{10}y + (\cos t)y^2$.

Write $z = 1/y$. (That is, define $z(t) = 1/y(t)$ for all t .) Then $y = 1/z$ and $y' = -z'/z^2$. Substituting these expressions for y and y' in the original equation gives

$$-\frac{z'}{z^2} = -\frac{1}{10z} + (\cos t)\frac{1}{z^2}$$

which, when multiplied by $-z^2$, gives

$$z' = -\frac{1}{10}z + \cos t.$$

This is now a linear equation for z , and its general solution is

$$z(t) = \frac{1}{2}(\cos t + \sin t) + ce^{-t/10}. \quad (1)$$

Recalling that $y = 1/z$, we get a candidate for a general solution for y :

$$y(t) = \frac{1}{\frac{1}{2}(\cos t + \sin t) + ce^{-t/10}}. \quad (2)$$

□

Trouble with the example. That example went well, but there is a sticking point. Equation (1) is indeed a general solution for z , in a very satisfying way. For each $c \in \mathbb{R}$,

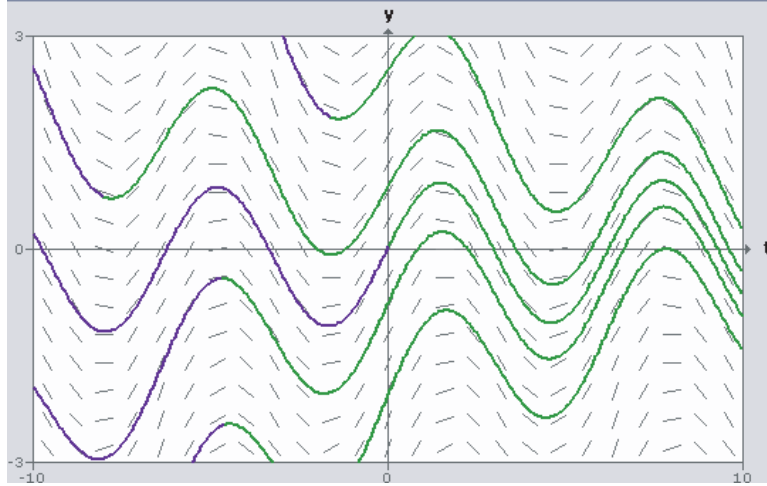


Figure 1: Every solution for z is valid for all t .

equation (1) gives a solution for z that is valid for all $t \in \mathbb{R}$, and it gives all the solutions for z .

But the original equation was about y , not z ; and as a general solution for y , equation (2) leaves something to be desired. For one thing, none of the solutions given by equation (2) is valid for all t . Instead, for each c we get only a patchwork of solutions, each valid only in an interval between points where the denominator of (2) is zero. For another thing, there is actually a solution for y that is valid for all t , namely the constant solution $y(t) = 0$. But this solution isn't given by equation (2) and doesn't correspond to any solution for z .

The trouble is that the relationship $z = 1/y$ is only well-behaved when z and y are both nonzero. When either is zero the substitution doesn't work.

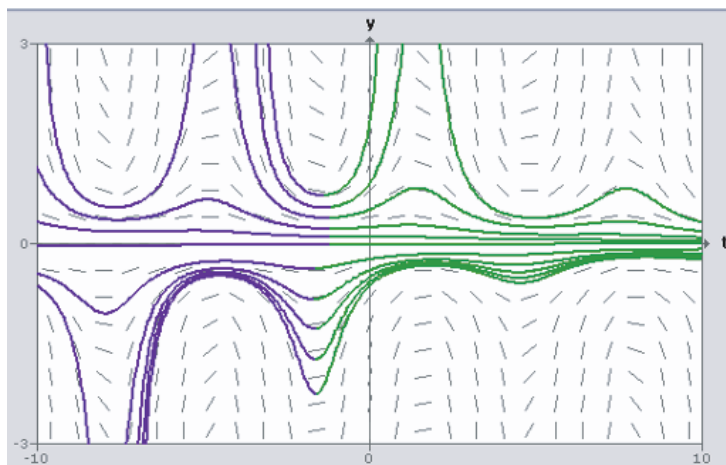


Figure 2: Solutions to the original equation in y have vertical asymptotes.

