Fields and waves in nature and engineering — the big picture:

Fundamental building blocks of matter — electrons and protons at atomic scales — interact with one another gravitationally and via "electromagnetic" forces. These interactions are most conveniently described in terms of suitably defined "vector fields" that permeate space and time, or simply the *space-time* $(x, y, z, t) \equiv (\mathbf{r}, t)$. Interactions attributed to particle *masses* can be formulated by *gravitational fields* $\mathbf{g}(\mathbf{r}, t)$ specified in reference frames where spatial coordinates $\mathbf{r} = (x, y, z)$ are defined. Far stronger interactions attributed to particle *charges*, on the other hand, are formulated in terms of a *pair* of vector fields, $\mathbf{E}(\mathbf{r}, t)$ and $\mathbf{B}(\mathbf{r}, t)$, known as *electric and magnetic fields*, respectively.

Electric and magnetic fields:

A particle with charge q and mass m as well as position and velocity vectors $\mathbf{r}$ and $\mathbf{v} = \frac{d\mathbf{r}}{dt}$ specified at an instant t *within* a measurement frame (or "lab" frame) will be accelerated in accordance with

$$m\frac{d\mathbf{v}}{dt} = q(\mathbf{E}(\mathbf{r}, t) + \mathbf{v} \times \mathbf{B}(\mathbf{r}, t)), \tag{1}$$

which is *Newton's 2nd law of motion*[1] for a particle under the influence of *Lorentz force*

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \tag{2}$$

In view of (1), the operational definitions of fields $\mathbf{E}(\mathbf{r}, t)$ and $\mathbf{B}(\mathbf{r}, t)$ arise from particle acceleration $\mathbf{a} = \frac{d\mathbf{v}}{dt}$ that can be measured in the lab frame: the electric field $\mathbf{E}$ is evidently force per unit stationary charge (i.e., $\mathbf{v} = 0$) whereas field $\mathbf{B}$ describes an additional force per charge in transport (i.e., $q\mathbf{v}$) that acts in a direction perpendicular to $\mathbf{v}$.

There are important differences between gravitational and electromagnetic interactions: Gravitational interactions are always attractive indicating that particle masses m that generate the gravitational field $\mathbf{g}(\mathbf{r}, t)$ must all have the same algebraic sign (taken to be positive by convention). Electromagnetic interactions, on the other hand, are attractive or repulsive depending on particle charges q which can be positive or negative. *By convention* a positive charge $q = e \approx 1.6 \times 10^{-19}$ C is attributed to the fundamental particle know as *proton*, while, again by convention, $q = -e$ for an *electron*, the sole companion of the proton within a hydrogen atom[2]. Protons and electrons are charged elementary building blocks[3] of all atoms (hydrogen as

[1]Valid so long as $|\mathbf{v}| \ll c$ where c is the speed of light in vacuum.

[2]Hydrogen atom exists as a consequence of mutual attraction between proton and electron counterbalanced by quantum mechanical constraints on allowed energy states — the constraints include the influence of short-lived *virtual* particle/anti-particle pairs interacting with the proton and electron in a sporadic manner.

[3]Atoms can also contain in their nuclei varying numbers of an uncharged particle known as the *neutron* which is responsible for different isotopes of chemical elements (e.g., the hydrogen isotope known as deuterium contains a neutron in addition to a proton and an electron). While neutrons have no net

well as atoms of heavier elements) that constitute the matter around us. In a collection of fundamental particles the total mass is always a monotonically increasing function of the number of particles. However, that is not the case with total charge since individual particle charges can be positive or negative. In fact, the net charge density $\rho(\mathbf{r}, t)$ found in macroscopic amounts of matter is typically close to zero as a result of having nearly equal numbers of protons and electrons in ordinary matter composed of charge-neutral atoms and molecules[4].

charge, they consist of charged sub-nuclear particles known as up ($\frac{2}{3}e$) and down ($-\frac{1}{3}e$) *quarks* whose motions within the neutron establish currents and a magnetic moment.

[4]The reason why intrinsically weaker gravity becomes dominant in the macro world.

Fields are relative:

Physical laws that we use today to describe our surroundings have been developed to have identical forms in all reference frames in uniform motion with respect to one another. For instance, Lorentz force law on a charge q is

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \text{ and } \mathbf{F}' = q(\mathbf{E}' + \mathbf{v}' \times \mathbf{B}') \tag{3}$$

in terms of unprimed and primed variables measured in two reference frames. Moreover, particle charge q and the speed of light c are assigned invariant[5] values in reference frames in relative motion (thus q' and c' are unnecessary to invoke in physical models). The ramifications of these restrictions constituting the **special theory of relativity** (first described by Einstein in 1905 and covered at UIUC in PHYS 325) are in full accord with experimental measurements. They are also well matched by **Newtonian relations** (approximate but more intuitive laws of dynamics covered in PHYS 211) if and when the relative speed of primed and unprimed frames is negligible compared to the speed of light c.

Since in Newtonian descriptions mass m and acceleration $\frac{d\mathbf{v}}{dt}$ have invariant values in all reference frames, it follows that if and when $|\mathbf{v}' - \mathbf{v}| \ll c$, then $\mathbf{F}' = \mathbf{F}$, in which case (3) implies

$$\mathbf{E}' + \mathbf{v}' \times \mathbf{B}' = \mathbf{E} + \mathbf{v} \times \mathbf{B}. \tag{4}$$

[5]Other "relativistic invariants" between different reference frames include particle (rest) masses and the so-called "spacetime interval" $\sqrt{t^2 - L^2/c^2} = \sqrt{t'^2 - L'^2/c^2}$ between two events ocurring at two locations and two times separated by a distance L and time-delay t, respectively. Relativistic invariants are the most prized physical quantitites to focus on in relativistic models (simply because they remain fixed in all reference frames). Note that distances $L \neq L'$ and time-delays $t \neq t'$ are not *relativistic* invariants!

Then, for a stationary charge in the primed frame, we have $\mathbf{v}' = 0$ and

$$\mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B}, \tag{5}$$

which indicates that force per unit stationary charge in the primed frame — i.e., the electric field in the primed frame — is a linear combination of electrical and magnetic forces exerted on the same charge as seen from another reference frame (unprimed) where the charge appears to have a vector velocity $\mathbf{v}$.

Thus, electric and magnetic fields needed in the formulation of charged particle interactions are not unrelated to one another — rather, they intermix in a manner that depends on the reference frame[6] being used for analysis purposes. Note that charges q which are stationary in one reference frame (and therefore carry no electrical current) will appear to be in motion in another frame and thus carry electrical currents I. It must therefore be evident that the equations for $\mathbf{E}$ and $\mathbf{B}$ in any reference frame must be cross-coupled and depend on both charge and current densities that are measured in the same frame.

[6]Given $\mathbf{E}$ and $\mathbf{B}$ measured in the lab, $\mathbf{E}'$ and $\mathbf{B}'$ measured by an observer moving through the lab with a constant velocity $\mathbf{v}$ are well approximated by $\mathbf{E}' \approx \mathbf{E} + \mathbf{v} \times \mathbf{B}$ and $\mathbf{B}' \approx \mathbf{B} - \frac{\mathbf{v}\times\mathbf{E}}{c^2}$ so long as $|\mathbf{v}| \ll c = 3 \times 10^8$ m/s, the *speed of light* in free space (shown by relativistic analysis discussed in PHYS 225 — exact transformation formulae are $\mathbf{E}'_\parallel = \mathbf{E}_\parallel$, $\mathbf{B}'_\parallel = \mathbf{B}_\parallel$, $\mathbf{E}'_\perp = \gamma(\mathbf{E}_\perp + \mathbf{v} \times \mathbf{B}_\perp)$, $\mathbf{B}'_\perp = \gamma(\mathbf{B}_\perp - \frac{\mathbf{v}\times\mathbf{E}_\perp}{c^2})$, where $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$).

Maxwell's field equations:

The required set of coupled equations governing $\mathbf{E}$ and $\mathbf{B}$ was "discovered" in 1864 by James Clerk Maxwell to be (first introduced in PHYS 212 in integral form and discussed throughout this course)

$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_o}$	Divergence eqn's	$\nabla \cdot \mathbf{B} = 0$
$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$	Curl eqn's	$\nabla \times \mathbf{B} = \mu_o \mathbf{J} + \mu_o \epsilon_o \frac{\partial \mathbf{E}}{\partial t}$

where

$$\mu_o \equiv 4\pi \times 10^{-7} \frac{\mathbf{H}}{\mathbf{m}} \ \textbf{and} \ \epsilon_o = \frac{1}{\mu_o c^2} \approx \frac{1}{36\pi \times 10^9} \frac{\mathbf{F}}{\mathbf{m}}$$

in mksA units and

$$c = \frac{1}{\sqrt{\mu_o \epsilon_o}} \approx 3 \times 10^8 \frac{\mathbf{m}}{\mathbf{s}}$$

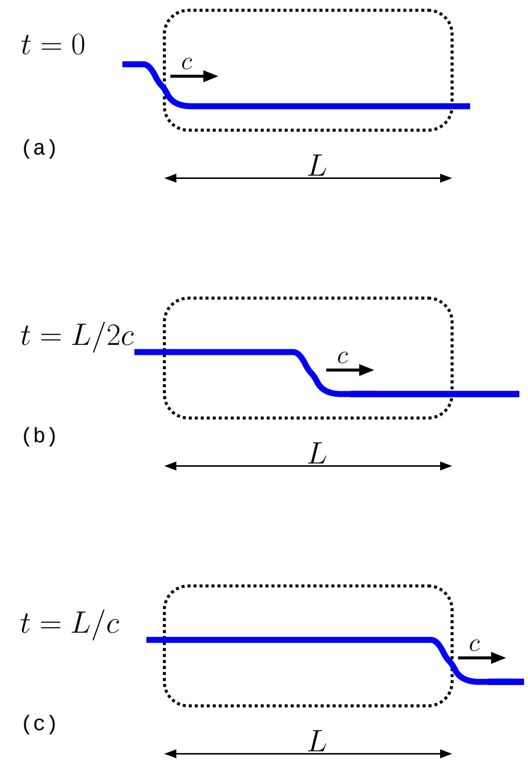
is the speed of light in free space. Furthermore $\rho = \rho(\mathbf{r}, t)$ refers to the net charge density and $\mathbf{J} = \mathbf{J}(\mathbf{r}, t)$ to the current density in the measurement frame, whereas $\nabla \cdot \mathbf{E}$ and $\nabla \times \mathbf{E}$ refer to the **divergence** and **curl** of vector field $\mathbf{E}$ generated by partial differentiation of the orthogonal components of $\mathbf{E}$ (concepts introduced in MATH 241 and reviewed in Lecture 4).

