# Chapter 2. Charges and Conductors

*This chapter starts our discussion of the very common situations when the electric charge distribution in space is not known a priori, but rather should be calculated in a self-consistent way together with the electric field it creates. The simplest situations of this kind involve conductors and lead to the so-called boundary problems in that the partial differential equations describing the field distribution have to be solved with appropriate boundary conditions. Such problems are also typical for other parts of electrodynamics* (*and indeed for other fields of physics as well*)*, so following tradition, I will use this chapter's material as a playground for a discussion of various methods of boundary problem solution, and the special functions most frequently encountered on that way.*

## 2.1. Polarization and screening

The basic principles of electrostatics outlined in Chapter 1 present the conceptually full solution of the problem of finding the electrostatic field (and hence Coulomb forces) induced by electric charges distributed over space with some density $\rho(\mathbf{r})$. However, in most practical situations, this function is not known but should be found self-consistently with the field. For example, if a sample of relatively dense material is placed into an external electric field, it is typically *polarized*, i.e. acquires some local charges of its own, which contribute to the total electric field $\mathbf{E}(\mathbf{r})$ inside, and even outside it – see Fig. 1a.

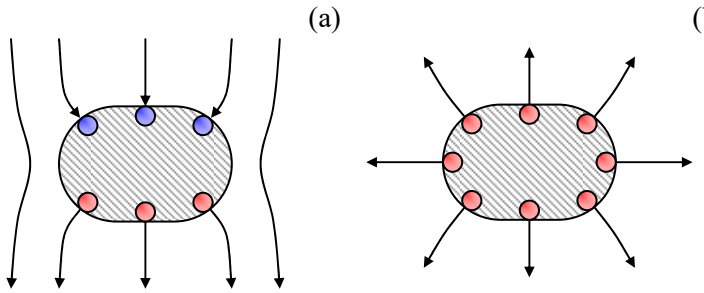



Fig. 2.1. Two typical electrostatic situations involving conductors: (a) polarization by an external field, and (b) re-distribution of the conductor's own charge over its surface – schematically. Here and below, the red and blue points denote charges of opposite signs.

The full solution of such problems should satisfy not only the fundamental Eq. (1.7) but also the so-called *constitutive relations* between the macroscopic variables describing the sample's material. Due to the atomic character of real materials, such relations may be very involved. In this part of my series, I will have time to address these relations, for various materials, only rather superficially,[1] focusing on their simple approximations. Fortunately, in most practical cases such approximations work very well.

In particular, for the polarization of good conductors, a very reasonable approximation is given by the so-called *macroscopic model*, in which the free charges in the conductor are treated as a charged continuum that is free to move under the effect of the force $\mathbf{F} = q\mathbf{E}$ exerted by the *macroscopic* electric field $\mathbf{E}$, i.e. the field averaged over space on the atomic scale – see also the discussion at the end of Sec.

[1] A more detailed discussion of the electrostatic field screening may be found, e.g., in SM Sec. 6.4. (Alternatively, see either Sec. 13.5 of J. Hook and H. Hall, *Solid State Physics*, 2nd ed., Wiley, 1991; or Chapter 17 of N. Ashcroft and N. Mermin, *Solid State Physics*, Brooks Cole, 1976.)

1.1. In electrostatics (which excludes the case dc currents, to be discussed in Chapter 4 below), there should be no such motion, so everywhere inside the conductor the macroscopic electric field should vanish:

$$\mathbf{E} = 0\,. \tag{2.1a}$$

This is the *electric field screening*[2] effect, meaning, in particular, that conductors' polarization in an external electric field has the extreme form shown (rather schematically) in Fig. 1a, with the field of the induced surface charges completely compensating the external field in the conductor's bulk. Note that Eq. (1a) may be rewritten in another, frequently more convenient form:

Conductor: macroscopic model

$$\phi = \text{const}\,, \tag{2.1b}$$

where $\phi$ is the *macroscopic* electrostatic potential related to the macroscopic field by Eq. (1.33).[3] (If a problem includes several unconnected conductors, the constant in Eq. (1b) may be specific for each of them.)

Now let us examine what we can say about the electric field in free space just *outside* a conductor, within the same macroscopic model. At close proximity, any smooth surface (in our current case, that of a conductor) looks planar. Let us integrate Eq. (1.28) over a narrow ($d << l$) rectangular loop $C$ encircling a part of such plane conductor's surface (see the dashed line in Fig. 2a), and apply it to the electric field vector $\mathbf{E}$ the well-known vector algebra equality – the *Stokes theorem*[4]

$$\int_S (\nabla \times \mathbf{E})_n \, d^2 r = \oint_C \mathbf{E} \cdot d\mathbf{r}\,, \tag{2.2}$$

where $S$ is any surface limited by the contour $C$.

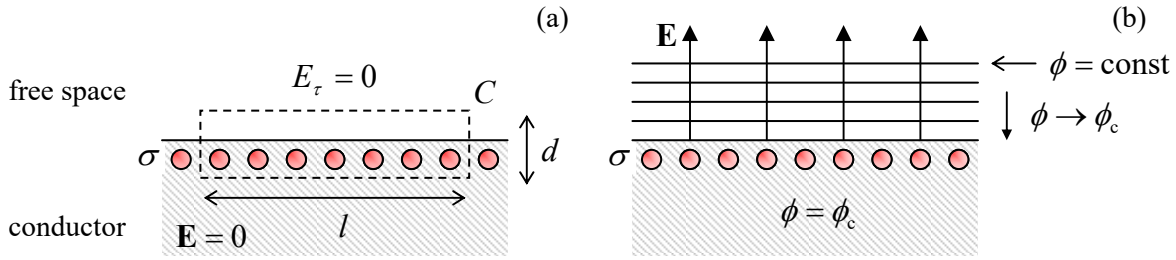



Fig. 2.2. (a) The surface charge layer at a conductor's surface, and (b) the electric field lines and equipotential surfaces near it.

In our current case, the contour is dominated by two straight lines of length $l$, so if $l$ is much smaller than the characteristic spatial scale of the field's changes but much larger than the interatomic distances, the right-hand side of Eq. (2) may be well approximated as $[(E_\tau)_{\text{in}} - (E_\tau)_{\text{out}}]\, l$, where $E_\tau$ is the tangential component of the corresponding macroscopic field, parallel to the surface. On the other hand, according to Eq. (1.28), the left-hand side of Eq. (2) equals zero. Hence, the macroscopic field's

[2] This term, used for the *electric* field, should not be confused with *shielding* – the term used for the description of *magnetic* field's reduction by magnetic materials – see Chapter 5 below.

[3] Since averaging of a function over space is a linear operation, any linear relation between genuine (microscopic) variables, including Eq. (1.33), is also valid for the corresponding macroscopic variables.

[4] See, e.g., MA Eq. (12.1).

component $E_\tau$ should be continuous at the surface, and to satisfy Eq. (1a) inside the conductor, the component has to vanish immediately outside it: $(E_\tau)_{\text{out}} = 0$. This means that the electrostatic potential immediately outside of a conducting surface cannot change along it. In other words, the equipotential surfaces outside a conductor should "lean" to the conductor's surface, with their potential values approaching the constant potential of the constant potential of the conductor – see Fig. 2b.

So, the electrostatic field just outside any conductor has to be normal to its surface. To find this normal field, we may apply the universal relation (1.24) to our macroscopic field **E**. Since in our current case $E_n = 0$ inside the conductor, we get

Surface charge density

$$\sigma = \varepsilon_0 \left(E_n\right)_{\text{out}} \equiv -\varepsilon_0 \left(\nabla \phi\right)_n \equiv -\varepsilon_0 \frac{\partial \phi}{\partial n}, \tag{2.3}$$

where $\sigma$ is the macroscopic areal density of the conductor's surface charge. Note that deriving this universal relation between the normal component of the field and the surface charge density, we have not used any cause-vs-effect arguments, so Eq. (3) is valid regardless of whether the surface charge is induced by an externally applied field (as in the case of conductor's polarization, shown in Fig. 1a), or the electric field is induced by the electric charge placed on the conductor and then self-redistributed over its surface (Fig. 1b), or it is some combination of both effects.

Before starting to use the macroscopic model for the solution of particular problems of electrostatics, let me use the balance of this section to briefly discuss its limitations. (The reader in a rush may skip this discussion and proceed to Sec. 2; however, I believe that every educated physicist has to understand when this model works, and when it does not.)

Since the argumentation which has led us to Eq. (1.24) and hence to Eq. (3) is valid for any thickness $d$ of the Gauss pillbox, within the macroscopic model, the whole surface charge is located within an infinitely thin surface layer. This is of course impossible physically: for one, this would require an infinite volumic density $\rho$ of the charge. In reality, the charged layer (and hence the region of the electric field's crossover from the finite value (3) to zero) has a nonzero thickness $\lambda$. At least three effects contribute to $\lambda$.

(i) *Atomic structure of matter*. Within each atom, and frequently between the adjacent atoms as well, the genuine ("microscopic") electric field is highly non-uniform. Thus, as was already stated above, Eq. (1) is valid only for the *macroscopic field*, i.e. the field averaged over distances of the order of the atomic size scale $a_0 \sim 10^{-10}$ m,[5] and cannot be applied to the field changes on that scale. As a result, the surface layer of charges cannot be much thinner than $a_0$.

(ii) *Thermal excitation*. According to Eq. (1.9), in the whole field-free bulk of a conductor, the net charge density, $\rho = e(n - n_e)$, [6] has to vanish, so the numbers of protons in atomic nuclei ($n$) and electrons ($n_e$) per unit volume have to be balanced. However, if an external electric field penetrates a conductor, free electrons can shift in or out of its affected part, depending on the field's contribution to their potential energy, $\Delta U = q_e\phi = -e\phi$. (Here the arbitrary constant in $\phi$ is chosen to give $\phi = 0$ well inside the conductor.) In classical statistics, this change is described by the Boltzmann distribution:[7]

[5] This scale originates from the quantum-mechanical effects of electron motion, characterized by the *Bohr radius* $r_B \equiv \hbar^2/m_e(e^2/4\pi\varepsilon_0) \approx 0.53\times10^{-10}$ m – see, e.g., QM Eq. (1.10). It also defines the scale $E_B = e/4\pi\varepsilon_0 r_B^2 \sim 10^{12}$ SI units (V/m) of the microscopic electric fields inside atoms. (Please note how large these fields are.)

[6] In this series, $e$ denotes the fundamental charge, $e \approx 1.6\times10^{-19}$ C > 0, so that the electron's charge equals $(-e)$.

[7] See, e.g., SM Sec. 3.1.

$$n_e(\mathbf{r}) = n\exp\left\{-\frac{U(\mathbf{r})}{k_B T}\right\}, \tag{2.4}$$

where $T$ is the absolute temperature in kelvins (K), and $k_B \approx 1.38\times10^{-23}$ J/K is the Boltzmann constant. As a result, the net charge density is

$$\rho(\mathbf{r}) = en\left(1-\exp\left\{\frac{e\phi(\mathbf{r})}{k_B T}\right\}\right). \tag{2.5}$$

The penetrating electric field polarizes the atoms as well. As will be discussed in the next chapter, such polarization results in the reduction of the electric field by a material-specific dimensionless factor $\kappa$ (larger, but typically not too much larger than 1), called the *dielectric constant*. As a result, the Poisson equation (1.41) takes the so-called *Poisson-Boltzmann* form,[8]

$$\frac{d^2\phi}{dz^2} = -\frac{\rho}{\kappa\varepsilon_0} = \frac{en}{\kappa\varepsilon_0}\left(\exp\left\{\frac{e\phi}{k_B T}\right\}-1\right), \tag{2.6}$$

where we have taken advantage of the 1D geometry of the system to simplify the Laplace operator, with the $z$-axis normal to the surface.

Even with this simplification, Eq. (6) is a nonlinear differential equation allowing an analytical but rather bulky solution. Since our current goal is just to estimate the field penetration depth $\lambda$, let us simplify the equation further by considering the low-field limit: $e|\phi| \sim e|E|\lambda << k_B T$. In this limit, we may extend the exponent into the Taylor series, and keep only two leading terms (of which the first one cancels with the following unity). As a result, Eq. (6) becomes linear,

$$\frac{d^2\phi}{dz^2} = \frac{en}{\varepsilon\varepsilon_0}\frac{e\phi}{k_B T}, \qquad \text{i.e. } \frac{d^2\phi}{dz^2} = \frac{1}{\lambda^2}\phi, \tag{2.7}$$

where the constant $\lambda$, in this case, is called the *Debye* (or "*Debye-Hückel*") *screening length* $\lambda_D$:

$$\lambda_D^2 \equiv \frac{\kappa\varepsilon_0 k_B T}{e^2 n}. \tag{2.8}$$

Debye screening length

As the reader certainly knows, Eq. (7) describes an exponential decrease of the electric potential, with the characteristic length $\lambda_D$: $\phi \propto \exp\{-z/\lambda_D\}$, where the $z$-axis is directed into the conductor. Plugging in the involved fundamental constants into Eq. (8), we get the following estimate: $\lambda_D[\text{m}] \approx 70\times(\kappa\times T[\text{K}]/n[\text{m}^{-3}])^{1/2}$. According to this formula, in semiconductors at room temperature, the Debye length may be rather substantial. For example, in silicon ($\kappa \approx 12$) doped to the free charge carrier concentration $n = 3\times10^{18}$ cm$^{-3}$ (the value typical for modern integrated circuits),[9] $\lambda_D \approx 2$ nm, still well

[8] This equation and/or its straightforward generalization to the case of charged particles (ions) of several kinds is also (especially in the theories of electrolytes and plasmas) called the *Debye-Hückel equation*.

[9] There is a good reason for making an estimate of $\lambda_D$ for this case: the electric field created by the gate electrode of a field-effect transistor, penetrating into doped silicon by a depth ~$\lambda_D$, controls the electric current in this most important electronic device – on whose back all our information technology rides. Because of that, $\lambda_D$ establishes the possible scale of semiconductor circuit shrinking, which is the basis of the well-known Moore's law. (Practically, the scale is determined by integrated circuit patterning techniques, and Eq. (8) may be used to find the proper charge carrier density $n$ and hence the necessary level of silicon doping – see, e.g., SM Sec. 6.4.)

above the atomic size scale $a_0$, thus justifying the estimate However, for typical good metals ($n \sim 10^{29}$ m$^{-3}$, $\kappa \sim 10$) the same formula gives $\lambda_D \sim 10^{-11}$ m, less than $a_0$. In this case, Eq. (8) should not be taken literally, because it is based on the assumption of a continuous charge distribution.

(iii) *Quantum statistics*. Actually, the last estimate is not valid for good metals (and highly doped semiconductors) for one more reason: their free electrons obey the quantum (*Fermi-Dirac*) statistics rather than the Boltzmann distribution (4).[10] As a result, at all realistic temperatures, the electrons form a degenerate quantum gas, occupying all available energy states below some energy level $\mathscr{E}_F >> k_B T$, called the *Fermi energy*. In these conditions, the screening of a relatively low electric field may be described by replacing Eq. (5) with

$$\rho \equiv e(n - n_e) = eg(\mathscr{E}_F)(-U) = -e^2 g(\mathscr{E}_F)\phi, \tag{2.9}$$

where $g(\mathscr{E})$ is the density of quantum states (per unit volume per unit energy) at the electron's energy $\mathscr{E}$. At the Fermi surface, the density is of the order of $n/\mathscr{E}_F$.[11] As a result, we again get the second of Eqs. (7), but with a different characteristic scale $\lambda$, defined by the following relation:

Thomas-Fermi screening length

$$\lambda_{TF}^2 \equiv \frac{\kappa\varepsilon_0}{e^2 g(\mathscr{E}_F)} \sim \frac{\kappa\varepsilon_0 \mathscr{E}_F}{e^2 n}, \tag{2.10}$$

and called the *Thomas-Fermi screening length*. Since for most good metals, $n$ is of the order of $10^{29}$ m$^{-3}$, and $\mathscr{E}_F$ is of the order of 10 eV, Eq. (10) typically gives $\lambda_{TF}$ close to a few $a_0$, and makes the Thomas-Fermi screening theory valid at least semi-quantitatively.

To summarize, the electric field penetration into good conductors is limited to a depth $\lambda$ ranging from a fraction of a nanometer to a few nanometers, so for problems with a characteristic linear size much larger than that scale, the macroscopic model (1) gives very good accuracy, and we will use them in the rest of this chapter. However, the reader should remember that in many situations involving semiconductors, as well as at some nanoscale experiments with metals, the electric field penetration should be taken into account.

Another important condition of the macroscopic model's validity is imposed on the electric field's magnitude, which is especially significant for semiconductors. Indeed, as Eq. (6) shows, Eq. (7) is only valid if $e|\phi| << k_B T$, so $|E| \sim |\phi|/\lambda_D$ should be much lower than $k_B T/e\lambda_D$. In the example given above ($\lambda_D \approx 2$ nm, $T = 300$ K), this means $|E| << E_t \sim 10^7$V/m $\equiv 10^5$V/cm – the value readily reachable in the lab. In larger fields, the field penetration becomes nonlinear, leading in particular to the very important effect of *carrier depletion*; it will be discussed in SM Sec. 6.4. For typical metals, such linearity limit, $E_t \sim \mathscr{E}_F/e\lambda_{TF}$ is much higher, $\sim 10^{11}$ V/m, but the model may be violated at lower fields by other effects, such as the impact-ionization leading to *electric breakdown*, which may start at $\sim 10^6$ V/m.

## 2.2. Capacitance

Let us start using the macroscopic model from systems consisting of charged conductors only, with no so-called *stand-alone* charges in the free space outside them.[12] Our goal here is to calculate the

[10] See, e.g., SM Sec. 2.8. For a more detailed derivation of Eq. (10), see SM Chapter 3.
[11] See, e.g., SM Sec. 3.3.

distributions of the electric field **E** and potential $\phi$ in space, and the distribution of the surface charge density $\sigma$ over the conductor surfaces. However, before doing that for particular situations, let us see if there are any integral measures of these distributions, which should be our primary focus.

The simplest case is of course a single conductor in the otherwise free space. According to Eq. (1b), all its volume should have the same electrostatic potential $\phi$, evidently providing one convenient global measure of the situation. Another integral measure is provided by the total charge

$$Q \equiv \int_V \rho\, d^3r \equiv \oint_S \sigma\, d^2r\,, \tag{2.11}$$

where the last integral is extended over the whole surface $S$ of the conductor. In the general case, what can we tell about the relation between $Q$ and $\phi$? At $Q = 0$, there is no electric field in the system, and it is natural (though not absolutely necessary) to select the arbitrary constant in the electrostatic potential to have $\phi = 0$ everywhere. Then, if the conductor is charged with a non-zero $Q$, according to the linear Eq. (1.7), the electric field at any point of space has to be proportional to that charge. Hence the electrostatic potential at all points, including its value $\phi$ inside the conductor, is also proportional to $Q$:

$$\phi = \wp Q\,. \tag{2.12}$$

The proportionality coefficient $\wp$, which depends on the conductor's size and shape, but on neither $\phi$ nor $Q$, is called its *reciprocal capacitance* (or, not too often, "electric elastance"). Usually, Eq. (12) is rewritten in a different form,

$$\boxed{Q = C\phi, \quad \text{with } C \equiv \frac{1}{\wp},} \tag{2.13}$$

Self-capacitance

where $C$ is called *self-capacitance*. (Frequently, $C$ is called just *capacitance*, but as we will see very soon, for more complex situations the latter term may be ambiguous.)

Before calculating $C$ for particular geometries, let us have a look at the electrostatic energy $U$ of a single conductor. To calculate it, of the several relations discussed in Chapter 1, Eq. (1.61) is most convenient, because all elementary charges $q_k$ are now parts of the conductor charge, and hence reside at the same potential $\phi$ – see Eq. (1b) again. As a result, the equality becomes very simple:

$$U = \frac{1}{2}\phi \sum_k q_k \equiv \frac{1}{2}\phi Q\,. \tag{2.14}$$

Moreover, using the linear relation (13), the same result may be re-written in two more forms:

$$\boxed{U = \frac{Q^2}{2C} = \frac{C}{2}\phi^2.} \tag{2.15}$$

Electro-static energy

We will discuss several ways to calculate $C$ in the next sections, and right now will have a quick look at just the simplest example for that we have calculated everything necessary in the previous chapter: a conducting sphere of radius $R$. Indeed, we already know the electric field distribution: according to Eq. (1), $E = 0$ inside the sphere, while Eq. (1.19), with $Q(r) = Q$, describes the field distribution outside it, because of the evident spherical symmetry of the surface charge distribution.

[12] In some texts, these charges are called "free". This term is somewhat misleading, because they may well be bound, i.e. unable to move freely.

Moreover, since the latter formula is exactly the same as for the point charge placed in the sphere's center, the potential's distribution in space may be obtained from Eq. (1.35) by replacing $q$ with the sphere's full charge $Q$. Hence, on the surface of the sphere (and, according to Eq. (1b), through its interior),

$$\phi = \frac{1}{4\pi\varepsilon_0}\frac{Q}{R}. \tag{2.16}$$

Comparing this result with the definition (13), for the sphere's self-capacitance we obtain a very simple formula[13]

C: shere

$$\boxed{C = 4\pi\varepsilon_0 R}. \tag{2.17}$$

This formula, which should be well familiar to the reader, is convenient to get some feeling of how large the SI unit of capacitance (1 *farad*, abbreviated as F) is: the self-capacitance of Earth ($R_E \approx 6.34\times10^6$ m) is below 1 mF! Another important note is that while Eq. (17) is not exactly valid for a conductor of arbitrary shape, it implies an important general estimate

$$C \sim 2\pi\varepsilon_0 a \tag{2.18}$$

where $a$ is the scale of the linear size of any conductor.[14]

Now proceeding to a system of *two* arbitrary conductors, we immediately see why we should be careful with the capacitance definition: one constant $C$ is insufficient to describe all electrostatic properties of such a system. Indeed, here we have two, generally different conductor potentials, $\phi_1$ and $\phi_2$, that may depend on both conductor charges, $Q_1$ and $Q_2$. Using the same arguments as for the single-conductor case, we may conclude that the dependence is always linear:

$$\begin{aligned} \phi_1 &= \wp_{11}Q_1 + \wp_{12}Q_2, \\ \phi_2 &= \wp_{21}Q_1 + \wp_{22}Q_2, \end{aligned} \tag{2.19}$$

but now has to be described by more than one coefficient. Actually, it turns out that there are three rather than four different coefficients in these relations, because

$$\wp_{12} = \wp_{21}. \tag{2.20}$$

This equality may be proved in several ways, for example, using the general *reciprocity theorem of electrostatics* (whose proof was the subject of Problem 1.17):

$$\int \rho_1(\mathbf{r})\phi^{(2)}(\mathbf{r})d^3r = \int \rho_2(\mathbf{r})\phi^{(1)}(\mathbf{r})d^3r, \tag{2.21}$$

[13] In the Gaussian units, using the standard replacement $4\pi\varepsilon_0 \to 1$, this relation takes an even simpler form: $C = R$, very easy to remember. Generally, in the Gaussian units (but not in the SI system!) the capacitance has the dimensionality of length, i.e. is measured in centimeters. Note also that a fractional SI unit, 1 picofarad ($10^{-12}$ F), is very close to the Gaussian unit: 1 pF = $[(1\times10^{-12})/(4\pi\varepsilon_0\times10^{-2})]$ cm ≈ 0.8998 cm. So, 1 pF is close to the capacitance of a metallic ball with a 1-cm radius, making this unit very convenient for human-scale systems.

[14] These arguments are somewhat insufficient to say which size should be used for $a$ in the case of narrow, extended conductors, e.g., a thin, long wire. Very soon we will see that in such cases the electrostatic energy, and hence $C$, depends mostly on the *larger* size of the conductor.

where $\phi^{(1)}(\mathbf{r})$ and $\phi^{(2)}(\mathbf{r})$ are the potential distributions induced, respectively, by two electric charge distributions, $\rho_1(\mathbf{r})$ and $\rho_2(\mathbf{r})$. In our current case, each of these integrals is limited to the volume (or, more exactly, the surface) of the corresponding conductor, where each potential is constant and may be taken out of the integral. As a result, Eq. (21) is reduced to

$$Q_1\phi^{(2)}(\mathbf{r}_1) = Q_2\phi^{(1)}(\mathbf{r}_2). \tag{2.22}$$

In terms of Eq. (19), $\phi^{(2)}(\mathbf{r}_1)$ is just $\wp_{12}Q_2$, while $\phi^{(1)}(\mathbf{r}_2)$ equals $\wp_{21}Q_1$. Plugging these expressions into Eq. (22), and canceling the product $Q_1Q_2$, we arrive at Eq. (20).

Hence the 2×2 matrix of coefficients $\wp_{jj'}$ (called the *reciprocal capacitance matrix*) is always symmetric, and using the natural notation $\wp_{11} \equiv \wp_1$, $\wp_{22} \equiv \wp_2$, $\wp_{12} = \wp_{21} \equiv \wp$, we may rewrite it in a simpler form:

$$\begin{pmatrix} \wp_1 & \wp \\ \wp & \wp_2 \end{pmatrix}. \tag{2.23}$$

Plugging the relation (19), in this new notation, into Eq. (1.61), we see that the full electrostatic energy of the system may be expressed as a quadratic form of its charges:

$$U = \frac{\wp_1}{2}Q_1^2 + \wp Q_1Q_2 + \frac{\wp_2}{2}Q_2^2. \tag{2.24}$$

It is evident that the middle term on the right-hand side of this equality describes the electrostatic coupling of the conductors. (Without it, the energy would be just a sum of two independent electrostatic energies of conductors 1 and 2.)[15] Still, even with this simplification, Eqs. (19) and (20) show that in the general case of arbitrary charges $Q_1$ and $Q_2$, the system of two conductors should be characterized by three, rather than just one coefficient ("the capacitance"). This is why we may attribute a single capacitance to the system only in some particular cases.

For practice, the most important of them is when the system as the whole is electrically neutral: $Q_1 = -Q_2 \equiv Q$. In this case, the most important function of $Q$ is the difference between the conductors' potentials, called the *voltage*:[16]

$$\boxed{V \equiv \phi_1 - \phi_2,} \tag{2.25}$$

Voltage: definition

For that function, the subtraction of two Eqs. (19) gives

$$\boxed{V = \frac{Q}{C}, \quad \text{with } C \equiv \frac{1}{\wp_1 + \wp_2 - 2\wp},} \tag{2.26}$$

Mutual capacitance

where the coefficient $C$ is called the *mutual capacitance* between the conductors – or, again, just "capacitance" if the term's meaning is absolutely clear from the context. The same coefficient describes

[15] This is why systems with $\wp \ll \wp_1, \wp_2$ are called *weakly coupled*, and may be analyzed using approximate methods – see, e.g., Fig. 4 and its discussion below.

[16] A word of caution: in condensed matter physics and electrical engineering, voltage is most commonly defined as the difference between *electrochemical* rather than *electrostatic* potentials. These two notions coincide if the conductors have equal *workfunctions* – for example, if they are made of the same material. In this course, this condition will be implied, and the difference between the two voltages ignored – to be discussed in detail in SM Sec. 6.3.

the electrostatic energy of the system. Indeed, plugging Eqs. (19) and (20) into Eq. (24), we see that both forms of Eq. (15) are reproduced if $\phi$ is replaced with $V$, $Q_1$ with $Q$, and with $C$ meaning the mutual capacitance:

Capacitor's energy

$$U = \frac{Q^2}{2C} = \frac{C}{2}V^2. \tag{2.27}$$

The best-known system for which the mutual capacitance $C$ may be readily calculated is the *plane* (or "parallel-plate") *capacitor*: a system of two conductors separated with a narrow plane gap of a constant thickness $d$ and an area $A \sim a^2 >> d^2$ – see Fig. 3.

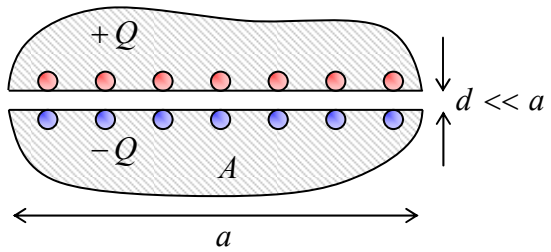



Fig. 2.3. Plane capacitor – schematically.

Since the surface charges that contribute to the opposite charges $\pm Q$ of the conductors of this system, attract each other, in the limit $d << a$ they sit entirely on the opposite surfaces limiting the gap, so there is virtually no electric field outside of the gap, while (according to the discussion in Sec. 1) inside the gap it is normal to the surfaces. According to Eq. (3), the magnitude of this field is $E = \sigma/\varepsilon_0$. Integrating this field across thickness $d$ of the narrow gap, we get $V \equiv \phi_1 - \phi_2 = Ed = \sigma d/\varepsilon_0$, so $\sigma = \varepsilon_0 V/d$. However, due to the constancy of the potential of each electrode, $V$ should not depend on the position in the gap area. As a result, $\sigma$ should be also constant over all the gap area $A$, regardless of the external geometry of the conductors (see Fig. 3 again), and hence $Q = \sigma A = \varepsilon_0 V/d$. Thus we may write $V = Q/C$, with

C: Plane capacitor

$$C = \frac{\varepsilon_0 A}{d}. \tag{2.28}$$

Let me offer a few comments on this well-known formula. First, it is valid even if the gap is not quite planar – for example, if it gently curves on a scale much larger than $d$, but retains its thickness. Second, Eq. (28), which is valid only if $A \sim a^2$ is much larger than $d^2$, ignores the nonuniform electric fields spreading to distances $\sim d$ beyond the gap edges. Such *fringe fields* result in an additional *stray capacitance* $C' \sim \varepsilon_0 a << C \sim \varepsilon_0 a \times (a/d)$.[17] Finally, the same condition ($A >> d^2$) assures that $C$ is much larger than the self-capacitance $C_j$ of each conductor – see Eq. (18).