It isn't hard to cope with these difficulties if we take time to analyze the original equation qualitatively, say by plotting the slope field and a few numerical solutions. The equation is of the form $y' = f(y, t)$ where $f(y, t) = y + (\cos t)y^2$ is continuous for all y and t . Its partial derivative $\frac{\partial f}{\partial y} = 1 + 2y \cos t$ is also continuous for all y and t , so there is a unique solution through every point of the $t - y$ plane. It's just that the nonzero solutions all have vertical asymptotes, and so are valid only on particular intervals. The substitution does most of the work of solving the system, but a common-sense examination is still necessary.

Some theory. In general, given an equation of the form $y' = f(y, t)$, we can define a substitution by a relationship of the form $z = h(y)$. Usually the first step is to invert the relationship to obtain $y = g(z)$, and then differentiate to obtain $y' = g'(z)z'$. We then substitute these expressions into the original equation to obtain a differential equation in z . Sometimes we can solve the equation for z , but even if we cannot, we may still regard the substitution as successful if the equation for z is simpler than the original equation.

Often we start with $y = g(z)$ and the inverse expression $z = h(y)$ is never needed.

The substitution works best when the functions g and h provide a one-to-one correspondence between y values and z values, and when both g and h have continuous derivatives everywhere. When these conditions hold, solutions for y always correspond to solutions for z on the same interval, and vice versa. When the conditions don't hold the substitution may still be helpful, but care must be taken.

Example 2 Solve the Bernoulli equation, $y' = a(t)y + b(t)y^n$.

The first example was an example of a Bernoulli equation with $n = 1$. This is the general form. The correct substitution is $z = y^{1-n}$, which converts the equation to a linear equation. Depending on the value of n , the results may be valid only for positive y and z . \square

Example 3 Solve the equation $y' = a(t)e^y + b(t)$.

A good substitution is $z = e^{-y}$, so $y = -\ln z$ and $y' = -z'/z$. The equation becomes $-z'/z = a(t)(1/z) + b(t)$ or

$$z' = -b(t)z - a(t), \tag{3}$$

which is linear. We are concerned only with solutions to (3) with positive z , because these are the only values that correspond to values of y . \square

Example 4 Solve the equation $y' = -e^y - 1$.

This is a special case of the previous example, with $a(t) = b(t) = -1$. Substituting $z = e^{-y}$, we get this version of (3):

$$z' = z + 1 \tag{4}$$

for which the solutions are $z(t) = ce^t - 1$. Each value of $c \in \mathbb{R}$ gives a solution for $z(t)$ that is valid for all t . But only positive values of c give solutions for y , since when $c \leq 0$, z is never positive. For a particular value $c > 0$ the corresponding solution for y is

$$y(t) = -\ln(ce^t - 1)$$

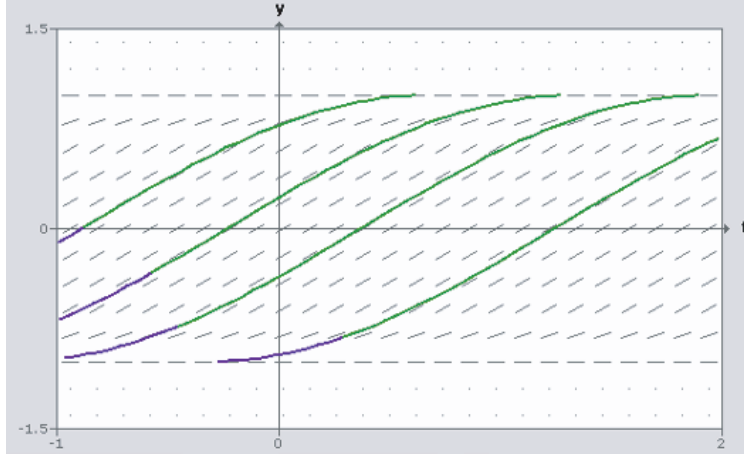


Figure 3: Example 5: Solutions for z .

which is valid only for $t > \ln(1/c)$. These are the only solutions for y . □

Example 5 Solve the equation $y' = y + \sqrt{e^{2t} - y^2}$.