Solutions of Maxwell's equations — waves and static fields (AC/DC):

Maxwell's partial differential equations shown above, describing the coupled dynamics of electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$ in response to space and time varying source fields ρ and $\mathbf{J}$, require an extended study to appreciate their full ramifications and predictions. All predictions of these equations have been experimentally verified and it has been found out that everything that is known and observed about electricity and magnetism can be explained in terms of these equations and their quantized forms.

One of their predictions, derived specifically in Lecture 18, is that they support *traveling wave solutions* of the form

$$\mathbf{E}(\mathbf{r},t) \propto \mathbf{B}(\mathbf{r},t) \propto \cos(2\pi f(t-\frac{z}{c})) \qquad (6)$$



in regions where $\mathbf{J}=\rho=0$. These are co-sinusoidal field perturbations having oscillation *frequencies* f, oscillation *periods* $T=\frac{1}{f}$, *wavelengths* $\lambda=\frac{c}{f}$, and they *travel* in 3D space with the speed of light c in free space. Since Maxwell's equations are linear, superpositions of co-sinusoidal waves with different wavelengths provide additional solutions — these can have arbitrary spatial variations and still travel at a fixed speed c. Any such field perturbation will travel across a region of size L during a time interval L/c as illustrated in the margin.

Another prediction of Maxwell's equations is that fields established by static — i.e., non-time-varying — charge and current densities $\rho=\rho(\mathbf{r})$ and $\mathbf{J}=\mathbf{J}(\mathbf{r})$ satisfy two separate sets of decoupled equations

Electrostatics		**Magnetostatics**
$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_o}$	Divergence eqn's	$\nabla \cdot \mathbf{B} = 0$
$\nabla \times \mathbf{E} = 0$	Curl eqn's	$\nabla \times \mathbf{B} = \mu_o \mathbf{J}$

shown in the left and right columns above — these were obtained by simply setting the terms $\partial\mathbf{E}/\partial t$ and $\partial\mathbf{B}/\partial t$ in the curl equations to zero. Independent "curl-free" static electric fields $\mathbf{E}(\mathbf{r})$ and "divergence-free" static magnetic fields $\mathbf{B}(\mathbf{r})$ satisfying these simplified equations are naturally far easier to determine than the coupled dynamic fields $\mathbf{E}(\mathbf{r}, t)$ and $\mathbf{B}(\mathbf{r}, t)$ to be encountered in response to time-varying sources $\rho(\mathbf{r}, t)$ and $\mathbf{J}(\mathbf{r}, t)$.

Quasi-static fields:

Even though in practical cases of interest (in physics and engineering) time-varying sources are the "rule" and static sources an "exception", learning to solve the simplified set of *electrostatics* and *magnetostatics* equations turns out to be invaluable. The reason is, static solutions often provide accurate approximations — known as *quasi-static* approximation — for time-varying field problems involving slowly-varying sources $\rho(\mathbf{r}, t)$ and $\mathbf{J}(\mathbf{r}, t)$.

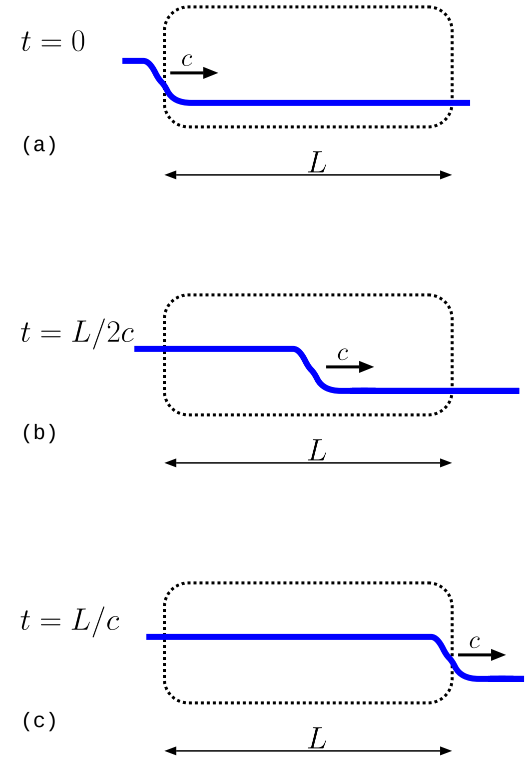
More specifically, if the source variation period T is much longer than the travel time L/c of field perturbations across a region of size L, that is, if

$$T \gg \frac{L}{c}, \tag{7}$$

then field calculations for the entire region can be done statically using the instantaneous (as opposed to retarded or previous) values of field sources ρ and $\mathbf{J}$. This is true because under the given condition source strengths will remain nearly constant over time intervals needed to communicate the new fields to the most distant corners of the region of interest. We can also re-state the same inequality (7) as

$$L \ll cT = \frac{c}{f} = \lambda \tag{8}$$

using the definition of wavelength λ introduced earlier. The indication is then, any system with a physical size L that is small in terms of wavelength λ of the applied field variations can be analyzed quasi-statically by starting from Maxwell's static equations.



Fields and circuits:

Lumped circuit analysis techniques introduced in ECE 110 and 210 constitute practical applications of the quasi-static approach suitable for "electrically small circuits" consisting of capacitors, inductors, and resistors and slowly varying AC sources. By contrast, the analysis of "electrically large circuits" with physical dimensions L approaching or exceeding λ requires taking a proper account of propagation time delays L/c in the system by developing a *distributed circuit* approach based on the full set of Maxwell's equations.

One practical application area where this need is most acute nowadays is in *chip* (integrated circuit) design and packaging suitable for high-speed computing[7]. While the physical dimensions of electronic chips and microcircuits are generally very small, such elements can still be *electrically large* in the sense that $L \sim \lambda$ because of reduced wavelengths λ at high clock speeds $f = 1/T$. Thus, even the computer engineers (CompE's) amongst us need to understand and learn how to mitigate (and take advantage of) the ramifications of Maxwell's equations.

[7] E.g., Taflove, "Why study electromagnetics", *IEEE APM*, **44**, 132, 2002.

Details and study plan:

So much for the big picture about fields and waves encountered in nature and engineering systems and circuits. Working details of how fields and wave effects can be computed and characterized will be provided in the remaining parts of these notes.

Over the course of 39 lectures we will develop and study, in succession, the equations and applications of electrostatics (Lectures 1-11), magnetism (Lectures 12-15), and electromagnetics (Lectures 16-39) with a focus on time varying (quasi-static as well as wave-like) phenomena.

ECE 329:

We start by finding out how the equations of *electrostatics* arise from the familiar Coulomb's law (like charges repel, unlike charges attract) and the idea of field superpositions. We learn to solve electrostatic problems using the notion of *electrostatic potential* (voltage) and develop the notions of *polarization*, *conduction*, *charge continuity*, and *capacitance* in quasi-static settings of practical importance.

Next we learn how magnetic fields arise from charges in motion (a relative concept depending on the reference frame of the observer) and develop the governing laws of *magnetostatics* (also an extension of Coulomb's law seen from different reference frames). The *vector potential* is introduced for magnetic field calculations from prescribed current configurations, and notions of *magnetization* and *inductance* are subsequently developed and applied in quasi-static settings.

Just like time-varying electric fields imply time-varying charge densities (or *vice versa*) in electro-quasi-statics (EQS), time-varying currents imply time-varying magnetic flux in magneto-quasi-statics (MQS). We also learn that time-varying magnetic-flux is accompanied by time varying electric fields — a key finding of Faraday's called *induced field* with paradigm shifting ramifications and applications — and requires the modification of curl-free electric field condition into a dynamic equation known as Faraday's law.

Finally, the full set of Maxwell's equations is reached after adding a time-varying electric field term to the curl equation of magnetostatics. This change, first introduced by Maxwell in order to make sure that the governing equations of electricity and magnetism are consistent with *conservation of charge*, acknowledges the two-way coupling and feedback between electric and magnetic fields: time-varying magnetic fields induce time-varying electric fields — Faraday effect — and time-varying electric fields in turn induce time-varying magnetic fields (call it the "Maxwell effect") in order to sustain electromagnetic field variations in regions far away from charges and current loops — that is the way nature seems to work (and here we are to observe all that thanks to Maxwell effect allowing us to be here).

A study of wave solutions of Maxwell's equations follows, including *plane TEM waves* in free space, *linear and circular polarized* waves, waves in *conducting media*, *normal incidence* of waves on planar interfaces of homogeneous regions, *energy* and *momentum transfer*, *guided waves* in two-wire *transmission-line* (TL) systems, *transient response* on TL circuits, *resonant* *oscillations* in TL cavities, sinusoidal *steady-state* analysis of TL's and *distributed circuits*, *Smith Chart* applications, and finally *losses* in TL systems.

That is the full scope of the 39 lectures of ECE 329 — the course ends with an intensive study of distributed circuit concepts based on transmission lines, a study that complements the lumped circuit techniques examined and mastered in earlier courses.

ECE 329 is only the first half of our first-pass study of the fields and waves topics essential in electrical engineering education. Important topics such as radiation and antennas (generation details of electromagnetic waves by time-varying currents) and dispersion (frequency dependence of wave propagation speeds in material media) are barely mentioned or not at all in ECE 329. These constitute the main topics of the follow-on course, ECE 350.

