CHAPTER 6

Contraction principle

45. Initial value problems and integral equations

45.1. Anharmonic oscillator. Consider the initial value problem for an oscillator for which the frequency depends on time and and the amplitude

$$
\begin{aligned}
(45.1)\qquad & u''(t) + (\omega^2 - \nu(t,u))u(t) = f(t)\,, \quad t > 0 \\
& u\Big|_{t=0} = u_0\,, \quad u'\Big|_{t=0} = u_1\,.
\end{aligned}
$$

Here it is assumed that $|\nu| < \omega^2$ (to have an oscillatory solution), and $f(t)$ is an external force. For example, one can think of a pendulum with non-small oscillations. Recall that if the *anharmonicity* ν vanishes, then the solution $u(t) = v(t)$ to the initial value problem can be found by the convolution

$$
\begin{aligned}
u(t) &= G(t) \star \Big(\theta(t)f(t) + u_0\delta'(t) + u_1\delta(t)\Big)\,, \\
G(t) &= \theta(t)\frac{\sin(\omega t)}{\omega}\,.
\end{aligned}
$$

If f is a regular distribution,

$$\int_0^T |f(t)|\,dt < \infty$$

for any $T > 0$, then

$$v(t) = \int_0^t G(t-\tau)f(\tau)\,d\tau + u_0\cos(\omega t) + u_1\,\frac{\sin(\omega t)}{\omega}\,, \quad t > 0$$

and $v \in C^1(t \geq 0)$ such that

$$\lim_{t\to 0^+} v(t) = u_0\,, \quad \lim_{t\to 0^+} v'(t) = u_1$$

Indeed, the integrand and its derivative are continuous in the parameter t and are bounded by an integrable function independent of the parameter t

$$|G(t-\tau)f(\tau)| \leq \frac{1}{\omega}\,|f(\tau)|\,, \quad |\partial_t G(t-\tau)f(\tau)| \leq |f(\tau)|$$

Therefore, $v(t)$ and $v'(t)$ are continuous for $t > 0$ and

$$v'(t) = \int_0^t \cos(\omega(t-\tau)) f(\tau)\, d\tau + u_1 \cos(\omega t) + u_0 \omega \sin(\omega t)\,, \quad t > 0$$

from which the limit properties follow.

If in addition, $f \in C^0(t \geq 0)$, then $v \in C^2(t \geq 0)$ by similar arguments ($\partial_t^2 G(t-\tau) f(\tau)$ is continuous and is bounded by an integrable function $\omega |f(\tau)|$ independent of t), and for $t > 0$

$$\begin{aligned} v''(t) &= f(t) - \int_0^t \sin(\omega(t-\tau)) f(\tau)\, d\tau - u_1 \omega \sin(\omega t) - u_0 \omega^2 \cos(\omega t) \\ &= f(t) - \omega^2 v(t) \end{aligned}$$

so that $v''(t) \to f(0) - \omega^2 u_0$ which means that v'' has a continuous extension to $t = 0$. Thus, with a continuous force, $v(t)$ is the classical solution to the initial value problem.

Suppose the anharmonicity $\nu(t, z)$ is a continuous function in the variable $t \geq 0$ and z, and, in addition, it is bounded

$$|\nu(t, z)| \leq \nu_0\,.$$

Therefore for any continuous amplitude $u(t)$, the anharmonicity $\nu(t, u(t))$ is a continuous and bounded function of time.

Let $u(t)$ be a solution to the non-linear initial value problem. Let us extend it by zero for $t < 0$. Then $u(t)$ is a regular distribution that satisfies the equation

$$u''(t) + \omega^2 u(t) = u_0 \delta'(t) + u_1 \delta'(t) + \theta(t) F(t)$$

where

$$F(t) = f(t) + \nu(t, u(t))\,.$$

By taking the convolution of both sides of this equation with the causal Green's function $G(t)$, it is concluded that any classical solution $u \in C^2(t \geq 0)$ to the non-linear initial value problem satisfies the *integral equation*:

$$\begin{aligned} u(t) &= G(t) * \Big(\theta(t) F(t) + u_0 \delta'(t) + u_1 \delta'(t) \Big) \\ &= v(t) + \int_0^t G(t-\tau) \nu\Big(\tau, u(\tau)\Big)\, u(\tau)\, d\tau \end{aligned}$$

where $v(t)$ is the solution to the associated linear problem. Thus, *every solution to the original initial value problem for a differential equation is a solution to the above integral equation.*

Let us investigate if the converse is true. First note that if $u(t)$ is continuous for $t \geq 0$, then the function

$$\hat{K}u(t) = \int_0^t G(t-\tau)\nu\Big(\tau, u(\tau)\Big)\, u(\tau)\, d\tau$$

is also continuous for $t \geq 0$. This follows from continuity of $\nu(t, u(t))$ and the Green's function G. Therefore the *integral operator* $\hat{K}$ maps the space of continuous function into itself:

$$\hat{K}: \quad C^0(t \geq 0) \to C^0(t \geq 0)$$

Suppose $u(t)$ is a continuous function that satisfies the integral equation

$$u(t) = v(t) + \hat{K}u(t)\,. \tag{45.2}$$

Let us show that u is also twice continuously differentiable and solves the said initial value problem. This means that *the initial value problem and the integral equation problem are equivalent.*

It is sufficient to investigate the differentiability of $\hat{K}u$ and find its derivatives to verify the differential equation. First note that $G \in C^\infty(t > 0)$ and

$$|\partial_t^n G(t-\tau)\nu(\tau, u(\tau))u(\tau)| \leq \omega^{n-1}|\nu(\tau, u(\tau))u(\tau)| \in \mathcal{L}[0, T]$$

by continuity of ν for any continuous u. Therefore the function $\hat{K}u$ is also from $C^\infty(t > 0)$ and

$$\begin{aligned}
\frac{d}{dt}\,\hat{K}u(t) &= \int_0^t \cos\Big(\omega(t-\tau)\Big)\nu\Big(\tau, u(\tau)\Big)\, u(\tau)\, d\tau\,,\\
\frac{d^2}{dt^2}\,\hat{K}u(t) &= \nu((t, u(t))u(t) - \omega \int_0^t \sin\Big(\omega(t-\tau)\Big)\nu\Big(\tau, u(\tau)\Big)\, u(\tau)\, d\tau\\
&= \nu((t, u(t))u(t) - \omega^2 \hat{K}u(t)
\end{aligned}$$

By taking the limit $t \to 0^+$ in (**45.2**) and in the first equation above u is shown to satisfy the initial conditions

$$\begin{aligned}
u\Big|_{t=0} &= v(0) + \lim_{t\to 0^+} \hat{K}u(t) = v(0) + 0 = u_0\,,\\
u'\Big|_{t=0} &= v'(0) + \lim_{t\to 0^+} \frac{d}{dt}\,\hat{K}u(t) = v'(0) + 0 = u_1
\end{aligned}$$

Using the explicit form of the derivatives of $\hat{K}u$, u is shown to satisfy the differential equation (**45.1**):

$$\begin{aligned} u''(t) &= v''(t) + \frac{d^2}{dt^2}\hat{K}u(t) \\ &= -\omega^2 v(t) + f(t) + \nu\Big(t, u(t)\Big)u(t) - \omega^2 \hat{K}u(t) \\ &= -\omega^2 v(t) + f(t) + \nu\Big(t, u(t)\Big)u(t) - \omega^2\Big(u(t) - v(t)\Big) \\ &= -\Big(\omega^2 - \nu(t, u(t))\Big)u(t) + f(t)\,, \quad t > 0 \end{aligned}$$

Thus, the problems are indeed equivalent.

45.2. The von Neumann series. The integral equation (**45.2**) is an example of a non-linear Volterra equation. There exists a general method for solving it that will be given later. Here a linear case is considered as it does not require any special theory beyond a basis analysis. So let ν be independent of the amplitude, but still depends on time. The operator $\hat{K}$ is linear in this case:

$$\hat{K}(c_1 u_1 + c_2 u_2) = c_1 \hat{K}u_1 + c_2 \hat{K}u_2$$

for any continuous functions $u_{1,2}$ and any constants $c_{1,2}$. As before, $\nu(t)$ is assumed to be continuous and bounded

$$|\nu(t)| < \nu_0\,.$$

To find a hint for solving (**45.2**), consider an approximate solution based on an approximation of an integral by a Riemann sum. In this case, the problem is reduced to a standard linear algebra problem.

45.2.1. A discretization approach. Suppose the solution is sought in an interval $[0, T]$. Let us partition this interval by $t_j = j\Delta t$, $j = 0, 1, ..., N$, $\Delta t = T/N$, and $t_{j-1} \le t_j^* \le t_j$ are sample points in partition intervals $[t_{j-1}, t_j]$, $j = 1, 2, ..., N$. Let $u_j = u(t_j^*)$ and $v_j = v(t_j^*)$. Then the integral is approximated by a Riemann sum:

$$\begin{aligned} u_j &= v_j + \int_0^{t_j^*} G(t_j^* - \tau)\nu(\tau)u(\tau)\,d\tau \\ &= v_j + \left(\int_0^{t_1^*} + \int_{t_1^*}^{t_2^*} + \cdots + \int_{t_{j-1}^*}^{t_j^*}\right) G(t_j^* - \tau)\nu(\tau)u(\tau)\,d\tau \\ &\approx v_j + \sum_{n=1}^{j} \Delta t_n^* G(t_j^* - t_n^*)\nu(t_n^*)u_n \end{aligned}$$

where $\Delta t_n^* = t_n^* - t_{n-1}^*$. The last term in the sum vanishes because $G(t_j^* - t_j^*) = 0$. So, the values of the solution at sample points can be approximated by the solution of the linear system

$$u_j + \sum_{n=1}^{N} K_{jn} u_n = v_j \,, \qquad K_{jn} = \begin{cases} G(t_j^* - t_n^*)\nu(t_n^*)\Delta t_n^* & , j > n \\ 0 & , j \le n \end{cases}$$

or in the matrix notations

$$(I - K)\mathbf{u} = \mathbf{v}$$

where I is the unit $N \times N$ matrix, K is the matrix with elements K_{jn}, and $\mathbf{u}$ and $\mathbf{v}$ are the N vector with components u_j and v_j, respectively. One can connect the points (t_j^*, u_j) in a plane by line segments to obtain a graph of a continuous function that approximates a solution to the integral equation.

A general solution to the matrix equation exists and is unique if the matrix $I - K$ is invertible. For the latter it is necessary and sufficient that the associated homogeneous equation $(I - K)\mathbf{u} = \mathbf{0}$ has only the trivial solution $\mathbf{u} = \mathbf{0}$. This is easy to verify by using the triangular structure of the matrix K and by writing the equation explicitly in the components:

$$\left.\begin{aligned} u_1 &= 0 \\ u_2 - K_{21}u_1 &= 0 \\ u_3 - K_{31}u_1 - K_{32}u_2 &= 0 \\ &\vdots \end{aligned}\right\} \quad \Rightarrow \quad u_j = 0\,.$$

Thus, the equation has a unique solution

$$\mathbf{u} = (I - K)^{-1}\mathbf{v}\,.$$

Recall a geometric series representation for a function $(1 - q)^{-1}$. Let us see if a *formal* geometric series

$$(I - K)^{-1}\mathbf{v} = \mathbf{v} + K\mathbf{v} + K^2\mathbf{v} + \cdots = \mathbf{v} + \sum_{n=1}^{\infty} K^n \mathbf{v}$$

converges to the solution. It turns out that this series is a finite sum. Indeed, K is a triangular matrix with diagonal elements being zero. Any matrix with these properties is nilpotent, and its Nth power is the zero matrix where N is the dimension of the matrix so that $K^n = 0$ for all $n \ge N$ in the series. Let us show that

$$\mathbf{u} = \mathbf{v} + K\mathbf{v} + K^2\mathbf{v} + \cdots + K^{N-1}\mathbf{v}$$

is the solution to the matrix equation:

$$\mathbf{v} + K\mathbf{u} = \mathbf{v} + K\mathbf{v} + K^2\mathbf{v} + \cdots + K^{N-1}\mathbf{v} + K^N\mathbf{v} = \mathbf{u}$$

because $K^N = 0$.

This offers the sought-after hint for solving a linear Volterra equation because repetitive actions of $\hat{K}$ is easy to calculate in the continuum case. As $\Delta t \to 0$ and $N \to \infty$, the action of $\hat{K}^n$ is given by an iterated integral and one would expect that the approximate solution converges to a solution to the integral equation. This is indeed so.

45.2.2. Continuous approach. The analysis given above motivates us to seek a solution to the integral equation in the form of a series

$$u(t) = v(t) + \hat{K}v(t) + \hat{K}^2 v(t) + \cdots = \sum_{n=0}^{\infty} \hat{K}^n v(t)$$

This series is called a *von Neumann series* for an operator $\hat{K}$. Here it is proved that

(i) *The von Neumann series for the integral Volterra operator converges uniformly to a continuous function from $C^0(t \geq 0)$ for any continuous function $v \in C^0(t \geq 0)$;*
(ii) *The sum of the series is a solution to the Volterra integral equation;*
(iii) *The solution is unique.*

Uniform convergence. Let us first investigate the convergence of the series. Put

$$m_0 = \sup_{[0,T]} |v(t)| < \infty$$

Here m_0 depends on T. The idea is to find upper bounds on the terms of the series and investigate the convergence of upper bounds. This would imply that the series *converges uniformly*. One has

$$|G(t-\tau)\nu(\tau)| \leq \frac{\nu_0}{\omega} \equiv \xi\,.$$

Using this inequality, one infers that

$$|\hat{K}v(t)| \leq \xi \int_0^t |v(\tau)| d\tau \leq \xi t\, m_0$$

$$|\hat{K}^2 v(t)| \leq \xi \int_0^t |\hat{K}v(\tau)|\, d\tau \leq \xi^2 m_0 \int_0^t \tau\, d\tau = \frac{\xi^2 t^2}{2}\, m_0$$

$$|\hat{K}^n v(t)| \leq \frac{(\xi t)^n}{n!}\, m_0$$

The latter relation is easy to prove by induction:

$$|\hat{K}^{n+1}v(t)| \leq \xi \int_0^t |\hat{K}^n v(\tau)|\, d\tau \leq \xi m_0 \int_0^t \frac{(\xi t)^n}{n!}\, d\tau$$
$$\leq \frac{(\xi t)^{n+1}}{(n+1)!}\, m_0$$

So the terms of the series are majorated by

$$|\hat{K}^n v(t)| \leq \frac{(\xi T)^n}{n!}\, m_0 = M_n$$

for all $t \in [0, T]$ and the series of the upper bounds converges

$$\sum_{n=0}^{\infty} M_n = e^{\xi T} m_0$$

This implies that the series converges uniformly on any interval $[0, T]$. Moreover, since the terms of the series are continuous functions, the uniform convergence also implies that the sum of the series is continuous on any interval $[0, T]$. As T is arbitrary, the sum of the series is a continuous function for all $t \geq 0$.

Verification of the integral equation. Let us check if the sum satisfies the integral equation. Put

$$u_N(t) = \sum_{n=0}^{N-1} \hat{K}^n v(t)$$

This is a sequence of partial sums, and

$$|u(t) - u_N(t)| \leq \sum_{n=N}^{\infty} |\hat{K}^n v(t)| \leq \Big(e^{\xi T} - \sum_{n=0}^{N-1} \frac{(\xi T)^n}{n!}\Big)\, m_0 = C_N(T)$$

for all $t \in [0, T]$. Since the power series for the exponential has infinite radius of convergence, $C_N(T) \to 0$ as $N \to \infty$ for any T. Therefore by taking the supremum in the left side of the above inequality

$$\lim_{N\to\infty} \sup_{[0,T]} |u(t) - u_N(t)| \leq \lim_{N\to\infty} C_N(T) = 0$$

This shows that the error of the approximation $u(t) \approx u_N(t)$ goes to zero *uniformly* on any interval $[0, T]$ with increasing N. It follows that

$$|\hat{K}(u - u_N)(t)| \leq \xi \int_0^t |u(\tau) - u_N(\tau)|\, d\tau \leq \xi C_N(T) \int_0^t d\tau$$
$$\leq \xi T C_N(T)$$

Therefore the image of the sequence of partial sums under the action of the integral operator, $\hat{K}u_N$, also converges uniformly to $\hat{K}u$:

$$\lim_{N\to\infty} \sup_{[0,T]} |\hat{K}(u - u_N)(t)| = 0$$

By taking the limit $N \to \infty$ in the identity

$$v + \hat{K}u_N = u_N + \hat{K}^N v$$

it is concluded that

$$v + \hat{K}u = u$$

because the sequences $\hat{K}u_N$, u_N, and $\hat{K}^N v$ converge uniformly to $\hat{K}u$, u, and 0, respectively, on any interval $[0,T]$. Thus, the series is indeed a solution to the integral equation.

Uniqueness of the solution. Let us show that the found solution is *unique*. Suppose that u and w are two solutions:

$$u = \hat{u} + v\,, \quad w = \hat{K}w + v$$

Then for any n

$$\hat{K}^n(u - w) = u - w \quad \Rightarrow \quad \sup_{[0,T]} |\hat{K}^n(u - w)(t)| = \sup_{[0,T]} |u(t) - w(t)|$$

On the other hand it has been shown that

$$\sup_{[0,T]} |\hat{K}^n(u - w)(t)| \le \frac{(\xi T)^n}{n!} \sup_{[0,T]} |u(t) - w(t)|$$

So, for any n

$$\sup_{[0,T]} |u(t) - w(t)| \le \frac{(\omega T)^n}{n!} \sup_{[0,T]} |u(t) - w(t)|\,.$$

But the factor $\frac{(\xi T)^n}{n!}$ is less than 1 for all n that are large enough and for any T. Therefore the inequality can hold if and only if

$$\sup_{[0,T]} |u(t) - w(t)| = 0$$

which is only possible if

$$u(t) = w(t)$$

and the solution is indeed unique.

The existence and uniqueness of a solution to the Volterra integral equation also implies that *the original initial value problem for the oscillator with a time-dependent frequency also has a unique solution.* A conversion of a problem in differential equation to an equivalent problem in integral equation is one of main mathematical tools to investigate the existence and uniqueness of a solution and also to obtain the solution by solving the integral equation.

