Final Examination: Differential Equations

§1. Exactness and Integrating Factors

I. Exact Equations

Consider a first-order ODE which is linear in the derivative of the unknown:

$$F(t,x) + G(t,x)\,x' = 0. \tag{$*$}$$

Here, suppose F and G are continuous functions in some rectangle $R = [a,b] \times [c,d]$ of the tx-plane. The equation $(*)$ is called **exact** if there is a function Φ defined on some larger open rectangle $R_\varepsilon = (a-\varepsilon, b+\varepsilon) \times (c-\varepsilon, d+\varepsilon)$, where $\varepsilon > 0$, such that throughout R_ε (and hence certainly throughout R) we have

$$\Phi_t = F, \qquad \Phi_x = G.$$

Here, $\Phi_t := \partial\Phi/\partial t$, etc. An exact equation may be solved immediately; the solutions $x(t)$ are defined implicitly by

$$\Phi(t,x) = C, \tag{$**$}$$

where C is a parameter. Thus, the solutions are the *level curves* of the function Φ in the tx-plane. We can see this as follows. Let $x(t)$ be a function satisfying $(**)$, that is, $\Phi(t,x(t)) \equiv C$, for all values of t in some interval $I \subseteq \mathbb{R}$. Differentiating both sides with respect to t, and using the multivariable version of the chain rule, we get

$$\Phi_t(t,x(t))\,\frac{\mathrm{d}t}{\mathrm{d}t} + \Phi_x(t,x(t))\,\frac{\mathrm{d}x}{\mathrm{d}t} \;\equiv\; 0,$$

or after putting $\Phi_t = F$ and $\Phi_x = G$,

$$F(t,x(t)) + G(t,x(t))\,x'(t) \;\equiv\; 0.$$

This says exactly that $x(t)$ satisfies $(*)$. Conversely, if $x(t)$ satisfies $(*)$, then the last two centered equations above are true of $x(t)$, and the chain rule may be used in reverse to conclude $\frac{\mathrm{d}}{\mathrm{d}t}\big[\Phi(t,x(t))\big] \equiv 0$. This may be integrated with respect to t, giving $\Phi(t,x(t)) \equiv C$, where C is a constant of integration. Thus, $x(t)$ also satisfies $(**)$.

The question remains: given only F and G, how can we tell if such a Φ (called a **potential**) actually exists? A hint is provided by the following consideration. If Φ does exist, and if it happens to be C^2, then the equality of the two mixed second-order partial derivatives of a C^2 function (a basic theorem of multivariable calculus) gives

$$F_x = (\Phi_t)_x = (\Phi_x)_t = G_t.$$

Thus, it is reasonable to ask if the condition $F_x = G_t$ is also *sufficient* to guarantee the existence of Φ.

We claim that if F and G are C^1 in the rectangle R_ε, and satisfy the identity

$$\boxed{F_x = G_t} \qquad \text{throughout } R_\varepsilon, \tag{$\dagger$}$$

then a potential Φ may be constructed as follows. Taking (t_0, x_0) to be any specific point in the interior of R, define

$$\boxed{\Phi(t,x) = \int_{t_0}^{t} F(\tau, x_0)\,\mathrm{d}\tau + \int_{x_0}^{x} G(t,\chi)\,\mathrm{d}\chi.} \tag{$\dagger\dagger$}$$

It's easy to show that the function Φ defined by $(\dagger\dagger)$ satisfies $\Phi_x \equiv G$. The first term doesn't involve x, so its x-derivative is 0. As for the second term, its x-derivative is simply $G(t,x)$, by the Fundamental Theorem of Calculus. So, indeed, $\Phi_x(t,x) \equiv G(t,x)$. More difficult is to see that $\Phi_t = F$. It's easy enough to differentiate the first term with respect to t, using the Fundamental Theorem of Calculus. The result is $F(t,x_0)$. To differentiate the second term with respect to t, we need an important trick called *differentiation under the integral sign*. Under suitable conditions (to be discussed shortly), this trick will allow us to compute as follows:

$$\frac{\partial}{\partial t}\int_{x_0}^{x} G(t,\chi)\,\mathrm{d}\chi = \int_{x_0}^{x} \frac{\partial}{\partial t}\big[G(t,\chi)\big]\,\mathrm{d}\chi = \int_{x_0}^{x} G_t(t,\chi)\,\mathrm{d}\chi = \int_{x_0}^{x} F_\chi(t,\chi)\,\mathrm{d}\chi = F(t,x) - F(t,x_0).$$

After moving $\partial/\partial t$ across the x-integration sign (the trick), we used our assumption (†) to write $G_t(t,\chi) = F_\chi(t,\chi)$, and we then used the Fundamental Theorem of Calculus to integrate $\partial F/\partial\chi$ with respect to χ, from $\chi = x_0$ to $\chi = x$. The overall t-derivative of $\Phi(t,x)$ is therefore $F(t,x_0) + [F(t,x) - F(t,x_0)] = F(t,x)$. That is, $\Phi_t(t,x) \equiv F(t,x)$. So the function Φ satisfies the definition of a potential, and we have solved the problem.

But when is it justifiable to move $\partial/\partial t$ across the integral sign in $\int_{x_0}^{x} G(t,\chi)\,\mathrm{d}\chi$?

The best answer is that G must be C^1 on R — an assumption to which we have already committed, for both F and G. However, the *proof* that this condition suffices involves some difficult technical details. Instead, let's content ourselves with the slightly stronger assumption that G is C^2 on R. This makes little difference in practice. Thus, our solution to the problem of constructing a potential will be valid at least whenever F is C^1 and G is C^2.

Assume G is C^2 on R. Write

$$K(t,x) := \int_{x_0}^{x} G(t,\chi)\,\mathrm{d}\chi \quad\text{ and }\quad L(t,x) := \int_{x_0}^{x} G_t(t,\chi)\,\mathrm{d}\chi.$$

If $h \neq 0$ is sufficiently small that $(t+h,x) \in R$, then we also have

$$K(t+h,x) = \int_{x_0}^{x} G(t+h,\chi)\,\mathrm{d}\chi.$$

Since all of these integrals have the same dummy variable and the same limits, we can combine as follows:

$$\frac{K(t+h,x) - K(t,x)}{h} - L(t,x) = \int_{x_0}^{x} \left[\frac{G(t+h,\chi) - G(t,\chi)}{h} - G_t(t,\chi)\right] \mathrm{d}\chi.$$

Taking absolute values, and using the elementary fact that $\left|\int_a^b f(u)\,\mathrm{d}u\right| \leq \int_{\min(a,b)}^{\max(a,b)} |f(u)|\,\mathrm{d}u$, we get

$$\left|\frac{K(t+h,x) - K(t,x)}{h} - L(t,x)\right| \leq \int_{\min(x_0,x)}^{\max(x_0,x)} \left|\frac{G(t+h,\chi) - G(t,\chi)}{h} - G_t(t,\chi)\right| \mathrm{d}\chi. \tag{‡}$$

By the Mean Value Theorem, $G(t+h,\chi) - G(t,\chi) = h \cdot G_t(t+\theta h,\chi)$, where $0 < \theta < 1$ (so that $t+\theta h$ lies between t and $t+h$). It follows that

$$\left|\frac{G(t+h,\chi) - G(t,\chi)}{h} - G_t(t,\chi)\right| = \left|G_t(t+\theta h,\chi) - G_t(t,\chi)\right|.$$

Using the Mean Value Theorem once more, this becomes

$$\left|(\theta h) \cdot G_{tt}(t+\phi\theta h,\chi)\right|, \qquad \text{where } 0 < \phi < 1.$$

Since G is C^2 on R, the function G_{tt} is continuous on R. Also, R is closed and bounded. So the Extreme Value Theorem tells us that there is a constant $M \geq 0$ such that $|G_{tt}| \leq M$ throughout R. Since $t+\phi\theta h$ is between t and $t+h$, and χ is between x_0 and x, the nature of rectangles guarantees that the point $(t+\phi\theta h,\chi)$ lies in R. Thus, regardless of the values of t, x_0, and x, the integrand on the right-hand side of (‡) is bounded above by

$$\left|(\theta h) \cdot G_{tt}(t+\phi\theta h,\chi)\right| \leq M|\theta h| \leq M|h|.$$

[Recall that $0 < \theta < 1$.] It follows that the right-hand side of (‡) is bounded above by

$$M|h||x - x_0|,$$

where $|x - x_0|$ is the length of the interval over which the integration is being done. But the above quantity clearly tends to 0 as $h \to 0$. Thus, looking at the left-hand side of (‡), we see that the partial difference quotient $\{K(t+h,x) - K(t,x)\}/h$ approaches $L(t,x)$ as $h \to 0$. The limit of this difference quotient is by definition $\partial K/\partial t$. Thus, $(\partial/\partial t)K = L$, or in different notation,

$$\frac{\partial}{\partial t}\int_{x_0}^{x} G(t,\chi)\,\mathrm{d}\chi = \int_{x_0}^{x} \frac{\partial}{\partial t}\left[G(t,\chi)\right]\mathrm{d}\chi.$$

II. Integrating Factors

An equation of the form $F(t,x) + G(t,x)\,x' = 0$ will rarely be exact. On the other hand, we may be able to convert it into an exact equation by multiplying it through by another function, $u(t,x)$, called an **integrating factor**. We have already seen the use of integrating factors, when we learned how to solve first-order *linear* equations (that is, linear in both x and x', rather than just in x').

Assuming that F, G, and u are sufficiently smooth on some open rectangle R_ε containing a given closed rectangle R, the new equation $u(t,x)F(t,x) + u(t,x)G(t,x)\,x' = 0$ will be exact provided that

$$(uF)_x = (uG)_t,$$

or

$$uF_x + u_x F = uG_t + u_t G$$

throughout R_ε. This gives a first-order linear *partial differential equation* (PDE) for the unknown $u(t,x)$, namely

$$\boxed{Gu_t - Fu_x = (F_x - G_t)u.} \tag{$*$}$$

Unfortunately, PDEs are generally harder to solve than ODEs; but we don't need the general solution for u — only one particular solution. In certain special cases, we may be able to find one. Again, we expect these cases to be rare. However, any time we can solve a differential equation exactly by an explicit analytic formula, it is notable and worth recording.

The idea goes as follows. Let $*$ represent some *known* binary operation (such as addition or multiplication), chosen in advance, and suppose this operation is differentiable as a function of t and x. Now imagine that equation $(*)$ has a solution (which need not be explicitly known) of the form

$$u(t,x) = f(t * x),$$

where f is a differentiable function of one variable. By the multivariable version of the chain rule,

$$u_t(t,x) = f'(t*x)\,\frac{\partial(t*x)}{\partial t}, \qquad \text{and} \qquad u_x(t,x) = f'(t*x)\,\frac{\partial(t*x)}{\partial x}.$$

Equation $(*)$ becomes

$$f'(t*x)\left[G(t,x)\,\frac{\partial(t*x)}{\partial t} - F(t,x)\,\frac{\partial(t*x)}{\partial x}\right] = \big[F_x(t,x) - G_t(t,x)\big]\,f(t*x).$$

Cross-dividing, we find that

$$\frac{F_x(t,x) - G_t(t,x)}{G(t,x)\,\frac{\partial}{\partial t}(t*x) - F(t,x)\,\frac{\partial}{\partial x}(t*x)} = \frac{f'(t*x)}{f(t*x)}. \tag{$**$}$$

Thus, the existence of an integrating factor of the form $f(t*x)$ forces the quantity on the left-hand side of $(**)$, which involves only the known functions F, G, and $*$, to be functionally dependent on $t*x$. For the right-hand side of $(**)$ is clearly a function of $t*x$ *only* — not of t or x individually. Since F and G are given, and $*$ is known, we can just compute the combination appearing on the left-hand side of $(**)$, and ask whether or not the result is of the form $\phi(t*x)$, where ϕ is some reasonable function (continuous, say). *If it is*, then we have $f'(t*x)/f(t*x) = \phi(t*x)$ where ϕ is a single-variable function we will already have computed explicitly as part of our test. From this equation, we can determine f, and hence u.

Moreover, our logic is reversible. If we cancel our assumption that an integrating factor of the correct form exists, and assume *only* that the combination of F and G appearing on the left-hand side of $(**)$ comes out to be $\phi(t*x)$ for some continuous single-variable function ϕ, then we will show shortly that there is a single-variable function f satisfying $f'(t*x)/f(t*x) = \phi(t*x)$, that is, satisfying $(**)$. But $(**)$ may be cross-multiplied and rearranged into $(*)$, with 'u' now being *defined* by $u(t,x) := f(t*x)$. Since this u satisfies $(*)$, it's an integrating factor.

The only remaining question is how to find f from our knowledge of ϕ.

Temporarily writing $y := t * x$, our problem boils down to finding a single-variable function $f(y)$ that satisfies the separable differential equation $f'(y)/f(y) = \phi(y)$, with ϕ known. That is, $\frac{\mathrm{d}}{\mathrm{d}y}\big[\ln f(y)\big] = \phi(y)$. The solution is

$$f(y) = \exp\left[\int \phi(y)\,\mathrm{d}y\right] = \exp\left[\int_c^y \phi(z)\,\mathrm{d}z\right],$$

where c is a real parameter. We only need one example of such an f, so c may be chosen to equal any convenient number, say c_0 (which will be a known constant), provided only that c_0 belongs to the domain of ϕ. Thus, our integrating factor is given by

$$\boxed{u(t,x) = \exp\left[\int_{c_0}^{t*x} \phi(z)\,\mathrm{d}z\right]}. \tag{†}$$

What this formula means can be described in words as follows. Take the function ϕ, which has been discovered while testing the relevant combination of F and G, and write it as a function of the auxiliary variable z only. Next, antidifferentiate with respect to z. We get a new function of z. In this new function, substitute first $z = t * x$, and then $z = c_0$, and subtract the two results. Finally, exponentiate.

Once the integrating factor $u(t,x)$ has been computed, we still have to solve the exact differential equation $u(t,x)F(t,x) + u(t,x)G(t,x)\,x' = 0$. We do this by finding a potential Φ using the formula derived in subsection I.

The binary operations $*$ that tend to be encountered most frequently are t, x, $t+x$, tx, and t^2+x^2, though there are certainly other useful ones. Let's summarize in a table the computational test that must be applied to F and G in each one of these common cases.

In each row of the table, a specific binary operation $*$ is selected, and the expression on the left-hand side of $(**)$ is updated to reflect the specific partial derivatives of $t * x$ with respect to t and x. The test itself consists in asking whether the given combination of F and G is, or is not, a pure function of $t * x$. This question is represented below by the usual symbol: a question mark over an equals sign. The question is answered by computing the relevant combination of F and G and attempting to simplify it algebraically until it is clear whether or not it depends only on $t * x$. If it does, then in discovering this fact we will also be determining the function ϕ.

$t * x$	Integrating Factor Test
t	$\dfrac{F_x - G_t}{G} \overset{?}{=} \phi(t)$
x	$\dfrac{F_x - G_t}{-F} \overset{?}{=} \phi(x)$
$t+x$	$\dfrac{F_x - G_t}{G - F} \overset{?}{=} \phi(t+x)$
tx	$\dfrac{F_x - G_t}{xG - tF} \overset{?}{=} \phi(tx)$
t^2+x^2	$\dfrac{F_x - G_t}{2tG - 2xF} \overset{?}{=} \phi(t^2+x^2)$

III. An Example

Consider the equation

$$(x - tx^2) + (t + t^2x^2)x' = 0,$$

in which $F(t,x) = x - tx^2$ and $G(t,x) = t + t^2x^2$. This is not exact, since

$$F_x(t,x) = 1 - 2tx \quad\text{ and }\quad G_t(t,x) = 1 + 2tx^2$$

are not identically equal. However, notice that

$$F_x(t,x) - G_t(t,x) = (1 - 2tx) - (1 + 2tx^2) = -2tx(1+x),$$

while

$$x\,G(t,x) - t\,F(t,x) = x(t + t^2x^2) - t(x - tx^2) = t^2x^2(x+1).$$

Dividing these expressions, the factor $x+1$ cancels, and we get

$$\frac{F_x(t,x) - G_t(t,x)}{x\,G(t,x) - t\,F(t,x)} = \frac{-2tx}{t^2x^2} = -\frac{2}{tx}.$$

This is indeed of the form $\phi(tx)$, where

$$\phi(y) := -\frac{2}{y}.$$

By the test in the fourth row of our table, there must be an integrating factor of the form $u(t,x) = f(tx)$. To find it, we integrate $\phi(z)$ with respect to z, from (say) 1 to y, getting $\big[-2\ln z\big]_1^y = -2\ln y = \ln y^{-2}$. Exponentiating, we get

$$f(y) = y^{-2} = \frac{1}{y^2}.$$

The integrating factor is

$$u(t,x) = f(tx) = \frac{1}{t^2x^2}.$$

[At this point, it's fair to ask: what, if anything, motivated the guess that we should compute $xG - tF$? A safe answer is that we can just try all the combinations of F and G in our table (and this table can be extended further if desired) until we find something that works; if we don't, then we give up on integrating factors and try some other method. But in this instance it seems clear that the combination tx plays a special role in the construction of F and G, suggesting that we skip directly to the fourth row of our table; and indeed, we find that the test is passed.]

Now multiplying our equation through by the integrating factor yields the exact equation

$$\left(\frac{x - tx^2}{t^2x^2}\right) + \left(\frac{t + t^2x^2}{t^2x^2}\right)x' = 0,$$

or after simplifying,

$$\left(\frac{1 - tx}{t^2x}\right) + \left(\frac{1 + tx^2}{tx^2}\right)x' = 0.$$

We view this as $\hat{F}(t,x) + \hat{G}(t,x)\,x' = 0$, with $\hat{F}(t,x) = (1-tx)/(t^2x)$ and $\hat{G}(t,x) = (1+tx^2)/(tx^2) = 1 + 1/(tx^2)$. Both these functions are C^2 (in fact C^∞) on any rectangle R interior to one of the four quadrants of the tx-plane, so that the coordinate axes $t = 0$ and $x = 0$ are avoided. Let's illustrate the solution for the interior of quadrant I, for which we may use $t_0 = 1$ and $x_0 = 1$. Our formula for the potential gives

$$\begin{aligned}
\Phi(t,x) &= \int_1^t \hat{F}(\tau,1)\,\mathrm{d}\tau + \int_1^x \hat{G}(t,\chi)\,\mathrm{d}\chi \\
&= \int_1^t \frac{1-\tau}{\tau^2}\,\mathrm{d}\tau + \int_1^x \left(1 + \frac{1}{t\chi^2}\right)\mathrm{d}\chi \\
&= \left[-\frac{1}{\tau} - \ln\tau\right]_{\tau=1}^{\tau=t} + \left[\chi - \frac{1}{t\chi}\right]_{\chi=1}^{\chi=x} \\
&= \left(-\frac{1}{t} - \ln t\right) - (-1) + \left(x - \frac{1}{tx}\right) - \left(1 - \frac{1}{t}\right) \\
&= x - \frac{1}{tx} - \ln t.
\end{aligned}$$

The solution is given implicitly by $\Phi(t,x) = C$, i.e.,

$$x - \frac{1}{tx} - \ln t = C.$$

This can be solved more-or-less explicitly for x in terms of t. We can multiply through by x and rearrange the result into $x^2 - (C + \ln t)x - t^{-1} = 0$. By the quadratic formula, we get two solutions corresponding to each value of C:

$$x^{\pm}(t) = \frac{(C + \ln t) \pm \sqrt{(C + \ln t)^2 + 4t^{-1}}}{2}.$$

IV. Problems

1. Find the general solution of the equation $x^2 + (3tx + x^2 - 1)x' = 0$ in the region $x > 0$.

2. Show that an equation of the form $x\,p(tx) + t\,q(tx)\,x' = 0$, where p and q are any two differentiable functions of one variable, has an integrating factor of the form $f(tx)$.

3. Use the result of problem 2 to solve the equation $x(1 + tx) + t[(tx)^2 - 1]\,x' = 0$.

§2. First-Order Equations with a Variable Isolated and a Variable Missing

I. Theory

The simplest of all differential equations is

$$x' = f(t),$$

where f is a Riemann integrable function. The solution in this case is obtained trivially, by integrating:

$$x = C + \int_{t_0}^{t} f(\tau)\,\mathrm{d}\tau.$$

Here, $t_0 \in I$ is some convenient 'initial time', and C is an arbitrary constant of integration. Notice that the equation $x' = f(t)$ has the simple property that out of the three relevant variables t, x, x', one is *isolated* (here, x') and one is *missing* (here, x). In this subsection, we ask whether other ODE's with this property can be solved in some relatively straightforward way. Of course, x' cannot be the missing variable, since any such equation would not be an ODE.

Consider first the equation

$$t = f(x'). \tag{$*$}$$

Here we assume that f is C^1 on its domain. If f is one-to-one and has a known inverse function g (that is, g is some composition of recognizable functions, each of which can be computed by standard computer algebra systems), then we can of course rewrite $(*)$ in the equivalent form $x' = g(t)$. The function g has the derivative $g'(t) = 1/f'(g(t))$, which exists and varies continuously with t on any interval in which $f'(g(t)) \neq 0$. If the time interval I is taken to have this property, then the equation $x' = g(t)$ is of the kind already considered: x' is isolated, and x is missing.

