Chapter 3
Boundary value problems

June 3, 2021

Boundary value problems are introduced, with emphasis on variational and numerical methods.

Contents

1 Definition of problem

1.1 General problem

Electrostatics reduces to Poisson's equation

$$-\nabla^2\Phi(\mathbf{r}) = \frac{\rho(\mathbf{r})}{\epsilon_0}$$

As with any partial differential equation (PDE), boundary conditions (BCs) have to be specified. The last Chapter implicitly used the BC that $\Phi(\mathbf{r}) \to 0$ as $\mathbf{r} \to \infty$. Now we deal with a more general version: to solve Poisson's equation in a volume V (often a finite volume), with Φ specified on the boundary $S = \partial V$. Mathematically,

$$\begin{aligned} -\nabla^2\Phi(\mathbf{r}) &= \frac{\rho(\mathbf{r})}{\epsilon_0}, \qquad \mathbf{r} \in V \\ \Phi(\mathbf{r}) &= \phi(\mathbf{r}), \qquad \mathbf{r} \in S \end{aligned} \tag{1}$$

where ϕ is a prescribed function on S. Specification of Φ on S defines a *Dirichlet problem*. We shall not discuss the *Neumann problem* in which the normal derivative is specified instead. It would in general be inconsistent to specify both Φ and the normal derivative on the surface. The reason is simple: with just the values of Φ on the surface, there is already a unique solution (see below), whose normal derivative would not in general agree with that prescribed.

1.2 Reduction to homogeneous differential equation

Let $\Phi = \Phi_1 + \Phi_2$, where the former is any function that satisfies the PDE but not necessarily the BC, e.g., as given by the superposition formula

$$\Phi_1(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\int \frac{\rho(\mathbf{s})}{|\mathbf{r}-\mathbf{s}|}\, d^3s$$

It remains to determine Φ_2, which satisfies

$$\begin{aligned} \nabla^2\Phi_2(\mathbf{r}) &= 0, \qquad \mathbf{r} \in V \\ \Phi_2(\mathbf{r}) &= \phi_2(\mathbf{r}), \qquad \mathbf{r} \in S \end{aligned}$$

where

$$\phi_2(\mathbf{r}) \quad = \quad \phi(\mathbf{r}) - \Phi_1(\mathbf{r}), \qquad \mathbf{r} \in S$$

Henceforth we deal only with boundary value problems for which the PDE is *homogeneous*.

1.3 Various methods

Many sophisticated techniques have been developed for such boundary value problems, typically for special geometries (e.g., spherical or cylindrical) and leading to various special functions. Standard textbooks may be consulted. These notes instead introduce two other classes of methods: (a) minimum principle and (b) numerical methods especially those based on discretization. The former leads to existence and uniqueness proofs, and points the way to approximate methods. The latter is important in this age of powerful computers; but we only sketch the key ideas, without going into the details of numerical schemes.

2 Minimum principle

2.1 Formulation and proof

Thus, consider the problem of a homogeneous PDE, known as Lapalce's equation:

$$\nabla^2 \Phi(\mathbf{r}) \quad = \quad 0, \qquad \mathbf{r} \in V \tag{2}$$

$$\Phi(\mathbf{r}) \quad = \quad \phi(\mathbf{r}), \qquad \mathbf{r} \in S \tag{3}$$

We claim that it is equivalent to the following minimum principle.

Take any Φ that satisfies (3) but not necessarily (2). Calculate the positive number

$$I[\Phi] \quad = \quad \frac{1}{2} \int_V (\boldsymbol{\nabla}\Phi)^2 \, d^3r \tag{4}$$

Then the Φ which minimizes I satisfies (2). Before proceeding with the proof, we make two remarks. (a) The argument of I is not a number (or a few numbers) but a function. Thus I is called a *functional*. (b) Up to a factor of ϵ_0^{-1}, I is the same as the electrostatic energy, so we can think of the method heuristically as minimizing the energy.[1]

To prove this theorem, let Φ be the minimum and make an infinitesimal change

$$\Phi(\mathbf{r}) \quad \mapsto \quad \Phi(\mathbf{r}) + \delta\Phi(\mathbf{r})$$

staying within the class of functions that satisfy the BC; thus

$$\delta\Phi(\mathbf{r}) \quad = \quad 0, \qquad \mathbf{r} \in S \tag{5}$$

The first-order variations around a minimum must be zero:

$$\begin{aligned} 0 \quad &= \quad \delta I = \int_V (\boldsymbol{\nabla}\Phi) \cdot (\boldsymbol{\nabla}\delta\Phi) \, d^3r \\ &= \quad -\int_V (\nabla^2\Phi) \, \delta\Phi \, d^3r \end{aligned} \tag{6}$$

in which we have integrated by parts and used the condition (5) to discard the surface term. But since $\delta\Phi$ is arbitrary, this shows that $\nabla^2 \Phi = 0$.

Problem 1
Fill in the missing steps in (6) by starting from the identity

$$\boldsymbol{\nabla} \cdot [(\boldsymbol{\nabla}\Phi) \, \delta\Phi] \quad = \quad (\nabla^2\Phi) \, \delta\Phi + (\boldsymbol{\nabla}\Phi) \cdot (\boldsymbol{\nabla}\delta\Phi)$$

Integrate over V and use Gauss' theorem to convert the LHS to a surface integral. §

2.2 Existence and uniqueness

The minimization principle allows a proof of existence and uniqueness.

Existence
Since $I \geq 0$, there must be a minimum, and the minimum satisfies the PDE. This proves the existence of a solution. In fact, this property will lead to approximate methods by searching (within some restricted domain) for the approximate minimum.

Uniqueness
Suppose Φ_1 and Φ_2 are two different solutions to (2) and (3). Then $\Phi = \Phi_1 - \Phi_2 \neq 0$ is a solution to

$$\nabla^2 \Phi(\mathbf{r}) \quad = \quad 0, \qquad \mathbf{r} \in V \tag{7}$$

$$\Phi(\mathbf{r}) \quad = \quad 0, \qquad \mathbf{r} \in S \tag{8}$$

and must lead to minimum I within the class of functions that satisfy the BC (8). Now consider the variation

$$\Phi \quad \mapsto \quad (1 - \delta) \, \Phi$$

which stays within the same class (8); this would not be the case if the BC is not homogeneous. Then

$$I \quad \mapsto \quad (1 - 2\delta) \, I$$

showing that the original is not an extremum, a contradiction.

[1] This statement is strictly true only if Φ is the correct potential. That is why we say "heuristically".

3 An example

3.1 Statement of problem

Consider the space V between two concentric spherical surfaces of radii a and b, i.e., $a \le r \le b$ (**Figure 1**). The inner surface is at $\Phi = 1$ and the outer surface is at $\Phi = 0$. We want to find the potential $\Phi(r)$ for all r. In particular, we want to find (a) $\Phi(r)$ for $r = r_m = (a+b)/2$; and (b) the capacitance C of the system, with the spherical surfaces regarded as the two poles. For numerical work, take $a = R$, $b = 3R$ and express the capacitance as $C = \gamma\pi\epsilon_0 R$, where γ is a pure number to be determined.

3.2 Exact solution

In this case, there is an exact solution. The potential goes as $1/r$ up to a multiplicative and an additive constant, and obviously

$$\begin{aligned}
\Phi(r) &= \left(\frac{1}{r} - \frac{1}{b}\right)\left(\frac{1}{a} - \frac{1}{b}\right)^{-1} \\
\frac{d\Phi}{dr} &= -\frac{1}{r^2}\left(\frac{1}{a} - \frac{1}{b}\right)^{-1} \\
I &= \frac{1}{2}\int \left(\frac{d\Phi}{dr}\right)^2 4\pi r^2\, dr \\
&= 2\pi\,\frac{ab}{b-a} = 3\pi R \qquad (9)
\end{aligned}$$

All integrals in this example are understood to be over the interval $a < r < b$.