The opportunities opened by the last fact for electronic engineering and experimental physics practice are rather astonishing. For example, a very realistic 3-nm layer of high-quality aluminum oxide, which may provide nearly perfect electric insulation between two thin conducting films, with an area of 0.1 m$^2$ (a typical area of silicon wafers used in the semiconductor industry) provides $C \sim 1$ mF,[18] larger than the self-capacitance of the whole planet Earth!

[17] The exact value of $C'$ depends on the shape of the conductors. In a rare case when it has been calculated analytically, two thin round concentric disks of radius $R$, $C' = \varepsilon_0 R\,[\ln(16\pi R/d) - 1]$.

[18] Just as in Sec. 1, for the estimate to be realistic, I took into account the additional factor $\kappa$ (for aluminum oxide, close to 10) which should be included in the numerator of Eq. (28) to make it applicable to dielectrics – see Chapter 3 below.

In a plane capacitor with $d << a$, the electrostatic coupling of the two conductors is evidently very *strong*. As an opposite example of a *weakly* coupled system, let us consider two conducting spheres of the same radius $R$, separated by a much larger distance $d$ (Fig. 4).

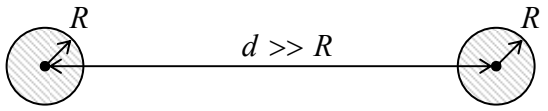



Fig. 2.4. A system of two far-separated, similar conducting spheres.

In this case, the diagonal components of the matrix (23) may be approximately found from Eq. (16), i.e. by neglecting the coupling altogether:

$$\wp_1 = \wp_2 \approx \frac{1}{4\pi\varepsilon_0 R}. \tag{2.29}$$

Now, if we had just one sphere (say, number 1), the electric potential at distance $d$ from its center would be given by Eq. (16): $\phi = Q_1/4\pi\varepsilon_0 d$. If we move to this point a small ($R << d$) sphere without its own charge, we may expect that its potential should not be too far from this result, so $\phi_2 \approx Q_1/4\pi\varepsilon_0 d$. Comparing this expression with the second of Eqs. (19) (taken for $Q_2 = 0$), we get

$$\wp \approx \frac{1}{4\pi\varepsilon_0 d} << \wp_{1,2}. \tag{2.30}$$

From here and Eq. (26), the mutual capacitance

$$C \approx \frac{1}{\wp_1 + \wp_2} \approx 2\pi\varepsilon_0 R. \tag{2.31}$$

We see that (somewhat counter-intuitively), in this limit $C$ does not depend substantially on the distance between the spheres, i.e. does *not* describe their electrostatic coupling. The off-diagonal coefficients of the *reciprocal* capacitance matrix (20) play this role much better – see Eq. (30).

Now let us consider the case when only one conductor of the two is charged, for example, $Q_1 \equiv Q$, while $Q_2 = 0$. Then Eqs. (19)-(20) yield

$$\phi_1 = \wp_1 Q_1. \tag{2.32}$$

Now, we may follow Eq. (13) and define $C_1 \equiv 1/\wp_1$ (and $C_2 \equiv 1/\wp_2$), just to see that such *partial capacitances* of the conductors of the system differ from its mutual capacitance $C$ – cf. Eq. (26). For example, in the case shown in Fig. 4, $C_1 = C_2 \approx 4\pi\varepsilon_0 R \approx 2C$.

Finally, let us consider one more frequent case when one of the conductors carries a certain charge (say, $Q_1 = Q$), but the potential of its counterpart is sustained constant, say $\phi_2 = 0$.[19] (This condition is especially easy to implement if the second conductor is much larger than the first one. Indeed, as the estimate (18) shows, in this case, it would take a much larger charge $Q_2$ to make the potential $\phi_2$ comparable with $\phi_1$.) In this case the second of Eqs. (19), with the account of Eq. (20), yields $Q_2 = -(\wp/\wp_2)Q_1$. Plugging this relation into the first of those equations, we get

[19] In electrical engineering, such a constant-potential conductor is called the *ground*. This term stems from the fact that in many cases the electrostatic potential of the (weakly) conducting ground at the Earth's surface is virtually unaffected by laboratory-scale electric charges.

$$Q_1 = C_1^{\text{ef}}\phi_1, \qquad \text{with } C_1^{\text{ef}} \equiv \left(\mathcal{P}_1 - \frac{\mathcal{P}^2}{\mathcal{P}_2}\right)^{-1} \equiv \frac{\mathcal{P}_2}{\mathcal{P}_1\mathcal{P}_2 - \mathcal{P}^2}. \tag{2.33}$$

Thus, this *effective capacitance* of the first conductor is generally different from both its partial capacitance $C_1$ and the mutual capacitance $C$ of the system, emphasizing again how accurate one should be using the term "capacitance" without a qualifier.

Note also that none of these capacitances is equal to any element of the matrix reciprocal to the matrix (23):

$$\begin{pmatrix} \mathcal{P}_1 & \mathcal{P} \\ \mathcal{P} & \mathcal{P}_2 \end{pmatrix}^{-1} = \frac{1}{\mathcal{P}^2 - \mathcal{P}_1\mathcal{P}_2}\begin{pmatrix} -\mathcal{P}_2 & \mathcal{P} \\ \mathcal{P} & -\mathcal{P}_1 \end{pmatrix}. \tag{2.34}$$

Because of this reason, this *physical capacitance matrix*, which expresses the vector of conductor charges via the vector of their potentials, is less convenient for most applications than the reciprocal capacitance matrix (23). The same conclusion is valid for multi-conductor systems, which are most conveniently characterized by an evident generalization of Eq. (19). Indeed, in this case, even the mutual capacitance between two selected conductors may depend on the electrostatic conditions of other components of the system.

Logically, at this point I would need to discuss the particular, but practically very important case when the regions where the electric field between each pair of conductors is most significant do not overlap – such as in the example shown in Fig. 5a. In this case, the system's properties may be discussed using the *equivalent-circuit* language, representing each such region as a *lumped* (localized) *capacitor*, with a certain mutual capacitance $C$, and the whole system as some connection of these capacitors by conducting "wires", whose length and geometry are not important – see Fig. 5b.

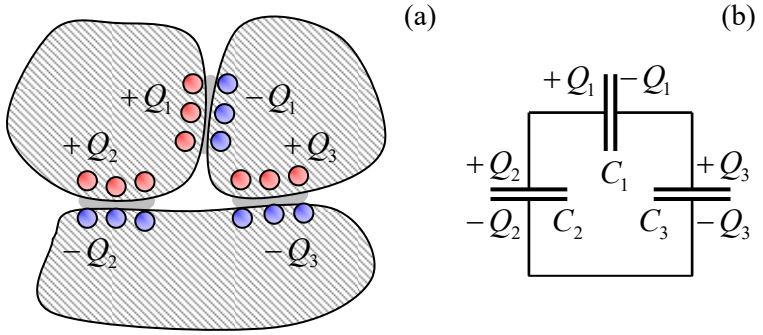



Fig. 2.5. (a) A simple system of conductors, with three well-localized regions of high electric field (and hence surface charge) concentration, and (b) its representation with an equivalent circuit of three lumped capacitors.

Since the analysis of such equivalent circuits is covered in typical introductory physics courses, I will save time by skipping their discussion. However, since such circuits are very frequently met in physical experiment and electrical engineering practice, I would urge the reader to self-test their understanding of this topic by solving a couple of problems offered at the end of this chapter,[20] and if their solution presents any difficulty, review the corresponding section in an undergraduate textbook.

[20] These problems have been selected to emphasize the fact that not every circuit may be reduced to the simplest connections of the component capacitors and/or their groups in parallel and/or in series.

## 2.3. The simplest boundary problems

In the general case when the electric field distribution in the free space between the conductors cannot be easily found from the Gauss law or a particular symmetry, the best approach is to try to solve the differential Laplace equation (1.42), with the boundary conditions (1b):

$$\nabla^2\phi = 0, \qquad \phi\Big|_{S_k} = \phi_k\,, \tag{2.35}$$

Typical boundary problem

where $S_k$ is the surface of the $k^{\text{th}}$ conductor of the system. After this *boundary problem* has been solved, i.e. the spatial distribution $\phi(\mathbf{r})$ has been found at all points outside the conductors, it is straightforward to use Eq. (3) to find the surface charge density, and finally the total charge

$$Q_k = \oint_{S_k} \sigma d^2 r \tag{2.36}$$

of each conductor, and hence any component of the reciprocal capacitance matrix. As an illustration, let us implement this program for three very simple problems.

(i) Plane capacitor (Fig. 3). In this case, the easiest way to solve the Laplace equation is to use the linear (Cartesian) coordinates with one axis (say, $z$) normal to the conductor surfaces – see Fig. 6.

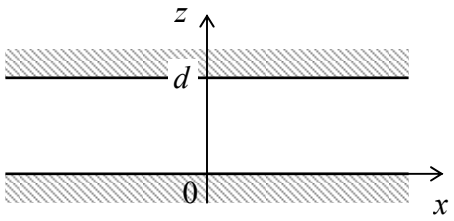



Fig. 2.6. The plane capacitor as the system for the simplest illustration of the boundary problem (35) and its solution.

In these coordinates, the Laplace operator is just the sum of three second derivatives.[21] It is evident that due to the problem's translational symmetry within the [$x$, $y$] plane, deep inside the gap (i.e. at any lateral distance from the edges much larger than $d$) the electrostatic potential may only depend on the coordinate normal to the gap surfaces: $\phi(\mathbf{r}) = \phi(z)$. For such a function, the derivatives over $x$ and $y$ vanish, and the boundary problem (35) is reduced to a very simple ordinary differential equation

$$\frac{d^2\phi}{dz^2}(z) = 0, \tag{2.37}$$

with boundary conditions

$$\phi(0) = 0, \quad \phi(d) = V. \tag{2.38}$$

(For the sake of notation simplicity, I have used the discretion of adding a constant to the potential, to make one of the potentials vanish, and also the definition (25) of the voltage $V$.) The general solution of Eq. (37) is a linear function: $\phi(z) = c_1 z + c_2$, whose constant coefficients $c_{1,2}$ may be readily found from the boundary conditions (38). The final solution is

$$\phi = V\frac{z}{d}. \tag{2.39}$$

[21] See, e.g. MA Eq. (9.1).

From here the only nonzero component of the electric field is

$$E_z = -\frac{d\phi}{dz} = -\frac{V}{d}, \tag{2.40}$$

and the surface charge of the capacitor plates is

$$\sigma = \varepsilon_0 E_n = \mp\varepsilon_0 E_z = \pm\varepsilon_0 \frac{V}{d}, \tag{2.41}$$

where the upper and lower signs correspond to the upper and lower plates, respectively. Since $\sigma$ does not depend on $x$ and $y$, we can get the full charges $Q_1 = -Q_2 \equiv Q$ of the surfaces by its multiplication by the gap area $A$, giving us again the already obtained result (28) for the mutual capacitance $C \equiv Q/V$. I believe that this calculation, though very easy, may serve as a good illustration of the boundary problem solution approach, which will be used below for more complex cases.

(ii) Coaxial-cable capacitor. *Coaxial cable* is a system of two round cylindrical, coaxial conductors, with the cross-section shown in Fig. 7.

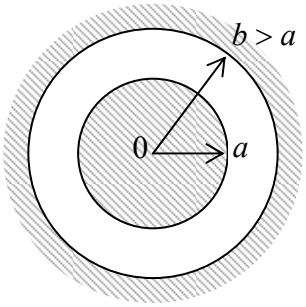



Fig. 2.7. The cross-section of a coaxial cable.

Evidently, in this case, the cylindrical coordinates $\{\rho, \varphi, z\}$, with the $z$-axis coinciding with the common axis of the cylinders, are most convenient.[22] Due to the axial symmetry of the problem, in these coordinates $\mathbf{E}(\mathbf{r}) = \mathbf{n}_\rho E(\rho)$, $\phi(\mathbf{r}) = \phi(\rho)$, so in the general expression for the Laplace operator[23] we may take $\partial/\partial\varphi = \partial/\partial z = 0$. As a result, only the radial term of the operator survives, and the boundary problem (35) takes the form

$$\frac{1}{\rho}\frac{d}{d\rho}\left(\rho\frac{d\phi}{d\rho}\right) = 0, \qquad \phi(a) = V, \qquad \phi(b) = 0. \tag{2.42}$$

The sequential double integration of this ordinary linear differential equation is elementary (and similar to that of the Poisson equation in spherical coordinates, carried out in Sec. 1.3), giving

$$\rho\frac{d\phi}{d\rho} = c_1, \qquad \phi(\rho) = c_1\int_a^\rho \frac{d\rho''}{\rho''} + c_2 = c_1 \ln\frac{\rho}{a} + c_2. \tag{2.43}$$

The constants $c_{1,2}$ may be found using boundary conditions (42):

[22] I am sorry for using, for the 2D radius, the same letter $\rho$ as for the volumic density of charge. (Both notations are too common to refuse.) I do not believe this may lead to confusion, because the letter will not be used in two different meanings during any particular discussion.

[23] See, e.g., MA Eq. (10.3).

$$c_2 = V, \qquad c_1 \ln\frac{b}{a} + c_2 = 0, \tag{2.44}$$

giving $c_1 = -V/\ln(b/a)$, so Eq. (43) takes the following form:

$$\phi(\rho) = V\left[1 - \frac{\ln(\rho/a)}{\ln(b/a)}\right]. \tag{2.45}$$

Next, for our axial symmetry, the general expression for the gradient of a scalar function is reduced to its radial derivative, so

$$E(\rho) \equiv -\frac{d\phi(\rho)}{d\rho} = \frac{V}{\rho \ln(b/a)}. \tag{2.46}$$

This expression, plugged into Eq. (2), allows us to find the density of the conductors' surface charge. For example, for the inner electrode

$$\sigma_a = \varepsilon_0 E(a) = \frac{\varepsilon_0 V}{a \ln(b/a)}, \tag{2.47}$$

so its full charge (per unit length of the system) is

$$\frac{Q}{l} = 2\pi a \sigma_a = \frac{2\pi\varepsilon_0 V}{\ln(b/a)}. \tag{2.48}$$

(It is straightforward to check that the charge of the outer electrode is equal and opposite.) Hence, by the definition of the mutual capacitance, its value per unit length is

$$\boxed{\frac{C}{l} \equiv \frac{Q}{lV} = \frac{2\pi\varepsilon_0}{\ln(b/a)}}. \tag{2.49}$$

C: Coaxial cable

This expression shows that the total capacitance $C$ is proportional to the systems length $l$ (if $l >> a, b$), while being only logarithmically dependent on is the dimensions of its cross-section. Since the logarithm of a large argument is an extremely slow function (sometimes called a *quasi-constant*), if the external conductor is made very large ($b >> a$), the capacitance diverges, but very weakly. Such *logarithmic divergence* may be cut by any minuscule additional effect, for example by the finite length $l$ of the system. This fact yields the following very useful estimate of the self-capacitance of a *single* round wire of radius $a$:

$$C \approx \frac{2\pi\varepsilon_0 l}{\ln(l/a)}, \qquad \text{for } l >> a. \tag{2.50}$$

On the other hand, if the gap $d$ between the conductors is very narrow: $d \equiv b - a << a$, then $\ln(b/a) \equiv \ln(1 + d/a)$ may be approximated as $d/a$, and Eq. (49) is reduced to $C \approx 2\pi\varepsilon_0 al/d$, i.e. to Eq. (28) for the plane capacitor, of the appropriate area $A = 2\pi al$.

(iii) Spherical capacitor. This is a system of two conductors, with a *central* cross-section similar to that of the coaxial cable (Fig. 7), but now with spherical rather than axial symmetry. This symmetry implies that we may be better off using spherical coordinates, so the potential $\phi$ depends only on one of them: the distance $r$ from the common center of the conductors: $\phi(\mathbf{r}) = \phi(r)$. As we already know from Sec. 1.3, in this case the general expression for the Laplace operator is reduced to its first (radial) term,

so the Laplace equation takes the simple form (1.47). Moreover, we have already found the general solution of this equation – see Eq. (1.50):

$$\phi(r) = \frac{c_1}{r} + c_2, \tag{2.51}$$

Now acting exactly as above, i.e. determining the (only essential) constant $c_1$ from the boundary condition $\phi(a) - \phi(b) = V$, we get

$$c_1 = V\left(\frac{1}{a} - \frac{1}{b}\right)^{-1}, \qquad \text{so that} \quad \phi(r) = \frac{V}{r}\left(\frac{1}{a} - \frac{1}{b}\right)^{-1} + c_2. \tag{2.52}$$

Next, we can use the spherical symmetry to find the electric field, $\mathbf{E}(\mathbf{r}) = \mathbf{n}_r E(r)$, with

$$E(r) = -\frac{d\phi}{dr} = \frac{V}{r^2}\left(\frac{1}{a} - \frac{1}{b}\right)^{-1}, \tag{2.53}$$

and hence its values on conductors' surfaces, and then the surface charge density $\sigma$ from Eq. (3). For example, for the inner conductor's surface,

$$\sigma_a = \varepsilon_0 E(a) = \varepsilon_0 \frac{V}{a^2}\left(\frac{1}{a} - \frac{1}{b}\right)^{-1}, \tag{2.54}$$

so, finally, for the full charge of that conductor, we get the following result:

$$Q = 4\pi a^2 \sigma = 4\pi\varepsilon_0\left(\frac{1}{a} - \frac{1}{b}\right)^{-1} V. \tag{2.55}$$

(Again, the charge of the outer conductor is equal and opposite.) Now we can use the definition (26) of the mutual capacitance to get the final result:

C: Spherical capacitor

$$C \equiv \frac{Q}{V} = 4\pi\varepsilon_0\left(\frac{1}{a} - \frac{1}{b}\right)^{-1} \equiv 4\pi\varepsilon_0 \frac{ab}{b-a}. \tag{2.56}$$

For $b >> a$, it coincides with Eq. (17) for the self-capacitance of the inner conductor. On the other hand, if the gap $d$ between two conductors is narrow, $d \equiv b - a << a$, then

$$C = 4\pi\varepsilon_0 \frac{a(a+d)}{d} \approx 4\pi\varepsilon_0 \frac{a^2}{d}, \tag{2.57}$$

i.e. the capacitance approaches that of the planar capacitor of the area $A = 4\pi a^2$ – as it should.

All this seems (and indeed is) very straightforward, but let us contemplate what was the reason for such easy successes. In each of the cases (i)-(iii) we have managed to find such coordinates that both the Laplace equation and the boundary conditions involved only one of them. The necessary condition for the former fact is for the coordinates to be *orthogonal*. This means that the three vector components of the local differential $d\mathbf{r}$, due to small variations of the new coordinates (say, $dr$, $d\theta$, and $d\varphi$ for the spherical coordinates), are mutually perpendicular.

## 2.4. Using other orthogonal coordinates

The cylindrical and spherical coordinates used above are only the simplest examples of the *curvilinear orthogonal* (or just "orthogonal") coordinates, and that approach may be extended to other

coordinate systems of this type. As an example, let us calculate the self-capacitance of a thin, round conducting disk. The cylindrical or spherical coordinates would not give much help here, because while they have the appropriate axial symmetry, they would make the boundary condition on the disk too complicated: involving two coordinates, either $\rho$ and $z$, or $r$ and $\theta$. Help comes from noting that the flat disk, i.e. the area with $z = 0$, $r < R$, may be viewed as the limiting case of an *axially-symmetric ellipsoid* (or "degenerate ellipsoid", or "ellipsoid of rotation", or "spheroid") – the surface formed by rotation of the usual ellipse about one of its major axes – which would be also the symmetry axis of the disk – in Fig. 8, the $z$-axis.

–

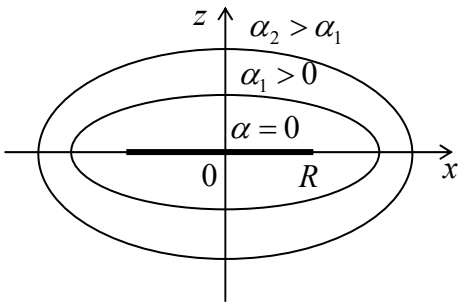



Fig. 2.8. Solving the disk's capacitance problem. (The cross-section of the system by the vertical plane $y = 0$.)

Analytically, this ellipsoid may be described by the following equation:

$$\frac{x^2 + y^2}{a^2} + \frac{z^2}{b^2} = 1, \tag{2.58}$$

where $a$ and $b$ are the so-called *major semi-axes*, whose ratio determines the ellipse's *eccentricity* – the degree of its "squeezing". For our problem, we will only need *oblate* ellipsoids with $a \geq b$; according to Eq. (58), they may be represented as surfaces of constant $\alpha$ in the *oblate spheroidal* (also called "degenerate ellipsoidal") *coordinates* $\{\alpha, \beta, \varphi\}$ that are related to the Cartesian coordinates as follows:[24]

$$\begin{aligned} x &= R\cosh\alpha\sin\beta\cos\varphi, & 0 \leq \alpha < \infty, \\ y &= R\cosh\alpha\sin\beta\sin\varphi, \quad \text{with} & 0 \leq \beta \leq \pi, \\ z &= R\sinh\alpha\cos\beta, & 0 \leq \varphi \leq 2\pi. \end{aligned} \tag{2.59}$$

Such spheroidal coordinates are an evident generalization of the spherical coordinates, which correspond to the limit $\alpha >> 1$ (i.e. $r >> R$). In the opposite limit, the surface of constant $\alpha = 0$ describes our thin disk of radius $R$, with the coordinate $\beta$ describing the distance $\rho \equiv (x^2 + y^2)^{1/2} = R\sin\beta$ of its point from the $z$-axis. It is almost evident (and easy to prove) that the curvilinear coordinates (59) are also orthogonal; the Laplace operator in them is:

$$\nabla^2 = \frac{1}{R^2(\cosh^2\alpha - \sin^2\beta)} \times \left[ \begin{aligned} &\frac{1}{\cosh\alpha}\frac{\partial}{\partial\alpha}\left(\cosh\alpha\frac{\partial}{\partial\alpha}\right) \\ &+ \frac{1}{\sin\beta}\frac{\partial}{\partial\beta}\left(\sin\beta\frac{\partial}{\partial\beta}\right) + \left(\frac{1}{\sin^2\beta} - \frac{1}{\cosh^2\alpha}\right)\frac{\partial^2}{\partial\varphi^2} \end{aligned} \right]. \tag{2.60}$$

Though this expression may look a bit intimidating, let us notice that since in our current problem, the boundary conditions depend only on $\alpha$:[25]

[24] For solution of some problems, it is convenient to use Eqs. (59) with $-\infty < \alpha < +\infty$ and $0 \leq \beta \leq \pi/2$.

[25] I have called the disk's potential $V$, to distinguish it from the potential $\phi$ at an arbitrary point of space.

$$\phi\big|_{\alpha=0} = V, \qquad \phi\big|_{\alpha=\infty} = 0, \tag{2.61}$$

there is every reason to assume that the electrostatic potential in all space is a function of $\alpha$ alone; in other words, that all ellipsoids $\alpha$ = const are the equipotential surfaces. Indeed, acting on such a function $\phi(\alpha)$ by the Laplace operator (60), we see that the two last terms in the square brackets vanish, and the Laplace equation (35) is reduced to a simple ordinary differential equation

$$\frac{d}{d\alpha}\left[\cosh\alpha \frac{d\phi}{d\alpha}\right] = 0. \tag{2.62}$$

Integrating it twice, just as we did in the three previous problems, we get

$$\phi(\alpha) = c_1 \int \frac{d\alpha}{\cosh\alpha}. \tag{2.63}$$

This integral may be readily worked out using the substitution $\xi \equiv \sinh\alpha$ (which gives $d\xi \equiv \cosh\alpha\, d\alpha$, i.e. $d\alpha = d\xi/\cos\alpha$, and $\cosh^2\alpha = 1 + \sinh^2\alpha \equiv 1 + \xi^2$):

$$\phi(\alpha) = c_1 \int_0^{\sinh\alpha} \frac{d\xi}{1+\xi^2} + c_2 = c_1 \tan^{-1}(\sinh\alpha) + c_2. \tag{2.64}$$

The integration constants $c_{1,2}$ may be simply found from the boundary conditions (61), and we arrive at the following final expression for the electrostatic potential:

$$\phi(\alpha) = V\left[1 - \frac{2}{\pi}\tan^{-1}(\sinh\alpha)\right] \equiv \frac{2V}{\pi}\tan^{-1}\left(\frac{1}{\sinh\alpha}\right). \tag{2.65}$$

This solution satisfies both the Laplace equation and the boundary conditions. Mathematicians tell us that the solution of any boundary problem of the type (35) is *unique*, so we do not need to look any further.

Now we may use Eq. (3) to find the surface density of electric charge, but in the case of a thin disk, it is more natural to add up such densities on its top and bottom surfaces at the same distance $\rho = (x^2 + y^2)^{1/2}$ from the disk's center. The densities are evidently equal, due to the problem symmetry about the plane $z = 0$, so the total density is $\sigma = 2\varepsilon_0 E_n|_{z=+0}$. According to Eq. (65), and the last of Eqs. (59), the electric field on the upper surface is

$$E_n\big|_{z=+0} = -\frac{\partial\phi}{\partial z}\Big|_{z=+0} = -\frac{\partial\phi(\alpha)}{\partial(R\sinh\alpha\cos\beta)}\Big|_{\alpha=0} = \frac{2}{\pi}V\frac{1}{R\cos\beta} = \frac{2}{\pi}V\frac{1}{(R^2-\rho^2)^{1/2}}, \tag{2.66}$$

and we see that the charge is distributed over the disk very nonuniformly:

$$\sigma = \frac{4}{\pi}\varepsilon_0 V \frac{1}{(R^2-\rho^2)^{1/2}}, \tag{2.67}$$

with a singularity at the disk edge. Below we will see that such singularities are very typical for sharp edges of conductors. Fortunately, in our current case the divergence is integrable, giving a finite disk charge:

$$Q = \int_{\substack{\text{disk}\\ \text{surface}}} \sigma d^2\rho = \int_0^R \sigma(\rho) 2\pi\rho d\rho = \frac{4}{\pi}\varepsilon_0 V \int_0^R \frac{2\pi\rho d\rho}{(R^2-\rho^2)^{1/2}} = 4\varepsilon_0 VR \int_0^1 \frac{d\xi}{(1-\xi)^{1/2}} = 8\varepsilon_0 RV. \qquad (2.68)$$

Thus, for the disk's self-capacitance we get a very simple result,

$$C = 8\varepsilon_0 R \equiv \frac{2}{\pi} 4\pi\varepsilon_0 R, \qquad (2.69)$$

a factor of $\pi/2 \approx 1.57$ lower than that for the conducting sphere of the same radius, but still complying with the general estimate (18).

Can we always find such a "good" system of orthogonal coordinates? Unfortunately, the answer is *no*, even for highly symmetric geometries. This is why the practical value of this approach is limited, and other, more general methods of boundary problem solution are clearly needed. Before proceeding to their discussion, however, let me note that in the case of 2D problems (i.e. cylindrical geometries[26]), the orthogonal coordinate method gets much help from the following *conformal mapping* approach.

Let us consider a pair of Cartesian coordinates $\{x, y\}$ of the cylinder's cross-section plane as a complex variable $\boldsymbol{z} \equiv x + iy$,[27] where $i$ is the imaginary unit ($i^2 = -1$), and let $\boldsymbol{w}(\boldsymbol{z}) = u + iv$ be an *analytic complex function* of $\boldsymbol{z}$.[28] For our current purposes, the most important property of an analytic function is that its real and imaginary parts obey the following *Cauchy-Riemann relations*:[29]

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \qquad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}. \qquad (2.70)$$

For example, for the function

$$\boldsymbol{w} = \boldsymbol{z}^2 \equiv (x+iy)^2 \equiv (x^2 - y^2) + 2ixy, \qquad (2.71)$$

whose real and imaginary parts are

$$u \equiv \mathrm{Re}\,\boldsymbol{w} = x^2 - y^2, \qquad v \equiv \mathrm{Im}\,\boldsymbol{w} = 2xy, \qquad (2.72)$$

we immediately see that $\partial u/\partial x = 2x = \partial v/\partial y$, and $\partial v/\partial x = 2y = -\partial u/\partial y$, in accordance with Eq. (70).

Let us differentiate the first of Eqs. (70) over $x$ again, then change the order of differentiation, and after that use the latter of those equations:

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial}{\partial x}\frac{\partial u}{\partial x} = \frac{\partial}{\partial x}\frac{\partial v}{\partial y} = \frac{\partial}{\partial y}\frac{\partial v}{\partial x} = -\frac{\partial}{\partial y}\frac{\partial u}{\partial y} = -\frac{\partial^2 u}{\partial y^2}, \qquad (2.73)$$

---

[26] Let me remind the reader that the term *cylindrical* describes any surface formed by a translation, along a straight line, of an *arbitrary* curve, and hence more general than the usual circular cylinder. (In this terminology, for example, a prism is also a cylinder of a particular type, formed by a translation of a polygon.)

