Natural Sciences Tripos, Part IA

Lent Term 2024

Mathematical Methods II, Course B

Prof Natalia Berloff

N.G.Berloff@damtp.cam.ac.uk

Acknowledgment: These notes are a modified

version of those produced by Dr R. Rafikov, Dr S. Colwell, and others.

Course website

Up to date information can be found on Moodle

https://www.vle.cam.ac.uk/

Schedule

Ordinary differential equations. First-order equations: separable equations; linear equations, integrating factors. Examples involving substitution. Second-order linear equations with constant coefficients; $\exp(\lambda x)$ as a trial solution, including degenerate case. Superposition. Particular integrals and complementary functions. Constants of integration and a number of necessary boundary/initial conditions. Particular integrals by trial solutions. Examples including radioactive sequences. Resonance, transients and damping. [6 lectures]

Differentiation of functions of several variables. Differentials, chain rule. Exact differentials, illustrations including Maxwell's relations. Scalar and vector fields. Gradient of a scalar as a vector field. Directional derivatives. Unconditional stationary values; classification using Hessian matrix. Conditional stationary values, Lagrange multipliers, examples with two or three variables. Boltzmann distribution as an example. [8 lectures]

Parameterised curves. Line integral of a vector field. Conservative and non-conservative vector fields. Surface integrals and flux of a vector field over a surface. Divergence of a vector field. ∇^2 as div grad. Curl. Divergence and Stokes's theorems. [5 lectures]

Orthogonality relations for sine and cosine. Fourier series; examples. [2 lectures]

Extended examples distributed through the course. [3 lectures]

Recommended textbook

K. F. Riley, M. P. Hobson and S. J. Bence (2006). *Mathematical Methods for Physics and Engineering*, 3rd edition. Cambridge University Press

M L Boas Mathematical Methods in the Physical Sciences, 3rd edition. Wiley, 1983 (paperback and hardback)

E Kreyszig Advanced Engineering Mathematics, 9th edition. Wiley, 2005 (8th edition 1999)

G Stephenson Mathematical Methods for Science Students, 2nd edition. Prentice Hall/Pearson, 1973 (paperback)

A Jeffrey Mathematics for Engineers and Scientists, 6th edition. Chapman & Hall, 2004 (paperback)

1 ORDINARY DIFFERENTIAL EQUATIONS

1.1 Introduction

1.1.1 Familiar examples

1. A body of mass m falls under gravity (g) with a drag force proportional to the square of the velocity (with constant of proportionality μ). The downward velocity v varies with the time t according to Newton's second law

   $$\text{mass} \times \text{acceleration} = \text{force}$$

   $$m\frac{\mathrm{d}v}{\mathrm{d}t} = mg - \mu v^2.$$

2. Two chemical species A and B react to form a third species C: $\mathrm{A} + \mathrm{B} \to \mathrm{C}$. Their concentrations a, b and c depend on time according to the rate equations

   $$\frac{\mathrm{d}a}{\mathrm{d}t} = -kab,$$
   $$\frac{\mathrm{d}b}{\mathrm{d}t} = -kab,$$
   $$\frac{\mathrm{d}c}{\mathrm{d}t} = +kab,$$

   where k is the rate constant ($k > 0$).

In Ex 1 we have an equation giving the derivative of an unknown function $v(t)$ in terms of $v(t)$ itself.

In Ex 2 we have three such equations for three unknown functions $a(t)$, $b(t)$ and $c(t)$.

1.1.2 Classification

In the mathematical study of *ordinary differential equations* (ODEs) it is helpful to introduce a standard notation and to call the unknown function $y(x)$.

- x is called the *independent variable*;
- y (which depends on x) is called the *dependent variable.*

In Example 1 we could convert $v(t)$ into $y(x)$ by relabelling the variables:

$$m\frac{\mathrm{d}y}{\mathrm{d}x} = mg - \mu y^2.$$

The first derivative of $y(x)$ can be written as $\frac{\mathrm{d}y}{\mathrm{d}x}$ or $y'(x)$. Similarly, the second derivative can be written as $\frac{\mathrm{d}^2y}{\mathrm{d}x^2}$ or $y''(x)$. The nth derivative is $\frac{\mathrm{d}^ny}{\mathrm{d}x^n}$ or $y^{(n)}(x)$. When it is understood that y is a function of x, we may not write the (x).

(When the independent variable is time t, it is more usual to write $\dot{y}$, $\ddot{y}$, etc., instead of y', y'', etc.)

The general form of a first-order ODE is

$$F(y', y, x) = 0,$$

i.e. a functional relationship between the derivative of the unknown function, the function itself and the independent variable. We can often solve this equation for y' and write it in the alternative form

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

Similarly, the general form of an nth-order ODE is

$$F(y^{(n)}, y^{(n-1)}, \ldots, y, x) = 0,$$

i.e. a functional relationship between the first n derivatives of the unknown function, the function itself and the independent variable. We can often write this in the form

$$y^{(n)} = f(y^{(n-1)}, \ldots, y, x).$$

These are called *ordinary* differential equations because the unknown function $y(x)$ depends on only one independent variable x and so the derivatives appearing in the equations are ordinary derivatives rather than partial derivatives (see Section 2 later).

Example 2 is actually a *system* of ODEs. We shall not say much about systems of ODEs in this course. However, we can note that in this case

$$\frac{\mathrm{d}b}{\mathrm{d}t} - \frac{\mathrm{d}a}{\mathrm{d}t} = 0 \quad \Rightarrow \quad b - a = \text{constant} = b_0 - a_0,$$

where a_0 and b_0 are the initial values of a and b, and so

$$\frac{\mathrm{d}a}{\mathrm{d}t} = -ka(a + b_0 - a_0),$$

which is a first-order ODE for a alone.

Differential equations can be solved by

- *analytical* methods (exact mathematics);
- *numerical* methods (approximation on a computer).

In this course, we discuss some of the more important analytical methods.

1.2 First-order equations

1.2.1 Directly integrable (exact) equations

The simplest type of differential equation is of the form

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x).$$

This equation can be integrated directly:

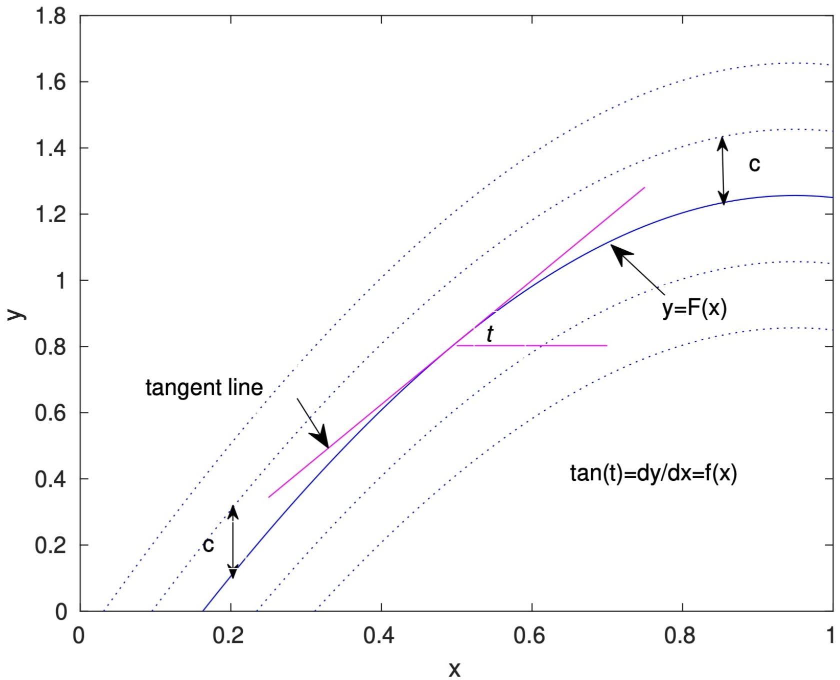
$$y = \int f(x)\,\mathrm{d}x.$$

In this case, the ODE is solved in the sense that it is reduced to the problem of finding the integral of a given function.

If $F(x)$ is an antiderivative of $f(x)$, i.e. $F'(x) = f(x)$, then the general solution is

$$y = F(x) + c,$$

where c is an arbitrary constant. Recall that the notation $\int f(x)\,\mathrm{d}x$ means an *indefinite integral*, which always includes an arbitrary additive constant.



We obtain a family of solutions, distinguished by the value of the parameter c. Each one represents a curve in the (x, y) plane. The differential equation tells us the slope of the curve at each point. In this case the slope depends only on x.

This method does not work with the more general first-order ODE

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(y, x)$$

because the right-hand side contains the unknown function and cannot be integrated directly with respect to x.

There are examples of exact differential equations (see Section 2.2.10 later) in which the terms can be grouped in such a way that the equation can be integrated directly.

In many other examples there is a way to rewrite the ODE so that it can be integrated directly. The most important cases are those of *separable* equations and *linear* equations.

1.2.2 Separable equations

A *separable* first-order ODE is one in which the two variables appear in separate factors, i.e. of the general form

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{f(x)}{g(y)}.$$

(We could have called the right-hand side $f(x)h(y)$, but the above notation simplifies the solution.)

Multiply both sides of the equation by $g(y)$:

$$g(y)\frac{\mathrm{d}y}{\mathrm{d}x} = f(x).$$

The equation can now be integrated directly with respect to x, using integration by substitution on the left-hand side:

$$\int g(y)\frac{\mathrm{d}y}{\mathrm{d}x}\,\mathrm{d}x = \int f(x)\,\mathrm{d}x$$

$$\int g(y)\,\mathrm{d}y = \int f(x)\,\mathrm{d}x.$$

Again, the ODE is solved in the sense that it is reduced to the problem of finding the integrals of two given functions. If $F(x)$ is an integral of $f(x)$ and $G(y)$ is an integral of $g(y)$, then the general solution is

$$G(y) = F(x) + c,$$

where c is an arbitrary constant.