By considering the energy, and noticing that the capacitor is charged to a voltage of one unit, we have

$$\begin{aligned}
\frac{1}{2}C &= U = \frac{\epsilon_0}{2}\int E^2\, dV \\
C &= 2\epsilon_0 I
\end{aligned}$$

So the exact parameter is $\gamma = 6$, and the potential at the midpoint is obtained from (9)

$$\Phi(r_m) = \frac{(1/2) - (1/3)}{1 - (1/3)} = 0.25 \qquad (10)$$

3.3 Using a trial function

Pretend we do not know the exact solution. We can approximate by a trial function, say a linear function that interpolates between the two surfaces:

$$\begin{aligned}
\Phi &= (b-r)(b-a)^{-1} \\
\frac{d\Phi}{dr} &= -(b-a)^{-1}
\end{aligned}$$

$$\begin{aligned}
I &= \frac{1}{2}\int \left(\frac{d\Phi}{dr}\right)^2 4\pi r^2\, dr \\
&= \frac{1}{2}(b-a)^{-2}\int 4\pi r^2\, dr \\
&= \frac{2\pi}{3}\,\frac{b^3 - a^3}{(b-a)^2} = \frac{13\pi}{3}R
\end{aligned}$$

Thus we find $\gamma = 26/3 = 8.67$, which is (considerably) larger than the exact answer. The potential at the midpoint is

$$\Phi(r_m) = \frac{3-2}{3-1} = 0.5$$

also not very accurate.

3.4 Optimization

The previous estimate was not so good, because we just guessed a solution, without any optimization. Let us try to do better by including one parameter. Given the need to maintain the BCs, add a term that vanishes at both ends:

$$\begin{aligned}
\Phi &= (b-r)(b-a)^{-1} + \alpha(b-r)(r-a)a^{-2} \\
\frac{d\Phi}{dr} &= -(b-a)^{-1} + \alpha[-2r + (a+b)]a^{-2}
\end{aligned}$$

where in the extra term a factor a^{-2} has been inserted to make α dimensionless. The dependence of I on α must take the form

$$I = I_0 + 2\alpha I_1 + \alpha^2 I_2$$

with I_0 being the answer as in Section 3.3, and I_1 not expected to be zero. Thus we can optimize over α, giving the choice $\alpha = -I_1/I_2$, with the corresponding minimum value of I

$$I = I_0 - \frac{I_1^2}{I_2}$$

definitely smaller than I_0.

Problem 2
Take the case $a = R$, $b = 3R$ and evaluate I_0, I_1, I_2. Find the minimum value of I and hence the parameter γ in the estimate of C. Also give an estimate for $\Phi(r_m)$. Compare with the exact solution and the less accurate estimate in the last subsection. §

For the trial function (Section 3.3) / optimized solution (Problem 2 in Section 3.4) and exact solution (Section 3.2), we have respectively $\gamma = 8.67, 6.35, 6$; $\Phi(r_m) = 0.5, 0.283, 0.25$. So adding one parameter gets a reasonably accurate solution.

4 Discretization

4.1 One dimension

Start with 1D. A function $\Phi(x)$ is sampled at discrete points $x = na$, on a grid of spacing a (dots in **Figure 2a**). For a small enough, the discrete values (and the interpolation between them) would provide a good approximation.

The first derivative is most naturally defined midway between two grid points (crosses in **Figure 2b**); or, more conveniently, on a *link*. Mathematically this means we take, for any grid point x

$$\begin{aligned} \frac{d\Phi(x+a/2)}{dx} &= a^{-1}\left[\Phi(x+a) - \Phi(x)\right] \\ \frac{d\Phi(x-a/2)}{dx} &= a^{-1}\left[\Phi(x) - \Phi(x-a)\right] \end{aligned} \quad (11)$$

The second derivative is most naturally defined on the grid points (dots on the line in **Figure 2c**). Mathematically this means

$$\begin{aligned} \frac{d^2\Phi(x)}{dx^2} &= a^{-1}\left[\frac{d\Phi(x+a/2)}{dx} - \frac{d\Phi(x-a/2)}{dx}\right] \\ &= a^{-2}\left[\Phi(x+a) - 2\Phi(x) + \Phi(x-a)\right] \\ &= -2a^{-2}\left[\Phi(x) - \tilde{\Phi}(x)\right] \end{aligned} \quad (12)$$

where

$$\tilde{\Phi}(x) \equiv \frac{1}{2}\left[\Phi(x+a) + \Phi(x-a)\right] \quad (13)$$

is the average value for the two nearest neighbors (nn).

4.2 Two dimensions

Now generalize to 2D. The analog to (12) is

$$\nabla^2\Phi(x,y) = -4a^{-2}\left[\Phi(x,y) - \tilde{\Phi}(x,y)\right] \quad (14)$$

where $\tilde{\Phi}(x,y)$ is the average over the four nn, shown by the crosses in **Figure 3**:

$$\begin{aligned} \tilde{\Phi}(x,y) \equiv \frac{1}{4}&\left[\Phi(x+a,y) + \Phi(x-a,y)\right. \\ &\left. + \Phi(x,y+a) + \Phi(x,y-a)\right] \end{aligned} \quad (15)$$

The formula for 3D would involve the average over six nn in three directions. We shall continue the examples with 2D only.

4.3 Laplace's equation

Therefore Laplace's equation has a simple meaning:

> A function Φ satisfies Laplace's equation if the value of Φ at each point agrees exactly with the average over its nn.

Write the above condition formally as

$$\boxed{\Phi = F(\Phi)} \quad (16)$$

where Φ (written without argument or index) denotes the column vector of the values at the grid points, and F is a vector function of the vector argument, which simply replaces the value at each grid point by the average of the nn:

$$F(\Phi) \equiv \tilde{\Phi}$$

Example 1
Start with a trivial example. Refer to **Figure 4**. The boundary values of Φ are as given, and there is only one unknown interior value to be determined. The answer is simply the average of the nn:

$$\Phi = \frac{1}{4}(4+6+3+5) = 4.5$$

The "corner" values on the boundary do not matter. Also, if the problem comes from discretizing a continuous problem with a small grid size a, then there should be very small differences between neighboring values. We exaggerate these differences in this and later examples to make the arithmetic more obvious. §

Example 2
Refer to **Figure 5**, with two unknown interior values Φ_1 and Φ_2. Agreement with the nn average requires

$$\begin{aligned} \Phi_1 &= \frac{1}{4}(4+\Phi_2+3+7) \\ &= 3.50 + 0.25\,\Phi_2 \\ \Phi_2 &= \frac{1}{4}(5+6+4+\Phi_1) \\ &= 3.75 + 0.25\,\Phi_1 \end{aligned} \quad (17)$$

which can be written more systematically as a matrix equation

$$\begin{pmatrix} 1 & -0.25 \\ -0.25 & 1 \end{pmatrix}\begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix} = \begin{pmatrix} 3.50 \\ 3.75 \end{pmatrix}$$

Note that the RHS comes from the prescribed boundary values. §

Problem 3
Find Φ_1, Φ_2. Answer: $\Phi_1 = 71/15 = 4.733$, $\Phi_2 = 74/15 = 4.933$. §

Problem 4
Solve the problem in **Figure 6**, where there are three unknown interior values. §

General case
Consider Example 2, and rewrite the equations as

$$\begin{aligned}\Phi_1 &= \frac{1}{4}\Phi_2 + \frac{1}{4}(4+3+7)\\ \Phi_2 &= \frac{1}{4}\Phi_1 + \frac{1}{4}(5+6+7)\end{aligned}$$

Based on this, it is easy to write down the general case in d dimensions with a total of N internal points arbitrarily labelled as $1,\ldots,N$. For two sites i and j, define $n(i,j)=1$ if i,j are nn, and 0 otherwise. Then the condition is

$$\begin{aligned}\Phi = F(\Phi) &= N\,\Phi + B\\ \Phi_i &= N_{ij}\,\Phi_j + B_i\\ N_{ij} &= n(i,j)/(2d)\\ B_i &= \sum_{j\in S} n(i,j)\,\phi_j/(2d) \qquad (18)\end{aligned}$$

where ϕ_j are the prescribed boundary values. Thus B_i is zero if site i is not nn to at least one boundary site. Then, the problem is reducible to an $N\times N$ matrix equation

$$\begin{aligned}A\,\Phi &= B\\ \Phi &= A^{-1}\,B \qquad (19)\end{aligned}$$

with

$$A = I - N \qquad (20)$$

There are ready numerical packages that can invert fairly large matrices, especially those are (a) symmetric, and (b) sparse, in the sense that many elements of A are zero.