45.3. Higher dimensional generalizations. Consider the Cauchy problem for a non-linear wave equation

$$u''_{tt} - c^2\Delta_x u = \lambda F(u,x,t)\,, \quad t>0\,,$$
$$u\Big|_{t=0} = u_0(x)\,, \quad u'_t\Big|_{t=0} = u_1(x)\,, \quad x\in\mathbb{R}^N$$

where λ is a numerical parameter and F is a continuous function of all variables. It is also assumed that u_0 and u_1 are smooth enough in order for a solution to exist in the class $C^2(t>0)\cap C^1(t\geq 0)$ when $F=0$. Suppose that the problem has a solution $u(x,t)$. It is extended by zero for all $t<0$. The extended function is a regular distribution that satisfies the equation

$$\Big(\frac{\partial^2}{\partial t^2} - c^2\Delta_x\Big)u(x,t) = \delta(t)\cdot u_1(x) + \delta'(t)\cdot u_0(x) + \lambda\theta(t)F(u(x,t),x,t)$$

Let $G(x,t)$ be the causal Green's function for the wave operator

$$\Big(\frac{\partial^2}{\partial t^2} - c^2\Delta_x\Big)G(x,t) = \delta(t)\cdot\delta(x)\,, \quad G(x,t)=0\,, \quad t<0\,.$$

By taking the convolution with both sides of this equation, it is shown that the classical solution for the Cauchy problem also satisfies the integral equation

$$u(x,t) = v(x,t) + \lambda G(x,t) * \theta(t)F(u(x,t),x,t)\,,$$
$$v(x,t) = G(x,t) * u_1(x) + \frac{\partial}{\partial t}G(x,t) * u_0(x)$$

where, in the latter equation, the convolution is taken with respect to x. The convolutions are classical convolution because $u_{0,1}$ and F are continuous functions. Recall the classical wave potentials.

For example, if $N=1$, then the integral equation is deduced from the d'Alembert formula:

$$u(x,t) = v(x,t) + \frac{\lambda}{2c}\int_0^t \int_{x-c(t-\tau)}^{x+c(t-\tau)} F(u(y,\tau),y,\tau)\,dy\,d\tau$$

$$v(x,t) = \frac{1}{2c}\int_{x-ct}^{x+ct} u_1(y)\,dy + \frac{1}{2}\Big(u_0(x+ct) + u_0(x-ct)\Big)$$

Similarly, if $N = 3$, then the integral equation is obtained from the Kirchhoff formula

$$u(x,t) = v(x,t) + \frac{\lambda}{4\pi c^2} \int\limits_{|y-x|<ct} \frac{F(u(y, t - \frac{|x-y|}{c}), y, t - \frac{|x-y|}{c})}{|x-y|}\, d^3y\,,$$

$$v(x,t) = \frac{1}{4\pi c^2 t} \int\limits_{|y-x|=ct} u_1(y)\, dS_y + \frac{1}{4\pi c^2}\frac{\partial}{\partial t}\frac{1}{t} \int\limits_{|y-x|=ct} u_0(y)\, dS_y\,.$$

Using the smoothness properties of the wave potentials established in Chapter 3, one can show that the non-linear operator

$$\hat{K}u(x,t) = G(x,t) * F(u(x,t), x, t)$$

maps any continuous function $u \in C^0(t \geq 0)$ to $C^0(t \geq 0)$ if $F(u,x,t)$ is a continuous function of all arguments.

Suppose that the integral equation has a solution in the class of continuous functions $C^0(t \geq 0)$. Then under some smoothness conditions on F that follow from the smoothness properties of the wave potentials, one can show that this solution is also from the class $C^2(t > 0) \cap C^1(t \geq 0)$ and solves the Cauchy problem. A verification of this assertion in general is left to the reader as an exercise. For example, for $N = 1$, it is sufficient that $F(u,x,t)$ is continuously differentiable for $t \geq 0$ in all three variables.

Thus, the non-linear Cauchy problem is equivalent to the integral equation

$$u = v + \lambda \hat{K} u\,.$$

It is clear that the approach can be used to obtain integral equations that are equivalent to Cauchy problems for other equations such as the heat equation, Maxwell's equations, Schroedinger equation, or the transfer equation.

45.4. Perturbation theory. A traditional approach, often employed for solving integral equations in physics, is based on the so-called *perturbation theory.* Consider a sequence of function defined recursively

$$u_n = v + \lambda \hat{K} u_{n-1}\,, \quad u_0 = v\,.$$

If λ is small enough, then the sequence is expected to converge to a solution in some sense. A proof of this assertion is far from obvious and will be discussed later. Here the idea is illustrated by a simple example where the convergence of the perturbation sequence is not difficult to establish.

45.4.1. An example. Consider the Cauchy problem for a 2D wave equation where

$$F(u,x,t) = \nu(x,t)u$$

and the function $\nu(x,t)$ is continuous and bounded so that

$$|\nu(x,t)| \le \xi^2\,, \quad (x,t) \in \mathbb{R} \times [0,T] = \Omega$$

for any $T > 0$. Of course, ξ may depend on T. In this case, the operator $\hat{K}$ is linear and the perturbation sequence becomes the von Neumann series. Its convergence can be studied by means similar to the case of an oscillator. The function $v(x,t)$ is also assumed to be continuous and bounded:

$$\sup_{\Omega} |v(x,t)| = m_0 < \infty\,.$$

Recall that $v(x,t)$ is the solution to the Cauchy problem for $\lambda = 0$. By d'Alemberts formula, it is bounded if the initial data are bounded. Then for all $(x,t) \in \Omega$

$$\begin{aligned}
|\hat{K}v(x,t)| &\le \frac{\xi^2}{2c}\int_0^t \int_{x-c(t-\tau)}^{x+c(t-\tau)} |v(y,\tau)|\,dy\,d\tau \\
&\le \frac{\xi^2 m_0}{2c}\int_0^t \int_{x-c(t-\tau)}^{x+c(t-\tau)} dy\,d\tau = \frac{\xi^2 t^2}{2}\,m_0\,, \\
|\hat{K}^2 v(x,t)| &\le \frac{\xi^2}{2c}\int_0^t \int_{x-c(t-\tau)}^{x+c(t-\tau)} |\hat{K}v(y,\tau)|\,dy\,d\tau \\
&\le \frac{\xi^4}{4c}\,m_0 \int_0^t \tau^2 \int_{x-c(t-\tau)}^{x+c(t-\tau)} dy d\tau = \frac{\xi^4 t^4}{4!}\,m_0\,, \\
|\hat{K}^n v(x,t)| &\le \frac{(\xi t)^{2n}}{(2n)!}\,m_0
\end{aligned}$$

The latter inequality is easy to prove by mathematical induction. Therefore the terms of the von Neumann series are uniformly bounded by

$$|\lambda^n \hat{K}^n v(x,t)| \le \lambda^n \frac{(\xi T)^{2n}}{(2n)!}\,m_0\,, \quad (x,t) \in \Omega$$

and the series of the bounds converges for any λ

$$m_0 \sum_{n=0}^{\infty} \lambda^n \frac{(\xi T)^{2n}}{(2n)!} = m_0 \cosh\left(\sqrt{|\lambda|}\,\xi T\right)$$

This implies that the von Neumann series converges uniformly to some continuous function on Ω.

Following the reasoning of Section **45.2**, it is not difficult to show that the sum of the von Neumann series is a unique solution to the integral equation. This also proves that the corresponding Cauchy problem has a unique solution (under additional assumptions on smoothness of the function $\nu(x,t)$) that is given by the von Neumann series. Thus, the use of perturbation theory is justified for any λ. It will be shown later that under some additional assumptions on $F(u,x,t)$, a similar result holds for F that is non-linear in u.

45.5. Exercises.

2. Cauchy problem for wave equations. Let

$$F(u,x,t) = \nu(x,t)u\,, \quad |\nu(x,t)| \leq \omega\,, \quad x \in \mathbb{R}^N\,, \quad t \geq 0\,,$$

for some constant ω.
(i) Use the reasoning for the case $N=1$ to show that for $N=2,3$

$$|\hat{K}v(x,t)| = |G(x,t) \star F(v(x,t),x,t)| \leq M_N(t,\omega) \sup_{\Omega} |v(x,t)|$$

where $\Omega = [0,T] \times \mathbb{R}^N$. Find an explicit form $M_N(t,\omega)$.
(ii) Use the above result to investigate the uniform convergence of the von Neumann series:

$$u(x,t) = v(x,t) + \sum_{n=1}^{\infty} \lambda^n \hat{K}^n v(x,t)$$

in $[0,T] \times \mathbb{R}^N$. Find all values of λ for which the series converges uniformly.
(iii) Show that the von Neumann series is a unique solution to the integral equation.
(iv) Find sufficient conditions on smoothness of the function $\nu(x,t)$ under which the integral equation and the Cauchy problem are equivalent, and prove that the Cauchy problem has a unique solution given by the von Neumann series for $N=2,3$.

2. Cauchy problem for a non-linear transfer equation. Consider the Cauchy problem for the transfer equation

$$\frac{1}{c}\frac{\partial u(x,t)}{\partial t} + (s,\nabla_x)u(x,t) + \alpha u(x,t) = F(u(x,t),x,t)\,, \quad t > 0\,,$$
$$u\Big|_{t=0} = u_0(x)$$

where s and x are vectors from $\mathbb{R}^N$, $|s| = 1$, $c > 0$, $\alpha > 0$, and $u_0 \in C^1$. Assume that $F(u, x, t)$ is a continuous function of u, x, and $t \geq 0$.
(i) Use the causal Green's function for the transfer operator to show that if $u(x, t)$ is a solution to the Cauchy problem, then it is also a solution to the integral equation

$$\begin{aligned} u(x,t) &= v(x,t) + \hat{K}u(x,t)\,, \\ \hat{K}u(x,t) &= \int_0^t e^{-\alpha c(t-\tau)} F\Big(u(x - cs(t-\tau), \tau), x - cs(t-\tau), \tau\Big)\, d\tau\,, \\ v(x,t) &= e^{-\alpha ct} u_0(x - cst) \end{aligned}$$

(ii) Show that $\hat{K}u$ is continuous for $t \geq 0$ if u is continuous for $t \geq 0$.
(iii) Find conditions on $F(u, x, t)$ under which a continuous solution of the integral equation is from class $C^1(t > 0) \cap C^0(t \geq 0)$ and solves the Cauchy problem, thus establishing the equivalence of the problems.
(iv) Suppose that

$$F(u, x, t) = \nu(x, t)u$$

where $\nu(x, t)$ is a continuous and bounded function. In this case, the operator $\hat{K}$ is linear. Prove that the von Neumann series $\sum \hat{K}^n v$ converges uniformly on $\mathbb{R}^N \times [0, T]$.
(v) Prove that the sum of the von Neumann series is a continuous function that satisfy the integral equation and that the corresponding Cauchy problem has a unique solution given by the von Neumann series

46. Wave scattering

Integral equations appear in other applications in physics. One example that is especially significant is about wave scattering. There are many physically different waves (electromagnetic, elastic, sound, quantum, etc). However the mathematical tool for studying their scattering is the same and is known as the Lippmann-Schwinger integral equation. Here its derivation is given for a scalar wave equation in $\mathbb{R}^3$. It describes, e.g., sound waves.

46.1. Formulation of the problem. In this section t is real and $\mathbf{x} \in \mathbb{R}^3$. In a homogeneous medium, the wave equation

$$\frac{1}{c^2}\, u''_{tt}(t, \mathbf{x}) - \Delta_x(t, \mathbf{x}) = 0$$

has plane wave solutions

$$u(t, \mathbf{x}) = u_0 e^{-i\omega t + i(\mathbf{k}, \mathbf{x})}\,, \quad \omega = c|\mathbf{k}|$$

The amplitude $u(t, \mathbf{x})$ has a constant value on any plane perpendicular to the vector $\mathbf{k}$ at any fixed moment of time. With changing t, any such plane propagates with the rate c in the direction of the vector $\mathbf{k}$. The constant c is the wave speed, and depends on the physical properties of the medium. If there is an inhomogeneity in the medium, then c is no longer a constant and depends on the position. Any plane propagating in an inhomogeneous medium will be distorted, and this distortion is called a scattered wave, which might intuitively be thought of as a linear combination of plane waves with different wave vectors $\mathbf{k}$. The frequency ω does not change in the scattering process. So, the solution is sought in the form

$$u(t, \mathbf{x}) = e^{-i\omega t}\Big(u_i(\mathbf{x}) + u_s(\mathbf{x})\Big)$$

where $u_i(\mathbf{x})$ and $u_s(\mathbf{x})$ are the amplitudes of the incident and scattered waves. The incident wave satisfies the wave equation without any inhomogeneities:

$$\Delta u_i(\mathbf{x}) + k^2 u_i(\mathbf{x}) = 0\,, \quad k^2 = \frac{\omega^2}{c_0^2}$$

where c_0 is the wave speed in a homogeneous medium. In particular, one can take a plane wave $u_i(\mathbf{x}) = u_0 e^{i(\mathbf{k}, \mathbf{x})}$ where $|\mathbf{k}| = k$. The problem is to find $u_s(\mathbf{x})$ if c is no longer a constant.

Put

$$V(\mathbf{x}) = \frac{\omega^2}{c_0^2} - \frac{\omega^2}{c^2(\mathbf{x})} = k^2\Big(1 - \frac{c_0^2}{c^2(\mathbf{x})}\Big)$$

Then the wave equation for the scattered wave has the form

$$(\Delta + k^2)u_s = Vu_i + Vu_s\,. \tag{46.1}$$

Suppose that the scattering inhomogeneity occupies a bounded region in space, that is, the function V has a bounded support

$$\operatorname{supp} V = \bar{\Omega} \subset B_R$$

so that the scattered wave satisfies the Helmholtz equation as the incident wave

$$(\Delta + k^2)u_s = 0\,, \qquad \mathbf{x} \in \Omega_c = \mathbb{R}^3 \setminus \bar{\Omega}$$

in the complement of Ω. By physical reasons, the scattered wave should be propagating away from a scattering structure. Not every solution to the Helmholtz equation would have this property.

46.1.1. Sommerfeld radiation condition. Recall that regular distributions

$$u_\pm(t, \mathbf{x}) = -\frac{e^{\pm irk}}{4\pi r}\, e^{-i\omega t} = G_\pm(\mathbf{x})e^{-i\omega t}\,, \quad r = |\mathbf{x}|$$

are solutions to the wave equation with a harmonic point-like source:

$$\Big(\frac{1}{c_0^2}\frac{\partial^2}{\partial t^2} - \Delta_x\Big)u_\pm(t, \mathbf{x}) = e^{-i\omega t}\delta(\mathbf{x})$$

Note that $u_\pm$ are smooth functions that satisfy the homogeneous wave equation in any open region that does not contain the origin $\mathbf{x} = \mathbf{0}$. The phase of the solution u_+ remains constant on the sphere of radius $r = r_0 + ct$ that is increasing with increasing time t at the rate c. A physical interpretation of this property is that the function u_+ describes a spherical wave propagating away from the point source. Similarly, the solution u_- describes a spherical waves collapsing onto the point source. Therefore a scattering wave should be a combination of waves like u_+.

The amplitude of an outgoing wave generated by an extended monochromatic source $\rho e^{-i\omega t}$ is given by

$$u_+(t, \mathbf{x}) = e^{-i\omega t}\,(G_+ * \rho)(\mathbf{x}) = e^{-i\omega t}\int_\Omega \frac{e^{ik|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}\,\rho(\mathbf{y})d^3y$$

which can be interpreted as a combination of outgoing spherical waves. The support Ω of sources is bounded. So in the asymptotic region $|\mathbf{x}| = r \to \infty$, one has

$$|\mathbf{x} - \mathbf{y}| = r\Big|\hat{\mathbf{x}} - \frac{\mathbf{y}}{r}\Big| = r\Big(1 + O\Big(\frac{1}{r}\Big)\Big)\,,$$

where $\hat{\mathbf{x}}$ is the unit vector in the direction of $\mathbf{x}$. Therefore

$$u_+(t,\mathbf{x}) = -e^{-i\omega t}\,\frac{e^{ikr}}{4\pi r}\Big(1 + O\Big(\frac{1}{r}\Big)\Big)$$

for large r. This wave satisfies the asymptotic condition

$$\lim_{r\to\infty} r\Big(\frac{\partial u_+}{\partial r} - iku_+\Big) = 0\,.$$

Note that the convolution $G_- * \rho$ does not satisfies this condition because of the different sign in the exponential.

Thus, it is demanded that the scattered wave satisfies this asymptotic condition

$$\lim_{r\to\infty} r\Big(\frac{\partial u_s}{\partial r} - iku_s\Big) = 0\,. \tag{46.2}$$

It is called the *Sommerfeld radiation (or outgoing wave) condition.*

46.2. Lippmann-Schwinger equation. Suppose that u_s is a solution to (**46.1**) that satisfies the Sommerfeld condition (**46.2**). Then owing to the analogy with a non-homogeneous Helmholtz equation, u_s is sought in the form

$$u_s(\mathbf{x}) = (G_+ * \tilde{\rho})(\mathbf{x})\,,$$

where the density $\tilde{\rho}$ is to be found and it is supported in $\bar{\Omega}$. Then

$$Vu_i + Vu_s = (\Delta + k^2)u_s = (\Delta + k^2)G_+ * \tilde{\rho} = \delta * \tilde{\rho} = \tilde{\rho}$$

It follows that if, without loss of generality, one sets

$$\tilde{\rho} = Vu_i + Vw$$

where w is an unknown function, then

$$Vu_s = Vw$$

so that w is equal to u_s in $\bar{\Omega}$. Therefore restricting the relation $u_s = G_+ * \tilde{\rho}$ to $\bar{\Omega}$, it is concluded that w satisfies the equation

$$w = G_+ * (Vu_i) + G_+ * (Vw)$$

Suppose that V is bounded, then the linear integral operator

$$\hat{K}w(\mathbf{x}) = G_+ * (Vw)(\mathbf{x}) = -\int_\Omega \frac{e^{ik|\mathbf{x}-\mathbf{y}|}V(\mathbf{y})}{4\pi|\mathbf{x}-\mathbf{y}|}\,w(\mathbf{y})d^3y$$

maps a continuous function on $\bar{\Omega}$ to a continuous function on $\bar{\Omega}$:

$$\hat{K}:\; C^0(\bar{\Omega}) \;\to\; C^0(\bar{\Omega})$$

This follows from properties of functions defined by potential-like integrals studied earlier. Then $w \in C^0(\bar{\Omega})$ is a solution to the integral equation

$$w = f + \hat{K}w\,, \quad f = \hat{K}(Vu_i) \in C^0(\bar{\Omega})$$

This integral equation is known as the *Lippmann-Schwinger equation* for scattered waves.

Now suppose that $w \in C^0(\bar{\Omega})$ is a solution to the Lippmann-Schwinger equation. Then

$$u_s(\mathbf{x}) = -\int_\Omega \frac{e^{ik|\mathbf{x}-\mathbf{y}|}V(\mathbf{y})}{4\pi|\mathbf{x}-\mathbf{y}|}\Big(u_i(\mathbf{y}) + w(\mathbf{y})\Big)d^3y\,.$$

If V vanishes at the boundary of Ω smoothly enough, then u_s can have continuous partial derivatives everywhere in space as one might infer from the theory of potential-like integrals. In other words, under some assumptions on smoothness of V, one can show that u_s is twice continuously differentiable and satisfies the original differential equation (**46**.**1**) and the Sommerfeld outgoing wave condition (**46**.**2**). The incident and scattered waves in the region occupied by the scatterers play the role of sources for the scattered wave in the asymptotic region. If u_s is known in Ω, then u_s is known in the whole space. The existence and uniqueness of a solution to the integral equation implies the existence and uniqueness of a solution to the scattering problem.

