PART I

ORDINARY DIFFERENTIAL EQUATIONS AND THEIR APPROXIMATIONS

CHAPTER 1

FIRST-ORDER SCALAR EQUATIONS

In this chapter we study the basic properties of first-order scalar ordinary differential equations and their solutions. The first and larger part is devoted to linear equations and various of their basic properties, such as the principle of superposition, Duhamel's principle, and the concept of stability. In the second part we study briefly nonlinear scalar equations, emphasizing the new behaviors that emerge, and introduce the very useful technique known as the principle of linearization. The scalar equations and their properties are crucial to an understanding of the behavior of more general differential equations.

1.1 Constant coefficient linear equations

Consider a complex function y of a real variable t. One of the simplest differential equations that y can obey is given by

$$\frac{dy}{dt} = \lambda y, \tag{1.1}$$

Introduction to Numerical Methods for Time Dependent Differential Equations, First Edition.
By Heinz-O. Kreiss and Omar E. Ortiz.

where $\lambda \in \mathbb{C}$ is constant. We want to solve the initial value problem, that is, we want to determine a solution for $t \geq 0$ with given initial value

$$y(0) = y_0 \in \mathbb{C}. \tag{1.2}$$

Clearly,

$$y(t) = e^{\lambda t} y_0 \tag{1.3}$$

is the solution of (1.1), (1.2). Let us discuss the solution under different assumptions for the λ constant. In Figures 1.1 and 1.2 we illustrate the solution for $y_0 = 1 + 0.4i$ and different values of λ.

1. $\lambda \in \mathbb{R}, \lambda < 0$. In this case both the real and imaginary parts of the solution decay exponentially. If $|\lambda| \gg 1$, the decay is very rapid.

2. $\lambda \in \mathbb{R}, \lambda > 0$. The solution grows exponentially. The growth is slow if $|\lambda| \ll 1$. For example, for $\lambda = 0.01$ we have, by Taylor expansion,

   $$e^{0.01t} = 1 + 0.01t + 0.0001t^2 + \cdots .$$

   On the other hand, if $\lambda \gg 1$, the solution grows very rapidly.

3. $\lambda = i\xi, \xi \in \mathbb{R}$. In this case, the *amplitude* $|y(t)|$ of the solution is constant in time,

   $$|y(t)| = |e^{i\xi t}||y_0| = |y_0|.$$

   If the complex initial data y_0 is written as

   $$y_0 = y_{0\mathcal{R}} + i y_{0\mathcal{I}}, \quad y_{0\mathcal{R}}, y_{0\mathcal{I}} \in \mathbb{R},$$

   the solution is

   $$\begin{aligned} y(t) &= e^{i\xi t} y_0 \\ &= \big(\cos(\xi t) + i \sin(\xi t)\big)(y_{0\mathcal{R}} + i y_{0\mathcal{I}}) \\ &= \big(y_{0\mathcal{R}} \cos(\xi t) - y_{0\mathcal{I}} \sin(\xi t)\big) + i\big(y_{0\mathcal{R}} \sin(\xi t) + y_{0\mathcal{I}} \cos(\xi t)\big), \end{aligned}$$

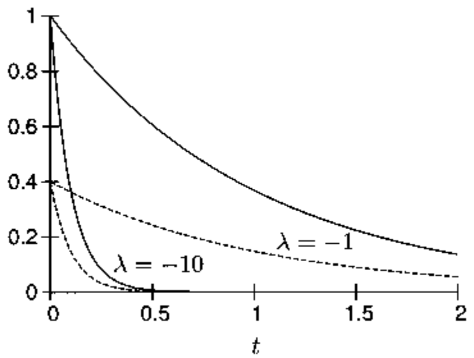

Figure 1.1 Exponentially decaying solutions. Re y shown as solid lines and Im y as dashed lines.

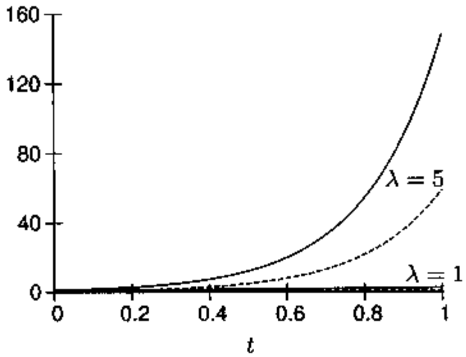

Figure 1.2 Exponentially growing solutions. $\operatorname{Re} y$ is shown as solid line and $\operatorname{Im} y$ as a dashed line.

which defines the real part $y_{\mathcal{R}}(t)$ and imaginary part $y_{\mathcal{I}}(t)$ of the solution. Both parts are oscillatory functions of t. The solution is highly oscillatory if $|\xi| \gg 1$. Figure 1.3 shows the solution for $\lambda = 2i$ and $y_0 = 1 + 0.4i$. Another representation of the solution is obtained if we write the initial data in amplitude-phase form,

$$y_0 = e^{i\alpha}|y_0|, \quad -\pi < \alpha \le \pi.$$

One calls the modulus $|y_0|$ the *amplitude* of y_0 and the principal argument α the *phase* of y_0. The solution becomes

$$y(t) = e^{i\xi t}y_0 = e^{i(\xi t+\alpha)}|y_0| = \big(\cos(\xi t + \alpha) + i\sin(\xi t + \alpha)\big)|y_0|. \quad (1.4)$$