Problem 5
Solve the problem in **Figure 7**, where there are six unknown interior values. Label these interior points as shown and write out A and B. Use MATLAB or some such package to invert A and hence determine Φ. §

5 Flow and relaxation

5.1 Curvature flow

Next introduce a class of *flow and relaxation methods*, first in a continuous form, and later in a discrete form more relevant for actual calculations. The whole point is to start with some Φ which satisfies the BC but not necessarily the PDE, and gradually change it (or let it evolve or "flow" with "time") until it converges to the correct answer.

So introduce a "time" parameter t and let $\Phi(\mathbf{r},t)$ evolve according to

$$\frac{\partial\Phi(\mathbf{r},t)}{\partial t} = \lambda\,\nabla^2\Phi(\mathbf{r},t) \qquad (21)$$

with $\lambda>0$, and subject to the BC that Φ starts off correctly and remains unchanged on S. The *fixed point* of the evolution, i.e., where Φ no longer changes, must be the correct solution. But we need to prove that Φ will approach the fixed point rather than move away from it.

To see this, consider

$$I(t) = \frac{1}{2}\int_V [\boldsymbol{\nabla}\Phi(\mathbf{r},t)]^2\; d^3r$$

and its rate of change

$$\begin{aligned}\frac{dI}{dt} &= \int_V (\boldsymbol{\nabla}\Phi)\cdot[\boldsymbol{\nabla}(\partial\Phi/\partial t)]\; d^3r\\ &= -\int_V (\nabla^2\Phi)(\partial\Phi/\partial t)\, d^3r\\ &= -\lambda\int_V (\nabla^2\Phi)^2\, d^3r \le 0\end{aligned}$$

where in going to the second line we have integrated by parts; the surface term vanishes because $\partial\Phi/\partial t = 0$ on S. Thus, provided $\lambda>0$, I will always decrease in the evolution, and will approach the minimum that we seek.[2]

Refer to one dimension; the analog of $\nabla^2\Phi(\mathbf{r})$ would be $d^2\Phi(x)/dx^2$, and the second derivative indicates curvature — a linear function would have a graph which is not curved. Thus we call this method *curvature flow*. Incidentally, the proof of the Poincare conjecture makes use of a similar idea, Ricci flow, in which a metric $g_{\mu\nu}$ is evolved in a manner proportional to the Ricci tensor $R_{\mu\nu}$, and the ultimate solution sought is a fixed point of the flow.

5.2 Discrete form

Next consider the curvature flow method for a discrete spatial lattice; all we have to do is to rewrite the RHS of (21) using (14), leading to

$$\frac{d\Phi(\mathbf{r},t)}{dt} = -\mu\left[\Phi(\mathbf{r},t) - \tilde{\Phi}(\mathbf{r},t)\right] \qquad (22)$$

where

$$\mu = 4\lambda a^{-2} > 0$$

This means we decrease (increase) Φ if it is larger (smaller) than the average of the nn values — in other words shift it *towards* the average value. It is then obvious that Φ will converge to a solution such that at every site Φ is equal to the average of the nn values — the solution to the PDE.

[2] This is plausible enough for physicists; mathematicians would want a more rigorous proof. But it should be easy to show that I does not get trapped in a local minimum that is not the correct solution. Refer to the uniqueness theorem.

5.3 Relaxation method

This idea can be taken one step further, using discrete time steps, and in each time step changing the value to the average of nn values:

$$\Phi(\mathbf{r}, t{+}1) \quad = \quad \tilde{\Phi}(\mathbf{r}, t) \tag{23}$$

This algorithm is called the *relaxation method*.

To illustrate this, go back to Example 2 and start off with $\Phi_1 = \Phi_2 = 0$. We just iterate (17). For example, the first iteration gives

$$\begin{aligned} \Phi_1 &= 3.50 + 0.25 \times 0 = 3.50 \\ \Phi_2 &= 3.75 + 0.25 \times 0 = 3.75 \end{aligned}$$

and the next iteration gives

$$\begin{aligned} \Phi_1 &= 3.50 + 0.25 \times 3.75 = 4.438 \\ \Phi_2 &= 3.75 + 0.25 \times 3.50 = 4.625 \end{aligned}$$

Problem 6
Carry out two more iterations. §

In fact it is easy to automate the iteration, as shown in the spreadsheet `relax m1.xls`. The iteration converges very rapidly to the correct answer.

Problem 7
Adapt the spreadsheet to solve Problem 4 by the relaxation method. §

It is necessary to examine the situation for large problems in order to appreciate the advantage of the relaxation method. Imagine a 3D problem, in which space is divided into 100 segments in each direction (in other words, 1 percent spatial resolution, which is not extravagant). Then there are $N = 10^6$ sites, and the matrix A has $N^2 = 10^{12}$ entries,[3] well beyond the reach of most computers. But the relaxation method deals only with the values at the N sites, one evolution step at a time; a number such as $N \sim 10^6$ is manageable.

6 Finite elements*

**This Section is more advanced, and can be skipped.*

The *finite elements* method does not require a regular grid, and is suited to irregular shapes, and/or situations where one may want to have a finer grid in one region and a coarser grid in another.

Consider the 2D problem of solving (2) with BC (3) on a domain V with boundary S. This problem is equivalent to minimizing $I[\Phi]$ as defined by (4). Cover V by a patchwork of *elements*, each a triangle. Then

$$I \quad = \quad \sum_\gamma I_\gamma$$

where I_γ is the corresponding integral over the element labelled by γ.

Consider one element, and let its three vertices be labelled for the moment as $i = 1, 2, 3$, with coordinates (x_i, y_i) and function values Φ_i. Represent $\Phi(\mathbf{r})$ within this element as a linear function[4]

$$\Phi(x, y) \quad = \quad A + Bx + Cy$$

Applying this at the three points gives

$$A + Bx_i + Cy_i \quad = \quad \Phi_i$$

for $i = 1, 2, 3$, or, in matrix form

$$P \begin{pmatrix} A \\ B \\ C \end{pmatrix} \quad = \quad \begin{pmatrix} \Phi_1 \\ \Phi_2 \\ \Phi_3 \end{pmatrix}$$

where

$$P \quad = \quad \begin{pmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{pmatrix}$$

The solution is

$$\begin{pmatrix} A \\ B \\ C \end{pmatrix} \quad = \quad Q \begin{pmatrix} \Phi_1 \\ \Phi_2 \\ \Phi_3 \end{pmatrix}$$

where $Q = P^{-1}$.

The integral is then

$$I_\gamma \quad = \quad \frac{1}{2} V_\gamma \left(B_\gamma^2 + C_\gamma^2 \right)$$

where the labels on B and C have been restored, and V_γ is the "volume" of the element.

The matrix P, its inverse Q and the "volumes" V_γ for each element depend only on the grid, and are readily evaluated. Since B_γ and C_γ are linear in the potentials, I must be a quadratic in the values of Φ at the vertices. All the coefficients depend only on the geometry of the elements, and can be evaluated once the grid is chosen.

For most elements, all three vertices involve internal points with (unknown) function values Φ_i. But some elements may involve vertices on the

[3] Only half are independent, but that does not alter the qualitative aspects of the argument.

[4] The constants A, B, C are different for each element, and should therefore carry a subscript γ, which is omitted for notational simplicity.

boundary with (known) function values ϕ_m. Thus the most general quadratic is

$$I = \sum_{ij} \frac{1}{2} M_{ij} \Phi_i \Phi_j + \sum_{im} N_{im} \Phi_i \phi_m + \ldots \quad (24)$$

where i, j range over internal points, m ranges over boundary points, and terms involving two boundary points have been omitted because these do not contribute to the minimization. The coefficients are determined by the grid points and assumed to have been evaluated.