Note:

- There is no need to introduce *two* arbitrary constants here because they combine into one.
- It may or may not be possible to go further and express y explicitly as a function of x.
- It is permissible to separate the variables by writing

  $$g(y)\,\mathrm{d}y = f(x)\,\mathrm{d}x$$

  and then integrate to find the solution as above. Although $\frac{\mathrm{d}y}{\mathrm{d}x}$ does not mean $\mathrm{d}y$ divided by $\mathrm{d}x$, we will later give a meaning to relations such as this between *differentials* (Section 2.2.6)

Example:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x^2+1}{y+1}.$$

Multiply by $y+1$ to separate the variables:

$$(y+1)\frac{\mathrm{d}y}{\mathrm{d}x} = x^2+1.$$

Integrate with respect to x:

$$\int (y+1)\frac{\mathrm{d}y}{\mathrm{d}x}\,\mathrm{d}x = \int (y+1)\,\mathrm{d}y = \int (x^2+1)\,\mathrm{d}x$$

$$\tfrac{1}{2}y^2 + y = \tfrac{1}{3}x^3 + x + c. \tag{1}$$

This is an *implicit solution* for y in terms of x. In this particular case it is possible to obtain an *explicit solution*:

$$\tfrac{1}{2}(y+1)^2 = \tfrac{1}{3}x^3 + x + c + \tfrac{1}{2}$$

$$y = -1 \pm \sqrt{\tfrac{2}{3}x^3 + 2x + 2c + 1}.$$

Note the choice of sign in this solution, to be discussed in the next section.

We can *check* by substitution that this solution satisfies the differential equation:

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{\mathrm{d}}{\mathrm{d}x}\left(-1 \pm \sqrt{\tfrac{2}{3}x^3 + 2x + 2c + 1}\right) \\
&= \pm\frac{1}{2}\frac{2x^2+2}{\sqrt{\frac{2}{3}x^3 + 2x + 2c + 1}} \\
&= \frac{x^2+1}{y+1}. \qquad \checkmark
\end{aligned}$$

It is often much easier to check a solution than to obtain it in the first place, so this is well worth doing.

Example:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x^2+1}{1+\cos y}.$$

Separate the variables:

$$\int (1+\cos y)\,\mathrm{d}y = \int (x^2+1)\,\mathrm{d}x$$

$$y + \sin y = \tfrac{1}{3}x^3 + x + c.$$

In this case, we cannot find y in terms of x analytically but could do so approximately using a computer.

1.2.3 Initial condition

In the solutions written above an arbitrary constant of integration appears. Each value of the constant gives a different solution. The expression for the solution including the arbitrary constant is called the *general solution*.

To specify a unique solution for a first-order ODE we require one extra piece of information, usually the value of y at some value of x. This extra piece of information is usually called an *initial condition* (or *boundary condition*).

In Example 1 in Section 1.1.1 (falling mass with drag force), we need to know the value of the velocity v at some initial time t_0 to specify the solution uniquely for $t > t_0$.

In the previous example of a separable equation,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x^2+1}{y+1},$$

let us add the initial condition that $y = 1$ at $x = 0$.

As before, separating variables and integrating leads to

$$\tfrac{1}{2}y^2 + y = \tfrac{1}{3}x^3 + x + c \tag{1}$$

and then

$$y = -1 \pm \sqrt{\tfrac{2}{3}x^3 + 2x + 2c + 1}.$$

Substituting the initial values $y = 1$ and $x = 0$ gives us an equation for c:

$$1 = -1 \pm \sqrt{2c + 1}.$$

We must choose the plus sign; then we find $c = \frac{3}{2}$. The specific solution of the equation, subject to the given initial condition, is therefore

$$y = -1 + \sqrt{\frac{2}{3}x^3 + 2x + 4}.$$

Note that we could have found the value of c at an earlier stage, by substituting the initial values $y = 1$ and $x = 0$ into the implicit form (1) of the general solution.

1.2.4 Linear equations

A *linear* first-order ODE is one of the general form

$$\frac{\mathrm{d}y}{\mathrm{d}x} + p(x)y = f(x),$$

in which the dependent variable y and its derivative appear linearly. If the term $f(x)$ is absent, the equation is *homogeneous*. With the $f(x)$ term it is *inhomogeneous*.

- ▶ Linearity is an essential concept in mathematics. A linear function of x is one of the form $ax+b$, where a and b are constants. If $b=0$ this is called a homogeneous linear function, otherwise it is inhomogeneous. A linear function of x and y is one of the form $ax+by+c$, etc.

The homogeneous case

$$\frac{\mathrm{d}y}{\mathrm{d}x} + p(x)y = 0$$

is separable:

$$\int \frac{\mathrm{d}y}{y} = -\int p(x)\,\mathrm{d}x$$

$$\ln y = -\int p(x)\,\mathrm{d}x$$

$$y = \exp\left[-\int p(x)\,\mathrm{d}x\right].$$

If $P(x)$ is an integral of $p(x)$, then the general solution is

$$y = \exp[-P(x) + c] = C \exp[-P(x)],$$

where $C = \exp(c)$. Since the indefinite integral $\int p(x)\,\mathrm{d}x$ includes an arbitrary additive constant (c), the solution y involves an arbitrary *multiplicative* constant (C).

The inhomogeneous case

$$\frac{\mathrm{d}y}{\mathrm{d}x} + p(x)y = f(x) \tag{3}$$

can be integrated if we first multiply the equation by a suitable factor $\mu(x)$, known as an *integrating factor*:

$$\mu(x)\frac{\mathrm{d}y}{\mathrm{d}x} + \mu(x)p(x)y = \mu(x)f(x). \tag{3a}$$

We choose $\mu(x)$ so that the left-hand side is the derivative of something with respect to x. Noticing that as

$$\frac{\mathrm{d}}{\mathrm{d}x}[\mu(x)y] = \mu(x)\frac{\mathrm{d}y}{\mathrm{d}x} + \frac{\mathrm{d}\mu(x)}{\mathrm{d}x}y$$

the first term of the left-hand side of (3a) is part of the derivative of the product $\mu(x)y$, we choose $\mu(x)$ so that the second term is the other part. i.e.

$$\mu(x)p(x)y = \frac{\mathrm{d}\mu}{\mathrm{d}x}\,y.$$

This can be done without knowing what y is (because equation 3 is linear). Our condition for the integrating factor is a separable ODE:

$$\int \frac{\mathrm{d}\mu}{\mu(x)} = \int p(x)\,\mathrm{d}x$$

$$\ln \mu(x) = \int p(x)\,\mathrm{d}x$$

$$\mu(x) = \exp\left[\int p(x)\,\mathrm{d}x\right]. \tag{4}$$

Note that $\mu(x)$ is just the reciprocal of the solution of the homogeneous equation (2). Again, the integrating factor contains an arbitrary multiplicative constant. In this case it does not matter what value we choose for this constant.

Having multiplied equation (3) by $\mu(x)$, we now have

$$\frac{\mathrm{d}}{\mathrm{d}x}[\mu(x)y] = \mu(x)f(x),$$

which can be integrated directly:

$$\mu(x)y = \int \mu(x)f(x)\,\mathrm{d}x.$$

If we included an arbitrary multiplicative constant in $\mu(x)$, it would cancel out at this point. However, the indefinite integral on the right-hand side does contain an arbitrary additive constant.

Note that we cannot cancel $\mu(x)$ itself on both sides of this equation, because on the right-hand side it appears inside an integral. To avoid possible confusion, especially if we want to divide by $\mu(x)$, we may prefer to introduce a dummy integration variable (t, say) and write

$$y(x) = \frac{1}{\mu(x)} \int^x \mu(t) f(t)\,\mathrm{d}t. \tag{5}$$

Example:

$$\frac{\mathrm{d}y}{\mathrm{d}x} + 3y = \mathrm{e}^{-2x}$$

with initial condition $y = 2$ at $x = 0$.

First compare with the standard form $\frac{\mathrm{d}y}{\mathrm{d}x} + p(x)y = f(x)$:

$$p(x) = 3, \qquad f(x) = \mathrm{e}^{-2x}.$$

Calculate the integrating factor:

$$\mu(x) = \exp\left[\int p(x)\,\mathrm{d}x\right] = \exp\left(\int 3\,\mathrm{d}x\right) = \mathrm{e}^{3x}.$$

(As noted above, we can discard the arbitrary constant in this expression.) Multiply the ODE by the integrating factor:

$$\mathrm{e}^{3x}\frac{\mathrm{d}y}{\mathrm{d}x} + 3\,\mathrm{e}^{3x}y = \mathrm{e}^{3x}\,\mathrm{e}^{-2x}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{3x}y\right) = \mathrm{e}^{x}.$$

Integrate with respect to x:

$$\mathrm{e}^{3x}y = \mathrm{e}^{x} + c.$$

General solution:

$$y = \mathrm{e}^{-2x} + c\,\mathrm{e}^{-3x}.$$

Apply the initial condition $y = 2$ at $x = 0$:

$$2 = 1 + c.$$

Thus $c = 1$ and the specific solution is

$$y = \mathrm{e}^{-2x} + \mathrm{e}^{-3x}.$$

Check that the solution satisfies the ODE:

$$\begin{aligned}\frac{\mathrm{d}y}{\mathrm{d}x} + 3y &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{-2x} + \mathrm{e}^{-3x}\right) + 3\left(\mathrm{e}^{-2x} + \mathrm{e}^{-3x}\right)\\ &= -2\,\mathrm{e}^{-2x} - 3\,\mathrm{e}^{-3x} + 3\,\mathrm{e}^{-2x} + 3\,\mathrm{e}^{-3x}\\ &= \mathrm{e}^{-2x}. \qquad \checkmark\end{aligned}$$

Note that we generally remember the formula (4) for the integrating factor, but we don't substitute p and f into the general formula (5) for the solution which can also be written:

$$y(x) = \exp\left[-\int^{x} p(t)\,\mathrm{d}t\right]\int^{x} f(s)\; e^{\int^{s} p(w)\mathrm{d}w}\mathrm{d}s.$$

1.2.5 Solution by substitution

Many first-order ODEs can be solved by making a substitution (a change of variables) that reduces the equation to a separable one or a linear one. Finding such a substitution (if it exists) may require intelligent guesswork. The aim is to simplify the nonlinearity in the equation, reducing it to either a separable or a linear form. Two types of equation that can be solved by substitution are:

- the *homogeneous differential equation*
$$\frac{\mathrm{d}y}{\mathrm{d}x} = f\left(\frac{y}{x}\right); \tag{6}$$
- the *Bernoulli differential equation*
$$\frac{\mathrm{d}y}{\mathrm{d}x} + p(x)y = q(x)y^n. \tag{7}$$

Equation (6) is known as a *homogeneous* equation, although this word has more than one meaning in the context of ODEs. In this case it refers to the fact that the equation is unchanged when y is replaced by αy *and* x is replaced by αx, i.e. if the (x, y) plane is stretched by the same amount in both directions.