The real part of the solution with $\lambda = i$, $|y_0| = 1$, and $\alpha = -\pi/4$ is shown in Figure 1.4.

$$\operatorname{Re} y = \cos(t - \pi/4)$$

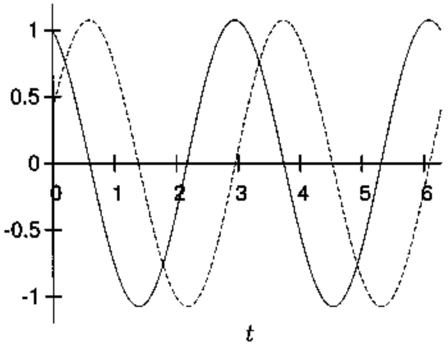

Figure 1.3 Oscillatory solution. $\operatorname{Re} y$ is shown as a solid line and $\operatorname{Im} y$ as a dashed line.

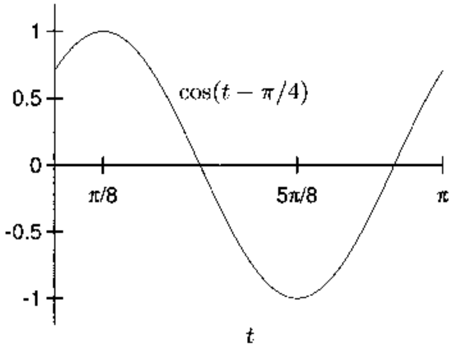

Figure 1.4 Real part of the solution (1.4).

4. *The general case.* Let

$$\lambda = \eta + i\xi, \quad y_0 = e^{i\alpha}|y_0|, \quad \xi, \eta \in \mathbb{R}.$$

The solution is given by

$$y(t) = e^{(\eta+i\xi)t} e^{i\alpha}|y_0| = e^{i(\xi t+\alpha)} e^{\eta t}|y_0|;$$

thus,

$$|y(t)| = e^{\eta t}|y_0|, \quad y(t) = e^{i(\xi t+\eta)}|y(t)|.$$

Therefore, depending on the sign of η, the amplitude $|y(t)|$ of the solution grows, decays, or remains constant. The phase $\xi t + \eta$ is a linear function of t and changes rapidly if $|\xi|$ is large.

Next, consider the inhomogeneous problem

$$\begin{aligned} \frac{dy}{dt} &= \lambda y + a e^{\mu t}, \\ y(0) &= y_0, \end{aligned} \tag{1.5}$$

where λ, a, μ, and y_0 are complex constants. Regardless of the initial condition at first, we look for a *particular solution* of the form

$$y_P(t) = e^{\mu t} A, \quad A = \text{const.} \tag{1.6}$$

Introducing (1.6) into the differential equation (1.5) gives us

$$\mu e^{\mu t} A = \lambda e^{\mu t} A + a e^{\mu t},$$

that is,

$$(\mu - \lambda) A = a.$$

If $\mu \neq \lambda$ we obtain the particular solution

$$y_P(t) = e^{\mu t}\frac{a}{\mu - \lambda}.$$

On the other hand, if $\mu = \lambda$, the procedure above is not successful and we try to find a solution of the form[1]

$$y_P(t) = te^{\mu t}A. \tag{1.7}$$

Introducing (1.7) into the differential equation gives us

$$e^{\mu t}A + \mu te^{\mu t}A = \lambda te^{\mu t}A + ae^{\mu t}.$$

The last equation is satisfied if we choose $A = a$; recall that $\lambda = \mu$ by assumption. Let us summarize our results.

Lemma 1.1 *The function*

$$y_P(t) = \begin{cases} \dfrac{ae^{\mu t}}{\mu - \lambda} & \text{if } \mu \neq \lambda \\ \\ ate^{\mu t} & \text{if } \mu = \lambda \end{cases}$$

is a solution of the differential equation $dy/dt = \lambda y + ae^{\mu t}$.

Note that the particular solution $y_P(t)$ does not adjust, in general, to the initial data given (i.e., $y_P(0) \neq y_0$). The initial value problem (1.5) can now be solved in the following way. We introduce the dependent variable u by

$$y = y_P + u.$$

Initial value problem (1.5) yields

$$\begin{aligned} \frac{dy_P}{dt} + \frac{du}{dt} &= \lambda(y_P + u) + ae^{\mu t}, \\ y_P(0) + u(0) &= y_0, \end{aligned}$$

and, since $dy_P/dt = \lambda y_P + ae^{\mu t}$, we obtain

$$\begin{aligned} \frac{du}{dt} &= \lambda u, \\ u(0) &= y_0 - y_P(0). \end{aligned}$$

Thus, $u(t)$ satisfies the corresponding homogeneous differential equation, and (1.3) yields

$$u(t) = e^{\lambda t}(y_0 - y_P(0)).$$

---

[1] The exceptional case, $\mu = \lambda$, is called the *case of resonance*.

The complete solution

$$y(t) = y_P(t) + u(t)$$

consists of two parts, $y_P(t)$ and $u(t)$. The function $y_P(t)$ is also called the *forced solution* because it has essentially the same behavior as that of the *forcing* $ae^{\mu t}$. The other part, $u(t)$, is often called the *transient solution* since it converges to zero for $t \to \infty$ if $\operatorname{Re} \lambda < 0$.