46.3. Von Neumann series for the Lippmann-Schwinger equation. An approximate solution to the Lippmann-Schwinger equation can be obtained by replacing the integral by a suitable Riemann sum just like in the case of the Volterra equation. If I_n is a partition cell in Ω and U_n is the value of u_s in I_u, then the equation is reduced to the same linear algebra problem as in the case of the Volterra equation where the matrix elements of $\mathbf{K}$ are

$$K_{jn} = \int_{I_n} \frac{e^{ik|\mathbf{x}_j-\mathbf{y}|}V(\mathbf{y})}{4\pi|\mathbf{x}_j-\mathbf{y}|}\,d^3y$$

Note that for the diagonal elements K_{jj}, the integral converges (the singularity at $\mathbf{y} = \mathbf{x}_j$ is integrable). In contrast to the Volterra case, the matrix $\mathbf{K}$ is a general matrix and, hence, the inverse of $\mathbf{I} - \mathbf{K}$ may or may not exist. For example, if the matrix $\mathbf{K}$ has a unit eigenvalue, then $\mathbf{I} - \mathbf{K}$ is not invertible and the solution may not exist and even if it exists it is not unique (recall the Fredholm alternative in the linear algebra). This suggests that more sophisticated methods are needed for analyzing the Lippmann-Schwinger equation than a von Neumann series. However, the von Neumann series is not totally useless in this case.

Let $|V(\mathbf{x})| \le V_0$. Then one can show that the von Neumann series converges uniformly on $\bar{\Omega}$ to a unique solution *if* V_0 *is small enough.* Suppose that the amplitude of an incident wave is bounded (which is always the case in any physical situation)

$$m_0 = \sup |u_i(\mathbf{x})|$$

Let R be the diameter of $\bar{\Omega}$ (the largest distance between two points in $\bar{\Omega}$). Then for any $\mathbf{x}$ in $\bar{\Omega}$, the region $\bar{\Omega}$ is contained in the ball $B_R(\mathbf{x})$ centered at $\mathbf{x}$. Therefore the following estimate holds

$$|\hat{K}u_i(\mathbf{x})| \le \frac{m_0 V_0}{4\pi} \int_\Omega \frac{d^3y}{|\mathbf{x}-\mathbf{y}|} \le \frac{m_0 V_0}{4\pi} \int_{B_R(\mathbf{x})} \frac{d^3y}{|\mathbf{x}-\mathbf{y}|} = \frac{1}{2}\, m_0 V_0 R^2$$

where the last integral can be evaluated in spherical coordinates. Since this inequality is true for any $\mathbf{x}$, it is concluded that

$$\sup |\hat{K}u_i(\mathbf{x})| \le \frac{1}{2}\, m_0 V_0 R^2\,.$$

Similarly, replacing u_i by $\hat{K}u_i$ in the above analysis, it is concluded that

$$|\hat{K}^2 u_i(\mathbf{x})| \le \frac{1}{2}\, V_0 R^2 \sup |\hat{K}u_i(\mathbf{x})| = m_0 \Big(\frac{V_0 R^2}{2}\Big)^2$$

Using the mathematical induction, the terms of the von Neumann series are shown to be uniformly bounded

$$|\hat{K}^n u_i(\mathbf{x})| \le \sup |\hat{K}^n u_i(\mathbf{x})| \le m_0 \Big(\frac{V_0 R^2}{2}\Big)^n$$

Therefore the von Neumann series

$$u_s(\mathbf{x}) = u_i(\mathbf{x}) + \sum_{n=1}^{\infty} \hat{K}^n u_i(\mathbf{x})$$

converges uniformly to a unique solution of the Lippmann-Schwinger equation, provided

$$\frac{V_0 R^2}{2} < 1\,.$$

In particular, if $c(\mathbf{x})$ does not deviate from c_0 by orders in magnitude, this condition is ensured if the dimensionless parameter kR is small enough. Note that $\lambda = 2\pi/k$ is the wavelength. For this reason, the stated approximation is called a *long wavelength approximation*, and it is applicable when dimensions of the scattering region are sufficiently small as compared to the wavelength of the incident wave.

46.4. Scattering by point-like particles. [1] Suppose that $\Omega = B_R$ is a ball of radius R centered at the origin, and $V = V_0$ has a constant value in it. Let us assume that

$$kR \ll 1$$

so that the long wavelength approximation is applicable. The parameter k defines a characteristic length over which the phase factor e^{ikr} has a significant variation. If r lies in an interval $[r_0 - R, r_0 + R]$, then the phase factor is close to constant e^{ikr_0}. By setting $r_0 = 0$,

$$\hat{K}w(\mathbf{x}) = \frac{V_0}{4\pi}\Big(\int_{|\mathbf{y}|<R} \frac{w(\mathbf{y})}{|\mathbf{y}|}\, d^3y + O(kR)\Big)\,.$$

In the leading order of this approximation $\hat{K}w$ is a constant function. Therefore, the solution to the integral equation should be sought among constant functions in this approximation. Let

$$w(\mathbf{x}) = w_s$$

Then using the spherical coordinates to evaluate the integral

$$\hat{K}w(\mathbf{x}) = \frac{1}{2}V_0R^2w_s$$

The integral equation becomes an algebraic equation in the long wavelength approximation:

$$w_s = \frac{1}{2}V_0R^2w_s + \frac{1}{2}V_0R^2u_0$$

where $u_0 = u_i(\mathbf{0})$, because the incident wave does not vary much within the scatterer (owing to the condition $kR \ll 1$). The solution reads

$$w_s = \frac{\frac{1}{2}V_0R^2}{1-\frac{1}{2}V_0R^2}\, u_0$$

The scattered wave is given by

$$u_s(\mathbf{x}) = \frac{V_0u_0}{1-\frac{1}{2}V_0R^2}\int\limits_{|\mathbf{y}|<R} \frac{e^{ik|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}\, d^3y\,, \quad |\mathbf{x}| > R$$

In the leading order of the small parameter kR, the term $\frac{1}{2}V_0R^2$ can be neglected in the denominator. By the integral mean value theorem, the integral is proportional to the volume of the scatterer so that

$$u_s(\mathbf{x}) \sim \frac{4\pi R^3V_0}{3}\,\frac{e^{ik|\mathbf{x}|}}{4\pi|\mathbf{x}|}\,, \quad |\mathbf{x}| \to \infty$$

[1]To be revised to include a systematic expansion of the solution in powers of kR, leading order (s-wave), next-to-leading order (p-wave), etc.

The result can be extended to a collection of small particles of radii R_j, $kR_j \ll 1$, centered at points $\mathbf{x}_j$, assuming that all particles lie in ball of some radius (to make the support of the scatterers bounded). In this case, Ω consists of non-overlapping balls $B_{R_j}(\mathbf{x}_j)$, and V is assumed to take a constant value V_j on each ball. So, the solution to the integral equation should be sought as a piece-wise constant function

$$w(\mathbf{x}) = w_j\,, \quad |\mathbf{x}-\mathbf{x}_j| < R_j\,, \quad j = 1, 2, ..., N$$

where N is total number of scatterers because

$$e^{ik|\mathbf{x}-\mathbf{y}|} = e^{ik|\mathbf{x}_j-\mathbf{y}|} + O(kR_j)\,, \quad \mathbf{x} \in B_{R_j}(\mathbf{x}_j)\,,$$

where $\mathbf{y}$ lies in any of the balls. Therefore, in this approximation the action of $\hat{K}$ on such w returns a function that has constant values on each scatterer:

$$\hat{K}w(\mathbf{x}) = \sum_{j=1}^{N} K_{nj} w_j\,, \quad |\mathbf{x}-\mathbf{x}_n| < R_n$$

$$K_{nj} = V_j \int\limits_{|\mathbf{y}-\mathbf{x}_j|<R_j} \frac{e^{ik|\mathbf{x}_n-\mathbf{y}|}}{|\mathbf{x}_n-\mathbf{y}|}\, d^3y$$

If u_{0j} is the value of the incident wave amplitude on the jth scatterer, then the unknown constants w_j satisfy the linear equation

$$w_n = \sum_{j=1}^{N} K_{nj} w_j + \sum_{j=1}^{N} K_{nj} u_{0j}$$

If $\boldsymbol{w}$ and $\boldsymbol{u}_0$ are $N-$dimensional vectors with components w_j and u_{0j}, respectively, and $\boldsymbol{K}$ is the matrix with matrix elements K_{nj}, then the equation can be cast in the matrix form

$$\boldsymbol{w} = \boldsymbol{K}\boldsymbol{w} + \boldsymbol{K}\boldsymbol{u}_0$$

It has a unique solution

$$\boldsymbol{w} = (\boldsymbol{I} - \boldsymbol{K})^{-1}\boldsymbol{K}\boldsymbol{u}_0$$

provided the inverse matrix exists. The diagonal elements of $\boldsymbol{K}$ are of order $(kR_j)^2$ because just like in the single particle case

$$K_{jj} = \frac{1}{2} V_j R_j^2 \sim (kR_j)^2$$

The non-diagonal matrix elements contain an extra small factor

$$K_{nj} = \frac{e^{ik|\mathbf{x}_n-\mathbf{x}_j|}}{4\pi|\mathbf{x}_n-\mathbf{x}_j|} \cdot \frac{4\pi R_j^3}{3} V_j = K_{jj} e^{ikD_{jn}} \frac{2R_j}{3D_{jn}}\,, \quad j \neq n$$

where $D_{jn} = |\mathbf{x}_n - \mathbf{x}_j|$ is the distance between the jth and nth particles. So, if D_{jn} is of order of the wavelength of the incident wave or greater, $D_{jn}l \sim 1$, then

$$|K_{jn}| \sim (kR_j)^3$$

so that $\boldsymbol{K}$ is close to the diagonal matrix and its Frobenius norm is approximately the largest diagonal element:

$$\|\boldsymbol{K}\| = \frac{1}{2}\max_j\{V_j R_j^2\} + O((kR)^3) < 1$$

for small enough $R = \max_j\{R_j\}$. Hence, the inverse matrix in question exists. In particular,

$$\boldsymbol{w} + \boldsymbol{u}_0 = (\boldsymbol{I} - \boldsymbol{K})^{-1}\boldsymbol{K}\boldsymbol{u}_0 + \boldsymbol{u}_0 = (\boldsymbol{I} - \boldsymbol{K})^{-1}\boldsymbol{u}_0$$

Therefore the scattered amplitude in the asymptotic region is given by

$$u_s(\mathbf{x}) = \sum_{j=1}^{N} V_j S_j \int\limits_{|\mathbf{y}-\mathbf{x}_j|<R_j} \frac{e^{ik|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}\, d^3y\,, \quad |\mathbf{x}| > \max_j\{|\mathbf{x}_j| + R_j\}$$

where S_j is the jth component of the vector $\boldsymbol{w} + \boldsymbol{u}_0$:

$$S_j = \Big((\boldsymbol{I} - \boldsymbol{K})^{-1}\boldsymbol{u}_0\Big)_j\,.$$

Using the integral mean value theorem to evaluate the integral and obtain the asymptotic scattered amplitude

$$u_s(\mathbf{x}) \sim \sum_{j=1}^{N} \frac{e^{ik|\mathbf{x}-\mathbf{x}_j|}}{4\pi|\mathbf{x}-\mathbf{x}_j|}\,\rho_j\,, \qquad \rho_j = \frac{4\pi}{3}\, R_j^3 V_j S_j$$

In the asymptotic region, $|\mathbf{x}| \to \infty$, the scattered wave is a superposition of spherical waves emitted by a point-like source of strength ρ_j and located at a point $\mathbf{x}_j$.

46.5. Remarks. The examples of the initial value problem for an anharmonic oscillator and the wave scattering illustrate a general approach to study some problems formulated in terms of differential equations with some boundary or initial conditions by converting them into equivalent integral equations by means of suitable Green's functions for the differential operators involved. In many instances, a solution or its suitable approximation is easier to find for the integral equation than for the original problem. In what follows, a general problem of solving the equation

$$u = \lambda\hat{K}u + f$$

where $\hat{K}$ is an operator acting in some functional space will be analyzed, where λ is a numerical parameter. Since u belongs to an infinite dimensional space, no matrix (finite-dimensional) analogy might be deceiving when investigating the existence of the inverse

$$u = (I - \lambda\hat{K})^{-1}f$$

In physics, one often employ a *perturbation theory* which amounts to seeking the solution as a power series in the parameter λ:

$$u = u_0 + \lambda u_1 + \lambda^2 u_2 + \cdots$$

In th operator theory, the inverse $(I - \lambda\hat{K})^{-1}$ is called the *resolvent* of the operator $\hat{K}$. Its existence or non-existence for some λ depends on fundamental properties of $\hat{K}$. As a consequence, the operator theory provides a justification of the perturbation theory and answers the question for what λ the perturbation series converges to a solution.

47. Linear spaces

DEFINITION **47.1**. (Linear space)
A collection of elements X is called a vector space *and its elements are called* vectors *if there are two operations on vectors,* addition *and* scalar multiplication*, that have the following properties*
Addition*:* $u + v \in X\,, \forall u, v \in X$

$$\begin{aligned}
&(1a) \quad u + v = v + u\,,\\
&(2a) \quad u + (v + w) = (u + v) + w\,,\\
&(3a) \quad \exists!\ 0 \in X:\ u + 0 = u\,,\\
&(4a) \quad \forall u \in X\ \exists!\ (-u) \in X:\ u + (-u) = 0
\end{aligned}$$

Scalar multiplication*:* $\alpha u \in X\,,\ \forall u \in X\,,\ \forall \alpha \in \mathbb{F}$

$$\begin{aligned}
&(1m) \quad \alpha(\beta u) = (\alpha\beta) u\,,\\
&(2m) \quad 1\,u = u\,,\\
&(3m) \quad (\alpha + \beta) u = \alpha u + \beta u\,,\\
&(4m) \quad \alpha(u + v) = \alpha u + \alpha v
\end{aligned}$$

If $\mathbb{F} = \mathbb{R}$, then X is a real vector space*, and if $\mathbb{F} = \mathbb{C}$, then X is a* complex vector space*.*

By the commutative and distributive properties, (1a) and (2a), the order of term in any linear combination of vectors is not relevant and it will be written as

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n$$

It also follows from the axioms that

$$u \overset{(2m)}{=} 1\,u = (1 + 0)u \overset{(3m)}{=} 1\,u + 0\,u = u + 0\,u \quad \overset{(3a)}{\Rightarrow} \quad 0\,u = 0 \in X$$

In turn, it follows from the latter that

$$0 = 0\,u = [1 + (-1)]\,u = u + (-1)u \quad \overset{(4a)}{\Rightarrow} \quad (-1)u = -u$$

If X contains at least one element $u \neq 0$, then X has infinitely many elements αu, where $\alpha \in \mathbb{F}$.

DEFINITION **47.2**. (Linear manifold)
A subset $M \subset X$ of a linear space X that is closed with respect to the addition and scalar multiplication is called a linear manifold*, that is,*

$$u + v \in M\,, \quad \alpha u \in M \quad \forall u, v \in M\,, \quad \forall \alpha \in \mathbb{F}$$

Any two linear manifolds in X have at least one common element, the zero element $0 \in X$, because $0\,u = 0$ for any $u \in X$.

47.1. Examples of linear spaces.

Example 1. The space of continuous functions is a linear space,

$$f:\ \Omega \to \mathbb{F}\,, \quad f \in C^0(\Omega)$$

if the addition of functions is defined by

$$(f+g)(x) = f(x) + g(x)\,, \quad \forall x \in \Omega$$

and the scalar multiplication is defined by

$$(\alpha f)(x) = \alpha f(x)\,, \quad \forall x \in \Omega$$

Clearly, the space $X = C^0(\Omega)$ is closed with respect to the addition and scalar multiplication, and all the axions of these operations are fulfilled. Note that the zero function $g(x) = 0$ is the only continuous function that has the property $f(x) + g(x) = 0$.

Consider $X = C^0([-a, a])$. Then the even and odd functions form two linear manifolds with the only common element being the zero function.

Example 2. Let $\mathcal{P}_n$ be a set of all polynomials of degree at least n of one real variable. Let $\mathcal{P}$ be a set of all polynomials. Then $\mathcal{P}$ is a linear space with addition and scalar multiplication defined as for real-valued functions, $\mathcal{P}_n$ is a linear manifold in $\mathcal{P}$. Note also that $\mathcal{P}$ is a linear manifold in $C^0(\mathbb{R})$.

Example 3. Consider a linear homogeneous differential equation

$$Lu(x) = u^{(n)}(x) + a_1(x)u^{(n-1)}(x) + \cdots + a_{n-1}(x)u'(x) + a_n(x)u(x) = 0$$

The set of its solutions a vector space because $L(u+v) = Lu + Lv = 0$ and $L(\alpha u) = \alpha Lu = 0$ if $Lu = Lv = 0$ for any number α. This set is also a linear manifold in the linear space $C^n(\mathbb{R})$.

47.2. Dimension of a linear space.

Definition **47.3**. (Span of a set)
Let $U \subset X$ be a collection of vectors. The set of all linear combinations of elements of A is called the span *of A:*

$$v = \alpha_1 u_{j_1} + \alpha_2 u_{j_2} + \cdots + \alpha_n u_{j_n} \in \mathrm{Span}\, A$$

for any choice of coefficients α_k, $k = 1, 2, ..., n$ and any finite subset elements of U, $\{u_{j_k}\}_{k=1}^n \subset U$.

If U is a finite set, then its span is a collection of all linear combinations of elements of U.

Definition 47.4. (Linearly independent set)
A finite set $U = \{u_1, u_2, ..., u_n\}$ *is called* linearly dependent *if* $\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n = 0$ *for some* α_j, $j = 1, 2, ..., n$, *that do not vanish simultaneously. If* $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$ *is the only solution to the equation:*

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n = 0$$

then the set U *is called* linearly independent. *An infinite set* U *is linearly independent if any finite subset of* U *is linearly independent.*

For example, the set $U = \{1, x, x^2, ..., x^n\}$ is linearly independent in $\mathcal{P}_n$ or $C^(\mathbb{R})$. The set $U = \{x^n\}_0^\infty$ is linearly independent in $\mathcal{P}$ and $C^(\mathbb{R})$.

Definition 47.5. (Dimension of a linear space)
The dimension *of a linear space* X *is equal to* n, $\dim X = n$, *if* X *has a set of* n *linearly independent vectors, but any set of* $n+1$ *vectors is linearly dependent. A vector space is said to have* infinite dimension, $\dim X = \infty$, *if for each positive integer* n *one can find* n *linearly independent vectors in* X.

For example, $\dim \mathcal{P}_n = n + 1$ because the linearly independent set $\{x^k\}_0^n$ has $n+1$ elements. Evidently, $\dim \mathcal{P} = \infty$ and $\dim C^0 = \infty$.