Equation (6) is solved by the substitution $y = u(x)x$:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}(ux) = \frac{\mathrm{d}u}{\mathrm{d}x}x + u = f(u)$$
$$\frac{\mathrm{d}u}{\mathrm{d}x} = \frac{f(u) - u}{x}.$$

This is a separable equation that can be solved as in Section 1.2.2.

Example:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x^2 + y^2}{xy} \qquad \text{with initial condition } y = 1 \text{ at } x = 1.$$

The equation is homogeneous (in the xy sense!) because the right-hand side in unchanged when x is replaced by αx and y is replaced by αy, as both numerator and denominator are multiplied by α^2. Substitute $y = ux$:

$$\frac{\mathrm{d}u}{\mathrm{d}x} x + u = \frac{x^2 + u^2 x^2}{xux} = \frac{1 + u^2}{u}$$

$$\frac{\mathrm{d}u}{\mathrm{d}x} x = \frac{1}{u} \Rightarrow$$

$$\int u \,\mathrm{d}u = \int \frac{\mathrm{d}x}{x}$$

$$\tfrac{1}{2} u^2 = \ln x + c.$$

Apply the initial condition $y = 1$ at $x = 1$ (so $u = 1$ there):

$$\tfrac{1}{2} = c.$$

Thus (plus sign needed)

$$u = \sqrt{2 \ln x + 1} \Rightarrow y = x\sqrt{2 \ln x + 1}.$$

Equation (7) is known as a *Bernoulli equation*. If $n = 0$ or $n = 1$, the equation is already linear. In all other cases it is solved by the substitution $z = y^{1-n}$. This works because

$$\begin{aligned}\frac{\mathrm{d}z}{\mathrm{d}x} &= (1-n)y^{-n}\frac{\mathrm{d}y}{\mathrm{d}x} \\ &= (1-n)y^{-n}[-p(x)y + q(x)y^n] \\ &= (1-n)[-p(x)z + q(x)],\end{aligned}$$

which is a linear first-order ODE for z.

Example:

$$\frac{\mathrm{d}y}{\mathrm{d}x} + xy = x^3y^2.$$

This is a Bernoulli equation with $n = 2$, so let $z = y^{-1}$:

$$\frac{\mathrm{d}z}{\mathrm{d}x} = -\frac{1}{y^2}\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{1}{y^2}(-xy + x^3y^2) = \frac{x}{y} - x^3 = xz - x^3.$$

Write in the standard form of a linear first-order ODE:

$$\frac{\mathrm{d}z}{\mathrm{d}x} - xz = -x^3.$$

Integrating factor:

$$\mu(x) = \exp\left[\int(-x)\,\mathrm{d}x\right] = \mathrm{e}^{-x^2/2}.$$

Multiply ODE by IF:

$$\mathrm{e}^{-x^2/2}\left(\frac{\mathrm{d}z}{\mathrm{d}x} - xz\right) = \frac{\mathrm{d}}{\mathrm{d}x}\left(z\,\mathrm{e}^{-x^2/2}\right) = -x^3\,\mathrm{e}^{-x^2/2}.$$

Integrate:

$$\begin{aligned} z\,\mathrm{e}^{-x^2/2} &= -\int x^3\,\mathrm{e}^{-x^2/2}\,\mathrm{d}x \\ &= \int x^2\left(-x\,\mathrm{e}^{-x^2/2}\right)\,\mathrm{d}x \\ &= x^2\,\mathrm{e}^{-x^2/2} - \int 2x\,\mathrm{e}^{-x^2/2}\,\mathrm{d}x \\ &= (x^2+2)\,\mathrm{e}^{-x^2/2} + c. \end{aligned}$$

Thus

$$z = x^2 + 2 + c\,\mathrm{e}^{x^2/2}$$

and so

$$y = \frac{1}{x^2 + 2 + c\,\mathrm{e}^{x^2/2}}.$$

Other examples of the change of variables.

1.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = F(ax + by + c)$$

where F(z) is an arbitrary function

$$u(x) = ax + by + c$$

$$\frac{\mathrm{d}u}{\mathrm{d}x} = a + b\frac{\mathrm{d}y}{\mathrm{d}x} = a + bF(u) \qquad \Rightarrow \qquad \frac{\mathrm{d}u}{a + b\ F(u)} = \mathrm{d}x$$

$$x = \int^{ax+by+c} \frac{\mathrm{d}u}{a + b\ F(u)}.$$

2.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = yf(e^{\alpha x}y^{\beta})$$

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}ln(y)}{\mathrm{d}x} = f(e^{\alpha x}y^{\beta}) = f(e^{\alpha x + \beta\ln(y)})$$

Introduce new function $g(z) = f(e^z)$, $\Rightarrow$

$$\frac{\mathrm{d}ln(y)}{\mathrm{d}x} = g(\alpha x + \beta ln(y))$$

Introduce $u(x) = ln(y)$ substitution $\Rightarrow$

$$\frac{\mathrm{d}u}{\mathrm{d}x} = g(\alpha x + \beta u)$$

which is an equation of the previous type.

Solutions in a parametric form (non-examinable)

Suppose $F(y', y, x) = 0$ - can't be resolved w.r.t. y'.

1. Suppose that F is such that

   (a) $y(x)$ doesn't appear explicitly ($F(y', x) = 0$),

   (b) resolvable w.r.t. x. $\quad x = f(y')$.

   Substitution $s = y'$:

   $$s(x) = y' = \frac{\mathrm{d}y}{\mathrm{d}x}; \qquad \frac{\mathrm{d}x}{\mathrm{d}s} = f'(s) = \frac{\mathrm{d}f}{\mathrm{d}s}$$

   $$y = \int s(x)dx = \int s(x)\frac{\mathrm{d}x}{\mathrm{d}s}ds = \int sf'(s)ds$$

   i.e.

   $$x = f(s), \qquad y = \int sf'(s)ds$$

   where s is a parameter.

   If lucky, can get $s(y)$ and hence $x = f(s(y))$, i.e. solution in implicit form.

2. Suppose F is such that

   (a) x doesn't appear explicitly ($F(y', y) = 0$),

   (b) resolvable w.r.t. y: $y = f(y')$

   Substitution $r = y'$.

$$r(y) = y' = \frac{\mathrm{d}y}{\mathrm{d}x}; \qquad \frac{\mathrm{d}y}{\mathrm{d}r} = f'(r) = \frac{\mathrm{d}f}{\mathrm{d}r}$$

$$x = \int \frac{1}{r(y)} dy = \int \frac{1}{r(y)} \frac{\mathrm{d}y}{\mathrm{d}r} dr = \int \frac{1}{r} f'(r) dr$$

   i.e.

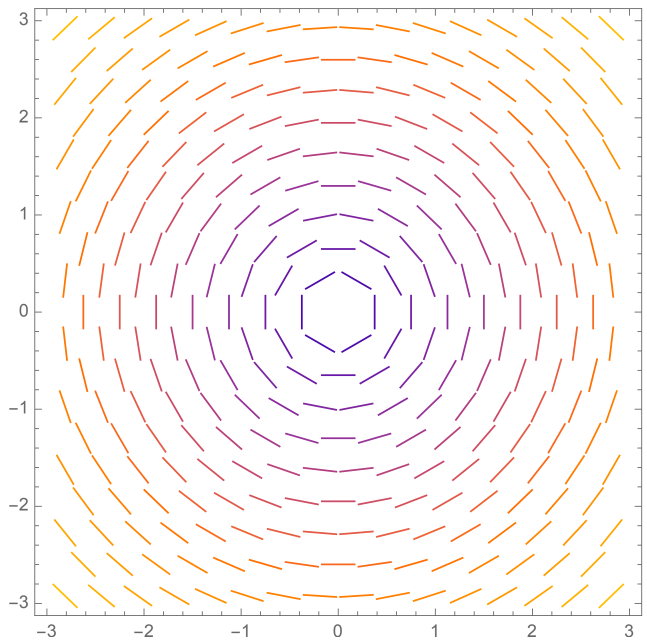
$$y = f(r), \qquad x = \int r^{-1} f'(r) dr$$

   which is the solution in a parametric form.

1.2.6 Additional remarks

Often, first-order ODEs cannot be solved analytically, e.g. $y' = x + y^3$ (note that this equation is neither linear nor separable, and substitutions do not help).

A first-order ODE can be interpreted *geometrically*. The equation $F(y', y, x) = 0$ defines the slope of the solution curve at every point of the (x, y) plane. There is a one-parameter family of solution curves, with one passing through each point where f is well defined. This is consistent with the appearance of a single arbitrary constant in the general analytical solution, and with the fact that **one** initial condition must be specified to determine a unique solution.



1.2.7 Application: radioactive sequences

This is a scientifically interesting example and demonstrates how to solve a simple coupled equation system.
In a radioactive rock, isotope A decays into isotope B at a rate proportional to the number a of remaining nuclei of A. Isotope B decays into isotope C at a rate proportional to the number b of remaining nuclei of B. (This is the sequence A $\to$ B $\to$ C.) Determine expressions for $a(t)$ and $b(t)$.