Finally, we want to show how we can solve the initial value problem

$$\begin{aligned} \frac{dy}{dt} &= \lambda y + F(t), \\ y(0) &= y_0, \end{aligned} \tag{1.8}$$

with a general forcing $F(t)$. We can solve this problem by applying a procedure known as *Duhamel's principle.*

1.1.1 Duhamel's principle

Lemma 1.2 *The solution of* (1.8) *is given by*

$$y(t) = e^{\lambda t} y_0 + \int_0^t e^{\lambda(t-s)} F(s)\, ds. \tag{1.9}$$

Proof: Define $y(t)$ by formula (1.9). Clearly, $y(0) = y_0$ (i.e., the initial condition is satisfied). Also, $y(t)$ is a solution of the differential equation, because

$$\frac{dy}{dt} = \lambda e^{\lambda t} y_0 + F(t) + \lambda \int_0^t e^{\lambda(t-s)} F(s)\, ds = \lambda y(t) + F(t).$$

This proves the lemma. ∎

Exercise 1.1 *Prove that the solution* (1.9) *is the unique solution of* (1.8).

We shall now discuss the relation between the solution to inhomogeneous equation (1.8) and the homogeneous equation

$$\frac{du}{dt} = \lambda u. \tag{1.10}$$

We consider (1.10) with initial condition $u = u(s)$ at a time $s > 0$. At a later time $t \geq s$ the solution is

$$u(t) = e^{\lambda(t-s)} u(s).$$

Thus, $e^{\lambda(t-s)}$ is a factor that connects $u(t)$ with $u(s)$. We will call it the *solution operator* and use the notation

$$S(t,s) = e^{\lambda(t-s)}, \quad \text{i.e.,} \quad u(t) = S(t,s)u(s). \tag{1.11}$$

The solution operator has the following properties:

$$\begin{aligned} S(t,0) &= e^{\lambda t}, \quad S(t,t) = 1, \\ S(t,t_1)S(t_1,s) &= S(t,s). \end{aligned} \tag{1.12}$$

Now we can show that the solution of inhomogeneous equation (1.8) can be expressed in terms of the solution of homogeneous equation (1.10). Then (1.9) becomes

$$y(t) = S(t,0)y(0) + \int_0^t S(t,s)F(s)\,ds. \tag{1.13}$$

In a somewhat loose way, we may consider the integral as a "sum" of many terms $S(t,s_j)F(s_j)\Delta s$; think of approximating the integral by a Riemann sum. Then (1.13) expresses the solution of inhomogeneous problem (1.8) as a weighted superposition of solutions $t \to S(t,s)$ of homogeneous equation (1.10). The idea of expressing the solution of an inhomogeneous problem via solutions of the homogeneous equation is very useful. As we will see, it generalizes to systems of equations, to partial differential equations, and also to difference approximations. It is known as *Duhamel's principle.*

Exercise 1.2 *Use Duhamel's principle to derive a representation for the solution of*

$$\begin{aligned} \frac{dy}{dt} &= \lambda(t)y + F(t), \\ y(0) &= y_0. \end{aligned}$$

Exercise 1.3 *Consider the inhomogeneous initial value problem*

$$\begin{aligned} \frac{dy}{dt} &= \lambda y + P_n(t)\, e^{\mu t}, \\ y(0) &= y_0, \qquad \lambda, \mu, y_0 \in \mathbb{C}, \end{aligned}$$

where $P_n(t)$ is a polynomial of degree n with complex coefficients. Show that the solution to the problem is of the form

$$y(t) = e^{\lambda t}\tilde{y}_0 + Q_m(t)e^{\mu t},$$

where $Q_m(t)$ is a polynomial of degree m with $m = n$ in the nonresonance case ($\mu \neq \lambda$) and $m = n + 1$ in the resonance case ($\mu = \lambda$). Determine $\tilde{y}_0$ in each case.

We now want to consider scalar equations with smooth variable coefficients, which leads to the next principle.

1.1.2 Principle of frozen coefficients

In many applications the problem with smooth variable coefficients can be *localized,* that is, it can be decomposed in many constant coefficient problems (by using a partition of unity). Then by solving all these constant coefficient problems, one can construct an approximate solution to the original variable coefficient problem. The approximation can be as good as one wants. The general theory concludes that if all relevant constant coefficient problems have a solution, the variable coefficient problem also has a solution. This procedure is known as the *principle of frozen coefficients*. We do not go into it more deeply here.

1.2 Variable coefficient linear equations

1.2.1 Principle of superposition

The initial value problem

$$\begin{aligned}\frac{dy}{dt} &= a(t)y + F(t),\\ y(0) &= y_0,\end{aligned} \tag{1.14}$$

is an example of a linear problem. It has the following properties:

1. Let $y(t)$ be a solution of (1.14). Let σ be a constant and replace $F(t)$ and y_0 by $\sigma F(t)$ and σy_0, respectively. In other words, consider the new problem

   $$\begin{aligned}\frac{d\tilde{y}}{dt} &= a(t)\tilde{y} + \sigma F(t),\\ \tilde{y}(0) &= \sigma y_0.\end{aligned} \tag{1.15}$$

   Multiplying (1.14) by σ gives us

   $$\begin{aligned}\frac{d(\sigma y)}{dt} &= a(t)(\sigma y) + \sigma F(t),\\ (\sigma y(0)) &= \sigma y_0.\end{aligned}$$