ECE 350:

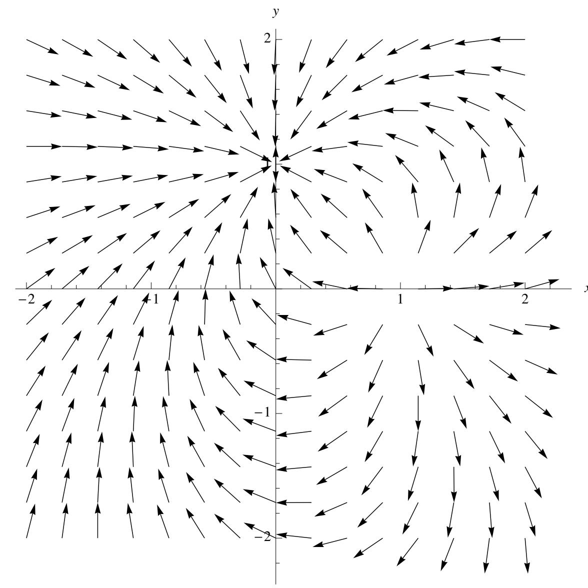
ECE 350 starts with the discussion of electromagnetic *radiation theory* and *transmission antennas*, continues with propagation and wave guidance effects (including *dispersion*, *phase and group velocities*, *Doppler shifts*, *oblique incidence*, *evanescence* and *tunneling* effects, *guided modes* in parallel-plate, rectangular, and dielectric slab *waveguides*), treats *cavity* fluctuations (including *resonant modes*, *blackbody radiation* in 3D cavities, *thermal noise*), and concludes with a discussion of *antenna reception* (including effective area, available power, link equations).

Beyond ECE 329 and 350:

Students having gone through ECE 329 and 350 will find themselves ready to encounter higher level courses in our curriculum focusing on different application areas and frequency regimes of the implications of Maxwell's equations. It is a life-long endeavor to master these relationships which have precipitated the scientific upheavals of the 20th century (relativity and quantum mechanics) and have remained intact and essential despite the upheavals unlike most aspects of classical physics. Our high speed electronics and communication networks and devices are intrinsically and fundamentally based on fields and wave concepts. Progress and innovation in these areas will require a deep understanding of fields and waves and how they interact with novel materials and structures.

Learn the basics and then go and invent the next thing!

1 Vector fields and Lorentz force



- Interactions between charged particles can be described and modeled[8] in terms of ***electric*** and ***magnetic*** **fields** just like gravity can be formulated in terms of ***gravitational fields*** of massive bodies.
  - In general, charge carrier dynamics and electromagnetic field variations[9] account for all electric and magnetic phenomena observed in nature and engineering applications.
- Electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$ generated by charge carriers — *electrons* and *protons* at microscopic scales — permeate all space with proper time delays, and combine additively.
  - Consequently we associate with each location of space having Cartesian coordinates

$$(x, y, z) \equiv \mathbf{r}$$

a pair of time-dependent *vectors*

$$\mathbf{E}(\mathbf{r}, t) = (E_x(\mathbf{r}, t), E_y(\mathbf{r}, t), E_z(\mathbf{r}, t))$$

[8] Interactions can also be formulated in terms of *past locations* (i.e., trajectories) of charge carriers. Unless the charge carriers are stationary — i.e., their past and present locations are the same — this formulation becomes impractically complicated compared to field based descriptions.

[9] Time-varying fields can exist even in the absence of charge carriers as we will find out in this course — light propagation in vacuum is a familiar example of this.

and

$$\mathbf{B}(\mathbf{r},t) = (B_x(\mathbf{r},t), B_y(\mathbf{r},t), B_z(\mathbf{r},t))$$

that we refer to as $\mathbf{E}$ and $\mathbf{B}$ for brevity (dependence on position $\mathbf{r}$ and time t is *implied*).

- Field vectors $\mathbf{E}$ and $\mathbf{B}$ and electric charge and current densities ρ and $\mathbf{J}$ — describing the distribution and motions of charge carriers — are related by (i.e., satisfy) a coupled set of linear constraints known as **Maxwell's equations,** shown in the margin.
  - Maxwell's equations are expressed in terms of divergence and curl of field vectors — recall MATH 241 — or, equivalently, in terms of closed surface and line integrals of the fields enclosing arbitrary volumes V and surfaces S in 3D space, as you have first seen in PHYS 212.
    - Maxwell's equations were "discovered" as a consequence of experimental and theoretical studies led by 19th century scientists including Gauss, Ampere, Faraday, and Maxwell.

  They remain intact and essential despite the scientific upheavals (paradigm shifts) of 20th century: relativity and quantum physics[10].

Maxwell's equations:

$$\begin{aligned} \nabla\cdot\mathbf{E} &= \frac{\rho}{\epsilon_o} \\ \nabla\cdot\mathbf{B} &= 0 \\ \nabla\times\mathbf{E} &= -\frac{\partial\mathbf{B}}{\partial t} \\ \nabla\times\mathbf{B} &= \mu_o\mathbf{J} + \mu_o\epsilon_o\frac{\partial\mathbf{E}}{\partial t}. \end{aligned}$$

such that

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v}\times\mathbf{B}),$$

with

$$\mu_o \equiv 4\pi\times10^{-7}\,\frac{\mathbf{H}}{\mathbf{m}},$$

and

$$\epsilon_o = \frac{1}{\mu_o c^2} \approx \frac{1}{36\pi\times10^9}\,\frac{\mathbf{F}}{\mathbf{m}},$$

in mksA units, where

$$c = \frac{1}{\sqrt{\mu_o\epsilon_o}} \approx 3\times10^8\,\frac{\mathbf{m}}{\mathbf{s}}$$

is the speed of light in free space.

(In Gaussian-cgs units $\frac{\mathbf{B}}{c}$ is used in place of $\mathbf{B}$ above, while $\epsilon_o = \frac{1}{4\pi}$ and $\mu_o = \frac{1}{\epsilon_o c^2} = \frac{4\pi}{c^2}$.)

[10]Fields are are utilized in different ways in classical and quantum electrodynamics, but Maxwell's field equations remain the same under both paradigms. Relativity theory is an updated model of space and time relations developed to achieve consistency with the implications of Maxwell's equations.

Given the charge and current densities ρ and $\mathbf{J}$, Maxwell's equations can be *solved* for the fields $\mathbf{E}$ and $\mathbf{B}$.

- Field solutions $\mathbf{E}$ and $\mathbf{B}$ in turn determine how a "test charge" q with mass m, position $\mathbf{r}$, and velocity $\mathbf{v} \equiv \dot{\mathbf{r}} = \frac{d\mathbf{r}}{dt}$ *accelerates* in accordance with **Lorentz force**

  $$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$$

  and Newton's 2nd law $\mathbf{F} = \frac{d}{dt}m\mathbf{v}$ (in classical electrodynamics). As such

  - **electric field E** at any location $\mathbf{r}$ is the vector force per stationary (i.e., $\mathbf{v} = 0$) unit charge (i.e., $q = 1$ C),
  - **magnetic field B** describes an additional force per unit charge which is experienced by charges in motion ($\mathbf{v} \neq 0$) in the reference frame — typically called the "lab frame" — where $\mathbf{E}$ and $\mathbf{v}$ are measured.

Lorentz force

Units in mksA system:

- q[=]C=sA,
- $\mathbf{E}$[=]N/C=V/m,
- $\mathbf{B}$[=]V.s/m^2=Wb/m^2=T,
- ρ[=]C/m^3,
- $\mathbf{J}$[=]A/m^2,

where C, N, V, Wb, and T are abbreviations for *Coulombs*, *Newtons*, *Volts*, *Webers*, and *Teslas*, respectively.

Charge q is quantized in units of $e = 1.602 \times 10^{-19}$ C, a relativistic invariant.

Since Lorentz force equation has the same form in all *inertial* reference frames[11] (like all laws of physics, including Maxwell's equations) while the charge velocity $\mathbf{v}$ is clearly frame-of-reference dependent, it follows that the values of fields $\mathbf{E}$ and $\mathbf{B}$ must also be dependent on the reference frame[12].

[11]Coordinate systems in which particles not subjected to any force — or, if general relativistic effects are to be retained, particles subjected to gravitational forces only — follow linearly varying trajectories.

[12]Given $\mathbf{E}$ and $\mathbf{B}$ measured in the lab, $\mathbf{E}'$ and $\mathbf{B}'$ measured by an observer moving through the lab with a constant velocity $\mathbf{v}$ are well approximated by $\mathbf{E}' \approx \mathbf{E} + \mathbf{v} \times \mathbf{B}$ and $\mathbf{B}' \approx \mathbf{B} - \frac{\mathbf{v}\times\mathbf{E}}{c^2}$ so long as $|\mathbf{v}| \ll c = 3 \times 10^8$ m/s, the *speed of light* in free space.

- Charge carrier positions $\mathbf{r}$, velocities $\dot{\mathbf{r}}$, and accelerations $\ddot{\mathbf{r}} = \frac{\mathbf{F}}{m}$, *as well as* forces $\mathbf{F}$, fields $\mathbf{E}$ and $\mathbf{B}$, and current density $\mathbf{J}$ are all described, in general, in terms of **3D vectors**.