Thus, let's suppose that f is either not one-to-one, or else has an inverse function that is not known explicitly as a composition of recognizable functions. To proceed toward a solution, we introduce a new auxiliary variable p, subject to the constraint that

$$p = x' = \frac{\mathrm{d}x}{\mathrm{d}t}. \tag{$**$}$$

Then our equation $(*)$ becomes

$$t = f(p).$$

Since t is expressed as a function of p, our aim will be to express x as a function of p. If this can be achieved, then we may replace the special variable p — whose meaning $(\mathrm{d}x/\mathrm{d}t)$ has already been fixed — with a generic parameter, say θ, which is independent of both t and x. This will yield *parametric equations* of the form $(t, x) = (h(\theta), \ell(\theta))$. As θ varies freely through some interval J in the common domain of the functions h and ℓ, a curve in the tx-plane is generated. While we have no explicit equation for this curve of the form $F(t, x) = 0$, we can use a computer graphics system to plot it, and from the plot we can read off various aspects of the implicitly defined solutions $x(t)$. The parametric equations should also contain a parameter C arising from integration, so we get the general solution of $(*)$. For any fixed value of C, one solution curve in the tx-plane will be generated by allowing θ to vary through J.

With this in mind, we use the chain rule and the equations $p = \mathrm{d}x/\mathrm{d}t$ and $t = f(p)$ to compute as follows:

$$\frac{\mathrm{d}x}{\mathrm{d}p} = \frac{\mathrm{d}x}{\mathrm{d}t} \cdot \frac{\mathrm{d}t}{\mathrm{d}p} = p \cdot \frac{\mathrm{d}}{\mathrm{d}p}\big[f(p)\big] = p\,f'(p).$$

Now we can integrate both sides with respect to p (using 'ρ', say, as a dummy variable) to get

$$x = C + \int_{p_0}^{p} \rho\,f'(\rho)\,\mathrm{d}\rho.$$

Here, p_0 is some convenient point, and C is a constant of integration. We have achieved our goal of expressing x as a function of p. Replacing p with the generic parameter θ, we have the parametric equations

$$\boxed{\begin{cases} t = f(\theta) \\ x = C + \displaystyle\int_{p_0}^{\theta} \rho\, f'(\rho)\, \mathrm{d}\rho \end{cases}} \tag{†}$$

To plot solution curves, first fix a value of C. Then let θ vary through some interval J (where $p_0 \in J$) lying within the common domain of f and f' — which is simply the domain of f, since f is assumed C^1. For each of a very large number of closely spaced values of θ in J, plot the corresponding points (t, x) using (†). Connect these points with a smooth interpolating curve (such as a cubic spline). The more values of θ one uses — and the more closely spaced those values are — the more accurate this interpolating curve will be. In this sense, (†) is the general solution of (∗).

The next case is an equation of the form

$$x' = f(x).$$

This is separable, so a suitable method of solution is already known to us. In particular, the general solution is defined implicitly by the relation

$$t = C + \int_{x_0}^{x} \frac{1}{f(\chi)}\, \mathrm{d}\chi,$$

which is valid near any point x_0 that is not a root of f (i.e., for which $f(x_0) \neq 0$). If f has roots, then each one represents a constant function of t that also solves the equation $x' = f(x)$, since both sides will be identically zero for all t. These are called **singular solutions**, since they don't fit into the general solution for any particular value of the parameter C.

While $x' = f(x)$ was an easy case, we still want to see what can be said about the equation

$$x = f(x'). \tag{††}$$

Again, if f has a known inverse function g, then this becomes $x' = g(x)$, are we are back to the separable case. Otherwise, we need a new method for addressing (††). Once again, we introduce the auxiliary variable p, subject to the constraint $p = x' = \mathrm{d}x/\mathrm{d}t$, so that our equation (††) becomes

$$x = f(p).$$

If we can express t as a function of p, then we will once again be able to replace p with a generic parameter θ, and produce parametric equations for solution curves $(t, x) = (h(\theta), \ell(\theta))$ in the tx-plane. To achieve this, we make the following computation:

$$\frac{\mathrm{d}t}{\mathrm{d}p} = \frac{\mathrm{d}t}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}p} = \left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^{-1}\frac{\mathrm{d}x}{\mathrm{d}p} = p^{-1}\frac{\mathrm{d}}{\mathrm{d}p}\big[f(p)\big] = \frac{f'(p)}{p}.$$

Integrating both sides with respect to p (using ρ as a dummy variable), we get

$$t = C + \int_{p_0}^{p} \frac{f'(\rho)}{\rho}\, \mathrm{d}\rho.$$

Here, p_0 must be chosen within the domain of f (and hence the domain of f'), but we must ensure that $p_0 \neq 0$, so that the integral above will not be improper at its lower limit of integration. Replacing the specific variable p with the unrelated parameter θ, which can vary freely through any interval J in the domain of f not containing 0, we get the parametric equations

$$\boxed{\begin{cases} t = C + \displaystyle\int_{p_0}^{\theta} \frac{f'(\rho)}{\rho}\, \mathrm{d}\rho \\ x = f(\theta) \end{cases}} \tag{‡}$$

For each fixed value of C, these give rise to a solution of curve for (††) in the tx-plane.

II. An Example

Consider the equation

$$x = x' \ln x'.$$

Putting $p = x'$, this becomes

$$x = p \ln p.$$

Here, $f(p) = p \ln p$, and therefore $f'(p) = 1 + \ln p$. Also, using $p_0 = 1$, we have

$$t = C + \int_1^p \frac{f'(\rho)}{\rho}\, \mathrm{d}\rho = C + \int_1^p \frac{1 + \ln\rho}{\rho}\, \mathrm{d}\rho.$$

In this case, we can actually antidifferentiate the integrand in terms of elementary functions. Breaking up the fraction into two terms, the first term $1/\rho$ antidifferentiates to $\ln \rho$. The second term is amenable to a change of variable. Let $u = \ln \rho$, so that $\mathrm{d}u = (1/\rho)\, \mathrm{d}\rho$. Then the second term of the integrand (including the differential) becomes $\{(\ln \rho)/\rho\}\, \mathrm{d}\rho = u\, \mathrm{d}u$. The antiderivative of this term is then $\frac{1}{2}u^2 = \frac{1}{2}(\ln \rho)^2$. We now have

$$t = C + \ln p + \tfrac{1}{2}(\ln p)^2.$$

Finally, replace p by the generic parameter θ, which may vary freely in any subinterval of the domain of ln not containing 0. The largest such subinterval is the domain of ln itself, namely $J = (0, \infty)$. Hence the solution is

$$\begin{cases} t = C + \ln\theta + \frac{1}{2}(\ln\theta)^2 \\ x = \theta \ln \theta \end{cases} \qquad (\theta > 0).$$

III. Problems

1. Solve $t = (x')^3 - x'$.
2. Solve $x = x' + (x')^2$.

§3. The Legendre Transformation

I. Theory

Given a first-order ODE of the form

$$f(t, x, x') = 0, \tag{$*$}$$

there is an interesting change of variables due to Legendre that may be of use in obtaining solutions. It tends to work when the function f in $(*)$ depends on t and x in very simple ways. On the other hand, x' may enter into equation $(*)$ in complicated ways. As an example of an equation with these qualities, consider for instance $t + x + x' \sin x' = 0$. This cannot be solved explicitly for x' in terms of t and x; nor is it of any of the forms considered in the last subsection.

We introduce the variable p as usual, subject to the constraint that

$$p = \frac{\mathrm{d}x}{\mathrm{d}t}. \tag{$**$}$$

In terms of the variables t, x, p, we will now define a new set of variables: T, X, P. The definitions are as follows:

$$T := p, \qquad X := tp - x, \qquad P := t. \tag{$\dagger$}$$

Since the combination $tp - x$ occurs here, let's look at the same combination of the capitalized variables, $TP - X$. A simple calculation yields

$$TP - X = pt - (tp - x) = x.$$

To this observation, we add reiterations of the first and last definitions in (†), getting the three equations

$$t = P, \qquad x = TP - X, \qquad p = T. \tag{$\dagger\dagger$}$$

These same three equations also result from definitions (†) on formally replacing all the lower-case letters by their capital versions and vice versa. Moreover, the constraint (∗∗) can also have all its letters capitalized and still yield a true statement. For the chain rule and the fact that $p = x'$ may be used to compute that

$$\frac{\mathrm{d}X}{\mathrm{d}T} = \frac{\mathrm{d}X}{\mathrm{d}t}\,\frac{\mathrm{d}t}{\mathrm{d}T} = \frac{\mathrm{d}X}{\mathrm{d}t}\left(\frac{\mathrm{d}T}{\mathrm{d}t}\right)^{-1} = \frac{\mathrm{d}(tp-x)}{\mathrm{d}t}\left(\frac{\mathrm{d}p}{\mathrm{d}t}\right)^{-1} = (p + tp' - x')(p')^{-1} = (tp')(p')^{-1} = t = P.$$

That is,

$$P = \frac{\mathrm{d}X}{\mathrm{d}T}. \tag{‡}$$

Thus, the transformation taking each lower-case letter to its capital version and each capital letter to its lower-case version *preserves the relationships* (∗∗) *and* (†). This mapping is called the **Legendre transformation**.

Applying the Legendre transformation to our original equation (∗), i.e,, $f(t, x, p) = 0$, we get the new equation

$$\boxed{f(P, TP - X, T) = 0.} \tag{\#}$$

The advantage of (#) over (∗) is that if f is a relatively simple function of its first two arguments, then (#) is simple in the derivative $P = \mathrm{d}X/\mathrm{d}T$, while being perhaps more complicated in the new independent variable T (which now appears in the last argument slot of f).

Assume that the new equation (#) can be solved for X as a function of T. Say the general solution is

$$X = \Phi_C(T),$$

where C is a parameter. Notice that

$$t = P = \frac{\mathrm{d}X}{\mathrm{d}T} = \frac{\mathrm{d}}{\mathrm{d}T}\big[\Phi_C(T)\big] = \Phi'_C(T),$$

while

$$x = TP - X = T\left(\frac{\mathrm{d}X}{\mathrm{d}T}\right) - X = T\,\Phi'_C(T) - \Phi_C(T).$$

In these equations, let's replace the variable T (which has the specific meaning p, i.e., x') by a generic parameter θ. We then get parametric equations for a curve in the tx-plane, namely

$$\boxed{\begin{cases} t = \Phi'_C(\theta) \\ x = \theta\,\Phi'_C(\theta) - \Phi_C(\theta). \end{cases}} \tag{\#\#}$$

Here, θ is free to run through some interval J in the common domain of Φ_C and Φ'_C, provided such an interval exists (if not, then the Legendre method is not applicable). This curve in the tx-plane — actually a *family* of curves, one for each value of C — implicitly defines a solution $x(t)$ of the original equation (∗) near any value of $\theta \in J$ for which

$$\frac{\mathrm{d}x/\mathrm{d}\theta}{\mathrm{d}t/\mathrm{d}\theta} = \frac{\theta\,\Phi''_C(\theta)}{\Phi'_C(\theta)} \neq 0.$$

In this sense, (##) may be regarded as the general solution of the original equation (∗).

II. An Example

Consider the equation mentioned earlier:

$$t + x + x' \sin x' = 0.$$

Another way to write this is

$$t + x + p \sin p = 0, \qquad \text{where } p := \frac{\mathrm{d}x}{\mathrm{d}t}.$$

Applying the Legendre transformation, we get

$$P + (TP - X) + T \sin T = 0, \qquad \text{where } P := \frac{\mathrm{d}X}{\mathrm{d}T}.$$

After writing $\dot{X} := X^{\boldsymbol{\cdot}} := \mathrm{d}X/\mathrm{d}T$, this says

$$\dot{X} + T\dot{X} - X + T\sin T = 0,$$

or

$$(T+1)\dot{X} - X = -T\sin T.$$

Dividing through by $T+1$,

$$\dot{X} - \left(\frac{1}{T+1}\right)X = -\frac{T\sin T}{T+1}.$$

This is a first-order linear ODE for X as a function of T. An integrating factor is given by

$$U = \exp\left[\int \frac{-1}{T+1}\,\mathrm{d}T\right] = \exp\big[-\ln(T+1)\big] = (T+1)^{-1} = \frac{1}{T+1}.$$

Multiplying through by U, the left-hand side becomes $U\dot{X} + \dot{U}X = (UX)^{\boldsymbol{\cdot}}$. We can then integrate both sides with respect to T (using 'τ', say, as a dummy variable), getting

$$UX = C - \int_0^T \frac{\tau\sin\tau}{(\tau+1)^2}\,\mathrm{d}\tau.$$

Here, C is a constant of integration. The integrand here cannot be antidifferentiated in terms of elementary functions, so we leave the integral sign in place. Multiplying through by $1/U = T+1$, we have

$$X = \Phi_C(T) = C(T+1) - (T+1)\int_0^T \frac{\tau\sin\tau}{(\tau+1)^2}\,\mathrm{d}\tau.$$

Our parametric expression for the family of solution curves of the original equation in the tx-plane is now seen to be

$$\begin{cases} t = \Phi_C'(\theta) = C - \displaystyle\int_0^\theta \frac{\tau\sin\tau}{(\tau+1)^2}\,\mathrm{d}\tau - (\theta+1)\frac{\theta\sin\theta}{(\theta+1)^2} \\[2ex] x = \theta\,\Phi_C'(\theta) - \Phi_C(\theta) = \theta\left[C - \displaystyle\int_0^\theta \frac{\tau\sin\tau}{(\tau+1)^2}\,\mathrm{d}\tau - \frac{\theta\sin\theta}{\theta+1}\right] - C(\theta+1) + (\theta+1)\displaystyle\int_0^\theta \frac{\tau\sin\tau}{(\tau+1)^2}\,\mathrm{d}\tau \end{cases}$$

or after some simplification,

$$\begin{cases} t = C - \displaystyle\int_0^\theta \frac{\tau\sin\tau}{(\tau+1)^2}\,\mathrm{d}\tau - \frac{\theta\sin\theta}{\theta+1} \\[2ex] x = -C + \displaystyle\int_0^\theta \frac{\tau\sin\tau}{(\tau+1)^2}\,\mathrm{d}\tau - \frac{\theta^2\sin\theta}{\theta+1} \end{cases} \qquad (\theta > -1).$$

For any fixed C, the curve obtained as (t,x) varies with $\theta \in (-1,\infty)$ is a solution of our equation. By means of the parametric formulas given above, these curves may be plotted for different values of C with the aid of a computer, and the graphical and numerical properties of the solutions $x(t)$ may be inferred.

III. Problems

1. Show that the Legendre transformation can be used to solve the equation $at + bx = f(x')$ for any continuous function f, where a and b are nonzero real constants.

   [*Remark.* If f is a one-to-one function with an explicitly known inverse function g, then $at + bx = f(x')$ may be written in the equivalent form $x' = g(at + bx)$. In this case, a significantly simpler method may be applied to solve the equation. This method will be described later. The Legendre method is necessary when f is either not one-to-one, or else has an inverse function that is not a composition of recognizable standard functions.]

2. Use the result of the preceding problem to solve $t - 2x = (x')^3 - x'$.

§4. Bernoulli Equations

I. Theory

Bernoulli discovered a change of variables that reduces the special first-order equation

$$x' = f(t)\,x + g(t)\,x^{\alpha} \tag{$*$}$$

to a first-order *linear* equation in a new unknown. Equation $(*)$ is called a **Bernoulli equation**. Here, the functions f and g are continuous on a common interval I, and α is a real constant satisfying $\alpha \neq 1$. We note that the case $\alpha = 1$ is actually a separable equation, and thus amenable to exact solution, for it reduces to $x' = [f(t) + g(t)]x$.

To solve the Bernoulli equation, we multiply through by $x^{-\alpha}$, getting

$$x^{-\alpha}\,x' = f(t)\,x^{1-\alpha} + g(t). \tag{$**$}$$

The form of this equation suggests making the change of variable

$$y := x^{1-\alpha}.$$

The chain rule gives

$$y' = (1-\alpha)x^{1-\alpha-1}\,x' = (1-\alpha)x^{-\alpha}\,x'.$$

For brevity, let's write $\beta := 1 - \alpha$. By our assumption on α, we have $\beta \neq 0$. Moreover, $y = x^{\beta}$ and $y' = \beta x^{-\alpha}\,x'$.

If we multiply $(**)$ through by β, we get the equivalent equation $\beta x^{-\alpha}\,x' = \beta f(t)\,x^{\beta} + \beta g(t)$. In terms of y,

$$y' = \beta f(t)\,y + \beta g(t). \tag{$\dagger$}$$

This is a first-order linear equation in y, and can be solved for y in terms of t, by means of the integrating factor

$$u(t) = \exp\left[-\beta\int_{t_0}^{t} f(\tau)\,\mathrm{d}\tau\right].$$

Say the general solution is

$$y = \Phi_C(t),$$

where C is a parameter. Then, since $y = x^{\beta}$, we have $x = y^{1/\beta}$, or

$$\boxed{x = \Big[\Phi_C(t)\Big]^{1/(1-\alpha)}.}$$

This is the general solution of the Bernoulli equation $(*)$.

II. An Example

Take for example

$$x' = t^2 x + t\sqrt{x}\,;$$

a Bernoulli equation with $f(t) = t^2$, $g(t) = t$, and $\alpha = \frac{1}{2}$ (i.e., $x^{\alpha} = x^{1/2} = \sqrt{x}$). Multiply through by $\beta = 1 - \alpha = \frac{1}{2}$, and divide through by $x^{\alpha} = \sqrt{x}$, so the equation becomes

$$\frac{x'}{2\sqrt{x}} = \frac{t^2}{2}\sqrt{x} + \frac{t}{2}.$$

Letting $y := \sqrt{x}$, this says precisely that

$$y' = \frac{t^2}{2}\,y + \frac{t}{2}.$$

Writing this in the usual form of a linear equation, $y' - (t^2/2)y = t/2$, we see that an integrating factor is given by

$$u = \exp\left[-\int_0^t \frac{\tau^2}{2}\,\mathrm{d}\tau\right] = e^{-t^3/6}.$$

Multiplying through by u and integrating from $\tau = 0$ to $\tau = t$, we get

$$e^{-t^3/6} y = C + \int_0^t \frac{\tau}{2}\, e^{-\tau^3/6}\, \mathrm{d}\tau.$$

The integrand cannot be antidifferentiated in terms of elementary functions, so we leave the integral sign in. Now

$$y = Ce^{t^3/6} + \int_0^t \frac{\tau}{2}\, e^{(t^3-\tau^3)/6}\, \mathrm{d}\tau.$$

Since $y = \sqrt{x}$, we have $x = y^2$. The solution of our original problem is seen to be

$$x = \left[Ce^{t^3/6} + \int_0^t \frac{\tau}{2}\, e^{(t^3-\tau^3)/6}\, \mathrm{d}\tau \right]^2.$$

III. Problems

1. Solve $x' = tx + t^2x^3$.

§5. First-Order Analytic Equations: Series Solutions

I. Theory

In this subsection, we consider a first-order equation $x' = f(t,x)$ in which the slope function f is **analytic**. This means that f can be expanded as a convergent *double power series* about each point in its domain. Taking the center of the expansion to be the origin $(0,0)$ of the tx-plane, such a series looks like this:

$$f(t,x) = \sum_{m,n\geq 0} a_{m,n}\, t^m x^n. \tag{$*$}$$

We assume this series converges within a rectangle about $(0,0)$, say for $|t| < A$ and $|x| < B$, with $A > 0$ and $B > 0$. The precise meaning of convergence for doubly indexed series will be discussed shortly.

Our goal will be to prove that *there is an analytic solution* of the equation $x' = f(t,x)$ satisfying the initial condition $x(0) = 0$. It is given by a convergent power series

$$x(t) = \sum_{n\geq 0} c_n\, t^n, \tag{$**$}$$

where the coefficient sequence $(c_n)_{n\geq 0}$ can be computed recursively, from c_0 up to c_N (for any N), in terms of the known values $a_{m,n}$. In some rare but important cases, the recursion can even be solved for c_n as an explicit function of n, so that the *full* coefficient sequence is known, rather than just some initial segment of it. We will also develop a lower bound for the radius of convergence R of the series $(**)$, so that the solution will be known to exist on some interval $|t| < R$ with $R > 0$.

In principle, given enough time, we can compute *any* finite number of terms of $(**)$. Thus, what we're really doing in most cases is computing *polynomial approximations* to the true solution, of arbitrarily high degree. We expect these to approximate the solution well at times t close to $t_0 = 0$, but to decrease in accuracy far from this initial time. At any fixed t, the approximation $x_n(t)$ of the solution $x(t)$ given by the n-th partial sum of $(**)$ will certainly tend improve in accuracy as n increases. However, there is little we can say in general about the absolute error $|x(t) - x_n(t)|$. [We will revisit this question of error bounds in the more specific context of *linear* equations later.]

An obvious advantage of the series technique is its great generality. The vast majority of slope functions one actually encounters in applications are analytic. So this technique yields numerical information about the solutions of a vast array of ODE's. One might object that the *Picard method*, in which the solution of $x' = f(t,x)$ is constructed as the limit of a sequence of iterated integrals [recall: $x_{n+1}(t) = x_0 + \int_{t_0}^t f(\tau, x_n(\tau))\, \mathrm{d}\tau$ for $n \geq 0$, with $x_0(t) \equiv x_0$] already gives a scheme for approximating the solution $x(t)$, and it is even more general: it only requires that f be C^1 with respect to x, rather than analytic in t and x. The answer to this objection is clear: the Picard iterates, while very useful as a theoretical tool for proving the existence of solutions, are virtually impossible to actually *compute*. They are therefore useless in applications, where concrete numerical information is desired. By contrast, a polynomial function whose coefficients can be computed easily by a recursion scheme gives us much more concrete information. So this technique is certainly worth developing in its own right.