   Thus, (1.15) solves the new problem and, using uniqueness, the solution (1.15) is

   $$\tilde{y}(t) = \sigma y(t).$$

2. Consider (1.14) with a set of two forcing functions $F_1(t), F_2(t)$ and two initial data y_{01}, y_{02}. Denote the resulting solutions by $y_1(t)$, $y_2(t)$, respectively:

   $$\begin{aligned}\frac{dy_1}{dt} &= a(t)y_1(t) + F_1(t),\\ y_1(0) &= y_{01},\\ \frac{dy_2}{dt} &= a(t)y_2(t) + F_2(t),\\ y_2(0) &= y_{02}.\end{aligned}$$

Adding the equations, we find that

$$\begin{aligned}\frac{d(y_1+y_2)}{dt} &= a(t)(y_1+y_2)+(F_1(t)+F_2(t)),\\ y_1(0)+y_2(0) &= y_{01}+y_{02}.\end{aligned}$$

Thus, the sum

$$\tilde{y}(t)=y_1(t)+y_2(t)$$

is a solution of

$$\begin{aligned}\frac{d\tilde{y}}{dt} &= a(t)\tilde{y}+(F_1(t)+F_2(t)),\\ \tilde{y}(0) &= y_{01}+y_{02}.\end{aligned}$$

To summarize, for a linear problem such as (1.14), we can use superposition of solutions to obtain new solutions. This property of linear systems, known as the *superposition principle,* can be used to compose solutions with complicated forcing functions out of solutions of simpler problems. Consider, for example,

$$\frac{dy}{dt}=-y+\sin^2(t),\quad y(0)=y_0. \tag{1.16}$$

Since

$$\sin^2(t)=\left(\frac{1}{2i}(e^{it}-e^{-it})\right)^2=-\frac{1}{4}(e^{2it}-2+e^{-2it})$$

consists of three terms, we solve three problems:

$$\begin{aligned}\frac{dy_1}{dt} &= -y_1-\frac{1}{4}e^{2it},\\ \frac{dy_2}{dt} &= -y_2+\frac{1}{2},\\ \frac{dy_3}{dt} &= -y_3-\frac{1}{4}e^{-2it}.\end{aligned}$$

We do not yet impose initial conditions, so that we can choose simple particular solutions. By Lemma 1.1, particular solutions of the equations above are

$$y_{1P}=-\frac{e^{2it}}{4(1+2i)},\quad y_{2P}=\frac{1}{2},\quad y_{3P}=-\frac{e^{-2it}}{4(1-2i)}.$$

Using the superposition principle, we find that

$$\begin{aligned}\tilde{y}=y_{1P}+y_{2P}+y_{3P} &= \frac{1}{2}-\frac{1}{4}\left(\frac{e^{2it}}{1+2i}+\frac{e^{-2it}}{1-2i}\right)\\ &= \frac{1}{2}-\frac{1}{20}\left((1-2i)e^{2it}+(1+2i)e^{-2it}\right)=\frac{1}{2}-\frac{1}{10}\cos(2t)-\frac{1}{5}\sin(2t)\end{aligned}$$

solves

$$\frac{d\tilde{y}}{dt} = -\tilde{y} + \sin^2(t).$$

Clearly,

$$\tilde{y}(0) = \tfrac{1}{2} - \tfrac{1}{10} = \tfrac{2}{5}.$$

Therefore, the solution of (1.16) is given by

$$y(t) = \tilde{y}(t) + \sigma e^{-t},$$

where σ is determined by the initial condition

$$y(0) = \tilde{y}(0) + \sigma = \tfrac{2}{5} + \sigma = y_0, \quad \text{i.e.,} \quad \sigma = y_0 - \tfrac{2}{5}.$$

The superposition principle relies only on linearity; it holds for any linear equation or system of linear equations, both ordinary and partial differential equations. An equation is linear if the dependent variable and its derivatives appear linearly only (i.e., as the first power), in the equation.

Exercise 1.4 *Solve the initial value problem*

$$\begin{aligned} \frac{dy}{dt} &= -2y + (1+t^2)\sin(t)\cos(t), \\ y(0) &= 1. \end{aligned}$$

1.2.2 Duhamel's principle for variable coefficients

We want to discuss now the solution of problem (1.14) in terms of Duhamel's principle. To this end we discuss the solution operator in a more abstract setting.

Consider first an initial value problem for the homogeneous equation associated with (1.14):

$$\begin{aligned} \frac{dy}{dt} &= a(t)y, \\ y(0) &= y_0. \end{aligned} \tag{1.17}$$

The *solution operator* for problem (1.17) is given by

$$S(t_2, t_1) = \exp\Big(\int_{t_1}^{t_2} a(s)\, ds \Big). \tag{1.18}$$

Clearly,

$$y(t) = S(t, 0)y_0$$

is the solution of (1.17). With this solution operator, Duhamel's principle [see equation (1.13)] generalizes to our variable coefficient problem (1.14):

$$y(t) = S(t, 0)y(0) + \int_0^t S(t, r)F(r)\, dr. \tag{1.19}$$

This can be proved in terms of general properties of the solution operator.