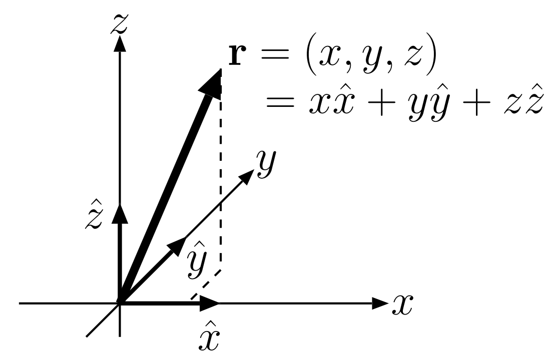
- In **Cartesian coordinates** such vectors and vector functions (of position $\mathbf{r}$ and/or time t) can be expressed in terms of mutually ***orthogonal*** **unit vectors** $\hat{x}$, $\hat{y}$, and $\hat{z}$ as in

  $$\mathbf{r} = (x, y, z) = x\hat{x}+y\hat{y}+z\hat{z} \quad \text{and} \quad \mathbf{E} = (E_x, E_y, E_z) = E_x\hat{x}+E_y\hat{y}+E_z\hat{z} \quad \text{etc.,}$$

  where

  - $|\mathbf{r}| \equiv \sqrt{x^2 + y^2 + z^2}$ and $|\mathbf{E}| \equiv \sqrt{E_x^2 + E_y^2 + E_z^2}$ etc., are **vector magnitudes**,
  - $\hat{r} \equiv \frac{\mathbf{r}}{|\mathbf{r}|}$ and $\hat{E} \equiv \frac{\mathbf{E}}{|\mathbf{E}|}$ etc., are associated **unit vectors**,
  - with ***dot products***
    - $\hat{r} \cdot \hat{r} = 1$, $\hat{E} \cdot \hat{E} = 1$, $\hat{x} \cdot \hat{x} = 1$, etc., but
    - $\hat{x} \cdot \hat{y} = \hat{x} \cdot \hat{z} = \hat{y} \cdot \hat{z} = 0$
  - and ***cross products***
    - $\hat{x} \times \hat{y} = \hat{z}$,
    - $\hat{y} \times \hat{z} = \hat{x}$,
    - $\hat{z} \times \hat{x} = \hat{y}$,

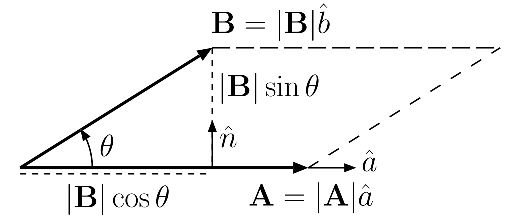
  adopting a **right-handed convention** (see the margin note in the next page).



UNIT VECTORS AND A POSITION VECTOR IN RIGHT-HANDED CARETESIAN COORDINATES

- Recall that
  - *Dot product* $\mathbf{A}\cdot\mathbf{B}$ is defined as $|\mathbf{A}|$ times $|\mathbf{B}|$ times the cosine of angle θ between $\mathbf{A}$ and $\mathbf{B}$.
    - Thus dot product is zero when angle θ is $90°$, as in the case of $\hat{x}$ and $\hat{y}$, etc.
  - *Cross product* $\mathbf{A}\times\mathbf{B}$ is defined as a vector with a magnitude $|\mathbf{A}|$ times $|\mathbf{B}|$ times the sine of angle θ between $\mathbf{A}$ and $\mathbf{B}$ and a direction orthogonal to both $\mathbf{A}$ and $\mathbf{B}$ in a **right-handed** sense (see margin note) .
    - Thus cross product is zero when the vectors cross multiplied are collinear ($\theta = 0°$) or anti-linear ($\theta = 180°$).

Right handed convention: cross product vector points in the direction indicated by the thumb of your ***right hand*** when you rotate your fingers from vector $\mathbf{A}$ toward vector $\mathbf{B}$ through angle θ you decide to use.



$\mathbf{A}\cdot\mathbf{B} = |\mathbf{A}||\mathbf{B}|\cos\theta$
DOT PRODUCT:product of projected vector lengths

$\mathbf{A}\times\mathbf{B} = |\mathbf{A}||\mathbf{B}|\sin\theta\hat{a}\times\hat{n}$
CROSS PRODUCT: right-handed perpendicular area vector of the parallelogram formed by co-planar vectors

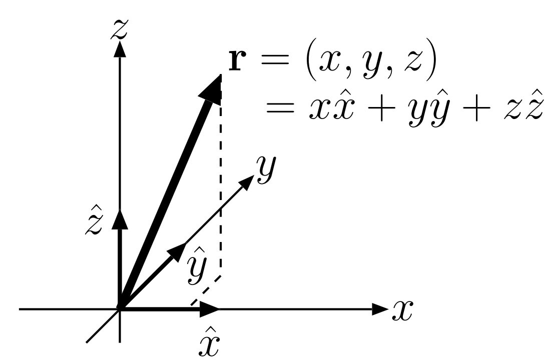
Example 1: Given the vectors $\mathbf{v} = (5, 10, 0)$ and $\mathbf{B} = (0, 0, 2)$ compute the cross and dot products $\mathbf{v}\times\mathbf{B}$ and $\mathbf{v}\cdot\mathbf{B}$.

Solution: Since we can also write $\mathbf{v} = 5\hat{x} + 10\hat{y}$ and $\mathbf{B} = 2\hat{z}$, it follows that

$$\mathbf{v}\times\mathbf{B} = (5\hat{x} + 10\hat{y}) \times 2\hat{z} = 10\hat{x}\times\hat{z} + 20\hat{y}\times\hat{z} = -10\hat{y} + 20\hat{x}.$$

Alternatively, using the well known determinant method for cross products,

$$\mathbf{v}\times\mathbf{B} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ 5 & 10 & 0 \\ 0 & 0 & 2 \end{vmatrix} = \hat{x}(10\cdot 2 - 0\cdot 0) - \hat{y}(5\cdot 2 - 0\cdot 0) + \hat{z}(5\cdot 0 - 10\cdot 0) = 20\hat{x} - 10\hat{y}.$$



UNIT VECTORS AND A POSITION VECTOR IN RIGHT-HANDED CARETESIAN COORDINATES

Also, $\mathbf{v}\cdot\mathbf{B} = (5,10,0)\cdot(0,0,2) = 5\cdot 0 + 10\cdot 0 + 0\cdot 2 = 0$.

Having three non-colinear force measurements $\mathbf{F}_i$ corresponding to three distinct test particle velocities $\mathbf{v}_i$ is sufficient to determine the fields E and B at any location in space produced by distant sources as illustrated by this example.

Example 2: A particle with charge $q = 1$ C passing through the origin $\mathbf{r} = (x,y,z) = 0$ of the lab frame is observed to accelerate with forces

$$\mathbf{F}_1 = 2\hat{x},\quad \mathbf{F}_2 = 2\hat{x} - 6\hat{z},\quad \mathbf{F}_3 = 2\hat{x} + 9\hat{y}\,\mathrm{N}$$

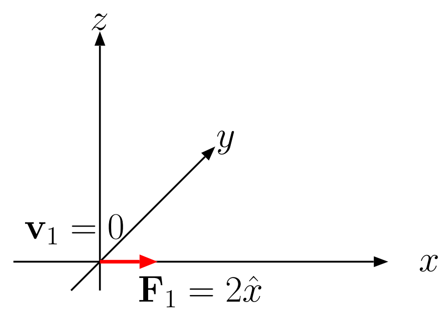
when the velocity of the particle is

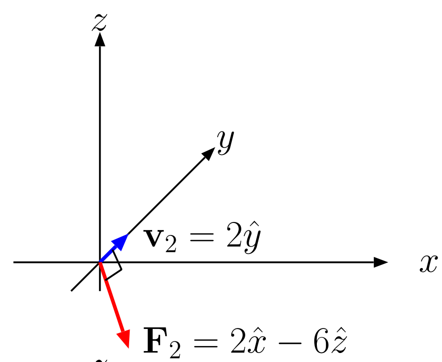
$$\mathbf{v}_1 = 0,\quad \mathbf{v}_2 = 2\hat{y},\quad \mathbf{v}_3 = 3\hat{z}\,\frac{\mathrm{m}}{\mathrm{s}},$$

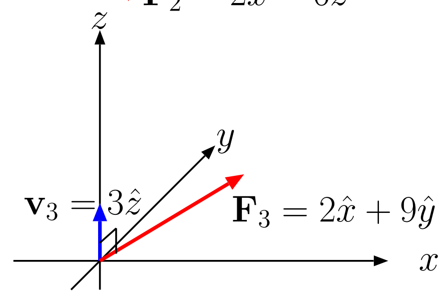
in turns. Use the Lorentz force equation

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v}\times\mathbf{B})$$

to determine the fields $\mathbf{E}$ and $\mathbf{B}$ at the origin.







Solution: Using the Lorentz force formula first with $\mathbf{F} = \mathbf{F}_1$ and $\mathbf{v} = \mathbf{v}_1$, we note that

$$2\hat{x} = (1)(\mathbf{E} + 0\times\mathbf{B}),$$

which implies that

$$\mathbf{E} = 2\hat{x}\,\frac{\mathrm{N}}{\mathrm{C}} = 2\hat{x}\,\frac{\mathrm{V}}{\mathrm{m}}.$$

Next, we use

$$\mathbf{v}\times\mathbf{B} = \frac{\mathbf{F}}{q} - \mathbf{E} = \frac{\mathbf{F}}{q} - 2\hat{x}$$

with $\mathbf{F}_2 = 2\hat{x} - 6\hat{z}$ and $\mathbf{v}_2 = 2\hat{y}$, as well as $\mathbf{E} = 2\hat{x}$ V/m, to obtain

$$2\hat{y}\times\mathbf{B} = -6\hat{z} \;\Rightarrow\; \hat{y}\times\mathbf{B} = -3\hat{z};$$

likewise, with $\mathbf{F}_3 = 2\hat{x} + 9\hat{y}$ and $\mathbf{v}_3 = 3\hat{z}$,

$$3\hat{z} \times \mathbf{B} = 9\hat{y} \quad \Rightarrow \quad \hat{z} \times \mathbf{B} = 3\hat{y}.$$