Assume that the decay constant for A is k_a and the decay constant for B is k_b ($\neq k_a$). Then $a(t)$ and $b(t)$ satisfy the equations

$$\frac{\mathrm{d}a}{\mathrm{d}t} = -k_a a, \qquad \frac{\mathrm{d}b}{\mathrm{d}t} = k_a a - k_b b.$$

Note that the equation for $a(t)$ decouples from that for $b(t)$. The solution is

$$a(t) = a_0 \exp(-k_a t),$$

where $a_0 = a(0)$.
Substitute into the equation for $b(t)$:

$$\frac{\mathrm{d}b}{\mathrm{d}t} + k_b b = k_a a_0 \exp(-k_a t).$$

This is a linear first-order ODE. Multiply by the integrating factor $\exp(k_b t)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\exp(k_b t) b\right] = k_a a_0 \exp[(k_b - k_a)t].$$

Integrate:

$$\exp(k_b t)b = \frac{k_a a_0}{k_b - k_a}\exp[(k_b - k_a)t] + C$$

$$b(t) = \frac{k_a a_0}{k_b - k_a}\exp(-k_a t) + C\exp(-k_b t).$$

Suppose that the initial condition is $b(0) = 0$. Then

$$0 = \frac{k_a a_0}{k_b - k_a} + C$$

and so

$$b(t) = \frac{k_a a_0}{k_b - k_a}[\exp(-k_a t) - \exp(-k_b t)].$$

At large times one of the exponential terms dominates and $b(t)$ decays exponentially at the slower of the rates k_a and k_b. The ratio b/a is given by

$$\frac{b(t)}{a(t)} = \frac{k_a}{k_b - k_a}\{1 - \exp[(k_a - k_b)t]\}.$$

Note that

- the ratio b/a is independent of a_0;
- if $k_a < k_b$ (i.e. if A decays more slowly than B) then the ratio b/a approaches the equilibrium value $k_a/(k_b - k_a)$ at large times.

This analysis has an application in *radiometric dating*. A specific example is uranium–thorium dating, based on the sequence

$$^{234}\mathrm{U} \to {}^{230}\mathrm{Th} \to {}^{226}\mathrm{Ra} \qquad (\mathrm{A} \to \mathrm{B} \to \mathrm{C}).$$

The half-life of ^{234}U is 245000 years, while that of ^{230}Th is 75000 years. The condition $k_a < k_b$ therefore applies. When materials such as coral form they contain uranium but not thorium (which is insoluble). The initial condition $b = 0$ is therefore naturally set at formation. The ratio between ^{230}Th and ^{234}U can be measured in the sample, and the difference from the equilibrium value can be used to deduce the time since formation, i.e.

$$t = -\left(\frac{1}{k_b - k_a}\right) \ln\left[1 - \left(\frac{k_b - k_a}{k_a}\right)\frac{b(t)}{a(t)}\right].$$

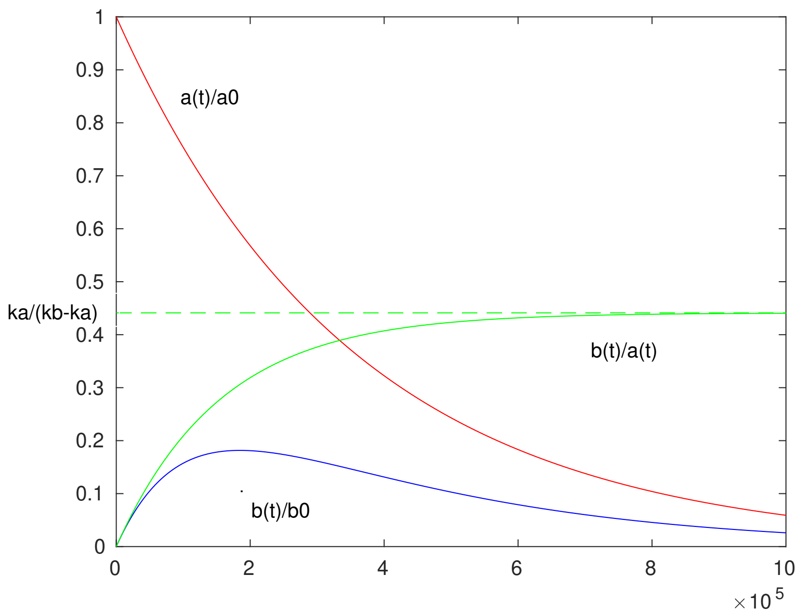


Figure 1: Plot a/a0, b/a0 and b/a versus t over one million years

1.3 Second-order equations

1.3.1 Linear second-order ODEs

Ordinary differential equations of second order are encountered at least as frequently as those of first order. Among other things, they describe oscillatory phenomena.

Second-order ODEs are more difficult to solve and some of the simple categories we have identified for first-order ODEs, e.g. separable equations, do not extend to second-order equations.

In this course we focus on linear equations, which are the ones most frequently encountered and also the easiest to solve.

A *linear* second-order ODE is one of the general form

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + p(x)\frac{\mathrm{d}y}{\mathrm{d}x} + q(x)y = f(x), \tag{1}$$

in which the dependent variable y and its first and second derivatives appear linearly. The term $f(x)$ on the right-hand side is often regarded as a *forcing* term, as it forces the solution to be non-zero.

- If $f(x)$ is absent, the equation is *homogeneous* or *unforced.*
- With the $f(x)$ term it is *inhomogeneous* or *forced.*

We write equation (1) in the symbolic form

$$Ly = f(x),$$

where L is the *linear differential operator*

$$L = \frac{\mathrm{d}^2}{\mathrm{d}x^2} + p(x)\frac{\mathrm{d}}{\mathrm{d}x} + q(x).$$

Like the differential operators $\frac{\mathrm{d}}{\mathrm{d}x}$ and $\frac{\mathrm{d}^2}{\mathrm{d}x^2}$, L can act (operate) on any function of x written to the right of it (provided that it is twice differentiable) and produce another function of x.

L is a *linear* operator because

$$L(\alpha u) = \alpha L u$$

and

$$L(u + v) = Lu + Lv$$

where u and v are any functions of x, and α is any constant.

Exercise: Verify these properties.

If α and β are constants and $w = \alpha u + \beta v$ then

$$\begin{aligned}
L(w) &= L(\alpha u + \beta v) \\
&= \frac{\mathrm{d}^2}{\mathrm{d}x^2}(\alpha u + \beta v) + p(x)\frac{\mathrm{d}}{\mathrm{d}x}(\alpha u + \beta v) + q(x)(\alpha u + \beta v) \\
&= \alpha[\frac{\mathrm{d}^2}{\mathrm{d}x^2}u + p(x)\frac{\mathrm{d}}{\mathrm{d}x}u + q(x)u] + \alpha[\frac{\mathrm{d}^2}{\mathrm{d}x^2}v + p(x)\frac{\mathrm{d}}{\mathrm{d}x}v + q(x)v] \\
&= \alpha L(u) + \beta l(v).
\end{aligned}$$

1.3.2 Principle of superposition

This is an important general principle that applies to any type of linear equation (not necessarily a differential equation). It comes in two parts.

Regarding the homogeneous equation $Ly = 0$:

- if u and v both satisfy the homogeneous equation $Ly = 0$, then αu and $u + v$ also satisfy the homogeneous equation;
- similarly, any *linear combination* (or *superposition*) of solutions of the homogeneous equation, such as $\alpha u + \beta v$, is also a solution.

These properties follow from the definition of a linear operator.

Regarding the inhomogeneous equation $Ly = f$:

- a *particular integral* y_p is any solution of the inhomogeneous equation $Ly = f$ (here 'integral' means 'solution');
- the *complementary function* y_c is the general solution of the homogeneous equation $Ly = 0$;
- the *general solution* of the inhomogeneous equation $Ly = f$ is the sum of the particular integral and the complementary function.

To see why $y_\mathrm{p} + y_\mathrm{c}$ satisfies the inhomogeneous equation:

$$L(y_\mathrm{p} + y_\mathrm{c}) = Ly_\mathrm{p} + Ly_\mathrm{c} = f + 0 = f. \qquad \checkmark$$

To see why this is the general solution:

$$L(y - y_\mathrm{p}) = Ly - Ly_\mathrm{p} = f - f = 0 \quad \Rightarrow \quad y - y_\mathrm{p} = y_\mathrm{c}.$$

1.3.3 Homogeneous equation with constant coefficients

Let $p(x) = 2a$ and $q(x) = b$. The 2 is inserted for later convenience.

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + 2a\frac{\mathrm{d}y}{\mathrm{d}x} + by = 0. \tag{2}$$

The standard method of solving this equation is to try a solution of the form $y = \mathrm{e}^{\lambda x}$, where λ is a constant. Substituting this trial solution into the left-hand side of equation (2), we find

$$\lambda^2\,\mathrm{e}^{\lambda x} + 2a\lambda\,\mathrm{e}^{\lambda x} + b\,\mathrm{e}^{\lambda x} = (\lambda^2 + 2a\lambda + b)\,\mathrm{e}^{\lambda x}.$$

Therefore equation (2) is satisfied by $y = \mathrm{e}^{\lambda x}$ if and only if λ satisfies the quadratic equation

$$\lambda^2 + 2a\lambda + b = 0.$$

We have reduced the differential equation to an algebraic one, known as the *auxiliary equation*. (This procedure works because the derivatives of $\mathrm{e}^{\lambda x}$ are proportional to $\mathrm{e}^{\lambda x}$, and the equation has constant coefficients.) The roots of the auxiliary equation are

$$\lambda = -a \pm \sqrt{a^2 - b}.$$

Call these roots λ_1 and λ_2, and assume for now that they are distinct. Therefore, according to the principle of superposition,

$$y = C_1\,\mathrm{e}^{\lambda_1 x} + C_2\,\mathrm{e}^{\lambda_2 x}$$

is a solution of equation (2). It contains two arbitrary constants and is in fact the general solution. Note that, if one of the roots of the auxiliary equation is $\lambda = 0$, then the corresponding solution is just a constant.

1.3.4 Case of complex roots

The roots of the auxiliary equation may be complex, even if a and b are real numbers and the ODE is a real equation. This happens if $a^2 - b < 0$. In this case the roots can be written as

$$\lambda = -a \pm \mathrm{i}\sqrt{b-a^2}.$$

with $b - a^2 > 0$. It is still true that $\mathrm{e}^{\lambda_1 x}$ and $\mathrm{e}^{\lambda_2 x}$ are solutions of the ODE, but neither of these is a real function. The real and imaginary parts of $\mathrm{e}^{\lambda_1 x}$ are

$$\mathrm{e}^{-ax}\cos\left(\sqrt{b-a^2}\,x\right), \qquad \mathrm{e}^{-ax}\sin\left(\sqrt{b-a^2}\,x\right),$$

and these both satisfy the ODE. (The real and imaginary parts of $\mathrm{e}^{\lambda_2 x}$ are the same as those of $\mathrm{e}^{\lambda_1 x}$ except that the imaginary part has the opposite sign. This is because λ_2 is the complex conjugate of λ_1.)