The first thing to see is that there is no loss of generality in taking $(0,0)$ to be the center of the expansion $(*)$, or equivalently, in taking the initial condition to be $x(0) = 0$. If instead we wanted to study solutions of $x' = f(t,x)$ satisfying $x(t_0) = x_0$, then we would need a series expansion of f about the point (t_0, x_0), whose general term would be of the form $b_{m,n}\,(t-t_0)^m(x-x_0)^n$, summed over all $m,n \geq 0$. In this case, consider a shifted time variable $T := t - t_0$ and a shifted space variable $X := x - x_0$. By the chain rule, note that

$$\frac{\mathrm{d}X}{\mathrm{d}T} = \frac{\mathrm{d}X}{\mathrm{d}t}\,\frac{\mathrm{d}t}{\mathrm{d}T} = \frac{\mathrm{d}(x-x_0)}{\mathrm{d}t}\,\frac{\mathrm{d}(T+t_0)}{\mathrm{d}T} = \frac{\mathrm{d}x}{\mathrm{d}t}\cdot 1 = \frac{\mathrm{d}x}{\mathrm{d}t}.$$

Define a new function

$$g(u,v) := f(u+t_0,\, v+x_0).$$

Then note that $g(t-t_0,\, x-x_0) = f(t,x)$. Therefore, assuming $x(t)$ is a solution of $\mathrm{d}x/\mathrm{d}t = f(t,x)$, we have

$$\frac{\mathrm{d}X}{\mathrm{d}T} = \frac{\mathrm{d}x}{\mathrm{d}t} = f(t,x) = g(t-t_0,\, x-x_0) = g(T,X).$$

The new equation, $\mathrm{d}X/\mathrm{d}T = g(T,X)$, has an analytic slope function with series expansion

$$g(T,X) = g(t-t_0,\, x-x_0) = f(t,x) = \sum_{m,n\geq 0} b_{m,n}\,(t-t_0)^m(x-x_0)^n = \sum_{m,n\geq 0} b_{m,n}\,T^m X^n.$$

This is centered at $(T,X) = (0,0)$. Finally, our old initial condition $x(t_0) = x_0$ becomes $X(0) = 0$, since

$$X\big|_{T=0} = (x-x_0)\big|_{t-t_0=0} = \big(x\big|_{t=t_0}\big) - x_0 = x(t_0) - x_0 = 0.$$

Thus, our original problem — where the slope function f was expanded as a series about (t_0, x_0), and the solution curve passes through (t_0, x_0) — has been replaced by a new problem in which the slope function g is expanded as a series about $(0,0)$, and the solution curve passes through $(0,0)$. As such a transformation is always possible, we may as well assume for the sake of simplicity that f is expanded about $(0,0)$ from the outset.

A technical point we should get sorted out is what it means for a doubly indexed series to converge. Given a doubly indexed sequence $(a_{m,n} \mid m \geq 0,\ n \geq 0)$, define a **partial double sum** as

$$S_{m,n} := \sum_{i=0}^{m}\sum_{j=0}^{n} a_{i,j}.$$

Suppose there is a number $S \in \mathbb{R}$ such that $S_{m,n} \to S$ as $(m,n) \to (\infty,\infty)$, where the latter means that for any $\varepsilon > 0$, there is some $N = N(\varepsilon)$ such that $|S_{m,n} - S| < \varepsilon$ whenever $m \geq N$ and $n \geq N$. In this case, S is called the **sum** of the double series, and we write

$$S = \sum_{m,n\geq 0} a_{m,n}.$$

An important relation between the terms $a_{m,n}$ and the partial double sums $S_{m,n}$ is that

$$a_{m,n} = S_{m,n} - S_{m-1,n} - S_{m,n-1} + S_{m-1,n-1},$$

as is easily checked. If the double series $\sum a_{m,n}$ is assumed to converge, then all four terms on the right-hand side of this relation approach S, where S is the sum, as $(m,n) \to (\infty,\infty)$. Thus,

$$a_{m,n} \to S - S - S + S = 0 \qquad \text{as } (m,n) \to (\infty,\infty).$$

This is the **Divergence Test for Double Series**. We will use it below.

Consider now a **double power series**,

$$\sum a_{m,n}\,t^m x^n.$$

This is said to be **absolutely convergent** at (t,x) if the double series $\sum |a_{m,n}||t|^m|x|^m$ is convergent in the sense defined above. We define a **biradius of convergence** for the double power series as a pair (α,β) — which is not in general unique – such that whenever $|t|<\alpha$ and $|x|<\beta$, the double power series is *absolutely* convergent. Thus, absolute convergence is built into the definition of "convergent double power series". Saying that a double power series has a biradius of convergence (α,β) where both α and β are positive is already saying that the series is not just convergent, but absolutely convergent, on the open rectangle $(-\alpha,\alpha)\times(-\beta,\beta)$.

On any rectangle of (absolute) convergence, say $(-\alpha,\alpha)\times(-\beta,\beta)$, a double power series can be written as a pair of nested singly indexed series, as follows:

$$\sum_{m,n\geq 0} a_{m,n}\, t^m x^n = \sum_{m\geq 0}\Bigg(\sum_{n\geq 0} a_{m,n}\, t^m x^n\Bigg) = \sum_{n\geq 0}\Bigg(\sum_{n\geq 0} a_{m,n}\, t^m x^n\Bigg).$$

In other words, for any fixed m the series $\sum_{n\geq 0} a_{m,n}\, t^m x^n$ converges, for any fixed n the series $\sum_{m\geq 0} a_{m,n}\, t^m x^n$ converges, and the two nested series are both equal to the sum $f(t,x)$ of the double series. The reason is as follows. The absolute convergence of the double series means that the double series $\sum |a_{m,n}||t|^m|x|^n$ converges. All the terms of this series are nonnegative. For any fixed m, consider $\sum_n |a_{m,n}||t|^m|x|^n$. The ordinary partial sum $\sum_{n=0}^N |a_{m,n}||t|^m|x|^n$ is bounded above by the double partial sum $\sum_{m=0}^M\sum_{n=0}^N |a_{m,n}||t|^m|x|^n$, since the former is a sub-sum of the latter, and the excess terms are all nonnegative. But the double partial sum is bounded above by the full sum $\sum_{m,n\geq 0} |a_{m,n}||t|^m|x|^n$, since the double partial sums increase toward their limit as m and n increase. So the partial sums of $\sum_n |a_{m,n}||t|^m|x|^n$ for a fixed m are bounded above, and increase with n. Therefore they must converge. So the inner series is absolutely convergent, and hence also convergent in the ordinary sense. The same argument works for the other inner series, for any fixed n. Finally, it's a question of whether we can exchange the two limit processes $m\to\infty$ and $n\to\infty$, which occur in different orders in the two nested series. The answer is this: if (1) $S_{m,n}\to S$ as $(m,n)\to(\infty,\infty)$, (2) $\lim_n S_{m,n}$ exists for each fixed m, and (3) $\lim_m S_{m,n}$ exists for each fixed n, then we have $S=\lim_m(\lim_n S_{m,n})=\lim_n(\lim_m S_{m,n})$. The proof is left as an easy exercise in "epsilonics".

Since singly indexed series can be differentiated term-by-term, we have

$$\frac{\partial^p f}{\partial t^p} = \frac{\partial^p}{\partial t^p}\Bigg[\sum_{m\geq 0}\Bigg(\sum_{n\geq 0} a_{m,n}\, t^m x^n\Bigg)\Bigg] = \sum_{m\geq 0}\frac{\partial^p}{\partial t^p}\Bigg[\Bigg(\sum_{n\geq 0} a_{m,n}\, x^n\Bigg)t^m\Bigg].$$

Here, the quantity $\sum_{n\geq 0} a_{m,n}\, x^n$ does not depend on t, to compute the indicated p-th order partial derivative, we simply differentiate t^m with respect to t a total of p times in each term. This gives

$$\frac{\partial^p f}{\partial t^p} = \sum_{m\geq 0}\Bigg[\Bigg(\sum_{n\geq 0} a_{m,n}\, x^n\Bigg) m(m-1)\cdots(m-p+1)\, t^{m-p}\Bigg] = \sum_{m,n\geq 0}[m]_p\, a_{m,n}\, t^{m-p}x^n,$$

where by definition, $[m]_p := m(m-1)\cdots(m-p+1)$. Here, the use of the double series notation at the end is justified because we could have made the same calculation with the summation signs interchanged. Treating the resulting double series as our new starting point, we can now run the same argument, but performing a q-fold differentiation with respect to x. This gives

$$\frac{\partial^{p+q} f}{\partial t^p\,\partial x^q} = \frac{\partial^q}{\partial x^q}\left[\frac{\partial^p f}{\partial t^p}\right] = \sum_{m,n\geq 0}[m]_p\,[n]_q\, a_{m,n}\, t^{m-p}x^{n-q}.$$

Notice that while we've written $m,n\geq 0$, the quantity $[m]_p$ is zero when $m<p$, and similarly $[n]_q=0$ when $n<q$. So in actually, the sum is over all $m\geq p$ and all $n\geq q$. Into the above result, let's plug $t=0$ and $x=0$. It follows that all terms in which either t or x have a positive exponent will vanish. The only surviving term is that in which $m=p$ and $n=q$. In this case, $[m]_p=[p]_p=p!$, and $[n]_q=[q]_q=q!$. Hence,

$$\left[\frac{\partial^{p+q} f}{\partial t^p\,\partial x^q}\right]_{0,0} = p!\,q!\,a_{p,q},$$

where the two subscripted 0's indicate the substitutions $t=0$ and $x=0$. This gives the formula for computing the coefficients in the series expansion of f, based on the partial derivatives of f at $(0,0)$. Namely,

$$\boxed{a_{p,q} = \frac{1}{p!\,q!}\left[\frac{\partial^{p+q} f}{\partial t^p\,\partial x^q}\right]_{0,0}.} \tag{†}$$

Notice in particular that $a_{0,0} = f(0,0)$. Also, $a_{1,0} = f_t(0,0)$ and $a_{0,1} = f_x(0,0)$, while $a_{1,1} = f_{tx}(0,0)$. On the other hand, $a_{2,0} = \frac{1}{2} f_{tt}(0,0)$, and $a_{0,2} = \frac{1}{2} f_{xx}(0,0)$. Etc.

We know from the Picard theory that $(*)$ has a solution $x(t)$ satisfying $x(0) = 0$, defined on some interval $|t| < h$, where $h > 0$. What we don't yet know is that this function is analytic at $t = 0$. However, *if* this function is to have a convergent power series expansion about $t = 0$, then Taylor's Theorem tells us that there is no choice as to what the coefficients must be. Namely:

$$x(0) + x'(0)\,t + \frac{x''(0)}{2!}\,t^2 + \cdots + \frac{x^{(n)}(0)}{n!}\,t^n + \cdots . \tag{††}$$

Let's introduce the notation

$$c_n := \frac{x^{(n)}(0)}{n!}.$$

We do not actually know yet whether or not the series $\sum_{n \geq 0} c_n\, t^n$ converges, so we cannot yet say whether or not $x(t)$ is represented by this series (i.e., whether or not $x(t)$ is analytic at $t = 0$). It is our aim to prove the convergence of this series for all t in some interval $|t| < R$, where $R > 0$.

Clearly $c_0 = x(0) = 0$, by the initial condition. Now directly from $(*)$, we also have

$$c_1 = x'(0) = f(0, x(0)) = f(0,0) = a_{0,0}.$$

We can also take the identity $x'(t) \equiv f(t, x(t))$ on $|t| < h$ and differentiate both sides, getting

$$x''(t) = f_t(t, x(t)) + f_x(t, x(t))\, x'(t), \tag{‡}$$

with the aid of the chain rule. The new identity is still valid for $|t| < h$. Thus,

$$2!\, c_2 = x''(0) = f_t(0,0) + f_x(0,0)\, x'(0) = a_{1,0} + a_{0,1} a_{0,0}.$$

That is,

$$c_2 = \frac{a_{1,0} + a_{0,1} a_{0,0}}{2}.$$

Differentiating (‡) again, we get

$$x'''(t) = \big[f_{tt}(t, x(t)) + f_{tx}(t, x(t))\, x'(t)\big] + \big[f_x(t, x(t))\, x''(t) + \{f_{xt}(t, x(t)) + f_{xx}(t, x(t))\, x'(t)\}\, x'(t)\big].$$

Plugging in $t = 0$ gives

$$x'''(0) = f_{tt}(0,0) + f_{tx}(0,0)\, x'(0) + f_x(0,0)\, x''(0) + \{f_{xt}(0,0) + f_{xx}(0,0)\, x'(0)\}\, x'(0),$$

or (using our previously derived expressions for $x'(0)$ and $x''(0)$),

$$3!\, c_3 = x'''(0) = 2a_{2,0} + a_{1,1} a_{0,0} + a_{0,1}(a_{1,0} + a_{0,1} a_{0,0}) + \{a_{1,1} + (2a_{0,2}) a_{0,0}\} a_{0,0}.$$

That is,

$$c_3 = \frac{2a_{2,0} + a_{1,1} a_{0,0} + a_{0,1} a_{1,0} + a_{0,1}^2 a_{0,0} + a_{0,0} a_{1,1} + 2a_{0,2} a_{0,0}^2}{6}.$$

Continuing in this same way indefinitely, we find that for every n,

$$c_n = P_n(a_{0,0}; a_{1,0}, a_{0,1}; a_{2,0}, a_{1,1}, a_{0,2};\ \ldots\ ; a_{n,0}, \ldots, a_{0,n}), \tag{‡‡}$$

where P_n is a multivariable polynomial in the indicated arguments, *which has exclusively positive real coefficients.* The positivity of the coefficients is crucial to the argument, so let's make sure we understand where it is coming from. In every case, the last step in solving for c_n is to divide by $n!$, and obviously this doesn't affect the positivity of the coefficients. Now in computing $n!\, c_n$, we're always using the chain rule to differentiate some identity we've already derived. The chain rule only involves two operations: addition and multiplication. Since the positive reals are closed under these two operations, we see that every new differentiation keeps the coefficients of the $a_{p,q}$'s positive. It's also clear that no $a_{p,q}$ with $p + q > n$ is involved in the formula for c_n, since these come from partial derivatives of f

involving a total of more than n differentiations, while c_n is computed using just n differentiations (via the chain rule).

The importance of the positivity of the coefficients of P_n comes from the following observation. Suppose we consider a new differential equation $x' = \phi(t,x)$, where ϕ is analytic at $(0,0)$, and has the expansion

$$\phi(t,x) = \sum_{m,n\geq 0} b_{m,n}\, t^m x^n.$$

Suppose also that we have an *term-by-term comparison*,

$$|a_{m,n}| \leq b_{m,n} \qquad \text{for all } m \geq 0 \text{ and } n \geq 0. \tag{\#}$$

Finally, suppose that the equation $x' = \phi(t,x)$ can be *solved explicitly* subject to the initial condition $x(0) = 0$, and that the unique solution (which is now explicitly known) can be expanded as a convergent power series

$$\sum_{n\geq 0} d_n\, t^n, \tag{\#\#}$$

with radius of convergence $R > 0$. By exactly the same calculations with repeated differentiation that we performed above, we would find that

$$d_n = P_n(b_{0,0}; b_{1,0}, b_{0,1}; b_{2,0}, b_{1,1}, b_{0,2};\, \ldots\, ; b_{n,0}, \ldots, b_{0,n}),$$

where P_n is the exact same multivariable polynomial as occurred previously. Here, finally, is where we use the positivity of the coefficients of P_n. For each n, we now have

$$\begin{aligned}
|c_n| &= |P_n(a_{0,0}; a_{1,0}, a_{0,1};\, \ldots\, ; a_{n,0}, \ldots, a_{0,n})| \\
&\leq P_n\big(|a_{0,0}|; |a_{1,0}|, |a_{0,1}|;\, \ldots\, ; |a_{n,0}|, \ldots, |a_{0,n}|\big) \\
&\leq P_n(b_{0,0}; b_{1,0}, b_{0,1};\, \ldots\, ; b_{n,0}, \ldots, b_{0,n}) \\
&= d_n.
\end{aligned}$$

The first inequality here uses the triangle inequality to push the absolute value sign into all the terms of P_n; but in so doing, the actual coefficients of P_n do not change, because they are already positive. Hence, the result is the exact same polynomial P_n, but with $|a_{p,q}|$ in place of $a_{p,q}$. The second inequality works because of the term-by-term comparison (#). These inequalities can be multiplied by one another, multiplied by the (positive) coefficients of P_n (which does not reverse them), and added together, producing other true inequalities.

But now we see something very useful. Since

$$|c_n| \leq d_n \qquad \text{for all } n,$$

and since the series $\sum_{n\geq 0} d_n\, t^n$ is known to converge for all values of t satisfying $|t| < R$, the Comparison Test from calculus shows that the series $\sum_{n\geq 0} c_n\, t^n$ also converges for all t with $|t| < R$. That is, our series (††) converges whenever $|t| < R$, so that the Picard solution $x(t)$ of $x' = f(t,x)$ is actually analytic at $t = 0$, as desired.

To complete the argument, we still need to find an acceptable function $\phi(t,x)$. We claim that the function

$$\phi(t,x) := \frac{M}{(1 - t/A)(1 - x/B)} \tag{\%}$$

will work, where M is a positive constant we will describe shortly. A technicality that arises here is that for this argument to work, we need to take our values of A and B to be slightly smaller than the true radii $\alpha > 0$ and $\beta > 0$ in a biradius. Thus, the series for f converges absolutely on the *boundary* of the rectangle $[-A, A] \times [-B, B]$, and not just in the interior. In particular, the series converges absolutely when $(t,x) = (A,B)$, the upper-right corner point of the rectangle.

Notice that ϕ is M times the product of two geometric series: one in t/A, and the other in x/B. Thus,

$$\phi(t,x) = M\Bigg(\sum_{m\geq 0} \frac{t^m}{A^m}\Bigg)\Bigg(\sum_{n\geq 0} \frac{x^n}{B^n}\Bigg) = \sum_{m,n\geq 0} \bigg(\frac{M}{A^m B^n}\bigg)\, t^m x^n. \tag{\%\%}$$

This expansion is valid for all (t, x) with $|t| < A$ and $|x| < B$, since these are precisely the conditions under which the common ratios t/A and x/B of the geometric series will satisfy the requirements $|t/A| < 1$ and $|x/B| < 1$. Thus,

$$b_{m,n} := \frac{M}{A^m B^n}.$$

To prove the term-by-term comparison (#), we need to prove that there is some $M > 0$ such that for all m and n,

$$|a_{m,n}|\, A^m B^n \le M. \tag{@}$$

For (@) says precisely that $|a_{m,n}| \le b_{m,n}$. Since $A > 0$ and $B > 0$, (@) says that $|a_{m,n}\, A^m B^n| \le M$ for some M. But this is easy to deduce from the Divergence Test for Double Series. We know that the series expansion of f converges absolutely at (A, B). Thus, the double series $\sum |a_{m,n}\, A^m B^n|$ converges. By the Divergence Test, we must have $|a_{m,n}\, A^m B^n| \to 0$ as $(m, n) \to (\infty, \infty)$. It follows that for some large N, all the terms with $m \ge N$ and $n \ge N$ have $|a_{m,n}\, A^m B^n| \le 1$ (say). This leaves three "regions" of subscript-pairs (m, n) unaccounted for. There's the finite "rectangle" where $0 \le m < N$ and $0 \le n < N$; but this contains only finitely many pairs (N^2, to be exact), and thus we can set P to be the maximum of $|a_{m,n}\, A^m B^n|$ on these pairs. We also have two "infinite strips". One is the set of pairs (m, n) such that $0 \le m < N$ but $n \ge N$. For each fixed m, however, the series $\sum_n |a_{m,n}\, A^m B^n|$ converges. So, by the Divergence Test for ordinary series, there is some N_m (dependent on m) such that $|a_{m,n}\, A^m B^n| \le 1$ when $n \ge N_m$. Since there are only finitely many relevant values of m (from 0 to $N - 1$), we can let $N^* := \max(N_0, N_1, \ldots, N_{N-1})$, and we then see that for *any* value $m \in \{0, 1, \ldots, N - 1\}$, as long as $n \ge N^*$, we have $|a_{m,n}\, A^m B^n| \le 1$. This still leaves all the pairs (m, n) with $0 \le m < N$ and $N \le n < N^*$, but again there are only finitely many of these, and we can let the maximum value of $|a_{m,n}\, A^m B^n|$ on these pairs be Q. A similar analysis can be done for the other infinite strip. Letting M denote the maximum of the separate upper bounds we have on each of the "regions", we see that $|a_{m,n}\, A^m B^n| \le M$ for *all* m and n. This is what we wanted.