Substitute $\mathbf{B} = B_x\hat{x} + B_y\hat{y} + B_z\hat{z}$ in above relations to obtain

$$\hat{y} \times (B_x\hat{x} + B_y\hat{y} + B_z\hat{z}) = -B_x\hat{z} + B_z\hat{x} = -3\hat{z}$$

and

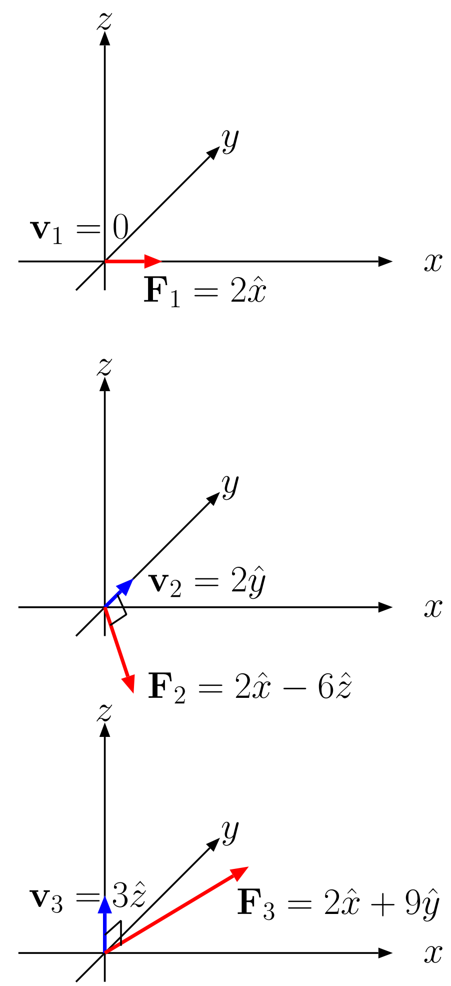
$$\hat{z} \times (B_x\hat{x} + B_y\hat{y} + B_z\hat{z}) = B_x\hat{y} - B_y\hat{x} = 3\hat{y}.$$

Matching the coefficients of $\hat{x}$, $\hat{y}$, and $\hat{z}$ in each of these relations we find that

$$B_x = 3\,\frac{\mathrm{Wb}}{\mathrm{m}^2}, \quad \text{and} \quad B_y = B_z = 0.$$

Hence, vector

$$\mathbf{B} = 3\hat{x}\,\frac{\mathrm{Wb}}{\mathrm{m}^2}.$$



- In your first homework you will be asked to do a sequence of vector exercises, including problems on volume, surface, and line integrals of vector or scalar functions of space (i.e., "fields"). These problems should be worked out with the help of your PHYS 212 and/or MATH 241 texts and notes.
  - **This course assumes a background of PHYS 212 and MATH 241 (on electromagnetic fields and vector calculus) as well**

as ECE 210 (lumped circuits and linear systems concepts including time- and frequency-domain approaches and phasors).

- The main objective of the course is to build up a **firm understanding of electromagnetic field concepts** introduced in PHYS 212, and to learn how to use **Maxwell's equations under static and time-varying conditions** associated with unguided (i.e., wireless) and guided (mainly transmission lines) electromagnetic waves. The study of **guided waves** is the key to extend the familiar **lumped-circuit** concepts into the realm of **distributed circuits**. This is the first half of a sequence of core electromagnetics courses in our curriculum, the second course being the 3-of-5 elective ECE 350.

Prerequisites:
MATH 241
PHYS 212
ECE 210

  – Topical outline:

    1. **Static electric fields, potential, polarization, quasi-static applications** (10 lectures)
    2. **Static currents and magnetic fields** (3 lectures)
    3. **Time-varying fields and Maxwell's eqns** (4 lectures)
    4. **Plane wave solutions of Maxwell's eqns** (9 lectures)
    5. **Guided waves in transmission lines and distributed circuits** (13 lectures)

Follow-on:
ECE 350

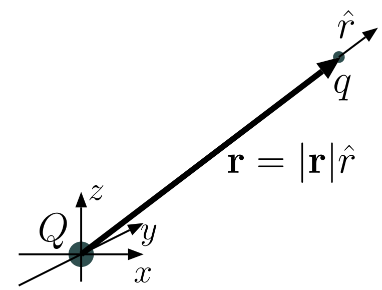
2 Static electric fields — Coulomb's and Gauss's laws

Static electric fields $\mathbf{E}(\mathbf{r})$ are produced by static (non-time-varying) distribution of charges and obey and the electrostatic laws shown in the margin where $\rho(\mathbf{r})$ denotes the net charge density in 3D volume. Over the next few lectures we will find out how these laws emerge from *Coulomb's law.*

Laws of electrostatics:

$$\nabla \cdot \mathbf{E} = \rho/\epsilon_o$$
$$\nabla \times \mathbf{E} = 0$$

At the most elementary level, each stationary point charge (electron or proton) Q is surrounded by its radially directed electrostatic field $\mathbf{E}$ given by Coulomb's law, and in the presence of multiple charges the field vectors of all the charges are added vectorially (linear superposition holds) to obtain a superposition field $\mathbf{E}$.



Force exerted by Q on q:

$$\mathbf{F} = q\mathbf{E}$$

with electric field

$$\mathbf{E} = \frac{Q}{4\pi\epsilon_o|\mathbf{r}|^2}\hat{r}$$

With multiple Q's superpose multiple E's

- **Coulomb's law** specifies the electric field of a stationary charge Q at the origin as

  $$\mathbf{E}(\mathbf{r}) = \frac{Q}{4\pi\epsilon_o r^2}\hat{r}$$

  as a function of position vector $\mathbf{r} = (x, y, z)$, where $\epsilon_o \approx \frac{1}{36\pi\times 10^9}$ F/m is a scaling constant known as *permittivity of free space*,

  $$r = |\mathbf{r}| = \sqrt{x^2 + y^2 + z^2}$$

  is radial distance from the charge, and $\hat{r} = \frac{\mathbf{r}}{r}$ radial unit vector pointing away from the charge.

  – This Coulomb field $\mathbf{E}(\mathbf{r})$ will exert a force $\mathbf{F} = q\mathbf{E}(\mathbf{r})$ on any stationary "test charge" q brought within distance r of Q (see figure in the margin).

If $qQ > 0$, force $\mathbf{F}$ is repulsive (directed along $\hat{r}$), if $qQ < 0$ it is attractive — *like charges repel, unlike charges attract.*

The existence of a Coulomb field accompanying each charge carrier in its rest frame[1] is taken to be a fundamental property of charge carriers (established by measurements).

- When multiple static charges Q_n are present in a region, the force on a stationary test charge q can be described as $q\mathbf{E}$ in terms of a superposition field

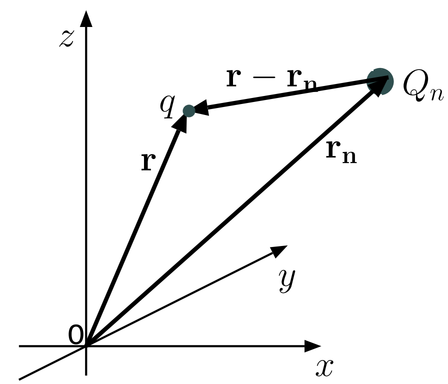
$$\mathbf{E} = \sum_n \frac{Q_n}{4\pi\epsilon_o r_n^2}\hat{r}_n$$

written in terms of the magnitudes and directions of vectors $\mathbf{r}_n$ pointing from each Q_n to q.

  – Equivalently, we can write

$$\mathbf{E}(\mathbf{r}) = \sum_n \frac{Q_n}{4\pi\epsilon_o|\mathbf{r}-\mathbf{r}_n|^2}\frac{\mathbf{r}-\mathbf{r}_n}{|\mathbf{r}-\mathbf{r}_n|},$$

where $\mathbf{r}$ and $\mathbf{r}_n$ now denote the locations of q and Q_n with respect to a common origin — this form is more convenient when static electric field $\mathbf{E}$ is to be calculated for an arbitrary location $\mathbf{r}$ (independent of the test charge notion).



Position vectors of charges are referenced with respect to a common origin O

[1]In non-inertial rest frames charge carriers will also produce an *additional* field proportional to the *acceleration* of free particles observed in such frames (e.g., Boyer, *Am. J. Phys.*, **47**, 129, 1979; Gupta and Padmanabhan, *Phys. Rev. D*, **57**, 7241, 1998).

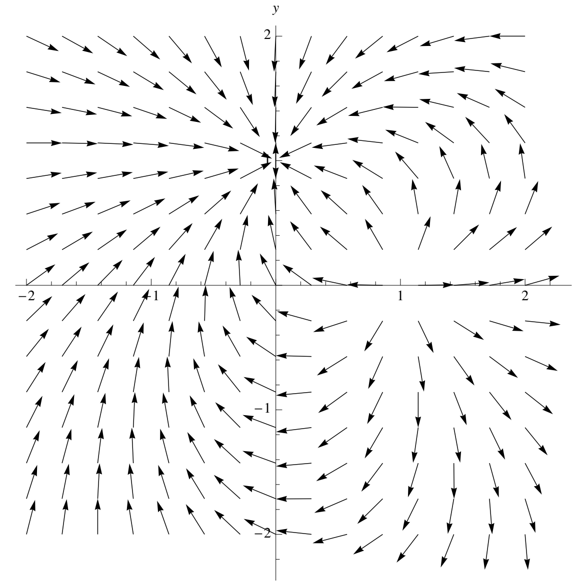
Example 1: Charges $Q_1 = 4\pi\epsilon_o$ and $Q_2 = -2Q_1$ are located at coordinates $\mathbf{r}_1 = (1,0,0) = \hat{x}$ and $\mathbf{r}_2 = (0,1,0) = \hat{y}$, respectively. What is the expression for $\mathbf{E}(\mathbf{r})$ and what is the explicit value of vector $\mathbf{E}(0)$?