The general solution can then be written in the alternative, and manifestly real, form

$$y = \mathrm{e}^{-ax}\left[D_1\cos\left(\sqrt{b-a^2}\,x\right) + D_2\sin\left(\sqrt{b-a^2}\,x\right)\right].$$

In the special case $a = 0$, we have purely oscillatory solutions with no exponential growth or decay:

$$y = D_1\cos\left(\sqrt{b}\,x\right) + D_2\sin\left(\sqrt{b}\,x\right).$$

This occurs when the ODE is the familiar one that describes simple harmonic motion,

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + by = 0,$$

with $b > 0$.

1.3.5 Case of equal roots

The roots of the auxiliary equation may be equal. This happens if $a^2 - b = 0$, in which case both roots are equal to $-a$. We then have only one solution of the form $y = \mathrm{e}^{\lambda x}$, with $\lambda = -a$. This is sometimes called the *degenerate case*.

It turns out that a second solution in this case is $y = x\,\mathrm{e}^{\lambda x}$. With this trial solution,

$$\begin{aligned}\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + 2a\frac{\mathrm{d}y}{\mathrm{d}x} + by &= 2\lambda\,\mathrm{e}^{\lambda x} + \lambda^2 x\,\mathrm{e}^{\lambda x} + 2a(\mathrm{e}^{\lambda x} + \lambda x\,\mathrm{e}^{\lambda x}) + bx\,\mathrm{e}^{\lambda x} \\ &= (\lambda^2 + 2a\lambda + b)x\,\mathrm{e}^{\lambda x} + (2\lambda + 2a)\,\mathrm{e}^{\lambda x}.\end{aligned}$$

So $y = x\,\mathrm{e}^{\lambda x}$ is a solution of equation (2) if and only if $\lambda^2 + 2a\lambda + b = 0$ *and* $2\lambda + 2a = 0$. These are precisely the conditions for the auxiliary equation to have a repeated root $\lambda = -a$.

The general solution in this case is

$$y = (C_1 + C_2 x)\,\mathrm{e}^{\lambda x},$$

with $\lambda = -a$.

Note that, if the roots of the auxiliary equation are both $\lambda = 0$, then the general solution is $y = C_1 + C_2 x$. In this case the ODE is just $\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = 0$.

1.3.6 Initial conditions or boundary conditions

For first-order ODEs, we found that a general solution contains one arbitrary constant. The value of that constant can be determined by supplying an initial condition, i.e. the value of y at a particular value of x.
For second-order ODEs, the general solution contains two arbitrary constants. Their values can be determined by supplying two additional pieces of information. These could be

- *initial conditions*: the values of y *and* $\mathrm{d}y/\mathrm{d}x$ at one value of x;

- *boundary conditions*: conditions that apply at two different values of x. The most common type of boundary conditions for second-order ODEs are those that specify the value of y *or* $\mathrm{d}y/\mathrm{d}x$ at two different values of x. (More generally, they might specify the value of some combination of y and $\mathrm{d}y/\mathrm{d}x$.)

The names *initial conditions* and *boundary conditions* suggest that time is involved in the first case and space in the second case. Indeed, these are the ways that physical problems are usually posed.

- In an *initial-value problem* we are given complete information about a system at some initial time and we calculate its evolution into the future. (In some cases the past evolution can also be calculated.)

- In a *boundary-value problem* time is not involved, usually because we are considering a steady state of some kind. The spatial boundaries of the system impose constraints at different locations that must be satisfied simultaneously.

1.3.7 Inhomogeneous equation with constant coefficients

We now turn to the problem of solving an inhomogeneous second-order ODE with constant coefficients. This is an equation of the general form

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + 2a\frac{\mathrm{d}y}{\mathrm{d}x} + by = f(x), \tag{3}$$

where $f(x)$ is a given function. We write this equation in the symbolic form

$$Ly = f(x)$$

using the linear differential operator

$$L = \frac{\mathrm{d}^2}{\mathrm{d}x^2} + 2a\frac{\mathrm{d}}{\mathrm{d}x} + b.$$

According to the principle of superposition, to find the general solution of equation (3) we find any solution y_{p} (a *particular integral*) and add it to the general solution y_{c} of the homogeneous equation $Ly = 0$ (the *complementary function*). The result contains two arbitrary constants, which come from the complementary function. They can be determined by applying initial or boundary conditions.

We have just seen how to find the general solution of a homogeneous linear second-order ODE with constant coefficients. Therefore we know how to find the complementary function in this case.

The usual method for finding a particular integral is to try a solution of a similar form to the given function $f(x)$ but containing some adjustable parameters. The trial solution is substituted into the equation and, if it works, the values of the parameters can be found.

To understand how this works, consider the action of the operator

$$L = \frac{\mathrm{d}^2}{\mathrm{d}x^2} + 2a\frac{\mathrm{d}}{\mathrm{d}x} + b$$

on various functions:

$$L\,1 = b$$

$$L\,x = 2a + bx$$

$$L\,x^2 = 2 + 4ax + bx^2$$

$$L\,x^n = n(n-1)x^{n-2} + 2anx^{n-1} + bx^n$$

$$L\,\mathrm{e}^{kx} = (k^2 + 2ak + b)\,\mathrm{e}^{kx}$$

$$L\,\cos(kx) = (-k^2 + b)\cos(kx) - 2ak\sin(kx)$$

$$L\,\sin(kx) = (-k^2 + b)\sin(kx) + 2ak\cos(kx).$$

We see from this that

- when L acts on a polynomial function of degree n, it produces another polynomial function of degree n (unless $b = 0$);
- when L acts on an exponential function, it produces a multiple of that exponential function;
- when L acts on a linear combination of cosine and sine functions, it produces another linear combination of those functions.

Here is another aspect of the principle of superposition:

- if $Ly_1 = f_1$ and $Ly_2 = f_2$, then $L(y_1 + y_2) = f_1 + f_2$;
- if $Ly_1 = f_1$ then $L(\alpha y_1) = \alpha f_1$.

In other words, *the solution depends linearly on the forcing.* We can use this property to construct particular integrals. Exercise: Find a particular integral of $y' + 3y' + y = x^3 + x^2 + x + 1$.

For example, if $f(x)$ is a polynomial function of degree n, i.e.

$$f(x) = c_n x^n + c_{n-1} x^{n-1} + \cdots + c_1 x + c_0,$$

then we would try a particular integral of a similar form, i.e.

$$y_{\mathrm{p}} = d_n x^n + d_{n-1} x^{n-1} + \cdots + d_1 x + d_0.$$

Substituting this trial solution into $Ly_{\mathrm{p}} = f(x)$ and comparing the coefficients of each power of x, we would find the values of $d_n, d_{n-1}, \ldots, d_1, d_0$.

However, this solution fails if $b = 0$, i.e. if 0 is a root of the auxiliary equation, because then Ly_{p} is a polynomial of degree $< n$. The problem in this case is that the constant polynomial 1 is a solution of the homogeneous equation and so the constant term d_0 is not a useful part of the particular integral. In order to match all the coefficients in this case, we need to introduce a different degree of freedom and let

$$y_{\mathrm{p}} = d_{n+1} x^{n+1} + d_n x^n + \cdots + d_1 x.$$

Even this fails if $a = b = 0$, i.e. if 0 is a repeated root of the auxiliary equation. Then both 1 and x are solutions of the homogeneous equation and so neither d_0 nor d_1 is useful. In this case we need to let

$$y_{\mathrm{p}} = d_{n+2}x^{n+2} + d_{n+1}x^{n+1} + \cdots + d_2x^2.$$

(In fact, in this case, the ODE is just $y'' = f$ and can be solved by integrating twice, which does indeed give a polynomial of degree $n + 2$.)

Suppose now we wish to solve $Ly_{\mathrm{p}} = \mathrm{e}^{kx}$. We have seen that

$$L\,\mathrm{e}^{kx} = (k^2 + 2ak + b)\,\mathrm{e}^{kx}. \tag{4}$$

A solution of $Ly_{\mathrm{p}} = \mathrm{e}^{kx}$ is therefore $y_{\mathrm{p}} = (k^2 + 2ak + b)^{-1}\,\mathrm{e}^{kx}$.

However, this solution fails if $k^2 + 2ak + b = 0$, i.e. if k is a root of the auxiliary equation. In this case e^{kx} is a solution of the homogeneous equation and so it cannot be a particular integral. Now consider

$$L(x\,\mathrm{e}^{kx}) = (k^2 + 2ak + b)x\,\mathrm{e}^{kx} + (2k + 2a)\,\mathrm{e}^{kx}. \tag{5}$$

If $k^2 + 2ak + b = 0$ then we see that a solution of $Ly_{\mathrm{p}} = \mathrm{e}^{kx}$ is $y_{\mathrm{p}} = (2k + 2a)^{-1}x\,\mathrm{e}^{kx}$.

Even this fails if $2k + 2a = 0$, i.e. if k is a repeated root of the auxiliary equation. Then both e^{kx} and $x\,\mathrm{e}^{kx}$ are solutions of the homogeneous equation. So consider

$$L(x^2\,\mathrm{e}^{kx}) = (k^2 + 2ak + b)x^2\,\mathrm{e}^{kx} + 2(2k + 2a)x\,\mathrm{e}^{kx} + 2\,\mathrm{e}^{kx}.$$

If $k^2 + 2ak + b = 0$ *and* $2k + 2a = 0$ then we see that a solution of $Ly_{\mathrm{p}} = \mathrm{e}^{kx}$ is $y_{\mathrm{p}} = \frac{1}{2}x^2\,\mathrm{e}^{kx}$.

Summary on particular integrals:

- If $f(x)$ is a polynomial, try a general polynomial of the same degree for y_{p}. But if 1 is a solution of the homogeneous equation, try one degree higher. If 1 and x are both solutions of the homogeneous equation, try two degrees higher.
- If $f(x) = c\,\mathrm{e}^{kx}$, try $y_{\mathrm{p}} = d\,\mathrm{e}^{kx}$. But if e^{kx} is a solution of the homogeneous equation, try $y_{\mathrm{p}} = dx\,\mathrm{e}^{kx}$ instead. If $x\,\mathrm{e}^{kx}$ is also a solution of the homogeneous equation, try $y_{\mathrm{p}} = dx^2\,\mathrm{e}^{kx}$ instead.
- If $f(x) = c_1 \cos(kx) + c_2 \sin(kx)$, try $y_{\mathrm{p}} = d_1 \cos(kx) + d_2 \sin(kx)$. But if $\cos(kx)$ and $\sin(kx)$ are solutions of the homogeneous equation, try $y_{\mathrm{p}} = d_1 x \cos(kx) + d_2 x \sin(kx)$ instead.