All we need to do now is show that the equation $x' = \phi(t, x)$ can be solved explicitly, and that its solution can be expanded as a convergent power series about $t = 0$. The remarkable thing is that because of the specific form of ϕ given in (%), the equation in question is separable. We can write it as

$$(1 - x/B)\, x' = \frac{M}{1 - t/A},$$

or

$$(1 - x/B)\, \mathrm{d}x = \frac{M}{1 - t/A}\, \mathrm{d}t.$$

Integration gives

$$x - \frac{x^2}{2B} = -MA \ln(1 - t/A) + C,$$

where C is a constant of integration. Since $x = 0$ when $t = 0$, and since $\ln 1 = 0$, it's clear that we must take $C = 0$ to get the initial condition satisfied. Now we can rearrange our solution as follows:

$$x^2 - 2Bx - MAB \ln(1 - t/A) = 0.$$

By the quadratic formula,

$$x = \frac{2B \pm \sqrt{4B^2 + 4MAB \ln(1 - t/A)}}{2} = B - \sqrt{B^2 + MAB \ln(1 - t/A)}.$$

We have to choose the minus sign here, since when $t = 0$ we get $x(0) = B - \sqrt{B^2} = 0$ with the minus sign, whereas we would have gotten $x(0) = 2B > 0$ with the plus sign. Let's factor B^2 out of the radical sign as B, getting

$$x = B - B\sqrt{1 + M(A/B) \ln(1 - t/A)}. \tag{\$}$$

Now the expression $\ln(1 - u)$ has a convergent power series expansion in powers of u, provided only that $|u| < 1$. For us, $u := t/A$, so we need to have $|t| < A$, which is an assumption we have already put into effect. Also, $\sqrt{1 + v} = (1 + v)^{1/2}$ has a convergent power series expansion in powers of v (Newton's Binomial Series), provided only that $|v| < 1$. We need to have $v := M(A/B) \ln(1 - u)$ satisfy $|v| < 1$. This will happen if $|ln(1 - u)| < B/MA$, or $-B/MA < \ln(1 - u) < B/MA$. Exponentiating (and noting that the exponential function is everywhere increasing), we find that $e^{-B/MA} < 1-u < e^{B/MA}$. Putting $u = t/A$, and isolating t, we find that our solution ($) of $x' = \phi(t,x)$ and $x(0) = 0$ has a convergent power series representation for all t in the range

$$A(1 - e^{B/MA}) < t < A(1 - e^{-B/MA}).$$

Here, the lower bound is negative, since e raised to a positive power is always strictly greater than 1. The upper bound is positive, since e raised to a negative power is strictly less than 1. Thus, our lower estimate for the radius of convergence of the solution (and hence also of the analytic solution of $x' = f(t,x)$) is

$$\boxed{R = A\ \min\left(e^{B/MA} - 1,\ 1 - e^{-B/MA}\right) > 0.} \tag{\$\$}$$

We caution that this lower bound for the radius of convergence is difficult to use in practice, since it is generally difficult to find A, B, and M, based merely on a given formula for f. The important point, however, is that ($$) furnishes a proof that the convergence of the analytic solution is occurring on a nontrivial interval about $t = 0$.

A point in favor of the preceding method of proof is that it extends to higher dimensions (and higher order equations) with very little extra work. If we have a system of coupled first-order ODE's, or equivalently a single first-order vector ODE

$$\mathbf{x}' = \mathbf{f}(t, \mathbf{x}),$$

where each component function f_j of $\mathbf{f}$ is analytic about $(0, \mathbf{0}) \in \mathbb{R} \times \mathbb{R}^d$ in a sense exactly analogous to that described above for the case of $d = 1$. If we pair with this equation the initial condition $\mathbf{x}(0) = \mathbf{0} \in \mathbb{R}^d$, then we can still take as a **dominant** for each component function $f_j(t, \mathbf{x})$ the auxiliary function

$$\phi(t, \mathbf{x}) := \frac{M}{(1 - t/A)(1 - x_1/B)(a - x_2/B) \cdots (1 - x_d/B)},$$

where $(A, B, \ldots, B)$ is a **multiradius of convergence** for the multivariable series expansion of f_j, and where M is an upper bound for $\left|a_{m,n_1,\ldots,n_d}\, A^m B^{n_1+\cdots+n_d}\right|$ over all combinations of subscripts $(m, n_1, \ldots, n_d)$. Also define

$$\mathbf{\Phi}(t, \mathbf{x}) := (\phi(t, \mathbf{x}),\ \phi(t, \mathbf{x}),\ \ldots,\ \phi(t, \mathbf{x})),$$

and consider the vector ODE $\mathbf{x}' = \mathbf{\Phi}(t, \mathbf{x})$. Because all the component functions are the same, and because ϕ is a symmetric function of the x_i's (i.e., invariant under permutations of the x_i's), all d of the scalar equations in this vector system are essentially the same. If $y(t)$ is a solution of the scalar differential equation

$$y' = \psi(t, y) := \frac{M}{(1 - t/A)(1 - y/B)^d}, \tag{$***$}$$

then take

$$\mathbf{y}(t) := (y(t), y(t), \ldots, y(t)).$$

This will be a solution of

$$\mathbf{y}' = \mathbf{\Phi}(t, \mathbf{y}),$$

as you may readily check (set $x_j = y$ for all j). But the equation ($***$) is again separable. Its explicit solution is

$$y = B - B\ \sqrt[d+1]{1 + (d+1)M(A/B)\ln(1 - t/A)}.$$

Once again using the power series expansion of the natural logarithm, as well as Newton's Binomial Series for $(1+v)^{1/(d+1)}$ in powers of v, we find that y can be expanded in a convergent power series on some interval $|t| < R$, where $R > 0$. This means that the solution $\boldsymbol{y}(t)$ of the vector ODE $\mathbf{y}' = \mathbf{\Phi}(t, \mathbf{y})$ is analytic (in the sense that each of its component functions is analytic) in a positive-radius interval about $t = 0$. A term-by-term comparison argument will now establish that our original problem $\mathbf{x}' = \mathbf{f}(t, \mathbf{x})$ has an analytic solution $\mathbf{x}(t) = (x_1(t), \ldots, x_d(t))$ which also converges on $|t| < R$.

The argument goes like this. Each component $x_j(t)$ has its n-th Taylor coefficient $c_{j,n} := {x_j}^{(n)}(0)/n!$ dominated in absolute value by the n-th coefficient in the series for $y(t)$, namely $d_n := y^{(n)}(0)/n!$. In other words, we have $|c_{j,n}| \le d_n$ for all n, where $j \in \{1, 2, \ldots, d\}$ is fixed. We get this by expressing $c_{j,n}$ as a polynomial, with *positive* *coefficients*, in the $a_{m,n_1,\ldots,n_d}$ (which come from the multi-index series expansion of f), exactly as in the case $d = 1$. But because M has been chosen to satisfy $\left|a_{m,n_1,\ldots,n_d}\, A^m B^{n_1+\cdots+n_d}\right| \le M$, we also have

$$\left|a_{m,n_1,\ldots,n_d}\right| \le M A^{-m} B^{-(n_1+\cdots+n_d)} =: b_{m,n_1,\ldots,n_d}.$$

Finally, these $b_{m,n_1,\ldots,n_d}$'s are none other than the coefficients in the multi-index series expansion of

$$\phi(t,\mathbf{x}) = M\,(1-t/A)^{-1}(1-x_1/B)^{-1}\cdots(1-x_d/B)^{-1},$$

as a product of geometric series in $d+1$ distinct variables. Since d_n is the same polynomial function of the $b_{m,n_1,\ldots,n_d}$ as $c_{j,n}$ is of the $a_{m,n_1,\ldots,n_d}$, and since that polynomial function has exclusively positive coefficients, the fact that $\left|a_{m,n_1,\ldots,n_d}\right| \le b_{m,n_1,\ldots,n_d}$ implies that $|c_{j,n}| \le d_n$. Thus, the Comparison Test lets us deduce the convergence of the Taylor series for $x_j(t)$ from the already-known convergence of the Taylor series of $y(t)$, on the same positive-radius interval. We see, therefore, that $\mathbf{x}(t)$ is analytic at $t = 0$, as desired.

Next, recall from several points during class that a higher order scalar equation of the form

$$x^{(d)} = f(t, x, x', \ldots, x^{(d-1)}) \tag{†††}$$

can be treated as a first-order vector ODE by introducing auxiliary variables — namely, define $\mathbf{x} := (x_1, x_2, \ldots, x_d)$, where

$$x_1 := x, \qquad x_2 := x', \qquad x_3 := x'', \qquad \text{etc.} \qquad x_d := x^{(d-1)}.$$

Then we have $\mathbf{x}' = \mathbf{f}(t,\mathbf{x})$, where by definition

$$\mathbf{f}(t, x_1, x_2, \ldots, x_{d-1}, x_d) := (x_2, x_3, \ldots, x_d, f(t,\mathbf{x})).$$

Here, notice that *every* component function will be analytic provided only that f is assumed to be analytic. The reason is that the first $d-1$ component functions are just *polynomials* — and hence convergent power series — in the x_j's. Appealing to our discussion of analytic vector ODE's, it therefore follows that there is an analytic solution of the higher order scalar ODE (†††) satisfying a given set of initial conditions, such as the standardized conditions

$$x(0) = x'(0) = x''(0) = \cdots = x^{(d-1)}(0) = 0.$$

Again, the radius of convergence of the solution $x(t)$ is some positive number R.

Coming back now to the first-order scalar case, $x' = f(t,x)$, one point we haven't addressed yet is how to actually compute the coefficients c_n of our series solution. There are two approaches. We have already seen one way, where repeated differentiation is used to express the coefficients c_n of the solution as polynomials in the coefficients $a_{m,n}$ of the slope function f. This method always works, but except in special cases it usually becomes prohibitively laborious. The other approach is to substitute the ansatz $\sum_{n\ge 0} c_n\, t^n$ for x in the equation $x' = f(t,x)$. By equating coefficients of like powers of t on the two sides of the equation, we hope to get a recursive relation that will determine c_{n+1} in terms of $c_0, c_1, \ldots, c_n$. This is not always possible. It works only when the double series for $f(t,x)$ is of finite degree in x for each fixed power of t. That is, it works when

$$f(t,x) = \sum_{m\ge 0} \sum_{n=0}^{d(m)} a_{m,n}\, t^m x^n = \sum_{m\ge 0} P_m(x)\, t^m,$$

where $P_m(x) = \sum_{n=0}^{d(m)} a_{m,n}\, x^n$ is a polynomial for each fixed m. Without this assumption, c_{n+1} will be expressible as a function of $c_0, c_1, \ldots, c_n$, but this function will not be *computable*, since it will involve infinitely many basic operations that have to be performed on the finite set of values $c_0, c_1, \ldots, c_n$. This infinitary process does converge, but its exact value will not be obtainable in a finite time through computations, so it is not useful. When this method works, however, it is much more efficient than the differentiation method.

Without going into too much detail, it isn't difficult to simplify the equation $x' = f(t,x)$, in the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{\ell\ge 0} c_\ell\, t^\ell\right) = \sum_{m\ge 0} \sum_{n=0}^{d(m)} a_{m,n}\, t^m \left(\sum_{k\ge 0} c_k\, t^k\right)^{n},$$

by performing all the operations indicated. What we get is the following:

$$\sum_{\ell\geq 0}(\ell+1)\,c_{\ell+1}\,t^\ell = \sum_{\ell\geq 0}\left[\sum_{k=0}^{\ell}\sum_{n=0}^{d(\ell-k)} a_{\ell-k,n}\left\{\sum_{i_1+i_2+\cdots+i_n=k} c_{i_1}c_{i_2}\cdots c_{i_n}\right\}\right]t^\ell.$$

Setting equal the coefficient of t^ℓ on each side, and solving for $c_{\ell+1}$, we get the desired recursion

$$\boxed{c_{\ell+1} = \frac{1}{\ell+1}\sum_{k=0}^{\ell}\sum_{n=0}^{d(\ell-k)} a_{\ell-k,n}\left[\sum_{i_1+i_2+\cdots+i_n=k} c_{i_1}c_{i_2}\cdots c_{i_n}\right]} \qquad (\ddagger\ddagger\ddagger)$$

To check that this is a valid and computable recursion, first notice that all the summations are finite. In particular, the innermost summation is over all n-tuples $(i_1, i_2, \ldots, i_n)$ of nonnegative integers for which $i_1 + i_2 + \cdots + i_n = k$, with k fixed. There are only finitely many such n-tuples, so that summation is finite, and hence computable. The middle summation and the outermost summation are obviously finite, and hence computable. For validity, we must check that no c_j with $j \geq \ell+1$ appears on the right-hand side of ($\ddagger\ddagger\ddagger$). For $c_{\ell+1}$ must be computable from $c_0, c_1, \ldots, c_\ell$ alone. To see that this is the case, notice that each i_s is $\leq k$, as they all add up to k; but k is $\leq \ell$, since k runs from 0 to ℓ in the outermost summation sign. It follows that $i_s \leq \ell$ for each s, and therefore $c_{i_s} \in \{c_0, c_1, \ldots, c_\ell\}$. Since these are the only c_j's appearing on the right-hand side of the recursion, the validity has been verified.

II. An Example

Consider the equation

$$x' = e^{-tx}.$$

If the slope function were $e^{-(t+x)}$ instead of e^{-tx}, then we could factor it as $e^{-t}e^{-x}$, and the equation would be separable. But this is not the case, and some thought should convince you that we have not encountered any exact technique capable of solving the above ODE. We can use the standard exponential series to write

$$x' = \sum_{m\geq 0}\frac{(-1)^m}{m!}\,t^m x^m.$$

Notice that the series expansion here satisfies the computability condition described above, with $P_m(x) := \{(-1)^m/m!\}\,x^m$ being a "one-term polynomial", of degree $d(m) = m$. Notice also that

$$a_{m,n} = \begin{cases} (-1)^n/n! & m = n \\ 0 & m \neq n \end{cases}$$

Putting all this into the recursion ($\ddagger\ddagger\ddagger$) above, we find that

$$c_{\ell+1} = \frac{1}{\ell+1}\sum_{k=0}^{\ell}\frac{(-1)^{\ell-k}}{(\ell-k)!}\left[\sum_{i_1+i_2+\cdots+i_{\ell-k}=k} c_{i_1}c_{i_2}\cdots c_{i_{\ell-k}}\right].$$

Let's compute some of the coefficients with this formula. First, c_0 is unconstrained, so write $c_0 = C$. Taking $\ell = 0$,

$$c_1 = \frac{1}{0+1}\,\frac{(-1)^{0-0}}{(0-0)!}\sum(\;) = 0.$$

Here, $\sum(\;)$ indicates an "empty sum", which is defined to be 0 by convention. The reason is that $\ell = 0$ forces $k = 0$, so that $\ell - k = 0$; but then the i_s's should form a 0-tuple, which is an empty list. Next, take $\ell = 1$ to get

$$c_2 = \frac{1}{1+1}\sum_{k=0}^{1}\frac{(-1)^{1-k}}{(1-k)!}\left[\sum_{i_1+i_2+\cdots+i_{1-k}=k} c_{i_1}c_{i_2}\cdots c_{i_{1-k}}\right] = \frac{1}{2}\left[-\frac{1}{1!}\sum_{i_1=0} c_{i_1} + \frac{1}{0!}\sum(\;)\right] = -\frac{1}{2}c_0 = -\frac{C}{2}.$$

Next, in somewhat less detail,

$$c_3 = \frac{1}{3}\left[\frac{1}{2!}\sum_{i_1+i_2=0} c_{i_1}c_{i_2} - \frac{1}{1!}\sum_{i_1=1} c_{i_1} - \frac{1}{0!}\sum(\;)\right] = \frac{1}{3}\left[\frac{1}{2}c_0c_0 - c_1\right] = \frac{c_0^2}{6} - \frac{c_1}{3} = \frac{C^2}{6}.$$

Again,

$$c_4 = \frac{1}{4}\left[-\frac{1}{3!}\sum_{i_1+i_2+i_3=0} c_{i_1}c_{i_2}c_{i_3} + \frac{1}{2!}\sum_{i_1+i_2=1} c_{i_1}c_{i_2} - \frac{1}{1!}\sum_{i_1=2} c_{i_1} + \frac{1}{0!}\sum(\ )\right] = \frac{1}{4}\left[-\frac{c_0^3}{6} + \frac{c_1c_0+c_0c_1}{2} - c_2\right] = \frac{3C-C^3}{24}.$$

Thus, our general solution has a convergent series expansion starting out as

$$x = C - \frac{C}{2}t^2 + \frac{C^2}{6}t^3 + \frac{3C-C^3}{24}t^4 + \cdots.$$

Although we have yet to find a coefficient that doesn't vanish with C, such a coefficient must eventually come up, since otherwise taking $C=0$ would yield the identically zero function as a solution, which is obviously wrong (if $x \equiv 0$, then $x' \equiv 0$ but $e^{-tx} \equiv 1$). There is no obvious pattern to the coefficients as polynomials in C, so we just have to compute as many as we have patience (or computer power) to deal with.

What about estimating the radius of convergence of our series solution? Consider the formula

$$R = A\,\min\left(e^{B/MA} - 1,\, 1 - e^{-B/MA}\right).$$

The exponential series always converges, so we can take A and B to be any convenient values. An easy estimate can be had by taking $A=B=1$. Then we need M to be an upper bound for

$$|a_{m,n}\,A^mB^n| = |a_{m,n}| = \begin{cases} 1/m! & n=m \\ 0 & n \neq m \end{cases}$$

It's clear, then, that $M=1$ works. Thus, a very conservative (but easily obtained) lower estimate for the radius of convergence is

$$R \geq 1\cdot\min(e-1,\,1-e^{-1}) = 1-e^{-1} \approx 0.63.$$

Better estimates (i.e., larger values of R) are not too difficult to obtain, but this illustrates the basic idea.

III. Problems

1. Expand to at least the fourth power of t the general series solution of $x' = 1/(1-tx)$, and give a valid (if crude) lower estimate on the radius of convergence of the series solution.

   [*Hint.* You may use the general recursion (‡‡‡) if you wish. However, you may want to rewrite the equation in the form $x' - txx' = 1$, and substitute the ansatz $x = \sum_{\ell\geq 0} c_\ell\, t^\ell$ directly into this modified equation. This will involve multiplying the series for x' and x, and then multiplying in the extra factor of t by simply raising every exponent by 1.

   Be careful in subtracting the two terms on the left-hand side; you will have to re-index one of these series so that their general terms are "like terms" that can be combined. This may involve separating off a small number of terms from the beginning of one of the series, so that the lower limits of summation can be made to match before combining them into one series. The total coefficient of t^0 on the left-hand side should then be set equal to 1 (the constant on the right-hand side), while the total coefficient of every positive power of t on the left-hand side should be set equal to 0. From this, get a recursion for $c_{\ell+1}$ in terms of $c_0, c_1, \ldots, c_\ell$.]

§6. Riccati Equations

I. Theory

The **Riccati equation** is

$$x' = f(t)\,x^2 + g(t)\,x + h(t), \tag{$*$}$$

where f, g, h are continuous functions defined on a common interval I. Notice that $(*)$ is not linear; however, the nonlinearity is of the simplest possible kind: the only obstruction is the presence of a *quadratic* term in x. This makes the Riccati equation an important case study in one-dimensional nonlinear dynamics. Nonlinear ODE's can often be approximated by Riccati ODE's, after neglecting all terms of a Taylor expansion in powers of x of degree > 2.

The Riccati equation is not susceptible to exact solution in closed form except in certain special cases, but there is still much that can be said about its solutions. The first thing we're going to prove is that there is a change of variable that reduces $(*)$ to a *linear* equation of order 2 in a new unknown. Thus, there is a tradeoff between the order of the differential equation and the degree of the slope function in x. The Riccati equation is of order 1 and degree 2 in x, while the associated linear equation will be of order 2 and degree 1 in x. The existence of this association between a Riccati equation and a linear equation is mainly of theoretical rather than practical interest.

We'll also see how the general solution of $(*)$ can be obtained by an integration once a particular solution is known. As for how one might find a particular solution, there is a technique that searches for polynomial solutions when f, g, h are rational functions. There is no guarantee that a polynomial solution exists, but often one does.

The first step is to reduce $(*)$ to another Riccati equation in which the quadratic term has a coefficient of 1. To do this, the term $f(t)\,x^2$ has to become y^2 for some new variable y, and the obvious choice for such a y is

$$y := fx.$$

Here, and in what follows, we suppress the notation of dependence on t for the sake of brevity. Then $y^2 = f^2x^2$, so to make this term appear we have to multiply $(*)$ through by f. This produces $fx' = (fx)^2 + g(fx) + fh = y^2 + gy + fh$. To deal with the term fx', we can compute

$$y' = (fx)' = f'x + fx',$$

showing that

$$fx' = y' - f'x = y' - (f'/f)(fx) = y' - (f'/f)y.$$

Thus, $(*)$ becomes $y' - (f'/f)y = y^2 + gy + fh$, or after combining like terms,

$$\boxed{y' = y^2 + \big\{g(t) + [f'(t)/f(t)]\big\}\,y + \{f(t)h(t)\}} \tag{$**$}$$

This is a Riccati equation in y, where the leading coefficient function is 1.