The last term looks bad, but the rest of the equation seems harmless. The right side is meaningful only when $|y| < e^t$. The equation $y' = y$ has the solution $y = e^t$, and linear equations of the form $y' = y + b(t)$ have solutions of the form $y = e^t$ times some integral, so perhaps this equation has a solution of the form $y = e^t z$. If it does, then $y' = e^t z + e^t z'$ and we can substitute to get the equation $e^t z + e^t z' = e^t z + \sqrt{e^{2t} - (e^t z)^2}$ or, after simplifying,

$$z' = \sqrt{1 - z^2}.$$

Can you guess the general solution to this one? If not, then the expression $\sqrt{1 - z^2}$ might suggest that sines and cosines are involved. So we might make a further substitution $z = \sin u$ for some function u . Then $z' = (\cos u)u'$ and $\sqrt{1 - z^2} = \cos u$, so the equation becomes

$$u' = 1$$

which has the general solution $u(t) = t - c$. Reversing the substitutions gives $z = \sin(t - c)$ and

$$y = e^t \sin(t - c).$$

Actually, only the upward-sloping parts of these functions work, because in the original equation we interpret the square root to be non-negative. We have, however, found all of the solutions to the original equation.

This example was contrived to work well, but it does illustrate some ways in which good substitutions might be found. □

A bit more theory. Note that in the last example, the substitution was of the form $y = g(z, t)$ rather than just $y = g(z)$. In this case, $y' = g_1(z, t)z' + g_2(z, t)$, where g_1 and g_2 are the partial derivatives of g (also written $\partial g/\partial z$ and $\partial g/\partial t$).

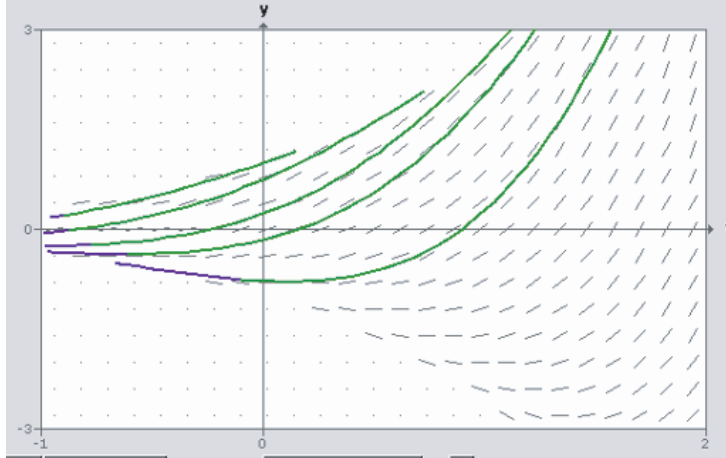


Figure 4: Example 5: Solutions for y — upward branches only.

2 Solving the general first-order linear equation with a substitution

The general first-order linear differential equation is $y' = a(t)y + b(t)$. We can solve it by a substitution like the one in the previous example.

Example 6 Solve the general first-order equation $y' = a(t)y + b(t)$.

The substitution we need is $y(t) = e^{A(t)}z(t)$, where A is any function whose derivative is $A'(t) = a(t)$. Now $y'(t) = a(t)e^{A(t)}z(t) + e^{A(t)}z'(t)$ and the equation becomes

$$a(t)e^{A(t)}z(t) + e^{A(t)}z'(t) = a(t)e^{A(t)}z(t) + b(t)$$

or, on simplifying,

$$z'(t) = b(t)e^{-A(t)}. \quad (5)$$

The last equation is really just an integral. If B is any function whose integral is $b(t)e^{-A(t)}$, then the general solution to (5) is $z(t) = B(t) + c$ and the general solution to the original equation is

$$y(t) = e^{A(t)}(B(t) + c).$$

□

3 The “magic bullet” substitutions in 2×2 linear systems

Next we consider a family of substitutions in homogeneous, linear 2×2 systems of equations with real constant coefficients. These are systems of the form $Y' = AY$ where

$Y(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$ and A is a square matrix with real constant entries,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Each of these substitutions converts the system into a system of much simpler form, both algebraically and geometrically.