Example: Find the solution of

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + 5\frac{\mathrm{d}y}{\mathrm{d}x} + 6y = 6x$$

subject to the initial conditions $y = 1$ and $\mathrm{d}y/\mathrm{d}x = 0$ at $x = 0$.

First find the complementary function by solving the auxiliary equation

$$\lambda^2 + 5\lambda + 6 = 0.$$

This has distinct roots $\lambda_1 = -2$ and $\lambda_2 = -3$. The complementary function is therefore

$$y_\mathrm{c} = C_1\,\mathrm{e}^{-2x} + C_2\,\mathrm{e}^{-3x}.$$

Since the right-hand side of the ODE is $6x$, we try a particular integral of the form $y_\mathrm{p} = cx + d$. This satisfies the ODE if

$$0 + 5c + 6(cx + d) = 6x,$$

i.e. if $c = 1$ and $d = -5/6$. Thus

$$y_\mathrm{p} = x - \frac{5}{6}.$$

The general solution is then

$$y = y_\mathrm{c} + y_\mathrm{p} = C_1\,\mathrm{e}^{-2x} + C_2\,\mathrm{e}^{-3x} + x - \frac{5}{6}.$$

Now apply the initial conditions:

$$1 = C_1 + C_2 - \frac{5}{6},$$

$$0 = -2C_1 - 3C_2 + 1.$$

Thus $C_1 = 9/2$ and $C_2 = -8/3$. The solution is therefore

$$y = \frac{9}{2}\,\mathrm{e}^{-2x} - \frac{8}{3}\,\mathrm{e}^{-3x} + x - \frac{5}{6}.$$

1.4 Connections

The standard methods we have seen for first- and second-order ODEs seem entirely different. Let us now see some connections between them.

We solved the inhomogeneous first-order linear ODE

$$\frac{\mathrm{d}y}{\mathrm{d}x} + p(x)y = f(x)$$

by the integrating-factor method. But it can also be solved by finding a particular integral and the complementary function. The complementary function is the general solution of the homogeneous equation,

$$y_{\mathrm{c}} = \exp\left[-\int p(x)\,\mathrm{d}x\right],$$

which we can also write as

$$y_{\mathrm{c}} = \frac{C}{\mu(x)}$$

where $\mu(x)$ is the integrating factor that we identified previously, and C is an arbitrary constant. Supposing that we can find a particular integral y_{p} by intelligent guesswork, the general solution is

$$y = y_{\mathrm{p}} + \frac{C}{\mu(x)}. \tag{1}$$

Let's compare this with the general solution we found by the integrating-factor method:

$$y(x) = \frac{1}{\mu(x)} \int^x \mu(t) f(t)\, \mathrm{d}t.$$

The indefinite integral contains an arbitrary additive constant, which agrees with the appearance of C in equation (1). So we see that the methods give the same answer, and the integrating-factor method provides a systematic way to find a particular integral.

Now let's consider an inhomogeneous second-order linear ODE with constant coefficients,

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + 2a \frac{\mathrm{d}y}{\mathrm{d}x} + by = f(x).$$

Let λ_1 and λ_2 be the roots of the auxiliary equation,

$$\lambda^2 + 2a\lambda + b = (\lambda - \lambda_1)(\lambda - \lambda_2) = 0.$$

Then we can factorize the differential operator and write the ODE as

$$\left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_1\right)\left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_2\right) y = f(x). \tag{2}$$

Exercise: Check that this expands out correctly. Note that it doesn't matter in which order the factors are written; differential operators don't usually commute with each other, but they do here because of the constant coefficients.

To solve this equation, let us introduce a function $z(x)$ defined by

$$\left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_2\right) y = z(x). \tag{3}$$

Then the ODE (2) states that

$$\left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_1\right) z = f(x). \tag{4}$$

In this way we have replaced the second-order ODE (2) with two coupled first-order ODEs (3) and (4). They are still linear and have constant coefficients. They can each be solved by the integrating-factor method: first (4) for $z(x)$, then (3) for $y(x)$. Provided that we can do the integrals, we don't need to try to guess a particular integral. There is therefore a systematic way to find a particular integral. However, the standard method, based on guessing the form of the particular integral and matching coefficients, is often easier.

1.5 Application: the damped oscillator

1.5.1 Mathematical model

A standard form of the equation for a forced, damped oscillator is

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + 2\gamma\frac{\mathrm{d}x}{\mathrm{d}t} + \omega_0^2 x = f(t). \tag{1}$$

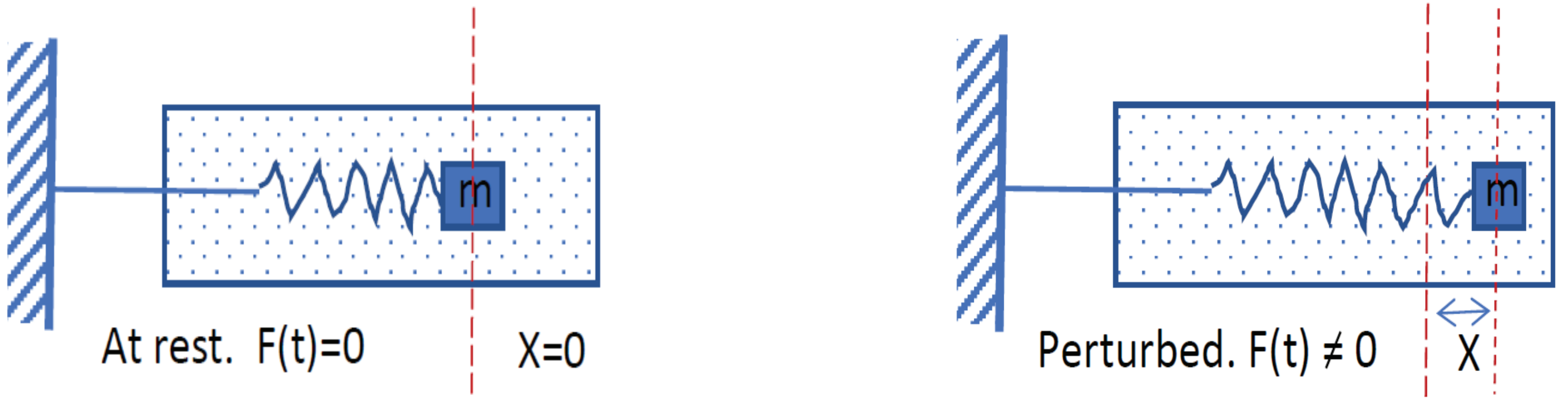
This equation arises in many contexts, most commonly in mechanical systems where it derives from Newton's second law, but also in the analysis of electrical circuits and other systems capable of oscillation.
We'll call

- $x(t)$ the displacement,
- $\mathrm{d}x/\mathrm{d}t$ the velocity,
- $\mathrm{d}^2x/\mathrm{d}t^2$ the acceleration,
- γ the damping coefficient,
- ω_0 the natural frequency,
- $f(t)$ the force.

We'll assume that $\omega_0 > 0$ and $\gamma \geqslant 0$. (Different notation for γ and ω_0 may be encountered.)
Note that we are solving for $x(t)$, not for $y(x)$, so the notation has to be adapted. As is often the case, the independent variable is time t.

1.5.2 Example of physical derivation

A body of mass m is connected to a spring of spring constant k. It also experiences a resistive force proportional to its velocity (with constant of proportionality μ) and a time-dependent external force $F(t)$.



In the absence of the external force, the system has an equilibrium position. Let $x(t)$ be the displacement from this position.

$$F_s = -kx \qquad F_d = -\mu v = -\mu \frac{dx}{dt}$$

Then Newton's second law gives:

mass $\times$ acceleration $=$ force

$$m\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = F_s + F_d + F(t) = -kx - \mu\frac{\mathrm{d}x}{\mathrm{d}t} + F(t).$$

We can rearrange this in the form

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2}+\frac{\mu}{m}\frac{\mathrm{d}x}{\mathrm{d}t}+\frac{k}{m}x=\frac{F(t)}{m}.$$

This agrees with our standard form, equation (1), if we identify $\gamma=\mu/(2m)$, $\omega_0=\sqrt{k/m}$ and $f(t)=F(t)/m$. (This means that the 'force' $f(t)$ in our standard model is really the force per unit mass.)

1.5.3 The energy equation

(This is not a way of solving the equation of motion, but assists in the interpretation of the solutions.)

Multiplying equation (1) by $\mathrm{d}x/\mathrm{d}t$, we have

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2}\frac{\mathrm{d}x}{\mathrm{d}t}+2\gamma\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2+\omega_0^2x\frac{\mathrm{d}x}{\mathrm{d}t}=f(t)\frac{\mathrm{d}x}{\mathrm{d}t},$$

which implies

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{1}{2}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2+\frac{1}{2}\omega_0^2x^2\right]=-2\gamma\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2+f(t)\frac{\mathrm{d}x}{\mathrm{d}t}.$$

The expression inside the square brackets represents the energy of the oscillating system (in the mechanical system it is the kinetic energy (E_k) plus potential energy (E_p) per unit mass).

$$v=\frac{\mathrm{d}x}{\mathrm{d}t}\qquad\frac{1}{2}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2=\frac{v^2}{2}=E_k$$

$$\frac{1}{2}\omega_0^2 x^2 = \frac{kx^2}{2m} = \frac{1}{m}.\frac{kx^2}{2} = E_p$$

The equation states that the energy changes as a result of the damping force and the external force. Damping always reduces the energy (except when the damping coefficient or the velocity is zero). The external force can do either positive or negative work on the system, depending on whether the force has the same sign as the velocity or the opposite sign.

If there is no external force then the energy can only decay and the velocity must tend to zero (except when the damping coefficient vanishes). Sustained motion is possible if there is an external force. In fact the motion cannot cease if the force is time-dependent.

1.5.4 Free oscillations

In the absence of external forcing, we have

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + 2\gamma\frac{\mathrm{d}x}{\mathrm{d}t} + \omega_0^2 x = 0.$$

This is a homogeneous linear second-order ODE with constant coefficients.