It is not difficult to show that (1.18) has the following properties:

1. $S(t,t) = I$. Here I represents the identity operator [i.e., $Iv(t) = v(t)$].
2. Let $t \geq t_1 \geq 0$. Then

$$S(t,t_1)S(t_1,0) = S(t,0).$$

3. $S(t,r)$ is a smooth function of t and

$$\frac{\partial S(t,r)}{\partial t} = a(t)S(t,r).$$

We shall use these properties to prove that (1.19) solves (1.14). Since $S(0,0) = I$, we have $y(0) = y_0$. Also,

$$\begin{aligned}\frac{dy}{dt} &= S_t(t,0)y(0) + F(t) + \int_0^t S_t(t,r)F(r)\,dr \\ &= a(t)\Big(S(t,0)y(0) + \int_0^t S(t,r)F(r)dr\Big) + F(t) \\ &= a(t)y(t) + F(t).\end{aligned}$$

Therefore, $y(t)$ given by (1.19) is the solution of (1.14).

Exercise 1.5 *Find the solution operator and, using Duhamel's principle, the solution of the following initial value problems.*

(*a*)

$$\begin{aligned}\frac{dy}{dt} &= \frac{y}{1+t} + t, \\ y(0) &= 1.\end{aligned}$$

(*b*)

$$\begin{aligned}\frac{dy}{dt} &= -2ty + 2te^{-2t^2}, \\ y(0) &= 0.\end{aligned}$$

(*c*)

$$\begin{aligned}\frac{dy}{dt} &= \cos(t)y + \sin(t)\cos(t), \\ y(0) &= 2.\end{aligned}$$

1.3 Perturbations and the concept of stability

Given a problem and perturbations to it, we want to know what effect the perturbations have on the solution.

As an example, consider the initial value problem

$$\begin{aligned} \frac{dy}{dt} &= \lambda y - e^{-t}, \\ y(0) &= \frac{1}{\lambda + 1} \end{aligned} \tag{1.20}$$

with $\lambda \neq -1$. (The exceptional case of resonance, $\lambda = -1$, can be treated with slight modifications.)

The solution of (1.20) is the decaying function

$$y(t) = \frac{e^{-t}}{\lambda + 1}.$$

Now consider the same differential equation with perturbed initial data

$$\begin{aligned} \frac{d\tilde{y}}{dt} &= \lambda \tilde{y} - e^{-t}, \\ \tilde{y}(0) &= \frac{1}{\lambda + 1} + \varepsilon, \end{aligned} \tag{1.21}$$

where $0 < \varepsilon \ll 1$ is a small constant. Let $w(t) = \tilde{y}(t) - y(t)$ denote the difference between the perturbed and original solutions. Subtracting (1.20) from (1.21), we obtain

$$\begin{aligned} \frac{dw}{dt} &= \lambda w, \\ w(0) &= \tilde{y}(0) - y(0) = \varepsilon, \end{aligned}$$

whose solution is

$$w(t) = e^{\lambda t}\varepsilon, \quad \text{i.e.,} \quad |w(t)| = e^{\operatorname{Re}\lambda t}|\varepsilon|. \tag{1.22}$$

Depending on the sign of $\operatorname{Re}\lambda$, there are three possibilities.

1. $\operatorname{Re}\lambda < 0$. In this case the perturbation term $w(t)$ decays exponentially with time and the solution of the perturbed problem converges to the solution of the original problem as t increases. In Figure 1.5 the solid line represents the solution with $\lambda = -7/12$, and the dashed line is the solution with perturbed initial data.

2. $\operatorname{Re}\lambda = 0$. The perturbation $w(t)$ does not decrease with time, but it does not grow either (see Figure 1.6).

3. $\operatorname{Re}\lambda > 0$. The perturbation grows exponentially in time. Figure 1.7 shows the perturbed and unperturbed solutions for $\lambda = 1$.

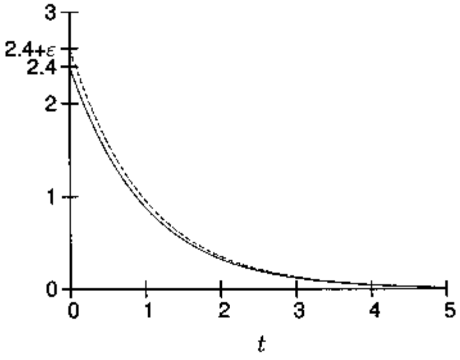

Figure 1.5 Decaying perturbation.

In the latter case it will be very difficult to compute the original solution accurately in long time intervals. For example, if $\lambda = 1$ and $\varepsilon = 10^{-10}$, then

$$w(T) = e^T 10^{-10} = 1 \quad \text{if} \quad T = \log_e 10^{10} = 10 \log_e 10 \simeq 25.$$

Therefore, if the calculation introduces an error $\varepsilon = 10^{-10}$ at $t = 0$, this error will grow to $w(T) = 1$ at about $T = 25$. This growth holds even if no further errors, except the original error $\varepsilon = 10^{-10}$ at $t = 0$, are introduced.

In applications the initial data and the forcing are never given exactly. Therefore, if $\operatorname{Re}\lambda > 0$, one cannot guarantee that the answer computed is close to the correct answer. For $\operatorname{Re}\lambda < 0$, the situation is the opposite: Initial errors in the data are wiped out. Problems corresponding to $\operatorname{Re}\lambda < 0$ are called *strongly stable*. If $\operatorname{Re}\lambda = 0$, the problem is *stable* but not strongly stable, and if $\operatorname{Re}\lambda > 0$, the problem is *unstable*.