Minimization with respect to Φ_i then leads to

$$\sum_j M_{ij} \Phi_j + \sum_m N_{im} \phi_m = 0$$

or, in obvious matrix notation,

$$M\,\Phi = K \quad (25)$$

where

$$K_i = -\sum_m N_{im} \phi_m$$

reducing the problem to inverting a matrix.

The apparently tedious intermediate steps are related only to the geometry of the grid points. This step is now fairly well automated by software packages that are "universal". You may get a sense of what is involved by working out a problem with just a small number of elements.

It is straightforward to generalize to higher dimensions, and to other problems reducible to the minimization of quadratic integrals.

7 T Matrix method*

**This Section is more advanced, and can be skipped.*

7.1 Boundary value problem with dielectrics

Although the macroscopic version of EM in the presence of material media is not yet formally introduced and will not be a focus of this course, it is fair to assume some prior familiarity. This Section sketches the main ideas behind a class of techniques for boundary value problems in the presence of a linear dielectric.

Class of problems
A dielectric body occupies a volume V (typical linear dimension R), without loss of generality centered at the origin; the dielectric constant is κ inside V and unity outside (vacuum). A uniform external field $\mathbf{E}^0$ is imposed far away. What is the field $\mathbf{E}(\mathbf{r})$ everywhere? Let $S = \partial V$ be the boundary of the dielectric.

The problem is specified by

$$\nabla^2 \Phi = 0 \quad (26)$$

inside V and also outside V, with the following three BCs.

- Φ is continuous across S.
- The normal component of $\mathbf{D}$ is continuous across S.
- Far away, the field approaches $\Phi^0(\mathbf{r}) = -\mathbf{E}^0 \cdot \mathbf{r}$.

Problem 8
Take the special case where the dielectric is a sphere of radius R. Show that the field $\mathbf{E}$ inside is uniform and of course in the same direction as the external field, with magnitude E given by

$$\frac{E}{E^0} = \frac{3}{\kappa + 2} \quad (27)$$

Hint: Let the external field be in the z direction, and separate variables in spherical coordinates. Only $P_\ell(\cos\theta)$ with $\ell = 1$ is involved, and the allowed radial functions are r^ℓ and $r^{-\ell-1}$. §

General equations
In obvious notation,

$$\begin{aligned} \mathbf{E}(\mathbf{r}) &= -\boldsymbol{\nabla}\Phi(\mathbf{r}) \\ \mathbf{D}(\mathbf{r}) &= \kappa(\mathbf{r})\epsilon_{\mathbf{0}}\,\mathbf{E}(\mathbf{r}) \end{aligned} \quad (28)$$

where κ is the position-dependent dielectric constant:

$$\kappa(\mathbf{r}) = \begin{cases} \kappa & : \ \mathbf{r} \in V \\ 1 & : \ \mathbf{r} \notin V \end{cases}$$

Thus, the system is governed by

$$\boxed{\boldsymbol{\nabla} \cdot [\kappa(\mathbf{r}) \boldsymbol{\nabla} \Phi(\mathbf{r})] = 0} \quad (29)$$

and

$$\Phi(\mathbf{r}) \to \Phi^0(\mathbf{r}) \quad \text{as } |\mathbf{r}| \to \infty \quad (30)$$

Problem 9
Show that (29) can be derived by minimizing

$$I = \frac{1}{2} \int \kappa(\mathbf{r}) \, (\boldsymbol{\nabla}\Phi(\mathbf{r}))^2 \, dV \quad (31)$$

where the volume integral is over all space and the minimization is carried out over all Φ satisfying the BC (30). §

Problem 10
In (31), of course $\boldsymbol{\nabla}\Phi$ has to be finite, i.e., Φ has to be continuous. This implies that, across the boundary, the *tangential* component of $\mathbf{E}$ is continuous.

By integrating (29) over a "pillbox" across S, show that the *normal* component of $\mathbf{D}$ is continuous. §

Discretization?
For a general shape, there is no analytic solution. One might be tempted to try a discretization scheme such as those introduced earlier in this Chapter, just adding factors of κ in suitable places. But there are difficulties in this case.

- Since the BC is imposed at infinity, a huge volume would have to be included, say of linear dimension $L \gg R$. Discretize with a grid of length $a \ll R$. The number of points required will be
$$N_1 \quad \sim \quad (L/a)^3$$
If one takes (quite modestly) $L/R \sim R/a \sim 100$, then $L/a \sim 10^4$, $N_1 \sim 10^{12}$ and in any matrix method, the number of elements is $(N_1)^2 \sim 10^{24}$ — which is not practical.
- The potential Φ is not smooth across the boundary; in particular, the normal component of $\boldsymbol{\nabla}\Phi$ is not continuous. Using a grid to approximate a non-smooth function is tricky.

This Section introduces a class of methods that overcome these difficulties.

7.2 Integral equation

Rewrite (29) as

$$-\nabla^2\Phi(\mathbf{r}) \quad = \quad \boldsymbol{\nabla}\cdot[\mu(\mathbf{r})\,\boldsymbol{\nabla}\Phi(\mathbf{r})] \qquad (32)$$

where

$$\mu(\mathbf{r}) \quad = \quad \begin{cases} \kappa - 1 & : \quad \mathbf{r} \in V \\ 0 & : \quad \mathbf{r} \notin V \end{cases}$$

Think of the RHS of (32) as the source; the formal solution is

$$\Phi(\mathbf{r}) \quad = \quad \Phi^0(\mathbf{r}) + \int G(\mathbf{r}-\mathbf{s})\,\partial_i\left[\mu(\mathbf{s})\,\partial_i\Phi(\mathbf{s})\right]\,dV$$

where G is the Green's function for the operator $-\nabla^2$, the volume integral is for $\mathbf{s}$ over all space, and a homogeneous solution Φ^0 has been added.

Problem 11
Show that the integral above vanishes as $\mathbf{r} \to \infty$. Hence show that Φ^0 is the potential corresponding to the externally imposed field. §

Integrate by parts to obtain

$$\Phi(\mathbf{r}) \quad = \quad \Phi^0(\mathbf{r}) + \int \partial_i G(\mathbf{r}-\mathbf{s})\;\left[\mu(\mathbf{s})\,\partial_i\Phi(\mathbf{s})\right]\,dV$$

Two properties have been used in the evaluation. (a) The surface term at infinity vanishes, because $\mu(\mathbf{s}) = 0$ there. (b) The minus sign coming from integration by parts cancels another one in converting the derivative with respect of $\mathbf{s}$ into a derivative with respect to the argument of G, namely $\mathbf{r}-\mathbf{s}$.

Because of the factor $\mu(\mathbf{s})$, the integral can now be restricted to V, and $\mu = \kappa - 1$ can be taken out of the integral:

$$\boxed{\Phi(\mathbf{r}) = \Phi^0(\mathbf{r}) + \mu\int_V \partial_i G(\mathbf{r}-\mathbf{s})\,\partial_i\Phi(\mathbf{s})\,dV} \qquad (33)$$

Consider this for $\mathbf{r} \in V$. Then (33) gives a consistency condition on Φ making reference *only* to V, i.e., the inside. All reference to the outside has been removed, and the BC at infinity is expressed through the function Φ^0. Therefore the two difficulties mentioned earlier have been resolved.

If the inside is discretized with grid size a, then the number of points required is

$$N_2 \sim (R/a)^3 \quad \ll \quad N_1$$

which is a major improvement. For the example cited, $N_2 \sim 3\times 10^6$ and the number of elements of the matrix is $(N_2)^2 \sim 10^{12}$ compared with $(N_1)^2 \sim 10^{24}$.

From this point onwards, there are several different approaches, and two are introduced.