[27] The complex variable $\boldsymbol{z}$ should not be confused with the (real) 3rd spatial coordinate $z$! We are considering 2D problems now, with the potential independent of $z$.

[28] An analytic (or "holomorphic") function may be defined as one that may be expanded into the Taylor series in its complex argument, i.e. is infinitely differentiable in the given point. (Almost all "regular" functions, such as $\boldsymbol{z}^n$, $\boldsymbol{z}^{1/n}$, $\exp \boldsymbol{z}$, $\ln \boldsymbol{z}$, etc., and their linear combinations are analytic at all $\boldsymbol{z}$, maybe besides certain special points.) If the reader needs to brush up on their background on this subject, I can recommend a popular textbook by M. Spiegel *et al*., *Complex Variables*, 2nd ed., McGraw-Hill, 2009.

[29] These relations may be used, in particular, to prove the Cauchy integral formula – see, e.g., MA Eq. (15.1).

and similarly for $v$. This means that the sum of second-order partial derivatives of each of the real functions $u(x, y)$ and $v(x, y)$ is zero, i.e. that both functions obey the 2D Laplace equation. This mathematical fact opens a nice way of solving problems of electrostatics for (relatively simple) 2D geometries. Imagine that for a particular boundary problem we have found a function $\boldsymbol{w}(\boldsymbol{z})$ for that either $u(x, y)$ or $v(x, y)$ is constant on all electrode surfaces. Then all lines of constant $u$ (or $v$) represent equipotential surfaces, i.e. the problem of the potential distribution has been essentially solved.

As a simple example, let us consider a problem important for practice: the *quadrupole electrostatic lens* – a system of four cylindrical electrodes with hyperbolic cross-sections, whose boundaries are described by the following relations:

$$x^2 - y^2 = \begin{cases} +a^2, & \text{for the left and right electrodes,} \\ -a^2, & \text{for the top and bottom electrodes,} \end{cases} \tag{2.74}$$

voltage-biased as shown in Fig. 9a.

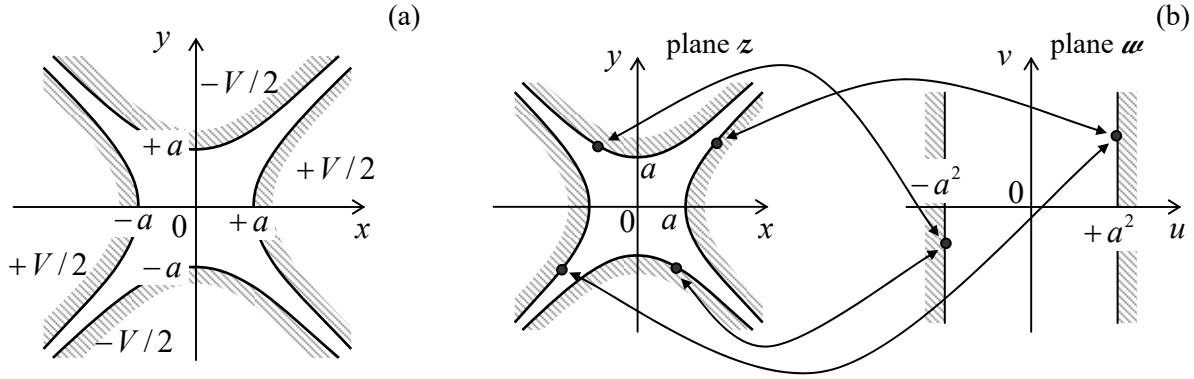



Fig. 2.9. (a) The quadrupole electrostatic lens' cross-section and (b) its conformal mapping.

Comparing these relations with Eqs. (72), we see that each electrode surface corresponds to a constant value of the real part $u(x, y)$ of the function given by Eq. (71): $u = \pm a^2$. Moreover, the potentials of both surfaces with $u = +a^2$ are equal to $+V/2$, while those with $u = -a^2$ are equal to $-V/2$. Hence we may conjecture that the electrostatic potential at each point is a function of $u$ alone; moreover, a simple linear function,

$$\phi = c_1 u + c_2 \equiv c_1 (x^2 - y^2) + c_2 , \tag{2.75}$$

is a valid (and hence the unique) solution of our boundary problem. Indeed, it does satisfy the Laplace equation, while the constants $c_{1,2}$ may be readily selected in a way to satisfy all the boundary conditions shown in Fig. 9a:

$$\phi = \frac{V}{2} \frac{x^2 - y^2}{a^2} . \tag{2.76}$$

so the boundary problem has been solved.

According to Eq. (76), all equipotential surfaces are hyperbolic cylinders, similar to those of the electrode surfaces. What remains is to find the electric field at an arbitrary point inside the system:

$$E_x = -\frac{\partial \phi}{\partial x} = -V \frac{x}{a^2}, \quad E_y = -\frac{\partial \phi}{\partial y} = V \frac{y}{a^2}. \tag{2.77}$$

These formulas show, in particular, that if charged particles (e.g., electrons in an electron-optics system) are launched to fly ballistically through such a lens, along the $z$-axis, they experience a force pushing them toward the symmetry axis and proportional to the particle's deviation from the axis (and thus equivalent in action to an optical lens with a positive refraction power) in one direction, and a force pushing them out (negative refractive power) in the perpendicular direction. One can show that letting the particles fly through several such lenses, with alternating voltage polarities, in series, enables beam focusing.[30]

Hence, we have reduced the 2D Laplace boundary problem to that of finding the proper analytic function $\boldsymbol{w}(\boldsymbol{z})$. This task may be also understood as that of finding a *conformal map*, i.e. a correspondence between components of any point pair, $\{x, y\}$ and $\{u, v\}$, residing, respectively, on the initial Cartesian plane $\boldsymbol{z}$ and the plane $\boldsymbol{w}$ of the new variables. For example, Eq. (71) maps the real electrode configuration onto a plane capacitor of an infinite area (Fig. 9b), and the simplicity of Eq. (75) is due to the fact that for the latter system the equipotential surfaces are just parallel planes $u$ = const.

For more complex geometries, the suitable analytic function $\boldsymbol{w}(\boldsymbol{z})$ may be hard to find. However, for conductors with piece-linear cross-section boundaries, substantial help may be obtained from the following *Schwarz-Christoffel integral*

$$\boldsymbol{w}(\boldsymbol{z}) = \text{const} \times \int \frac{d\boldsymbol{z}}{(\boldsymbol{z}-x_1)^{k_1}(\boldsymbol{z}-x_2)^{k_2}...(\boldsymbol{z}-x_{N-1})^{k_{N-1}}}. \qquad (2.78)$$

that provides a conformal mapping of the interior of an *arbitrary* $N$-sided polygon onto the plane $\boldsymbol{w} = u + iv$, onto the upper half ($y > 0$) of the plane $\boldsymbol{z} = x + iy$. In Eq. (78), $x_j$ ($j = 1, 2, N - 1$) are the points of the $y = 0$ axis (i.e., of the boundary of the mapped region on plane $\boldsymbol{z}$) to which the corresponding polygon vertices are mapped, while $k_j$ are the exterior angles at the polygon vertices, measured in the units of $\pi$, with $-1 \le k_j \le +1$ – see Fig. 10.[31] Of the points $x_j$, two may be selected arbitrarily (because their effects may be compensated by the multiplicative constant in Eq. (78), and the additive constant of integration), while all the others have to be adjusted to provide the correct mapping.

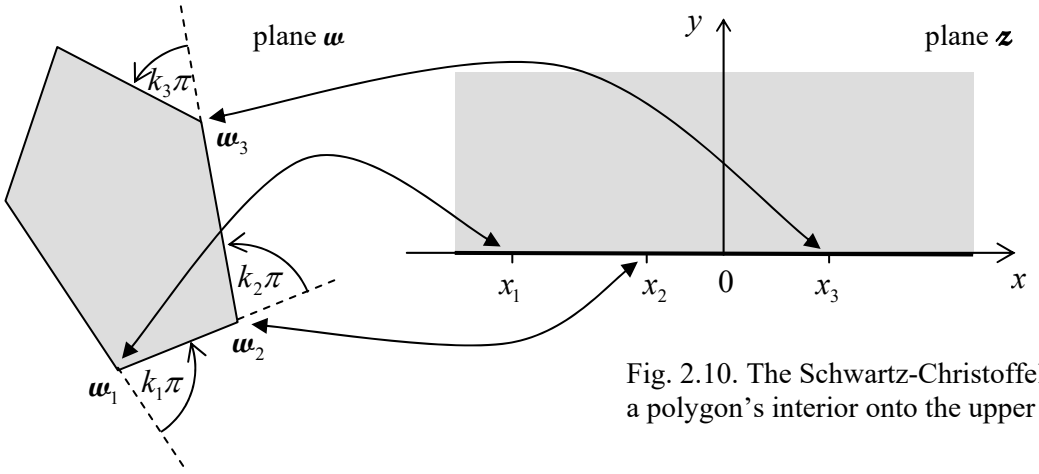



Fig. 2.10. The Schwartz-Christoffel mapping of a polygon's interior onto the upper half-plane.

[30] See, e.g., textbook by P. Grivet, *Electron Optics*, 2nd ed., Pergamon, 1972.

[31] The integral (78) includes only $(N - 1)$ rather than $N$ poles because a polygon's shape is fully determined by $(N - 1)$ positions $\boldsymbol{w}_j$ of its vertices and $(N - 1)$ angles $\pi k_j$. In particular, since the algebraic sum of all external angles of a polygon equals $2\pi$, the last angle parameter $k_j = k_N$ is uniquely determined by the set of the previous ones.

In the general case, the complex integral (78) may be hard to tackle. However, in some important cases, in particular those with right angles ($k_j = \pm ½$) and/or with some points $\boldsymbol{w}_j$ at infinity, the integrals may be readily worked out, giving explicit analytical expressions for the mapping functions $\boldsymbol{w}(\boldsymbol{z})$. For example, let us consider a semi-infinite strip defined by restrictions $-1 \le u \le +1$ and $0 \le v$, on the $\boldsymbol{w}$-plane – see the left panel of Fig. 11.

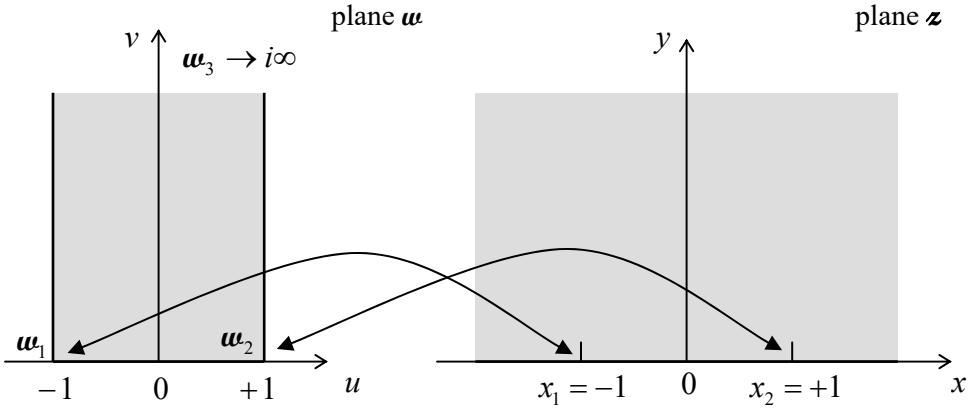



Fig. 2.11. A semi-infinite strip mapped onto the upper half-plane.

The strip may be considered as a triangle, with one vertex at the infinitely distant vertical point $\boldsymbol{w}_3 = 0 + i\infty$. Let us map the polygon onto the upper half of plane $\boldsymbol{z}$, shown on the right panel of Fig. 11, with the vertex $\boldsymbol{w}_1 = -1 + i\,0$ mapped onto the point $\boldsymbol{z}_1 = -1 + i\,0$, and the vertex $\boldsymbol{w}_2 = +1 + i\,0$ mapped onto the point $\boldsymbol{z}_2 = +1 + i\,0$. Since the external angles at these vertices are equal to $+\pi/2$, and hence $k_1 = k_2 = +½$, Eq. (78) is reduced to

$$\boldsymbol{w}(\boldsymbol{z}) = \text{const} \times \int \frac{d\boldsymbol{z}}{(\boldsymbol{z}+1)^{1/2}(\boldsymbol{z}-1)^{1/2}} \equiv \text{const} \times \int \frac{d\boldsymbol{z}}{(\boldsymbol{z}^2-1)^{1/2}} \equiv \text{const} \times i \int \frac{d\boldsymbol{z}}{(1-\boldsymbol{z}^2)^{1/2}}. \qquad (2.79)$$

This complex integral may be worked out, just as for real $\boldsymbol{z}$, with the substitution $\boldsymbol{z} = \sin\xi$, giving

$$w(\boldsymbol{z}) = \text{const}' \times \int^{\sin^{-1}\boldsymbol{z}} d\xi = c_1 \sin^{-1}\boldsymbol{z} + c_2. \qquad (2.80)$$

Determining the constants $c_{1,2}$ from the required mapping, i.e. from the conditions $\boldsymbol{w}(-1 + i\,0) = -1 + i\,0$ and $\boldsymbol{w}(+1+ i\,0) = +1 + i\,0$ (see the arrows in Fig. 11), we finally get[32]

$$\boldsymbol{w}(\boldsymbol{z}) = \frac{2}{\pi}\sin^{-1}\boldsymbol{z}, \quad \text{i.e. } \boldsymbol{z} = \sin\frac{\pi\boldsymbol{w}}{2}. \qquad (2.81a)$$

Using the well-known expression for the sine of a complex argument,[33] we may rewrite this elegant result in either of the following two forms for the real and imaginary components of $\boldsymbol{z}$ and $\boldsymbol{w}$:

$$u = \frac{2}{\pi}\sin^{-1}\frac{2x}{\left[(x+1)^2+y^2\right]^{1/2}+\left[(x-1)^2+y^2\right]^{1/2}}, \quad v = \frac{2}{\pi}\cosh^{-1}\frac{\left[(x+1)^2+y^2\right]^{1/2}+\left[(x-1)^2+y^2\right]^{1/2}}{2},$$

[32] Note that this function differs only by a linear transformation of variables from the function $\boldsymbol{z} = c\cosh\boldsymbol{w}$, which is the canonical form of the definition of the so-called *elliptic* (not ellipsoidal!) orthogonal coordinates.

[33] See, e.g., MA Eq. (3.5).

$$x = \sin\frac{\pi u}{2}\cosh\frac{\pi v}{2}, \qquad y = \cos\frac{\pi u}{2}\sinh\frac{\pi v}{2}. \tag{2.81b}$$

It is amazing how perfectly the last formula manages to keep $y \equiv 0$ at the different borders of our $\boldsymbol{w}$-region (Fig. 11): at its side borders ($u = \pm 1$, $0 \le v < \infty$), this is performed by the first multiplier, while at the bottom border ($-1 \le u \le +1$, $v = 0$), the equality is enforced by the second multiplier.

This mapping may be used to solve several electrostatics problems with the geometry shown in Fig. 11a; probably the most surprising of them is the following one. A straight gap of width $2t$ is cut in a very thin conducting plane, and voltage $V$ is applied between the resulting half-planes – see the bold straight lines in Fig. 12.

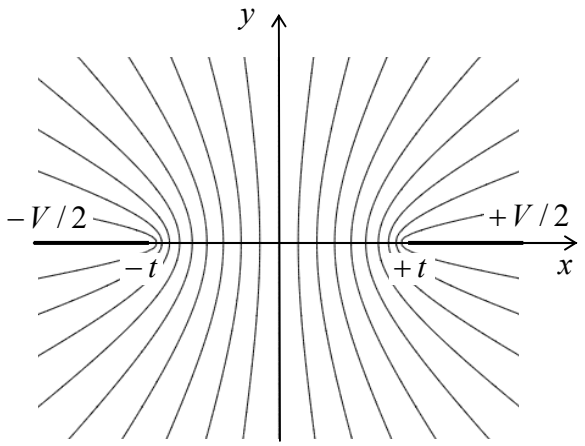



Fig. 2.12. The equipotential surfaces of the electric field between two thin conducting semi-planes (or rather their cross-sections by the plane $z$ = const).

Selecting a Cartesian coordinate system with the $z$-axis directed along the cut, the $y$-axis normal to the plane, and the origin in the middle of the cut (Fig. 12), we can write the boundary conditions of this Laplace problem as

$$\phi = \begin{cases} +V/2, & \text{for } x > +t,\ y = 0, \\ -V/2, & \text{for } x < -t,\ y = 0. \end{cases} \tag{2.82}$$

(Due to the problem's symmetry, we may expect that in the middle of the gap, i.e. at $-t < x < +t$ and $y = 0$, the electric field is parallel to the plane and hence $\partial\phi/\partial y = 0$.) The comparison of Figs. 11 and 12 shows that if we normalize our coordinates $\{x, y\}$ to $t$, Eqs. (81) provide the conformal mapping of our system onto the plane $\boldsymbol{z}$ to a plane capacitor on the plane $\boldsymbol{w}$, with the voltage $V$ between two conducting planes located at $u = \pm 1$. Since we already know that in that case $\phi = (V/2)u$, we may immediately use the first of Eqs. (81b) to write the final solution of the problem:[34]

$$\phi = \frac{V}{2}u = \frac{V}{\pi}\sin^{-1}\frac{2x}{\left[(x+t)^2 + y^2\right]^{1/2} + \left[(x-t)^2 + y^2\right]^{1/2}}. \tag{2.83}$$

The thin lines in Fig. 12 show the corresponding equipotential surfaces;[35] it is evident that the electric field concentrates at the gap edges, just as it did at the edge of the thin disk (Fig. 8). Let me

[34] This result may be also obtained by the Green's function method, to be discussed in Sec. 10 below.

[35] Another graphical representation of the electric field distribution, by *field lines*, is less convenient. (It is more useful for the magnetic field, which may be represented by a scalar potential only in particular cases, so there is no surprise that the field lines were introduced only by Michael Faraday in the 1830s.) As a reminder, the field

leave the remaining calculation of the surface charge distribution and the mutual capacitance between the half-planes (per unit length of the system in the $z$-direction) for the reader's exercise.

## 2.5. Variable separation – Cartesian coordinates

The general approach of the methods discussed in the last two sections was to satisfy the Laplace equation by a function of a single variable that also satisfies the boundary conditions. Unfortunately, in many cases this cannot be done – at least, using reasonably simple functions. In this case, a very powerful method called the *variable separation*,[36] may work, typically producing "semi-analytical" results in the form of series (infinite sums) of either elementary or well-studied special functions. Its main idea is to look for the solution of the boundary problem (35) as the sum of partial solutions,

$$\phi = \sum_k c_k \phi_k \ , \tag{2.84}$$

where each function $\phi_k$ satisfies the Laplace equation, and then select the set of coefficients $c_k$ to satisfy the boundary conditions. More specifically, in the variable separation method, the partial solutions $\phi_k$ are looked for in the form of a product of functions, each depending on just one spatial coordinate.

Let us discuss this approach on the classical example of a rectangular box with conducting walls (Fig. 13), with the same potential (that I will take for zero) at all its sidewalls and the lower lid, but a different potential $V$ at the top lid ($z = c$). Moreover, to demonstrate the power of the variable separation method, let us carry out all the calculations for a more general case when the top lid's potential is an arbitrary 2D function $V(x, y)$.[37]

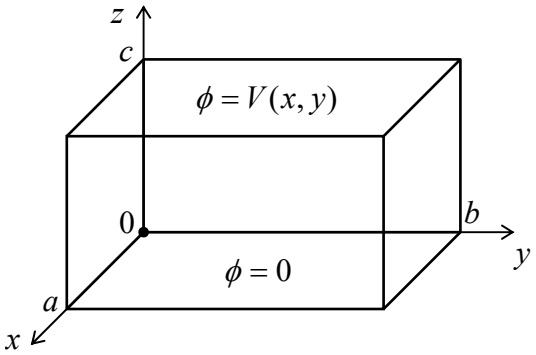



Fig. 2.13. The standard playground for the variable separation discussion: a rectangular box with five conducting, grounded walls and a fixed potential distribution $V(x, y)$ on the top lid.

For this geometry, it is natural to use the Cartesian coordinates $\{x, y, z\}$, representing each of the partial solutions in Eq. (84) as the following product

---

line is the curve to which the field vectors are tangential at each point. Hence the electric field lines are always normal to the equipotential surfaces, so it is always straightforward to sketch them, if desirable, from the equipotential surface pattern – like the one shown in Fig. 12.

[36] This method was already discussed in CM Sec. 6.5 and then used also in Secs. 6.6 and 8.4 of that course. However, it is so important that I need to repeat its discussion in this part of my series, for the benefit of the readers who have skipped the Classical Mechanics course for whatever reason.

[37] Such voltage distributions may be implemented in practice, for example, using the so-called *mosaic electrodes* consisting of many electrically-insulated and individually-biased panels.

$$\phi_k = X(x)Y(y)Z(z). \tag{2.85}$$

Plugging it into the Laplace equation expressed in the Cartesian coordinates,

$$\frac{\partial^2\phi_k}{\partial x^2} + \frac{\partial^2\phi_k}{\partial y^2} + \frac{\partial^2\phi_k}{\partial z^2} = 0, \tag{2.86}$$

and dividing the result by *XYZ*, we get

$$\frac{1}{X}\frac{d^2X}{dx^2} + \frac{1}{Y}\frac{d^2Y}{dy^2} + \frac{1}{Z}\frac{d^2Z}{dz^2} = 0. \tag{2.87}$$

Here comes the punch line of the variable separation method: since the first term of this sum may depend only on *x*, the second one only of *y*, etc., Eq. (87) may be satisfied everywhere in the volume only if each of these terms equals a constant. In a minute we will see that for our current problem (Fig. 13), these constant *x*- and *y*-terms have to be negative; hence let us denote these *variable separation constants* as $(-\alpha^2)$ and $(-\beta^2)$, respectively. Now Eq. (87) shows that the constant *z*-term has to be positive; denoting it as $\gamma^2$ we get the following relation:

$$\alpha^2 + \beta^2 = \gamma^2. \tag{2.88}$$

Now the variables are separated in the sense that for the functions $X(x)$, $Y(y)$, and $Z(z)$ we got separate ordinary differential equations,

$$\frac{d^2X}{dx^2} + \alpha^2 X = 0, \qquad \frac{d^2Y}{dy^2} + \beta^2 Y = 0, \qquad \frac{d^2Z}{dz^2} - \gamma^2 Z = 0, \tag{2.89}$$

which are related only by Eq. (88) for their constant parameters.

Let us start with the equation for $X(x)$. Its general solution is the sum of functions $\sin\alpha x$ and $\cos\alpha x$, multiplied by arbitrary coefficients. Let us select these coefficients to satisfy our boundary conditions. First, since $\phi \propto X$ should vanish at the back vertical wall of the box (i.e., with the coordinate origin choice shown in Fig. 13, at $x = 0$ for any *y* and *z*), the coefficient at $\cos\alpha x$ should be zero. The remaining coefficient (at $\sin\alpha x$) may be included in the general factor $c_k$ in Eq. (84), so we may take *X* in the form

$$X = \sin\alpha x. \tag{2.90}$$

This solution satisfies the boundary condition at the opposite wall ($x = a$) only if the product $\alpha a$ is a multiple of $\pi$, i.e. if $\alpha$ is equal to any of the following numbers (commonly called *eigenvalues*):[38]

$$\alpha_n = \frac{\pi}{a}n, \qquad \text{with } n = 1, 2, ... \tag{2.91}$$

(Terms with negative values of *n* would not be linearly-independent from those with positive *n* and may be dropped from the sum (84). The value $n = 0$ is formally possible but would give $X = 0$, i.e. $\phi_k = 0$, at

[38] Note that according to Eqs. (91)-(92), as the spatial dimensions *a* and *b* of the system are increased, the distances between the adjacent eigenvalues tend to zero. This fact implies that for spatially infinite systems, the eigenvalue spectra are continuous, so the sums of the type (84) become integrals; however, the general approach remains the same. A few problems of this type are provided in Sec. 9 for the reader's exercise.

any $x$, i.e. no contribution to sum (84), so it may be dropped as well.) Now we see that we indeed had to take $\alpha$ real, i.e. $\alpha^2$ positive – otherwise, instead of the oscillating function (90), we would have a sum of two exponential functions, which cannot equal zero at two independent points of the $x$-axis.

Since the equation (89) for function $Y(y)$ is similar to that for $X(x)$, and the boundary conditions on the walls perpendicular to axis $y$ ($y = 0$ and $y = b$) are similar to those for $x$-walls, the absolutely similar reasoning gives

$$Y = \sin \beta y, \qquad \beta_m = \frac{\pi}{b} m, \qquad \text{with } m = 1, 2, \ldots, \tag{2.92}$$

where the integer $m$ may be selected independently of $n$. Now we see that according to Eq. (88), the separation constant $\gamma$ depends on two integers $n$ and $m$, so the relationship may be rewritten as

$$\gamma_{nm} = \left[\alpha_n^2 + \beta_m^2\right]^{1/2} = \pi \left[\left(\frac{n}{a}\right)^2 + \left(\frac{m}{b}\right)^2\right]^{1/2}. \tag{2.93}$$

The corresponding solution of the differential equation for $Z$ may be represented as a linear combination of two exponents $\exp\{\pm\gamma_{nm} z\}$, or alternatively of two hyperbolic functions, $\sinh\gamma_{nm} z$ and $\cosh\gamma_{nm} z$, with arbitrary coefficients. At our choice of coordinate origin, the latter option is preferable because $\cosh\gamma_{nm} z$ cannot satisfy the zero boundary condition at the bottom lid of the box ($z = 0$). Hence we may take $Z$ in the form

$$Z = \sinh \gamma_{nm} z, \tag{2.94}$$

which automatically satisfies that condition.

Now it is the right time to merge Eqs. (84)-(85) and (90)-(94), replacing the temporary index $k$ with the full set of possible eigenvalues, in our current case of two integer indices $n$ and $m$:

Variable separation in Cartesian coordinates (example)

$$\phi(x, y, z) = \sum_{n,m=1}^{\infty} c_{nm} \sin\frac{\pi n x}{a} \sin\frac{\pi m y}{b} \sinh \gamma_{nm} z, \tag{2.95}$$

where $\gamma_{nm}$ is given by Eq. (93). This solution satisfies not only the Laplace equation but also the boundary conditions on all walls of the box, besides the top lid, for arbitrary coefficients $c_{nm}$. The only job left is to choose these coefficients from the top-lid requirement:

$$\phi(x, y, c) \equiv \sum_{n,m=1}^{\infty} c_{nm} \sin\frac{\pi n x}{a} \sin\frac{\pi m y}{b} \sinh \gamma_{nm} c = V(x, y). \tag{2.96}$$

It may look bad to have just one equation for the infinite set of coefficients $c_{nm}$. However, the decisive help comes from the fact that the functions of $x$ and $y$ that participate in Eq. (96), form *full*, *orthogonal* sets of 1D functions. The last term means that the integrals of the products of the functions with different integer indices over the region of interest equal zero. Indeed, direct integration gives

$$\int_0^a \sin\frac{\pi n x}{a} \sin\frac{\pi n' x}{a} dx = \frac{a}{2} \delta_{nn'}, \tag{2.97}$$

where $\delta_{nn'}$ is the Kronecker symbol, and similarly for $y$ (with the evident replacements $a \to b$, and $n \to m$). Hence, a fruitful way to proceed is to multiply both sides of Eq. (96) by the product of the basis functions, with arbitrary indices $n'$ and $m'$, and integrate the result over $x$ and $y$:

$$\sum_{n,m=1}^{\infty} c_{nm} \sinh \gamma_{nm} c \int_0^a \sin\frac{\pi n x}{a} \sin\frac{\pi n' x}{a} dx \int_0^b \sin\frac{\pi m y}{b} \sin\frac{\pi m' y}{b} dy = \int_0^a dx \int_0^b dy\, V(x,y) \sin\frac{\pi n' x}{a} \sin\frac{\pi m' y}{b} . \quad (2.98)$$

Due to Eq. (97), all terms on the left-hand side of the last equation, besides those with $n = n'$ and $m = m'$, vanish, and (replacing $n'$ with $n$, and $m'$ with $m$, for notation brevity) we finally get

$$c_{nm} = \frac{4}{ab \sinh \gamma_{nm} c} \int_0^a dx \int_0^b dy\, V(x,y) \sin\frac{\pi n x}{a} \sin\frac{\pi m y}{b} . \quad (2.99)$$

The relations (93), (95), and (99) give the complete solution of the posed boundary problem; we can see both good and bad news here. The first bit of bad news is that in the general case, we still need to work out the integrals (99) – formally, the infinite number of them. In some cases, it is possible to do this analytically, in one shot. For example, if the top lid in our problem is a single conductor, i.e. has a constant potential $V_0$, we may take $V(x,y) = V_0 = \text{const}$, and both 1D integrations are elementary; for example

$$\int_0^a \sin\frac{\pi n x}{a} dx = \frac{a}{\pi n} \int_0^{\pi n} \sin\xi d\xi = \frac{a}{\pi n} \times \begin{cases} 2, & \text{for } n \text{ odd}, \\ 0, & \text{for } n \text{ even}, \end{cases} \quad (2.100)$$

and similarly for the integral over $y$, so

$$c_{nm} = \frac{16 V_0}{\pi^2 nm \sinh \gamma_{nm} c} \times \begin{cases} 1, & \text{if both } n \text{ and } m \text{ are odd}, \\ 0, & \text{otherwise}. \end{cases} \quad (2.101)$$

The second bad news is that even on such a happy occasion, we still have to sum up the series (95), so our result may only be called analytical with some reservations because in most cases we need to perform numerical summation to get the final numbers or plots.