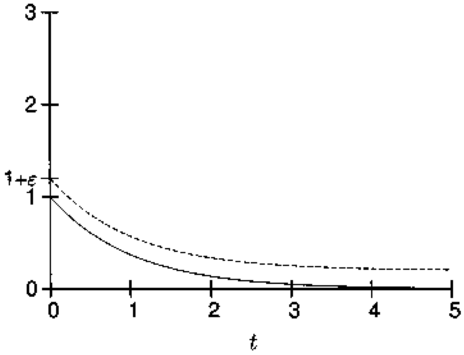

Figure 1.6 Non-decaying perturbation.

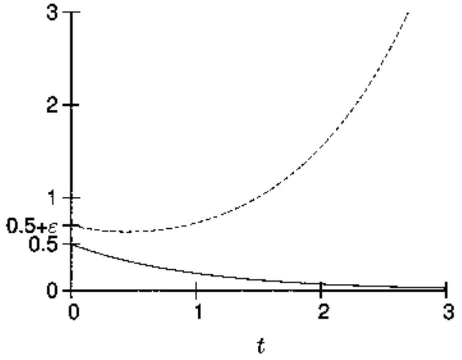

Figure 1.7 Exponentially growing perturbation.

Next, let us perturb the forcing and consider

$$\begin{aligned} \frac{d\tilde{y}}{dt} &= \lambda\tilde{y} - e^{-t} + \varepsilon G(t), \\ \tilde{y}(0) &= \frac{1}{\lambda + 1} \end{aligned} \tag{1.23}$$

instead of (1.20). The error term $w(t) = \tilde{y}(t) - y(t)$ solves

$$\begin{aligned} \frac{dw}{dt} &= \lambda w + \varepsilon G(t), \\ w(0) &= 0, \end{aligned} \tag{1.24}$$

and we obtain by Duhamel's principle,

$$w(t) = \varepsilon \int_0^t e^{\lambda(t-s)} G(s)\, ds.$$

Therefore, $w(t)$ satisfies the estimate

$$|w(t)| \le \varepsilon \int_0^t |e^{\lambda(t-s)}||G(s)|ds \le \varepsilon \max_{0\le s\le t} |G(s)| \int_0^t e^{\operatorname{Re}\lambda(t-s)} ds$$

$$= \varepsilon \max_{0\le s\le t} |G(s)| \begin{cases} \dfrac{e^{\operatorname{Re}\lambda t} - 1}{\operatorname{Re}\lambda} & \text{if } \operatorname{Re}\lambda \ne 0 \\ t & \text{if } \operatorname{Re}\lambda = 0 \end{cases} \tag{1.25}$$

We arrive at essentially the same conclusions as those for the perturbed initial data:

1. If the problem is strongly stable (i.e., $\operatorname{Re}\lambda < 0$), the perturbation of the solution is bounded by

$$|w(t)| \le \frac{\varepsilon}{|\operatorname{Re}\lambda|} \max_{0\le s\le t} |G(s)|;$$

that is, the difference of the solutions is of the same order as the perturbation of the forcing.

2. If the problem is stable but not strongly stable (i.e., $\operatorname{Re}\lambda = 0$), we obtain

$$|w(t)| \le \varepsilon \int_0^t |G(s)|\, ds \le \varepsilon t \max_{0\le s\le t} |G(t)|.$$

   Thus, the difference in the solutions can be estimated in terms of the integrated effect of the perturbation. This effect typically grows linearly with time. In most applications one can handle such situations and obtain accurate solutions by keeping the perturbations sufficiently small. For example, if $\varepsilon = 10^{-10}$ and $|G(t)| \le 1$, it takes a very long time before the effect of the perturbation is noticed.

3. If $\operatorname{Re}\lambda > 0$, the effect of the perturbation grows exponentially in time. In a long time interval, this may change the true solution drastically.

There are no difficulties in generalizing this observation to linear equations with variable coefficients:

$$\begin{aligned} \frac{dy}{dt} &= \lambda(t)y + F(t), \\ y(0) &= y_0. \end{aligned} \tag{1.26}$$

The influence of perturbations of the forcing and of the initial data depends on the behavior of the solution operator

$$S(t,s) = e^{\int_s^t \lambda(\eta)d\eta}, \quad t \ge s.$$

Definition 1.3 *Consider the linear initial value problem* (1.26) *and its solution operator* $S(t,s)$. *The problem is called strongly stable, stable, or unstable if the solution operator satisfies, respectively, the following estimates:*

$$|S(t,s)| \le e^{-\delta(t-s)}, \quad |S(t,s)| \le \text{const.}, \quad \textit{or} \quad |S(t,s)| \ge e^{\delta(t-s)},$$

where δ *is a positive constant.*

Exercise 1.6 *Consider, instead of* (1.20), *the initial value problem for the resonance case*

$$\begin{aligned} \frac{dy}{dt} &= \lambda y + e^{\lambda t}, \\ y(0) &= 0, \end{aligned}$$

and the problem with perturbed initial data,

$$\begin{aligned} \frac{d\tilde{y}}{dt} &= \lambda \tilde{y} + e^{\lambda t}, \\ \tilde{y}(0) &= \varepsilon, \quad 0 < \varepsilon \ll 1. \end{aligned}$$

Show that the same conclusions of the nonresonance case can be drawn for $w(t) = \tilde{y} - y(t)$.