Definition 47.6. (A basis in a linear space)
A subset $\{v_1, v_2, ..., v_n\} \subset X$ *in a finite-dimensional linear space* X, $\dim X = n < \infty$, *is called a* basis *in* X *if any element of* X *is a unique linear combination of* v_j:

$$\forall u \in X \quad \exists! \{\alpha_1, \alpha_2, ..., \alpha_n\} : \quad u = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$$

The following theorem describes basis properties of a basis

Theorem 47.1. (properties of a basis)
The basis vectors are linearly independent. Any set of n *linearly independent vectors in an* $n-$*dimensional vector space is a basis.*

Clearly the set of monomials $\{x^k\}_0^n$ is a basis in the space $\mathcal{P}_n$ of all polynomials of degree at most n. What is a basis in an infinite dimensional linear space? A basis in an infinite dimensional space should contain infinitely many elements. For example, any element of the space $\mathcal{P}$ is a unique linear combination of monomials $\{x^k\}_0^\infty$. So, in this sense the set $\{x^k\}_0^\infty$ is a basis in $\mathcal{P}$. However, this picture is not so easy to extend to the space of continuous functions. Evidently, not every continuous function is a polynomial. So, one can try to consider linear combinations with arbitrary large number of terms. For example, a *convergent* power series

$$f(x) = \sum_{n=0}^{\infty} \alpha_n x^n$$

defines a continuous function in the interval of convergence. Can all continuous functions be defined by power series? Evidently no because a power series defines a function from class C^{∞}. Yet, the very concept of linear combinations with infinite number of terms (series) requires a definition of *convergence*, or the measure for how close two elements of a linear space are to one another.

48. Banach spaces

48.1. Normed spaces.

DEFINITION **48.1**. *A linear space X is a* normed space *if there exists a real-valued function* $\|\cdot\|: X \to \mathbb{R}$ *with the following properties*

$$\begin{aligned} &\text{(i)} && \|u\| \geq 0\,, \\ &\text{(ii)} && \|u\| = 0 \quad \Leftrightarrow \quad u = 0\,, \\ &\text{(iii)} && \|cu\| = |c|\|u\|\,, \quad \forall c \in \mathbb{C}\,, \\ &\text{(iv)} && \|u+v\| \leq \|u\| + \|v\| \end{aligned}$$

for all u and v from X.

The stated properties are called the *norm axioms*. Note that every normed space is a metric space if the distance between any two elements is defined by

$$d(u,v) = \|u - v\|$$

Clearly, the distance axioms are satisfied.

Space of bounded functions. Let $X = \mathcal{B}(\Omega)$ be a set of bounded functions on $\Omega \subset \mathbb{R}^n$:

$$|u(x)| \leq M\,, \quad \forall x \in \Omega$$

It is a linear space because a linear combination of bounded functions is bounded. If $|u(x)| \leq M_1$ and $|v(x)| \leq M_2$, then

$$|c_1 u(x) + c_2 v(x)| \leq |c_1||u(x)| + c_2|v(x)| \leq |c_1|M_1 + c_2|M_2|\,.$$

Put

$$\|u\|_\infty = \sup_\Omega |u(x)|$$

Clearly, the supremum exists for any bounded function. So, $\|\cdot\|_\infty$ is a real-valued function on the space of bounded functions. It is non-negative and

$$\|u\|_\infty = 0 \quad \Leftrightarrow \quad u(x) = 0$$

The third axiom is also satisfied. Finally

$$|u(x) + v(x)| \leq |u(x)| + |v(x)| \leq \|u\|_\infty + \|v\|_\infty\,, \quad \forall x \in \Omega$$

Since the above inequality holds for all x in Ω, one can take the supremum in the left side so that

$$\|u+v\|_\infty \leq \|u\|_\infty + \|v\|_\infty$$

Thus the space of bounded function equipped with the supremum norm is a normed space.

The set of bounded continuous functions $C^0(\Omega) \cap \mathcal{B}(\Omega)$ is a linear manifold in $\mathcal{B}(\Omega)$. Note that if Ω is not closed and bounded, then $C^0(\Omega)$ can contain unbounded continuous functions.

48.2. Space of square integrable functions. A function $f : \Omega \subset \mathbb{R}^N \to \mathbb{C}$ is said to be *square integrable* on Ω if

$$\int_\Omega |f(x)|^2 \, d^N x < \infty$$

The integral is understood in the Lebesgue sense.

PROPOSITION **48.1**. *The set of square integrable functions is a linear space*

Indeed, a function scaled by a complex number α is square integrable:

$$\int_\Omega |\alpha u(x)|^2 \, d^N x = |\alpha|^2 \int_\Omega |u(x)|^2 \, d^N x < \infty$$

If a and b are real, then

$$(a-b)^2 \geq 0 \quad \Rightarrow \quad ab \leq \frac{1}{2}\, a^2 + \frac{1}{2}\, b^2$$

The from the triangle inequality for complex numbers z and w:

$$|z+w| \leq |z| + |w|$$

it follows that

$$|z+w|^2 \leq |z|^2 + |w|^2 + 2|z||w| \leq 2|z|^2 + 2|w|^2$$

Therefore

$$\int_\Omega |u(x)+v(x)|^2 d^N x \leq 2\int_\Omega |u(x)|^2 \, d^N x + 2\int_\Omega |v(x)|^2 d^N x < \infty$$

Thus, a linear combination of square integrable functions is square integrable.

Put

$$\|u\|_2 = \sqrt{\int_\Omega |u(x)|^2 d^N x}$$

Clearly the first axiom is satisfied. The second axiom is not satisfied because

$$\|u\|_2 = 0 \quad \Rightarrow \quad u(x) = 0 \;\; a.e.$$

So, $u(x)$ is not the zero function.

The space $\mathcal{L}_2(\Omega)$. Consider an equivalence class that consists of square integrable functions that are equal to one another almost everywhere. The set of square integrable functions is split into such equivalence classes. Let $\mathcal{L}_2(\Omega)$ be the set of all such equivalence classes. In other words, two elements u and v of $\mathcal{L}_2(\Omega)$ are are two equivalence classes. They are are equal if any functions from the corresponding equivalence classes are equal almost everywhere

$$u = v \quad \text{in} \quad \mathcal{L}_2(\Omega) \quad \Leftrightarrow \quad u(x) = v(x) \quad a.e.$$

Similarly, u and v are distinct elements of $\mathcal{L}_2(\Omega)$ if the difference of any two functions from the equivalence classes u and v is not equal to zero almost everywhere. The set $\mathcal{L}_2(\Omega)$ is a linear space because a linear combination of functions that are zero almost everywhere is zero almost everywhere. In particular, the zero element in $\mathcal{L}_2(\Omega)$ is the collection of all functions that are zero almost everywhere. Hence,

$$\|u\|_2 = 0 \quad \Leftrightarrow \quad u = 0 \quad \text{in } \mathcal{L}_2(\Omega)$$

so that the second axiom is satisfied in $\mathcal{L}_2(\Omega)$.

PROPOSITION **48.2**. *The linear space $\mathcal{L}_2(\Omega)$ equipped with the norm $\|\cdot\|_2$ is a normed space.*

Let us prove first the following inequality:

$$\Big|\int_\Omega \mathrm{Re}\, u\bar{v}\, d^N x\Big| \le \|u\|_2 \|v\|_2$$

First, note that the inequality holds if $v = 0$. If $v \neq 0$, consider the function of a real variable t

$$h(t) = \|u + tv\|_2^2 = \|u\|_2^2 + t^2\|v\|_2^2 + 2\int_\Omega \mathrm{Re}\, u\bar{v}\, d^N x$$

By construction, $h(t) \ge 0$ for all real t and attains its absolute minimum when its derivative vanishes:

$$h'(t) = 0 \quad \Rightarrow \quad t = t^* = -\frac{1}{\|v\|_2^2}\int_\Omega \mathrm{Re}\, u\bar{v}\, d^N x$$

Therefore the condition $h(t^*) \ge 0$ implies that

$$\|u\|_2^2 - \frac{1}{\|v\|_2^2}\left(\int_\Omega \mathrm{Re}\, u\bar{v}\, d^N x\right)^2 \ge 0$$

The desired inequality follows from this one.

Now it is dot difficult to verify the triangle inequality

$$\|u+v\|_2^2 = \|u\|_2^2 + \|v\|_2^2 + 2\int_\Omega \mathrm{Re}\, u\bar{v}\, d^N x$$
$$= \|u\|_2^2 + \|v\|_2^2 + 2\|u\|_2\|v\|_2 = \Big(\|u\|_2 + \|v\|_2\Big)^2$$

Taking the square root of this inequality, the triangle inequelity is established:

$$\|u+v\|_2 \le \|u\|_2 + \|v\|_2 \,.$$

Thus, the space $\mathcal{L}_2(\Omega)$ is a normed space.

Space $C_2^0(\Omega)$. Consider a collection of all continuous functions on Ω that are also square integrable. This collection is denoted by

$$C_2^0(\Omega) = C^0(\Omega) \cap \mathcal{L}_2(\Omega)$$

This collection is closed relative to addition and multiplication by a number and, hence, it is a linear manifold in the space of square integrable functions. It is also a normed space as the subspace of $\mathcal{L}_2(\Omega)$.

Suppose that Ω is closed and bounded. Then a linear space $C^0(\Omega)$ is a subspace of $\mathcal{L}_2(\Omega)$ and also a subspace of bounded functions $\mathcal{B}(\Omega)$:

$$C^0(\Omega) \subset \mathcal{L}_2(\Omega)\,, \qquad C^0(\Omega) \subset \mathcal{B}(\Omega)\,.$$

Consequently, one can define two different norms on a space of continuous functions, the supremum and $\mathcal{L}_2$ norms. As is shown below, properties of these normed spaces are quite different, despite that they have identical elements.

48.3. Other examples of normed spaces.

Space l_p. Consider a set of all complex sequences $\{u_n\} \subset \mathbb{C}$ with elements that have summable powers

$$\sum_n |u_n|^p < \infty\,, \quad p \ge 1$$

Put

$$\|u\|_p = \left(\sum_n |u_n|^p\right)^{\frac{1}{p}}$$

One can show that the norm axioms are satisfied for $\|\cdot\|_p$. The corresponding norm space is denoted by l_p.

The space l_∞ consists of all bounded complex sequences:

$$\sup_n\{|u_n|\} < \infty$$

It becomes a normed space if the norm is defined by

$$\|u\|_\infty = \sup_n\{|u_n|\}$$

By the necessary condition for convergence of a series $\sum |u_n|^p$, the sequence $\{u_\}$ must converge to 0

$$\lim_{n\to\infty} u_n = 0$$

Every convergent sequence in a complex plane is bounded:

$$|u_n| \leq M < \infty$$

for all n. Hence every element of l_p also belongs to l_∞

$$l_p \subset l_\infty\,, \quad p \geq 1\,.$$

By the comparison test the existence of $\|u\|_1$ implies the existence of $\|u\|_p$:

$$\sum_n |u_n| < \infty \quad \Rightarrow \quad \sum_n |u_n|^p < \infty$$

Indeed, since $|u_n| \to 0$, one can say that $|u_n| < 1$ for all $n > N$ for some integer N so that for any $p \geq 1$, $|u_n|^p < |u_n|$ for all $n > N$, and the existence $\|u\|_p$ follows from the comparison test. Therefore every sequence of l_1 also belongs to l_p:

$$l_1 \subset l_p \subset l_\infty\,, \quad p \geq 1\,.$$

One can show that

$$\lim_{p\to\infty} \|u\|_p = \|u\|_\infty\,, \quad \forall u \in l_1$$

This explains the notation l_∞.

Space $\mathcal{L}_p(\Omega)$. Consider a collection of functions whose power is Lebesgue integrable on Ω:

$$\int_\Omega |u(x)|^p d^N x < \infty\,, \quad p \geq 1$$

Just like in the case of space of $\mathcal{L}_2(\Omega)$, consider the equivalence classes of such functions, the functions u and v are equivalent if $u(x) = v(x)$ almost everywhere. Then space of such equivalence classes become a normed space if the norm is defined by

$$\|u\|_p = \left(\int_\Omega |u(x)|^p d^N x\right)^{\frac{1}{p}}$$

One can show that the norm axioms are satisfied.

Let Ω be bounded and closed. Then any continuous function on Ω is an $\mathcal{L}_p$ function and also a bounded function, that is, the space of continuous functions on a closed bounded region in a Euclidean space is a subspace of both the space of bounded functions $\mathcal{B}(\Omega)$ and $\mathcal{L}_p(\Omega)$. On this space of continuous functions, one can define either any of $\mathcal{L}_p$ norms of the supremum norm. One can show that

$$\lim_{p\to\infty} \|u\|_p = \|u\|_\infty = \sup_\Omega |u(x)|$$

This explains the notation for the supremum norm.

48.4. Lebesgue spaces. A collection of all functions whose $p^{\rm th}$ power is Lebesgue integrable on Ω will be denoted by $\mathcal{L}_p(\Omega)$:

$$f \in \mathcal{L}_p(\Omega) \quad \Leftrightarrow \quad \int_\Omega |f(x)|^p \, d^N x < \infty\,, \quad p \geq 1$$

48.5. Hölder's inequality. Suppose that

$$f \in \mathcal{L}_p(\Omega)\,, \quad g \in \mathcal{L}_q(\Omega)\,, \quad \frac{1}{p} + \frac{1}{q} = 1\,, \quad p > 1\,,\ q > 1\,.$$

Then the product fg is Lebesgue integrable on Ω and[2]

$$\Big| \int_\Omega f(x) g(x)\, d^N x \Big| \leq \left(\int_\Omega |f(x)|^p \, d^N x \right)^{\frac{1}{p}} \left(\int_\Omega |g(x)|^q \, d^N x \right)^{\frac{1}{q}}$$

48.6. $\mathcal{L}_p(\Omega)$ is a linear space. A collection of functions is a linear space if any linear combination of its elements belongs to the collection. Clearly, the function $cf(x)$ belongs to $\mathcal{L}_p(\Omega)$ for any constant c if f is from $\mathcal{L}_p(\Omega)$. Therefore, in order to verify if $\mathcal{L}_p(\Omega)$ is a linear space, it is sufficient to show that $|f+g|^p$ is integrable if $|f|^p$ and $|g|^p$ are integrable. It was already shown that $\mathcal{L}$ is a linear space. For $p > 1$, note that the function z^p is convex for $z > 0$. Recall that a convex positive function $f(z)$, $z > 0$, has the following characteristic property

$$f(tz_1 + (1-t)z_2) \leq t f(z_1) + (1-t) f(z_2)\,, \quad 0 \leq t \leq 1\,,$$

which means that the secant line through any two points on the graph of $f(z)$ lies above the graph. In particular, putting $t = \frac{1}{2}$ and $f(z) = z^p$, one infers from the above inequality that

$$(z_1 + z_2)^p \leq 2^{p-1}(z_1^p + z_2^p)$$

Therefore

$$|f(x) + g(x)|^p \leq \Big(|f(x)| + |g(x)| \Big)^p \leq 2^{p-1} \Big(|f(x)|^p + |g(x)|^p \Big)$$

[2]A proof of Hölder's inequality can be found in: F. Riesz and B. Sz.-Nagy, Functional analysis, Sec. 21

This implies that $|f+g|^p$ is integrable if $|f|^p$ and $|g|^p$ are integrable, and $\mathcal{L}_p(\Omega)$ is a linear space.

48.7. Minkowski inequality. Put

$$\|f\|_p = \left(\int_\Omega |f(x)|^p \, d^N x\right)^{\frac{1}{p}}$$

for any function $f \in \mathcal{L}_p(\Omega)$. Then

$$\|f+g\|_p \le \|f\|_p + \|g\|_p$$

which is known as *Minkowski's inequality.* It follows from Hölder's inequality which can be stated in the form

$$\Big|\int f(x)g(x)\, d^N x\Big| \le \|f\|_p \|g\|_q\,, \quad q = \frac{p}{p-1}\,.$$

Indeed, using Hölder's inequality for the integral of products $|f||f+g|^{p-1}$ and $|g||f+g|^{p-1}$ one infers that

$$\begin{aligned}
\|f+g\|_p^p &= \int_\Omega \Big|f(x)+g(x)\Big|^p d^N x \\
&\le \int_\Omega \Big(|f(x)| + |g(x)|\Big)\Big|f(x)+g(x)\Big|^{p-1} d^N x \\
&\le \Big(\|f\|_p + \|g\|_p\Big)\Big\||f+g|^{p-1}\Big\|_q \\
&= \Big(\|f\|_p + \|g\|_p\Big)\|f+g\|_p^{p-1}
\end{aligned}$$

Minkowski's inequality follows from the latter inequality.

48.8. Topology on $\mathcal{L}_p(\Omega)$. A sequence $\{f_n\}$ in $\mathcal{L}_p(\Omega)$ is said to converge to a function $f \in \mathcal{L}_p(\Omega)$ if

$$\lim_{n\to\infty} \|f - f_n\|_p = 0$$

and in this case one writes

$$f_n \to f \quad \text{in } \mathcal{L}_p(\Omega) \quad \text{as } n \to \infty$$

and says that the sequence $\{f_n\}$ converges to f in the $\mathcal{L}_p$ topology. Suppose that $f_n \to f$ and $f_n \to g$ in $\mathcal{L}_p(\Omega)$. Then it follows from Minkowski's inequality that

$$\|f-g\|_p \le \|f-f_n\|_p + \|g-f_n\|_p$$

for any n. This implies that

$$\|f-g\|_p = 0$$

which, in turn, means that the functions f and g are equal almost everywhere

$$f(x) = g(x) \quad a.e.$$

but not identical. So, the limit function f can be altered on a set of measure zero, and the resulting function function will also be a limit function of the same sequence. The situation is quite different from spaces $C^p(\bar{\Omega})$ where the topology guarantees the uniqueness of the limit.

Let us take a function f from $\mathcal{L}_p$ and consider the set of functions that are equal to f almost everywhere. Then the whole space $\mathcal{L}_p$ can viewed as a collection of such sets. If f and g belongs to different sets, that is, they are not equal almost everywhere, then $\|f - g\|_p \neq 0$ and $\|f - g\|_p = 0$ otherwise. So, one can define $\mathcal{L}_p$ as a space of *equivalence classes* so that a function $f(x)$ represents an element $f \in \mathcal{L}_p$ in the sense that two elements f and g from $\mathcal{L}_p$ are equal if their representatives are equal almost everywhere:

$$f = g \quad \text{in } \mathcal{L}_p \quad \text{if } f(x) = g(x) \ a.e.$$

The terms of any convergent sequence $\{f_n(x)\}$ can be changed by adding functions that are equal to zero almost everywhere, and the resulting sequence would have the *same* elements in $\mathcal{L}_p$ if the latter is defined as the space of equivalence classes. Note that this space remains linear because the sum of two functions that are zero almost everywhere is equal to zero almost everywhere. Thus, if $\mathcal{L}_p$ is defined as the space of equivalence classes with respect to adding functions that are zero almost everywhere, then the limit function of a convergent sequence in $\mathcal{L}_p$ is unique.