Now the first *good* news. Computers are very efficient for both operations (95) and (99), i.e. for the summation and integration. (As was discussed in Sec. 1.2, random errors are averaged out at these operations.) As an example, Fig. 14 shows the plots of the electrostatic potential in a cubic box ($a = b = c$), with an equipotential top lid ($V = V_0 = \text{const}$), obtained by a numerical summation of the series (95), using the analytical expression (101). The remarkable feature of this calculation is a very fast convergence of the series; for the middle cross-section of the cubic box ($z/c = 0.5$), already the first term (with $n = m = 1$) gives an accuracy of about 6%, while the sum of four leading terms (with $n, m = 1, 3$) reduces the error to just 0.2%. (For a longer box, $c > a, b$, the convergence is even faster – see the discussion below.) Only very close to the corners between the top lid and the sidewalls, where the potential changes rapidly, several more terms are necessary to get a reasonable accuracy.

The related piece of good news is that our "semi-analytical" result allows its ultimate limits to be explored analytically. For example, Eq. (93) shows that for a very flat box (with $c << a, b$), $\gamma_{n,m} z \le \gamma_{n,m} c << 1$ at least for the lowest terms of series (95), with $n, m << c/a, c/b$. In this case, the sinh functions in Eqs. (96) and (99) may be well approximated with their arguments, and their ratio by $z/c$. So if we limit the summation to these terms, Eq. (95) gives a very simple result

$$\phi(x,y) \approx \frac{z}{c} V(x,y) , \quad (2.102)$$

which means that each elementary segment of the flat box behaves just as a plane capacitor. Only near the sidewalls, the higher terms in the series (95) are important, producing some deviations from Eq. (102). (For the general problem with an arbitrary function $V(x,y)$, this is also true in all regions where this function changes sharply.)

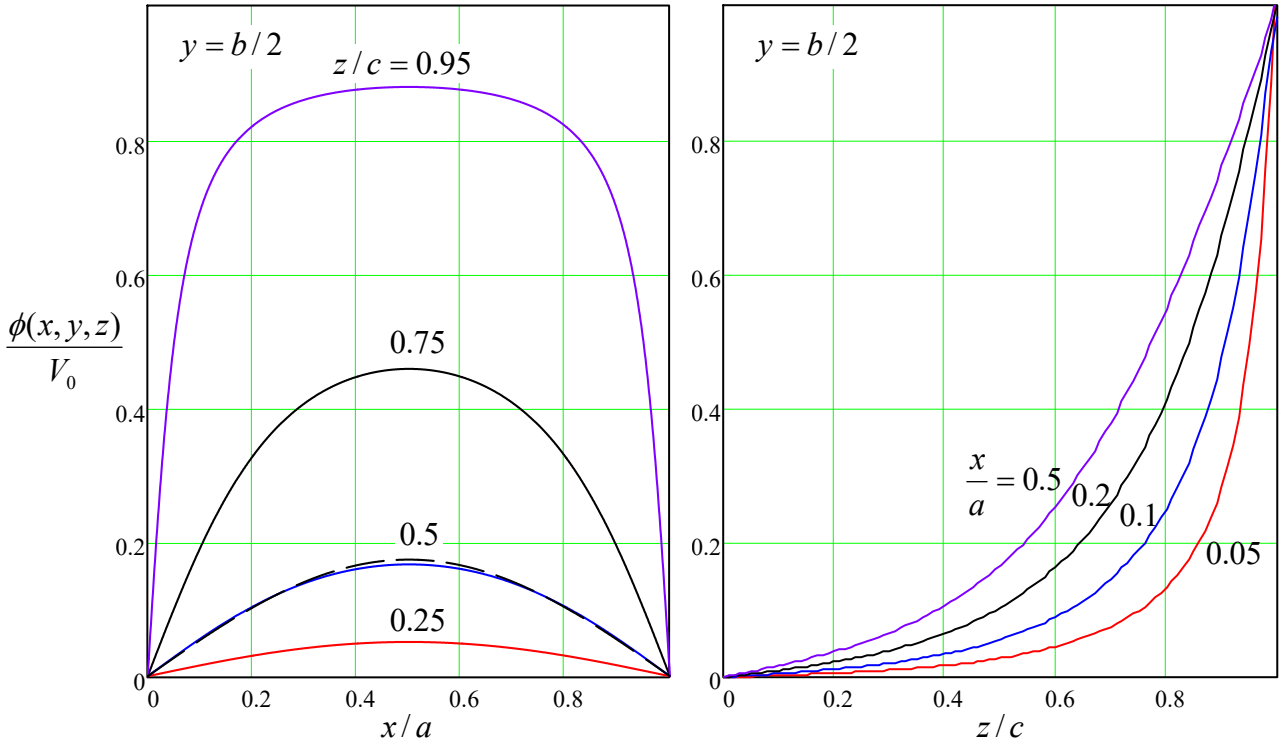



Fig. 2.14. The electrostatic potential's distribution inside a cubic box ($a = b = c$) with a constant voltage $V_0$ on the top lid (Fig. 13), calculated numerically from Eqs. (93), (95), and (101). The dashed line on the left panel shows the contribution of the main term of the series (with $n = m = 1$) to the full result, for $z/c = 0.5$.

In the opposite limit ($a, b << c$), Eq. (93) shows that on the contrary, $\gamma_{n,m}c >> 1$ for all $n$ and $m$. Moreover, the ratio $\sinh\gamma_{n,m}z/\sinh\gamma_{n,m}c$ drops sharply if either $n$ or $m$ is increased, provided that $z$ is not too close to $c$. Hence in this case a very good approximation may be obtained by keeping just the leading term, with $n = m = 1$, in Eq. (95), so the challenge of summation disappears. (As was discussed above, this approximation works reasonably well even for a cubic box.) In particular, for the constant potential of the upper lid, we can use Eq. (101) and the exponential asymptotic for both sinh functions, to get a very simple formula:

$$\phi = \frac{16}{\pi^2}\sin\frac{\pi x}{a}\sin\frac{\pi y}{b}\exp\left\{-\pi\frac{\left(a^2+b^2\right)^{1/2}}{ab}(c-z)\right\}. \qquad (2.103)$$

These results may be readily generalized to some other problems. For example, if all walls of the box shown in Fig. 13 have an arbitrary potential distribution, we may use the linear superposition principle to represent the electrostatic potential distribution as the sum of six partial solutions of the type of Eq. (95), each with one wall biased by the corresponding voltage, and all other grounded ($\phi = 0$).

To summarize, the results given by the variable separation method in the Cartesian coordinates are closer to what we could call a genuinely analytical solution than to a purely numerical solution.

Now, let us explore the issues that arise when this method is applied in other orthogonal coordinate systems.

## 2.6. Variable separation – polar coordinates

If a system of conductors is cylindrical, the potential distribution is independent of the $z$-coordinate along the cylinder axis: $\partial\phi/\partial z = 0$, and the Laplace equation becomes two-dimensional. If the conductor's cross-section is rectangular, the variable separation method works best in Cartesian coordinates $\{x, y\}$, and is just a particular case of the 3D solution discussed above. However, if the cross-section is circular, much more compact results may be obtained by using the polar coordinates $\{\rho, \varphi\}$. As we already know from Sec. 3(ii), these 2D coordinates are orthogonal, so the two-dimensional Laplace operator is a sum of two separable terms.[39] Requiring, just as we have done above, each component of the sum (84) to satisfy the Laplace equation, we get

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\phi_k}{\partial\rho}\right)+\frac{1}{\rho^2}\frac{\partial^2\phi_k}{\partial\varphi^2}=0. \tag{2.104}$$

In a full analogy with Eq. (85), let us represent each particular solution $\phi_k$ as a product $\mathcal{R}(\rho)\mathcal{F}(\varphi)$. Plugging this expression into Eq. (104) and then dividing all its parts by $\mathcal{R}\mathcal{F}/\rho^2$, we get

$$\frac{\rho}{\mathcal{R}}\frac{d}{d\rho}\left(\rho\frac{d\mathcal{R}}{d\rho}\right)+\frac{1}{\mathcal{F}}\frac{d^2\mathcal{F}}{d\varphi^2}=0. \tag{2.105}$$

Following the same reasoning as for the Cartesian coordinates, we get two separated ordinary differential equations

$$\rho\frac{d}{d\rho}\left(\rho\frac{d\mathcal{R}}{d\rho}\right)=\nu^2\mathcal{R}, \tag{2.106}$$

$$\frac{d^2\mathcal{F}}{d\varphi^2}+\nu^2\mathcal{F}=0, \tag{2.107}$$

where $\nu^2$ is the variable separation constant.

Let us start their analysis from Eq. (106), plugging into it a probe solution $\mathcal{R} = c\rho^\alpha$ where $c$ and $\alpha$ are some constants. The elementary differentiation shows that if $\alpha \neq 0$, the equation is indeed satisfied for any $c$, with just one requirement imposed on the constant $\alpha$, namely $\alpha^2 = \nu^2$. This means that the following linear superposition

$$\mathcal{R}=a_\nu\rho^{+\nu}+b_\nu\rho^{-\nu}, \qquad \text{for } \nu\neq 0, \tag{2.108}$$

with any constant coefficients $a_\nu$ and $b_\nu$, is also a solution of Eq. (106). Moreover, the general theory of linear ordinary differential equations tells us that the solution of a second-order equation like Eq. (106) may only depend on just two constant factors that scale two linearly independent functions. Hence, for all values $\nu^2 \neq 0$, Eq. (108) presents the *general* solution of that equation. The case when $\nu = 0$, in which the functions $\rho^{+\nu}$ and $\rho^{-\nu}$ are just constants and hence are *not* linearly independent, is special, but in this case, the integration of Eq. (106) is straightforward,[40] giving

[39] See, e.g., MA Eq. (10.3) with $\partial/\partial z = 0$.

[40] Actually, we have already performed it in Sec. 3 – see Eq. (43).

$$\mathcal{R} = a_0 + b_0 \ln \rho, \qquad \text{for } \nu = 0. \tag{2.109}$$

In order to specify the separation constant, let us explore Eq. (107), whose general solution is

$$\mathcal{F} = \begin{cases} c_\nu \cos \nu\varphi + s_\nu \sin \nu\varphi, & \text{for } \nu \neq 0, \\ c_0 + s_0 \varphi, & \text{for } \nu = 0. \end{cases} \tag{2.110}$$

There are two possible cases here. In many boundary problems solvable in cylindrical coordinates, the free-space region, in which the Laplace equation is valid, extends continuously around the origin point $\rho = 0$. In this region, the potential has to be continuous and uniquely defined, so $\mathcal{F}$ has to be a $2\pi$-periodic function of $\varphi$. For that, one needs the product $\nu(\varphi + 2\pi)$ to equal $\nu\varphi + 2\pi n$, with $n$ being an integer, immediately giving us a discrete spectrum of possible values of the variable separation constant:

$$\nu = n = 0, \pm 1, \pm 2, \ldots \tag{2.111}$$

In this case, both functions $\mathcal{R}$ and $\mathcal{F}$ may be labeled with the integer index $n$. Taking into account that the terms with negative values of $n$ may be summed up with those with positive $n$, and that $s_0$ has to equal zero (otherwise the $2\pi$-periodicity of function $\mathcal{F}$ would be violated), we see that the general solution of the 2D Laplace equation for such geometries may be represented as

Variable separation in polar coordinates

$$\phi(\rho,\varphi) = a_0 + b_0 \ln \rho + \sum_{n=1}^{\infty} \left( a_n \rho^n + \frac{b_n}{\rho^n} \right) \left( c_n \cos n\varphi + s_n \sin n\varphi \right). \tag{2.112}$$

Let us see how all this machinery works on the famous problem of a round cylindrical conductor placed into an electric field that is uniform and perpendicular to the cylinder's axis at large distances (see Fig. 15a), as if it is created by a large plane capacitor. First of all, let us explore the effect of the system's symmetries on the coefficients in Eq. (112). Selecting the coordinate system as shown in Fig. 15a, and taking the cylinder's potential for zero, we immediately get $a_0 = 0$.

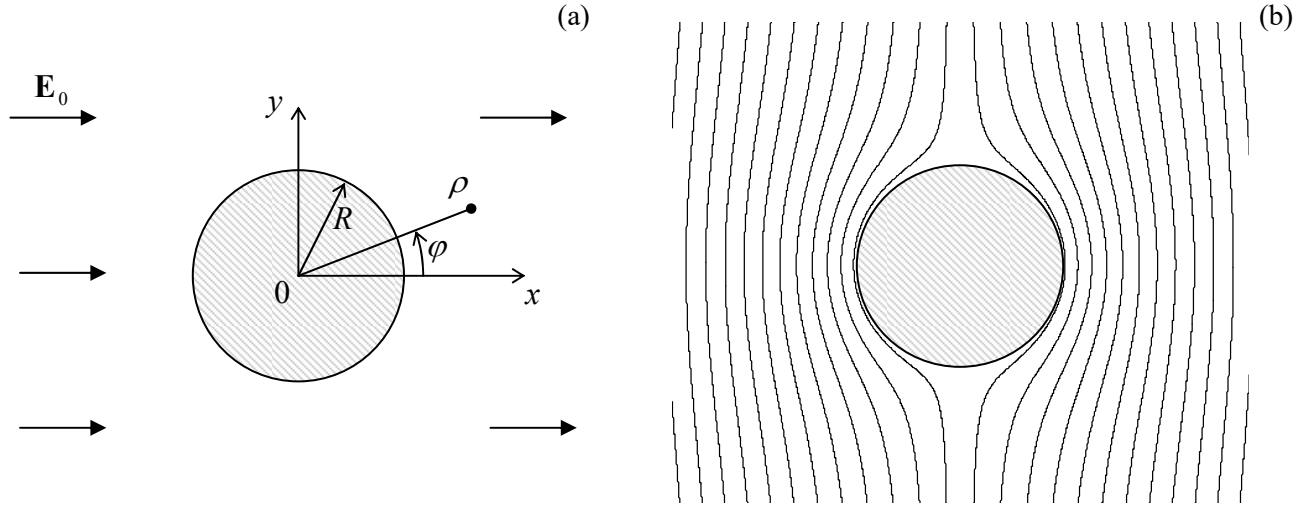



Fig. 2.15. A conducting cylinder inserted into an initially uniform electric field perpendicular to its axis: (a) the problem's geometry, and (b) the equipotential surfaces given by Eq. (117).

Moreover, due to the mirror symmetry about the plane [$x$, $z$], the solution has to be an even function of the angle $\varphi$, and hence all coefficients $s_n$ should also equal zero. Also, at large distances ($\rho \gg R$) from the cylinder, its effect on the electric field should vanish, and the potential should approach that of the uniform external field $\mathbf{E} = E_0\mathbf{n}_x$:

$$\phi \to -E_0 x \equiv -E_0 \rho \cos\varphi, \quad \text{for } \rho \to \infty . \tag{2.113}$$

This is only possible if in Eq. (112), $b_0 = 0$, and also all coefficients $a_n$ with $n \neq 1$ vanish, while the product $a_1c_1$ should be equal to $(-E_0)$. Thus the solution is reduced to the following form

$$\phi(\rho,\varphi) = -E_0 \rho \cos\varphi + \sum_{n=1}^{\infty} \frac{B_n}{\rho^n} \cos n\varphi , \tag{2.114}$$

in which the coefficients $B_n \equiv b_n c_n$ should be found from the boundary condition at $\rho = R$:

$$\phi(R,\varphi) = 0 . \tag{2.115}$$

This requirement yields the following equation,

$$\left( -E_0 R + \frac{B_1}{R} \right) \cos\varphi + \sum_{n=2}^{\infty} \frac{B_n}{R^n} \cos n\varphi = 0, \tag{2.116}$$

which should be satisfied for all $\varphi$. This equality, read backward, may be considered as an expansion of a function identically equal to zero into a series over mutually orthogonal functions $\cos n\varphi$. It is evidently valid if all coefficients of the expansion, including $(-E_0R + B_1/R)$, and all $B_n$ for $n \geq 2$ are equal to zero. Moreover, mathematics tells us that such expansions are unique, so this is the only possible solution of Eq. (116). So, $B_1 = E_0R^2$, and our final answer (valid only outside of the cylinder, i.e. for $\rho \geq R$), is

$$\phi(\rho,\varphi) = -E_0 \left( \rho - \frac{R^2}{\rho} \right) \cos\varphi \equiv -E_0 \left( 1 - \frac{R^2}{x^2 + y^2} \right) x . \tag{2.117}$$

This result, which may be graphically represented with the equipotential surfaces shown in Fig. 15b, shows a smooth transition between the uniform field (113) far from the cylinder, to the equipotential surface of the cylinder (with $\phi = 0$). Such smoothening is very typical for Laplace equation solutions. Indeed, as we know from Chapter 1, these solutions correspond to the lowest integral of the potential gradient's square, i.e. to the lowest potential energy (1.65) possible at the given boundary conditions.

To complete the problem, let us use Eq. (3) to calculate the distribution of the surface charge density over the cylinder's cross-section:

$$\sigma = \varepsilon_0 E_n \big|_{\text{surface}} \equiv -\varepsilon_0 \frac{\partial \phi}{\partial \rho}\Big|_{\rho = R} = \varepsilon_0 E_0 \cos\varphi \frac{\partial}{\partial \rho} \left( \rho - \frac{R^2}{\rho} \right)_{\rho = R} = 2\varepsilon_0 E_0 \cos\varphi. \tag{2.118}$$

This very simple formula shows that with the field direction shown in Fig. 15a ($E_0 > 0$), the surface charge is positive on the right-hand side of the cylinder and negative on its left-hand side, thus creating a field directed from the right to the left, which exactly compensates the external field inside the conductor, where the net field is zero. (Please take one more look at the schematic Fig. 1a.) Note also that the net electric charge of the cylinder is zero, in correspondence with the problem symmetry.

Another useful by-product of the calculation (118) is that the surface electric field equals $2E_0\cos\varphi$, and hence its largest magnitude is twice the field far from the cylinder. Such electric field concentration is very typical for all convex conducting surfaces.

The last observation gets additional confirmation from the second possible topology when Eq. (110) is used to describe problems with no angular periodicity. A typical example of this situation is a cylindrical conductor with a cross-section that features a corner limited by two straight-line segments (Fig. 16). Indeed, we may argue that at $\rho < R$ (where $R$ is the radial extension of the planar sides of the corner, see Fig. 16), the Laplace equation may be satisfied by a sum of partial solutions $\mathcal{R}(\rho)\mathcal{F}(\varphi)$, if the angular components of the products satisfy the boundary conditions on the corner sides. Taking (just for the simplicity of notation) the conductor's potential to be zero, and one of the corner's sides as the $x$-axis ($\varphi = 0$), these boundary conditions are

$$\mathcal{F}(0) = \mathcal{F}(\beta) = 0\,, \tag{2.119}$$

where the angle $\beta$ may be anywhere between 0 and $2\pi$ – see Fig. 16.

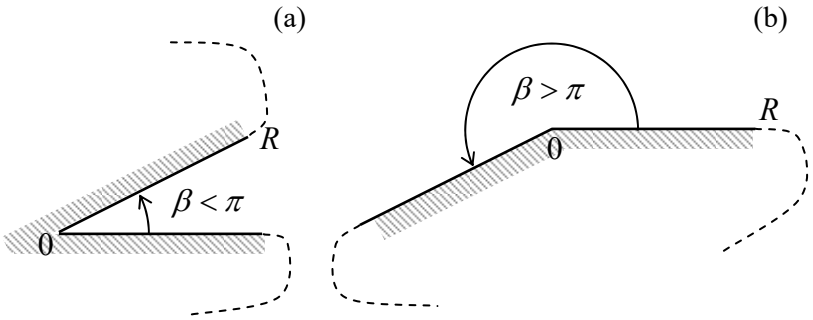



Fig. 2.16. The cross-sections of cylindrical conductors with (a) a corner and (b) a wedge.

Comparing this condition with Eq. (110), we see that it requires $s_0$ and all $c_\nu$ to vanish, and $\nu$ to take one of the values of the following discrete spectrum:

$$\nu_m\beta = \pi m, \qquad \text{with } m = 1, 2, \ldots. \tag{2.120}$$

Hence the full solution of the Laplace equation for this geometry takes the form

$$\phi = \sum_{m=1}^{\infty} a_m \rho^{\pi m/\beta} \sin\frac{\pi m\varphi}{\beta}, \qquad \text{for } \rho < R, \quad 0 \le \varphi \le \beta\,, \tag{2.121}$$

where the constants $s_\nu$ have been incorporated into $a_m$. The set of coefficients $a_m$ cannot be universally determined, because it depends on the exact shape of the conductor outside the corner, and the externally applied electric field. However, whatever the set is, in the limit $\rho \to 0$, the solution (121) is almost[41] always dominated by the term with the lowest $m = 1$:

$$\phi \to a_1 \rho^{\pi/\beta} \sin\frac{\pi}{\beta}\varphi\,, \tag{2.122}$$

because the higher terms tend to zero faster. This potential distribution corresponds to the surface charge density

[41] Exceptions are possible only for highly symmetric configurations when the external field is specially crafted to make $a_1 = 0$. In this case, the solution at $\rho \to 0$ is dominated by the first nonzero term of the series (121).

$$\sigma = \varepsilon_0 E_n\big|_{\text{surface}} = -\varepsilon_0 \left.\frac{\partial \phi}{\partial(\rho\varphi)}\right|_{\rho=\text{const},\ \varphi\to+0} = -\varepsilon_0 \frac{\pi a_1}{\beta}\rho^{(\pi/\beta-1)}. \tag{2.123}$$

(It is similar, with the opposite sign, on the opposite face of the angle.)

The result (123) shows that if we are dealing with a concave corner ($\beta < \pi$, see Fig. 16a), the charge density (and the surface electric field) tends to zero. On the other hand, at a "convex corner" with $\beta > \pi$ (actually, a wedge – see Fig. 16b), both the charge and the field's strength concentrate, formally diverging at $\rho \to 0$. (So, do not sit on a roof's ridge during a thunderstorm; rather hide in a ditch!) We have already seen qualitatively similar effects for the thin round disk and the split plane.

## 2. 7. Variable separation – cylindrical coordinates

Now, let us discuss how to generalize the approach discussed in the previous section to problems whose geometry is still axially symmetric, but where the electrostatic potential depends not only on the radial and angular coordinates but also on the axial coordinate: $\partial\phi/\partial z \neq 0$. The classical example of such a problem is shown in Fig. 17. Here the sidewall and the bottom lid of a hollow round cylinder are kept at a fixed potential (say, $\phi = 0$), but the potential $V$ fixed at the top lid is different. Evidently, this problem is qualitatively similar to the rectangular box problem solved above (Fig. 13), and we will also try to solve it first for the case of arbitrary voltage distribution over the top lid: $V = V(\rho, \varphi)$.

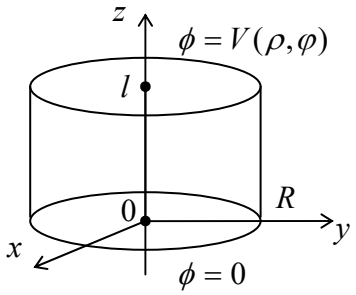



Fig. 2.17. A cylindrical volume with conducting walls.

Following the main idea of the variable separation method, let us require that each partial function $\phi_k$ in Eq. (84) satisfies the Laplace equation, now in the full cylindrical coordinates $\{\rho, \varphi, z\}$:[42]

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\phi_k}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2\phi_k}{\partial\varphi^2} + \frac{\partial^2\phi_k}{\partial z^2} = 0. \tag{2.124}$$

Plugging $\phi_k$ in the form of the product $\mathcal{R}(\rho)\mathcal{F}(\varphi)\mathcal{Z}(z)$ into Eq. (124) and then dividing all resulting terms by this product, we get

$$\frac{1}{\rho\mathcal{R}}\frac{d}{d\rho}\left(\rho\frac{d\mathcal{R}}{d\rho}\right) + \frac{1}{\rho^2\mathcal{F}}\frac{d^2\mathcal{F}}{d\varphi^2} + \frac{1}{\mathcal{Z}}\frac{d^2\mathcal{Z}}{dz^2} = 0. \tag{2.125}$$

Since the first two terms of Eq. (125) can only depend on the polar variables $\rho$ and $\varphi$, while the third term, only on $z$, at least that term should equal a constant. Denoting it (just like we did in the rectangular box problem) by $\gamma^2$, we get the following set of two equations:

[42] See, e.g., MA Eq. (10.3).

$$\frac{d^2\mathcal{Z}}{dz^2} = \gamma^2 \mathcal{Z}, \tag{2.126}$$

$$\frac{1}{\rho\mathcal{R}}\frac{d}{d\rho}\left(\rho\frac{d\mathcal{R}}{d\rho}\right) + \gamma^2 + \frac{1}{\rho^2\mathcal{F}}\frac{d^2\mathcal{F}}{d\varphi^2} = 0. \tag{2.127}$$

Now, multiplying all the terms of Eq. (127) by $\rho^2$, we see that the last term of the result, $(d^2\mathcal{F}/d\varphi^2)/\mathcal{F}$, may depend only on $\varphi$, and thus should equal a constant. Calling that constant $\nu^2$ (just as in Sec. 6 above), we separate Eq. (127) into an angular equation,

$$\frac{d^2\mathcal{F}}{d\varphi^2} + \nu^2\mathcal{F} = 0, \tag{2.128}$$

and a radial equation:

$$\frac{d^2\mathcal{R}}{d\rho^2} + \frac{1}{\rho}\frac{d\mathcal{R}}{d\rho} + \left(\gamma^2 - \frac{\nu^2}{\rho^2}\right)\mathcal{R} = 0. \tag{2.129}$$

We see that the ordinary differential equations for the functions $\mathcal{Z}(z)$ and $\mathcal{F}(\varphi)$ (and hence their solutions) are identical to those discussed earlier in this chapter. However, Eq. (129) for the radial function $\mathcal{R}(\rho)$ (called the *Bessel equation*) is more complex than in the 2D case and depends on two independent constant parameters, $\gamma$ and $\nu$. The latter challenge may be readily overcome if we notice that any change of $\gamma$ may be reduced to the corresponding re-scaling of the radial coordinate $\rho$. Indeed, introducing a dimensionless variable $\xi \equiv \gamma\rho$,[43] Eq. (129) may be reduced to an equation with just one parameter, $\nu$:

Bessel equation

$$\boxed{\frac{d^2\mathcal{R}}{d\xi^2} + \frac{1}{\xi}\frac{d\mathcal{R}}{d\xi} + \left(1 - \frac{\nu^2}{\xi^2}\right)\mathcal{R} = 0.} \tag{2.130}$$

Moreover, we already know that for angle-periodic problems, the spectrum of eigenvalues of Eq. (128) is discrete: $\nu = n$, with integer $n$.

Unfortunately, even in this case, Eq. (130), which is the canonical form of the Bessel equation, cannot be satisfied by a single "elementary" function. The solutions that we need for our current problem are called the *Bessel function of the first kind of order* $\nu$, commonly denoted as $J_\nu(\xi)$. Let me review in brief those properties of these functions that are most relevant to our problem – and many other problems discussed in this series.[44]

First of all, the Bessel function of a negative integer order is very simply related to that of the positive order:

$$J_{-n}(\xi) = (-1)^n J_n(\xi), \tag{2.131}$$

enabling us to limit our discussion to the functions with $n \geq 0$. Figure 18 shows four of these functions with the lowest positive $n$.

[43] Note that this normalization is specific for each value of the variable separation parameter $\gamma$. Also, please notice that the normalization is meaningless for $\gamma = 0$, i.e. for the case $Z(z) = \text{const}$. However, if we need partial solutions for this particular value of $\gamma$, we can always use Eqs. (108)-(109).

[44] For a more complete discussion of these functions, see the literature listed in MA Sec. 16, for example, Chapter 6 (written by F. Olver) in the famous collection compiled and edited by Abramowitz and Stegun, available online.