To complete the problem we need two initial conditions, i.e. the values of x and $\mathrm{d}x/\mathrm{d}t$ at some initial time t_0. A simple example of initial conditions is

$$x = a \quad \text{and} \quad \frac{\mathrm{d}x}{\mathrm{d}t} = 0 \quad \text{at} \quad t = 0, \tag{2}$$

which means that the system is released from rest at $t = 0$, with initial displacement a.

We seek solutions of the form $x = \mathrm{e}^{\lambda t}$ and obtain the auxiliary equation

$$\lambda^2 + 2\gamma\lambda + \omega_0^2 = 0,$$

which has roots

$$\lambda = -\gamma \pm \sqrt{\gamma^2 - \omega_0^2}\,.$$

There are four different cases to consider, depending on the magnitude of the damping coefficient γ:

- $\gamma = 0$ (*no damping*): two imaginary roots;
- $0 < \gamma < \omega_0$ (*weak damping*): two complex roots (negative real parts, opposite imaginary parts);
- $\gamma = \omega_0$ (*critical damping*): repeated (negative real) root;
- $\gamma > \omega_0$ (*strong damping*): two (negative) real roots.

1.5.5 Case of no damping, $\gamma = 0$

The general solution of the homogeneous equation in this case is

$$x = D_1 \cos(\omega_0 t) + D_2 \sin(\omega_0 t).$$

This corresponds to *simple harmonic motion.* In this case the ODE simplifies to $\mathrm{d}^2x/\mathrm{d}t^2 + \omega_0^2 x = 0$. Therefore ω_0 is the natural frequency of the system in the absence of damping.

- By 'frequency' we really mean angular frequency. This is a common usage. The true frequency is $\omega_0/(2\pi)$.

The constants D_1 and D_2 can be found by applying the initial conditions. In the example of equation (2), we find

$$a = D_1 \quad \text{and} \quad 0 = \omega_0 D_2.$$

The specific solution in this case is therefore

$$x = a\cos(\omega_0 t).$$

See blue line in Figure 2: $\frac{\gamma}{\omega} = 0$

1.5.6 Case of weak damping, $0 < \gamma < \omega_0$

(Also known as 'underdamped'.)

The general solution is

$$x = \mathrm{e}^{-\gamma t}[D_1 \cos(\nu t) + D_2 \sin(\nu t)], \qquad \nu = \sqrt{\omega_0^2 - \gamma^2}\,,$$

and describes decaying oscillations. The frequency ν is less than the natural frequency in the absence of damping and the amplitude of the oscillations decays proportional to $\mathrm{e}^{-\gamma t}$.

Applying the initial conditions (2), we find

$$a = D_1 \quad \text{and} \quad 0 = -\gamma D_1 + \nu D_2.$$

The specific solution in this case is therefore

$$x = a\,\mathrm{e}^{-\gamma t}\left[\cos(\nu t) + \frac{\gamma}{\nu}\sin(\nu t)\right], \qquad \nu = \sqrt{\omega_0^2 - \gamma^2}\,.$$

Note that, as the damping coefficient γ is increased towards its critical value ω_0, the frequency of the oscillations tends to zero. If γ is only slightly less than ω_0, there is no time to observe an oscillation before it is damped out.

See magenta line in Figure 2: $\frac{\gamma}{\omega} = 0.2$

1.5.7 Case of critical damping, $\gamma = \omega_0$

This is the case of equal roots of the auxiliary equation. The general solution is

$$x = (C_1 + C_2 t)\,\mathrm{e}^{-\gamma t}.$$

The constants C_1 and C_2 can be found by applying the initial conditions. In the example of equation (2), we find

$$a = C_1 \quad \text{and} \quad 0 = -\gamma C_1 + C_2.$$

The specific solution in this case is therefore

$$x = a(1 + \gamma t)\,\mathrm{e}^{-\gamma t}.$$

There are no oscillations in this case. In fact the velocity

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -a\gamma^2 t\,\mathrm{e}^{-\gamma t}$$

never changes sign, so the displacement tends monotonically to zero.

See green line in Figure 2: $\frac{\gamma}{\omega} = 1$

1.5.8 Case of strong damping, $\gamma > \omega_0$

(Also known as 'overdamped'.)

The general solution is

$$x = C_1\, \mathrm{e}^{\lambda_1 t} + C_2\, \mathrm{e}^{\lambda_2 t}, \qquad \lambda_{1,2} = -\gamma \pm \sqrt{\gamma^2 - \omega_0^2}\,.$$

Applying the initial conditions (2), we find

$$a = C_1 + C_2 \quad \text{and} \quad 0 = C_1\lambda_1 + C_2\lambda_2,$$

and so

$$C_1 = \frac{\lambda_2 a}{\lambda_2 - \lambda_1}, \qquad C_2 = -\frac{\lambda_1 a}{\lambda_2 - \lambda_1}.$$

The specific solution in this case is therefore

$$x = \frac{a(\lambda_2\, \mathrm{e}^{\lambda_1 t} - \lambda_1\, \mathrm{e}^{\lambda_2 t})}{\lambda_2 - \lambda_1}.$$

(In fact this form of the solution is valid for no, weak or strong damping. In the cases of no or weak damping the exponentials are complex and it is more natural to write the solution in a manifestly real form using cos and sin as we did earlier. The solution in the case of critical damping can also be obtained from this expression by considering the limit $\gamma \to \omega_0$ and applying l'Hôpital's rule.)

See brown line in Figure 2: $\frac{\gamma}{\omega} = 2$

Another way to write this solution is (**exercise**)

$$x = a\,\mathrm{e}^{-\gamma t}\left[\cosh(\sigma t) + \frac{\gamma}{\sigma}\sinh(\sigma t)\right], \qquad \sigma = \sqrt{\gamma^2 - \omega_0^2}\,.$$

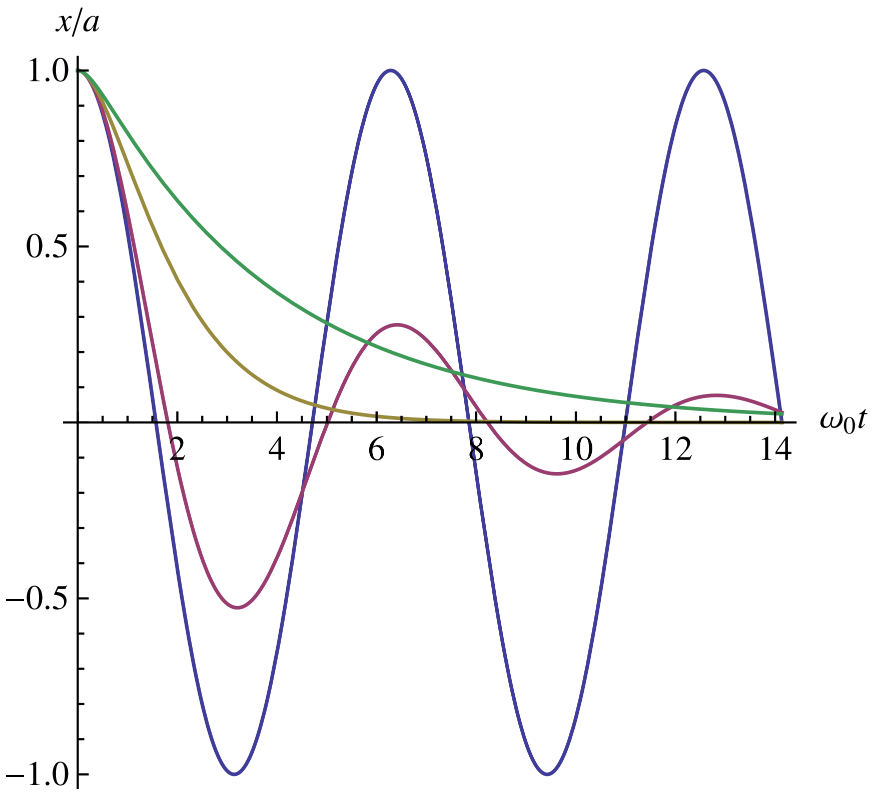


Figure 2: Free oscillations starting from $x(0) = a$, $\dot{x}(0) = 0$ with $\gamma/\omega_0 = 0$, 0.2, 1 and 2.

1.5.9 Forced oscillations and resonance

We now consider the response of the forced, damped oscillator

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + 2\gamma\frac{\mathrm{d}x}{\mathrm{d}t} + \omega_0^2 x = f(t) \tag{3}$$

in the case of harmonic forcing,

$$f(t) = f_0\cos(\omega t).$$

Here $\omega > 0$ is the forcing frequency, which may differ from the natural frequency ω_0 of the oscillator, and $f_0 > 0$ is the forcing amplitude. (We have chosen a cosine but the starting phase of the forcing is not important.)

We apply the standard method of particular integral + complementary function. We have already discussed the complementary functions in the four cases of no, weak, critical and strong damping.

Try a particular integral of the form

$$x_\mathrm{p} = B\cos(\omega t) + C\sin(\omega t) \tag{4}$$

and substitute into equation (3):

$$\begin{aligned}&-\omega^2[B\cos(\omega t) + C\sin(\omega t)] + 2\gamma\omega[-B\sin(\omega t) + C\cos(\omega t)]\\ &\qquad + \omega_0^2[B\cos(\omega t) + C\sin(\omega t)] = f_0\cos(\omega t).\end{aligned}$$

Compare coefficients of $\cos(\omega t)$ and $\sin(\omega t)$:

$$(\omega_0^2-\omega^2)B+2\gamma\omega C=f_0,$$

$$(\omega_0^2-\omega^2)C-2\gamma\omega B=0.$$

Solve to obtain

$$B=\frac{(\omega_0^2-\omega^2)f_0}{(\omega_0^2-\omega^2)^2+(2\gamma\omega)^2},$$

$$C=\frac{(2\gamma\omega)f_0}{(\omega_0^2-\omega^2)^2+(2\gamma\omega)^2}.$$

The particular integral (4) represents an oscillation of constant amplitude and with the same frequency as the forcing, although not generally in phase with it.

If $\gamma>0$ then the complementary function ultimately decays and the particular integral represents the long-term (long-time) response of the oscillator.