Distinct real eigenvalues. We first consider the case in which A has distinct real eigenvalues, say λ_1 and λ_2 , and corresponding eigenvectors

$$V_1 = \begin{pmatrix} k_1 \\ \ell_1 \end{pmatrix} \text{ and } V_2 = \begin{pmatrix} k_2 \\ \ell_2 \end{pmatrix}.$$

These eigenvectors are necessarily linearly independent, so that when we form the matrix

$$Q = \begin{pmatrix} k_1 & k_2 \\ \ell_1 & \ell_2 \end{pmatrix}$$

we find that Q is invertible. We choose this matrix because it has the special property that

$$AQ = Q \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

To see that this is true, write Q as a row vector consisting of two column vectors: $Q = (V_1 \ V_2)$. Then

$$AQ = (AV_1 \ AV_2) = (\lambda_1 V_1 \ \lambda_2 V_2) = (V_1 \ V_2) \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = Q \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Now consider the substitution given by

$$Z = QY$$

where Z is the vector function given by $Z(t) = \begin{pmatrix} z(t) \\ w(t) \end{pmatrix}$. Then $Z' = QY'$ and substitution gives

$$QZ' = AQZ = Q \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} Z$$

and since Q is invertible we can cancel it from both sides to give

$$Z' = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} Z.$$

This is a much easier system! It is completely decoupled; it amounts to just $z' = \lambda_1 z$ and $w' = \lambda_2 w$. These have the solutions $z = c_1 \exp(\lambda_1 t)$ and $w = c_2 \exp(\lambda_2 t)$. The general solution in vector form is

$$Z = \begin{pmatrix} z \\ w \end{pmatrix} = c_1 e^{\lambda_1 t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 e^{\lambda_2 t} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Translating back to the original variables using $Y = QZ$ gives

$$Y = \begin{pmatrix} x \\ y \end{pmatrix} = c_1 e^{\lambda_1 t} \begin{pmatrix} k_1 \\ \ell_1 \end{pmatrix} + c_2 e^{\lambda_2 t} \begin{pmatrix} k_2 \\ \ell_2 \end{pmatrix}$$

or, in another form,

$$\begin{aligned} x &= c_1 k_1 e^{\lambda_1 t} + c_2 k_2 e^{\lambda_2 t} \\ y &= c_1 \ell_1 e^{\lambda_1 t} + c_2 \ell_2 e^{\lambda_2 t}. \end{aligned}$$

This is, in fact, the general solution to the system $Y' = AY$.

We could have gotten these solutions in other ways (mainly by guessing, or by using someone else's good guess from a century ago). The advantage of the substitution is to show what is really happening. The functions z and w amount to a change of coordinates from the original y and z . In the $z - w$ coordinates the straight-line solutions are along the axes, giving us a completely standard-looking sink, source, or saddle.

This substitution works just as well when $\lambda_1 = \lambda_2$, as long as we can find two linearly independent eigenvectors. We address the case in which no such eigenvectors exist below, after the complex case.