Solution: Field $\mathbf{E}$ due to Q_1 and Q_2 at an arbitrary point $\mathbf{r}$ can be obtained as

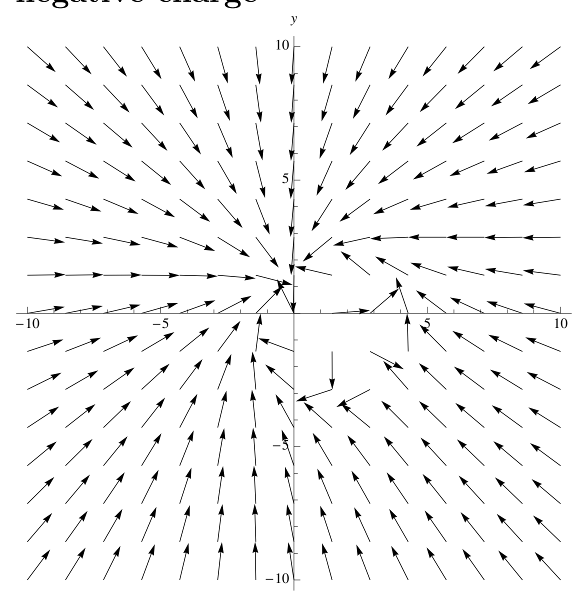
$$\begin{aligned}\mathbf{E}(\mathbf{r}) &= \frac{Q_1(\mathbf{r}-\mathbf{r}_1)}{4\pi\epsilon_o|\mathbf{r}-\mathbf{r}_1|^3} + \frac{Q_2(\mathbf{r}-\mathbf{r}_2)}{4\pi\epsilon_o|\mathbf{r}-\mathbf{r}_2|^3} \\ &= \frac{(\mathbf{r}-\hat{x})}{|\mathbf{r}-\hat{x}|^3} - \frac{2(\mathbf{r}-\hat{y})}{|\mathbf{r}-\hat{y}|^3} = \frac{(x-1,y,z)}{|(x-1,y,z)|^3} - \frac{2(x,y-1,z)}{|(x,y-1,z)|^3}\,\mathrm{V/m}.\end{aligned}$$

At the origin where $\mathbf{r} = (0,0,0)$, this result gives

$$\mathbf{E}(0,0,0) = \frac{(-1,0,0)}{|(-1,0,0)|^3} - \frac{2(0,-1,0)}{|(0,-1,0)|^3} = -\hat{x} + 2\hat{y}\,\mathrm{V/m}.$$



Field map of a dipole plus a negative charge



- The **vector map** shown in the margin depicts samples of unit vectors $\hat{E}(\mathbf{r}) \equiv \frac{\mathbf{E}(\mathbf{r})}{|\mathbf{E}(\mathbf{r})|}$ for the field $\mathbf{E}(\mathbf{r})$ obtained in Example 1 on a suitable grid established on xy-plane — such plots are useful or visualization purposes. Note that arrows emanate out of the positive charge at $(x,y) = (1,0)$ and converge upon the negative charge at $(x,y) = (0,1)$.
  - Electrostatic fields can be alternatively visualized in terms of so-called **field lines** or **flux lines**, continuous curves which are drawn tangential to unit vectors $\hat{E}(\mathbf{r})$ at every position $\mathbf{r}$. Try tracing out the flux lines over the vector map shown in the margin!

- According to Coulomb's law, electrostatic field of a charge Q placed at the origin points out in the radial direction $\hat{r}$ away from the origin and has a magnitude

$$E_r = \frac{Q}{4\pi\epsilon_o r^2}$$

that depends on radial distance r, but it does not depend on direction $\hat{r}$. The product of E_r with ϵ_o and the surface area of a sphere at radius r, namely, $S = 4\pi r^2$, yields

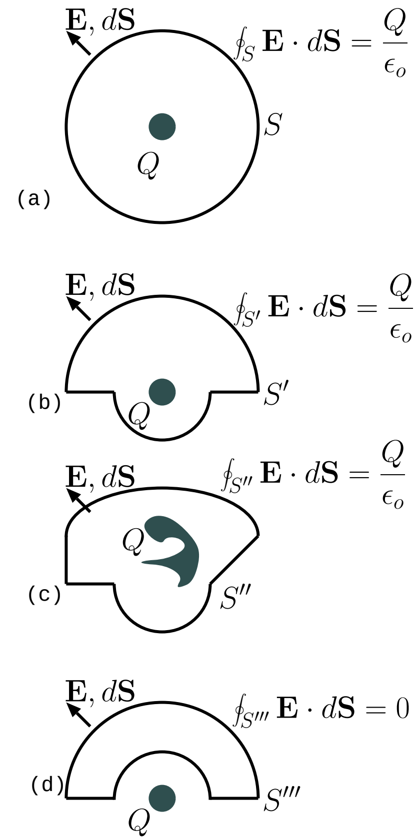
$$\epsilon_o E_r S = Q$$

independent of the radius of the sphere. Let's re-write the same result as

$$\epsilon_o \oint_S \mathbf{E} \cdot d\mathbf{S} = Q,$$

where

  - the "closed surface integral" $\oint_S \mathbf{E} \cdot d\mathbf{S}$ is called the **flux** of $\mathbf{E}$ over surface S bounding the volume $V = \frac{4\pi}{3} r^3$,
  - which in turn denotes the limiting value of the sum of dot products $\mathbf{E}_j \cdot \Delta\mathbf{S}_j$ computed over all surface elements of S having incremental areas $|\Delta\mathbf{S}_j|$ and unit vectors $\Delta\mathbf{S}_j / |\Delta\mathbf{S}_j|$ pointing away from volume V — the limiting value is obtained as all $|\Delta\mathbf{S}_j|$ approach zero (i.e., with increasingly finer subdivision of S into $|\Delta\mathbf{S}_j|$ elements).



Surface integral depends only on the net amount of charge contained within the surface --- charges outside the surface don't matter; surface shape doesn't matter; also charge motion within the surface does not matter.

Although we obtained the equality $\epsilon_o \oint_S \mathbf{E} \cdot d\mathbf{S} = Q$ above only for a spherical surface S centered about charge Q, we can easily convince ourselves — see the sketches on the right — that the equality should hold even when we distort the shape of surface S and/or displace Q away from the center so long as we do not move Q outside of S. All such variations are permitted because of inverse r-square dependence of the Coulomb's law and additive nature of fields, and if Q is moved outside the surface then the surface integral (flux) simply goes to zero.

- Hence, given an arbitrary shaped volume V enclosed by an arbitrary shaped surface S and including a net electrical charge Q_V, and defining a **displacement** field

  Displacement
  $\mathbf{D} = \epsilon_o \mathbf{E} [=] \frac{\mathrm{F}}{\mathrm{m}} \frac{\mathrm{V}}{\mathrm{m}} = \frac{\mathrm{C}}{\mathrm{m}^2}$

  $$\mathbf{D} \equiv \epsilon_o \mathbf{E},$$

  we obtain

  $$\oint_S \mathbf{D} \cdot d\mathbf{S} = Q_V, \quad \text{Gauss's law}$$

  a constraint known as **Gauss's law**. At this stage, the introduction of $\mathbf{D}$ is simply a notational convenience.

  Gauss's law

Gauss's law offers an alternative to *implementing an explicit sum of Coulomb fields* for calculating static field distributions $\mathbf{E}$ or $\mathbf{D} = \epsilon_o \mathbf{E}$ — the alternative method can be used when charge distributions have simplifying symmetry properties as will be illustrated in the next set of examples.

Also, later on we will learn that Gauss's law is valid even when charges Q_V within volume V are non-static (i.e., in motion), a condition under which Coulomb's law is no longer valid.

Example 2: Charged particles Q are located uniformly along the z-axis with an average line density of λ C/m extending from $z = -\infty$ to $+\infty$. We will compute the electrostatic field $\mathbf{E}$ of this charge distribution at a distance r from z-axis.

Having an average charge density of λ C/m implies that individual charges Q are spaced from one another by a distance $\Delta z = \frac{Q}{\lambda}$ along the z-axis. Assuming that charge locations are $z = n\Delta z$, where n is any integer, and using Coulomb's law, we find that

$$\mathbf{E}(\mathbf{r}) = \sum_{n=-\infty}^{\infty} \frac{Q}{4\pi\epsilon_o|\mathbf{r} - \hat{z}n\Delta z|^2}\frac{\mathbf{r} - \hat{z}n\Delta z}{|\mathbf{r} - \hat{z}n\Delta z|} = \sum_{n=-\infty}^{\infty} \frac{\lambda\Delta z(\mathbf{r} - \hat{z}n\Delta z)}{4\pi\epsilon_o|\mathbf{r} - \hat{z}n\Delta z|^3},$$

which, for position $\mathbf{r} = r(\hat{x}\cos\phi + \hat{y}\sin\phi)$ on xy-plane, at a distance r to the z-axis, reduces to

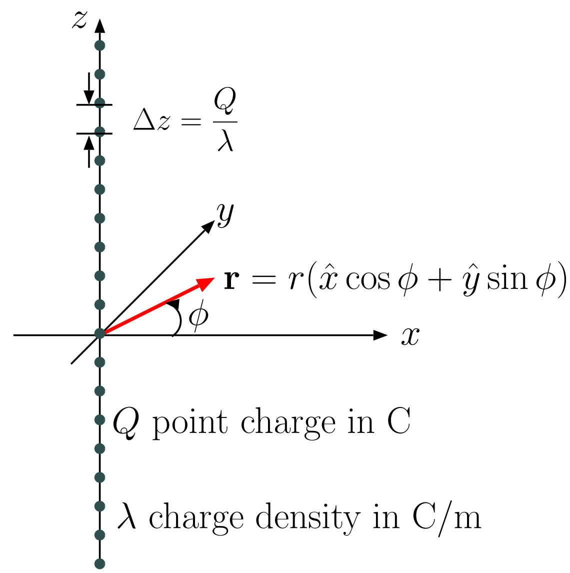
$$\mathbf{E} = \sum_{n=-\infty}^{\infty} \frac{\lambda r(\hat{x}\cos\phi + \hat{y}\sin\phi)}{4\pi\epsilon_o(r^2 + n^2\Delta z^2)^{3/2}}\Delta z \quad \text{(microscopic field)}$$

because the $\hat{z}$ component of $\mathbf{E}$ proportional to $n\Delta z$ cancels out (as a result of summation) due to symmetry in n. This field is "purely radial" in the direction

$$\hat{r} \equiv \hat{x}\cos\phi + \hat{y}\sin\phi$$

perpendicular to z-axis, and it can be evaluated, for $r \gg \Delta z$, as an integral (remember that sums of infinitesimals are *in effect* definite integrals)

$$\hat{r}\int_{-\infty}^{\infty} \frac{\lambda r}{4\pi\epsilon_o(r^2 + z^2)^{3/2}}dz = \hat{r}\frac{\lambda r}{4\pi\epsilon_o}\underbrace{\int_{-\infty}^{\infty} \frac{dz}{(r^2 + z^2)^{3/2}}}_{2/r^2} = \hat{r}\frac{\lambda}{2\pi\epsilon_o r}. \quad \text{(macroscopic field)}$$



- The result
$$\mathbf{E} = \hat{r}\frac{\lambda}{2\pi\epsilon_o r}$$
obtained above, valid for $r \gg \Delta z$, and labelled as ***macroscopic field***, also represents at any r (and z) the *space average* of the ***microscopic field*** taken over small volumes having dimensions of many Δz's (inter-particle separations).