7.3 Reduction to surface integral

In (33), change the first derivative as

$$\partial_i = \frac{\partial}{\partial r_i} = -\frac{\partial}{\partial s_i}$$

and pull the $\partial_i\Phi(\mathbf{s})$ factor inside the derivative, since the difference involves $\nabla^2\Phi = 0$. Then

$$\Phi(\mathbf{r}) = \Phi^0(\mathbf{r}) - \mu\int_V \frac{\partial}{\partial s_i}\left[G(\mathbf{r}-\mathbf{s})\,\partial_i\Phi(\mathbf{s})\right]\,dV$$

which can be converted to a surface integral

$$\Phi(\mathbf{r}) \quad = \quad \Phi^0(\mathbf{r}) - \mu\int_S G(\mathbf{r}-\mathbf{s})\,\varphi(\mathbf{s})\,dS \qquad (34)$$

where

$$\varphi(s) \quad = \quad \mathbf{n}\cdot\boldsymbol{\nabla}\Phi(s)$$

is the normal derivative (on the inside). Notice two properties. (a) Φ everywhere is expressed as an integral over the surface values of φ once the latter is known. (b) The surface values satisfy the consistency condition

$$\boxed{\varphi(\mathbf{r}) = \varphi^0(\mathbf{r}) - \mu\int_S K(\mathbf{r},\mathbf{s})\varphi(\mathbf{s})\,dS} \qquad (35)$$

where

$$K(\mathbf{r},\mathbf{s}) \;=\; \mathbf{n}\cdot\boldsymbol{\nabla} G(\mathbf{r}-\mathbf{s})$$

and $\mathbf{n}$ is the unit normal at $\mathbf{r}$.[5] This has a structure similar to (33), except that it refers to a surface rather than a volume. If the surface is discretized with grid size a, then the number of points is

$$N_3 \sim (R/a)^2 \ll N_2$$

which is another major gain in efficiency. Again, using the numbers in our example, $N_3 \sim 10^4$ and the number of matrix elements would be $(N_3)^2 \sim 10^8$.

Note how, in this example, a huge reduction in the scale of the matrix has been achieved

$$10^{24} \mapsto 10^{12} \mapsto 10^8$$

Discretization
Only a sketch of the discretization procedure is provided. Divide the surface into patches labelled by $\alpha = 1, 2, ...N$ (where $N \sim N_3$), with centers at $\mathbf{r}_\alpha$. Then for any function F on the surface,

$$\int_S F(\mathbf{s})\, dS \;\approx\; \sum_\beta F_\beta\, \Delta_\beta$$

where F_β is the function value at position $\mathbf{r}_\beta$ and Δ_β is the area of the patch β. Applying this to (35) at $\mathbf{r}_\alpha$ gives

$$\varphi_\alpha \;=\; \varphi^0_\alpha - \mu \sum_\beta M_{\alpha\beta}\, \varphi_\beta \qquad (36)$$

with the shorthand

$$\begin{aligned} \varphi_\alpha &= \varphi(\mathbf{r}_\alpha) \\ \varphi^0_\alpha &= \varphi^0(\mathbf{r}_\alpha) \\ M_{\alpha\beta} &= K(\mathbf{r}_\alpha, \mathbf{r}_\beta)\Delta_\beta \end{aligned}$$

In obvious matrix notation,

$$\begin{aligned} \varphi &= \varphi^0 - \mu M\, \varphi \\ \varphi &= T\, \varphi^0 \\ T &= (1+\mu M)^{-1} \end{aligned} \qquad (37)$$

In other words, the whole problem is reduced to inverting an $N \times N$ matrix, where $N \sim N_3$. Matrix inversion can be handled by a number of ready numerical packages — but only because the size of the problem has been drastically reduced.

Discussion
First, the problem can be approached in a number of slightly different ways, but the key idea is always the same, as described above. Second, the details of the numerical scheme still need to be worked out, but that is mostly book keeping. Third, evaluation of the matrices M and T depends only on the geometry of the grid points, without reference to the potential values.

In honesty, the method described above is not used that often in electrostatics. But the analogous T matrix method is extremely useful in scattering (at a definite frequency). The only significant change is that G is replaced by the analogous Green's function for the operator

$$-\left[\nabla^2 + (\omega/c)^2\right]$$

In the case of EM scattering, two further complications (neither of which is conceptually difficult) also need to be handled: (a) both the electric field and the magnetic field have to be evaluated, and these are not reducible to a scalar potential; and (b) there is another length scale, namely $\lambda = 2\pi c/\omega$. A review of the T matrix method can be found in Ref. [1].

7.4 Expansion in basis functions

Motivation
There is another way to handle a consistency condition such as (35). Instead of representing the function φ in terms of its values at N_3 discrete points, expand it as

$$\varphi(\mathbf{s}) \;=\; \sum_n a_n\, f_n(\mathbf{s})$$

and likewise for φ^0, where $\{f_n\}$ is a set of basis functions. In practice, one truncates the sum to say N terms, and therefore obtain, very similarly to (37), a set of equations, in obvious schematic notation,

$$\begin{aligned} a &= a^0 - \mu M\, a \\ a &= T\, a^0 \\ T &= (1+\mu M)^{-1} \end{aligned} \qquad (38)$$

where the expression for M is left as an exercise. The point here is that functions such as φ or Φ are smooth — after all Φ satisfies Laplace's equation, which heuristically states that the function at any point is equal to the average of the nearest neighbors. It is therefore possible to express many (say $N_3 \sim 10^4$) function values in terms of relatively few (say $N_4 < 10$) coefficients in an expansion. The matrix problem gets further reduced to a much more manageable size.

However, for expansion in terms of basis functions, there is not much extra labor in using the

[5] This makes K a function of the two position variables and not just a function of the difference.

volume integral rather than the surface integral, and we continue the formalism based on (33).

Expansion
Do a Taylor series expansion in terms of Cartesian coordinates[6]

$$\Phi(\mathbf{r}) \quad = \quad a_i r_i + a_{ij} r_i r_j + a_{ijk} r_i r_j r_k + \ldots$$

where without loss of generality the constant term has been omitted.

Now take a derivative of (33) and evaluate at $\mathbf{r} = 0$:

$$a_j \quad = \quad a_j^0 + \mu \int_V \partial_j \partial_i G(\mathbf{s}) \partial_i \Phi(\mathbf{s}) \, dV$$

Then put the expansion into the integrand:

$$a_j = a_j^0 + \mu \int_V \partial_j \partial_i G(\mathbf{s}) \left(a_i + a_{ij} s_j + \ldots \right) \, dV \quad (39)$$

If all terms are kept in the parentheses, this expression is exact. Now truncate this sum by keeping only the first term; in other words, approximate the potential inside V as a linear function. Then

$$a_j \quad = \quad a_j^0 - \mu M_{ji} a_i \quad (40)$$

$$M_{ji} \quad = \quad - \int_V \partial_j \partial_i G(\mathbf{s}) \, dV \quad (41)$$

The same type of matrix equation as before is obtained, but now the matrix is only 3×3.

The shape factor
Within this approximation, the answer depends only on the matrix M in (41), to be called the *shape factor*, for reasons to be explained.

- A simple count shows that M is dimensionless. Therefore it depends only on the shape of V and not its size: if V is scaled up by some factor, M is unchanged.
- The trace is exactly unity.
- Therefore for any object with cubic symmetry (including of course a sphere), $M_{ij} = (1/3)\delta_{ij}$.
- Thus, the result (27) follows trivially — provided the assumption of uniform field inside V can be justified.

We have formulated the analogous problem for the more complicated case of elasticity [2], and formulas for M are given therein.[7]

Higher-order terms
We indicate only briefly and schematically how higher-order terms can be accommodated. As an example, keep one more term in (39). Then (40) would acquire additional terms on the RHS proportional to a_{jk}, with coefficients proportional to

$$\int_V \partial_j \partial_i G(\mathbf{s}) s_k \, dV$$

With these extra unknowns a_{jk}, more equations are needed. These are obtained by taking one more derivative of (33) and evaluating at $\mathbf{r} = 0$:

$$\begin{aligned} a_{jk} \quad &= \quad a_{jk}^0 + \mu \int_V \partial_k \partial_j \partial_i G(\mathbf{s}) \partial_i \Phi(\mathbf{s}) \, dV \\ &= \quad a_{jk}^0 + \mu \int_V \partial_k \partial_j \partial_i G(\mathbf{s}) \\ &\qquad \times \left(a_i + a_{im} s_m + \ldots \right) dV \\ &= \quad a_{jk}^0 + (\ldots)_{jki} a_i \\ &\qquad + (\ldots)_{jkim} a_{im} + \ldots \end{aligned} \quad (42)$$

where the coefficients are shown schematically. If now we keep only to the terms shown, there are just enough equations to solve for the unknowns a_j and a_{jk}, again by inverting a matrix of modest size.