Next, we want to introduce some variable u related to y which will make the quadratic term y^2 cancel out from both sides of $(**)$. The left-hand side of $(**)$ is y', so we need to relate *differentiation* to *squaring*. This suggests looking at the quotient rule, where the denominator gets squared. Writing $y = v/u$, where for the moment v and u are independent, the quotient rule gives $y' = (uv' - u'v)/u^2$. At the same time, $y^2 = v^2/u^2$. Thus, y^2 will cancel with one of the terms of y' (namely, the second term) *if* it so happens that $-u'v/u^2 = v^2/u^2$. This involves having $v = -u'$, and thus $y = v/u = -u'/u$. So we have discovered the promising substitution

$$y = -u'/u = -(\ln u)', \qquad \text{i.e.,} \qquad \boxed{u := \exp\left[-\int_{t_0}^{t} y(\tau)\,\mathrm{d}\tau\right] = \exp\left[-\int_{t_0}^{t} f(\tau)\,x(\tau)\,\mathrm{d}\tau\right].} \tag{$\dagger$}$$

Once v is set equal to $-u'$, and once y^2 is subtracted away from both sides of $(**)$, all that remains on the left-hand side of $(**)$ is the term in our expression for y' that did *not* cancel out, namely $uv'/u^2 = u(-u')'/u^2 = -u''/u$. So $(**)$ becomes

$$-u''/u = [g + (f'/f)](-u'/u) + (fh).$$

Multiply this through by u, and then transpose the term u'' over to the other side, to get

$$\boxed{u'' - \big\{g(t) + [f'(t)/f(t)]\big\}\,u' + \{f(t)h(t)\}\,u = 0.} \tag{$\dagger\dagger$}$$

This is a second-order linear homogeneous equation in u (with nonconstant coefficients). From a solution $u(t)$ of $(\dagger\dagger)$ we can go back to $y(t)$ via $y(t) = -u'(t)/u(t)$, and then back to $x(t)$ via

$$\boxed{x(t) = \frac{y(t)}{f(t)} = -\frac{u'(t)}{f(t)\,u(t)}.}$$

This is a solution of $(*)$. In this sense, $(*)$ and $(\dagger\dagger)$ are equivalent.

Next, suppose f, g, h are analytic at $t = 0$, with the minimum of their radii of convergence being $R > 0$. Write

$$f(t) = \sum_{n\geq 0} \alpha_n\, t^n, \qquad g(t) = \sum_{n\geq 0} \beta_n\, t^n, \qquad h(t) = \sum_{n\geq 0} \gamma_n\, t^n.$$

We search for a series solution of $(*)$ in the form

$$x = \sum_{n\geq 0} c_n\, t^n.$$

By standard results on power series already studied in class, we have

$$x' = \sum_{n\geq 0} (n+1)c_{n+1}\, t^n$$

(after re-indexing the term-by-term derivative of the series for x), and

$$x^2 = \sum_{n\geq 0} \left(\sum_{i=0}^{n} c_i\, c_{n-i} \right) t^n.$$

Also,

$$f(t)\, x^2 = \sum_{n\geq 0} \left[\sum_{j=0}^{n} \alpha_{n-j} \left(\sum_{i=0}^{j} c_i\, c_{j-i} \right) \right] t^n,$$

and

$$g(t)\, x = \sum_{n\geq 0} \left[\sum_{j=0}^{n} \beta_{n-j}\, c_j \right] t^n.$$

Setting equal the coefficient of t^n in x' and the coefficient of t^n in $f(t)\, x^2 + g(t)\, x + h(t)$, we find that for all $n \geq 0$,

$$(n+1)\, c_{n+1} = \sum_{j=0}^{n} \left[\alpha_{n-j} \left(\sum_{i=0}^{j} c_i\, c_{j-i} \right) + \beta_{n-j}\, c_j \right] + \gamma_n.$$

This gives a recursion relation for the coefficients $(c_n)_{n\geq 0}$, namely

$$\boxed{c_{n+1} = \frac{1}{n+1} \sum_{j=0}^{n} \left[\alpha_{n-j} \left(\sum_{i=0}^{j} c_i\, c_{j-i} \right) + \beta_{n-j}\, c_j \right] + \frac{\gamma_n}{n+1} \qquad (n \geq 0).} \tag{‡}$$

Notice that since $0 \leq i \leq j \leq n$, only $c_0, c_1, \ldots, c_n$ appear on the right-hand side. So the recursion is successful in defining a sequence. The relation (‡) determines $c_1, c_2, c_3, \ldots$ in terms of c_0, but does not constrain c_0 at all. This means that the base case of the recursion is

$$\boxed{c_0 = C,} \tag{‡‡}$$

where C is a parameter. Together, (‡) and (‡‡) completely determine the unique solution $x(t)$ such that $x(0) = C$.

The convergence of the series for x is a direct consequence of the theory laid out in the last subsection. For the Riccati equation is clearly an analytic ODE whenever the coefficient functions f, g, h are analytic functions of t. After all, the term $f(t)\, x^2$ involves terms of the form $\alpha_m\, t^m x^2$; and these may be viewed as having the form $a_{m,n}\, t^m x^n$, provided we define $a_{m,n} := \alpha_m$ when $n = 2$ and $a_{m,n} := 0$ when $n \neq 2$. The terms $g(t)\, x$ and $h(t) = h(t)\, x^0$ may be viewed likewise. Thus, the general theory of convergence for series solutions of an analytic ODE applies here.

The next thing we'd like to see is that *if* a particular solution of $(*)$ is known, then the general solution can be found by an integration. Let the particular solution in question be called $x_0(t)$. It is then reasonable to guess that for any other solution $x(t)$, the *difference*

$$y(t) := x(t) - x_0(t)$$

satisfies some other (hopefully simpler) differential equation. Thus, we wish to substitute $x = x_0 + y$ into $(*)$. To do this we need to compute that $x' = x_0' + y'$, and also that $x^2 = x_0^2 + 2x_0 y + y^2$. Thus, $(*)$ becomes

$$x_0' + y' = f\cdot(x_0^2 + 2x_0 y + y^2) + g\cdot(x_0 + y) + h. \tag{\#}$$

Since x_0 satisfies $(*)$, we have $x_0' = f x_0^2 + g x_0 + h$. So these terms cancel from (#), are we are left with the equation

$$\boxed{y' = [2x_0(t) f(t) + g(t)]\, y + f(t)\, y^2} \tag{\#\#}$$

Here, notice that $x_0(t)$ is simply a known function of t, so the entire function in the square brackets is a known function of t. This is a very nice result, since (##) is just a Bernoulli equation with $\alpha = 2$. We already know how to solve it. Once the general solution for y is written down, say as

$$y = \Phi_C(t),$$

then the general solution of the Riccati equation $(*)$ is just

$$\boxed{x = x_0(t) + \Phi_C(t).} \tag{@}$$

While the preceding theorem is remarkable, it is not useful in practical work unless we can find a particular solution somehow. This is generally not possible for arbitrary continuous functions f, g, h, except by means of a series solution, say with $c_0 = 1$. The problem with this is we will not know our series solution beyond a certain term, so the function in square brackets in the Bernoulli equation (##) will not be fully known. However, those special cases in which f, g, h are *rational functions* of t tend to have a disproportionate importance in applications. In such cases, it is often possible to find a polynomial solution of $(*)$.

First, by clearing denominators in $(*)$ — i.e., multiplying through by the least common multiple of the polynomial denominators of f, g, h — we can put the Ricatti equation in the form

$$A(t)\, x' = B(t)\, x^2 + C(t)\, x + D(t), \tag{@@}$$

where A, B, C, D are polynomials, say of degrees a, b, c, d respectively. Let $x = P(t)$ be a polynomial solution (i.e., assume that one exists), of degree p. Note that $x' = P'(t)$ has degree $p - 1$, while $x^2 = [P(t)]^2$ has degree $2p$. Since

$$B(t)[P(t)]^2 + C(t)P(t) + D(t) - A(t)P'(t) \equiv 0,$$

where the left-hand side is a polynomial, the "identity theorem" for polynomials tells us that *all* coefficients of the polynomial $BP^2 + CP + D - AP'$ must be zero. The four individual polynomials in this combination — namely BP^2, CP, D, and $-AP'$ — have degrees

$$b + 2p, \qquad c + p, \qquad d, \qquad a + p - 1,$$

respectively. If one of these degrees were strictly greater than the other three, the leading term of the corresponding polynomial would have no counterbalancing term in any of the other polynomials, with which to cancel, so that the sum $BP^2 + CP + D - AP'$ would have at least one nonzero coefficient. That would be a contradiction. We conclude that of the four degrees displayed above, *at least two must be equal.*

Let's look at all the possibilities in a systematic way. If $b + 2p = c + p$, then $p = c - b$. If $b + 2p = d$, then $p = \frac{1}{2}(d - b)$. If $b + 2p = a + p - 1$, then $p = a - b - 1$. If $c + p = d$, then $p = d - c$. If $c + p = a + p - 1$, then a wrinkle occurs, because we cannot determine the value of p; rather, all we can say is that in this case we have $c = a - 1$. We'll come back to that case momentarily. Moving ahead, if $d = a + p - 1$, then $p = d - a + 1$.

The case when $c = a - 1$ can be analyzed further as follows. If $b + 2p$ is greater than the two equal values $c + p$ and $a + p - 1$, then the highest-degree term comes from BP^2, and must cancel with an equal-degree term in D (since CP and $-AP'$ have too low a degree to help). That is, we must in this case have $b + 2p = d$, or $p = \frac{1}{2}(d - b)$. The same analysis holds if we begin by supposing that d is greater than the two equal values $c + p$ and $a + p - 1$. Otherwise, $d \le c + p$ and $b + 2p \le c + p$, giving the lower bound $d - c \le p$ and the upper bound $p \le c - b$. Even in the case when $p = \frac{1}{2}(d - b)$, this value of p lies between the bounds $d - c$ and $c - b$; in fact, it is their average. We conclude that if $c = a - 1$, we have a finite range of possible values for p, from $d - c$ to $c - b$.

Summarizing, we have discovered that *if* a polynomial solution P exists, then its degree p is strongly constrained as follows. There are two cases, according as c is or is not equal to $a - 1$. Namely,

$$\boxed{c \ne a - 1 \implies p \in \left\{c - b,\ \tfrac{1}{2}(d - b),\ a - b - 1,\ d - c,\ d - a + 1\right\};} \tag{\$}$$

$$\boxed{c = a - 1 \implies d - c \le p \le c - b.} \tag{\$\$}$$

We now describe an algorithm that will *usually* find a polynomial solution P if one exists. The idea is to write

$$P(t) = ct^p + Q(t),$$

where c is a constant to be determined, and $\deg Q \leq p-1$. Thus, c is the leading coefficient of P. Here, p will taken as one of the finitely many possible candidate degrees we have determined above. For each candidate degree, the algorithm is tried. If no polynomial solution is found during this pass of the algorithm, we move to the next candidate degree. The aim of the algorithm is to find (1) the value of c, and (2) a Riccati equation for $Q(t)$. If successful, we can then further write $Q(t) = \hat{c}t^q + R(t)$, where $\deg R \leq q-1$, and run the algorithm once more. This should find (1) the value of $\hat{c}$, and (2) a Riccati equation for $R(t)$. If successful, we do it again, and so on. We caution that it's possible for a polynomial solution to exist, but for this algorithm to fail to determine one of its coefficients. So the technique is imperfect in this regard, though still useful in many circumstances. (Actually, the flaw can be fixed, but to explain the resolution would take us too far afield for present purposes.)

Here's how each loop of the algorithm works. We substitute $x = ct^p + Q$ into (@@). Suppressing dependence on t in the functions $A(t), B(t), \ldots$ for the sake of brevity, we get

$$A(pct^{p-1} + Q') = B(c^2t^{2p} + 2ct^p\,Q + Q^2) + C(ct^p + Q) + D. \tag{\%}$$

Let's write the leading term of A as αt^a, of B as βt^b, of C as γt^c, of D as δt^d, and of Q as κt^q. Then, in equation (%), after A, B, C have been distributed over the various sums by which they are multiplied, the leading terms of the individual polynomial summands are as follows (in order of appearance, from left to right):

$$(\alpha pc)t^{a+p-1},\;\; (\alpha q\kappa)t^{a+q-1},\;\; (\beta c^2)t^{b+2p},\;\; (2\beta c\kappa)t^{b+p+q},\;\; (\beta\kappa^2)t^{b+2q},\;\; (\gamma c)t^{c+p},\;\; (\gamma\kappa)t^{c+q},\;\; \delta t^d.$$

The second of these cannot be the term of highest degree on the left-hand side of (%), since $q < p$, so that $a+q-1 < a+p-1$. Thus, the highest-degree term on the left-hand side is $(\alpha pc)t^{a+p-1}$. In a similar way, the term of highest degree on the right-hand side is either $(\beta c^2)t^{b+2p}$, $(\gamma c)t^{c+p}$, or δt^d — and any two of these, or all three, may be tied for highest. For all the terms involving q in the exponent have lower degree than some other term on the right-hand side whose exponent involves only p. Now to get all the terms of highest degree to cancel, we must take whichever of the terms $-(\alpha pc)t^{a+p-1}$, $(\beta c^2)t^{b+2p}$, $(\gamma c)t^{c+p}$, and δt^d have the maximum degree (which may be as few as two, or as many as all four), combine them, and set the combined coefficient equal to 0. This will always produce either a quadratic or a linear equation for c. But the equation can fail to determine a value of c in two circumstances. First, we may get a quadratic equation for c that has no real roots. Second, our equation may involve only the two terms with coefficients $-\alpha pc$ and γc, giving an equation $(\gamma - \alpha p)c = 0$. If this occurs, and it happens that $\gamma \neq \alpha p$, then the only solution for c would be $c = 0$, which is not useful (since c is supposed to be the *leading* coefficient of P, and should therefore be nonzero). If this situation obtains while $\gamma = \alpha p$, then there are infinitely many possible values of c, but no one value is determined. This turns out to be a fixable flaw, but the details are not worth getting into here. After all, the very specific condition $\gamma = \alpha p$ is not expected to occur often.

Assuming we have been successful in assigning a specific numerical value to c, it is now easy to rearrange (%) into a Riccati equation for Q. To wit, we have

$$\boxed{AQ' = BQ^2 + \big(2ct^p\,B + C\big)Q + \big(D + ct^p\,C + c^2t^{2p}\,B - pct^{p-1}\,A\big)} \tag{\%\%}$$

Here, c is now a known constant, so the functions in the two sets of parentheses are known functions. Obviously this algorithm is going to be quite difficult to carry out by hand for a problem of even moderate complexity, but it is straightforward to program into a computer.

II. An Example

Consider the equation

$$x' = x^2 + (-2t)x + (t^2+1).$$

The coefficient functions are rational functions, and in fact polynomials. We have $A(t) = 1$, hence $a = \deg A = 0$; $B(t) = 1$, hence $b = \deg B = 0$; $C(t) = -2t$, hence $c = \deg C = 1$; and finally $D(t) = t^2+1$, hence $d = \deg D = 2$.

Since

$$c \neq a-1 \qquad (\text{for } c = 1 \text{ while } a-1 = -1),$$

the candidates for degrees are limited to

$$p \in \left\{c-b,\ \tfrac{1}{2}(d-b),\ a-b-1,\ d-c,\ d-a+1\right\} = \left\{1,\ 1,\ -1,\ 1,\ 3\right\}.$$

But -1 is not a possible degree for a polynomial, so in fact

$$p \in \{1, 3\}.$$

Let's try the smaller of these, for simplicity. So we're looking for a solution of the form

$$x = P(t) = ct + \kappa,$$

where $\kappa = \kappa t^0 = Q(t)$, in the notation employed above. Plugging this ansatz into the given equation, we find after combining like terms that

$$(c^2 - 2c + 1)t^2 + 2\kappa(c-1)t + (\kappa^2 + 1 - c) \equiv 0.$$

Our general algorithm tells us to set the coefficient of the highest occurring power of t, namely t^2, equal to 0. This gives $c^2 - 2c + 1 = 0$, which happens to factor as $(c-1)^2 = 0$. So $c = 1$, and all that remains is to see if we can determine κ. Now we can of course write down a Riccati equation for $Q(t) \equiv \kappa$, but in this simple example it's easier to simply update the last displayed equation with the value $c = 1$, getting

$$\kappa^2 = 0.$$

So we must take $\kappa = 0$. Our candidate for a polynomial solution is now

$$x = P(t) = ct + \kappa = 1t + 0 = t.$$

It is easy to check this in the original Riccati equation, and it works.

Now that we have our particular solution $x_0(t) = t$, we set $x = x_0 + y = t + y$ in the equation, getting

$$1 + y' = (t^2 + 2ty + y^2) + (-2t)(t + y) + (t^2 + 1),$$

or

$$y' = y^2 + (2t - 2t)y = y^2.$$

This is a Bernoulli equation, and can be solved as such, but it also happens to be separable. We can write it as $y^{-2}\,\mathrm{d}y = \mathrm{d}t$, and integrate to get $-y^{-1} = t - C$ (it doesn't matter if we add or subtract the constant of integration). It follows that $y^{-1} = C - t$, and hence that

$$y = \frac{1}{C - t}.$$

Finally, the general solution of the given Riccati equation is $x = x_0 + y$, that is,

$$x = t + \frac{1}{C - t}.$$

The particular solution $x_0(t) = t$ doesn't fit into the general solution for any value of C, but is a limiting case of the general solution, approached by it as $C \to \pm\infty$ (for each fixed t). Thus, $x_0(t)$ is a singular solution.

III. Problems

1. A Riccati equation in which the coefficient functions f, g, h are *constant functions*, say α, β, γ, is also separable: $x' = \alpha x^2 + \beta x + \gamma = q(x)$. To solve this as a separable equation involves integrating $1/q(x)$, which can be done with partial fraction expansions. However, if the discriminant $\Delta := \beta^2 - 4\alpha\gamma$ is nonnegative, then there is at least one real constant r_0 such that $q(r_0) = 0$, and we see that the constant function $x_0(t) \equiv r_0$ is a solution of the given ODE. (This fits into the general theory of singular solutions for separable ODE's, which are always constant functions.) In this case, we have a particular solution x_0 of our Riccati equation, and we can find the general solution easily (without need of partial fraction expansions).

   Apply these observations to the solution of the equation $x' = x^2 - x - 2$.

2. Solve $x' = t^3x^2 + (-2t^4)x + (t^5 + 1)$. [*Hint.* First try to find a polynomial solution of degree 1.]

§7. Envelopes, Orthogonal Curve Families, and Clairaut Equations

I. Curve Families and their Envelopes

Let $\mathscr{F}$ be a **one-parameter family of curves** in the tx-plane, parametrized by $C \in \mathbb{R}$. For a given value of C, let the unique curve in $\mathscr{F}$ corresponding to this value of C be given by

$$\mathscr{C}_C = \{(t, x) \mid F(t, x, C) = 0\},$$

where F is a function of three arguments, which we will assume is C^2 with respect to all three arguments. That is,

$$F_1 := F_t = \frac{\partial F}{\partial t}, \qquad F_2 := F_x = \frac{\partial F}{\partial x}, \qquad F_3 := F_C = \frac{\partial F}{\partial C}$$

are each C^1 throughout the domain of F. For simplicity, let's assume that the domain of F is all of $\mathbb{R}^3$, i.e., all of txC-space. Thus, for a fixed C, the set of points (t, x) satisfying the equation $F(t, x, C) = 0$ is the curve $\mathscr{C}_C \in \mathscr{F}$.

A smooth curve *not* belonging to $\mathscr{F}$, say

$$\mathscr{E} = \{(t, x) \mid G(t, x) = 0\},$$

with G being C^1 on its domain, is called an **envelope** of $\mathscr{F}$ if at each point $(t_0, x_0) \in \mathscr{E}$, there is precisely one value $C \in \mathbb{R}$ for which $\mathscr{C}_C$ is tangent to $\mathscr{E}$ at (t_0, x_0). This means firstly that $\mathscr{E}$ and $\mathscr{C}_C$ *intersect* at (t_0, x_0), i.e.,

$$G(t_0, x_0) = 0 \quad \text{and} \quad F(t_0, x_0, C) = 0, \tag{$*$}$$

and secondly that $\mathscr{E}$ and $\mathscr{C}_C$ have the *same tangent slope* at (t_0, x_0), i.e.,

$$-\frac{F_1(t_0, x_0, C)}{F_2(t_0, x_0, C)} = -\frac{G_1(t_0, x_0)}{G_2(t_0, x_0)}. \tag{$**$}$$

The two fractions here may be considered equal to ∞ (and hence equal to each other) if both denominators are 0 while both numerators are nonzero.

The reason these particular fractions represent the slopes of the respective tangent lines is because, treating x as an implicit function of t defined by the equation $G(t, x) = 0$, we have the identity $G(t, x(t)) \equiv 0$ on the domain of the implicit function (a small interval); now differentiating with respect to t and using the chain rule gives

$$G_1(t, x(t)) + G_2(t, x(t))\, x'(t) \equiv 0.$$

Plugging in $t = t_0$, and noting that $x(t_0) = x_0$ [since $(t_0, x_0) \in \mathscr{E}$], we have

$$G_1(t_0, x_0) + G_2(t_0, x_0)\, x'(t_0) = 0.$$

Writing $m_{\mathscr{E}}(t_0, x_0) := x'(t_0)$ for the slope of the tangent line of $\mathscr{E}$ at (t_0, x_0), we have

$$m_{\mathscr{E}}(t_0, x_0) = x'(t_0) = -\frac{G_1(t_0, x_0)}{G_2(t_0, x_0)},$$

provided the denominator is nonzero. If the denominator is zero while the numerator is nonzero, then we can make the convention that $m_{\mathscr{E}}(t_0, x_0) = \infty$. We will assume for simplicity that *no* point $(t_0, x_0) \in \mathscr{E}$ satisfies $G_1 = G_2 = 0$. So the slope of the tangent is always unambiguously defined, with ∞ as a possible value (for a vertical line). What $(**)$ says is that

$$m_{\mathscr{C}_C}(t_0, x_0) = m_{\mathscr{E}}(t_0, x_0).$$

To summarize, $\mathscr{E}$ is an envelope of $\mathscr{F}$ if every point (t_0, x_0) of $\mathscr{E}$ determines a unique value of C, which we may denote by $C_0 = C(t_0, x_0)$, such that the curves $\mathscr{E}$ and $\mathscr{C}_{C_0}$ intersect at (t_0, x_0) in a *mutually tangent fashion.*

Even less formally, an envelope of $\mathscr{F}$ is a fixed curve, not in $\mathscr{F}$, to which *every member of $\mathscr{F}$ is tangent.*

Assume an envelope $\mathscr{E}$ exists, and let $C(t, x)$ be its associated rule for assigning to each point $(t, x) \in \mathscr{E}$ a unique value of the parameter C. We call the envelope **smoothly adapted** to $\mathscr{F}$ if the function $C(t, x)$ is C^1 on its domain, and if $C(t,x)$ always changes at a nonzero rate as (t,x) moves along $\mathscr{E}$. The latter means that if $x(t)$ is a function defined implicitly by $G(t,x)=0$, so that the graph of $x(t)$ lies along the curve $\mathscr{E}$, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigl[C(t,x(t))\bigr] = C_1(t,x(t)) + C_2(t,x(t))\,x'(t) = C_1(t,x(t)) - C_2(t,x(t))\,\frac{G_1(t,x(t))}{G_2(t,x(t))} \neq 0. \tag{$***$}$$

This condition ensures that changing the point of tangency along $\mathscr{E}$ will also change the parameter value C corresponding to the unique $\mathscr{F}$-curve tangent to $\mathscr{E}$ at that point, at a *nonzero relative rate*.