  – In such a spatial average the *rapidly varying* structure of microscopic field (in particular at small r, caused by the discrete nature of charge distribution) is smoothed out as if electrical charge were distributed in space with a continuous density of λ C/m.

  – In realistic applications involving colossal numbers of charge carriers (of the order of 10^{23} in macroscopic chunks of solids) it is practical (and desirable) to focus our attention on macroscopic rather than microscopic fields.

We next illustrate how to obtain the macroscopic field $\mathbf{E} = \hat{r}\frac{\lambda}{2\pi\epsilon_o r}$ directly by using Gauss's law.

Solution using Gauss's law: We first notice that macroscopic electric field of a charge distribution along the z-axis having an average charge density of λ C/m should be pointing in radial direction $\hat{r}$ away from the z-axis (why?).

Also its magnitude E_r should be independent of azimuth angle ϕ by symmetry.

As a consequence, we can apply Gauss's law

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = Q_V$$

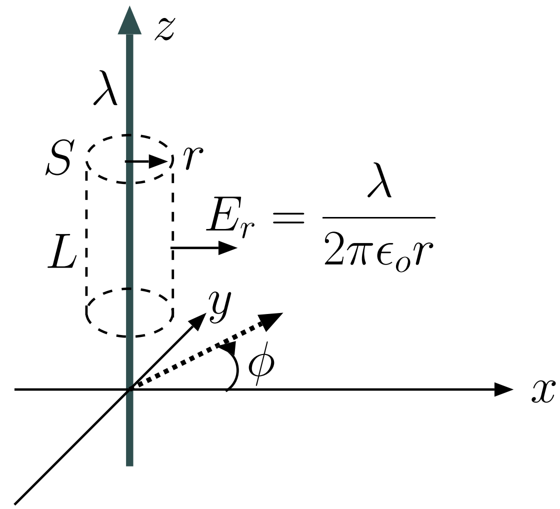
as

$$\epsilon_o E_r 2\pi r L = \lambda L$$

over the surface S of a cylindrical volume V of some length L and radius r centered about the z-axis as shown in the margin — notice that our "clever" choice of surface S in this problem resulted in the evaluation of the flux integral in Gauss's law without doing any calculus.

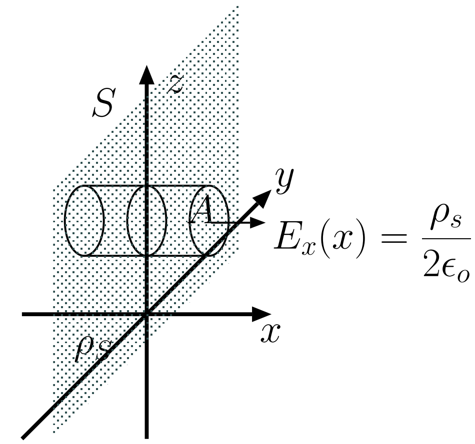
Clearly, this leads to (as obtained before using a line integral)

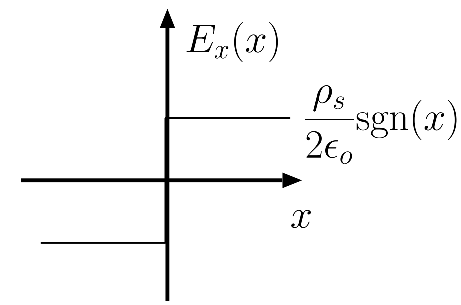
$$E_r = \frac{\lambda}{2\pi\epsilon_o r} \text{ and } \mathbf{E} = \frac{\lambda}{2\pi\epsilon_o r}\hat{r}.$$



3 Gauss's law and static charge densities

We continue with examples illustrating the use of Gauss's law in macroscopic field calculations:





Example 1: Point charges Q are distributed over $x = 0$ plane with an average surface charge density of ρ_s C/m^2. Determine the macroscopic electric field $\mathbf{E}$ of this charge distribution using Gauss's law.

Solution: First, invoking Coulomb's law, we convince ourselves that the field produced by surface charge density ρ_s C/m^2 on $x = 0$ plane will be of the form $\mathbf{E} = \hat{x}E_x(x)$ where $E_x(x)$ is an odd function of x because y- and z-components of the field will cancel out due to the symmetry of the charge distribution. In that case we can apply Gauss's law over a cylindrical integration surface S having circular caps of area A parallel to $x = 0$, and obtain

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = Q_V \quad \Rightarrow \quad \epsilon_o E_x(x)A - \epsilon_o E_x(-x)A = A\rho_s,$$

which leads, with $E_x(-x) = -E_x(x)$, to

$$E_x(x) = \frac{\rho_s}{2\epsilon_o} \text{ for } x > 0.$$

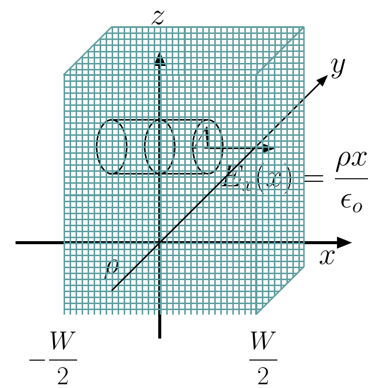
Hence, in vector form

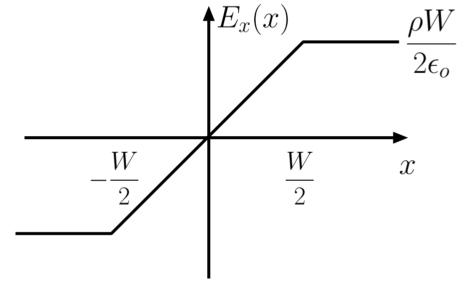
$$\mathbf{E} = \hat{x}\frac{\rho_s}{2\epsilon_o}\mathrm{sgn}(x),$$

where $\mathrm{sgn}(x)$ is the signum function, equal to ± 1 for $x \gtrless 0$.

Note that the macroscopic field calculated above is discontinuous at $x = 0$ plane containing the surface charge ρ_s, and points away from the same surface on both sides.

Example 2: Point charges Q are distributed throughout an infinite slab of width W located over $-\frac{W}{2} < x < \frac{W}{2}$ with an average charge density of ρ C/m^3. Determine the macroscopic electric field $\mathbf{E}$ of the charged slab inside and outside.





Solution: Symmetry arguments based on Coulomb's law once again indicates that we expect a solution of the form $\mathbf{E} = \hat{x}E_x(x)$ where $E_x(x)$ is an odd function of x.

In that case, applying Gauss's law with a cylindrical surface S having circular caps of area A parallel to $x = 0$ extending between $-x$ and $x < \frac{W}{2}$, we obtain

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = Q_V \quad \Rightarrow \quad \epsilon_o E_x(x)A - \epsilon_o E_x(-x)A = \rho 2xA,$$

which leads, with $E_x(-x) = -E_x(x)$, to

$$E_x(x) = \frac{\rho x}{\epsilon_o} \text{ for } 0 < x < \frac{W}{2}.$$

For $x > \frac{W}{2}$,

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = Q_V \quad \Rightarrow \quad \epsilon_o E_x(x)A - \epsilon_o E_x(-x)A = AW\rho,$$

leading to

$$E_x(x) = \frac{\rho W}{2\epsilon_o} \text{ for } x > \frac{W}{2}.$$

These results can be combined as

$$\mathbf{E} = \hat{x}E_x(x) = \begin{cases} -\hat{x}\frac{\rho W}{2\epsilon_o}, & \text{for } x < -\frac{W}{2} \\ \hat{x}\frac{\rho x}{\epsilon_o}, & \text{for } -\frac{W}{2} < x < \frac{W}{2} \\ \hat{x}\frac{\rho W}{2\epsilon_o}, & \text{for } x > \frac{W}{2}. \end{cases}$$

Note that the field solution depicted in the margin in terms of $E_x(x)$ plot is a continuous function of x as opposed to the discontinuous $E_x(x)$ solution obtained in Example 1 for the macroscopic field of a surface charge.

- In future calculations of electrostatic fields, we can use our previous results, namely

  - Coulomb field
  $$\mathbf{E} = \hat{r}\frac{Q}{4\pi\epsilon_o r^2} \text{ of a point charge } Q,$$
  - Field
  $$\mathbf{E} = \hat{r}\frac{\lambda}{2\pi\epsilon_o r} \text{ of constant line density } \lambda,$$
  - Field
  $$\mathbf{E} = \hat{x}\frac{\rho_s}{2\epsilon_o}\mathrm{sgn}(x) \text{ of constant surface density } \rho_s,$$
  - Field
  $$\mathbf{E} = \hat{x}\frac{\rho x}{\epsilon_o} \text{ of constant volume density } \rho$$

  as building blocks — that is, the above field equations can be superposed to determine the field structure of charge distributions $\rho(x, y, z)$ that can be expressed as superpositions of simpler charge distributions with known field structures. Some examples...