1.4 Nonlinear equations: the possibility of blow-up

Nonlinearities in the equation can produce a solution that blows up in finite time, that is, a solution that does not exist for all times. Consider, for example, the nonlinear initial value problem given by

$$\begin{aligned} \frac{dy}{dt} &= y^2, \quad t \geq 0, \\ y(0) &= y_0. \end{aligned} \tag{1.27}$$

For $y_0 = 0$ the solution is $y = 0$ for all times. Therefore, assume that $y_0 \neq 0$ in the following. To calculate the solution, we write the differential equation in the form

$$\frac{1}{y^2}\frac{dy}{dt} = 1$$

and integrate:

$$\int_0^t \frac{1}{y^2(s)}\frac{dy}{ds}\,ds = t.$$

The change of variables $y(s) = v$ gives us

$$\int_{y_0}^{y(t)} \frac{dv}{v^2} = t,$$

and we obtain

$$\frac{1}{y_0} - \frac{1}{y(t)} = t, \quad \text{i.e.,} \quad y(t) = \frac{1}{(1/y_0) - t}.$$

For $y_0 > 0$ the solution blows up at $t = 1/y_0$ (see Figure 1.8). This blow-up or divergence of the solution at a finite time is a consequence of the nonlinearity, that is, the term y^2 on the right-hand side of the equation. This behavior cannot occur in

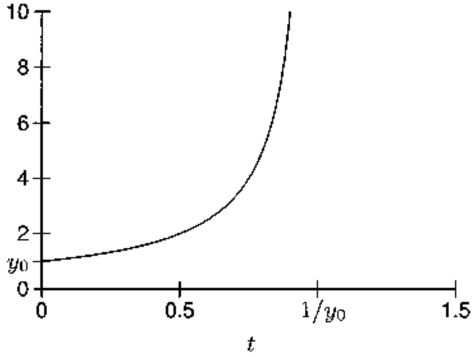

Figure 1.8 $y_0 > 0$.

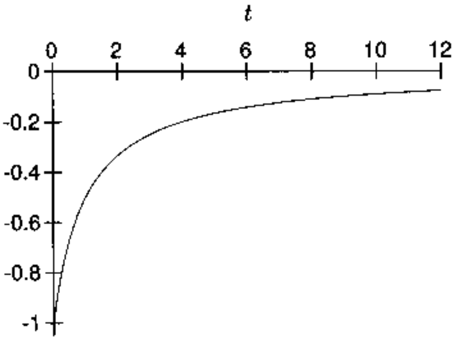

Figure 1.9 $y_0 < 0$.

a linear problem. On the other hand, if $y_0 < 0$, the solution $y(t)$ exists for all $t \geq 0$ and converges to zero for $t \to \infty$ (see Figure 1.9).

Consider now the more general problem

$$\begin{aligned} \frac{dy}{dt} &= f(y,t), \\ y(t_0) &= y_0. \end{aligned} \tag{1.28}$$

We give here without proof a simple version of the classical *existence and uniqueness theorem* for scalar ordinary differential equations (see, e.g., [3], chapt. 5).

Theorem 1.4 *If* $f(y,t)$ *and* $\partial f(y,t)/\partial y$ *are continuous functions in a rectangle* $\Omega = [y_0 - b, y_0 + b] \times [t_0 - a, t_0 + a]$, $a, b > 0$, *and* $|f(y,t)| \leq M$ *on* Ω, *there exists a unique, continuously differentiable solution* $y(t)$ *to the problem* (1.28) *in the interval* $|t - t_0| \leq \Delta t = \min\{a, b/M\}$.

Remark 1.5 *The time interval of existence depends on how large one can choose the rectangle, and so on the initial point* (y_0, t_0). *The solution can be continued to the future by solving the equation with new initial conditions starting at the point* $(y(t_0 + \Delta t), t_0 + \Delta t)$. *If one tries to continue the solution as much as possible, there are two possibilities:*

1. *One can continue the solution to arbitrarily large times, that is, the solution exists for all times* $t \geq 0$.

2. *There is a finite time* $T_0 = T_0(y_0) > 0$ *such that the solution exists for all times* $t < T_0$ *but not at* T_0. *In this case the solution blows up at* T_0; *that is,* $\lim_{t \to T_0^-} |y(t)| = \infty$.

Exercise 1.7 *Show that if a smooth solution* $y(t)$ *to* (1.28) *blows up at finite time, its derivative* dy/dt *blows up at the same time.* Hint: *Use the mean value theorem.*

Exercise 1.8 *Show that the converse of exercise* (1.7) *is false. To this end, consider the initial value problem*

$$\begin{aligned} \frac{dy}{dt} &= \frac{1}{2-y}, \\ y(0) &= 1. \end{aligned}$$

Explicitly find the solution $y(t)$ *and check that* $dy/dt \to \infty$ *when* $t \to (\frac{1}{2})^-$ *and that, nevertheless,* $y(t)$ *stays bounded.*

Exercise 1.9 *Is it possible that the solution of the real equation*

$$\begin{aligned} \frac{dy}{dt} &= \sin(y), \\ y(0) &= y_0 \end{aligned}$$

blows up at a finite time? Explain the Answer.