Exact solution for certain shapes
If only a constant field is imposed at spatial infinity, then $a_{jk}^0 = 0$ and (42) would give the trivial solution $a_{jk} = 0$ if

$$\int_V \partial_k \partial_j \partial_i G(\mathbf{s}) \, dV \quad = \quad 0 \quad (43)$$

When this condition is satisfied (and likewise for higher derivatives), all the higher Taylor coefficients vanish, and the assumption of constant field is valid.

It turns out that (43) and the analog for higher derivatives hold exactly if V is any ellipsoid. This somewhat surprising identity is shown in Ref. [2], and is the reason why the problem is exactly soluble for ellipsoids. For other shapes, it is necessary to solve these larger matrix systems — which is slightly tedious but doable.

A Another view of relaxation method

The relaxation method can be viewed in a different way. In Section 4, the key equation (16) was solved by inverting an $N \times N$ matrix. But equations of the type

$$\Phi \quad = \quad F(\Phi)$$

can be solved by the repeated iteration of

$$\Phi \quad \mapsto \quad F(\Phi) \quad (44)$$

[6] In doing so, we have not made use of the simplification that is possible because Φ satisfies Laplace's equation. However, the cost is modest for the first few terms.

[7] In that paper, M_{ij} is given in terms of another tensor $t_{ijk\ell}$.

Evidently, provided that the iteration converges, the limiting value is a solution to the original equation. Such an iteration is exactly the relaxation method. Seen in this light, the power of the method (provided it converges) is even more apparent — we do not even need the function F to be linear.

Problem 12
As a trivial example of such iterative methods, imagine that you have a calculator (or a spreadsheet) that can do elementary arithmetic but does not have a square root function. How would you find say $\sqrt{2}$?
(a) Write $\sqrt{2} = a + x$, where a is some convenient initial estimate and x is the correction to be found. Show that x satisfies

$$x \quad = \quad F(x) = \frac{1}{2a}(2 - a^2 - x^2)$$

(b) Take $a = 1$ and an initial guess $x = 0$. Iterate to find $\sqrt{2}$ to 4 places after the decimal. It is suggested you do this on a spreadsheet, without invoking the square root function.
(c) Discuss the convergence for different choices of a. §

References

[1] MI Mishchenkoa, LD Travisa and DW Mackowskib, "T-matrix computations of light scattering by nonspherical particles: A review", J Quant Spect and Rad Trans **55**, 535–575 (1996).
DOI: 10.1016/0022-4073(96)00002-7

[2] FC Chen and K Young, "Inclusions of arbitrary shape in an elastic medium", J Math Phys **18**, 1412–1416 (1977).
DOI: 10.1063/1.523438