Metric or distance in $\mathcal{L}_p$. The space $\mathcal{L}_p(\Omega)$ defined as the space of equivalence classes can be converted into a metric space by defining the distance function

$$d(f, g) = \|f - g\|_p$$

It satisfies the distance axioms. First, $d(f, g) \geq 0$ and $d(f, g) = 0$ if and only if $f = g$ (which is true because $f(x) = g(x)$ $a.e.$). The distance function is symmetric, $d(f, g) = d(g, f)$, and satisfies the triangle inequality,

$$d(f, g) \leq d(f, h) + d(g, h)$$

which follows from Minkowski's inequality.

48.9. Space of square integrable functions $\mathcal{L}_2$. . In a real Euclidean space one can define a distance by $|x-y|$ and an angle between vectors via the dot product

$$(x,y)=\sum_j x_j y_j$$

so that

$$-1\le \frac{(x,y)}{|x||y|}\le 1$$

by the Cauchy-Schwartz inequality. Owing to this property, the cosine of the angle between x and y can be defined as the dot product of unit vectors $x/|x|$ and $y/|y|$ parallel x and y respectively. In addition, the dot product is a linear function in both x and y.

Among all spaces $\mathcal{L}_p$, only the space of square integrable functions $\mathcal{L}_2$ admits a dot or inner product with such properties:

$$\langle f,g\rangle=\int_\Omega f(x)g(x)\,d^N x$$

where f and g are real and belong to $\mathcal{L}_2(\Omega)$. By Hölder's inequality with $p=q=\frac{1}{2}$

$$|\langle f,g\rangle|\le \|f\|_2\|g\|_2$$

so that the angle between two square integrable functions can be defined as the root of the equation

$$\cos(\theta)=\frac{\langle f,g\rangle}{\|f\|_2\|g\|_2}\,,\quad 0\le\theta\le\pi\,,$$

by the analogy with a real Euclidean space. In particular, f and g are orthogonal if their inner product vanishes. For example, the functions $f_n(x)=\sin(\pi nx)$, where $n=1,2,...$, are orthogonal in $\mathcal{L}_2(-1,1)$:

$$\langle f_n,f_m\rangle=\int_{-1}^{1}\sin(\pi nx)\sin(\pi mx)\,dx=\delta_{nm}$$

Orthogonal sets in $\mathcal{L}_2$ play a fundamental role in the Fourier analysis and quantum physics. It will be shown later that there are so called *complete* orthogonal sets $\{\phi_n\}$ in $\mathcal{L}_2$ such any element f can be expanded into a Fourier series

$$f(x)\sim\sum_{n=1}^{\infty}f_n\phi_n(x)\,,\quad f_n=\frac{\langle f,\phi_n\rangle}{\|\phi_n\|_2^2}$$

that converges to f in the $\mathcal{L}_2$ topology (that is, almost everywhere):

$$\lim_{m\to\infty}\Big\|f-\sum_{n=1}^{m}f_n\phi_n\Big\|_2=0$$

The Fourier series resembles the expansion of a vector in a Euclidean space over an orthogonal basis. In this sense, the space of square integrable functions is an infinite dimensional generalization of a Euclidean space. The space $\mathcal{L}_2$ is a Hilbert space that is used in quantum physics.

48.10. Complete normed spaces. A sequence $\{u_n\}$ in a normed space X is said to *converge* to $u \in X$ if

$$\lim_{n\to\infty} \|u - u_n\| = 0$$

In this case, one also writes

$$u_n \to u \quad \text{in} \ \ X$$

A sequence $\{u_\}\subset X$ is called a *Cauchy sequence* in X for any $\varepsilon >$ one can find an integer N such that

$$\|u_n - u_m\| \le \varepsilon \qquad n \ge N\,,\ m \ge N$$

In other words, if the numbers $\|u_n - u_m\|$ can be made arbitrary small for large enough n and m and remains arbitrary small for all large enough n and m, then $\{u_n\}$ is a Cauchy sequence.

Euclidean spaces $\mathbb{R}^N$ (or $\mathbb{C}^N$) are normed spaces, and every Cauchy sequence has a limit in them. This is no longer true for a general normed space.

PROPOSITION **48.3**. *Every convergent sequence in a normed space is a Cauchy sequence. But the converse is false.*

Let $u_n \to u$ in X. Fix $\varepsilon > 0$. Then there is an integer N such that

$$\|u - u_n\| < \varepsilon\,, \quad \forall n > N$$

Then for any $n > N$ and $m > N$, by the triangle inequality

$$\|u_n - u_m\| = \|u_n - u + u - u_m\| \le \|u - u_n\| + \|u - u_m\| < 2\varepsilon$$

Since ε is arbitrary, this implies that $\{u_n\}$ is a Cauchy sequence. To show that the converse is false, it is sufficient to give an example.

Consider the normed space $X = C_2^0([-1,1]) \subset \mathcal{L}_2(-1,1)$. Put

$$u_n(x) = \frac{1}{2} + \frac{1}{\pi}\arctan(nx) \in C_2^0([-1,1])$$

This sequence of continuous functions has a pointwise limit

$$\lim u_n(x) = u(x) = \begin{cases} \frac{1}{2} & ,\ x = 0 \\ \frac{1}{2} + \frac{1}{2}\text{sign}(x) = \theta(x) & ,\ x \neq 0 \end{cases}$$

where $\theta(x)$ is the step function. Note that $u(x)$ is square integrable but not continuous at $x = 0$. So, the limit function is not from X. Let us show that $\{u_n\}$ is a Cauchy sequence in X. Note that

$$|u_n(x) - u(x)| \leq |u_n(x)| + |u(x)| \leq 2$$

and a constant function is integrable on $(-1, 1)$. Therefore by the Lebesgue dominated convergence theorem

$$\lim_{n\to\infty} \int_{-1}^{1} |u_n(x) - u(x)|^2 dx = \int_{-1}^{1} \lim_{n\to\infty} |u_n(x) - u(x)|^2 dx = 0$$

This means that $u_n \to u$ in $\mathcal{L}_2(-1, 1)$ and, hence, $\{u_n\}$ is a Cauchy sequence in $\mathcal{L}_2(-1, 1)$. Since all u_n are continuous functions, the sequence $\{u_n\}$ is also a Cauchy sequence in the subspace $C_2^0([-1, 1]) \subset \mathcal{L}_2(-1, 1)$. However, the sequence converges in $\mathcal{L}_2(-1, 1)$ but has no limit in $C_2^0([-1, 1])$.

Thus, a limit point of a Cauchy sequence, if it exists, can be in a larger space. A vague analogy can be made with limits of an open set in a Euclidean space. A limit point of a Cauchy sequence in an open set can lie in the boundary of the set.

Banach space. A normed space is called *complete* if every Cauchy sequence has a limit in it. A complete normed space is also called a *Banach space*.

For example, the space of continuous square integrable function, $C_2^0(\Omega)$, is not a Banach space.

PROPOSITION **48.4**. *The space of bounded functions with the supremum norm is a Banach space*

Let $\{u_n\} \subset \mathcal{B}(\Omega)$ be a Cauchy sequence. Fix $\varepsilon > 0$. Then for any $x \in \Omega$ one can find an integer N such that

$$|u_n(x) - u_m(x)| \leq \|u_n - u_m\|_\infty < \varepsilon$$

for all $n \geq N$ and $m \geq N$. This implies that the numerical sequence $\{u_n(x)\}$ is a Cauchy sequence for every $x \in \Omega$. Therefore there exists a function $u(x)$ such that

$$\lim_{n\to\infty} u_n(x) = u(x)$$

In other words, the pointwise limit exists for any Cauchy sequence in $\mathcal{B}(\Omega)$.

Let us show that $u \in \mathcal{B}(\Omega)$, that is, $u(x)$ is bounded. Note that any Cauchy sequence in a normed space is bounded. Indeed, put $m = N$ so that

$$\|u_N - u_n\| < \varepsilon\,, \quad \forall n \geq N$$

Put

$$A = \max_{k=1,2,\ldots,N} \{\|u_N - u_k\|, \varepsilon\}$$

Then

$$\|u_n\| = \|u_n - u_N + u_N\| \le A + \|u_N\| \equiv M < \infty$$

for all n. Therefore for any $x \in \Omega$ and all n

$$|u_n(x)| \le \|u_n\|_\infty \le M$$

By taking the limit

$$|u(x)| = \lim_{n\to\infty} |u_n(x)| \le M$$

for all x in Ω. Hence, $u(x)$ is bounded, and $u \in \mathcal{B}(\Omega)$.

Uniform convergence. The convergence with respect to the supremum norm is also called a *uniform convergence*. A sequence of functions $u_n(x)$ is said to converge *uniformly* to a function $u(x)$ if

$$\lim_{n\to\infty} \sup |u_n(x) - u(x)| = 0$$

THEOREM 48.1. *The limit of a uniformly convergent sequence of continuous functions is a continuous function.*

Let $u(x)$ be the pointwise limit of a sequence of bounded continuous functions $u_n(x)$,

$$\lim_{n\to\infty} u_n(x) = u(x)$$

Note that $u(x)$ exists by the above analysis and

$$\lim_{n\to\infty} \|u - u_n\|_\infty = 0$$

Let us show that for any $x \in \Omega$

$$\lim_{y\to x} u(y) = u(x)$$

Fix $\varepsilon > 0$. For any $x \in \Omega$ one can find an integer n (independent of x) such that

$$|u(x) - u_n(x)| \le \|u - u_n\|_\infty < \varepsilon$$

because the sequence converges uniformly. For any x and y, one has

$$\begin{aligned} |u(x) - u(y)| &\le |u(x) - u_n(x)| + |u_n(x) - u_n(y)| + |u(y) - u_n(y)| \\ &\le 2\varepsilon + |u_n(x) - u_n(y)| \end{aligned}$$

The function $u_n(x)$ is continuous at any point in Ω. Therefore there exists $\delta > 0$ such that

$$|u_n(x) - u_n(y)| < \varepsilon\,, \quad \text{whenever} \;\; |x - y| < \delta$$

This implies that

$$|u(x) - u(y)| < 3\varepsilon\,, \quad \text{whenever} \quad |x - y| < \delta$$

or $u(x)$ is continuous at any point in Ω.

The theorem about uniform convergence and continuity implies that the subspace of continuous functions in the space of bounded functions, $C^0(\Omega) \cap \mathcal{B}(\Omega)$ is complete.

48.11. Series in the space of bounded continuous functions.

THEOREM **48.2**. *Let $\{u_n\} \subset C^0(\Omega)$ be a sequence of continuous functions on $\Omega \subset \mathbb{R}^N$ that are bounded, and the series of bounds converges:*

$$|u_n(x)| \le M_n\,, \qquad \sum_n M_n < \infty$$

Then the series $\sum u_n(x)$ converges to a continuous function on Ω:

$$u(x) = \sum_n u_n(x) \ \in \ C^0(\Omega)$$

Let

$$v_n(x) = \sum_{k=1}^{n} u_k(x) \ \in \ C^0(\Omega)$$

be a sequence of partial sums of the functional series, and

$$s_n = \sum_{k=1}^{n} M_k$$

be a sequence of partial sums of the series of the bounds. Since $\{s_n\}$ converges by the hypothesis, it is a Cauchy sequence. For every $x \in \Omega$, $\{v_n(x)\}$ is also a Cauchy sequence. Indeed, let $n > m$ so that

$$\begin{aligned} |v_n(x) - v_m(x)| &\le |u_n(x)| + |u_{n-1}(x) + \cdots + |u_{m+1}(x)| \\ &\le M_n + M_{n-1} + \cdots + M_{m+1} = |s_n - s_m| \end{aligned}$$

This shows that $|v_n(x) - v_m(x)|$ can be made arbitrary small for all sufficiently large n and m because $\{s_n\}$ is a Cauchy sequence. Therefore, the functional series converges for any x. Put

$$u(x) = \sum_n u_n(x)\,, \qquad \lim_{n\to\infty} v_n(x) = u(x)\,, \quad x \in \Omega$$

It follows that for any $x \in \Omega$

$$|u(x) - v_n(x)| \le \sum_{k=n+1}^{\infty} M_n \equiv R_n$$

and, hence, by taking the supremum

$$\sup_{\Omega} |u(x) - v_n(x)| = \|u - v_n\|_\infty \le R_n$$

Since the series of the bounds converges, $R_n \to 0$ as $n \to \infty$, and therefore v_n converges to u uniformly on Ω. By the completeness of the space $C^0(\Omega) \cap \mathcal{B}(\Omega)$, the limit function is continuous because the terms of the sequence are continuous.

48.12. Other examples of Banach spaces. Consider a linear space $C^1[a,b]$ of all continuously differentiable functions on an interval $[a,b]$. Define the norm on it as the sum of supremum norms of the function and its derivative:

$$\|u\|_{C^1} = \|u\|_\infty + \|u'\|_\infty$$

This is a Banach space. Let $\{u_n\}$ be a Cauchy sequence in $C^1[a,b]$. Then it is also a Cauchy sequence with respect to the supremum norm because

$$\|u\|_\infty \le \|u\|_{C^1}$$

Therefore it converges to some continuous function $u \in C^0[a,b]$. The sequence of derivatives $\{u_n'\}$ is also a Cauchy sequence with respect to the supremum norm (replace u by u' in the above inequality). Therefore it converges to a continuous function $v \in C^0[a,b]$. By the fundamental theorem of Calculus

$$\int_a^x u_n'(y)\, dy = u_n(x) - u_n(a)\,, \quad a \le x \le b\,.$$

Since every convergent sequence is bounded, $|u_n'(y)| \le \|u_n'\|_\infty \le M$ for all n, and a constant function is integrable on $[a,b]$, it is concluded by the Lebesgue dominated convergence theorem that

$$\lim_{n\to\infty} \int_a^x u_n'(y)\, dy = \int_a^x \lim_{n\to\infty} u_n'(y)\, dy = \int_a^x v(y)\, dy = u(x) - u(a)$$

Since v is continuous, by the fundamental theorem of Calculus, v is continuously differentiable on $[a,b]$ and

$$v(x) = u'(x)$$

The space $C^p[a,b]$. Put

$$\|u\|_{C^p} = \|u\|_\infty + \|u'\|_\infty + \cdots + \|u^{(p)}\|_\infty$$

It is not difficult to verify that $C^p[a,b]$ equipped with this norm is a normed space. It is also a Banach space. A proof follows the same line of reasonings as in the case of $p = 1$ applied inductively.

The space $C^p(\Omega)$. Let Ω be closed and bounded in $\mathbb{R}^N$. Put

$$\|u\|_{C^p} = \|u\|_\infty + \|Du\|_\infty + \cdots + \|D^p u\|_\infty$$

where

$$\|D^\alpha\|_\infty = \max_{j_1, j_2, \ldots, j_\alpha} \sup \left| \frac{\partial^\alpha u}{\partial x_{j_1} \cdots \partial x_{j_\alpha}} \right|$$

Then $C^p(\Omega)$ is a Banach space. If $\{u_n\}$ is a Cauchy sequence, then it is also a Cauchy sequence with respect to the supremum norm and, hence, it converges uniformly to a continuous function $u \in C^0(\Omega)$. The sequence of the gradients $\{\nabla u_n\}$ also converges uniformly on Ω to some continuous vector field $F \in C^0(\Omega)$. Let C be a smooth curve in Ω going from a point A to a point B in Ω. Then by the fundamental theorem for line integrals

$$\int_C (\nabla u_n(x), dx) = u_n(B) - u_n(A)$$

The sequence of the gradients is bounded as it converges so that $|\nabla u_n| \leq N\|Du_n\|_\infty \leq M$ and a constant function is integrable on any smooth curve (of a finite length). Therefore by the Lebesgue dominated convergence theorem

$$\begin{aligned} \lim_{n\to\infty} \int_C (\nabla u_n(x), dx) &= \int_C \lim_{n\to\infty} (\nabla u_n(x), dx) \\ &= \int_C (F(x), dx) = u(B) - u(A) \end{aligned}$$

This shows that the line integral of the vector field F is independent of C. Therefore F is conservative and its potential is $u(x)$, that is, $F = \nabla u$. The argument can be repeated for every component of the gradient ∇u_n to show that $D^2 u_n$ converges to $D^2 u$, etc.

Spaces $\mathcal{L}_p(\Omega)$ and l_p. The space l_p consists of all complex sequences $u = \{a_n\}_1^\infty$ with summable p^{th} powers of terms:

$$\sum_{n=1}^{\infty} |a_n|^p < \infty\,, \quad p \geq 1$$

The norm is defined by

$$\|u\|_{l_p} = \Big(\sum_{n=1}^{\infty} |a_n|^p\Big)^{\frac{1}{p}}$$

Similarly, the space $\mathcal{L}_p(\Omega)$ consists of all complex valued function whose p^{th} power is integrable

$$\int_\Omega |u(x)|^p \, d^n x < \infty \,, \quad p \geq 1 \,.$$

and the $\mathcal{L}_p-$norm is defined by

$$\|u\|_p = \Big(\int_\Omega |u(x)|^p \, d^n x \Big)^{\frac{1}{p}} \,.$$

One can show that these space are Banach spaces. The case $p = 2$ will be discussed later in detail (in the framework of Hilbert spaces).

49. Contraction principle

49.1. Transformations or operators. Let S be a subset of a normed space X. A *transformation* T is a mapping of S into X, that is, it is a rule that assigns a unique element $Tu \in X$ to every element $u \in S$. One can also define transformations (or operators) that map one normed space X to another normed space Y.

Continuous transformations. A transformation T is said to be *continuous* at $u \in S$ if for any sequence $\{u_n\}$ that converges to u in S, the image sequence $\{Tu_n\}$ converges to Tu in X:

$$\lim_{n\to\infty} \|u_n - u\| = 0 \quad \Rightarrow \quad \lim_{n\to\infty} \|Tu_n - Tu\| = 0$$

If the transformation is continuous at every element in S, then the transformation is called *continuous on* S.

Lipschitz continuous transformations. A transformation T is called *Lipschitz continuous on* S if there exists a number ρ such that

$$\|Tu - Tv\| \leq \rho\|u - v\|\,, \qquad u, v \in S$$

and ρ is independent of u and v. Clearly, a Lipschitz continuous transformation is continuous on S, but the converse is not true. So, the Lipschitz continuity is a stronger condition on transformations than the continuity.

Contractions. If a transformation T is Lipschitz continuous on S and, in addition, $\rho < 1$, then T is called a *contraction*. Note that that the norm defines a natural distance $d(u,v) = \|u - v\|$. A contraction reduces the distance between elements:

$$d(Tu, Tv) \leq \rho d(u, v) < d(u, v)$$

An element u is called a *fixed point* of a transformation T if

$$Tu = u$$

Let $S = X = \mathbb{R}$ and $\|u\| = |u|$. Put

$$T_1 u = u\,, \quad T_2 u = \frac{1}{2}\,u\,, \quad T_3 u = |u|\,, \quad T_4 u = \sqrt{|u|}$$

Then all of these transformations are continuous. The transformations $T_{1,2,3}$ are also Lipschitz continuous, while T_4 is not Lipschitz continuous. Only T_2 is a contraction.