1.5 Principle of linearization

Consider the initial value problem

$$\begin{aligned} \frac{dy}{dt} &= \lambda y + y^2 + F(t), \\ y(0) &= 0. \end{aligned} \tag{1.29}$$

Assume that

$$F(t) = \cos(t) - \lambda \sin(t) - \sin^2(t).$$

A simple calculation shows that the solution of (1.29) is given by

$$y(t) = \sin(t).$$

Let ε with $0 \le \varepsilon \ll 1$ be a small constant and consider the perturbed problem

$$\begin{aligned} \frac{d\tilde{y}}{dt} &= \lambda \tilde{y} + \tilde{y}^2 + F(t) + \varepsilon G(t), \\ \tilde{y}(0) &= 0. \end{aligned} \tag{1.30}$$

Here $G(t)$ is a smooth function with

$$|G(t)| + \left|\frac{dG}{dt}\right| \le 1. \tag{1.31}$$

By Section 1.3 we expect that in some time interval $0 \le t \le T$,

$$\tilde{y}(t) = \sin(t) + \mathcal{O}(\varepsilon).^2$$

---

[2] From now on we frequently use the notation $\mathcal{O}$; for a precise definition, consult Section A.2.

Therefore, we make the following change of variables:

$$\tilde{y}(t) = \sin(t) + \varepsilon u(t). \tag{1.32}$$

Introducing (1.32) into (1.30) gives us

$$\cos(t) + \varepsilon \frac{du}{dt} = \lambda \sin(t) + \varepsilon\lambda u + \sin^2(t) + 2\varepsilon \sin(t)u + \varepsilon^2 u^2 + F(t) + \varepsilon G(t).$$

The form of F gives us

$$\begin{aligned} \frac{du}{dt} &= (\lambda + 2\sin t)u + \varepsilon u^2 + G(t), \\ u(0) &= 0. \end{aligned} \tag{1.33}$$

We expect that $|u| \le 1$ in some time interval $0 \le t \le T$, and therefore we can neglect the quadratic term εu^2 to obtain the *linearized equation*

$$\begin{aligned} \frac{d\tilde{u}}{dt} &= (\lambda + 2\sin t)\tilde{u} + G(t), \\ \tilde{u}(0) &= 0. \end{aligned} \tag{1.34}$$

In Section 1.3 we discussed the growth behavior of the solutions of (1.34). It depends on the solution operator

$$S(t, s) = e^{\lambda(t-s)+2\int_s^t \sin(r)\,dr}.$$

Exercise 1.10 *Prove, using the solution operator $S(t, s)$, that the problem is strongly stable for $\lambda < 0$.*

If the linearized equation is strongly stable (i.e., Re $\lambda < 0$), then

$$|\tilde{u}(t)| \le \text{const.} \max_{0\le s\le t} |G(s)|$$

and, using (1.31), $\tilde{u}$ is bounded for all times. In this case one can also show that, for sufficiently small ε,

$$|u - \tilde{u}| = \mathcal{O}(\varepsilon)$$

for all times. Thus, the linearized equation determines, to first approximation, the effect of the perturbation on the solution.

If the linearized equation is only stable, then

$$|\tilde{u}(t)| \le \text{const.} \int_0^t |G(s)|\,ds$$

and

$$|u - \tilde{u}| \le \text{const.}\, \varepsilon \int_0^t |G(s)|\,ds,$$

provided that

$$\varepsilon \max_{0\le s\le t} |\tilde{u}(s)|^2 \ll 1.$$

Therefore, the linearized equation describes the behavior of the perturbation in every time interval $0 \le t \le T$ with $T^2\varepsilon \ll 1$.

If the linearized equation is unstable,

$$|\tilde{u}(t)| \le \text{const.}\, e^{\operatorname{Re}\lambda t}$$

and the time interval where the linearized equation (1.34) is a good approximation of (1.33) is restricted to a time interval $0 \le t \le T$ with

$$T = \mathcal{O}(\log_e \varepsilon^{-1}).$$

This behavior is general in nature.

Consider the nonlinear equation

$$\begin{aligned} \frac{dy}{dt} &= f(y,t), \\ y(0) &= y_0. \end{aligned} \tag{1.35}$$

Assume that the solution $y(t)$ of this problem is known. Consider a perturbation

$$\begin{aligned} \frac{d\tilde{y}}{dt} &= f(\tilde{y},t) + \varepsilon G(t), \\ \tilde{y}(0) &= y_0 + \varepsilon\tilde{y}_0. \end{aligned} \tag{1.36}$$

We make the change of variables

$$\tilde{y} = y + \varepsilon u.$$

Since

$$f(y+\varepsilon u, t) = f(y,t) + \varepsilon\frac{\partial f}{\partial y}(y,t)u + \varepsilon^2\mathcal{O}(u^2),$$

we obtain

$$\begin{aligned} \frac{dy}{dt} + \varepsilon\frac{du}{dt} &= f(y,t) + \varepsilon\frac{\partial f}{\partial y}(y,t)u + \varepsilon^2\mathcal{O}(u^2) + \varepsilon G(t), \\ u(0) &= \tilde{y}_0. \end{aligned} \tag{1.37}$$

Neglecting the quadratic terms, we obtain the linearized equation

$$\begin{aligned} \frac{du}{dt} &= \frac{\partial f}{\partial y}(y,t)u + G(t), \\ u(0) &= \tilde{y}_0. \end{aligned} \tag{1.38}$$

The effect of the perturbation depends on the stability properties of (1.38).

Linearization is a very important tool because it is used to show that the nonlinear problem has a unique solution locally.