Given a smoothly adapted envelope, we can drop the subscripts on t_0, x_0, and $C_0 = C(t_0,x_0)$ (since they are arbitrary, with C_0 determined by t_0 and x_0), writing conditions $(*)$ and $(**)$ in the forms

$$F(t,x,C(t,x)) = 0 \tag{$\dagger$}$$

and

$$F_1(t,x,C(t,x))\cdot G_2(t,x) = F_2(t,x,C(t,x))\cdot G_1(t,x) \tag{$\dagger\dagger$}$$

for all (t,x) with $G(t,x)=0$. Letting $x(t)$ be an implicit function defined by $G(t,x)=0$, we get three identities in t:

$$G(t,x(t)) \equiv 0, \qquad F(t,x(t),C(t,x(t))) \equiv 0,$$

$$F_1(t,x(t),C(t,x(t)))\cdot G_2(t,x(t)) \equiv F_2(t,x(t),C(t,x(t)))\cdot G_1(t,x(t)). \tag{$\ddagger$}$$

Differentiating the first of these identities with respect to t gives

$$G_1(t,x(t)) + G_2(t,x(t))\,x'(t) \equiv 0. \tag{$\ddagger\ddagger$}$$

Doing the same for the second identity gives

$$F_1(t,x(t),C(t,x(t))) + F_2(t,x(t),C(t,x(t)))\,x'(t) + F_3(t,x(t),C(t,x(t)))\,\frac{\mathrm{d}}{\mathrm{d}t}\bigl[C(t,x(t))\bigr] \equiv 0.$$

Solving $(\ddagger\ddagger)$ for $x'(t)$ and plugging the result into the above equation will have the following effect, which we indicate briefly by suppressing arguments: the first two terms above become $F_1 + F_2(-G_1/G_2) = (F_1G_2 - F_2G_1)/G_2 \equiv 0$, using $(\ddagger)$ to identify the numerator as being identically zero. The above-displayed equation therefore reduces to

$$F_3(t,x(t),C(t,x(t)))\,\frac{\mathrm{d}}{\mathrm{d}t}\bigl[C(t,x(t))\bigr] \equiv 0.$$

By our assumption that the envelope is smoothly adapted, the factor $\frac{\mathrm{d}}{\mathrm{d}t}\bigl[C(t,x(t))\bigr]$ is always nonzero, and we can divide it away. All that remains is

$$F_3(t,x(t),C(t,x(t))) \equiv 0. \tag{$\#$}$$

Now look at the identities $(\dagger)$ and $(\#)$. They involve *only* the function F and its partial derivative with respect to the third argument; they make no mention of the function G, nor of any property of the function $C(t,x)$ beyond its output value.

Consideration of $(\dagger)$ and $(\#)$ leads to the following observation. In general, two smooth equations in three unknowns give rise to a *curve* of solution points, since there is one degree of freedom. Thus, considering

$$\boxed{F(t,x,C) = 0 \quad \text{and} \quad F_3(t,x,C) = 0} \tag{$\#\#$}$$

as a system of two equations in three unknowns, we should be able to "eliminate C between them" and get a relation between t and x, which will give us an equation for our envelope $\mathscr{E}$. If this works, it will give us a method for *finding* the envelope (and simultaneously establishing its existence), using nothing but the function F that defines $\mathscr{F}$.

The question is, when does it work? One set of sufficient conditions goes as follows. Assume that

$$\boxed{F_{33} \neq 0, \quad F_1F_{32} \neq F_2F_{31}, \quad \text{and} \quad F_2F_{33} \neq F_3F_{32}} \tag{@}$$

at all common solution points of $F(t,x,C)=0$ and $F_3(t,x,C)=0$. Then there is a smoothly adapted envelope $\mathscr{E}$ for the given curve family $\mathscr{F}$, which can be found by eliminating C between the two equations $(\#\#)$, provided it is algebraically possible to do so. Before discussing the proof, let's look at an example of this elimination process.

Consider the one-parameter family of curves $\mathscr{F}$ defined by

$$(t - C)^2 + x^2 = 1,$$

in which $F(t, x, C) := (t - C)^2 + x^2 - 1$. This is the family of all unit-radius circles in the tx-plane with centers on the t-axis. We compute

$$F_3(t, x, C) = F_C(t, x, C) = 2(t - C)(-1) = -2(t - C).$$

Setting $F_3(t, x, C) = 0$, we find that $C = C(t, x) = t$. Putting this into $F(t, x, C) = 0$, we get

$$(t - t)^2 + x^2 = 1, \quad \text{i.e.,} \quad x^2 = 1.$$

However, the equation $x^2 - 1 = 0$ factors as $(x - 1)(x + 1) = 0$, leading to two disconnected envelopes, namely

$$x = 1 \quad \text{and} \quad x = -1.$$

These are just the two horizontal lines tangent to all unit-radius circles centered on the t-axis, one from above and the other from below. For the upper envelope $x = 1$, we have $G(t, x) = x - 1$. We can see by direct computation that (using a self-explanatory notation)

$$C_1 - C_2 \frac{G_1}{G_2} = \partial_t[t] - \partial_x[t] \frac{\partial_t[x - 1]}{\partial_x[x - 1]} = 1 - 0 \cdot \frac{0}{1} = 1 \neq 0.$$

So the upper envelope is indeed smoothly adapted to the curve family. The lower envelope works similarly. As for the technical conditions (@), we compute in an abbreviated notation that

$$F_{33} = F_{CC} = -2(-1) = 2,$$

$$F_1F_{32} - F_2F_{31} = F_tF_{Cx} - F_xF_{Ct} = 2(t - C)(0) - (2x)(-2) = 4x,$$
$$F_2F_{33} - F_3F_{32} = F_xF_{CC} - F_CF_{Cx} = (2x)(2) - [-2(t - C)](0) = 4x.$$

None of these is zero on the set $\{x = \pm 1\}$, which is the common solution-set of the equations $F = 0$ and $F_3 = 0$. So our technical conditions (@) are satisfied.

Now let's prove the sufficiency of (@). Since $F_{33} \neq 0$ at the relevant points, the Implicit Function Theorem lets us solve the "second equation" $F_3(t, x, C) = 0$ for C as a function of t and x, at least locally, near a point in the common solution-set of $F = 0$ and $F_3 = 0$. Let this local implicit function be denoted by $C = C(t, x)$. Plugging this into the "first equation" $F(t, x, C) = 0$, we get $F(t, x, C(t, x)) = 0$. Near the relevant points, we claim that this equation can be solved locally for x as an implicit function of t. Notice that

$$\partial_x[F(t, x, C(t, x))] = F_x(t, x, C(t, x)) + F_C(t, x, C(t, x))C_x(t, x),$$

or in an abbreviated notation, $F_2 + F_3C_2$. However, by implicit differentiation with respect to x in the identity

$$F_3(t, x, C(t, x)) \equiv 0$$

(which is the equation that defined the function $C(t, x)$ in the first place), we find that

$$F_{32} + F_{33}C_2 \equiv 0.$$

Solving for C_2, we get $C_2 = -F_{32}/F_{33}$. We can plug this into our expression for the x-derivative of $F(t, x, C(t, x))$ to get

$$F_2 + F_3C_2 = \frac{F_2F_{33} - F_3F_{32}}{F_{33}}.$$

By the assumption (@), both the numerator and the denominator appearing here are nonzero on the relevant set of points. So $\partial_x[F(t, x, C(t, x))] \neq 0$ at such points, and the Implicit Function Theorem therefore determines x as a local function of t. Call this local function $x(t)$. Thus, by definition,

$$F(t, x(t), C(t, x(t))) \equiv 0.$$

To say this another way, define a new function

$$G(t,x) := F(t,x,C(t,x));$$

then we have the identity

$$G(t,x(t)) \equiv 0.$$

Notice however that G is defined only locally, near a common solution point of $F=0$ and $F_3=0$; it is not globally defined. On the other hand, there is such a G near *any* point in the common solution-set, and for each such G, the graph of $G(t,x)=0$ is contained within the common solution-set. These graphs overlap, and together they make up a smooth curve (smooth at each point, since each G is smooth), which comprises the desired envelope $\mathscr{E}$. This one envelope — which is a connected curve, by the overlapping of the local curves — may not cover the entire common solution-set, but several disjoint envelopes may do so. [In the example above, the common solution-set was covered by two disjoint envelopes, $x=1$ and $x=-1$.] A single globally defined function G can be found in the special case when there is an algebraic way to eliminate C between the two equations $F(t,x,C)=0$ and $F_3(t,x,C)=0$ (as in the example). Finally, we must only check that the envelope $\mathscr{E}$ is smoothly adapted. This involves checking that $C_1 - C_2(G_1/G_2) \neq 0$ at the relevant points. But we have the equations $C_1 = -F_{31}/F_{33}$ and $C_2 = -F_{32}/F_{33}$, by implicit differentiation in $F_3(t,x,C(t,x))=0$. Also, by implicit differentiation in the definition of G, we have $G_1 = F_1 + F_3C_1$ and $G_2 = F_2 + F_3C_2$. Putting everything together, some intense algebra with fractions (try it) shows that our middle assumption in (@), namely $F_1F_{32} \neq F_2F_{31}$, implies $C_1 - C_2(G_1/G_2) \neq 0$.

II. Differential Equation of a Curve Family; Singular Solutions

Let $\mathscr{F}$ be a one-parameter family of curves in the tx-plane, defined by a function $F(t,x,C)$ satisfying the assumptions of the preceding subsection. We can often derive a first-order ODE of which $\mathscr{F}$ is the general solution. That is, instead of starting with a first-order ODE and solving it to get a one-parameter family of solution curves, we start with a one-parameter family of curves and find a first-order ODE that each one of its member curves satisfies.

For any fixed C, say C_0, the curve $\mathscr{C}_{C_0}$ is defined as the solution-set in the tx-plane of

$$F(t,x,C_0) = 0. \tag{$*$}$$

Given a local function $x(t)$ defined implicitly by $F(t,x,C_0)=0$, so that $F(t,x(t),C_0)\equiv 0$, we can differentiate with respect to t to get

$$F_1(t,x(t),C_0) + F_2(t,x(t),C_0)\,x'(t) \equiv 0. \tag{$**$}$$

Note that there is no term involving a derivative with respect to C, since C_0 is just a constant. If C_0 can be eliminated between ($*$) and ($**$), then the result will be a relation between t, $x(t)$, and $x'(t)$. That is, we'll get a relation

$$H(t,x(t),x'(t)) \equiv 0. \tag{$\dagger$}$$

This shows that $x(t)$ is a solution of the first-order ODE

$$\boxed{H(t,x,x') = 0.} \tag{$\dagger\dagger$}$$

The algebraic nature of the functions involved must be favorable in order for this elimination process to have a chance of success. However, even under less favorable conditions, all this means is that the function H involved in the ODE ($\dagger\dagger$) is not explicitly known to us. It still exists, at least locally, by arguments similar to those used in the preceding subsection. On the other hand, since a differential equation of a completely unknown form is not particularly useful, we will just assume that the algebra is simple enough that H can be found explicitly.

For example, consider the one-parameter family of unit-radius circles centered on the t-axis, whose envelopes we found earlier. For any particular value C_0 of the parameter, the corresponding circle is

$$(t-C_0)^2 + x^2 = 1.$$

Replacing x with $x(t)$ gives the identity

$$(t-C_0)^2 + x(t)^2 \equiv 1.$$

Differentiating with respect to t, we get

$$2(t-C_0) + 2x(t)x'(t) \equiv 0.$$

Solve this last equation for C_0 to get $C_0 = t + x(t)x'(t)$, and plug this into the original equation to get

$$\bigl(t - \{t + x(t)x'(t)\}\bigr)^2 + x(t)^2 \equiv 1,$$

or after simplifying,

$$x(t)^2\,\{x'(t)^2 + 1\} \equiv 1.$$

That is, $x(t)$ satisfies the first-order ODE

$$x^2[(x')^2 + 1] = 1.$$

Since $x = 0$ is not possible in this equation (it would involve $0 = 1$), divide through by x^2 and rearrange to get

$$(x')^2 = \frac{1}{x^2} - 1 = \frac{1 - x^2}{x^2}.$$

Taking square roots gives two separable equations,

$$x' = \pm\frac{\sqrt{1 - x^2}}{x},$$

in other words, *under the temporary assumption* that $x^2 \neq 1$,

$$\frac{x}{\sqrt{1 - x^2}}\,\mathrm{d}x = \pm\,\mathrm{d}t.$$

Integrating both sides (with the aid of the substitution $u := 1 - x^2$) gives $\sqrt{1 - x^2} = C \pm t$, or $1 - x^2 = (C \pm t)^2$, or $(t \pm C)^2 + x^2 = 1$. So we recover the family $\mathscr{F}$ as the general solution.

Notice however that in separating the variables, we had to make the assumption that $x^2 \neq 1$. We now know the significance of this assumption. When $x^2 = 1$, we are on one of the envelopes, $x = 1$ or $x = -1$. And, importantly, *these also turn out to be solutions of the equation* $x^2[(x')^2 + 1] = 1$. We can see it by inspection: if $x \equiv \pm 1$, then $x' \equiv 0$, and the ODE says $(\pm 1)^2[0 + 1] = 1$, which is certainly true.

The lesson here is that whenever the general solution of a first-order ODE has an envelope, this envelope will be a **singular solution** of the ODE. Thus, we should be in the habit of checking the general solution for envelopes.

The fact that an envelope $G(t, x) = 0$ is also a solution of the differential equation $H(t, x, x') = 0$ is simple to prove. Let $x(t)$ be a function defined implicitly by $F(t, x, C_0) = 0$ for some fixed value C_0, say in some interval centered at t_0, and let $y(t)$ be a function defined implicitly by the envelope equation, $G(t, y) = 0$, also in some interval centered at t_0. Since the graph of $x(t)$ lies along the curve $\mathscr{C}_{C_0}$, and since every curve in $\mathscr{F}$ is a solution of the ODE $H(t, x, x') = 0$, we have $H(t, x(t), x'(t)) \equiv 0$ for all t sufficiently close to t_0. In particular, $H(t_0, x(t_0), x'(t_0)) = 0$. But the functions $x(t)$ and $y(t)$, whose graphs lie along $\mathscr{C}_{C_0}$ and $\mathscr{E}$ respectively near their point of tangency (t_0, x_0), must have "first-order contact" (intersection plus mutual tangency) at t_0, since $\mathscr{C}_{C_0}$ and $\mathscr{E}$ have first-order contact there. That is,

$$x(t_0) = y(t_0)\ [= x_0] \qquad \text{and} \qquad x'(t_0) = y'(t_0).$$

Since $H(t_0, x(t_0), x'(t_0)) = 0$, we therefore also have

$$H(t_0, y(t_0), y'(t_0)) = 0.$$

Had some other point t_1 been chosen as the center of the small intervals, instead of t_0, the same reasoning would have established that

$$H(t_1, z(t_1), z'(t_1)) = 0,$$

where $z(t)$ is implicitly defined by $G(t, z) = 0$ near t_1. However, if t_0 and t_1 are sufficiently close together that the domains of $y(t)$ and $z(t)$ overlap, with t_1 in the intersection, then the uniqueness clause of the Implicit Function Theorem forces $z(t) \equiv y(t)$ on the overlap interval, and hence $z(t_1) = y(t_1)$ and $z'(t_1) = y'(t_1)$. So in fact,

$$H(t_1, y(t_1), y'(t_1)) = 0.$$

By this reasoning, we can show that for *any* value of t in the domain of the local function $y(t)$, we have

$$H(t, y(t), y'(t)) = 0.$$

Thus, $y(t)$ is a solution of the ODE. Since the relation $G(t, y) = 0$ implicitly defines all such local functions $y(t)$, we see that the envelope $\mathscr{E} = \{(t, x) \mid G(t, x) = 0\}$ is a solution of the ODE. This is what we wished to prove.

III. Clairaut Equations

A special kind of one-parameter family is obtained if we suppose that all the member curves are straight lines. Let's call these **line families**.

Suppose that each possible value of the slope corresponds to *at most one* line in the line family $\mathscr{F}$. Another way to say this is that no two distinct lines in $\mathscr{F}$ are parallel. Rather, each line has its own unique slope, possessed by no other line in the family. In this case, we can use the slope as a parameter. The x-intercept (this is the tx-plane, remember) cannot be free to vary independently of the slope, since then we would have a two-parameter family of lines, rather than a one-parameter family. Thus, for a given slope C, the x-intercept must be determined by the value of C, say as $f(C)$, where f is a certain function. Thus, the equation defining the member curves of $\mathscr{F}$ is

$$x = Ct + f(C). \tag{$*$}$$

It's easy to derive the ODE of $\mathscr{F}$. Differentiating $(*)$ with respect to t, where x is viewed as an implicit function of t,

$$x' = C.$$

The term $f(C)$ is just a constant (for a fixed C), so its t-derivative is 0. Putting this result for C into $(*)$, we get

$$x = x't + f(x'). \tag{$**$}$$

Equation $(**)$ is called the **Clairaut equation**. It is typically written in the form

$$x - tx' = f(x').$$

The general solution of this equation is simply

$$\boxed{x - Ct = f(C),}$$

which we can write down immediately, with no computation at all.

We can extend the theory slightly to the **generalized Clairaut equation**

$$H(x - tx',\, x') = 0, \tag{$\dagger$}$$

where H is a function for which $H_1 \neq 0$ always (this is the partial derivative of H with respect to its first argument). The Implicit Function Theorem guarantees that $(\dagger)$ can be put into the form $(**)$ at least locally, but with different local functions f near different combinations of values for $x - tx'$ and x'. On the other hand, since for each local function f the general solution of $(**)$ is a line family, the general solution of $(\dagger)$ is the union of many line families (possibly uncountably many), and is therefore also a line family. Now the general solution for any given local function f is $x - Ct = f(C)$. But this is equivalent to $H(x - Ct,\, C) = 0$, by the way f was determined. This last equation no longer refers to the (unknown) implicit function f. So we can write it down immediately, based solely on our knowledge of H. Notice that for each fixed value of C, the equation $H(x - Ct,\, C) = 0$ is a relation between t and x, and determines a curve in the tx-plane (which we already know to be a line). Thus, we have the general solution of the generalized Clairaut equation $(\dagger)$:

$$\boxed{H(x - Ct,\, C) = 0.} \tag{$\dagger\dagger$}$$

As an example, consider the equation

$$(x - tx')^2 + (x')^2 - 1 = 0.$$

The left-hand side is a function of $x - tx'$ and x', so this is a generalized Clairaut equation. (The hardest part of this is being on guard to *recognize* such an equation in the first place.) So the solution is obtained immediately, as

$$(x - Ct)^2 + C^2 - 1 = 0,$$

or after some simplification,

$$x = Ct \pm \sqrt{1 - C^2}.$$

Now what about singular solutions? We can check for envelopes, as follows. Take the C-derivative of the equation $(x - Ct)^2 + C^2 - 1 = 0$; it gives

$$2(x - Ct)(-t) + 2C = 0.$$

This simplifies to $t(x - Ct) = C$, or after multiplying out,

$$C = \frac{tx}{1+t^2}.$$

Putting this expression for C into the general solution, we get

$$\left(x - \frac{t^2 x}{1+t^2}\right)^2 + \left(\frac{tx}{1+t^2}\right)^2 - 1 = 0.$$

Here, x^2 factors out, and when everything is simplified as much as possible, we find (try it):

$$x^2 - t^2 = 1.$$

This is the equation of a hyperbola. Thus, the two branches of this hyperbola are envelopes for our line family, and they also represent singular solutions of the given differential equation.