Closed sets in a Banach space. A set in a normed space is said to be *closed* if every Cauchy sequence in it converges to an element in the set. For example, the set of continuous bounded functions in the space of all bounded functions is closed in this space. However, a closed set in a Banach space is not generally a linear subspace because a closed set may not be closed under the addition of its elements and multiplication of them by a number.

Let $S_{m,M}$ denote a set of all continuous functions on a region Ω in a Euclidean space whose values lie between m and M:

$$m \le u(x) \le M\,, \quad x \in \Omega$$

This is a subset in the Banach space of bounded functions, $\mathcal{B}(\Omega)$, but it is not a linear subspace. For example, if u belongs to $S_{m,M}$, then $v(x) = 2u(x)$ does not belong to it. However, $S_{m,M}$ is closed in $\mathcal{B}(\Omega)$.

Let $\{u_n\}$ be a Cauchy sequence in $S_{m,M}$. Then by completeness of $\mathcal{B}(\Omega)$ there exists a bounded function $u(x)$ to which the sequence converges uniformly

$$\lim_{n\to\infty} \|u - u_n\|_\infty = 0$$

Since u_n are continuous, the limit function u is also continuous by the uniform convergence of the sequence. Let us show that $u(x) \le M$. Suppose that there exists $y \in \Omega$ such that $u(y) > M$. Then it follows from the inequality

$$u_n(y) \le M < u(y)$$

that for all n

$$\begin{aligned} 0 < u(y) - M &= u(y) - u_n(y) + u_n(y) - M \le u(y) - u_n(y) \\ &\le \|u - u_n\|_\infty \end{aligned}$$

This inequality implies that $\|u - u_n\|$ is *strictly* positive for all n and, hence, there is a positive number $\alpha = u(y) - M > 0$ such that

$$0 < \alpha \le \|u - u_n\|_\infty$$

for all n. But the right side of this inequality can be made arbitrary small (e.g., smaller than α). This contradiction implies that no such y exists and therefore $u(x) \le M$.

Let us show that $m \le u(x)$. Suppose that $u(y) < m$ for some y. Then

$$u(y) < m \le u_n(y)$$

and, hence, for all n

$$0 < m - u(y) \le u_n(y) - u(y) \le \|u_n - u\|_\infty$$

As in the previous case, this leads to a contradiction because $\|u - u_n\|$ can be made smaller than $m - u(y) > 0$ and, hence, $u(x) \geq m$ for any x. Thus, $S_{m,M}$ is a closed set in $\mathcal{B}(\Omega)$.

49.2. Contraction mapping theorem.

THEOREM **49.1**. *Let S be a closed set in a Banach space $\mathcal{B}$. Let a transformation $T: S \to S$ be a contraction. Then T has a unique fixed point in S, for any $u_0 \in S$, the sequence $u_n = T^n u_0$ converges to the fixed point.*

Uniqueness of the fixed point. Let us show first that if a fixed point exists, then it is unique. Let u and v be fixed points of T, that is,

$$Tu = u\,, \quad Tv = v$$

By the hypotheses, there exists $\rho < 1$ such that

$$\|u - v\| = \|Tu - Tv\| \leq \rho \|u - v\|$$

which is impossible unless $\|u - v\| = 0$ or $u = v$. Thus, the fixed is unique if it exists.

Existence of a fixed point. The idea is to show that $u_n = T^n u_0$ is a Cauchy sequence. Then by completeness of the set S, any Cauchy sequence has a limit in S. Fix $u_0 \in S$. Then

$$\begin{aligned}\|u_{n+1} - u_n\| &= \|Tu_n - Tu_{n-1}\| \leq \rho \|u_n - u_{n-1}\| \leq \rho^2 \|u_{n-1} - u_{n-2}\| \\ &\leq \rho^n \|u_1 - u_0\|\end{aligned}$$

For any integer $n > m$, the following chain inequalities holds

$$\begin{aligned}\|u_n - u_m\| &= \|(u_n - u_{n-1}) + (u_{n-1} - u_{n-2}) + \cdots + (u_{m+1} - u_m)\| \\ &\leq \|u_n - u_{n-1}\| + \|u_{n-1} - u_{n-2}\| + \cdots + \|u_{m+1} - u_m)\| \\ &\leq (\rho^{n-1} + \rho^{n-2} + \cdots + \rho^m)\|u_1 - u_0\| \\ &= \rho^m (1 + \rho + \cdots + \rho^{n-1-m})\|u_1 - u_0\| \\ &\leq \rho^m \Big(\sum_{k=0}^{\infty} \rho^k\Big)\|u_1 - u_0\| \\ &\leq \frac{\rho^m}{1-\rho}\|u_1 - u_0\|\end{aligned}$$

Since $\rho < 1$, $\rho^m \to 0$ as $m \to \infty$, and the distance $\|u_n - u_m\|$ can be made arbitrary small for all n and m that are large enough. This means that $u_n = T^n u_0$ is a Cauchy sequence for any choice of u_0. Therefore by completeness of the Banach space $\mathcal{B}$, there exists $u \in \mathcal{B}$ such that $T^n u_0 \to u$ in $\mathcal{B}$. But S is closed, and, hence, $u \in S$.

Fixed point as the limit point. Since T is a contraction on S, it is also continuous on S. By continuity of T,

$$u_n \to u \text{ in } S \quad \Rightarrow \quad Tu_n \to Tu \text{ in } S$$

Taking the limit $n \to \infty$ in the recurrence relation $u_{n+1} = Tu_n$, it is concluded that the limit point is also the fixed point of T:

$$\left.\begin{array}{r} u_{n+1} \to u \\ Tu_n \to Tu \end{array}\right\} \quad \Rightarrow \quad u = Tu$$

49.3. An extension of the contraction mapping theorem. It turns out that the conditions under which a transformation has a fixed point can be relaxed. It is sufficient that some power of the transformation is a contraction. This extension of the contraction mapping theorem is useful in applications.

THEOREM **49.2**. *Let S be closed in a Banach space $\mathcal{B}$. Let $T : S \to S \subset \mathcal{B}$ be a transformation such that its power $A = T^n$ is a contraction for some positive integer n. Then T has a unique fixed point $u \in S$, and for any $u_0 \in S$, $A^m u_0 \to u$ as $m \to \infty$.*

First, note that A has a unique fixed point in S, and for any $u_0 \in S$,

$$A^m u_0 \to u\,, \quad Au = u$$

Having found the fixed point u of A, put $u_0 = Tu$. Then

$$A^m u_0 = A^m T u_0 = T A^m u_0 = Tu$$

for all integers m. Therefore by taking the limit $m \to \infty$ in this relation, it is concluded that u is also a limit point of T:

$$Tu = \lim_{m\to\infty} A^m u_0 = u$$

Conversely, if u is a fixed point of T, then it is also a fixed point of A because

$$Tu = u \quad \Rightarrow \quad Au = T^m u = u$$

Thus, T and A have the same fixed points. But A has just one fixed point and, hence, T must have a unique fixed point.

Suppose that $n = 2$. Consider the sequence $T^k u_0$, $k = 0, 1, ...,$ for some u_0. Then the subsequence $T^{2m} u_0 = A^m u_0$ converges to the fixed point u. The other subsequence $T^{2m+1} u_0 = A^m T u_0$ also converges to the fixed point u because $A^m u_0$ converges to u for any choice of u_0. The argument can extended to any n (the sequence $T^k u_0$ is the union of sequences $A^m v_p$ where $v_p = T^p u_0$, $p = 0, 1, ..., n-1$). So, if a power of T is a contraction, then its fixed point can still be found as the limit of the sequence $T^k u_0$.

50. Non-linear Volterra integral equation

50.1. Volterra integral operator. Consider a space $C^0([a,b])$ of all continuous functions on an interval $[a,b]$. It is a Banach space with respect to the supremum norm because continuous functions attain its extreme values on closed and bounded sets in Euclidean spaces. Define a transformation or an operator on this space by

$$\hat{K}u(x) = \int_a^x K(x,y,u(y))\,dy$$

The kernel $K(x,y,z)$ is required to be Lipschitz continuous in the variable z

$$|K(x,y,z_1) - K(x,y,z_2)| \leq M(x,y)|z_1 - z_2|$$

where the function $M(x,y)$ is continuous in the rectangle $[a,b]\times[a,b]$. So, for any $z \in \mathbb{R}$, $K(x,y,z)$ is continuous in $[a,b]\times[a,b]$. The operator $\hat{K}$ is called a *non-linear Volterra integral operator.* If $K(x,y,z) = M(x,y)z$, then the Volterra operator is linear.

The integral equation

$$u = \hat{K}u + f$$

where f is a continuous function on $[a,b]$, is called the *Volterra equation.*

50.2. Properties of the Volterra operator. The Volterra operator maps a continuous function into a continuous function

$$\hat{K}\ :\quad C^0([a,b]) \to C^0([a,b])$$

and

$$|\hat{K}u(x) - \hat{K}v(x)| \leq m_0(x-a)\|u-v\|_\infty\,, \quad m_0 = \max M(x,y)\,,$$

The Volterra operator is Lipschitz continuous:

$$\|\hat{K}u - \hat{K}v\|_\infty \leq \rho\|u-v\|_\infty\,,$$

where

$$\rho = \max_{[a,b]} \int_a^x M(x,y)\,dy \leq m_0(b-a)\,.$$

The continuity of $\hat{K}u(x)$ follows from continuity of the Lebesgue integral

$$\lim_{x\to x_0} \hat{K}u(x) = \lim_{x\to x_0} \int_a^x K(x,y,u(y))\,dy = \lim_{x\to x_0} \int_a^{x_0} K(x,y,u(y))\,dy$$

and continuity of the kernel. Since $K(x, y, u(y))$ is continuous on $[a, b] \times [a, b]$ for any continuous $u(y)$, it is bounded by a constant

$$|K(x, y, u(y)| \le K_0 = \max_{[a,b]\times[a,b]} |K(x, y, u(y))|$$

and a constant function K_0 (independent of x) is integrable on any bounded interval. By the Lebesgue dominated convergence theorem, the limit can be moved into the integral

$$\lim_{x\to x_0} \int_a^{x_0} K(x, y, u(y))\, dy = \int_a^{x_0} K(x_0, y, u(y))\, dy = \hat{K}u(x_0)$$

This $\hat{K}u(x)$ is continuous on $[a, b]$.

Next, by the Lipschitz continuity of the kernel

$$\begin{aligned} |\hat{K}u(x) - \hat{K}v(x)| &\le \int_a^x M(x, y)|u(y) - v(y)|\, dy \\ &\le \|u - v\|_\infty \int_a^x M(x, y)\, dy \\ &\le \|u - v\|_\infty m_0 \int_a^x dy \\ &= m_0(x - a)\|u - v\|_\infty \end{aligned}$$

The right-hand side of the second inequality reaches its maximal value on a closed and bounded interval $[a, b]$ because it is a continuous function. Therefore for all $x \in [a, b]$:

$$|\hat{K}u(x) - \hat{K}v(x)| \le \rho\|u - v\|_\infty$$

where

$$\rho = \max_{[a,b]} \int_a^x M(x, y)\, dy \le m_0 \max_{[a,b]} \int_a^x dy = m_0(b - a)$$

Taking the maximum of the left-hand side (which exists by continuity of $\hat{K}u(x)$ for any continuous $u(x)$), it is concluded that the Volterra operator is Lipschitz continuous:

$$\|\hat{K}u - \hat{K}v\|_\infty \le \rho\|u - v\|_\infty$$

50.3. Solving the Volterra equation. Consider the transformation on the space of continuous functions $C^0([a, b])$ defined by the rule

$$Tu(x) = \hat{K}u(x) + f(x)\,, \quad f \in C^0([a, b])$$

By properties of the Volterra operator, the transformation maps the Banach space $C^0([a, b])$ into itself and is Lipschitz continuous

$$\|Tu - Tv\|_\infty = \|\hat{K}u - \hat{K}v\|_\infty \le \rho\|u - v\|_\infty$$

Then a solution to the Volterra equation is a fixed point of the transformation T. If T or its power $A = T^p$ is a contraction, then by the contraction mapping theorem, T has a unique fixed point that is the limit of the sequence $A^n u_0$ for some continuous u_0.

Let us show that $A = T^p$ is a contraction for some large enough p. By properties of the Volterra operator

$$|Tu(x) - Tv(x)| = |\hat{K}u(x) - \hat{K}v(x)| \le m_0(x-a)\|u-v\|_\infty$$

It follows from this inequality and the Lipschitz continuity of the kernel $K(x, y, z)$ in the variable z that

$$\begin{aligned}
|T^2u(x) - T^2v(x)| &= |T(Tu)(x) - T(Tv)(x)| \\
&= |\hat{K}(Tu)(x) - \hat{K}(Tv)(x)| \\
&\le \int_a^x |K(x,y,Tu(y)) - K(x,y,Tv(y)|\,dy \\
&\le m_0 \int_a^x |Tu(y) - Tv(y)|\,dy \\
&\le m_0\|u-v\|_\infty \int_a^x (y-a)\,dy \\
&\le \frac{m_0(x-a)^2}{2}\,\|u-v\|_\infty
\end{aligned}$$

Using the process, it is not difficult to show by mathematical induction that

$$|T^n u(x) - T^n v(x)| \le \frac{m_0(x-a)^n}{n!}\,\|u-v\|_\infty$$

Indeed, if the above inequality is true, then it is true for $n+1$:

$$\begin{aligned}
|T^{n+1}u(x) - T^{n+1}v(x)| &\le \int_a^x |K(x,y,T^n u(y)) - K(x,y,T^n v(y)|\,dy \\
&\le m_0 \int_a^x |T^n u(y) - T^n v(y)|\,dy \\
&\le m_0\|u-v\|_\infty \int_a^x \frac{(y-a)^n}{n!}\,dy \\
&\le \frac{m_0(x-a)^{n+1}}{(n+1)!}\,\|u-v\|_\infty
\end{aligned}$$

Therefore for all $x \in [a, b]$

$$|T^n u(x) - T^n v(x)| \le \frac{m_0(b-a)^n}{n!}\,\|u-v\|_\infty$$

Taking the maximum in the right-hand side, one infers that

$$\|T^n u - T^n v\| \leq \rho_n \|u - v\|_\infty\,, \quad \rho_n = \frac{m_0(b-a)^n}{n!}$$

The factorial grows faster than exponential so that

$$\lim_{n\to\infty} \rho_n = 0$$

Therefore there is an integer p such that

$$\rho_p = \frac{m_0(b-a)^p}{p!} < 1$$

and the transformation $A = T^p$ is a contraction.

Solving the Volterra equation. By the extended contraction principle the sequence of continuous functions

$$u_n(x) = T^n f(x)$$

converges to a function $u(x)$

$$\lim_{n\to\infty} u_n(x) = u(x)\,, \quad x \in [a,b]\,,$$

the convergence is uniform on $[a,b]$:

$$\lim_{n\to\infty} \|u_n - u\|_\infty = 0$$

and the limit function is the unique solution to the Volterra equation.

50.4. Applications to Cauchy problems. Consider the integral equation equivalent to the Cauchy problem for the wave equation discussed in Section **45.3**

$$u(x,t) = v(x,t) + \frac{\lambda}{4\pi c^2} \int\limits_{|y-x|<ct} \frac{F(u(y, t - \frac{|x-y|}{c}), y, t - \frac{|x-y|}{c})}{|x-y|}\, d^3y\,,$$

where $v(x,t)$ is continuous and has a bounded support (this is possible if the initial data have bounded supports by the Huygens principle). Suppose that the function $F(z,x,t)$ is Lipschitz continuous in the variable z:

$$|F(z_1,x,t) - F(z_2,x,t)| \leq \nu(x,t)|z_1 - z_2|$$

where $\nu(x,t)$ is a continuous and bounded function

$$\nu(x,t) \leq \nu_0\,, \quad x \in \mathbb{R}^3\,, \quad t \geq 0\,.$$

51. Non-linear boundary value problems

51.1. Fredholm integral equation. Consider an operator $\hat{K}$ acting on the space of continuous functions, $C^0(\bar{\Omega})$ where Ω is open and bounded in $\mathbb{R}^N$, by the rule

$$\hat{K}u(x) = \int_\Omega K(x, y, u(y))\, d^N y$$

In addition, the kernel $K(x, y, z)$ satisfies the same conditions as the Volterra kernel. It is Lipschitz continuous in the variable $z \in \mathbb{R}$:

$$|K(x, y, z_1) - K(x, y, z_2)| \le M(x, y)|z_1 - z_2|\,, \quad M \in C^0(\bar{\Omega} \times \bar{\Omega})$$

and for every z, $K(x, y, z)$ is continuous in $\bar{\Omega} \times \bar{\Omega}$. Then

$$\hat{K} : \; C^0(\bar{\Omega}) \to C^0(\bar{\Omega})$$

It follows from continuity of $K(x, y, z)$ in the variable x. The equation

$$u = \hat{K}u + f$$

where f is a continuous function on $\bar{\Omega}$, is called a *Fredholm integral equation*.

The transformation

$$Tu(x) = \hat{K}u(x) + f(x)$$

is Lipschitz continuous in $C^0(\bar{\Omega})$. Indeed, for any $x \in \bar{\Omega}$, one infers that

$$\begin{aligned}
|Tu(x) - Tv(x)| &\le \int_\Omega M(x, y)|u(y) - v(y)|\, d^N y \\
&\le \int_\Omega M(x, y)\, d^N y\, \|u - v\|_\infty \\
&\le \rho \|u - v\|_\infty \\
\rho &= \sup_\Omega \int_\Omega M(x, y)\, d^N y
\end{aligned}$$

By taking the supremum in the left-hand side,

$$\|Tu - Tv\|_\infty \le \rho \|u - v\|_\infty$$

If $\rho < 1$, then T is a contraction, and the Fredholm equation has a unique solution that is the fixed point of T. It turns out that Fredholm equations can be used to analyze and solve boundary value problems for non-linear (partial) differential equations as is shown below.

51.2. Sturm-Liouville operator. The Sturm-Liouville operator in an interval (a,b) is defined by the following rule

$$Lu(x) = -\Big(p(x)u'(x)\Big)' + q(x)u(x)\,, \quad a < x < b\,,$$

where the function $p(x)$ and $q(x)$ have the properties

$$p \in C^1[a,b]\,, \quad q \in C^0[a,b]\,, \quad p(x) > 0\,, \quad q(x) \ge 0\,.$$

It is further assumed that the functions on which the operator acts have the derivative that is continuously extendable to the endpoint of the interval:

$$u \in C^2(a,b) \cap C^1([a,b])$$

Define the boundary operators B_a and B_b by the rule

$$B_a(u) = \alpha_a u(a) - \beta_a u'(a) = 0\,, \quad B_b(u) = \alpha_b u(b) + \beta_b u'(b) = 0$$

where the numerical parameters satisfy the conditions

$$\alpha_j \ge 0\,, \quad \beta_j \ge 0\,, \quad \alpha_j + \beta_j > 0\,, \quad j = a, b$$

This condition states that α_j and β_j, $j = a, b$, are non-negative but cannot be zero simultaneously.