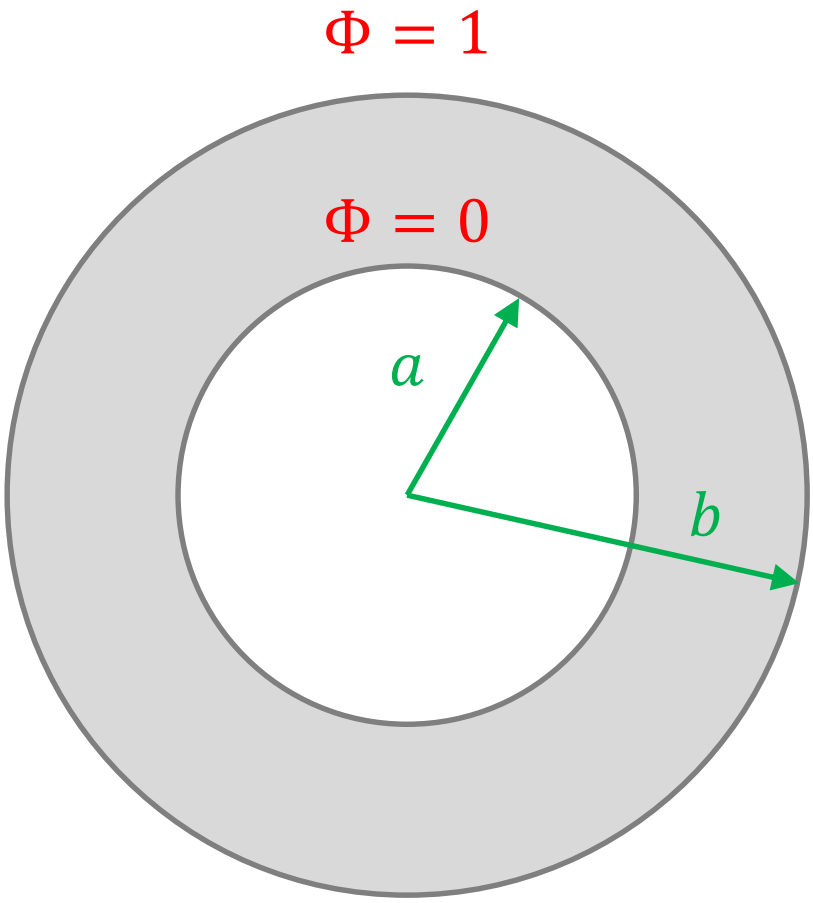

Figure 1

To find $\Phi(r)$ for $a \leq r \leq b$

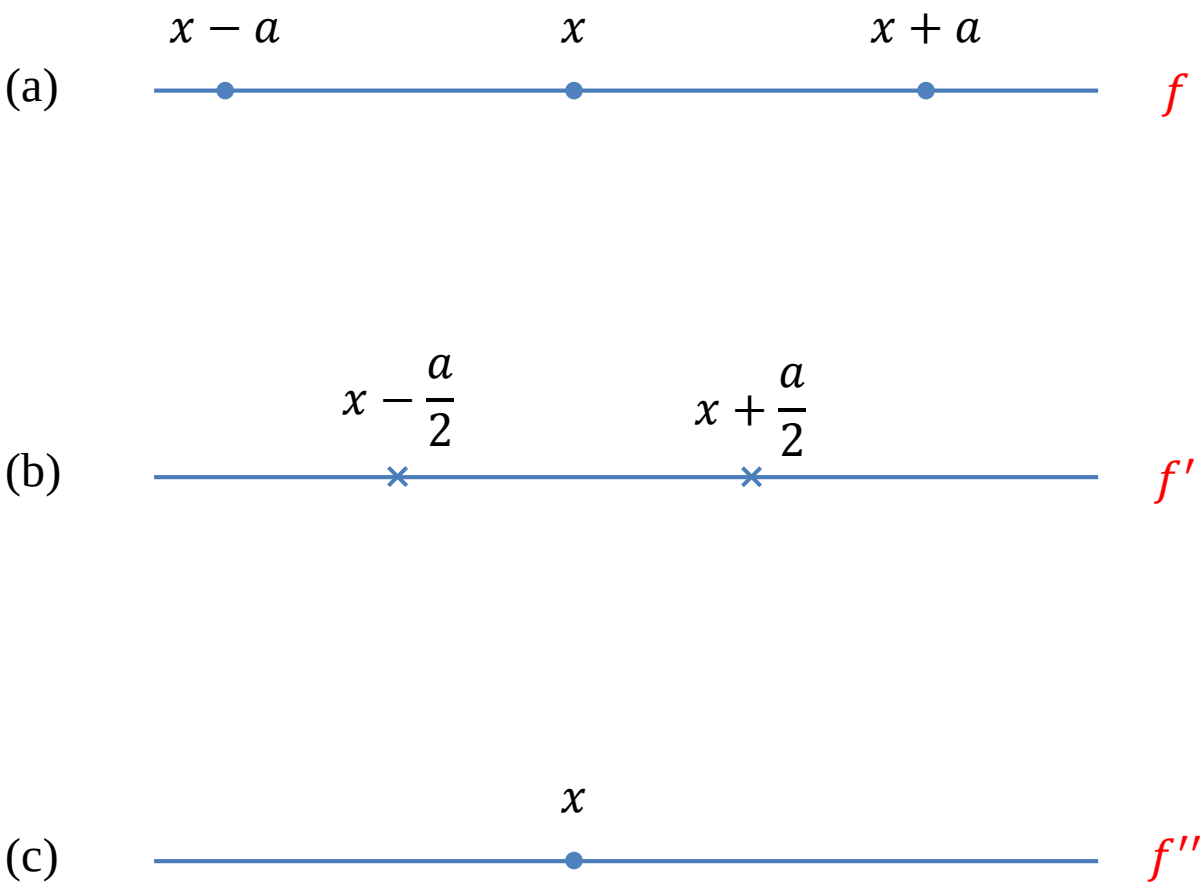

Figure 2

Different grid points for evaluating f, f', f''