Complex eigenvalues. Suppose now that A has complex eigenvalues. Since they are solutions to the characteristic polynomial of A , which has real coefficients, they must be a pair of complex conjugates: say, $\lambda_1 = \alpha + \beta i$ and $\lambda_2 = \alpha - \beta i$. The corresponding eigenvectors also consist of pairs of conjugates: say,

$$V_1 = \begin{pmatrix} k_1 + k_2 i \\ \ell_1 + \ell_2 i \end{pmatrix} \text{ and } V_2 = \begin{pmatrix} k_1 - k_2 i \\ \ell_1 - \ell_2 i \end{pmatrix}.$$

In this case we again use the matrix

$$Q = \begin{pmatrix} k_1 & k_2 \\ \ell_1 & \ell_2 \end{pmatrix},$$

although the components k_1, ℓ_1 , etc have a different meaning than before. (They are pieces of eigenvectors components, not the components themselves.) We can verify that Q satisfies the equation

$$AQ = Q \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}.$$

Now when we make the substitution $Y = QZ$, the system becomes

$$QZ' = AQZ = Q \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$$

which becomes, on cancelling the Q 's,

$$Z' = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}.$$

This system has a simple solution:

$$\begin{aligned}z(t) &= ce^{\alpha t} \sin(\beta(t - \phi)) \\w(t) &= ce^{\alpha t} \cos(\beta(t - \phi))\end{aligned}$$

What does this represent? If $\alpha = 0$ the pair (w, z) is moving in circles in the $w - z$ plane, counterclockwise if $\beta > 0$. If $\alpha > 0$, it is spiralling out, and if $\alpha < 0$ spiralling in, effectively moving along circles but constantly changing to larger or smaller circles, respectively.

We return to the original system by the relationship $Y = QZ$, so that

$$\begin{aligned}x(t) &= ce^{\alpha t} (k_1 \sin(\beta(t - \phi)) + k_2 \cos(\beta(t - \phi))) \\y(t) &= ce^{\alpha t} (\ell_1 \sin(\beta(t - \phi)) + \ell_2 \cos(\beta(t - \phi)))\end{aligned}$$

In the $x - y$ plane, the solutions move along ellipses, or they spiral in or out with respect to ellipses. The effect of the substitution is to turn the ellipses into circles, giving perhaps a clearer geometric understanding of the solutions.

Taking this case a bit further. In the complex-eigenvalue case we can go a bit further by converting to polar coordinates. Starting from the previous equation involving z and w , make the substitutions

$$\begin{aligned}z &= r \sin \theta \\w &= r \cos \theta\end{aligned}$$

so that $w' = r' \cos \theta - r\theta' \sin \theta$ and $z' = r' \sin \theta + r\theta' \cos \theta$. The equation becomes

$$\begin{aligned}r' \cos \theta - r\theta' \sin \theta &= \alpha r \cos \theta + \beta r \sin \theta \\r' \sin \theta + r\theta' \cos \theta &= -\beta r \cos \theta + \alpha r \sin \theta\end{aligned}$$

which simplifies, on solving for r' and θ' , to

$$\begin{aligned}r' &= \alpha r \\ \theta' &= \beta.\end{aligned}$$

This system is now completely decoupled. The solutions, $r(t) = r_0 e^{\alpha t}$ and $\theta(t) = \theta_0 + \beta t$, are almost as simple as they could be.

But we could take it a bit further still. We introduce a variant on polar coordinates, in which r is replaced by $s = \ln r$. The (s, θ) coordinate system doesn't allow us to represent the origin (which runs away to $s = -\infty$) but it has other advantages. If we make a final substitution of s for r , our original linear system becomes

$$\begin{aligned}s' &= \alpha \\ \theta' &= \beta\end{aligned}$$

and the solutions become entirely linear, $s(t) = s_0 + \alpha t$ and $\theta(t) = \theta_0 + \beta t$. That should be simple enough for everybody.

Those are more substitutions than we need to solve the equation, but they may be useful in applications, anyway. A typical analysis of an applied program might work this way: First, set up the differential equation and solve it by standard methods. Then, simplify the presentation by a series of substitutions as above. Then, see what the new coordinates (r, θ) or (s, θ) mean in terms of the original problem. It might turn out that the new coordinates have a hidden significance in the original context. That would be a nice benefit, but even if the new coordinates have no special meaning it might be possible to restate the problem in terms of the new coordinates. Now the setup and solution become trivial, and the analyst can present the results of the analysis on one page. As is the case with many art forms, by working hard enough the analyst can make the problem seem easy.