Boundary value problem for the Sturm-Liouville equation. Consider the following boundary value problem

$$\begin{aligned} &Lu(x) = g(x)\,, \quad x \in (a,b)\,, \quad f \in C^0([a,b]) \\ &B_a(u) = \gamma_a\,, \quad B_b(u) = \gamma_b \end{aligned}$$

This problem can be solved by the Green's function method.

Properties of the Sturm-Liouville operator. A homogeneous Sturm-Liouville equation, $Lu = 0$, has two linearly independent solutions as any second order differential equation (note that $p(x)$ does not vanish anywhere in $[a,b]$. If u_1 and u_2 are two linearly independent solutions, then their Wronskian does not vanish in $[a,b]$:

$$W(x) = u_1(x)u_2'(x) - u_1'(x)u_2(x) \ne 0\,, \quad x \in [a,b]$$

THEOREM 51.1. (Liouville-Ostrogradsky theorem)
Let L be a Sturm-Liouville operator in (a,b). Then for any two functions annihilated by L, the Wronskian does not vanish anywhere if it is not equal to zero at an endpoint of the interval, and

$$p(x)W(x) = p(a)W(a)$$

Let $Lu_1 = 0$ and $Lu_2 = 0$. Then

$$W'(x) = u_1(x)u_2''(x) - u_2(x)u_1''(x) = -\frac{p'(x)W(x)}{p(x)}$$

where the second derivative was expressed in terms of the derivatives and functions using the equation $Lu_{1,2} = 0$. It follows from the obtained relation that

$$p'(x)W(x) + p(x)W'(x) = 0 \quad \Rightarrow \quad p(x)W(x) = \text{const}$$

or

$$p(x)W(x) = p(a)W(a)$$

and $W(x) \neq 0$ if and only if $W(a) \neq 0$ because $p(x) > 0$.

Proposition **51.1**. *The homogeneous boundary value problem*

$$Lu = 0\,, \quad B_a(u) = 0\,, \quad B_b(u) = 0$$

has a non-trivial solution if and only if $q(x) = 0$ *and* $\alpha_a = \alpha_b = 0$*, and in this case,* $u(x)$ *is a constant.*

Let $u(x)$ be a non-trivial solution to the stated boundary value problem. Then by integration by parts

$$\begin{aligned}
0 &= \int_a^b u(x)Lu(x)\,dx \\
&= u(a)p(a)u'(a) - u(b)p(b)u'(b) \\
&\quad + \int_a^b p(x)\Big(u'(x)\Big)^2 dx + \int_a^b q(x)u^2(x)\,dx
\end{aligned}$$

The boundary term is non-negative. Indeed, using the boundary conditions

$$u(a)p(a)u'(a) = \begin{cases} \frac{\beta_a}{\alpha_a}p(a)(u'(a))^2 & , \ \alpha_a \neq 0 \\ 0 & , \ \alpha_a = 0 \end{cases} \ \geq 0$$

and similarly

$$-u(b)p(b)u'(b) = \begin{cases} \frac{\beta_b}{\alpha_b}p(b)(u'(b))^2 & , \ \alpha_b \neq 0 \\ 0 & , \ \alpha_b = 0 \end{cases} \ \geq 0$$

Thus, the boundary term is non-negative and vanishes only if $\alpha_a = \alpha_b = 0$ for a non-zero $u(x)$. The integral terms are also non-negative and must vanish independently

$$\int_a^b q(x)u^2(x)\,dx = 0 \quad \Rightarrow \quad q(x) = 0$$

if $u(x)$ is continuous and $u(x) \neq 0$. Similarly,

$$\int_a^b p(x)\Big(u'(x)\Big)^2\,dx = 0 \quad \Rightarrow \quad u'(x) = 0$$

because $p(x) > 0$. Thus, $u(x)$ must be a constant function.

Conversely, let $\alpha_a = \alpha_b = 0$ and $q = 0$. Then the boundary value problem

$$Lu = 0\,, \quad u'(a) = u'(b) = 0$$

has only a constant solution. Indeed, integrating the equation with $q = 0$, one infers that

$$p(x)u'(x) = C$$

for some constant C. Since $p(a) \neq 0$ and $u'(a) = 0$, it is concluded that $C = 0$ and, hence, $u'(x) = 0$ and only a constant solution is possible.

Proposition **51.2**. *The initial value problem*

$$Lu = 0\,, \quad u(a) = u'(a) = 0$$

has only the trivial solution $u(x) = 0$

If $u_1(x)$ and $u_2(x)$ are linearly independent solutions to $Lu = 0$, then a general solution is a linear combination $u = c_1u_1 + c_2u_2$. The initial conditions yield a linear system for the constants $c_{1,2}$:

$$c_1u_1(a) + c_2u_2(a) = 0\,, \quad c_1u_1'(a) + c_2u_2'(a) = 0$$

A homogeneous linear system has a non-trivial solution if and only if the determinant of its matrix is equal to zero, but this determinant is the Wronskian

$$W(a) = u_1(a)u_2'(a) - u_1'(a)u_2(a) \neq 0$$

that is not zero as u_1 and u_2 are linearly independent by the Liouville-Ostrogradsky theorem. Thus $c_1 = c_2 = 0$ and $u(x) = 0$.

Proposition **51.3**. *Let u_a and u_b be solutions to $Lu = 0$ such that u_a satisfies the left boundary condition, $B_a(u_a) = 0$ and u_b satisfies the right boundary condition $B_b(u_b) = 0$. Then u_a and u_b are linearly independent solutions to $Lu = 0$.*

By the Liouville-Ostrogradsky theorem, u_a and u_b are linearly dependent if and only if their Wronskian vanishes at $x = a$. The conditions $B_a(u_a) = 0$ and $W(a) = 0$ give a linear system for $u_a(a)$ and $u_a'(a)$:

$$\begin{aligned} B_a(u_a) &= \alpha_a u_a(a) - \beta_a u_a'(a) = 0 \\ W(a) &= u_a(a)u_b'(a) - u_a'(a)u_b(a) = 0 \end{aligned}$$

The system must have a non-trivial solution because otherwise $u_a(x) = 0$ because $u_a(a) = u_a'(a) = 0$. Therefore the determinant of the matrix of this system should vanish giving a conditions on $u_b(a)$ and $u_b'(a)$:

$$B_a(u_b) = \alpha_a u_b(a) - \beta_a u_b'(a) = 0$$

But the boundary value problem with $B_a(u_b) = B_b(u_b) = 0$ can only have a trivial solution $u_b(x) = 0$. Hence, $W(a) \neq 0$.

51.3. Solving the boundary value problem. Suppose that either $\alpha_a \neq 0$, or $\alpha_b \neq 0$, or $q(x) \neq 0$. A solution to the boundary value problem is sought as the sum

$$u(x) = u_0(x) + u_p(x)$$

where u_0 is the solution to the associate homogeneous problem

$$Lu_0 = 0\,, \quad B_a(u_0) = \gamma_a\,, \quad B_b(u_0) = \gamma_b$$

and u_p is a solution to the non-homogeneous problem with trivial boundary conditions

$$Lu_p = g\,, \quad B_a(u_p) = B_b(u_p) = 0$$

By linearity of the Sturm-Liouville operator and the boundary conditions, u satisfies the equation

$$Lu = L(u_0 + u_p) = Lu_0 + Lu_p = g$$

and the boundary conditions

$$B_j(u_0 + u_p) = B_j(u_0) + B_j(u_p) = \gamma_j\,, \quad j = a, b$$

The associate homogeneous problem. A solution to the homogeneous problem is sought in the form

$$u_0(x) = c_a u_a(x) + c_b u_b(x)$$

where u_a and u_b are linearly independent solutions such that $B_a(u_a) = 0$ and $B_b(u_b) = 0$. By linearity of $B_{a,b}$:

$$\gamma_a = B_a(u_0) = c_a B_a(u_a) + c_b B_a(u_b) = c_c B_a(u_b) \quad \Rightarrow c_b = \frac{\gamma_a}{B_a(u_b)}$$

$$\gamma_b = B_b(u_0) = c_a B_b(u_a) + c_b B_b(u_b) = c_a B_b(u_a) \quad \Rightarrow c_a = \frac{\gamma_b}{B_b(u_a)}$$

Note that $B_a(u_b) \neq 0$ and $B_b(u_a) \neq 0$ as is shown above. Thus, the associate homogeneous problem has a unique solution:

$$u_0(x) = \frac{\gamma_b}{B_b(u_a)}\, u_a(x) + \frac{\gamma_a}{B_a(u_b)}\, u_b(x)$$

The associated non-homogeneous problem. A solution is obtained by the method of variation of parameters. It is sought in the form

$$u_p(x) = v_a(x)u_a(x) + v_b(x)u_b(x)$$

where the unknown functions $v_{a,b}$ are subject to an additional condition

$$v_a'(x)u_a(x) + v_b'(x)u_b(x) = 0$$

Taking this condition into account, the boundary conditions require that

$$\begin{aligned} B_a(u_p) &= v_a(a)B_a(u_a) + v_b(a)B_a(u_b) = v_b(a)B_a(u_b) = 0 \\ &\Rightarrow v_b(a) = 0 \\ B_b(u_p) &= v_a(b)B_b(u_a) + v_b(b)B_b(u_b) = v_a(b)B_b(u_a) = 0 \\ &\Rightarrow v_a(b) = 0 \end{aligned}$$

The substitution of u_p into the equation yields

$$Lu_p = v_aLu_a + v_bLu_b - p(v_a'u_a' + v_b'u_b') = -p(v_a'u_a' + v_b'u_b') = g$$

Since $p(x) \neq 0$, this yields a second equation for the derivatives of the unknown functions

$$v_a'(x)u_a'(x) + v_b'(x)u_b'(x) = -\frac{g(x)}{p(x)}$$

The system has a unique solution because the determinant of the matrix of the system is the Wronskian

$$W(x) = u_a(x)u_b'(x) - u_a'(x)u_b(x) \neq 0$$

Therefore the unknown functions are solutions to the initial value problems

$$\begin{aligned} v_a'(x) &= \frac{g(x)u_b(x)}{p(x)W(x)} = \frac{g(x)u_b(x)}{p(a)W(a)}, \quad v_a(b) = 0, \\ v_b'(x) &= -\frac{g(x)u_a(x)}{p(x)W(x)} = -\frac{g(x)u_a(x)}{p(a)W(a)}, \quad v_b(a) = 0 \end{aligned}$$

so that

$$\begin{aligned} v_a(x) &= -\frac{1}{p(a)W(a)}\int_x^b u_b(y)g(y)\,dy \\ v_b(x) &= -\frac{1}{p(a)W(a)}\int_a^x u_a(y)g(y)\,dy \end{aligned}$$

The solution can written in the form

$$u_p(x) = \int_a^b G(x,y)g(y)\,dy\,,$$

$$G(x,y) = -\frac{1}{p(a)W(a)}\begin{cases} u_a(x)u_b(y) & , \quad a \le x \le y \\ u_b(x)u_a(y) & , \quad y \le x \le b \end{cases}$$

The function $G(x,y)$ is *the Green's function for the Sturm-Liouville operator.*

Properties of the Green's function. By construction, for any given $y \in (a,b)$, $G(x,y)$ is twice continuously differentiable in x and satisfies the homogeneous Sturm-Liouville equation:

$$L_x G(x,y) = 0\,, \quad x \ne y$$

and the zero boundary conditions in the variable x

$$B_{ax}(G(x,y)) = B_{bx}(G(x,y)) = 0\,, \quad y \in (a,b)$$

because $Lu_a = 0$, $B_a(u_a) = 0$ and $Lu_b = 0$, $B_b(u_b) = 0$. It is also continuous in $[a,b]\times[a,b]$. However, the classical derivative with respect to x does not exist at $x = y$, where the derivative has a jump discontinuity

$$\lim_{x\to y^-} p(x)D_xG(x,y) = -\frac{u_a'(y)u_b(y)}{p(a)W(a)}p(y)$$

$$\lim_{x\to y^+} p(x)D_xG(x,y) = -\frac{u_b'(y)u_a(y)}{p(a)W(a)}p(y)$$

so that the discontinuity jump is

$$\mathrm{disc}_{x=y} p(x)D_xG(x,y) = -\frac{p(y)W(y)}{p(a)W(a)} = -1$$

by the Liuoville-Ostrogradsky theorem. Therefore the distributional derivative of $p(x)D_xG(x,y)$ is given by

$$D_x\Big(p(x)D_xG(x,y)\Big) = -\delta(x-y) + \Big\{D_x\Big(p(x)D_xG(x,y)\Big)\Big\}$$

where $\{Df\}$ stands for the classical derivative wherever it exists. Therefore $G(x,y)$ is the fundamental solution for the Sturm-Liouville operator

$$L_xG(x,y) = \delta(x-y) + \{L_xG(x,y)\} = \delta(x-y)\,, \quad a < y < b\,,$$

that satisfies the boundary conditions

$$\alpha_a G(a,y) - \beta_a G_x'(a,y) = 0\,, \quad \alpha_b G(b,y) + \beta_b G_x'(b,y) = 0$$

Remark. If $\alpha_a = \alpha_b = 0$ and $q(x) = 0$, then the problem does not have solution for a generic g. In this case, the boundary conditions are

$$u_p'(a) = u_p'(b) = 0$$

If $u_p(x)$ is a solution, then

$$\begin{aligned}\int_a^b g(x)\,dx &= \int_a^n Lu_p(x)\,dx = -\int_a^b (p(x)u_p'(x))'\,dx \\ &= -p(x)u_p'(x)\Big|_a^b \\ &= 0\end{aligned}$$

by the boundary conditions. The vanishing integral of g is a necessary condition for the solution to exist in this case. If g satisfies this condition, then the solution exists and can still be written in the form

$$u_p(x) = c + \int_a^b G_0(x,y)g(y)\,dy$$

where c is a constant and the Green's function $G_0(x,y)$ can be obtained by integrating the equation

$$-(p(x)u_p'(x))' = g(x)$$

and using the boundary conditions. It is a fundamental solution for the Sturm-Liouville operator with appropriate boundary conditions The details are left to the reader as an exercise.

51.4. Non-linear Sturm-Liouville problem. Let the inhomogeneity in the Sturm-Liouville problem also depend non-linearly on the unknown function u:

$$Lu(x) = \lambda f(x, u(x))\,, \quad x \in (a,b)$$

where λ is a numerical parameter, and the boundary conditions remain the same as before. Suppose that $u \in C^2(a,b) \cap C^1[a,b]$ is a solution to the non-linear Sturm-Liouville problem. Then it is a solution to the Fredholm equation

$$u(x) = u_0(x) + \lambda \int_a^b G(x,y) f(y, u(y))\,dy = Tu(x)$$

where $G(x,y)$ is the Green's function for the Sturm-Liuoville operator. Conversely, by the properties of u_0 and $G(x,y)$, the function $Tu(x)$ is continuous on $[a,b]$ for any continuous u. Thus,

$$T:\ C^0[a,b] \to C^0[a,b]\,.$$

Suppose that u is a fixed point of T (a solution to the integral Fredholm equation). Using the explicit form of the Green's function

$$u = u_0 + v_a u_a + v_b u_b$$

where

$$\begin{aligned} v_a(x) &= -\frac{\lambda}{p(a)W(a)} \int_x^b u_b(y) f(y, u(y))\, dy \\ v_b(x) &= -\frac{\lambda}{p(a)W(a)} \int_a^x u_a(y) f(y, u(y))\, dy \end{aligned}$$

Therefore if $u \in C^0[a,b]$, then $v_{a,b} \in C^1[a,b]$, assuming that $f(x,z)$ is continuous in $[a,b] \times \mathbb{R}$. Therefore $u \in C^1[a,b]$ and

$$u' = u_0' + v_a u_a' + v_b u_b'$$

Recall that $v_a' u_a + v_b' u_b = 0$. It is not difficult to verify that the solution to the integral equation satisfies the boundary conditions:

$$B_a(u) = B_a(u_0) = \gamma_a\,, \quad B_b(u) = B_b(u_0) = \gamma_b$$

Let us see that u also is a solution to the Sturm-Liouville equation. Since u_0 and $u_{a,b}$ are from $C^2(a,b) \cap C^1[a,b]$, and $v_{a,b}$ are shown to be from $C^1[a,b]$, the above equation for u' shows that $u \in C^2(a,b)$ and

$$\begin{aligned} u'' &= u_0'' + v_a' u_a' + v_b' u_b' + v_a u_a'' + v_b u_b'' \\ &= u_0'' - \frac{f(x,u(x))}{p(x)W(x)} + v_a u_a'' + v_b u_b'' \end{aligned}$$

where the Liouville-Ostrogradsky theorem was used. It follows from this equation that u satisfies the equation

$$Lu(x) = f(x, u(x))\,.$$

Thus, the non-linear Sturm-Liouville problem and the integral Fredholm equation are equivalent.

Let us show that T is a contraction for a sufficiently small λ. Suppose that $f(x,z)$ is Lipschitz continuous in the variable z:

$$|f(x,z_1) - f(x,z_2)| \leq h(x)|z_1 - z_2|\,, \quad h \in C^0[a,b]$$

and continuous in the variable x for each z. Then, $f(x,u(x))$ is continuous for every continuous $u(x)$. Thus, the non-linear Strum-Liouville problem and the integral Fredholm problem are equivalent. It follows that for any $x \in [a, b]$ and any continuous u and v:

$$\begin{aligned}|Tu(x) - Tv(x)| &\le |\lambda| \int_a^b |G(x, y)h(y)|u(y) - v(y)|\, dy \\ &\le |\lambda| \int_a^b |G(x, y)h(y)\, dy \|u - v\|_\infty \\ &\le \frac{|\lambda|}{\lambda_0} \|u - v\|_\infty \\ \frac{1}{\lambda_0} &= \max_{[a,b]} \int_a^b |G(x, y)|h(y)\, dy\end{aligned}$$

Therefore

$$\|Tu - Tv\|_\infty \le \frac{|\lambda|}{\lambda_0} \|u - v\|_\infty$$

and T is a contraction if

$$|\lambda| < \lambda_0$$

So, the non-linear Sturm-Liouville problem can be solved by the contraction principle. The sequence

$$u_n(x) = T^n u_0(x)$$

converges uniformly to a unique solution u of the non-linear Sturm-Liouville problem:

$$\lim_{n\to\infty} \|u_n - u\|_\infty = 0$$

By the properties of the Green's function, the contraction sequence can also be viewed as a sequence of solutions to the linear Sturm-Liouville problems

$$Lu_n(x) = \lambda f(x, u_{n-1}(x))\,, \quad B_a(u_n) = \gamma_a\,, \quad B_b(u_n) = \gamma_b$$

If $f \sim u^p$, then terms of the sequence are partial sums for a power series in λ, a perturbation expansion in a small parameter λ. The radius of convergence of the perturbation series is proved to be no less than λ_0.