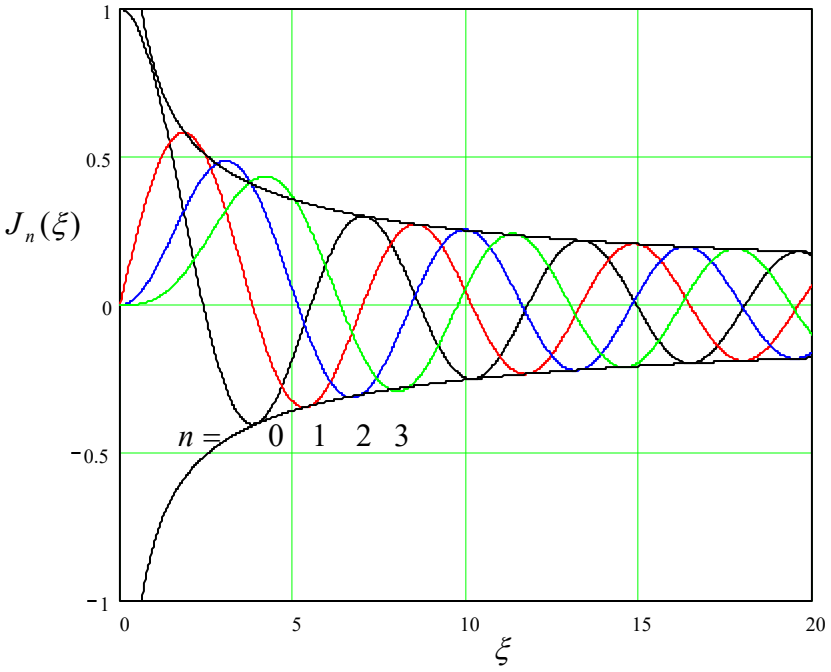



Fig. 2.18. Several Bessel functions $J_n(\xi)$ of integer order. The dashed lines show the envelope of the asymptotes (135).

As its argument is increased, each function is initially close to a power law: $J_0(\xi) \approx 1$, $J_1(\xi) \approx \xi/2$, $J_2(\xi) \approx \xi^2/8$, etc. This behavior follows from the Taylor series

$$J_n(\xi) = \left(\frac{\xi}{2}\right)^n \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(n+k)!}\left(\frac{\xi}{2}\right)^{2k}, \tag{2.132}$$

which is formally valid for any $\xi$, and may even serve as an alternative definition of the functions $J_n(\xi)$. However, the series is converging fast only at small arguments, $\xi << n$, where its leading term is

$$J_n(\xi)\Big|_{\xi\to 0} \to \frac{1}{n!}\left(\frac{\xi}{2}\right)^n. \tag{2.133}$$

At $\xi \approx n + 1.86n^{1/3}$, the Bessel function reaches its maximum[45]

$$\max_{\xi}\left[J_n(\xi)\right] \approx \frac{0.675}{n^{1/3}}, \tag{2.134}$$

and then starts to oscillate with a period gradually approaching $2\pi$, a phase shift that increases by $\pi/2$ with each unit increment of $n$, and an amplitude that decreases as $\xi^{-1/2}$. All these features are described by the following asymptotic formula:

$$J_n(\xi)\Big|_{\xi\to\infty} \to \left(\frac{2}{\pi\xi}\right)^{1/2} \cos\left(\xi - \frac{\pi}{4} - \frac{n\pi}{2}\right), \tag{2.135}$$

which starts to give a reasonable approximation soon after the function peaks – see Fig. 18.[46]

[45] These two approximations for the Bessel function peak are strictly valid for $n >> 1$, but may be used for reasonable estimates starting already from $n = 1$. For example, $\max_{\xi}[J_1(\xi)]$ is close to 0.58 and is reached at $\xi \approx 2.4$, just about 30% away from the values given by the asymptotic formulas.

Now we are ready for our case study (Fig. 17). Since the functions the $Z(z)$ have to satisfy not only Eq. (126) but also the bottom-lid boundary condition $Z(0) = 0$, they are proportional to $\sinh\gamma z$ – cf. Eq. (94). Then Eq. (84) becomes

$$\phi = \sum_{n=0}^{\infty}\sum_{\gamma} J_n(\gamma\rho)\left(c_{n\gamma}\cos n\varphi + s_{n\gamma}\sin n\varphi\right)\sinh\gamma z\,. \tag{2.136}$$

Next, we need to satisfy the zero boundary condition at the cylinder's side wall ($\rho = R$). This may be ensured by taking

$$J_n(\gamma R) = 0\,. \tag{2.137}$$

Since each function $J_n(x)$ has an infinite number of positive zeros (see Fig. 18 again), which may be numbered by an integer index $m = 1, 2, \ldots$, Eq. (137) may be satisfied with an infinite number of discrete values of the parameter $\gamma$:

$$\gamma_{nm} = \frac{\xi_{nm}}{R}, \tag{2.138}$$

where $\xi_{nm}$ is the $m$-th zero of the function $J_n(x)$ – see the top numbers in the cells of Table 1. (Very soon we will see what we need the bottom numbers for.)

Table 2.1. Approximate values of a few first zeros, $\xi_{nm}$, of a few lowest-order Bessel functions $J_n(\xi)$ (the top number in each cell), and the values of $dJ_n(\xi)/d\xi$ at these points (the bottom number).

| | $m = 1$ | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $n = 0$ | 2.40482<br>-0.51914 | 5.52008<br>+0.34026 | 8.65372<br>-0.27145 | 11.79215<br>+0.23245 | 14.93091<br>-0.20654 | 18.07106<br>+0.18773 |
| 1 | 3.83171<br>-0.40276 | 7.01559<br>+0.30012 | 10.17347<br>-0.24970 | 13.32369<br>+0.21836 | 16.47063<br>-0.19647 | 19.61586<br>+0.18006 |
| 2 | 5.13562<br>-0.33967 | 8.41724<br>+0.27138 | 11.61984<br>-0.23244 | 14.79595<br>+0.20654 | 17.95982<br>-0.18773 | 21.11700<br>+0.17326 |
| 3 | 6.38016<br>-0.29827 | 9.76102<br>+0.24942 | 13.01520<br>-0.21828 | 16.22347<br>+0.19644 | 19.40942<br>-0.18005 | 22.58273<br>+0.16718 |
| 4 | 7.58834<br>-0.26836 | 11.06471<br>+0.23188 | 14.37254<br>-0.20636 | 17.61597<br>+0.18766 | 20.82693<br>-0.17323 | 24.01902<br>+0.16168 |
| 5 | 8.77148<br>-0.24543 | 12.33860<br>+0.21743 | 15.70017<br>-0.19615 | 18.98013<br>+0.17993 | 22.21780<br>-0.16712 | 25.43034<br>+0.15669 |

Hence, Eq. (136) may be represented in a more explicit form:

Variable separation in cylindrical coordinates (example)

$$\phi(\rho,\varphi,z) = \sum_{n=0}^{\infty}\sum_{m=1}^{\infty} J_n\left(\xi_{nm}\frac{\rho}{R}\right)\left(c_{nm}\cos n\varphi + s_{nm}\sin n\varphi\right)\sinh\left(\xi_{nm}\frac{z}{R}\right). \tag{2.139}$$

[46] Eq. (135) and Fig. 18 clearly show the close analogy between the Bessel functions and the usual trigonometric functions, sine and cosine. To emphasize this similarity, and help the reader to develop more gut feeling of the Bessel functions, let me mention one result of the elasticity theory: while the sinusoidal functions describe, in particular, transverse standing waves on a guitar string, the functions $J_n(\xi)$ describe, in particular, transverse standing waves on an elastic round membrane (say, a round drum), with $J_0(\xi)$ describing their lowest (fundamental) mode – the only mode with a nonzero amplitude of the membrane center's oscillations.

Here the coefficients $c_{nm}$ and $s_{nm}$ have to be selected to satisfy the only remaining boundary condition – that on the top lid:

$$\phi(\rho,\varphi,l) \equiv \sum_{n=0}^{\infty}\sum_{m=1}^{\infty} J_n\left(\xi_{nm}\frac{\rho}{R}\right)\left(c_{nm}\cos n\varphi + s_{nm}\sin n\varphi\right)\sinh\left(\xi_{nm}\frac{l}{R}\right) = V(\rho,\varphi). \tag{2.140}$$

To use it, let us multiply both sides of Eq. (140) by the product $J_n(\xi_{nm'}\rho/R)\cos n'\varphi$, integrate the result over the lid area, and use the following property of the Bessel functions:

$$\int_0^1 J_n(\xi_{nm}s)J_n(\xi_{nm'}s)sds = \frac{1}{2}\left[J_{n+1}(\xi_{nm})\right]^2\delta_{mm'}. \tag{2.141}$$

As a small but important detour, the last relation expresses a very specific ("2D") orthogonality of the Bessel functions with different indices $m$ – do not confuse them with the function order indices $n$, please![47] Since it relates two Bessel functions of the same order $n$, it is natural to ask why its right-hand side contains the function with a different order $(n + 1)$. Some gut feeling of that may come from one more very important property of the Bessel functions, the so-called *recurrence relations*:[48]

$$J_{n-1}(\xi) + J_{n+1}(\xi) = \frac{2nJ_n(\xi)}{\xi}, \tag{2.142a}$$

$$J_{n-1}(\xi) - J_{n+1}(\xi) = 2\frac{dJ_n(\xi)}{d\xi}, \tag{2.142b}$$

which in particular yield the following formula (convenient for working out some Bessel function integrals):

$$\frac{d}{d\xi}\left[\xi^n J_n(\xi)\right] = \xi^n J_{n-1}(\xi). \tag{2.143}$$

Let us apply the recurrence relations at the special points $\xi_{nm}$. At these points, $J_n$ vanishes, and the system of two equations (142) may be readily solved to get, in particular,

$$J_{n+1}(\xi_{nm}) = -\frac{dJ_n}{d\xi}(\xi_{nm}), \tag{2.144}$$

so the square bracket on the right-hand side of Eq. (141) is just $(dJ_n/d\xi)^2$ at $\xi = \xi_{nm}$. Thus the values of the Bessel function derivatives at the zero points of the function, given by the lower numbers in the cells of Table 1, are as important for boundary problem solutions as the zeros themselves.

Now returning to our problem: since the angular functions $\cos n\varphi$ are also orthogonal – both to each other,

---

[47] The Bessel functions of the *same argument* but *different orders* are also orthogonal, but differently:

$$\int_0^{\infty} J_n(\xi)J_{n'}(\xi)\frac{d\xi}{\xi} = \frac{1}{n+n'}\delta_{nn'}.$$

[48] These relations provide, in particular, a convenient way for numerical computation of all $J_n(\xi)$ – after $J_0(\xi)$ has been computed. (The latter task is usually performed using Eq. (132) for smaller $\xi$ and an extension of Eq. (135) for larger $\xi$.) Note that most mathematical software packages, including all those listed in MA Sec. 16(iv), include ready subroutines for calculation of the functions $J_n(\xi)$ and other special functions used in this lecture series. In this sense, the conditional line separating these "special functions" from "elementary functions" is rather fine.

$$\int_0^{2\pi} \cos(n\varphi)\cos(n'\varphi)\, d\varphi = \pi\delta_{nn'}, \tag{2.145}$$

and to all functions $\sin n\varphi$, the integration over the lid area kills all terms of both series in Eq. (140), besides just one term proportional to $c_{n'm'}$, and hence gives an explicit expression for that coefficient. The counterpart coefficients $s_{n'm'}$ may be found by repeating the same procedure with the replacement of $\cos n'\varphi$ by $\sin n'\varphi$. This evaluation (left for the reader's exercise) completes the solution of our problem for an arbitrary lid potential $V(\rho,\varphi)$.

Still, before leaving the Bessel functions behind (for a while only :-), let me address two important issues. First, we have seen that in our cylinder problem (Fig. 17), the set of functions $J_n(\xi_{nm}\rho/R)$ with different indices $m$ (which characterize the degree of Bessel function's stretch along axis $\rho$) play a role similar to that of functions $\sin(\pi mx/a)$ in the rectangular box problem shown in Fig. 13. In this context, what is the analog of functions $\cos(\pi mx/a)$ – which may be important for some boundary problems? In a more formal language, are there any functions of the same argument $\xi \equiv \xi_{nm}\rho/R$, that would be linearly independent of the Bessel functions of the first kind, while satisfying the same Bessel equation (130)?

The answer is *yes*. For the definition of such functions, we first need to generalize our prior formulas for $J_n(\xi)$, and in particular Eq. (132), to the case of arbitrary, not necessarily real order $\nu$. Mathematics says that the generalization may be performed in the following way:

$$J_\nu(\xi) = \left(\frac{\xi}{2}\right)^\nu \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\Gamma(\nu+k+1)}\left(\frac{\xi}{2}\right)^{2k}, \tag{2.146}$$

where $\Gamma(s)$ is the so-called *gamma function* that may be defined as[49]

$$\Gamma(s) \equiv \int_0^{\infty} \xi^{s-1} e^{-\xi}\, d\xi. \tag{2.147}$$

The simplest, and the most important property of the gamma function is that for integer values of its argument, it gives the factorial of the number smaller by one:

$$\Gamma(n+1) = n! \equiv 1\cdot 2\cdot \ldots \cdot n, \tag{2.148}$$

so it is essentially a generalization of the notion of the factorial to all real numbers.

The Bessel functions defined by Eq. (146) satisfy, after the replacements $n \to \nu$ and $n! \to \Gamma(n+1)$, virtually all the relations discussed above, including the Bessel equation (130), the asymptotic formula (135), the orthogonality condition (141), and the recurrence relations (142). Moreover, it may be shown that $\nu \neq n$, functions $J_\nu(\xi)$ and $J_{-\nu}(\xi)$ are linearly independent of each other, and hence their linear combination may be used to represent the general solution of the Bessel equation. Unfortunately, as Eq. (131) shows, for $\nu = n$ this is not true, and a solution linearly independent of $J_n(\xi)$ has to be formed differently. The most common way to do that is first to define, for all $\nu \neq n$, the following functions:

$$Y_\nu(\xi) \equiv \frac{J_\nu(\xi)\cos\nu\pi - J_{-\nu}(\xi)}{\sin\nu\pi}, \tag{2.149}$$

[49] See, e.g., MA Eq. (6.7a). Note that $\Gamma(s) \to \infty$ at $s \to 0, -1, -2, \ldots$

called the *Bessel functions of the second kind*, or more often the *Weber functions*,[50] and then to follow the limit $\nu \to n$. At this, both the numerator and denominator of the right-hand side of Eq. (149) tend to zero, but their ratio tends to a finite value called $Y_n(x)$. It may be shown that the resulting functions are still the solutions of the Bessel equation and are linearly independent of $J_n(x)$, though are related just as those functions if the sign of $n$ changes:

$$Y_{-n}(\xi) = (-1)^n Y_n(\xi) . \tag{2.150}$$

Figure 19 shows a few Weber functions of the lowest integer orders.

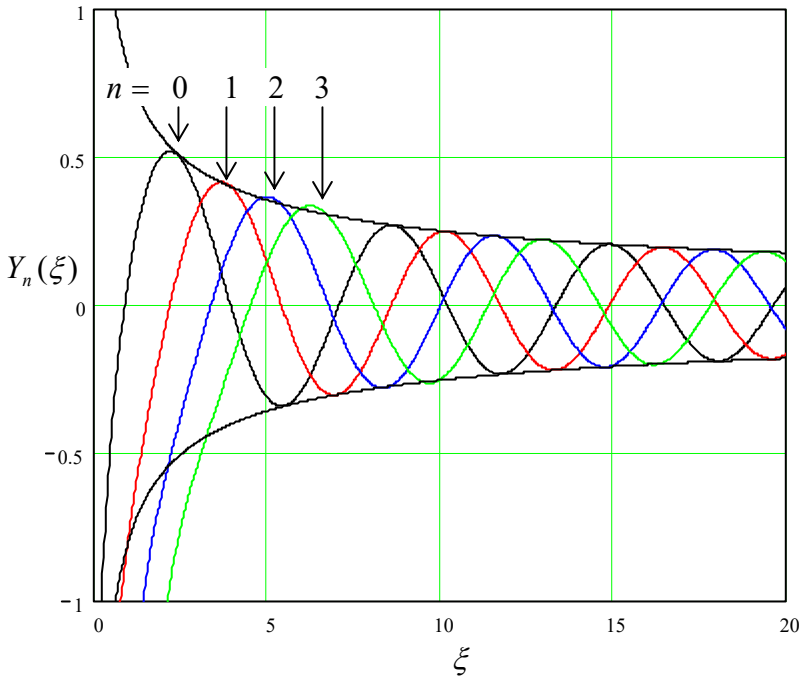



Fig. 2.19. A few Bessel functions of the second kind (a.k.a. the Weber functions, a.k.a. the Neumann functions).

The plots show that their asymptotic behavior is very similar to that of the functions $J_n(\xi)$:

$$Y_n(\xi) \to \left(\frac{2}{\pi\xi}\right)^{1/2} \sin\left(\xi - \frac{\pi}{4} - \frac{n\pi}{2}\right), \qquad \text{for } \xi \to \infty, \tag{2.151}$$

but with the phase shift necessary to make these Bessel functions orthogonal to those of the first order – cf. Eq. (135). However, for small values of argument $\xi$, the Bessel functions of the second kind behave completely differently from those of the first kind:

$$Y_n(\xi) \to \begin{cases} \dfrac{2}{\pi}\left(\ln\dfrac{\xi}{2} + \gamma\right), & \text{for } n = 0, \\ -\dfrac{(n-1)!}{\pi}\left(\dfrac{\xi}{2}\right)^{-n}, & \text{for } n \neq 0, \end{cases} \tag{2.152}$$

where $\gamma$ is the so-called *Euler constant*, defined as follows:

[50] Sometimes, they are called the *Neumann* functions and denoted as $N_\nu(\xi)$.

$$\gamma \equiv \lim_{n\to\infty}\left(1+\frac{1}{2}+\frac{1}{3}+...+\frac{1}{n}-\ln n\right)\approx 0.577157\ldots \tag{2.153}$$

As Eqs. (152) and Fig. 19 show, the functions $Y_n(\xi)$ diverge at $\xi \to 0$ and hence cannot describe the behavior of any physical variable, in particular the electrostatic potential.

One may wonder: if this is true, when do we need these functions in physics? Figure 20 shows an example of a simple boundary problem of electrostatics, whose solution by the variable separation method involves both functions $J_n(\xi)$ and $Y_n(\xi)$.

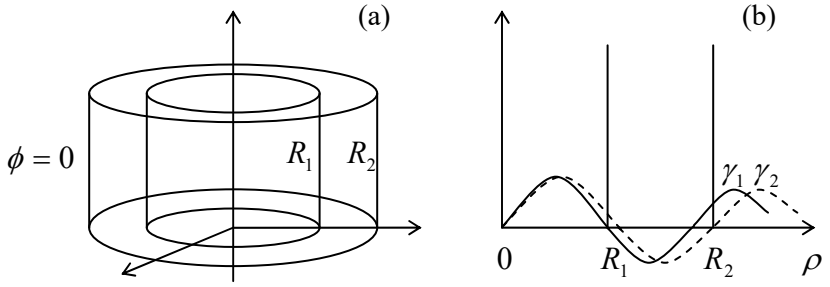



Fig. 2.20. A simple boundary problem that cannot be solved using just one kind of Bessel functions.

Here two round, conducting coaxial cylindrical tubes are kept at the same (say, zero) potential, but at least one of two lids has a different potential. The problem is almost completely similar to that discussed above (Fig. 17), but now we need to find the potential distribution in the free space between the tubes, i.e. for $R_1 < \rho < R_2$. If we use the same variable separation as in the simpler counterpart problem, we need the radial functions $\mathcal{R}(\rho)$ to satisfy two zero boundary conditions: at $\rho = R_1$ and $\rho = R_2$. With the Bessel functions of just the first kind, $J_n(\gamma\rho)$, it is impossible to do, because the two boundaries would impose two independent (and generally incompatible) conditions, $J_n(\gamma R_1) = 0$, and $J_n(\gamma R_2) = 0$, on one "stretching parameter" $\gamma$. The existence of the Bessel functions of the second kind immediately saves the day, because if the radial function solution is represented as a linear combination,

$$\mathcal{R} = c_J J_n(\gamma\rho) + c_Y Y_n(\gamma\rho), \tag{2.154}$$

two zero boundary conditions give two equations for $\gamma$ and the ratio $c \equiv c_Y/c_J$.[51] (Due to the oscillating character of both Bessel functions, these conditions would be typically satisfied by an infinite set of discrete pairs $\{\gamma, c\}$.) Note, however, that generally none of these pairs would correspond to zeros of either $J_n$ or $Y_n$, so having an analog of Table 1 for the latter function would not help much. Hence, even the simplest problems of this kind (like the one shown in Fig. 20) typically require the numerical solution of transcendental algebraic equations.

[51] A pair of independent linear functions, used for the representation of the general solution of the Bessel equation, may be also chosen differently, using the so-called *Hankel functions*

$$H_n^{(1,2)}(\xi) \equiv J_n(\xi) \pm iY_n(\xi).$$

For representing the general solution of Eq. (130), this alternative is completely similar, for example, to using the pair of complex functions $\exp\{\pm i\alpha x\} \equiv \cos\alpha x \pm i\sin\alpha x$ instead of the pair of real functions $\{\cos\alpha x, \sin\alpha x\}$ for the representation of the general solution of Eq. (89) for $X(x)$.

In order to complete the discussion of variable separation in the cylindrical coordinates, one more issue to address is the so-called *modified Bessel functions*: of the *first kind*, $I_\nu(\xi)$, and of the *second kind*, $K_\nu(\xi)$. They are two linearly independent solutions of the *modified Bessel equation*,

$$\frac{d^2\mathcal{R}}{d\xi^2}+\frac{1}{\xi}\frac{d\mathcal{R}}{d\xi}-\left(1+\frac{\nu^2}{\xi^2}\right)\mathcal{R}=0\,, \tag{2.155}$$

Modified Bessel equation

which differs from Eq. (130) "only" by the sign of one of its terms. Figure 21 shows a simple problem that leads (among many others) to this equation: a round thin conducting cylindrical pipe is sliced, perpendicular to its axis, to rings of equal height $h$, which are kept at equal but sign-alternating potentials.

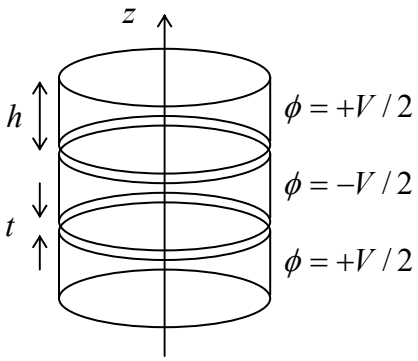



Fig. 2.21. A typical boundary problem whose solution may be conveniently described in terms of the modified Bessel functions.

If the system is very long (formally, infinite) in the $z$-direction, we may use the variable separation method for the solution of this problem, but now evidently need periodic (rather than exponential) solutions along the $z$-axis, i.e. linear combinations of $\sin kz$ and $\cos kz$ with various real values of the constant $k$. Separating the variables, we arrive at a differential equation similar to Eq. (129), but with the negative sign before the separation constant:

$$\frac{d^2\mathcal{R}}{d\rho^2}+\frac{1}{\rho}\frac{d\mathcal{R}}{d\rho}-(k^2+\frac{\nu^2}{\rho^2})\mathcal{R}=0\,. \tag{2.156}$$

The same radial coordinate's normalization, $\xi \equiv k\rho$, immediately leads us to Eq. (155), and hence (for $\nu = n$) to the modified Bessel functions $I_n(\xi)$ and $K_n(\xi)$.

Figure 22 shows the behavior of such functions, of a few lowest orders. One can see that at $\xi \to 0$ the behavior is virtually similar to that of the "usual" Bessel functions – cf. Eqs. (132) and (152), with $K_n(\xi)$ multiplied (by purely historical reasons) by an additional coefficient, $\pi/2$:

$$I_n(\xi)\to\frac{1}{n!}\left(\frac{\xi}{2}\right)^n, \qquad K_n(\xi)\to\begin{cases}-\left[\ln\left(\dfrac{\xi}{2}\right)+\gamma\right], & \text{for } n=0,\\ \dfrac{(n-1)!}{2}\left(\dfrac{\xi}{2}\right)^{-n}, & \text{for } n\neq 0,\end{cases} \tag{2.157}$$

However, the asymptotic behavior of the modified functions is very much different, with $I_n(x)$ exponentially growing, and $K_n(\xi)$ exponentially dropping at $\xi \to \infty$:

$$I_n(\xi) \to \left(\frac{1}{2\pi\xi}\right)^{1/2} e^{\xi}, \qquad K_n(\xi) \to \left(\frac{\pi}{2\xi}\right)^{1/2} e^{-\xi}. \tag{2.158}$$

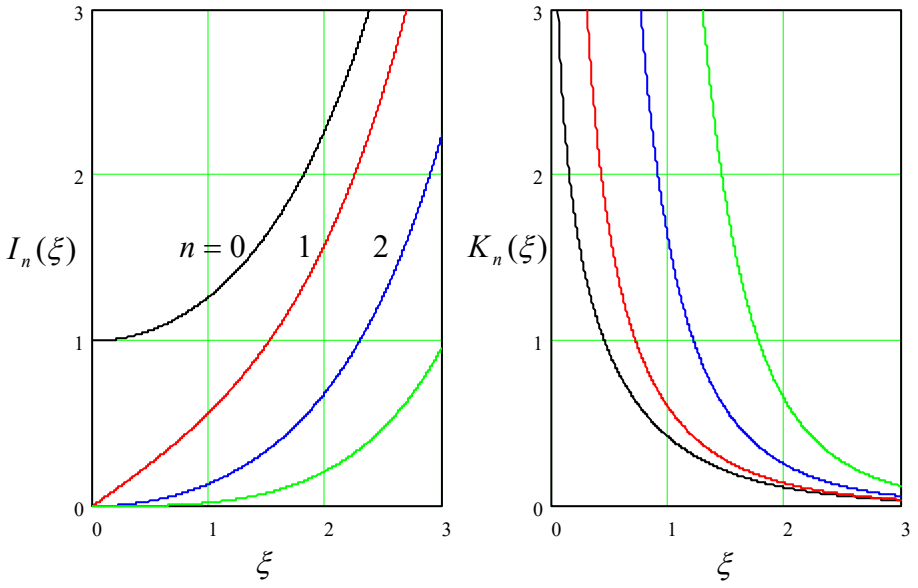



Fig. 2.22. The modified Bessel functions of the first kind (left panel) and the second kind (right panel).

This behavior is completely natural in the context of the problem shown in Fig. 21, in which the electrostatic potential may be represented as a sum of terms proportional to $I_n(\gamma\rho)$ inside the thin pipe, and of terms proportional to $K_n(\gamma\rho)$ outside it.

To complete our brief survey of the Bessel functions, let me note that all of them discussed so far may be considered as particular cases of *Bessel functions of the complex argument*, say $J_n(z)$ and $Y_n(z)$, or, alternatively, $H_n^{(1,2)}(z) \equiv J_n(z) \pm iY_n(z)$.[52] At that, the "usual" Bessel functions $J_n(\xi)$ and $Y_n(\xi)$ may be considered as the sets of values of these generalized functions on the real axis ($z = \xi$), while the modified functions as their particular case on the imaginary axis, i.e. at $z = i\xi$, also with real $\xi$:

$$I_\nu(\xi) = i^{-\nu} J_\nu(i\xi), \qquad K_\nu(\xi) = \frac{\pi}{2} i^{\nu+1} H_\nu^{(1)}(i\xi). \tag{2.159}$$

Moreover, this generalization of the Bessel functions to the whole complex plane $z$ enables the use of their values along other directions on that plane, for example under angles $\pi/4 \pm \pi/2$. As a result, one arrives at the so-called *Kelvin functions*:

$$\begin{aligned} &\mathrm{ber}_\nu \xi + i\,\mathrm{bei}_\nu \xi \equiv J_\nu(\xi e^{-i\pi/4}), \\ &\mathrm{ker}_\nu\, \xi + i\,\mathrm{kei}_\nu \xi \equiv i\frac{\pi}{2} H_\nu^{(1)}(\xi e^{-i3\pi/4}), \end{aligned} \tag{2.160}$$

which are also useful for some important problems in physics and engineering. Unfortunately, I do not have time/space to discuss these problems in this course.[53]

[52] These complex functions still obey the general relations (143) and (146), with $\xi$ replaced with $z$.

[53] In the QM part of this series we will run into the so-called *spherical Bessel functions* $j_n(\xi)$ and $y_n(\xi)$, which may be expressed via the Bessel functions of *semi-integer* orders. Surprisingly enough, these functions turn out to be simpler than $J_n(\xi)$ and $Y_n(\xi)$.

## 2. 8. Variable separation – spherical coordinates

The spherical coordinates are very important in physics, because of the (at least approximate) spherical symmetry of many physical objects – from nuclei and atoms, to water drops in clouds, to planets and stars. Let us again require each component $\phi_k$ of Eq. (84) to satisfy the Laplace equation. Using the full expression for the Laplace operator in spherical coordinates,[54] we get

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial \phi_k}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\phi_k}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2\phi_k}{\partial\varphi^2}=0\,. \tag{2.161}$$

Let us look for a solution of this equation in the following variable-separated form:

$$\phi_k=\frac{\mathcal{R}(r)}{r}\mathcal{P}(\cos\theta)\mathcal{F}(\varphi)\,, \tag{2.162}$$

Separating the variables one by one, starting from $\varphi$, just like this has been done in cylindrical coordinates, we get the following equations for the partial functions participating in this solution:

$$\frac{d^2\mathcal{R}}{dr^2}-\frac{l(l+1)}{r^2}\mathcal{R}=0\,, \tag{2.163}$$

$$\frac{d}{d\xi}\left[(1-\xi^2)\frac{d\mathcal{P}}{d\xi}\right]+\left[l(l+1)-\frac{\nu^2}{1-\xi^2}\right]\mathcal{P}=0\,, \tag{2.164}$$

$$\frac{d^2\mathcal{F}}{d\varphi^2}+\nu^2\mathcal{F}=0\,, \tag{2.165}$$

where $\xi\equiv\cos\theta$ is a new variable used in lieu of $\theta$ (so $-1\le\xi\le+1$), while $\nu^2$ and $l(l+1)$ are the separation constants. (The reason for the selection of the latter one in this form will be clear in a minute.)