More generally, to find possible envelopes of the generalized Clairaut equation, we must try to eliminate C between

$$\boxed{H(x - Ct,\, C) = 0 \quad \text{and} \quad -\, t\, H_1(x - Ct,\, C) + H_2(x - Ct,\, C) = 0,} \tag{‡}$$

the latter equation being the C-derivative of the former. This may be difficult or impossible, but there is one nice case where we can find an envelope explicitly. Say we have an ordinary Clairaut equation, $x - tx' = f(x')$, with f known. Also, suppose f' exists and is one-to-one, with a known inverse function. Then the C-derivative of the general solution $x - Ct = f(C)$ is simply $-t = f'(C)$, and we can solve this for C by applying $(f')^{-1}$ to both sides:

$$C = (f')^{-1}(-t).$$

Substituting this into $x = Ct + f(C)$, we get the equation of an envelope, namely

$$\boxed{x = t \cdot (f')^{-1}(-t) + f\big((f')^{-1}(-t)\big).} \tag{‡‡}$$

IV. Orthogonal Curve Families

A nice geometric problem goes as follows. Given a one-parameter curve family $\mathscr{F}$, find a second such family, say $\mathscr{G}$, such that for *any* curve $\mathscr{C}_C \in \mathscr{F}$ and for *any* curve $\mathscr{K}_K \in \mathscr{G}$, the two curves intersect at right angles. We write this condition as $\mathscr{C}_C \perp \mathscr{K}_K$. We also write $\mathscr{F} \perp \mathscr{G}$, and say that the two curve families are **orthogonal complements**. Curves are said to intersect at right angles if (1) they intersect, and (2) their tangent lines at this point of intersection are perpendicular.

Of course, we expect that the problem can't always be solved, either because there is no orthogonal complement, or because finding it is intractable. However, we can give a general procedure for characterizing the orthogonal complement by a differential equation, whenever it exists. Solving this differential equation may not be feasible, but numerical methods can be used to get approximate plots of the orthogonal system of curves.

The idea is simple. Let the original family $\mathscr{F}$ be the general solution of the first-order ODE

$$H(t, x, x') = 0. \tag{$*$}$$

Then consider the related ODE given by

$$\boxed{H(t, x, -1/x') = 0.} \tag{$**$}$$

If $\mathscr{G}$ is the general solution of $(**)$, then we claim that

$$\mathscr{F} \perp \mathscr{G}.$$

At any point (t_0, x_0), a solution $x(t)$ of $(*)$ satisfying $x(t_0) = x_0$ will have slope $x'(t_0)$, while a solution $y(t)$ of $(**)$ satisfying $y(t_0) = x_0$ will have slope $y'(t_0)$. We now have $H(t_0, x_0, x'(t_0)) = 0$, from $(*)$, and $H(t_0, x_0, -1/y'(t_0)) = 0$, from $(**)$ [using $x(t_0) = x_0 = y(t_0)$]. Assuming H has the property that $H_3 \neq 0$, so that its third argument is locally a function of its first two arguments, these equations force us to conclude that $x'(t_0) = h(t_0, x_0) = -1/y'(t_0)$, where

h is the local function giving the third argument in terms of the first two. Thus, the two slopes $x'(t_0)$ and $y'(t_0)$ are negative reciprocals. This means that the graphs of $x(t)$ and $y(t)$ have perpendicular tangent lines at (t_0, x_0). This is what's meant by orthogonal curves.

For an example, let $\mathscr{F}$ be the family of all parabolas with vertex at the origin in the tx-plane, namely

$$x = Ct^2.$$

Differentiating this with respect to t gives $x' = 2Ct$, or $C = x'/(2t)$. Putting this into the original one-parameter family, we get $x = [x'/(2t)]\, t^2$, or

$$2x = tx'.$$

This is the differential equation of $\mathscr{F}$. To find the differential equation of the orthogonal complement $\mathscr{G}$, replace x' by $-1/x'$. This results in

$$2x = -\frac{t}{x'},$$

or

$$2xx' = -t.$$

Separating the variables, we get $2x\,\mathrm{d}x = -t\,\mathrm{d}t$. This can be integrated immediately, and becomes $x^2 = C - \frac{1}{2}t^2$. Rearranging, we have the curve family

$$\mathscr{G}: \;\; t^2 + 2x^2 = C,$$

which we recognize as a family of similar *ellipses*, all centered at the origin.

V. Problems

1. Consider the general solution $\mathscr{F}$ of a first-order *linear* equation, $x' + f(t)\,x = g(t)$. Prove that $\mathscr{F}$ cannot have an envelope. Thus, singular solutions are a phenomenon exclusive to nonlinear equations. [*Hint.* The general solution will have the form $x = C\phi(t) + \psi(t)$, where ϕ is the reciprocal of an integrating factor. Compare this equation with its own C-derivative.]
2. Find the differential equation of the curve family $x = C + e^{t-C}$. Is there an envelope? If so, find it.
3. Find the general solution and envelopes (if any): $\sqrt{x - tx'} + \sqrt{x'} = 1$.
4. Consider the family $\mathscr{F}$ of all circles through the origin with center on the t-axis, namely $(t - C)^2 + x^2 = C^2$, or $t^2 + x^2 = 2Ct$. Find the orthogonal complement $\mathscr{G}$ of $\mathscr{F}$, and identify its members geometrically.

§8. Slope Function Dependent on a Linear Combination of the Variables

I. Theory

In this brief subsection, we show how to solve the equation

$$x' = f(at + bx + c), \tag{$*$}$$

where f is a continuous function, and where a, b, c are real constants. Since x' is isolated, the right-hand side has the interpretation of a "slope function", and this particular slope function is dependent on a linear combination of the two variables t and x, namely

$$at + bx.$$

The term c may be regarded as part of f, by replacing f with $\hat{f}(v) := f(v + c)$, but let's not bother with that.

If $a = b = 0$, then $(*)$ reduces to $x' = f(c)$ (a constant), and the solution is $x = f(c)\,t + C$. Thus, we assume henceforth that $a^2 + b^2 > 0$. If $b = 0$ and $a \neq 0$, then $(*)$ says $x' = f(at + c)$, and the solution is

$$x = C + \int_{t_0}^{t} f(a\tau + c)\,\mathrm{d}\tau.$$

If $a = 0$ and $b \neq 0$, then $(*)$ becomes $x' = f(bx + c)$, which is separable. The solution is defined implicitly by

$$t = C + \int_{x_0}^{x} \frac{1}{f(b\chi + c)}\,\mathrm{d}\chi.$$

The only new cases involve $a \neq 0$ and $b \neq 0$.

Let's make the substitution

$$\boxed{u := at + bx + c.} \tag{$**$}$$

Then, differentiating with respect to t (where x is regarded as a function of t) we have

$$u' = a + bx'.$$

By $(*)$, we can replace x' here with $f(at + bx + c)$, that is, with $f(u)$. We get

$$\boxed{u' = a + b\,f(u).} \tag{$\dagger$}$$

We see that $(\dagger)$ is separable, and its solutions are defined implicitly by the relation

$$t = C + \int_{u_0}^{u} \frac{1}{a + b\,f(\mu)}\,\mathrm{d}\mu.$$

To go back to the original variables, we simply replace u by $at + bx + c$. Thus, we get the relation

$$\boxed{t = C + \int_{u_0}^{at+bx+c} \frac{1}{a + b\,f(\mu)}\,\mathrm{d}\mu.} \tag{$\dagger\dagger$}$$

To clarify the meaning of $(\dagger\dagger)$, we antidifferentiate $1/\{a + b\,f(\mu)\}$ with respect to μ, then evaluate the resulting expression between $\mu = u_0$ (some convenient value in the domain of f) and $\mu = at + bx + c$. The final equation thus obtained will be a relation between t and x containing the parameter C. This is the general solution of $(*)$.

II. An Example

Consider the equation

$$x' = (8t - x + 1)^3.$$

Here, $f(v) = v^3$, while $a = 8$, $b = -1$, and $c = 1$. Setting $u := 8t - x + 1$, we have $u' = 8 - x'$, or $x' = 8 - u'$. Our equation therefore says $8 - u' = u^3$, or

$$u' = 8 - u^3.$$

We can separate the variables as follows:

$$\frac{1}{u^3 - 8}\,\mathrm{d}u = -\mathrm{d}t.$$

To integrate the left-hand side, note that by a common factoring formula, $u^3 - 8 = (u-2)(u^2 + 2u + 4)$. So we should have a partial fraction expansion

$$\frac{1}{u^3 - 8} = \frac{A}{u-2} + \frac{Bu + C}{u^2 + 2u + 4}.$$

Some easy algebra (clearing denominators and plugging in $u = 2$, $u = 0$, or $u = 1$) gives $A = \frac{1}{12}$, $B = -\frac{1}{12}$, and $C = -\frac{1}{3}$. Thus we have to compute the integral

$$\frac{1}{12} \int \left[\frac{1}{u-2} - \frac{u+4}{u^2 + 2u + 4} \right] \mathrm{d}u.$$

The first term gives $\frac{1}{12} \ln|u - 2|$. Rewrite the second term as

$$-\frac{1}{24} \int \frac{2u + 8}{u^2 + 2u + 4}\,\mathrm{d}u = -\frac{1}{24} \int \frac{2u + 2}{u^2 + 2u + 4}\,\mathrm{d}u - \frac{1}{4} \int \frac{1}{(u+1)^2 + (\sqrt{3})^2}\,\mathrm{d}u.$$

The first of these integrals has numerator equal to the derivative of the denominator, so that one gives a contribution of $-\frac{1}{24} \ln|u^2 + 2u + 4|$. The last integral is recognizable as an arctangent, namely

$$-\frac{1}{4\sqrt{3}} \tan^{-1}\left(\frac{u+1}{\sqrt{3}}\right).$$

In terms of t and u, the solution is

$$\frac{1}{12}\ln|u-2| - \frac{1}{24}\ln|u^2+2u+4| - \frac{1}{4\sqrt{3}}\tan^{-1}\left(\frac{u+1}{\sqrt{3}}\right) = C - t.$$

In terms of t and x, this becomes the rather unwieldy relation

$$\frac{1}{12}\ln|8t-x-1| - \frac{1}{24}\ln|64t^2+x^2-16tx+3| - \frac{1}{4\sqrt{3}}\tan^{-1}\left(\frac{8t-x}{\sqrt{3}}\right) = C - t.$$

III. Problems

1. Solve: $x' = \sqrt{t+x+1}$. [*Hint.* To integrate $1/(1+\sqrt{u})$, put $v := 1+\sqrt{u}$, and show that $\mathrm{d}u = 2(v-1)\,\mathrm{d}v$.]

§9. Homogeneous Equations

I. Theory

A **homogeneous equation** (not to be confused with a homogeneous *linear* equation) is a first-order ODE

$$x' = f(t,x) \tag{$*$}$$

in which the slope function f is **homogeneous of degree zero**, that is, satisfies

$$f(\lambda t, \lambda x) \equiv f(t,x) \quad \text{for all } \lambda > 0. \tag{$**$}$$

More generally, the function f is **homogeneous of degree** α if

$$f(\lambda t, \lambda x) \equiv \lambda^\alpha f(t,x) \quad \text{for all } \lambda > 0. \tag{$\dagger$}$$

Here, α is a fixed real number, called the **degree of homogeneity**. When $\alpha = 0$, we get $(**)$. A simple way to produce functions that are homogeneous of degree zero is by forming the ratio of two functions that are homogeneous of degree α. If g and h are homogeneous of degree α, then for any $\lambda > 0$, the ratio $f := g/h$ satisfies

$$f(\lambda t, \lambda x) = \frac{g(\lambda t, \lambda x)}{h(\lambda t, \lambda x)} = \frac{\lambda^\alpha g(t,x)}{\lambda^\alpha h(t,x)} = \frac{g(t,x)}{h(t,x)} = f(t,x).$$

The following method for solving $(*)$ is due to Euler. We will separately find solutions defined on $t > 0$ and on $t < 0$. Assuming $t > 0$, we can take $\lambda = 1/t > 0$, and we find that

$$x' = f(t,x) = f(\lambda t, \lambda x) = f(1,\, x/t).$$

This suggests making the substitution

$$\boxed{u := x/t, \qquad \text{i.e.,} \qquad x = tu.} \tag{$\dagger$}$$

Differentiating both sides of $x = tu$ with respect to t (and regarding x and u as unknown functions of t), we get

$$x' = u + tu'.$$

Thus, $(*)$ becomes

$$u + tu' = f(1,u).$$

Rearranging, we get the separable equation

$$\boxed{\frac{1}{f(1,u)-u}\,\mathrm{d}u = \frac{1}{t}\,\mathrm{d}t.} \tag{$\dagger\dagger$}$$

Since we've assumed that $t > 0$, we must integrate both sides from some $t_0 > 0$ to some $t > 0$. We get

$$\int_{u(t_0)}^{u(t)} \frac{1}{f(1,\mu)-\mu}\,\mathrm{d}\mu = \int_{t_0}^{t} \frac{1}{\tau}\,\mathrm{d}\tau = \ln t - \ln t_0.$$

The arbitrary parameter here is the value $u(t_0)$. By using some convenient lower limit u_0, we can take the parameter to be an additive constant of integration. We may as well absorb the term $-\ln t_0$ into this constant of integration. Finally, we replace $u(t)$ in this relation by $x(t)/t$, or simply x/t for brevity. This gives the general solution of $(*)$,

$$\boxed{\int_{u_0}^{x/t} \frac{1}{f(1,\mu)-\mu}\,\mathrm{d}\mu = C + \ln t \qquad (t>0)} \tag{‡}$$

To clarify, we antidifferentiate with respect to μ first, then evaluate the result from $\mu = u_0$ to $\mu = x/t$.

To solve in the regime $t < 0$, take $\lambda = -1/t > 0$, so that the equation $(*)$ becomes

$$x' = f(-1,\, -x/t).$$

The substitution $u := x/t$ still works; we get the separable equation

$$\boxed{\frac{1}{f(-1,-u)-u}\,\mathrm{d}u = \frac{1}{t}\,\mathrm{d}t.} \tag{‡‡}$$

This time, we integrate from some $t_0 < 0$ to some $t < 0$. The right-hand side this time integrates to $\ln|t| - \ln|t_0|$. The term $-\ln|t_0|$ will be absorbed into the additive constant of integration. We may also write $\ln|t| = \ln(-t)$. Our general solution of $(*)$ in this case is

$$\boxed{\int_{u_0}^{x/t} \frac{1}{f(-1,-\mu)-\mu}\,\mathrm{d}\mu = C + \ln(-t) \qquad (t<0)} \tag{\#}$$

Note finally that if f happens to be an *even function* of both of its arguments, i.e., if

$$f(-t,-x) \equiv f(t,x),$$

then formulas (‡) and (#) can be combined into a single formula,

$$\int_{u_0}^{x/t} \frac{1}{f(1,\mu)-\mu}\,\mathrm{d}\mu = C + \ln|t|,$$

which now holds for all $t \neq 0$. For $f(-1,-\mu) = f(1,\mu)$ by the evenness assumption, and $|t| = t$ for $t > 0$ while $|t| = -t$ for $t < 0$. This is a commonly encountered situation.

An interesting special case of the homogeneous equation is the **linear fractional equation**

$$x' = f\left(\frac{at+bx+p}{ct+dx+q}\right), \tag{\#\#}$$

where f is continuous and a, b, c, d, p, q are real constants with c, d, q not all zero. This is not actually homogeneous as it stands, but in most cases it can be reduced to a homogeneous equation by a suitable change of variables. The exception cases can be handled by other methods, as we'll see.

The first thing to notice is that when $c = d = 0$, so that $q \neq 0$, we have a slope function dependent on a linear combination of the variables: $(a/q)t + (b/q)x$. We already know how to solve that case. Thus, assume $c^2 + d^2 > 0$. In addition, the cases $a = c = 0$ (when the slope function depends only on x, yielding a separable equation) and $b = d = 0$ (when the slope function depends only on t, yielding to immediate integration) are already understood. So assume $a^2 + c^2 > 0$ and $b^2 + d^2 > 0$.

The main distinction now is between cases where $ad - bc = 0$ and cases where $ad - bc \neq 0$. In the former case, $ad = bc$, we claim that there is some constant r such that $a = rc$ and $b = rd$. To see this requires a bit of casework. First consider what happens if $a = 0$. Then $c \neq 0$, but $bc = ad = 0d = 0$. So we must have $b = 0$. It follows that $r = 0$ works; we have $a = 0 = 0c$ and $b = 0 = 0d$. The other possibility is that $a \neq 0$. If we had $c = 0$, then $d \neq 0$ (since $c^2 + d^2 > 0$). But now $ad = bc = b0 = 0$, a contradiction since $a \neq 0$ and $d \neq 0$. We conclude that $c \neq 0$. Also, if we had $d = 0$, then $b \neq 0$ (since $b^2 + d^2 > 0$). But now $bc = ad = a0 = 0$, a contradiction since $b \neq 0$ and

$c \neq 0$. Thus, $d \neq 0$. Finally, take the equation $ad = bc$ and divide both sides by the nonzero numbers c and d. We get $a/c = b/d$. Writing $r := a/c = b/d$, we find that $a = rc$ and $b = rd$. With this in mind, (##) becomes

$$x' = f\left(\frac{rc\,t + rd\,x + p}{ct + dx + q}\right) = f\left(\frac{r(ct + dx + q) + p - qr}{ct + dx + q}\right) = f\left(r + \frac{p - qr}{ct + dx + q}\right).$$

Defining a new function g by

$$g(v) := f\big(r + (p - qr)/v\big),$$

our differential equation becomes

$$x' = g(ct + dx + q).$$

This is once again the case of a slope function dependent on a linear combination of the variables.

The substantially new case is when $ad - bc \neq 0$. Here we will make a change of variables to try to reduce (##) to a homogeneous equation. The idea is to define a shifted time variable $T := t - \alpha$ and a shifted space variable $X := x - \beta$. Plugging $t = T + \alpha$ and $x = X + \beta$ into the linear fractional expression in (##), we get

$$\frac{at + bx + p}{ct + dx + q} = \frac{a(T + \alpha) + b(X + \beta) + p}{c(T + \alpha) + d(X + \beta) + q} = \frac{aT + bX + (a\alpha + b\beta + p)}{cT + dX + (c\alpha + d\beta + q)}.$$

We'll now try to choose α and β so that

$$\boxed{a\alpha + b\beta = -p \qquad \text{and} \qquad c\alpha + d\beta = -q.} \tag{@}$$

With these choices, our linear fractional expression will become simply

$$\frac{aT + bX}{cT + dX},$$

which resembles the original linear fractional expression, except without the constant terms p and q. In fact, it is now a ratio of two functions that are homogeneous of degree 1, so it is homogeneous of degree 0.

Can we choose α and β so as to make (@) true? These are two linear equations in two unknowns, which will have a unique solution if the determinant of its matrix of coefficients is nonzero. But this determinant is precisely $ad - bc$, which we are assuming is nonzero. The exact values of α and β can be found by easy algebra, and are

$$\boxed{\alpha = \frac{bq - pd}{ad - bc} \qquad \text{and} \qquad \beta = \frac{pc - aq}{ad - bc}.} \tag{@@}$$

What about the derivative term, x'? By the chain rule,

$$x' = \frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\mathrm{d}(X + \beta)}{\mathrm{d}t} = \frac{\mathrm{d}X}{\mathrm{d}t} = \frac{\mathrm{d}X}{\mathrm{d}T}\,\frac{\mathrm{d}T}{\mathrm{d}t} = \frac{\mathrm{d}X}{\mathrm{d}T}\,\frac{\mathrm{d}(t - \alpha)}{\mathrm{d}t} = \frac{\mathrm{d}X}{\mathrm{d}T} =: \dot{X},$$

where the over-dot is a shorthand notation for differentiation with respect to T. All in all, our original equation (##) now becomes

$$\boxed{\dot{X} = f\left(\frac{aT + bX}{cT + dX}\right).} \tag{\%}$$

This is a homogeneous first-order ODE. By Euler's method, it can be turned into the separable equation

$$T\,\dot{U} = f\left(\frac{a + bU}{c + dU}\right) - U,$$

where $U := X/T$. The solution proceeds in the usual way from here. Once the general solution is written out as a relation between T and X, we simply replace T with $t - \alpha$ and X with $x - \beta$, where α and β are the specific constants (@@). This will produce a relation between t and x, which is the general solution of (##).

II. An Example

Consider

$$x' = \frac{2tx}{t^2 + x^2}.$$

Notice that the slope function is homogeneous of degree zero, for

$$\frac{2(\lambda t)(\lambda x)}{(\lambda t)^2 + (\lambda x)^2} = \frac{\lambda^2(2tx)}{\lambda^2(t^2 + x^2)} = \frac{2tx}{t^2 + x^2}.$$

Dividing the fraction top and bottom by t^2 puts the equation in the form

$$x' = \frac{2(x/t)}{1 + (x/t)^2}.$$

Putting $u := x/t$, so that $x = tu$, and $x' = u + tu'$, the ODE becomes

$$u + tu' = \frac{2u}{1 + u^2}.$$

Subtracting u and combining, we find

$$tu' = \frac{2u}{1 + u^2} - \frac{u + u^3}{1 + u^2} = \frac{u - u^3}{1 + u^2} = \frac{u(1 - u)(1 + u)}{1 + u^2}.$$

Separating the variables yields

$$\frac{1 + u^2}{u(u - 1)(u + 1)}\,\mathrm{d}u = -\frac{1}{t}\,\mathrm{d}t. \qquad (\%\%)$$

[The minus sign is due to our changing $1 - u$ to $u - 1$ in the denominator.] To go further we need the partial fraction expansion

$$\frac{1 + u^2}{u(u - 1)(1 + u)} = \frac{A}{u} + \frac{B}{u - 1} + \frac{C}{u + 1}.$$

Easy algebra gives $A = -1$, $B = 1$, and $C = 1$. Thus, the integral of the left-hand side of (%%) is

$$-\ln|u| + \ln|u - 1| + \ln|u + 1| = \ln\left|\frac{u^2 - 1}{u}\right|.$$

On the right hand side of (%%), we get

$$-\ln|t| + C.$$

Setting these equal, and exponentiating both sides, we find that

$$\left|\frac{u^2 - 1}{u}\right| = e^C|t|^{-1}.$$

Here, e^C is an arbitrary positive constant. If we drop the absolute value signs, this will become an arbitrary nonzero constant, which we will still denote by C. Thus,

$$\frac{u^2 - 1}{u} = \frac{C}{t}.$$

This says $t(u^2 - 1) = Cu$. We may as well multiply through by t to get $t^2(u^2 - 1) = C(ut)$, or $(tu)^2 - t^2 = C(ut)$. Finally, recall that $x = ut$. So our general solution is

$$x^2 - t^2 = Cx.$$

More explicitly, we can use the quadratic formula to solve for x:

$$x = \frac{C \pm \sqrt{C^2 + 4t^2}}{2}.$$

III. Problems

1. Solve: $x' = \dfrac{x}{t} \ln \dfrac{t}{x}$. [*Hint.* To integrate $u \cdot \{\ln(1/u) - 1\}$, make the substitution $v := \ln(1/u) - 1$.]