51.5. Multi-variable generalizations. Let Ω be an open bounded region in $\mathbb{R}^N$. The operator

$$Lu = -\mathrm{div}\Big(p(x)\boldsymbol{\nabla} u(x)\Big) + q(x)u(x)$$

is called a Sturm-Liouville operator, where

$$p(x) > 0\,, \quad p \in C^1(\bar\Omega)\,, \quad q(x) \ge 0\,, \quad q \in C^0(\bar\Omega)$$

Let $\partial\Omega$ be smooth (or piecewise smooth) and $\mathbf{n}$ be the outward unit normal on $\partial\Omega$. Then the boundary value problem for the Sturm-Liouville operator is to find a function

$$u \in C^2(\Omega) \cap C^1(\bar{\Omega})$$

such that

$$\begin{aligned} &Lu(x) = f(x)\,, \quad x \in \Omega\,, \\ &\alpha(x)u(x) + \beta(x)\frac{\partial u}{\partial \mathbf{n}}\Big|_{\partial\Omega} = \gamma(x)\,, \quad x \in \partial\Omega \end{aligned}$$

where

$$\begin{aligned} &f \in C^0(\bar{\Omega})\,, \quad \alpha \in C^0(\partial\Omega)\,, \quad \beta \in C^0(\partial\Omega)\,, \quad \gamma \in C^0(\partial\Omega)\,, \\ &\alpha(x) \geq 0\,, \quad \beta \geq 0\,, \quad \alpha(x) + \beta(x) > 0\,, \end{aligned}$$

The last conditions mean that α and β are non-negative functions that do not vanish simultaneously anywhere in the boundary $\partial\Omega$.

The solution is sought in the form

$$u(x) = u_0(x) + u_p(x)$$

where u_0 is a solution to the associated homogeneous problem with $f(x) = 0$ and u_p is a solution to the associated problem with trivial boundary conditions, $\gamma(x) = 0$. It is possible to show that the latter problem has a unique solution if either $\alpha \neq 0$, or $\beta \neq 0$, or $q \neq 0$ (just like in the one-variable case), and there exists a Green's function $G(x, y)$ such that

$$\begin{aligned} &L_x G(x, y) = \delta(x - y)\,, \quad x, y \in \Omega\,, \\ &\alpha(x)G(x, y) + \beta(x)\frac{\partial G(x, y)}{\partial \mathbf{n}_x} = 0\,, \quad x \in \partial\Omega\,,\ y \in \Omega \end{aligned}$$

and

$$u_p(x) = \int_\Omega G(x, y) f(y)\, d^N y$$

Example: the Laplace operator in a ball. Let us find the Green's function for the Laplace operator in a ball $|x| < R$ in a three-dimensional Euclidean space that vanishes on the boundary of the ball:

$$\begin{aligned} &-\Delta_x G(x, y) = \delta(x - y)|,\,, \quad |x| < R\,,\ |y| < R\,, \\ &G(x, y)\Big|_{|x|=R} = 0\,, \quad |y| < R \end{aligned}$$

In physics, $G(x, y)$ is the electric potential created by a unit electric charge positioned at y in a conducting grounded sphere. It can be solved by the method of images. The solution is sought in the form

$$G(x,y) = \frac{1}{4\pi|x-y|} + g(x,y)$$

where $g(x,y)$ is a harmonic function in the ball $|x| < R$ that is chosen so that G fulfills the boundary conditions:

$$\Delta_x g(x,y) = 0\,, \qquad g(x,y)\Big|_{|x|=R} = -\frac{1}{4\pi|x-y|}\Big|_{|x|=R}.$$

Recall that the first term in G in a fundamental solution for the Laplace operator in $\mathbb{R}^3$.

The method of images assumes that it is possible to find a charge (or charges) outside $\bar{\Omega}$ such that the total potential vanishes on the boundary $\partial\Omega$. Since the potential of a point charge is a harmonic function everywhere except the position of the charge, G satisfies the same equation if g is the Coulomb potential of point charges outside $\bar{\Omega}$. It turns out that one image charge is sufficient if Ω is a ball. Let y^* be a position of the image charge. Then

$$g(x,y) = -\frac{q}{4\pi|x-y^*|}\,, \quad |y^*| > R\,, \quad q > 0$$

The charge must be positive $q > 0$ to make the zero potential on the sphere. By rotational symmetry, the charges and the center of the ball should lie in a line:

$$y^* = ky\,, \quad k > \frac{R}{|y|} > 1$$

Let θ be the angle between the position vectors $y \neq 0$ and x, where lies on the sphere $|x| = R$. Then by the cosine theorem

$$|x-y|^2 = R^2 + |y|^2 - 2R|y|\cos(\theta)$$

and

$$\begin{aligned}|x-y^*|^2 &= R^2 + |y^*|^2 - 2R|y^*|\cos(\theta)\\ &= R^2 + k^2|y|^2 - 2kR|y|\cos(\theta)\end{aligned}$$

The boundary condition requires that

$$|x-y|^2 = \frac{1}{q^2}\,|x-y^*|^2\,, \quad |x| = R$$

must hold for any choice of θ, which is equivalent to two conditions

$$R^2 + |y|^2 = \frac{1}{q^2}\Big(R^2 + k^2|y|^2\Big)\,, \quad 2R|y| = \frac{2kR|y|}{q^2}$$

It follows from the second condition that $q = \sqrt{k}$, and the first condition yields a quadratic equation for k whose root ($k > R/|y| > 1$) defines the *Kelvin's transform*

$$y^* = ky = \frac{R^2}{|y|^2}\, y$$

Thus,

$$G(x,y) = \frac{1}{4\pi|x-y|} - \frac{R|y|}{4\pi\Big||y|^2x - R^2y\Big|}$$

Note that this equation makes sense when $y \to 0$, while $x \neq y$. In this case,

$$\lim_{y\to 0} G(x,y) = \frac{1}{4\pi|x|} - \frac{1}{4\pi R}$$

In fact, this limit also holds in the sense of distributions. The assumption $y \neq 0$ made when deriving $G(x,y)$ can be dropped.

It is also known from the theory of harmonic functions that a harmonic function in a bounded open region is uniquely defined by its values on the boundary of the region. Therefore the associate homogeneous problem

$$\Delta u_0 = 0\,, \qquad u\Big|_{|x|=R} = \gamma$$

has a unique solution. It can be found by separating variables in spherical coordinates in the form of the Fourier series over spherical harmonics Y_{lm}. Thus, the boundary value problem for the Laplace operator has the solution

$$u(x) = u_0(x) + \int\limits_{|y|<R} \left(\frac{f(y)}{4\pi|x-y|} - \frac{R|y|f(y)}{4\pi\Big||y|^2x - R^2y\Big|}\right) d^3y$$

51.6. Non-linear generalizations. The multi-variable Sturm-Liouville problem can be made non-linear by making the inhomogeneity to be a function of position x and the amplitude $u(x)$:

$$Lu(x) = \lambda f(x, u(x))$$

where λ is a parameter. One can show that under suitable assumptions about smoothness of the function $f(x,z)$, the non-linear boundary value problem and the Fredholm equation

$$u(x) = u_0(x) + \lambda \int_\Omega G(x,y) f(y, u(y))\, d^Ny\,, \quad u \in C^0(\bar{\Omega})$$

are equivalent. So that a solution to the problem can be found by a perturbation theory based on the contraction principle that should converge for $|\lambda| < \lambda_0$ for some $\lambda_0 > 0$.

52. Exercises

1. Newton's mechanics. Consider a motion of a particle of unit mass along a line. If $x(t)$ is the position of a particle at a time t, then $x(t)$ satisfies the initial value problem

$$x''(t) = F\Big(t, x(t), x'(t)\Big)\,, \quad t > 0\,; \qquad x(0) = q_0\,, \quad x'(0) = p_0$$

where $x(t) \in C^2(t > 0) \cap C^1(t \geq 0)$ and the force $F(t, q, p) \in C^0$ (a continuous function of three variables). For a consistency of the theory, it is natural to expect that the initial value problem has a *unique solution that depends continuously on the initial data*. The latter means that small variations of the initial data should cause small variations of the solution. This is indeed so.

(i). Prove that if $x(t)$ is a solution to the initial value problem, then it satisfies the integro-differential equation

$$x(t) = \int_0^t (t-\tau)F\Big(\tau, x(\tau), x'(\tau)\Big)\,d\tau + q_0 + tp_0$$

Hint: Use a suitable Green's function for the operator $-d^2/dt^2$.

(ii) Prove that any solution to the integro-differential equation from class $C^1([0, T])$, where $0 \leq t \leq T$, is also a solution to the initial value problem, that is, the integro-differential equation and the initial value problem are equivalent.

(iii). Define the phase space variables $\mathbf{u} = (q, p) \in \mathbb{R}^2$ where $q = x(t)$ and $p = x'(t)$. Restate the above integro-differential equation as a system of Volterra non-linear equations in the phase space:

$$q(t) = \int_0^t K_1\Big(t, \tau, q(\tau), p(\tau)\Big)\,d\tau + q_0 t + p_0\,,$$

$$p(t) = \int_0^t K_2\Big(t, \tau, q(\tau), p(\tau)\Big)\,d\tau + p_0\,,$$

that is, express the functions of 4 variables $K_{1,2}(t, \tau, q, p)$ via t and $F(\tau, q, p)$, or using the vector notations

$$\mathbf{u}(t) = \int_0^t \mathbf{K}\Big(t, \tau, \mathbf{u}(\tau)\Big)\,d\tau + \mathbf{u}_0(t) \equiv \hat{\mathbf{K}}\mathbf{u}(t) + \mathbf{u}_0(t) \equiv \hat{\mathbf{T}}\mathbf{u}(t)$$

where $\mathbf{u}_0(t) = (q_0 + p_0 t, p_0)$ and $\hat{\mathbf{K}}$ and $\hat{\mathbf{T}}$ are non-linear integral operators on the linear space of vector-valued continuous functions.

(iv). Consider a linear space of bounded vector functions $\mathbf{v}(t) \in \mathbb{R}^n$, $t \in \Omega$. Put

$$\|\mathbf{v}\|_\infty = \sup_\Omega |\mathbf{v}(t)|$$

where $|\cdot|$ denotes the Euclidean norm. Show that $\|\cdot\|_\infty$ defines a norm in the linear space of vector-valued bounded functions.

(v). Use that bounded functions form a Banach space, to show that the linear space of vector-valued bounded functions (parametric curves) is also a Banach space with respect to the norm $\|\cdot\|_\infty$ (denoted by $\mathcal{B}$). In particular, show that the subset of continuous vector-valued functions on a closed bounded interval $\Omega = [a, b]$ is closed in the Banach space $\mathcal{B}$.

(vi). Suppose that the kernel of $\hat{\mathbf{K}}$ is continuous with respect to t and τ and Lipschitz continuous with respect the vector argument $\mathbf{u}$:

$$|\mathbf{K}(t, \tau, \mathbf{u}) - \mathbf{K}(t, \tau, \mathbf{v})| \le m_0 |\mathbf{u} - \mathbf{v}| , \qquad t, \tau \in [0, t_0]$$

where $t_0 > 0$ and the constant m_0 is independent of the vectors $\mathbf{u}$ and $\mathbf{v}$ (but m_0 may depend on t_0). Restate the above condition in terms of the force F. In particular, if F and its partial derivatives are continuous, state sufficient conditions on F and its partial derivatives in order for $\hat{\mathbf{K}}$ to be Lipschitz continuous.

(vii). Prove that a power of $\hat{\mathbf{T}}$ is a contraction on $\mathcal{B}$ with $[a, b] = [0, t_0]$, that is,

$$\|\mathbf{T}^n \mathbf{u} - \mathbf{T}^n \mathbf{v}\|_\infty \le \rho_n \|\mathbf{u} - \mathbf{v}\|_\infty , \qquad \rho_n < 1$$

for some positive integer n. Prove that the initial value problem in Newton's mechanics has a unique solution in any interval $[0, t_0]$
Hint: Use the same method (based on mathematical induction) as for the scalar non-linear Volterra integral equation.

(vii). Show that the solution to the initial value problem depends continuously on the initial data in the following sense. If $\mathbf{u}_1(t)$ and $\mathbf{u}_2(t)$ are two solutions corresponding to two functions $\mathbf{u}_{01}(t)$ and $\mathbf{u}_{02}(t)$ depending on the initial data (see Part (ii)), then

$$\|\mathbf{u}_1 - \mathbf{u}_2\|_\infty \le \frac{\sigma_n}{1 - \rho_n} \|\mathbf{u}_{01} - \mathbf{u}_{02}\|_\infty$$

where $\rho_n < 1$ is defined in Part (iv) (find an expression for σ_n via m_0 and t_0).

Hint: Define

$$\mathbf{T}_j\mathbf{u} = \hat{\mathbf{K}}\mathbf{u} + \mathbf{u}_{0j}\,, \quad j = 1, 2$$

so that $\mathbf{u}_j$ are fixed points of $\mathbf{T}_j$. Next show that

$$\begin{aligned} |\mathbf{T}_1\mathbf{u}(t) - \mathbf{T}_2\mathbf{v}(t)| &\le m_0 \int_0^t |\mathbf{u}(\tau) - \mathbf{v}(\tau)| d\tau + \|\mathbf{u}_{01} - \mathbf{u}_{02}\|_\infty \\ &\le m_0 t \|\mathbf{u} - \mathbf{v}\|_\infty + \|\mathbf{u}_{01} - \mathbf{u}_{02}\|_\infty \\ |\mathbf{T}_1^2\mathbf{u}(t) - \mathbf{T}_2^2\mathbf{v}(t)| &= |\mathbf{T}_1(\mathbf{T}_1\mathbf{u})(t) - \mathbf{T}_2(\mathbf{T}_2\mathbf{v})(t)| \\ &\le \frac{m_0^2 t^2}{2!} \|\mathbf{u} - \mathbf{v}\|_\infty + (m_0 t + 1)\|\mathbf{u}_{01} - \mathbf{u}_{02}\|_\infty \end{aligned}$$

where the last inequality was obtained by a subsequent use of the first two. Proceed by induction to derive a similar upper bound for $|\mathbf{T}_1^n\mathbf{u}(t) - \mathbf{T}_2^n\mathbf{v}(t)|$. Then set $\mathbf{u} = \mathbf{u}_1$ and $\mathbf{v} = \mathbf{u}_2$ to obtain the required result.

2. Initial value problem for the sine-Gordon equation. Consider the Cauchy problem

$$\begin{aligned} &u''_{tt} - c^2 u''_{xx} + \sin(u) = 0\,, \quad t > 0\,,\ x \in \mathbb{R}\,, \\ &u\Big|_{t=0} = u_0(x)\,, \quad u'_t\Big|_{t=0} = u_1(x) \end{aligned}$$

Reduce the problem to an integral equation. Show that the integral equation has a unique solution.

3. Special case of the Sturm-Liouville operator. Consider the boundary value problem

$$L_0 u(x) = -\Big(p(x)u'(x)\Big)' = f(x)\,, \quad x \in (a, b)\,, \quad u'(a) = u'(b) = 0$$

(i) Show that

$$u(x) = c + \int_a^b G(x, y) f(y)\, dy$$

where c is a constant and

$$\int_a^b f(x)\, dx = 0\,,$$

and find an explicit form of the Green's function $G(x, y)$ for the operator L_0.

(ii) Find $L_{0x}G(x, y)$ in the sense of distributions and find boundary conditions to which $G(x, y)$ satisfies at $x = a, b$ if $a < y < b$.

4. Non-linear boundary value problem for the second-derivative operator. Consider the nonlinear boundary value problem

$$u''(x) = \lambda u^4(x)\,, \quad x \in (0,1)\,, \quad u(0) = u(1) = 1\,,$$

where $u \in C^2(0,1) \cap C^0([0,1])$.

(i) Find the Green's function $G(x,y)$ of the operator of the second derivative such that

$$-\frac{d^2}{dx^2}G(x,y) = \delta(x-y)\,, \quad G(0,y) = G(1,y) = 0\,, \quad y \in (0,1)$$

(ii) Prove that if u is a solution to the boundary value problem, then it satisfies the Fredholm equation:

$$u(x) = 1 - \lambda \int_0^1 G(x,y)(u(y))^4 dy \equiv (Tu)(x)\,.$$

(iii) Show that $T:\ C^0([0,1]) \to C^0([0,1])$, that is, the function Tu is continuous on $[0,1]$ if u is continuous on $[0,1]$.

(iv) Prove that if $u \in C^0([0,1])$ is a solution to the Fredholm equation, then it is from class C^2 and a solution to the boundary value problem. That is, the Fredholm problem and the boundary value problems are equivalent.

(v) Prove that the subset

$$S = \{u \in C^0[0,1] \mid 0 \le u(x) \le 1\}$$

is closed in $C^0[0,1]$ but not a linear manifold in $C^0[0,1]$, and that the transformation T maps S into itself,

$$T:\ S \to S\,, \quad 0 \le \lambda \le 8\,.$$

Hint: Show that $G(x,y) \ge 0$ if $(x,y) \in [0,1] \times [0,1]$ and then show that $Tu(x) \le 1$ if $\lambda \ge 0$. Find a condition on λ under which $Tu(x) \le 1$.

(vi) Prove that T is a contraction if $\lambda \in [0,2)$.
Hints: Show first that

$$|(v(x))^4 - (u(x))^4| \le 4\|u-v\|_\infty\,, \qquad u, v \in S$$

Recall the identity $(a^4 - b^4) = (a-b)(a^3 + a^2b + ab^2 + b^3)$. Use this inequality to show that

$$\|Tv - Tu\|_\infty \le \frac{\lambda}{2}\,\|v-u\|_\infty$$

(vii) Prove that the nonlinear boundary value problem has a unique solution for $\lambda \in [0, 2)$ and for any $u_0 \in S$ the sequence $\{u_n\}$ generated by the iteration scheme

$$u_n'' = \lambda(u_{n-1})^4 , \quad u_n(0) = u_n(1) = 1 , \quad n = 1, 2, ...$$

converges uniformly to the solution of the boundary value problem.

(viii) Let $u_0(x) = 1$. Find $u_n(x)$ for $n = 1, 2, 3$.

5. Non-linear boundary value problem for an oscillator. Consider the non-linear boundary value problem

$$-u''(x) + \omega^2 u(x) = \lambda u^3(x) , \quad 0 < x < L , \quad u(0) = \gamma_0 , \ u(L) = \gamma_L$$

(i) Find an equivalent Fredholm integral equation for this problem.
(ii) Find an interval of values of the parameter λ for which the Fredholm equation can be solved by the contraction principle.
(iii) Find three first terms of the perturbation theory solution:

$$u(x) = u_0(x) + \lambda u_1(x) + \lambda^2 u_2(x) + O(\lambda^3) ,$$

if λ lies in the interval from part (ii).