One can see that Eq. (165) is very simple, and is absolutely similar to the Eq. (107) we have got for the polar and cylindrical coordinates. Moreover, the equation for the radial functions is *simpler* than in the cylindrical coordinates. Indeed, let us look for its partial solution in the form $cr^\alpha$ – just as we have done with Eq. (106). Plugging this solution into Eq. (163), we immediately get the following condition on the parameter $\alpha$:

$$\alpha(\alpha-1)=l(l+1)\,. \tag{2.166}$$

This quadratic equation has two roots, $\alpha=l+1$ and $\alpha=-l$, so the general solution of Eq. (163) is

$$\mathcal{R}=a_l r^{l+1}+\frac{b_l}{r^l}\,. \tag{2.167}$$

However, the general solution of Eq. (164) (called either the *general* or *associated Legendre equation*) cannot be expressed via what is usually called elementary functions.[55] Let us start its discussion from the axially-symmetric case when $\partial\phi/\partial\varphi=0$. This means $\mathcal{F}(\varphi)$ = const, and thus $\nu=0$, so Eq. (164) is reduced to the so-called *Legendre differential equation*:

[54] See, e.g., MA Eq. (10.9).

[55] Again, there is no generally accepted line between the "elementary" and "special" functions.

Legendre equation

$$\frac{d}{d\xi}\left[(1-\xi^2)\frac{d\mathcal{P}}{d\xi}\right]+l(l+1)\mathcal{P}=0\,. \tag{2.168}$$

One can readily verify that the solutions of this equation for integer values of $l$ are specific (*Legendre*) polynomials[56] that may be described by the following *Rodrigues' formula*:

Legendre polynomials

$$\mathcal{P}_l(\xi)=\frac{1}{2^l\,l!}\frac{d^l}{d\xi^l}(\xi^2-1)^l,\quad \text{with } l=0,1,2,\ldots. \tag{2.169}$$

According to this formula, the first few Legendre polynomials are pretty simple:

$$\begin{aligned}
\mathcal{P}_0(\xi)&=1,\\
\mathcal{P}_1(\xi)&=\xi,\\
\mathcal{P}_2(\xi)&=\frac{1}{2}\left(3\xi^2-1\right),\\
\mathcal{P}_3(\xi)&=\frac{1}{2}\left(5\xi^3-3\xi\right),\\
\mathcal{P}_4(\xi)&=\frac{1}{8}\left(35\xi^4-30\xi^2+3\right),\ldots,
\end{aligned} \tag{2.170}$$

though such explicit expressions become more and more bulky as $l$ is increased. As Fig. 23 shows, all these polynomials, which are defined on the [-1, +1] segment, end at the same point: $\mathcal{P}_l(+1) = +1$, while starting either at the same point or at the opposite point: $\mathcal{P}_l(-1) = (-1)^l$. Between these two endpoints, the $l^{\text{th}}$ Legendre polynomial has $l$ zeros. It is straightforward to use Eq. (169) to prove that these polynomials form a full, orthogonal set of functions, with the following normalization rule:

$$\int_{-1}^{+1}\mathcal{P}_l(\xi)\mathcal{P}_{l'}(\xi)d\xi=\frac{2}{2l+1}\delta_{ll'}, \tag{2.171}$$

so any function $f(\xi)$ defined on the segment [-1, +1] may be represented as a unique series over the polynomials.[57]

Thus, taking into account the additional division by $r$ in Eq. (162), the general solution of any axially symmetric Laplace problem may be represented as

Variable separation in spherical coordinates (for axial symmetry)

$$\phi(r,\theta)=\sum_{l=0}^{\infty}\left(a_l r^l+\frac{b_l}{r^{l+1}}\right)\mathcal{P}_l(\cos\theta)\,. \tag{2.172}$$

Note a strong similarity between this solution and Eq. (112) for the 2D Laplace problem in the polar coordinates. However, besides the difference in the angular functions, there is also a difference (by one) in the power of the second radial function, and this difference immediately shows up in problem solutions.

[56] Just for reference: if $l$ is not an integer, the general solution of Eq. (2.168) may be represented as a linear combination of the so-called *Legendre functions* (not polynomials!) *of the first and second kind*, $\mathcal{P}_l(\xi)$ and $\mathcal{Q}_l(\xi)$.

[57] This is why, at least for the purposes of this course, there is no good reason for pursuing (more complicated) solutions to Eq. (168) for non-integer values of $l$, mentioned in the previous footnote.

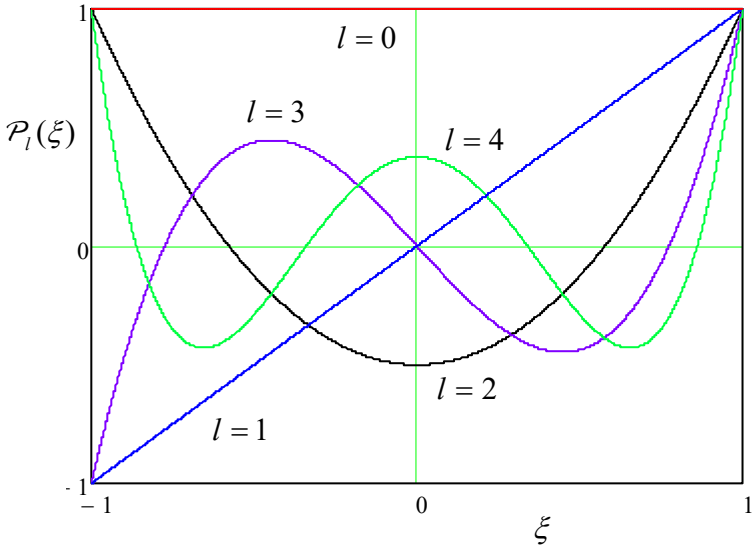



Fig. 2.23. A few lowest Legendre polynomials $\mathcal{P}_l(\xi)$.

Indeed, let us solve a problem similar to that shown in Fig. 15: find the electric field around a conducting sphere of radius $R$, placed into an initially uniform external field $\mathbf{E}_0$ (whose direction I will now take for the $z$-axis) – see Fig. 24a.

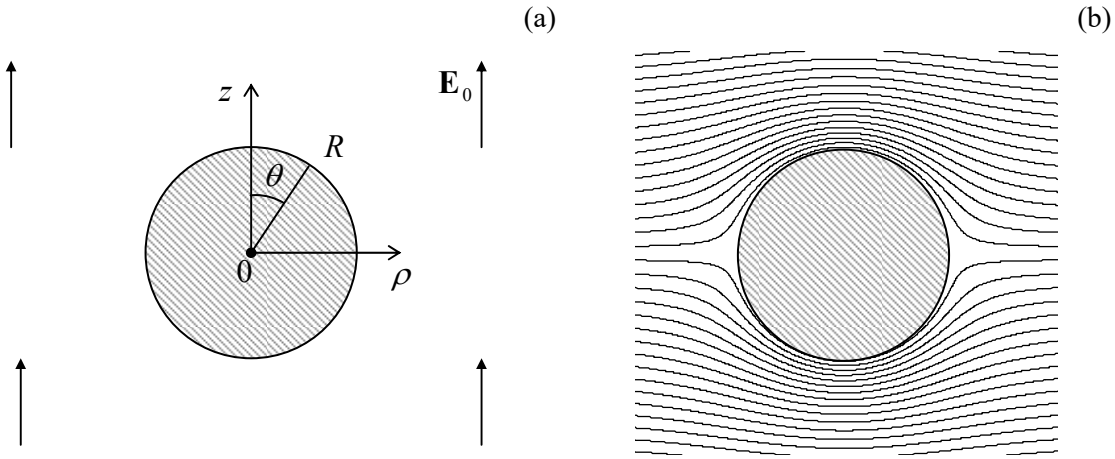



Fig. 2.24. Conducting sphere in a uniform electric field: (a) the problem's geometry, and (b) the equipotential surface pattern given by Eq. (176). The pattern is qualitatively similar but quantitatively different from that for the conducting cylinder in a perpendicular field – cf. Fig. 15.

If we select the arbitrary constant in the electrostatic potential so that $\phi|_{z=0} = 0$, then in Eq. (172) we should take $a_0 = b_0 = 0$. Now, just as has been argued for the cylindrical case, at $r >> R$ the potential should approach that of the uniform field:

$$\phi \to -E_0 z = -E_0 r \cos\theta , \tag{2.173}$$

so in Eq. (172), only one of the coefficients $a_l$ survives: $a_l = -E_0\delta_{l,1}$. As a result, from the boundary condition on the surface, $\phi(R,\theta) = 0$, we get the following equation for the coefficients $b_l$:

$$\left(-E_0 R + \frac{b_1}{R^2}\right)\cos\theta + \sum_{l\geq 2}\frac{b_l}{R^{l+1}}\mathcal{P}_l(\cos\theta) = 0 . \tag{2.174}$$

Now repeating the argumentation that led to Eq. (117), we may conclude that Eq. (174) is satisfied if

$$b_l = E_0 R^3 \delta_{l,1}, \tag{2.175}$$

so, finally, Eq. (172) is reduced to

$$\phi = -E_0\left(r - \frac{R^3}{r^2}\right)\cos\theta . \tag{2.176}$$

This distribution, shown in Fig. 24b, is very similar to Eq. (117) for the cylindrical case (cf. Fig. 15b, with the account for a different plot orientation), but with a different power of the radius in the second term. This difference leads to a quantitatively different distribution of the surface electric field:

$$E_n = -\frac{\partial\phi}{\partial r}\Big|_{r=R} = 3E_0\cos\theta , \tag{2.177}$$

so its maximal value is a factor of 3 (rather than 2) larger than the external field.

Now let me briefly (mostly just for the reader's reference) mention the Laplace equation solutions in the general case – with no axial symmetry. If the conductor-free space surrounds the origin from all sides, the solutions to Eq. (165) have to be $2\pi$-periodic, and hence $\nu = n = 0, \pm1, \pm2,\ldots$ Mathematics says that Eq. (164) with integer $\nu = n$ and a fixed integer $l$ has a solution only for a limited range of $n$:[58]

$$-l \le n \le +l . \tag{2.178}$$

These solutions are called a*ssociated Legendre functions* (generally, they are *not* polynomials). For $n \ge 0$, these functions may be defined via the Legendre polynomials, using the following formula:[59]

$$\mathcal{P}_l^n(\xi) = (-1)^n (1-\xi^2)^{n/2}\frac{d^n}{d\xi^n}\mathcal{P}_l(\xi) . \tag{2.179}$$

On the segment $\xi \in [-1, +1]$, each set of the associated Legendre functions with a fixed index $n$ and non-negative values of $l$ form a full, orthogonal set, with the normalization relation,

$$\int_{-1}^{+1}\mathcal{P}_l^n(\xi)\mathcal{P}_{l'}^n(\xi)d\xi = \frac{2}{2l+1}\frac{(l+n)!}{(l-n)!}\delta_{ll'}, \tag{2.180}$$

that is evidently a generalization of Eq. (171).

Since these relations may seem a bit intimidating, let me write down explicit expressions for a few $\mathcal{P}_l^n(\cos\theta)$ with the three lowest values of $l$ and $n \ge 0$, which are most important for applications.

$$l = 0: \qquad \mathcal{P}_0^0(\cos\theta) = 1; \tag{2.181}$$

[58] In quantum mechanics, the letter $n$ is typically reserved for the "principal quantum number", while the azimuthal functions are numbered by index $m$. However, here I will keep using $n$ as their index because, for this course's purposes, this seems more logical, in view of the similarity of the spherical and cylindrical functions.

[59] Note that some texts use different choices for the front factor (called the *Condon-Shortley phase*) in the functions $\mathcal{P}_l^m$, which do not affect the final results for the spherical harmonics $Y_l^m$.

$$l=1: \quad \begin{cases} \mathcal{P}_1^0(\cos\theta)=\cos\theta, \\ \mathcal{P}_1^1(\cos\theta)=-\sin\theta; \end{cases} \tag{2.182}$$

$$l=2: \begin{cases} \mathcal{P}_2^0(\cos\theta)=\left(3\cos^2\theta-1\right)/2, \\ \mathcal{P}_2^1(\cos\theta)=-2\sin\theta\cos\theta, \\ \mathcal{P}_2^2(\cos\theta)=-3\cos^2\theta. \end{cases} \tag{2.183}$$

The reader should agree there is not much to fear in these functions – they are just certain sums of products of $\cos\theta \equiv \xi$ and $\sin\theta \equiv (1-\xi^2)^{1/2}$. Fig. 25 shows the plots of a few lowest functions $\mathcal{P}_l^n(\xi)$.

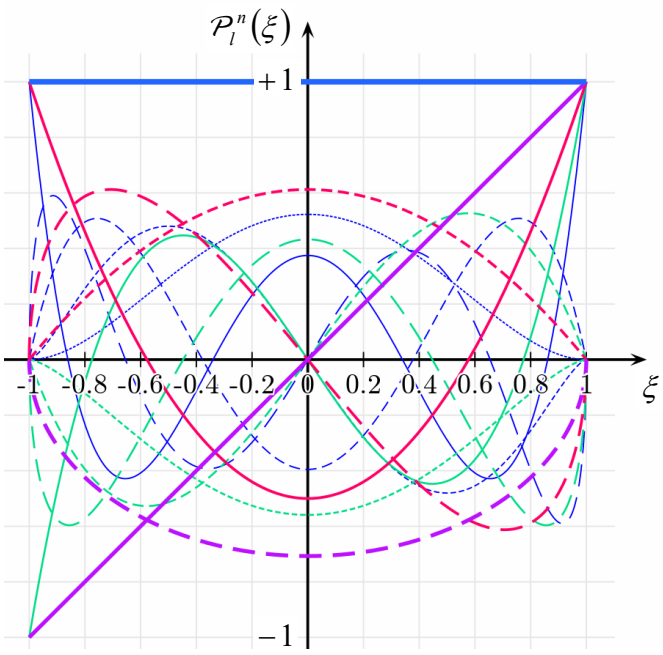



Fig. 2.25. A few lowest associated Legendre functions. (Adapted from an original by *Geek3*, available at https://en.wikipedia.org/wiki/Associated_Legendre_polynomials, under the GNU Free Documentation License.)

Using the associated Legendre functions, the general solution (162) to the Laplace equation in the spherical coordinates may be expressed as

Variable separation in spherical coordinates (general case)

$$\phi(r,\theta,\varphi)=\sum_{l=0}^{\infty}\left(a_l r^l+\frac{b_l}{r^{l+1}}\right)\sum_{n=0}^{l}\mathcal{P}_l^n(\cos\theta)\mathcal{T}_n(\varphi), \qquad \mathcal{T}_n(\varphi)=c_n\cos n\varphi+s_n\sin n\varphi\,. \tag{2.184}$$

Since the difference between the angles $\theta$ and $\varphi$ is somewhat artificial, physicists prefer to think not in terms of the functions $\mathcal{P}$ and $\mathcal{T}$ in separation, but directly about their products that participate in this solution.[60]

[60] In quantum mechanics, it is more convenient to use a slightly different alternative set of basic functions of the same problem, namely the following complex functions called the *spherical harmonics*:

$$Y_l^n(\theta,\varphi)\equiv\left[\frac{2l+1}{4\pi}\frac{(l-n)!}{(l+n)!}\right]^{1/2}\mathcal{P}_l^n(\cos\theta)e^{in\varphi},$$

which are defined for both positive and negative $n$ (within the limits $-l \le n \le +l$) – see, e.g., QM Secs. 3.6 and 5.6. (Note again that in that field, our index $n$ is traditionally denoted as $m$, and called the *magnetic quantum number*.)

As a rare exception for my courses, to save time I will skip giving an example of using the associated Legendre functions in electrostatics, because quite a few examples of these functions' applications will be given in the quantum mechanics part of this series.

## 2.9. Charge images

So far, we have discussed various methods of solution of the *Laplace* boundary problem (35). Let us now move on to the discussion of its generalization, the *Poisson* equation (1.41). We need it when besides conductors, we also have stand-alone charges with a known spatial distribution $\rho(\mathbf{r})$. (Its discussion will also allow us, better equipped, to revisit the Laplace problem in the next section.)

Let us start with a somewhat limited, but very useful *charge image* (or "image charge") *method*. Consider a very simple problem: a single point charge near a conducting half-space – see Fig. 26.

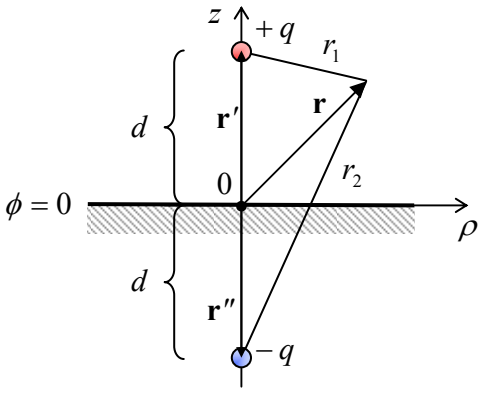



Fig. 2.26. The simplest problem readily solvable by the charge image method. The points' colors are used, as before, to denote the charges of the original (red) and opposite (blue) sign.

Let us prove that its solution, above the conductor's surface ($z \geq 0$), may be represented as:

$$\phi(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0}\left(\frac{q}{r_1} - \frac{q}{r_2}\right) \equiv \frac{q}{4\pi\varepsilon_0}\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|} - \frac{1}{|\mathbf{r}-\mathbf{r}''|}\right), \tag{2.185}$$

or in a more explicit form, using the cylindrical coordinates shown in Fig. 26:

$$\phi(\mathbf{r}) = \frac{q}{4\pi\varepsilon_0}\left\{\frac{1}{\left[\rho^2 + (z-d)^2\right]^{1/2}} - \frac{1}{\left[\rho^2 + (z+d)^2\right]^{1/2}}\right\}, \tag{2.186}$$

where $\rho$ is the distance of the field observation point $\mathbf{r}$ from the "vertical" line on which the charge is located. Indeed, this solution satisfies both the boundary condition $\phi = 0$ at the surface of the conductor ($z = 0$), and the Poisson equation (1.41), with the single $\delta$-functional source at point $\mathbf{r}' = \{0, 0, +d\}$ on its right-hand side, because the second singularity of the solution, at point $\mathbf{r}'' = \{0, 0, -d\}$, is outside the region of the solution's validity ($z \geq 0$). Physically, this solution may be interpreted as the sum of the fields of the actual charge ($+q$) at point $\mathbf{r}'$, and an equal but opposite charge ($-q$) at the "mirror image" point $\mathbf{r}''$ (Fig. 26). This is the basic idea of the charge image method.

Before moving on to more complex problems, let us discuss the situation shown in Fig. 26 in a little bit more detail, due to its fundamental importance. First, we can use Eqs. (3) and (186) to calculate the surface charge density:

$$\sigma = -\varepsilon_0 \frac{\partial \phi}{\partial z}\Big|_{z=0} = -\frac{q}{4\pi}\frac{\partial}{\partial z}\left\{\frac{1}{\left[\rho^2+(z-d)^2\right]^{1/2}} - \frac{1}{\left[\rho^2+(z+d)^2\right]^{1/2}}\right\}_{z=0} = -\frac{q}{4\pi}\frac{2d}{\left(\rho^2+d^2\right)^{3/2}}. \quad (2.187)$$

From this, the total surface charge is

$$Q = \int_S \sigma d^2 r = 2\pi \int_0^\infty \sigma(\rho)\rho d\rho = -\frac{q}{2}\int_0^\infty \frac{2d}{\left(\rho^2+d^2\right)^{3/2}}\rho d\rho. \quad (2.188)$$

This integral may be easily worked out using the substitution $\xi \equiv \rho^2/d^2$ (giving $d\xi = 2\rho d\rho/d^2$):

$$Q = -\frac{q}{2}\int_0^\infty \frac{d\xi}{(\xi+1)^{3/2}} = -q. \quad (2.189)$$

This result is very natural: the conductor brings as much surface charge from its interior to the surface as necessary to fully compensate for the initial charge ($+q$) and hence kill the electric field at large distances as efficiently as possible, hence reducing the total electrostatic energy (1.65) to the lowest possible value.

For a better feeling of this *polarization charge* of the surface, let us take our calculations to the extreme – to the $q$ equal to one elementary change $e$, and place a particle with this charge (for example, a proton) at a macroscopic distance – say 1 m – from the conductor's surface. Then, according to Eq. (189), the total polarization charge of the surface equals that of an electron, and according to Eq. (187), its spatial extent is of the order of $d^2 = 1$ m$^2$. This means that if we consider a much smaller part of the surface, $\Delta A << d^2$, its polarization charge magnitude $\Delta Q = \sigma \Delta A$ is much *less than one electron*! For example, Eq. (187) shows that the polarization charge of quite a macroscopic area $\Delta A = 1$ cm$^2$ right under the initial charge ($\rho = 0$) is $e\Delta A/2\pi d^2 \approx 1.6\times10^{-5}\ e$. Can this be true, or our theory is somehow limited to the charges $q$ much larger than $e$? (After all, the theory is substantially based on the approximate macroscopic model (1); maybe it is the culprit?)

Surprisingly enough, the answer to this question has become clear (at least to some physicists :-) only as late as in the mid-1980s when several experiments demonstrated, and theorists accepted (some of them rather grudgingly) that the usual polarization charge formulas are valid for elementary charges as well, i.e., such the polarization charge $\Delta Q$ of a macroscopic surface area may differ from a multiple of $e$. The underlying reason for this paradox is the physical nature of the polarization charge of a conductor's surface: as was discussed in Sec. 1, it is due not to new charged particles brought into the conductor (such charge would be in fact a multiple of $e$), but to a small *shift* of the free charges of a conductor by a very small distance from their equilibrium positions that they had in the absence of the external field induced by charge $q$. This shift is not quantized, at least on the scale relevant to our problem, and hence neither is $\Delta Q$.

This understanding has paved the way for the invention and experimental demonstration of several new devices including so-called *single-electron transistors*,[61] which may be used, in particular, for ultrasensitive measurement of polarization charges as small as $\sim10^{-6}\ e$. Another important class of single-electron devices is the dc and ac current standards based on the fundamental relation $I = -ef$,

[61] Actually, this term (for which the author of these notes may be blamed :-) is misleading: the operation of the "single-electron transistor" is based on the interplay of discrete charges (multiples of $e$) transferred between conductors, and *sub*-single-electron polarization charges – see, e.g., K. Likharev, *Proc. IEEE* **87**, 606 (1999).

where $I$ is the dc current carried by electrons transferred with the frequency $f$. The experimentally achieved[62] relative accuracy of such standards is of the order of $10^{-7}$, and is not too far from that provided by the competing approach based on a combination of the Josephson effect and the quantum Hall effect.[63]

Second, let us find the potential energy $U$ of the charge-to-surface interaction. For that, we may use the value of the electrostatic potential (185) at the point of the charge itself ($\mathbf{r} = \mathbf{r}'$), of course ignoring the infinite potential created by the real charge, so the remaining potential is that of the image charge

$$\phi_{\text{image}}(\mathbf{r}') = -\frac{1}{4\pi\varepsilon_0}\frac{q}{2d}. \tag{2.190}$$

Looking at the electrostatic potential's definition given by Eq. (1.31), it may be tempting to immediately write $U = q\phi_{\text{image}} = -(1/4\pi\varepsilon_0)(q^2/2d)$ [WRONG!], but this would be incorrect. The reason is that the potential $\phi_{\text{image}}$ is not independent of $q$, but is actually induced by this charge. This is why the correct approach is to calculate $U$ from Eq. (1.61), with just one term:

$$U = \frac{1}{2}q\phi_{\text{image}} = -\frac{1}{4\pi\varepsilon_0}\frac{q^2}{4d}, \tag{2.191}$$

giving twice lower energy than the wrong result cited above. To double-check Eq. (191), and also get a better feeling of the factor ½ that distinguishes it from the wrong guess, we can calculate $U$ as the integral of the force exerted on the charge by the conductor's surface charge (i.e., in our formalism, by the image charge):

$$U = -\int_{\infty}^{d} F(z)dz = \frac{1}{4\pi\varepsilon_0}\int_{\infty}^{d}\frac{q^2}{(2z)^2}dz = -\frac{1}{4\pi\varepsilon_0}\frac{q^2}{4d}. \tag{2.192}$$

This calculation clearly accounts for the gradual build-up of the force $F$, as the real charge is being brought from afar (where we have opted for $U = 0$) toward the surface.

This result has several important applications. For example, let us plot the electrostatic energy $U$ of an electron, i.e. a particle with charge $q = -e$, near a metallic surface, as a function of $d$. For that, we may use Eq. (191) until our macroscopic model (1) becomes invalid, and $U$ transitions to some negative constant value ($-\psi$) inside the conductor – see Fig. 27a. Since our calculation was for an electron with zero potential energy at infinity, at relatively low temperatures, $k_B T << \psi$, electrons in metals may occupy only the states with energies below $-\psi$ (the so-called *Fermi level*[64]). The positive constant $\psi$ is called the *workfunction* because it describes the smallest work needed to remove the electron from a metal. As was discussed in Sec. 1, in good metals the electric field screening takes place at interatomic distances $a_0 \sim 10^{-10}$ m. Plugging $d = 1\times10^{-10}$ m and $q = -e \approx -1.6\times10^{-19}$ C into Eq. (191), we get $\psi \approx 6\times10^{-19}$ J $\approx 3.5$ eV. This crude estimate is in surprisingly good agreement with the experimental values of the workfunction, ranging between 4 and 5 eV for most metals.[65]

[62] See, e.g., M. Keller *et al.*, *Appl. Phys. Lett.* **69**, 1804 (1996) ; F. Stein *et al.*, Metrologia **54**, 1 (2017).
[63] J. Brun-Pickard *et al.*, *Phys. Rev. X* **6**, 041051 (2016).
[64] More discussion of these states may be found in SM Secs. 3.3 and 6.3.
[65] More discussion of the workfunction, and its effect on the electrons' kinetics, is given in SM Sec. 6.3.

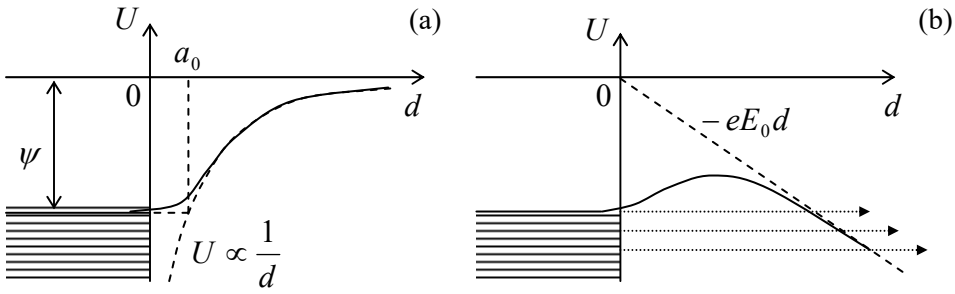



Fig. 2.27. (a) The origin of the workfunction, and (b) the field emission of electrons (schematically).

Next, let us consider the effect of an additional uniform external electric field $\mathbf{E}_0$ applied normally to a metallic surface, on this potential profile. For that, we may the potential energy that the field gives to the electron at distance $d$ from the surface, $U_{\text{ext}} = -eE_0 d$, to that created by the image charge. (As we know from Eq. (1.53), since the field $\mathbf{E}_0$ is independent of the electron's position, its recalculation into the potential energy does not require the coefficient ½.) As a result, the potential energy of an electron near the surface becomes

$$U(d) = -eE_0 d - \frac{1}{4\pi\varepsilon_0}\frac{e^2}{4d}, \qquad \text{for } d >> a_0, \tag{2.193}$$

with a similar crossover to $U = -\psi$ inside the conductor – see Fig. 27b. One can see that at the appropriate sign, and a sufficient magnitude of the applied field, it lowers the potential barrier that prevents electrons from leaving the conductor. At $E_0 \sim \psi/a_0$ (for metals, $\sim 10^{10}$ V/m), this suppression becomes so strong that electrons with energies at, and just below the Fermi level start quantum-mechanical tunneling through the remaining thin barrier. This is the *field electron emission* (or just "field emission") effect, which is used in vacuum electronics to provide efficient cathodes that do not require heating to high temperatures.[66]

Returning to the basic electrostatics, let us find some other conductor geometries where the method of charge images may be effectively applied. First, let us consider a right-angle corner (Fig. 28a). Reflecting the initial charge in the vertical plane, we get the image shown in the top left corner of that panel. This image makes the boundary condition $\phi$ = const satisfied on the vertical surface of the corner. However, for the same to be true on the horizontal surface, we have to reflect *both* the initial charge *and* the image charge in the horizontal plane, flipping their signs. The final configuration of four charges, shown in Fig. 28a, satisfies all boundary conditions. The resulting potential distribution may be readily written as an evident generalization of Eq. (185). From it, the electric field and electric charge distributions, and the potential energy and forces acting on the charge may be calculated exactly as above – an easy exercise left for the reader.