Now

$$x_{\mathrm{p}}=B\cos(\omega t)+C\sin(\omega t)=A\cos(\omega t-\phi)$$

where

$$B=A\cos\phi,\qquad C=A\sin\phi,$$

i.e.

$$A = \sqrt{B^2 + C^2}, \qquad \phi = \arctan\left(\frac{C}{B}\right).$$

Note that A represents the amplitude of the oscillation, while ϕ represents its phase shift relative to the forcing.

- $\phi > 0$ means that the oscillation lags behind the forcing;
- $\phi < 0$ means that the oscillation leads the forcing.

In our case, the amplitude of the oscillation is

$$A = \sqrt{B^2 + C^2} = \frac{f_0}{\sqrt{(\omega_0^2 - \omega^2)^2 + (2\gamma\omega)^2}}.$$

The tangent of the phase shift of the oscillation is

$$\tan\phi = \frac{C}{B} = \frac{2\gamma\omega}{\omega_0^2 - \omega^2}.$$

Note that

- the oscillation amplitude is proportional to the forcing amplitude, $A \propto f_0$, because this is a linear system;

- the amplitude is largest if the forcing frequency is close to the natural frequency and the damping coefficient is small;

- as the forcing frequency increases through the natural frequency, A reaches a peak and the phase lag ϕ increases through $\pi/2$.

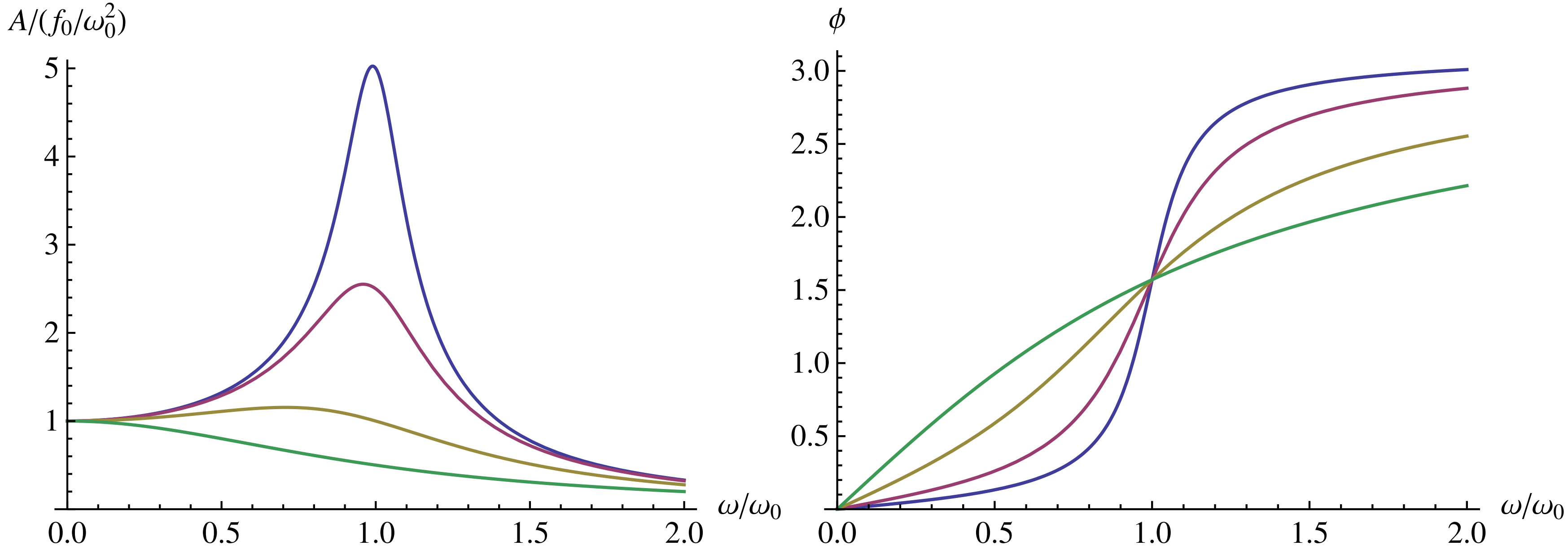


Figure 3: Amplitude and phase (in radians) of forced oscillations for $\gamma/\omega_0 = 0.1$, 0.2, 1 and 2.

This behaviour, in which a damped oscillator exhibits a large-amplitude response when forced close to its natural frequency, is known as *resonance*. The effect is strongest when the damping is very weak ($\gamma \ll \omega_0$).

1.5.10 Undamped resonance

There is one case in which the particular integral (4) does not work: when there is no damping ($\gamma = 0$) and the forcing frequency matches the natural frequency ($\omega = \omega_0$). This special problem is

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + \omega_0^2 x = f_0 \cos(\omega_0 t). \tag{5}$$

In this case the forcing function is a solution of the homogeneous equation, so we try a particular integral of the form

$$x_\mathrm{p} = Bt\cos(\omega_0 t) + Ct\sin(\omega_0 t)$$

and substitute into equation (5):

$$-2B\omega_0 \sin(\omega_0 t) + 2C\omega_0 \cos(\omega_0 t) = f_0 \cos(\omega_0 t).$$

This works if

$$B = 0, \qquad C = \frac{f_0}{2\omega_0},$$

and so

$$x_\mathrm{p} = \frac{f_0}{2\omega_0} t \sin(\omega_0 t).$$

This represents an oscillation that lags the forcing by $\pi/2$ and whose amplitude grows linearly in time. This is another manifestation of *resonance*. Energy is continually added to the oscillator by the external force and there is no means of dissipation.

In a real physical system the amplitude will be limited either by damping or by nonlinear effects not represented in equation (1).

1.5.11 Transients

In damped cases ($\gamma > 0$) the complementary function always tends to zero as $t \to \infty$. The long-term response is described by the particular integral, which is an oscillation of constant amplitude with the same frequency as the forcing. It contains no arbitrary constants and does not depend on the initial conditions.

The complementary function contains two arbitrary constants and does depend on the initial conditions. The fact that it decays means that the initial conditions are ultimately forgotten. The complementary function therefore describes a *transient response*.

In the undamped case ($\gamma = 0$) the complementary function is an undamped oscillation at the natural frequency ω_0. It does not decay and the initial conditions are never forgotten.

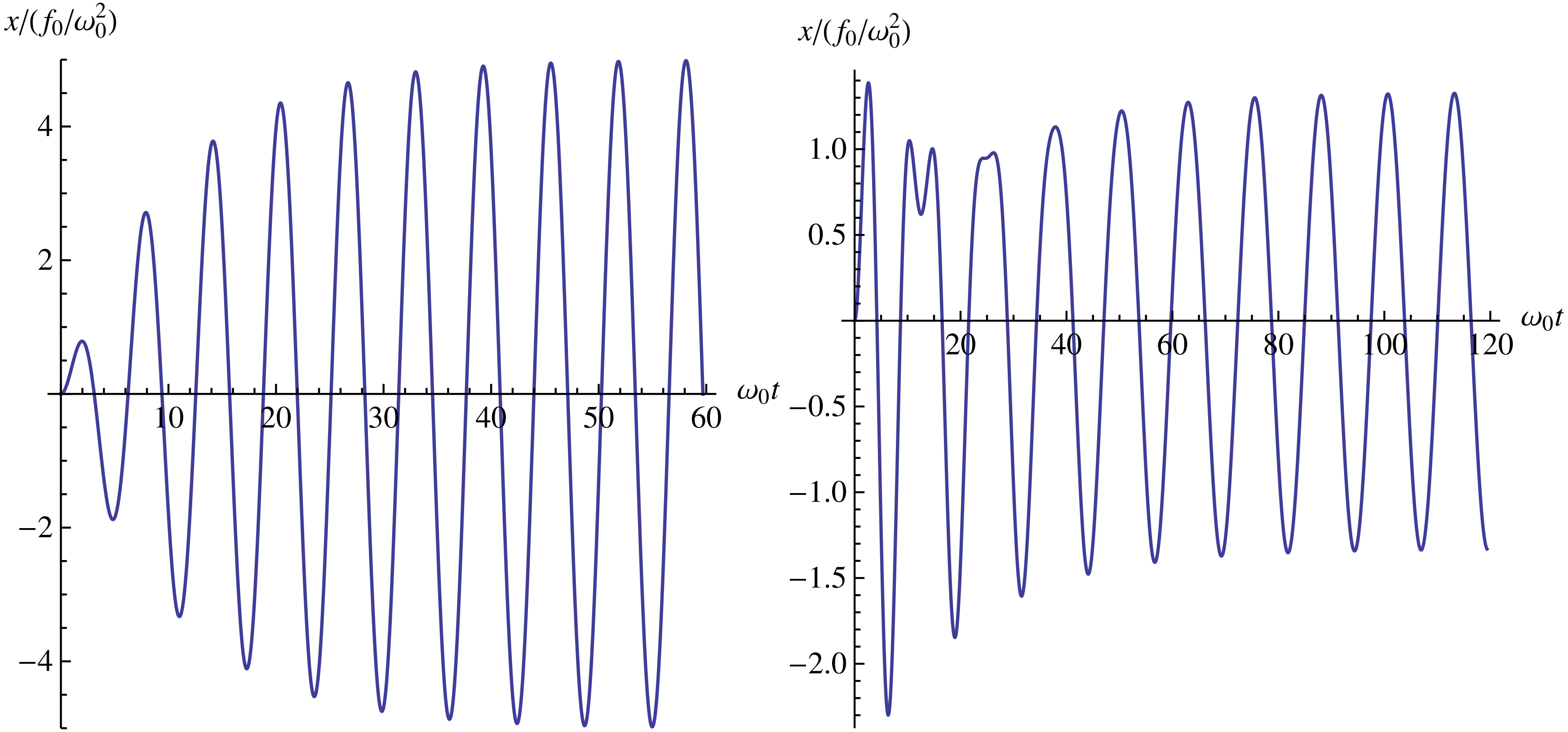


Figure 4: Transients leading to steady oscillation. Top: $\gamma/\omega_0 = 0.1$, $\omega = \omega_0$, $x(0) = \dot{x}(0) = 0$. Bottom: $\gamma/\omega_0 = 0.05$, $\omega = \omega_0/2$, $x(0) = \dot{x}(0) = 0$.