Repeated real eigenvalues. We now consider the case in which A has only one (repeated) eigenvalue, and only one eigenvector to go with it. Call the eigenvalue λ , and the eigenvector $V_1 = \begin{pmatrix} k_1 \\ \ell_1 \end{pmatrix}$.

In this case, we need to find a *generalized eigenvector*, a vector V_2 that satisfies the *generalized eigenvector equation* $(A - \lambda I)V_2 = V_1$. (We just put V_1 on the right side, in place of the zero vector that would go there if V_2 were an ordinary eigenvector.) When A has only one eigenvector, this equation is always solvable, and V_2 can be chosen to be independent of V_1 . Let $V_2 = \begin{pmatrix} k_2 \\ \ell_2 \end{pmatrix}$.

Now let

$$Q = \begin{pmatrix} k_1 & k_2 \\ \ell_1 & \ell_2 \end{pmatrix},$$

as before, and make the substitution $Y = QZ$. This time we get

$$AQ = Q \begin{pmatrix} \lambda & 1 \\ 1 & \lambda \end{pmatrix},$$

and the system becomes

$$Z' = \begin{pmatrix} \lambda & 1 \\ 1 & \lambda \end{pmatrix} Z.$$

The general solution is

$$Z = e^{\lambda t} \begin{pmatrix} c_1 + c_2 t \\ c_2 \end{pmatrix}$$

or in terms of Y ,

$$Y = e^{\lambda t} \begin{pmatrix} (c_1 + c_2 t)k_1 + c_2 k_2 \\ (c_1 + c_2 t)\ell_1 + c_2 \ell_2 \end{pmatrix}.$$

In all of these cases we are finding a matrix Q such that the matrix $Q^{-1}AQ$ has a desirable form. In effect we are carrying out a similarity transformation on A in order to put it into (the real version of) Jordan form.

4 The “magic bullet” substitutions applied to $n \times n$ systems

The same methods can be generalized directly to systems of higher order. Form the characteristic equation. Its roots are the eigenvalues of A . Find the corresponding eigenvectors. Form Q as follows:

- Each real eigenvector of A becomes a column of Q (avoiding scalar multiples, of course).
- From each pair of complex eigenvectors, pick one; form one column of Q from its real part, and one from its complex part.
- For each repeated eigenvalue with fewer eigenvectors than its multiplicity, find as many generalized eigenvectors, and include them as columns of Q .

In the case of repeated complex eigenvalues, find complex generalized eigenvectors and divide them into their real and complex parts. You may need to form generalized eigenvectors from generalized eigenvectors! In every step, you are simply putting A into its Jordan form. Having formed Q , make the substitution $Y = QZ$, and you will find that the resulting system for Z decouples just as in the two-dimensional case.

5 Using substitutions to bring an equilibrium point to the origin

We briefly illustrate cases in which a substitution is used to simplify a system in a small way, before the serious work begins.

Example 7 *Simplify the system $Y' = AY + B$ where B is an n -dimensional vector of constants.*

Make the substitution $Y = Z - A^{-1}B$. Now $Y' = Z'$ and the original equation becomes

$$Z' = A(Z - A^{-1}B) + B$$

or simply

$$Z' = AZ.$$

This is how we reduce a linear system with a (vector) constant term into a homogeneous system. \square

More generally, given any autonomous system with an equilibrium point, we can transform it into a system whose equilibrium point is at the origin.

Example 8 *Suppose that the system $Y' = F(Y)$ has an equilibrium point at $Y = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$. Move the equilibrium point to the origin.*

Substitute $Y = Z + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$. Now $Y' = Z'$ and the equation becomes

$$Z' = F\left(Z + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}\right).$$

By hypothesis, $F\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, so the transformed equation has an equilibrium point at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. □

In the last two examples, the substitutions represent only a small amount of progress toward a solution. Still, they simplify the equations in a significant way, eliminating one source of potential complications.

(end)