Next, consider a corner with the angle $\pi/4$ (Fig. 28b). Here we need to repeat the reflection operation not two but four times before we arrive at the final pattern of eight positive and negative charges. (Any attempt to continue this process would lead to overlap with the already existing charges.)

[66] The practical use of such "cold" cathodes is affected by the fact that, as it follows from our discussion in Sec. 4, any nanoscale irregularity of a conducting surface (a protrusion, an atomic cluster, or even a single "adatom" stuck to it) may cause a strong increase of the local field well above the applied uniform field $E_0$, making the electron emission reproducibility and stability in time significant challenges. In addition, the impact-ionization effects may lead to avalanche-type electric breakdown at dc fields as low as $\sim 3\times 10^6$ V/m.

This reasoning may be readily extended to corners of angles $\beta = \pi/n$, with any integer $n$, which require $2n$ charges (including the initial one) to satisfy all the boundary conditions.

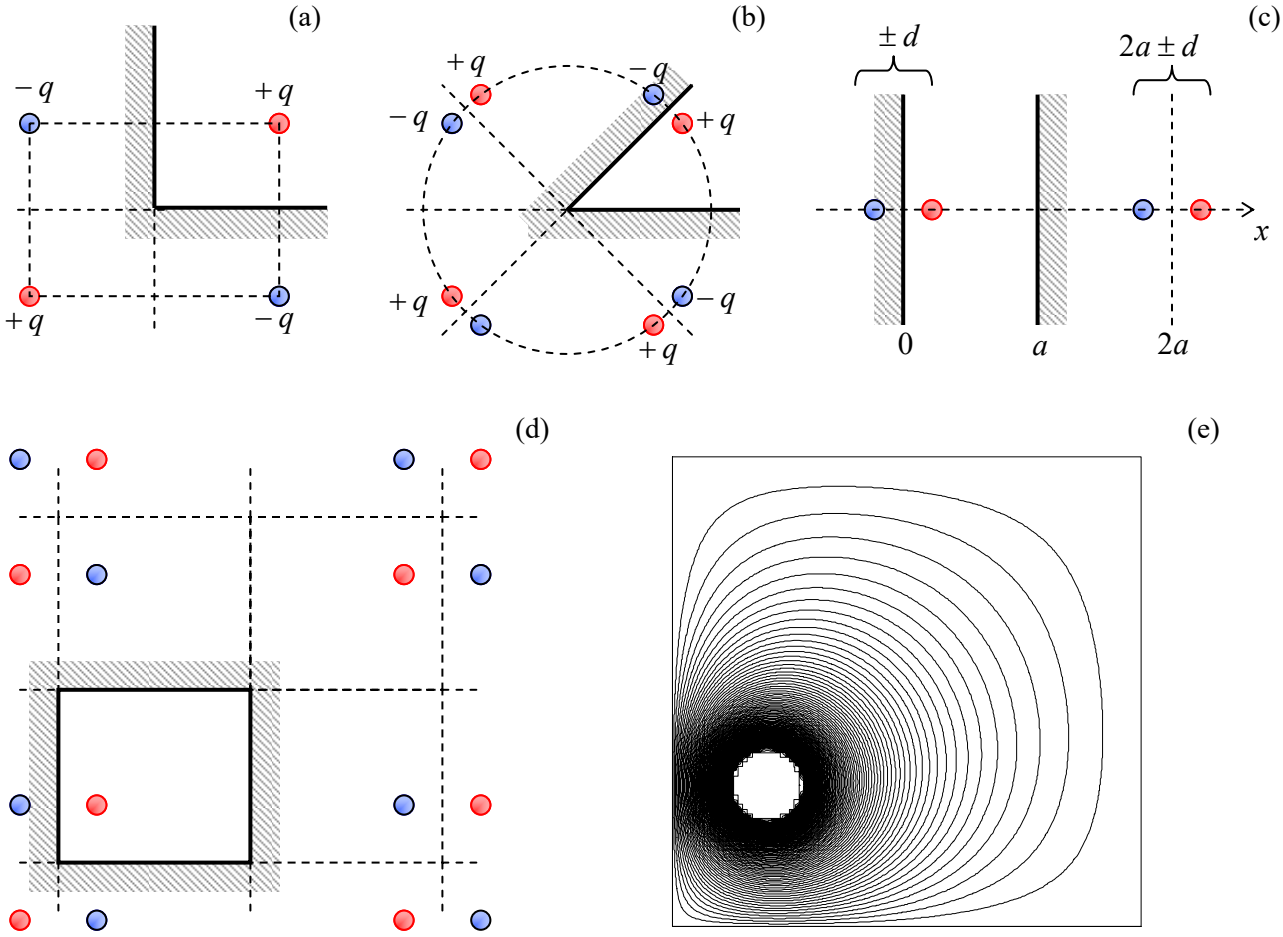



Fig. 2.28. The charge images for (a, b) the corners with angles $\pi/2$ and $\pi/4$, (c) a plane capacitor, and (d) a rectangular box; (e) typical equipotential surfaces for the last system.

Some configurations require an infinite number of images but are still tractable. The most important of them is a system of two parallel conducting surfaces, i.e. an unbiased plane capacitor of infinite area (Fig. 28c). Here the repeated reflection leads to an infinite system of charges $\pm q$ at points

$$x_j^{\pm} = 2aj \pm d\,, \tag{2.194}$$

where $d$ (with $0 < d < a$) is the position of the initial charge, and $j$ is an arbitrary integer. The resulting infinite sum for the potential of the real charge $q$, created by the field of its images,

$$\phi(d) = \frac{1}{4\pi\varepsilon_0}\left[-\frac{q}{2d} + \sum_{j\neq 0}\sum_{\pm}\frac{\pm q}{\left|d - x_j^{\pm}\right|}\right] \equiv -\frac{q}{4\pi\varepsilon_0}\left\{\frac{1}{2d} + \frac{d^2}{a^3}\sum_{j=1}^{\infty}\frac{1}{j\left[j^2 - (d/a)^2\right]}\right\}, \tag{2.195}$$

is converging (in its last form) very fast. For example, the exact value, $\phi(a/2) = -2\ln 2$ $(q/4\pi\varepsilon_0 a)$, differs by less than 5% from the approximation using just the first term of the sum.

The same method may be applied to 2D (cylindrical) and 3D rectangular conducting boxes that require, respectively, 2D or 3D infinite rectangular lattices of images; for example in a 3D box with sides $a$, $b$, and $c$, charges $\pm q$ are located at points (Fig. 28d)

$$\mathbf{r}_{jkl}^{\pm} = 2ja + 2kb + 2lc \pm \mathbf{r}', \tag{2.196}$$

where $\mathbf{r}'$ is the location of the initial (real) charge, and $j$, $k$, and $l$ are arbitrary integers. Figure 28e shows a typical result of the summation of the potentials of this charge set, including the real one, in a 2D box (within the plane of the real charge). One can see that the equipotential surfaces, concentric near the charge, are naturally leaning along the conducting walls of the box, which have to be equipotential.

Even more surprisingly, the image charge method works very efficiently not only for rectilinear geometries but also for spherical ones. Indeed, let us consider a point charge $q$ at distance $d$ from the center of a conducting, grounded sphere of radius $R$ (Fig. 29a), and try to satisfy the boundary condition $\phi = 0$ for the electrostatic potential on the sphere's surface using an imaginary charge $q'$ located at some point beyond the surface, i.e. inside the sphere.

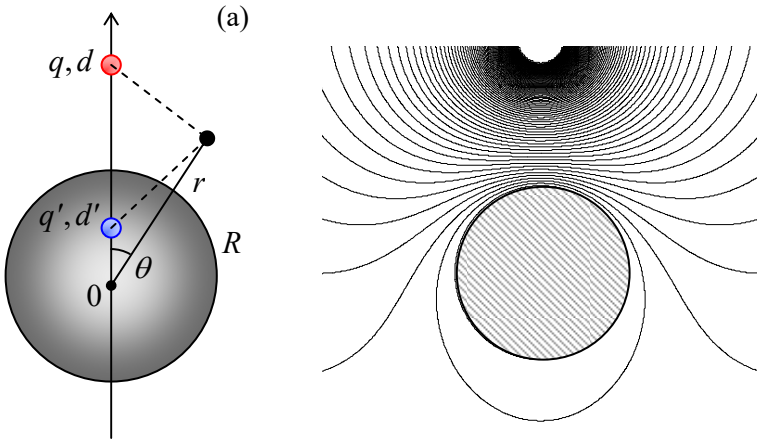



Fig. 2.29. Method of charge images for a conducting sphere: (a) the idea, and (b) the resulting potential distribution in the central plane containing the charge, for the particular case $d = 2R$.

From the problem's symmetry, it is clear that the point should be at the line passing through the real charge and the sphere's center, at some distance $d'$ from the center. Then the total potential created by the two charges at an arbitrary point of free space, i.e. at $r \geq R$ (Fig. 29a) is

$$\phi(r,\theta) = \frac{1}{4\pi\varepsilon_0}\left[\frac{q}{\left(r^2 + d^2 - 2rd\cos\theta\right)^{1/2}} + \frac{q'}{\left(r^2 + d'^2 - 2rd'\cos\theta\right)^{1/2}}\right]. \tag{2.197}$$

This expression shows that we can make the two involved fractions equal and opposite at all points on the sphere's surface (i.e. for any $\theta$ at $r = R$) if we take[67]

$$d' = \frac{R^2}{d}, \qquad q' = -\frac{R}{d}q. \tag{2.198}$$

Since the solution of any Poisson boundary problem is unique, Eqs. (197) and (198) give us the final solution for this problem. Fig. 29b shows a typical equipotential pattern following from this solution. It may be surprising how formulas that simple may describe such an elaborate field distribution.

[67] In geometry, such points with $dd' = R^2$, are referred to as the result of mutual *inversion* in a sphere of radius $R$.

Now we can calculate the total charge $Q$ on the grounded sphere's surface, induced by the external charge $q$. We could do this, as we have done for the conducting plane, by the brute-force integration of the surface charge density $\sigma = -\varepsilon_0 \partial\phi/\partial r|_{r=R}$. It is more elegant, however, to use the following Gauss law argument. Equality (197) is valid (at $r \geq R$) regardless of whether we are dealing with our real problem (charge $q$ and the conducting sphere) or with the equivalent charge configuration – with the point charges $q$ and $q'$, but no sphere at all. Hence, according to Eq. (1.16), the Gaussian integral over a surface with radius $r = R + 0$, and the total charge inside the sphere should be also the same. Hence we immediately get

$$Q = q' = -\frac{R}{d} q \,. \tag{2.199}$$

A similar argumentation may be used to calculate the charge-to-sphere interaction force:

$$F = qE_{\text{image}}(d) = q\frac{q'}{4\pi\varepsilon_0 (d-d')^2} = -\frac{q^2}{4\pi\varepsilon_0}\frac{R}{d}\frac{1}{(d-R^2/d)^2} = -\frac{q^2}{4\pi\varepsilon_0}\frac{Rd}{(d^2-R^2)^2}\,. \tag{2.200}$$

(Note that this expression is legitimate only at $d > R$.) At large distances, $d >> R$, this attractive force decreases as $1/d^3$. This unusual dependence arises because, as Eq. (199) specifies, the induced charge of the sphere, responsible for the force, is not constant but decreases as $1/d$. In the next chapter, we will see that such force is also typical for the interaction between a point charge and a *dipole*.

All previous formulas were for a sphere that is grounded to keep its potential equal to zero. But what if we keep the sphere galvanically insulated, so its *net charge* is fixed, for example, equals zero? Instead of solving this problem from scratch, let us use (again!) the almighty linear superposition principle. For that, we may add to the previous problem an additional charge, equal to $-Q = -q'$, to the sphere, and argue that this addition gives, at all points, an additional, spherically symmetric potential that does not depend on the potential induced by the external charge $q$, and was calculated in Sec. 1.2 – see Eq. (1.19). For the interaction force, such addition yields

$$F = \frac{qq'}{4\pi\varepsilon_0 (d-d')^2} + \frac{qq'}{4\pi\varepsilon_0 d^2} = -\frac{q^2}{4\pi\varepsilon_0}\left[\frac{Rd}{(d^2-R^2)^2} - \frac{R}{d^3}\right]. \tag{2.201}$$

At large distances, the two terms proportional to $1/d^3$ cancel each other, giving $F \propto 1/d^5$, so the potential energy of such interaction behaves as $U \propto 1/d^4$. Such a rapid force decay is due to the fact that the field of the uncharged sphere is equivalent to that of two (equal and opposite) induced charges $+q'$ and $-q'$, and the distance between them ($d - d' = d - R^2/d$) tends to zero at $d \to \infty$.

## 2.10. Green's functions

I have spent so much time/space discussing potential distributions created by a single point charge in various conductor geometries because for any of the geometries, the generalization of these results to the arbitrary distribution $\rho(\mathbf{r})$ of free charges is straightforward. Namely, if a single charge $q$, located at some point $\mathbf{r}'$, creates at point $\mathbf{r}$ the electrostatic potential

$$\phi(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0} qG(\mathbf{r},\mathbf{r}')\,, \tag{2.202}$$

then, due to the linear superposition principle, an arbitrary charge distribution creates the potential

$$\phi(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0}\sum_j q_j G(\mathbf{r},\mathbf{r}_j) = \frac{1}{4\pi\varepsilon_0}\int \rho(\mathbf{r}')G(\mathbf{r},\mathbf{r}')d^3r'. \quad (2.203)$$

Spatial Green's function

The function $G(\mathbf{r}, \mathbf{r}')$ is called the (spatial) *Green's function* – the notion very fruitful and hence popular in all fields of physics.[68] Evidently, as Eq. (1.35) shows, in the unlimited free space

$$G(\mathbf{r},\mathbf{r}') = \frac{1}{|\mathbf{r}-\mathbf{r}'|}, \quad (2.204)$$

i.e. the Green's function depends only on one scalar argument – the distance between the field-observation point $\mathbf{r}$ and the field-source (charge) point $\mathbf{r}'$. However, as soon as there are conductors around, the situation changes. In this course, I will only discuss Green's functions defined to vanish as soon as the radius-vector $\mathbf{r}$ points to the surface ($S$) of any conductor:[69]

$$G(\mathbf{r},\mathbf{r}')\Big|_{\mathbf{r}\in S} = 0. \quad (2.205)$$

With this definition, it is straightforward to deduce the Green's functions for the solutions of the last section's problems in which conductors were grounded, i.e. had potential $\phi = 0$. For example, for a semi-space $z \geq 0$ limited by a grounded conducting plane $z = 0$ (Fig. 26), Eq. (185) yields

$$G = \frac{1}{|\mathbf{r}-\mathbf{r}'|} - \frac{1}{|\mathbf{r}-\mathbf{r}''|}, \qquad \text{with } \boldsymbol{\rho}'' = \boldsymbol{\rho}' \text{ and } z'' = -z', \quad (2.206)$$

where $\boldsymbol{\rho}$ is the 2D radius vector. We see that in the presence of conductors (and, as we will see later, any other polarizable media), Green's function may depend not only on the difference $\mathbf{r} - \mathbf{r}'$, but on each of these two arguments in a specific way.

So far, this is just re-naming our old results. The really non-trivial result of the Green's function formalism in electrostatics is that, somewhat counter-intuitively, the knowledge of this function for a system with *grounded* conductors (Fig. 30a) enables the calculation of the field created by *voltage-biased* conductors (Fig. 30b), with the same geometry. To show this, let us use the so-called *Green's theorem* of the vector calculus.[70] The theorem states that for any two scalar, differentiable functions $f(\mathbf{r})$ and $g(\mathbf{r})$, and any volume $V$,

$$\int_V \left(f\nabla^2 g - g\nabla^2 f\right)d^3r = \oint_S \left(f\nabla g - g\nabla f\right)_n d^2r, \quad (2.207)$$

where $S$ is the surface limiting the volume. Applying the theorem to the electrostatic potential $\phi(\mathbf{r})$ and the Green's function $G$ (also considered as a function of $\mathbf{r}$), let us use the Poisson equation (1.41) to replace $\nabla^2\phi$ with $(-\rho/\varepsilon_0)$, and notice that $G$, considered as a function of $\mathbf{r}$, obeys the Poisson equation with the $\delta$-functional source:

$$\nabla^2 G(\mathbf{r},\mathbf{r}') = -4\pi\delta(\mathbf{r}-\mathbf{r}'). \quad (2.208)$$

[68] See, e.g., CM Sec. 5.1, QM Secs. 2.2 and 7.4, and SM Sec. 5.5. Note that the numerical coefficient in Eq. (202) (and hence all resulting formulas) is the matter of convention; this choice does not affect the final results.

[69] $G$ so defined is sometimes called the *Dirichlet function*.

[70] See, e.g., MA Eq. (12.3). Actually, this theorem is a ready corollary of the better-known divergence ("Gauss") theorem, MA Eq. (12.2).

(Indeed, according to its definition (202), this function may be formally considered as the field of a point charge $q = 4\pi\varepsilon_0$.) Now swapping the notation of the radius-vectors, $\mathbf{r} \leftrightarrow \mathbf{r}'$, and using the Green's function symmetry, $G(\mathbf{r}, \mathbf{r}') = G(\mathbf{r}', \mathbf{r})$,[71] we get

$$-4\pi\phi(\mathbf{r}) - \int_V \left(-\frac{\rho(\mathbf{r}')}{\varepsilon_0}\right) G(\mathbf{r},\mathbf{r}')d^3r' = \oint_S \left[\phi(\mathbf{r}')\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial n'} - G(\mathbf{r},\mathbf{r}')\frac{\partial \phi(\mathbf{r}')}{\partial n'}\right] d^2r' . \qquad (2.209)$$

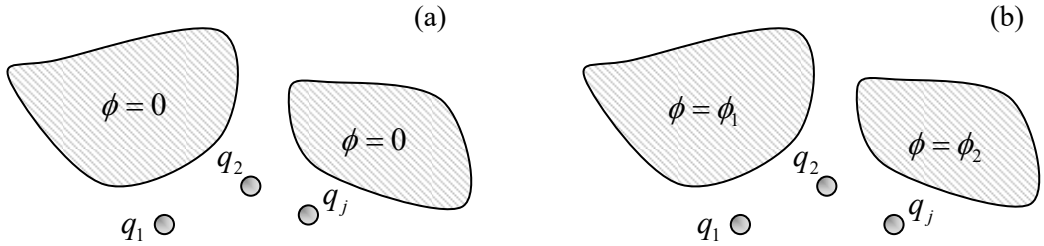



Fig. 2.30. Green's function method allows the solution of a simpler boundary problem (a) to be used for the solution of a more complex problem (b), for the same conductor geometry.

Let us apply this relation to the volume $V$ of *free space* between the conductors, and the boundary $S$ drawn immediately outside of their surfaces. In this case, by its definition, Green's function $G(\mathbf{r}, \mathbf{r}')$ vanishes at the conductor surface, i.e. at $\mathbf{r} \in S$ – see Eq. (205). Now changing the sign of $\partial n'$ (so it would be the outer normal for *conductors*, rather than free space volume $V$), dividing all terms by $4\pi$, and partitioning the total surface $S$ into the parts (numbered by index $j$) corresponding to different conductors (possibly, kept at different potentials $\phi_k$), we finally arrive at the famous result:[72]

Potential via Green's function

$$\phi(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0}\int_V \rho(\mathbf{r}')G(\mathbf{r},\mathbf{r}')d^3r' + \frac{1}{4\pi}\sum_k \phi_k \oint_{S_k} \frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial n'} d^2r' . \qquad (2.210)$$

While the first term on the right-hand side of this relation is a direct and evident expression of the superposition principle, given by Eq. (203), the second term is highly non-trivial: it describes the effect of conductors with *arbitrary* potentials $\phi_k$ (Fig. 30b), using the Green's function calculated for the similar system with *grounded* conductors, i.e. with all $\phi_k = 0$ (Fig. 30a). Let me emphasize that since our volume $V$ excludes conductors, the first term on the right-hand side of Eq. (210) includes only the stand-alone charges in the system (in Fig. 30, marked $q_1, q_2$, etc.), but not the surface charges of the conductors – which are taken into account, indirectly, by the second term.

In order to illustrate what a powerful tool Eq. (210) is, let us use to calculate the electrostatic field in two systems. In the first of them, a plane, circular, conducting disk of radius $R$, separated with a very thin cut from the remaining conducting plane, is biased with potential $\phi = V$, while the rest of the plane is grounded – see Fig. 31.

[71] This symmetry, evident for the particular cases (204) and (206), may be readily proved for the general case by applying Eq. (207) to the functions $f(\mathbf{r}) \equiv G(\mathbf{r}, \mathbf{r}')$ and $g(\mathbf{r}) \equiv G(\mathbf{r}, \mathbf{r}'')$. With this substitution, the left-hand side of that equality becomes equal to $-4\pi[G(\mathbf{r}'', \mathbf{r}') - G(\mathbf{r}', \mathbf{r}'')]$, while the right-hand side is zero, due to Eq. (205).

[72] In some textbooks, the sign before the surface integral is negative, because their authors use the outer normal to the *free-space* region $V$ rather than those *occupied by conductors* – as I do.

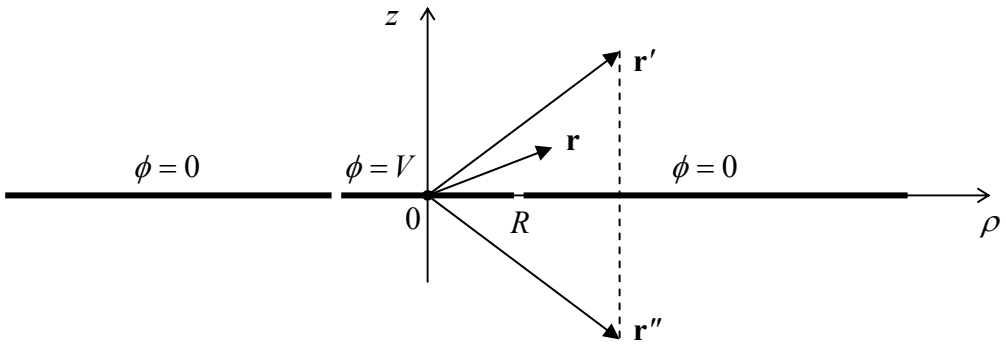



Fig. 2.31. A voltage-biased conducting disk separated from the rest of a conducting plane.

If the width of the gap between the disk and the rest of the plane is negligible, we may apply Eq. (210) without stand-alone charges, $\rho(\mathbf{r}') = 0$, and the Green's function for the uncut plane – see Eq. (206).[73] In the cylindrical coordinates, with the origin at the disk's center (Fig. 31), the function is

$$G(\mathbf{r},\mathbf{r}') = \frac{1}{\left[\rho^2 + \rho'^2 - 2\rho\rho'\cos(\varphi-\varphi') + (z-z')^2\right]^{1/2}} - \frac{1}{\left[\rho^2 + \rho'^2 - 2\rho\rho'\cos(\varphi-\varphi') + (z+z')^2\right]^{1/2}}. \tag{2.211}$$

(The sum of the first three terms under each square root in Eq. (211) is just the squared distance between the horizontal projections $\boldsymbol{\rho}$ and $\boldsymbol{\rho}'$ of the vectors $\mathbf{r}$ and $\mathbf{r}'$ (or $\mathbf{r}''$) correspondingly, while the last terms are the squares of their vertical displacements.)

Now we can readily calculate the derivative participating in Eq. (210), for $z \geq 0$:

$$\frac{\partial G}{\partial n'}\Big|_S = \frac{\partial G}{\partial z'}\Big|_{z'=+0} = \frac{2z}{\left(\rho^2 + \rho'^2 - 2\rho\rho'\cos(\varphi-\varphi') + z^2\right)^{3/2}}. \tag{2.212}$$

Due to the axial symmetry of the system, we may take $\varphi$ for zero. With this, Eqs. (210) and (212) yield

$$\phi = \frac{V}{4\pi}\oint_S \frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial n'} d^2r' = \frac{Vz}{2\pi}\int_0^{2\pi} d\varphi' \int_0^R \frac{\rho' d\rho'}{\left(\rho^2 + \rho'^2 - 2\rho\rho'\cos\varphi' + z^2\right)^{3/2}}. \tag{2.213}$$

This integral is not overly pleasing, but may be readily worked out at least for points on the symmetry axis ($\rho = 0$, $z \geq 0$):[74]

$$\phi = Vz\int_0^R \frac{\rho' d\rho'}{\left(\rho'^2 + z^2\right)^{3/2}} = \frac{V}{2}\int_0^{R^2/z^2} \frac{d\xi}{(\xi+1)^{3/2}} = V\left[1 - \frac{z}{\left(R^2 + z^2\right)^{1/2}}\right], \tag{2.214}$$

This result shows that if $z \to 0$, the potential tends to $V$ (as it should), while at $z >> R$,

$$\phi \to V\frac{R^2}{2z^2}. \tag{2.215}$$

[73] Indeed, if all parts of the cut plane are grounded, a narrow cut does not change the field distribution, and hence the Green's function, significantly.

[74] There is no need to repeat the calculation for $z \leq 0$: from the symmetry of the problem, $\phi(-z) = \phi(z)$.

Now let us use the same Eq. (210) to solve the (in :-)famous problem of the cut sphere (Fig. 32). Again, if the gap between the two conducting semi-spheres is very thin ($t << R$), we may use Green's function for the grounded (and uncut) sphere. For a particular case $\mathbf{r}' = d\mathbf{n}_z$, this function follows from Eqs. (197)-(198); generalizing the former relation for an arbitrary direction of vector $\mathbf{r}'$, we get

$$G = \frac{1}{\left[r^2 + r'^2 - 2rr'\cos\gamma\right]^{1/2}} - \frac{R/r'}{\left[r^2 + (R^2/r')^2 - 2r(R^2/r')\cos\gamma\right]^{1/2}}, \qquad \text{for } r, r' \geq R, \quad (2.216)$$

where $\gamma$ is the angle between the vectors $\mathbf{r}$ and $\mathbf{r}'$, and hence $\mathbf{r}''$ – see Fig. 32.

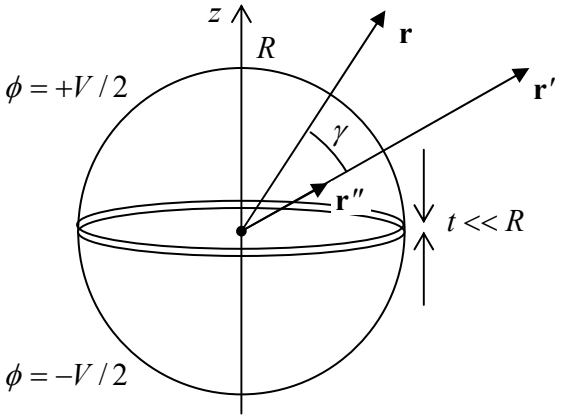



Fig. 2.32. A system of two separated, oppositely biased semi-spheres.

Now, calculating the Green's function's derivative,

$$\frac{\partial G}{\partial r'}\Big|_{r'=R+0} = -\frac{(r^2 - R^2)}{R\left[r^2 + R^2 - 2Rr\cos\gamma\right]^{3/2}}, \quad (2.217)$$

and plugging it into Eq. (210), we see that the integration is again easy only for the field on the symmetry axis (where $\mathbf{r} = z\mathbf{n}_z$, and $\gamma = \theta'$), giving:

$$\phi = \frac{V}{2}\left[1 - \frac{z^2 - R^2}{z\left(z^2 + R^2\right)^{1/2}}\right], \quad \text{for } \theta = 0. \quad (2.218)$$

For $z \to R$, this relation yields $\phi \to V/2$ (as it should), while for $z/R \to \infty$,

$$\phi \to \frac{3R^2}{4z^2}V. \quad (2.219)$$

As will be discussed in the next chapter, such a field is typical for an electric dipole.

## 2.11. Numerical methods

Despite the richness of analytical methods, for many boundary problems (especially in geometries without a high degree of symmetry), the numerical approach is the only way to the solution.

Though software packages offering their automatic numerical solution are abundant nowadays,[75] it is important for every educated physicist to understand "what is under the hood", at least because most universal programs exhibit mediocre performance in comparison with custom codes written for particular problems, and sometimes do not converge at all, especially for fast-changing (say, exponential) functions. The very brief discussion presented here[76] is a (hopefully, useful) fast glance under the hood, though it is certainly insufficient for professional numerical research work.

The simplest of the numerical approaches to the solution of partial differential equations, such as the Poisson or the Laplace equations (1.41)-(1.42), is the *finite-difference* method,[77] in which the sought continuous scalar function $f(\mathbf{r})$, such as the potential $\phi(\mathbf{r})$, is represented by its values in discrete points of a rectangular grid (frequently called *mesh*) of the corresponding dimensionality – see Fig. 33.