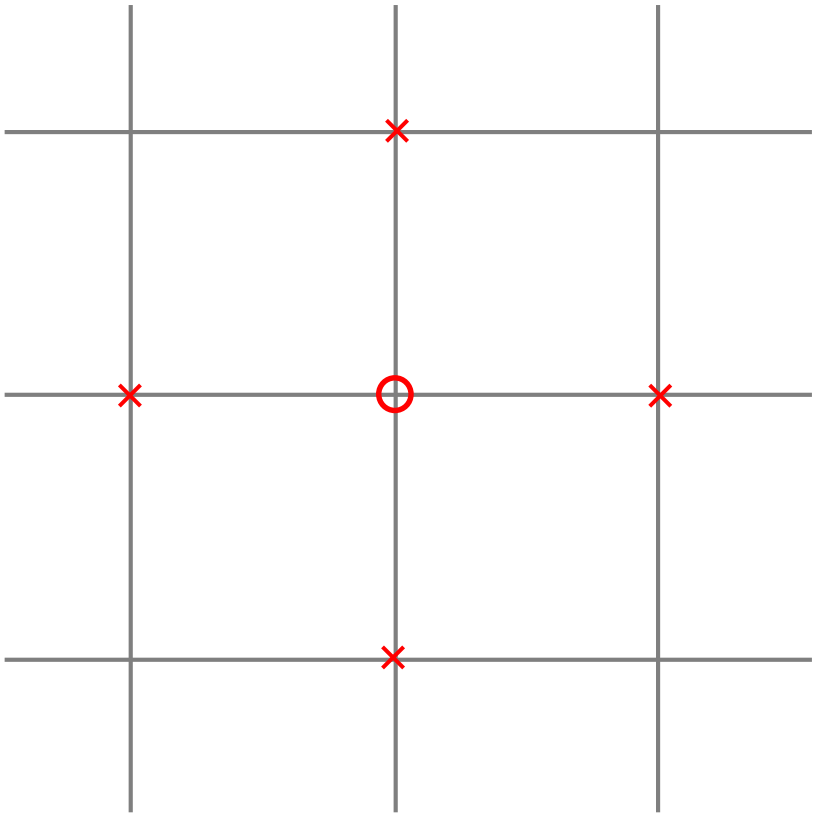

Figure 3

A site O and four nn x

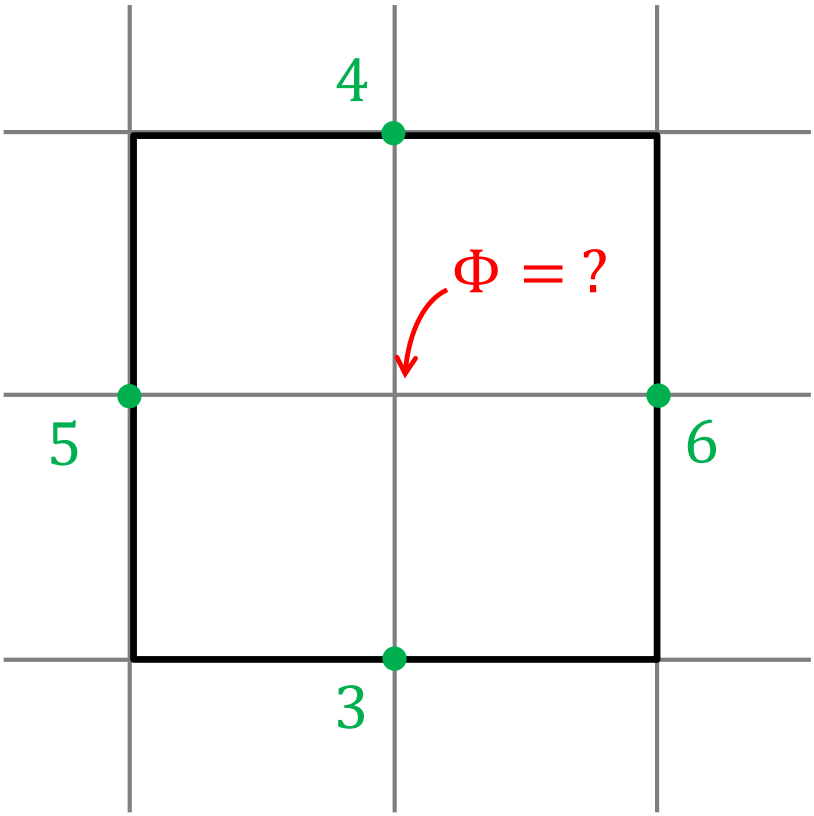

Figure 4

Find Φ from given boundary values

$$\Phi = \frac{1}{4}(4 + 6 + 3 + 5)$$

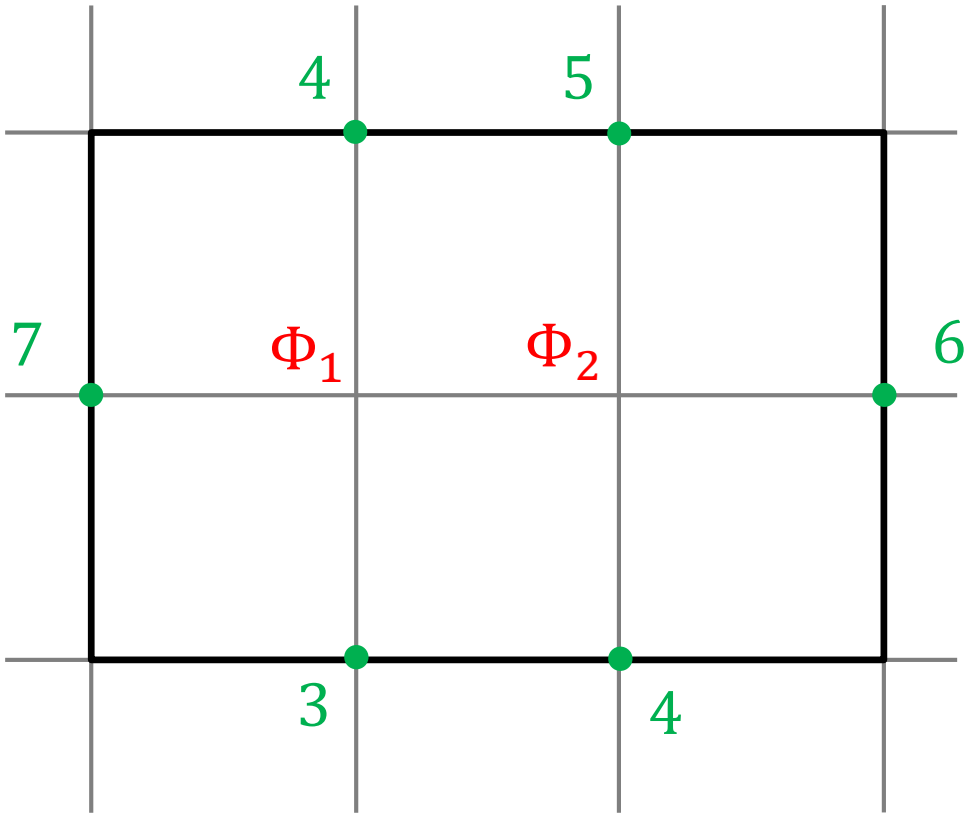

Figure 5

Find Φ_1, Φ_2 from the given boundary values

$$\Phi_1 = \frac{1}{4}(4 + \Phi_2 + 3 + 7)$$

$$\Phi_2 = \frac{1}{4}(5 + 6 + 4 + \Phi_1)$$

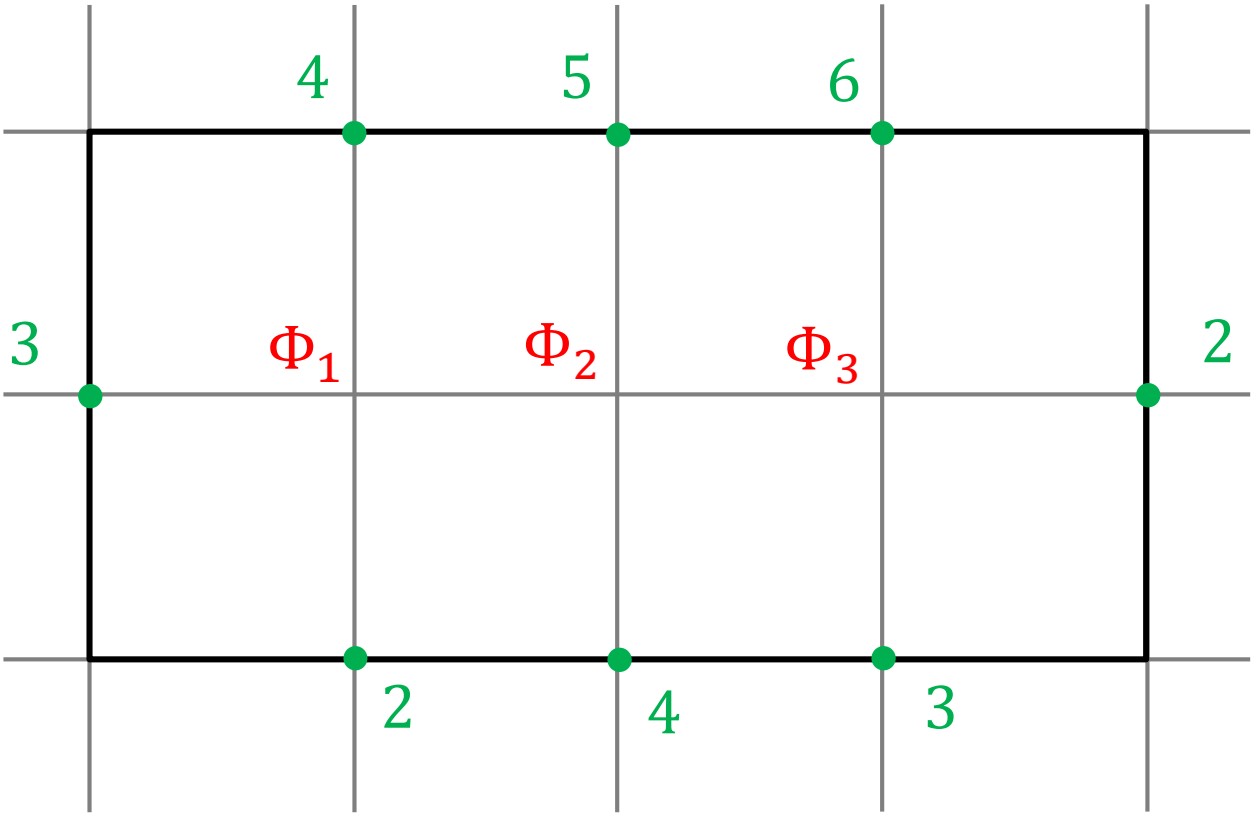

Figure 6

Find Φ_1, Φ_2, Φ_3 from given boundary values

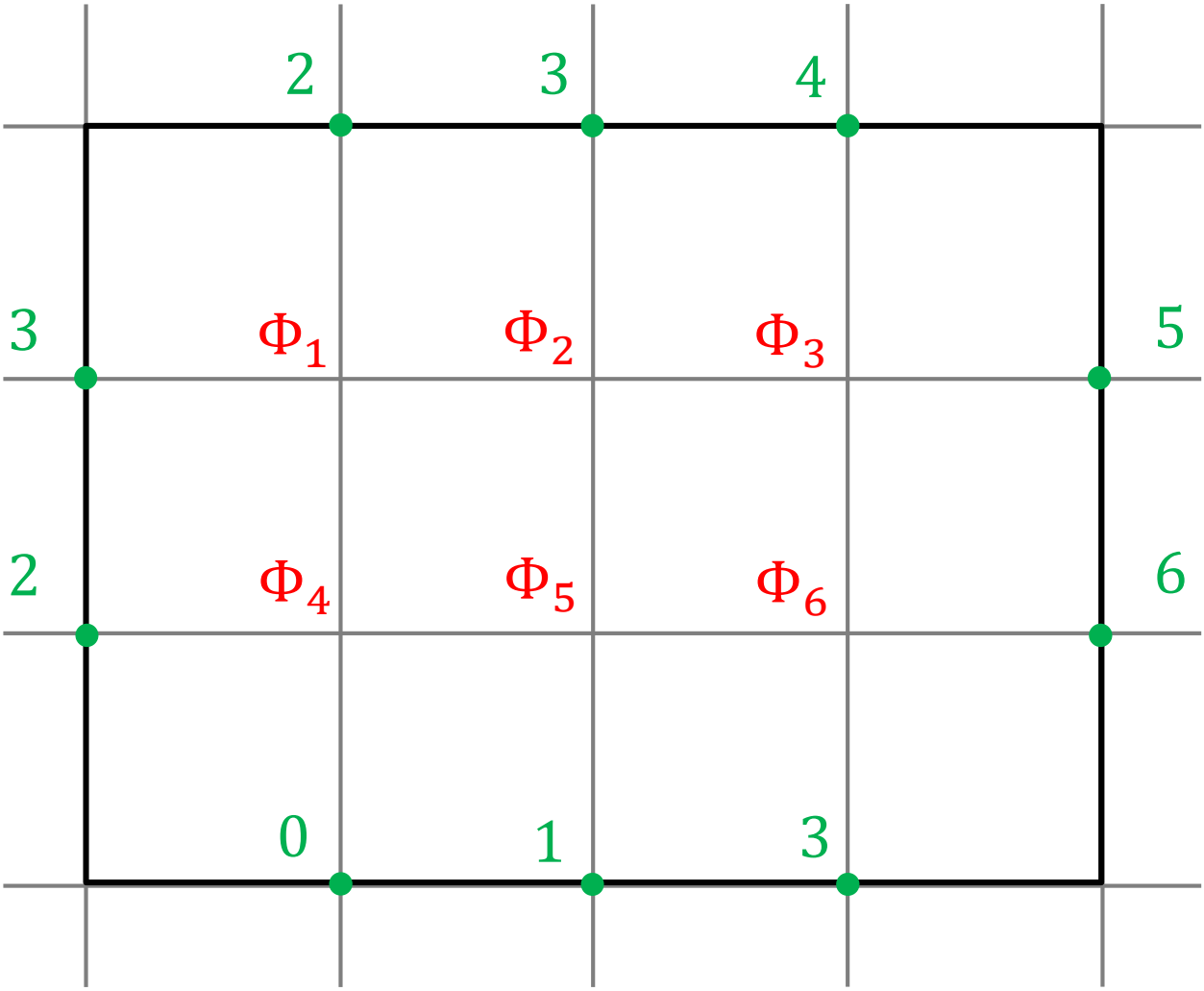

Figure 7

Find Φ_i from the boundary values