2. Solve: $x' = \dfrac{t + x - 1}{2t - x + 1}$.

§10. Series Solutions of Linear Equations: Ordinary Case

In class, much time was spent on the problem of solving n-th order linear equations with constant coefficients. We considered a linear differential operator

$$L[x] := a_0\, x^{(n)} + a_1\, x^{(n-1)} + \cdots + a_{n-1}\, x' + a_n\, x,$$

where $a_0, a_1, \dots, a_n \in \mathbb{R}$ with $a_0 \neq 0$. We then discussed two distinct problems:

(1) finding n linearly independent solutions $\psi_1(t)$, $\psi_2(t)$, ... , $\psi_n(t)$ — called **fundamental solutions** — of the homogeneous linear ODE $L[x] = 0$, and hence finding the general solution,

$$x(t) = c_1\psi_1(t) + c_2\psi_2(t) + \cdots + c_n\psi_n(t) \qquad (c_1, c_2, \dots, c_n \in \mathbb{R});$$

(2) finding a particular solution $\phi(t)$ of an *inhomogeneous* ODE of the form $L[x] = f$, where $f(t)$ is a given "driving function" (representing some input to a dynamical system, the solution $x(t)$ being the resulting output). For this, we developed two methods. *Variation of parameters* works for every Riemann integrable f (including those with finitely many jump discontinuities — an important case in electrical engineering). The other technique, called the *method of undetermined coefficients* (or the *annihilator method*), works only for driving functions that are superpositions of fundamental solutions for other linear differential operators with constant coefficients.

Of similar importance is the case of a linear equation with *nonconstant coefficients*. This is an equation — either homogeneous or inhomogeneous — based on an operator

$$L[x] := f_0\, x^{(n)} + f_1\, x^{(n-1)} + \cdots + f_{n-1}\, x' + f_n\, x, \tag{$*$}$$

where $f_0, f_1, \dots, f_n$ are given functions of t (some of which may be constant, though not all). These functions are assumed to be defined and continuous on a common interval I.

The only case of this kind we studied was the *first-order* linear equation ($n = 1$), either at an ordinary point or at a singular point. At an ordinary point t_0, where $f_0(t_0) \neq 0$, continuity implies that $f_0(t) \neq 0$ for all t in some interval $J \subseteq I$ centered at t_0. On this interval J, we can divide the equation $L[x] = f$ through by f_0, and we get a equation of the form $(*)$ that is **monic**: its leading coefficient function is 1. In the first-order case, this led to the equation $x' + p(t)\, x = q(t)$, which could always be solved by means of an integrating factor.

If t_0 was a singular point, so that $f_0(t_0) = 0$, the goal was to find solutions $x(t)$ defined on a "punctured interval", $J \setminus \{t_0\}$. To make progress with this, we turned to series solutions (or more precisely, to functions of the form $|t|^r\sigma(t)$, where $r \in \mathbb{R}$ and σ is analytic). This involves the basic assumption that the coefficient functions are analytic at $t = t_0$. Then, under certain favorable circumstances, we were able to produce a solution $x = |t|^r\sigma(t)$ in which the coefficients c_n of σ were determined by a recursion in terms of the prior-computed values $c_0, c_1, \dots, c_{n-1}$, with the initial coefficient c_0 free to take any real value C (so that the general solution is obtained). The favorable circumstances had to do with the nature of the singularity. Essentially, we required $f'(t_0) \neq 0$, so that the function f_0 crosses through its root t_0 "quickly", rather than being tangent to the t-axis and thus remaining close to 0 for a longer period of time. We called such a point t_0 a **regular singularity**.

The method of series solutions is usually the only game in town for solving linear equations with nonconstant coefficients. In fact, just as the convergent power series $\sum_{k\geq 0}(t^k/k!)$ is taken as the definition of the important transcendental function e^t, there are many other useful and important transcendental functions whose definition as convergent power series comes from the demand that they satisfy certain linear ODE's with nonconstant coefficients. Among these are the famous **Bessel functions** $J_\alpha(t)$ and $Y_\alpha(t)$, which find heavy use in physics and engineering, arising as solutions of the **Bessel equation** (which we'll look at shortly): $L[x] = t^2\, x'' + t\, x' + (t^2 - \alpha^2)x = 0$.

In this subsection, we're going to investigate the linear independence of series solutions to a homogeneous linear equation $L[x] = 0$ with nonconstant coefficients. Generally, if the order of L is n, then the recursion that determines the coefficients will leave $c_0, c_1, \ldots, c_{n-1}$ unconstrained, and will express c_j for $j \geq n$ as functions of these n unconstrained parameters. By taking $(c_0, c_1, c_2, \ldots, c_{n-1})$ successively equal to $(1, 0, 0, \ldots, 0, 0)$, $(0, 1, 0, \ldots, 0, 0)$, etc., $(0, 0, 0, \ldots, 1, 0)$, and $(0, 0, 0, \ldots, 0, 1)$, and then letting the recursion generate the rest of the coefficients, we'll produce n linearly independent series solutions of $L[x] = 0$. Let's call them $\psi_1, \psi_2, \ldots, \psi_n$. These solutions span the solution space, by exactly the same kind of reasoning employed in the constant coefficient theory.

With the fundamental solutions in hand, we can also use variation of parameters to solve a inhomogeneous equation $L[x] = f$. Our solution here will only be as accurate as our polynomial approximations to the fundamental solutions, which are known only as series with some large but finite number of coefficients computed (and the rest unknown).

First we consider the situation near an ordinary point t_0, which we can take to be $t_0 = 0$ without loss of generality (simply by shifting the time variable). In this case, there is no loss of generality in assuming that our linear differential operator L is monic, i.e., $f_0(t) \equiv 1$. Also, for $j = 1, 2, \ldots, n$, we have power series expansions

$$f_j(t) = \sum_{k \geq 0} \alpha_{j,k}\, t^k,$$

which we assume are all convergent in the same interval $|t| < R$, with $R > 0$. Since the functions f_j are known, so are the coefficients

$$\alpha_{j,k} = \frac{f_j{}^{(k)}(0)}{k!}$$

We search for solutions in the form

$$x = \sum_{k \geq 0} c_k\, t^k.$$

The ℓ-th derivative of this is computed term-by-term. After re-indexing, it becomes

$$x^{(\ell)} = \sum_{k \geq 0} [k + \ell]_\ell\; c_{k+\ell}\, t^k.$$

Recall that $[m]_p := m(m-1)(m-2)\cdots(m-p+1)$. Let's make the substitution $\ell = n - j$ above. This gives

$$x^{(n-j)} = \sum_{k \geq 0} [k + n - j]_{n-j}\; c_{k+n-j}\, t^k.$$

Our differential equation $L[x] = 0$, or $x^{(n)} + \sum_{j=1}^{n} f_j(t)\, x^{(n-j)}$, now becomes

$$\sum_{k \geq 0} \left\{ [k+n]_n\, c_{k+n} + \sum_{j=1}^{n} \sum_{\ell=0}^{k} [\ell + n - j]_{n-j}\; \alpha_{j,\,k-\ell}\, c_{\ell+n-j} \right\} t^k \;\equiv\; 0.$$

Since a power series can be identically zero only if all its coefficients are zero, we find that

$$\boxed{c_{k+n} = \frac{-1}{[k+n]_n} \sum_{j=1}^{n} \sum_{\ell=0}^{k} [\ell + n - j]_{n-j}\; \alpha_{j,\,k-\ell}\, c_{\ell+n-j}.} \qquad (**)$$

Notice that the only c_i's appearing on the right-hand side of $(**)$ are those with subscripts $i = \ell + n - j$, which we claim must be $< k + n$. This is equivalent to saying that $\ell - j < k$, or $\ell < j + k$. Now $\ell \leq k$ already, as ℓ runs from 0 to k in the inner summation sign. Also, j runs from 1 to n, so certainly $j \geq 1$. It follows that $\ell \leq k < 1 + k \leq j + k$, as required. This shows that $(**)$ is a valid recursion (the current case is computed in terms of the preceding cases only).

As for the convergence of the series solution $x = \sum_{k \geq 0} c_k\, t^k$, it follows from the general theory of analytic ODE's, studied in a prior subsection. Recall from that theory that a scalar ODE of any order, provided it is analytic, has a convergent series solution satisfying any suitable set of initial conditions.

The only question we might still ask about the convergence is whether we can say anything about the *radius* of convergence in this simple linear case. It can be shown that the radius of convergence for the series solution in this case is exactly the same as the common radius of convergence of all the coefficient functions $f_j(t)$, namely R. Thus we have the important principle that for linear ODE's, the solutions always exist (and have nice properties) on precisely the same time-interval on which the problem is meaningful. The same is far from true for nonlinear ODE's, as we have seen. The proof of this important fact will be omitted for now, for want of time, but it will appear here in a future draft.

A nice, simple example of the technique is furnished by the **Airy equation**,

$$x'' = tx.$$

Notice that this has nonconstant coefficients (specifically the coefficient function t), so our theory of constant coefficient LDO's cannot be applied. If the derivative were just x', instead of x'', then the equation would be separable, and trivial to integrate. However, we have no technique that can handle the Airy equation, apart from series solutions.

Let's put the ansatz $x = \sum_{k\geq 0} c_k\, t^k$ into the equation. First note that (after re-indexing)

$$x'' = \sum_{k\geq 0}(k+2)(k+1)\, c_{k+2}\, t^k.$$

Also,

$$tx = t\sum_{k\geq 0} c_k\, t^k = \sum_{k\geq 0} c_k\, t^{k+1} = \sum_{k\geq 1} c_{k-1}\, t^k.$$

In the last step here, we re-indexed, letting the "old k" become $k-1$ for the "new k". Notice that this move shifted the lower limit of summation. A decrease of one in the dummy variable must be counterbalanced by an increase of one in the starting index. The purpose in doing this is so that x'' and tx will both be expressed as series with general terms of the form $(\cdot)\, t^k$, making direct comparison of corresponding coefficients easier. Now, although the general terms in the two series involve like powers of t, the lower limits are not the same. This motivates to break a term off the series with more terms, as follows:

$$x'' = 2c_2 + \sum_{k\geq 1}(k+2)(k+1)\, c_{k+2}\, t^k.$$

For x'' and tx to be identical, it is now clear that we must have

$$c_2 = 0$$

(for there is no constant term in tx), and also by equating coefficients of t^k for $k \geq 1$,

$$(k+2)(k+1)\, c_{k+2} = c_{k-1} \qquad (k \geq 1).$$

Thus, we get the recursion

$$c_{k+2} = \frac{1}{(k+2)(k+1)}\, c_{k-1} \qquad (k \geq 1).$$

This is what's known as a **2-term linear recurrence relation** with a **jump** of 3. It's called this because only two terms of the coefficient sequence appear in the recursion, and they do so in a linear fashion: multiplied only by pure functions of k, rather than by any expressions involving other c_i's. The jump value is the difference between the two subscripts that appear; so for us, $(k+2)-(k-1) = 3$. This tells us that each term determines the term three units forward in the sequence. When k has its smallest value 1, the recursion says that $c_3 = (1/6)c_0$. When $k = 2$, we get $c_4 = (1/12)c_1$, and when $k = 3$, $c_5 = (1/20)c_2 = 0$ (since $c_2 = 0$). This can go on indefinitely, with all terms of the form $c_{2+3\ell}$ being 0. On the other hand, notice that c_0 and c_1 are unconstrained by the recurrence relation. So these are two independent parameters. This certainly makes sense, as the Airy equation is a second-order linear ODE, and should have a two-dimensional solution space.

In case of 2-term linear recurrence, we can actually solve more-or-less explicitly, though we may need to invent new notations for certain recursive functions so as to be able to write them in "closed form". An example of this is the notation '$n!$', which is defined recursively by $0! := 1$ and $(n+1)! := (n+1)(n!)$ for $n \geq 0$. When we write '$n!$', we are cheating a bit in order to have a "closed form" expression. Informally, however, we may view it as justified to introduce new notations of this kind when we believe that they will be of use not just in solving the problem at hand, but in a variety of other contexts as well. A notation of this kind that has come into common usage recently is the *skip-factorial function*, $n!!$, which does not mean $(n!)!$ — that would have to be written with parentheses — but rather means

$$n!! := n(n-2)(n-4)\cdots ,$$

where the product ends with a factor of 1 if n is odd, and with a factor of 2 if n is even. That is, $1!! := 1$ and $2!! := 2$, while $(2n+2)!! := (2n+2)(2n)!!$ and $(2n+3)!! := (2n+3)(2n+1)!!$ for $n \geq 0$. We can also have the *double-skip-factorial* $n!!! = n(n-3)(n-6)\cdots$, etc.

Continuing with our problem, let's take $k = 3\ell - 2$ in our recursion. We get

$$c_{3\ell} = \frac{1}{(3\ell)(3\ell-1)}\, c_{3\ell-3} = \frac{1}{(3\ell)(3\ell-1)} \cdot \frac{1}{(3\ell-3)(3\ell-4)}\, c_{3\ell-6} = \cdots .$$

This product will terminate when the final subscript is 0, giving

$$c_{3\ell} = \left[\frac{1}{(3\ell)(3\ell-1)} \cdot \frac{1}{(3\ell-3)(3\ell-4)} \cdots \frac{1}{3\cdot 2}\right] c_0 = \frac{c_0}{(3\ell)!!!\,(3\ell-1)!!!}.$$

Here, we have used the double-skip-factorial function to help us express the answer in closed form. A similar analysis starting with $k = 3\ell - 1$ will give

$$c_{3\ell+1} = \frac{c_1}{(3\ell+1)!!!\,(3\ell)!!!}$$

As we already know,

$$c_{3\ell+2} = \frac{c_2}{(3\ell+2)!!!\,(3\ell+1)!!!} = 0,$$

since $c_2 = 0$. We now have the following expansion for our series solution, after factoring c_0 from all the terms with "threeven" subscripts and factoring c_1 from the rest of the nonzero terms:

$$x \;=\; c_0\left[\sum_{\ell\geq 0} \frac{t^{3\ell}}{(3\ell)!!!\,(3\ell-1)!!!}\right] \;+\; c_1\left[\sum_{\ell\geq 0} \frac{t^{3\ell+1}}{(3\ell+1)!!!\,(3\ell)!!!}\right].$$

To make these expressions fully meaningful, we need to adopt the conventions that $0!!! = (-1)!!! = 1$. Of course, $1!!! = 1$; $2!!! = 2$; $3!!! = 3$; $4!!! = 4\cdot 1 = 4$; $5!!! = 5\cdot 2 = 10$; $6!!! = 6\cdot 3 = 18$; $7!!! = 7\cdot 4\cdot 1 = 28$; $8!!! = 8\cdot 5\cdot 2 = 80$; etc. Our convention gives us the correct constant terms in the above expansion. For example, looking at the first set of brackets, setting $\ell = 0$ gives $c_0\big[t^0/\{0!!!\,(-1)!!!\}\big] = c_0 \cdot 1 = c_0$, which should be the constant term of $x(t)$.

The series appearing in the brackets above actually converge for all $t \in \mathbb{R}$, by a simple application of the Ratio Test (try it). Let's write

$$\phi(t) := \sum_{\ell\geq 0} \frac{t^{3\ell}}{(3\ell)!!!\,(3\ell-1)!!!}, \qquad \psi(t) := \sum_{\ell\geq 0} \frac{t^{3\ell+1}}{(3\ell+1)!!!\,(3\ell)!!!}.$$

So the general solution of the Airy equation is

$$x(t) = c_0\,\phi(t) + c_1\,\psi(t).$$

Moreover, the two functions so defined are linearly independent. For if we set the above linear combination of ϕ and ψ equal to the identically zero function, then we get $x(t) \equiv 0$, forcing every coefficient of the power series to be zero (by the "Identity Theorem" for power series). In particular, the constant term c_0 must be zero, as must the coefficient c_1 of t^1. That is, $c_0 = c_1 = 0$. So only the trivial combination of ϕ and ψ can produce the identically zero function. We see that the two functions form a basis for the solution space of Airy's equation.

The argument just given regarding linear independence is actually perfectly general. The n-th order equation $L[x] = 0$, with its series solution given by the recursion $(**)$, will have its general solution $x(t)$ ultimately dependent on exactly n independent parameters — namely, the coefficients $c_0, c_1, \ldots, c_{n-1}$ — becuase the first coefficient to be constrained in any way by $(**)$ is c_n, when k has its smallest value of 0. Now the expression for c_n doesn't only involve c_0. Instead it involves $c_0, c_1, \ldots, c_{n-1}$. The important thing to notice is that the expression for c_n generated from $(**)$ by taking $k = 0$ is *linear* in $c_0, c_1, \ldots, c_{n-1}$. Thus, the term $c_n\, t^n$ can be split up into a number of terms of the forms $c_0(\cdot)\, t^n$, $c_1(\cdot)\, t^n$, etc., and $c_{n-1}(\cdot)\, t^n$, where the contents of the parentheses are purely numerical, and do not involve any of the c_i's. In a similar way, the expression for c_{n+1} generated by $(**)$ on taking $k = 1$ will involve $c_0, c_1, \ldots, c_{n-1}, c_n$ in a linear fashion; but since c_n is itself a linear expression in $c_0, c_1, \ldots, c_{n-1}$, we can eliminate c_n, and split the term $c_{n+1}\, t^{n+1}$ into terms like $c_0(\cdot)\, t^{n+1}$, $c_1(\cdot)\, t^{n+1}$, etc., and $c_{n-1}(\cdot)\, t^{n+1}$.

Now by organizing the terms of $x(t)$ into n groups, with those in the first group having the form $c_0(\cdot)\, t^*$ for some exponent '$*$', those in the second group having the form $c_1(\cdot)\, t^*$, etc., and those in the last group having the form $c_{n-1}(\cdot)\, t^*$, we can write

$$x(t) = c_0\, \psi_1(t) + c_1\, \psi_2(t) + \cdots + c_{n-1}\, \psi_n(t),$$

where $\psi_1, \psi_2, \ldots, \psi_n$ are now specific numerical functions, with no arbitrary parameters appearing in their coefficients. These n functions may be called **fundamental solutions** for the equation $L[x] = 0$. They are linearly independent, by the same argument we gave for the Airy equation.

Here it is in general. Set

$$x(t) = c_0\, \psi_1(t) + c_1\, \psi_2(t) + \cdots + c_{n-1}\, \psi_n(t) \;\equiv\; 0.$$

Since $x(t)$ is a power series, it can only be identically zero if all the coefficients are zero. On the other hand, $c_0, c_1, \ldots, c_{n-1}$ are precisely the coefficients of the first n terms of the series expansion for $x(t)$. So we have $c_0 = c_1 = \cdots = c_{n-1} = 0$, thus proving the linear independence. Since it's clear that the general solution is spanned by the fundamental solutions, we have a basis for the solution space (exactly as in the case of constant coefficients, but with a very different method for producing the fundamental solutions).

II. Problems

1. Consider the famous **Legendre equation**,

$$(1 - t^2)x'' - 2tx + \alpha(\alpha + 1)x = 0,$$

which arises in quantum mechanics when we search for so-called "spherical harmonics", for instance in the analysis of the electron orbits of a hydrogen atom. Here, $\alpha > 0$ is a fixed constant. Notice that there are two singular points for this equation, where the leading coefficient function vasnishes: $t = \pm 1$. However, focus on the ordinary point $t = 0$. Using series expansions in powers of t, find a pair of fundamental solutions of Legendre equation. Determine the coefficients of these two series solutions as precisely as possible; either by giving a formula for them, as we did with the Airy equation, or failing that, computing as many of them as you can. Finally, show that in the special case $\alpha = N$, where N is a positive integer, one of the series solutions actually terminates, giving an exact polynomial solution. A certain re-normalization of this polynomial solution (where the scale factor is chosen to make the graph pass through the point $(1, 1)$ in the tx-plane) is called the **Legendre polynomial of order** N, denoted P_N. [This has many remarkable properties we don't have time to go into here. One, which I can't resist showing you, is the famous **Rodrigues Formula**

$$P_N(t) = \frac{1}{2^N (N!)}\, \frac{\mathrm{d}^N}{\mathrm{d}t^N}\left[(t^2 - 1)^N\right].$$

There is a great deal more to be said about this subject, some of which you may yet see in advanced engineering courses.]

§11. Series Solutions of Linear Equations: Regular Singularities

I. Theory

UNDER CONSTRUCTION!

II. Examples

III. Problems