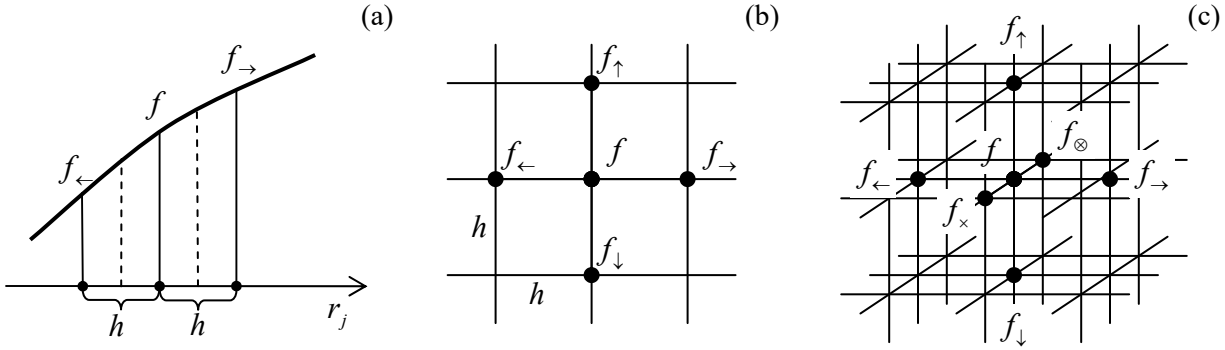



Fig. 2.33. The general idea of the finite-difference method in (a) one, (b) two, and (c) three dimensions.

Each partial second derivative of the function is approximated by the formula that readily follows from linear approximations of the function $f$ and then its partial derivatives – see Fig. 33a:

$$\frac{\partial^2 f}{\partial r_j^2} \equiv \frac{\partial}{\partial r_j}\left(\frac{\partial f}{\partial r_j}\right) \approx \frac{1}{h}\left(\frac{\partial f}{\partial r_j}\bigg|_{r_j+h/2} - \frac{\partial f}{\partial r_j}\bigg|_{r_j-h/2}\right) \approx \frac{1}{h}\left[\frac{f_\rightarrow - f}{h} - \frac{f - f_\leftarrow}{h}\right] = \frac{f_\rightarrow + f_\leftarrow - 2f}{h^2}, \quad (2.220)$$

where $f_\rightarrow \equiv f(r_j + h)$ and $f_\leftarrow \equiv f(r_j - h)$. (The relative error of this approximation is of the order of $h^4\partial^4 f/\partial r_j^4$.) As a result, the action of a 2D Laplace operator on the function $f$ may be approximated as

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \approx \frac{f_\rightarrow + f_\leftarrow - 2f}{h^2} + \frac{f_\uparrow + f_\downarrow - 2f}{h^2} = \frac{f_\rightarrow + f_\leftarrow + f_\uparrow + f_\downarrow - 4f}{h^2}, \quad (2.221)$$

and of the 3D operator, as

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \approx \frac{f_\rightarrow + f_\leftarrow + f_\uparrow + f_\downarrow + f_\otimes + f_\times - 6f}{h^2}. \quad (2.222)$$

(The notation used in Eqs. (221)-(222) should be clear from Figs. 33b and 33c, respectively.)

[75] See, for example, MA Secs. 16 (iii) and (iv).

[76] It is almost similar to that given in CM Sec. 8.5 and is reproduced here for the reader's convenience, illustrated with examples from this (EM) course.

[77] For more details see, e.g., R. Leveque, *Finite Difference Methods for Ordinary and Partial Differential Equations*, SIAM, 2007.

As a simple example, let us use this scheme to find the electrostatic potential distribution inside a cylindrical box with conducting walls and square cross-section $a \times a$, using an extremely coarse mesh with step $h = a/2$ (Fig. 34). In this case, our function, the electrostatic potential $\phi(x, y)$, equals zero at the side and bottom walls, and $V_0$ at the top lid, so, according to Eq. (221), the 2D Laplace equation may be approximated as

$$\frac{0 + 0 + V_0 + 0 - 4\phi}{(a/2)^2} = 0. \tag{2.223}$$

The resulting value for the potential in the center of the box is $\phi = V_0/4$.

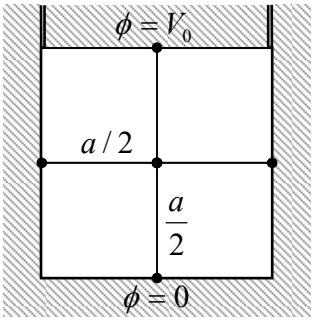



Fig. 2.34. Numerically solving an internal 2D boundary problem for a conducting, cylindrical box with a square cross-section, using a very coarse mesh (with $h = a/2$).

Surprisingly, this is the *exact* value! This may be proved either by solving this problem by the variable separation method, just as this has been done for a similar 3D problem in Sec. 5, or just from the following Green's-function argument. If all four walls of our 2D volume were biased to the voltage $V_0$, there would be no electric field in it at all, so the middle-point potential would be equal to $V_0$ as well. However, from the point of view of Eq. (210) with no bulk charge, $\rho(\mathbf{r}) = 0$, this result may be legitimately viewed as the linear superposition of the four contributions of the potentials $\phi_k = V_0$ of each wall. Since for this symmetric geometry, the corresponding geometrical factors are equal, the contribution of one wall, with $\phi_k = 0$ on all other walls (as in our current problem), has to equal $V_0/4$.

For a similar 3D problem (a cubic box), with a similar 3D mesh, Eq. (222) yields

$$\frac{0 + 0 + V_0 + 0 + 0 + 0 - 6\phi}{(a/2)^2} = 0, \tag{2.226}$$

so $\phi = V_0/6$. Using the same Green's-function argument, now for six walls of the cube, we see that this result is also exact! (This fact also follows from our variable-separation result expressed by Eqs. (95) and (99) with $a = b = c$.)

Though such exact results should be considered as a happy coincidence rather than the general law, they still show that numerical methods, even with relatively crude meshes, may be more computationally efficient than some "analytical" approaches, like the variable separation method with its infinite-sum results that, in most cases, require computers anyway – at least for the result's comprehension and analysis.

A more powerful (but also much more complex) approach is the *finite-element* method in which the discrete point mesh, typically with triangular cells, is (automatically) generated in accordance with the system geometry.[78] Such mesh generators provide higher point concentration near sharp convex

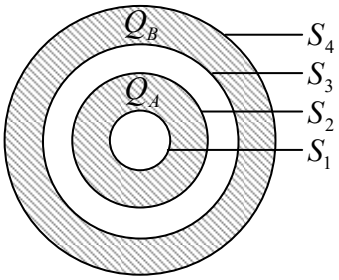
[78] See, e.g., CM Fig. 8.14.

parts of conductor surfaces, where the field concentrates and hence the potential changes faster, thus ensuring a better accuracy-to-speed tradeoff than the finite-difference methods on a uniform grid. The price to pay for this improvement is the algorithm's complexity which makes its adjustments much harder. Unfortunately, in this series, I do not have time for going into the details of that method and have to refer the reader to the special literature on this subject.[79]

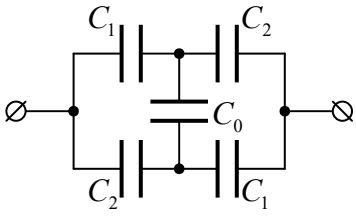
## 2.12. Exercise problems

2.1. Calculate the force (per unit area) exerted on a conducting surface by an external electric field normal to it. Compare the result with the electric field's definition given by Eq. (1.6), and comment.

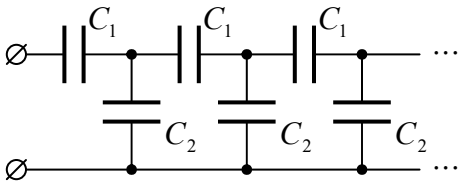
2.2. Electric charges $Q_A$ and $Q_B$ have been placed on two conducting concentric spherical shells – see the figure on the right. What is the full charge of each of the surfaces $S_1$-$S_4$?




2.3. Calculate the mutual capacitance between the terminals of the lumped-capacitor circuit shown in the figure on the right. Analyze and interpret the result.

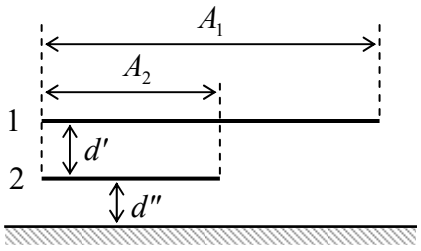



2.4. Calculate the mutual capacitance between the terminals of the semi-infinite lumped-capacitor circuit shown in the figure on the right, and find the law of the applied voltage's decay along the system. Analyze and interpret the result.



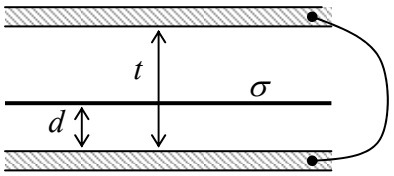

2.5. A system of two thin conducting plates is located over a ground plane as shown in the figure on the right, where $A_1$ and $A_2$ are the areas of the indicated plate parts, while *d'* and *d"* are the distances between them. Neglecting the fringe effects, calculate:

(i) the effective capacitance of each plate, and
(ii) their mutual capacitance.

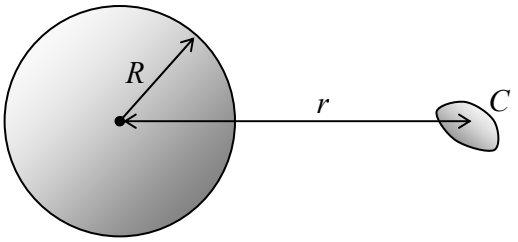



[79] See, e.g., either C. Johnson, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Dover, 2009, or T. Hughes, *The Finite Element Method*, Dover, 2000.

2.6. A wide and thin film carrying a uniformly distributed electric charge of areal density $\sigma$ is placed inside a plane capacitor whose plates are connected with a wire – see the figure on the right. Neglecting the fringe effects, calculate the surface charges of the plates and the net force exerted on the film (per unit area).



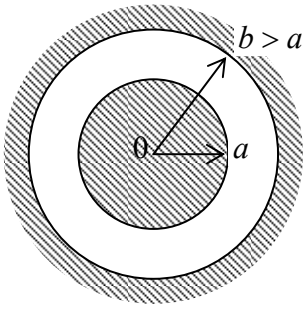

2.7. A relatively small conductor (possibly, of an irregular shape) with self-capacitance $C$ is located at distance $r$ from the center of a conducting sphere of radius $R$ – see the figure on the right. In the first approximation in $C$, find the reciprocal capacitance matrix of the system. Use the matrix to calculate its potential energy and the force of the conductor interaction for two cases:




(i) the conductor *charges* $Q$ are equal, and

(ii) the conductor *potentials* $\phi$ are kept equal.

2.8. Use the Gauss law to calculate the mutual capacitance of the following two-electrode systems, both with the cross-section shown in Fig. 7 (reproduced on the right):

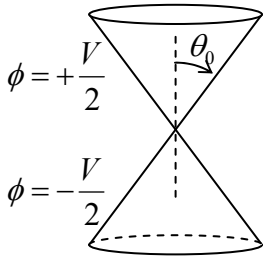



(i) a conducting sphere in the center of a spherical cavity inside another conductor, and

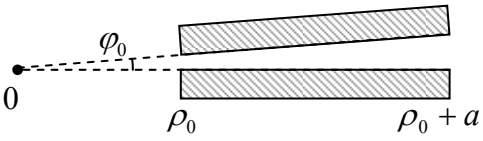
(ii) a long conducting round cylinder on the axis of a cylindrical cavity inside another conductor, i.e. a coaxial cable. (In this case, we speak about the capacitance per unit length).

Compare the results with those obtained in Sec. 2.2 using the Laplace equation.

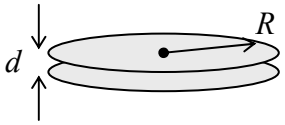
2.9. Calculate the electrostatic potential distribution around two barely separated conductors in the form of coaxial round cones (see the figure on the right), with voltage $V$ between them. Compare the result with that of a similar 2D problem, with the cones replaced with plane-face wedges. Can you calculate the mutual capacitances between the conductors in these systems? If not, can you estimate them?




2.10. Calculate the mutual capacitance between two rectangular planar electrodes of area $A = a \times l$, with a very small angle $\varphi_0$ between them – see the figure on the right.




2.11. Using the results for a single thin round disk, obtained in Sec. 4, consider a system of two such disks at a small distance $d << R$ from each other – see the figure on the right. In particular, calculate:

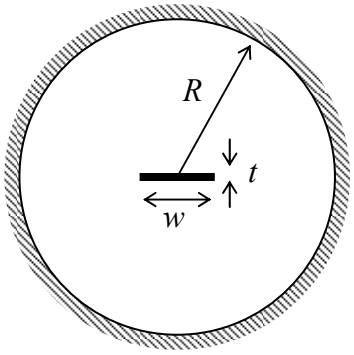



(i) the reciprocal capacitance matrix of the system,

(ii) the mutual capacitance between the disks,
(iii) the partial capacitance of one disk, and
(iv) the effective capacitance of one disk,

– all in the first nonvanishing approximation in $d/R << 1$. Compare the results (ii)-(iv) and interpret their similarities and differences.

2.12.* Calculate the mutual capacitance (per unit length) between two cylindrical conductors forming a system with the cross-section shown in the figure on the right, in the limit $t << w << R$.




*Hint*: You may like to use the elliptic coordinates mentioned in Sec. 4. They are defined by the following equality:

$$x + iy = c\cosh(\mu + i\nu), \qquad (*)$$

where $c$ is a constant.

2.13. Calculate the mutual capacitance (per unit length) between two similar, long, parallel wires, each with a round cross-section of radius $R$, whose axes are separated by distance $d > 2R$. Explore and interpret the result in the limits $R \to 0$ and $R \to 2d$.

*Hint*: You may like to use the 2D orthogonal *bipolar coordinates* $\{\tau, \sigma\}$ defined by the following relations with the Cartesian coordinates $\{x, y\}$:

$$x = a\frac{\sinh\tau}{\cosh\tau - \cos\sigma}, \qquad y = a\frac{\sin\sigma}{\cosh\tau - \cos\sigma}, \qquad \text{with } -\infty < \tau < +\infty, \quad -\pi \le \sigma \le +\pi.$$

In these coordinates, the Laplace operator is

$$\nabla^2 = \frac{1}{a^2}(\cosh\tau - \cos\sigma)\left(\frac{\partial^2}{\partial\tau^2} + \frac{\partial^2}{\partial\sigma^2}\right).$$

2.14. Formulate 2D electrostatic problems that may be simply solved using each of the following analytic functions of the complex variable $\boldsymbol{z} \equiv x + iy$:

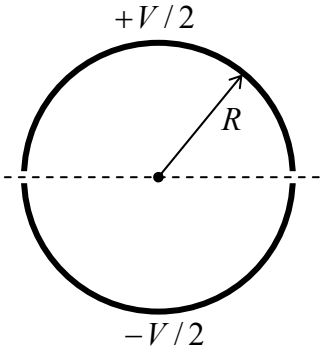
(i) $\boldsymbol{w} = \ln \boldsymbol{z}$,
(ii) $\boldsymbol{w} = \boldsymbol{z}^{1/2}$,
(iii) $\boldsymbol{w} = \boldsymbol{z} + 1/\boldsymbol{z}$,

and solve these problems.

2.15. On each side of a cylindrical volume with a rectangular cross-section $a \times b$, with no electric charges inside it, the electric field's component normal to the side's plane is constant, and also equal and opposite to that on the opposite side. Calculate the distribution of the electric potential inside the volume, provided that the magnitude of the normal components on the sides of length $b$ equals $E$. Suggest a practicable method to implement such potential distribution.

2.16. Complete the solution of the problem shown in Fig. 12, by calculating the distribution of the surface charge on the semi-planes. Can you calculate the mutual capacitance between the semi-planes (per unit length of the system)? If not, can you estimate it?

2.17. A straight, long, thin, round-cylindrical conducting pipe has been cut, along its axis, into two equal parts – see the figure on the right.




(i) Use the conformal mapping method to calculate the distributions of the electrostatic potential created by voltage $V$ applied between the two parts, both outside and inside the pipe, and of the surface charge.

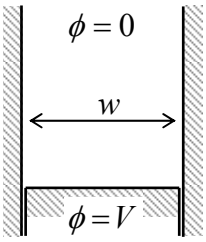
(ii)* Calculate the mutual capacitance between the pipe's halves (per unit length), taking into account a small width $2t << R$ of the cut.

*Hints*: In Task (i), you may like to use the complex function

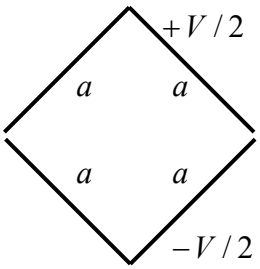
$$w = \ln \frac{R+z}{R-z},$$

while in Task (ii), you may use the solution of the previous problem.

2.18. A gap of constant width $w$ between two grounded conducting semi-spaces is closed, from one side, with a conducting plunger biased with voltage $V$, so that the cross-section of the system looks like the figure on the right shows. Use the variable separation method to calculate the distribution of the electrostatic potential within the gap.

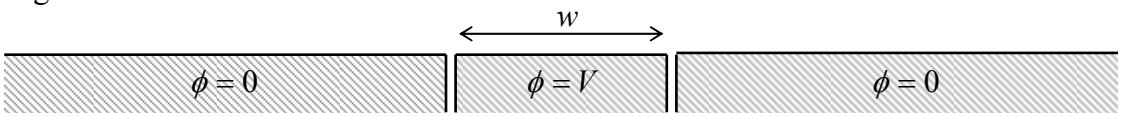



2.19. Use the variable separation method to calculate the electrostatic potential's distribution inside a very long thin-wall metallic box with a quadratic cross-section, cut and voltage-biased as shown in the figure on the right. (Assume that the cut's width is negligibly small.)



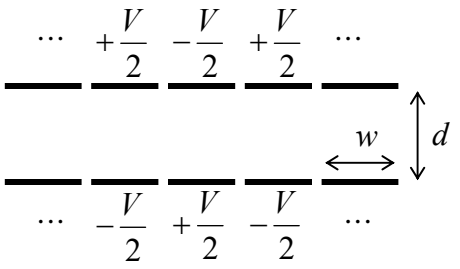

2.20. Solve Problem 17(i) by using the variable separation method, and compare the results.

2.21. Use the variable separation method to calculate the potential distribution above the plane surface of a conductor, with a strip of width $w$ separated by very thin cuts, and biased with voltage $V$ – see the figure below.




2.22. The previous problem is now modified: the cut-out and voltage-biased part of the conducting plane is now not a strip, but a square with side $w$. Calculate the potential distribution above the conductor's surface.

2.23. Each electrode of a large plane capacitor is cut into parallel long strips of equal width $w$, with very narrow gaps between them. These strips are kept at alternating potentials, as shown in the figure on the right. Use the variable separation method to calculate the electrostatic potential distribution in space, and explore the limit $w << d$.




2.24. Complete the cylinder problem started in Sec. 7 (see Fig. 17), for the cases when the top lid's voltage is fixed as follows:

(i) $V = V_0 J_1(\xi_{11}\rho/R)\sin\varphi$, where $\xi_{11} \approx 3.832$ is the first root of the Bessel function $J_1(\xi)$;

(ii) $V = V_0 = \text{const}$.

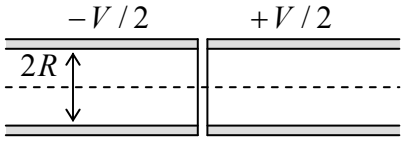
For both cases, calculate the electric field at the centers of the lower and upper lids. (For Task (ii), an answer including series and/or integrals is acceptable.)

2.25. For the infinitely long periodic system sketched in Fig. 21, assuming that $t << h, R$:

(i) calculate and sketch the electrostatic potential's distribution inside the system for various values of the ratio $R/h$, and

(ii) simplify the results for the limit $R/h \to 0$.

2.26. A long round cylindrical conducting pipe is split, with a very narrow cut normal to its axis, into two parts that are voltage-biased as the figure on the right shows. Use two different approaches to calculate the force exerted by the resulting field upon a charged particle flying along the pipe close to its axis. Can the system work as an electrostatic lens?




2.27. Use the variable separation method to find the potential distribution inside and outside of a thin spherical shell of radius $R$, with a fixed potential distribution on it: $\phi(R,\theta,\varphi) = V_0 \sin\theta\cos\varphi$.

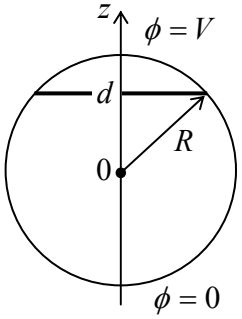
2.28. A thin spherical shell carries an electric charge with areal density $\sigma = \sigma_0\cos\theta$. Calculate the spatial distribution of the electrostatic potential and the electric field, both inside and outside the shell.

2.29. Use the variable separation method to solve the problem already considered in Sec. 10: calculate the potential distribution both inside and outside of a thin spherical shell of radius $R$, separated with a very thin cut along the central plane $z = 0$ into two halves, with voltage $V$ applied between them – see Fig. 32. Analyze the solution; in particular, compare the field at the $z$-axis, for $z > R$, with Eq. (218).

*Hint*: You may like to use the following integral of a Legendre polynomial with an odd index $l = 1, 3, 5\ldots = 2n - 1$:[80]

[80] As a reminder, the *double factorial* (also called "semifactorial") operator (!!) is similar to the usual factorial operator (!), but with the product limited to numbers of the same parity as its argument – in our particular case, of the odd numbers in the numerator and even numbers in the denominator.

$$I_n \equiv \int_0^1 P_{2n-1}(\xi)\,d\xi = \frac{1}{n!}\cdot\left(\frac{1}{2}\right)\cdot\left(-\frac{3}{2}\right)\cdot\left(-\frac{5}{2}\right)\cdot\ldots\cdot\left(\frac{3}{2}-n\right) \equiv (-1)^{n-1}\,\frac{(2n-3)!!}{2n(2n-2)!!}.$$




2.30. Calculate, up to terms $O(1/r^2)$, the long-range electric field induced by a split and voltage-biased conducting sphere – similar to that discussed in Sec. 10 (see Fig. 32) and in the previous problem, but with the cut's plane at an arbitrary distance $d < R$ from the center – see the figure on the right.

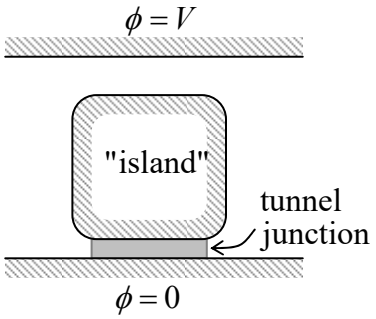
2.31. Calculate the field distribution in the simple electrostatic lens that was the subject of Problem 1.9, provided that the separation of the two field regions is provided by a thin conducting membrane, with a round hole of radius $R$.

*Hint*: You may like to use the fact that the general axially symmetric solution of the Laplace equation in the oblate ellipsoidal coordinates (59) may be represented in the following variable-separation form:

$$\phi = \sum_{n=0}^{\infty}\left[p_n \mathcal{P}_n(i\sinh\alpha) + q_n \mathcal{Q}_n(i\sinh\alpha)\right]\mathcal{P}_n(\cos\beta),$$

where $p_n$ and $q_n$ are constants, $\mathcal{P}_n$ are the Legendre polynomials (2.169), which are sometimes called the *Legendre functions of the first kind*, while $\mathcal{Q}_n$ are the *Legendre functions of the second kind* (briefly mentioned, in a different context, in Sec. 2.8) that may be defined by the following recurrence relations:

$$\mathcal{Q}_0(\xi) = \frac{1}{2}\ln\frac{1+\xi}{1-\xi}, \qquad \mathcal{Q}_1(\xi) = \mathcal{P}_1(\xi)\mathcal{Q}_0(\xi) - 1, \qquad \mathcal{Q}_{n>2}(\xi) = \frac{2n-1}{n}\xi\mathcal{Q}_{n-1}(\xi) - \frac{n-1}{n}\mathcal{Q}_{n-2}(\xi).$$

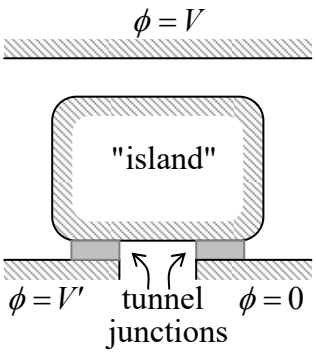



2.32. A small conductor (in this context, usually called a *single-electron island*) is placed between two conducting electrodes, with voltage $V$ applied between them. The gap between the island and one of the electrodes is so narrow that electrons may tunnel quantum-mechanically through this "junction" – see the figure on the right. Neglecting thermal excitation effects, calculate the equilibrium charge of the island as a function of $V$.

*Hint*: To solve this problem, you do not need to know much about the quantum-mechanical tunneling between conductors, besides that such tunneling of an electron, together with energy relaxation of the resulting excitations, may be considered a single inelastic (energy-dissipating) event.[81] At negligible thermal excitations, such an event takes place only if it decreases the total potential energy of the system.

[81] Strictly speaking, this statement, implying negligible quantum-mechanical coherence of the tunneling events, is correct only if the junction transparency is so low that its effective electric resistance is much higher than the fundamental quantum unit of resistance, $R_Q \equiv \pi\hbar/2e^2 \approx 6.5$ kΩ (see, e.g., QM Sec. 3.2). However, this condition is satisfied in most experimental tunnel junctions.

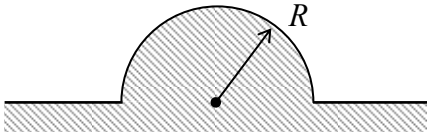
2.33. The system discussed in the previous problem is now generalized as the figure on the right shows. If the voltage $V'$ applied between the two bottom electrodes is sufficiently large, electrons can successively tunnel through two junctions of this system (called the *single-electron transistor*), carrying dc current between these electrodes. Neglecting thermal excitations, calculate the region of voltages $V$ and $V'$ where such a current is fully suppressed (*Coulomb-blocked*).




2.34. Use the charge image method to calculate the full surface charges induced in the plates of a very broad, voltage-unbiased plane capacitor of thickness $D$ by a point charge $q$ separated from one of the electrodes by distance $d$. Suggest at least one alternative method to obtain the same result.

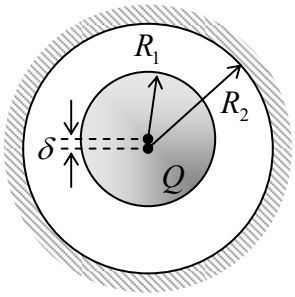
2.35. Use the charge image method to calculate the potential energy of the electrostatic interaction between a point charge placed in the center of a spherical cavity that had been carved inside a grounded conductor, and the cavity's walls. Looking at the result, could it be obtained in a simpler way (or ways)?

2.36. Use the method of charge images to find the Green's function of the system shown in the figure on the right, where the bulge on the (otherwise, plane) surface of a conductor has the shape of a semi-sphere of radius $R$.




2.37.* Use the spherical inversion expressed by Eq. (198), to develop an iterative method for a more and more precise calculation of the mutual capacitance between two similar conducting spheres of radius $R$, with their centers separated by distance $d > 2R$.

2.38.* A conducting sphere of radius $R_1$, carrying an electric charge $Q$, is placed inside a spherical cavity of radius $R_2 > R_1$, carved inside another bulk conductor. Calculate the electric force exerted on the sphere if its center is displaced by a small distance $\delta << R_1, R_2 - R_1$ from that of the cavity – see the figure on the right.




2.39. Within the simple models of the electric field screening in conductors, discussed in Sec. 2.1, analyze the partial screening of the electric field of a point charge $q$ by a planar conducting film of constant thickness $t << \lambda$, where $\lambda$ is (depending on charge carrier statistics) either the Debye or the Thomas-Fermi screening length – see, respectively, Eqs. (8) or (10). Assume that the distance $d$ between the charge and the film is much larger than $t$.

2.40. Prove the following expansion of the simplest Green's function (204) into a series over the Legendre polynomials:

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|}=\frac{1}{r_>}\sum_{l=0}^{\infty}\left(\frac{r_<}{r_>}\right)^l \mathcal{P}_l(\cos\theta),$$

where $r_>$ is the largest of the two scalars $r \equiv |\mathbf{r}| \geq 0$ and $r' \equiv |\mathbf{r}'| \geq 0$, while $r_<$ is the smallest of them.

2.41. Use the expansion that was the subject of the previous problem to confirm the analysis, in Sec. 2.9 of the lecture notes, of the system shown in Fig. 29: a grounded conducting sphere of radius $R$, and a point charge $q$ located at distance $d > R$ from its center.

2.42. Suggest a convenient definition of the Green's function for 2D electrostatic problems, and calculate it for:

(i) the unlimited free space, and
(ii) the free space above a conducting plane.

Use the latter result to re-solve Problem 21.

2.43. A conducting plane is separated into two parts with a very narrow straight cut, and voltage $V$ is applied between the resulting half-planes – see the figure below. Use the Green's function method to find the distribution of the electrostatic potential and the electric field everywhere in the space. Compare the result with Eq. (83). In hindsight, could the problem be solved in an even simpler way (or ways)?


2.44. Use the last result of Problem 42 and one of the conformal mappings discussed in Sec. 4 to find one more solution of Problem 18.

2.45. Calculate the 2D Green's functions for the free spaces:

(i) outside a round conducting cylinder, and
(ii) inside a round cylindrical hole in a conductor.

2.46. Solve Problem 17(i) using the Green's function method.

2.47. Solve the 2D boundary problem that was discussed in Sec. 11 (Fig. 34) by using:

(i) the finite difference method with the finer square mesh $h = a/3$, and
(ii) the variable separation method.

Compare the results at the mesh points, and comment.

This page is intentionally left blank