Example 3: Consider a pair of surface charges $\rho_s > 0$ and $-\rho_s$ C/m^2 of equal magnitudes placed on $x = -\frac{W}{2}$ and $x = \frac{W}{2}$ surfaces. Determine the electric field of this charge distribution depicted in the margin.

Solution: The field of charge density ρ_s C/m^2 on $x = -\frac{W}{2}$ plane should be

$$\mathbf{E}_+ = \hat{x}\frac{\rho_s}{2\epsilon_o}\mathrm{sgn}(x + \frac{W}{2}),$$

pointing away from the discontinuity surface at $x = -\frac{W}{2}$ on both sides. Likewise, the field of charge density $-\rho_s$ C/m^2 on $x = \frac{W}{2}$ plane should be

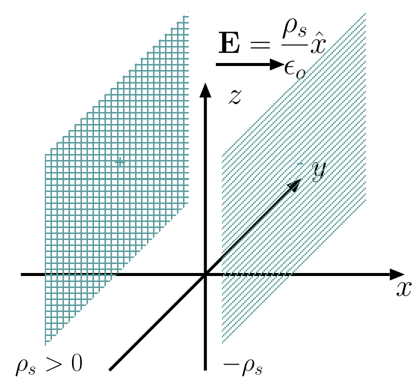
$$\mathbf{E}_- = -\hat{x}\frac{\rho_s}{2\epsilon_o}\mathrm{sgn}(x - \frac{W}{2}),$$

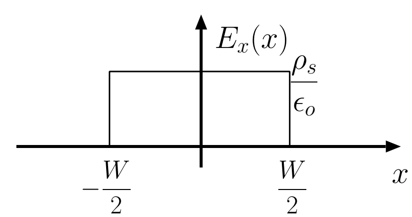
pointing toward $x = \frac{W}{2}$ surface from both sides. Superposing the two fields, we find that

$$\mathbf{E} = \mathbf{E}_+ + \mathbf{E}_- = \begin{cases} \hat{x}\frac{\rho_s}{\epsilon_o}, & \text{for } -\frac{W}{2} < x < \frac{W}{2}, \\ 0, & \text{otherwise,} \end{cases} = \hat{x}\frac{\rho_s}{\epsilon_o}\mathrm{rect}(\frac{x}{W})$$

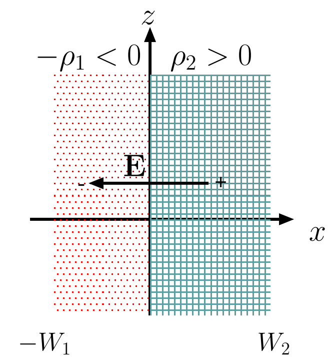
as depicted in the margin.

Note that the field lines of our solution point from positive charges on one surface to the negative charges resting on the other surface — this field has the structure of fields encountered in parallel plate capacitors that we will be studying soon.



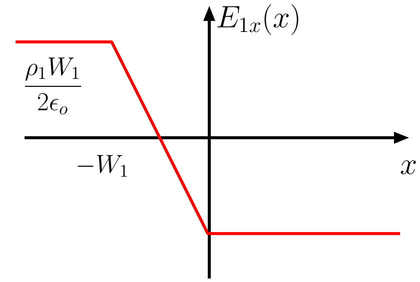


Example 4: An infinite charged slab of width W_1, located over $-W_1 < x < 0$, has a negative volumetric charge density of $-\rho_1$ C/m^3, $\rho_1 > 0$. A second slab of width W_2 and positive charge density ρ_2 is located over $0 < x < W_2$ as shown in the margin. Compute the electric field of this static charge configuration if $W_1\rho_1 = W_2\rho_2$, implying that the entire system is charge neutral (i.e., a net charge of zero).



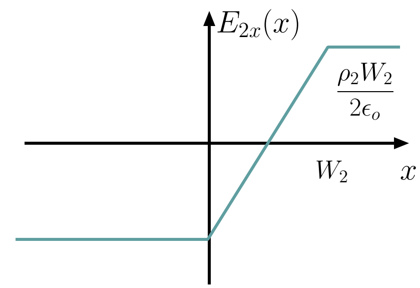
Solution: We note that the field of slab W_1 can be written as

$$\mathbf{E}_1 = \begin{cases} \hat{x}\frac{\rho_1 W_1}{2\epsilon_o}, & \text{for } x < -W_1 \\ -\hat{x}\frac{\rho_1(x+\frac{W_1}{2})}{\epsilon_o}, & \text{for } -W_1 < x < 0 \\ -\hat{x}\frac{\rho_1 W_1}{2\epsilon_o}, & \text{for } x > 0 \end{cases}$$



as depicted in the margin. Likewise, the field of slab W_2 is

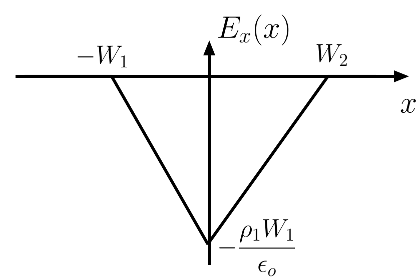
$$\mathbf{E}_2 = \begin{cases} -\hat{x}\frac{\rho_2 W_2}{2\epsilon_o}, & \text{for } x < 0 \\ \hat{x}\frac{\rho_2(x-\frac{W_2}{2})}{\epsilon_o}, & \text{for } 0 < x < W_2 \\ \hat{x}\frac{\rho_2 W_2}{2\epsilon_o}, & \text{for } x > W_2. \end{cases}$$



Note that field strengths $\frac{\rho_1 W_1}{2\epsilon_o}$ and $\frac{\rho_2 W_2}{2\epsilon_o}$ showing up in the expressions for $\mathbf{E}_1$ and $\mathbf{E}_2$ are equal because of the charge neutrality condition $W_1\rho_1 = W_2\rho_2$.

Consequently, when we superpose $\mathbf{E}_1$ and $\mathbf{E}_2$, the fields cancel out outside the region $-W_1 < x < W_2$, so that the total field becomes (as depicted in the margin)

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 = \begin{cases} -\hat{x}\frac{\rho_1(x+W_1)}{\epsilon_o}, & \text{for } -W_1 < x < 0 \\ \hat{x}\frac{\rho_2(x-W_2)}{\epsilon_o}, & \text{for } 0 < x < W_2 \\ 0, & \text{otherwise.} \end{cases}$$



- **Charge density** formalism which we find convenient to use for macroscopic field calculations can also be "adjusted" to describe the distributions of **isolated point charges** via the use of impulses or *delta functions* in space.

  – For example

  $$\rho(x, y, z) = Q\delta(x - x_o)\delta(y - y_o)\delta(z - z_o)$$

  can be regarded as a 3D volumetric charge density function representing a point charge Q located at a coordinate

  $$\mathbf{r} = (x, y, z) = (x_o, y_o, z_o) \equiv \mathbf{r}_o.$$

  ∘ This is justified because we can regard $\delta(x - x_o)$ to be zero everywhere except at $x = x_o$. By extension, the product

  $$\delta(x - x_o)\delta(y - y_o)\delta(z - z_o)$$

  is zero everywhere except at $\mathbf{r} = \mathbf{r}_o = (x_o, y_o, z_o)$ — therefore the density function $\rho(x, y, z)$ defined above behaves correctly to indicate the absence of charges everywhere with the exception of $\mathbf{r}_o$. Furthermore, the area property of the impulse implies that the volume integral of the charge density yields

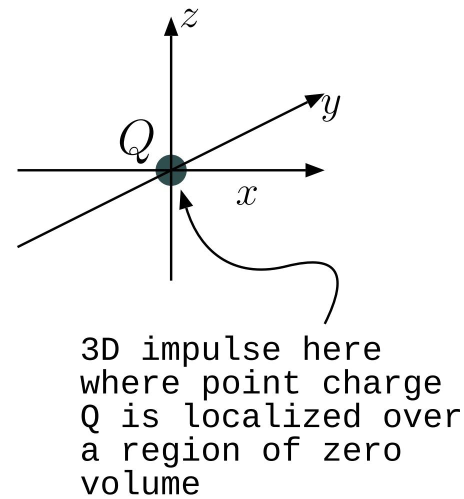
  $$\int \rho dV = \int\int\int Q\delta(x - x_o)\delta(y - y_o)\delta(z - z_o)dxdydz = Q$$

  as it should.

Gauss' Law in terms of charge density:

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = \int_V \rho dV$$

$$\rho(x, y, z) = Q\delta(x)\delta(y)\delta(z)$$



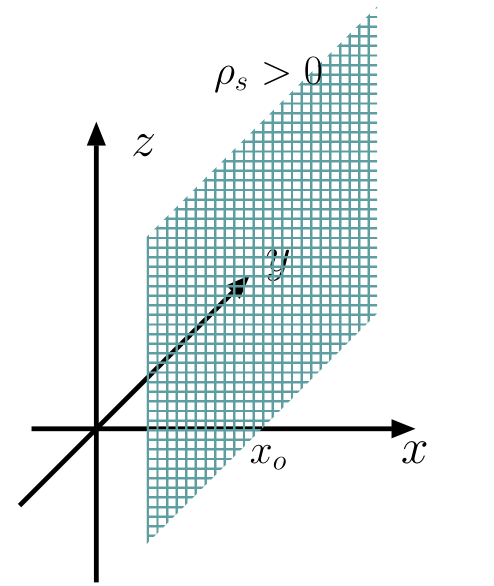
- $\circ$ Notice that the shifted impulses $\delta(x - x_o)$, etc., must have m^{-1} units in order to maintain dimensional consistency in the above expression.

- Another example is

$$\rho(x, y, z) = \rho_s(y, z)\delta(x - x_o)$$

representing a surface charge density of $\rho_s(y, z)$ C/m^2 on $x = x_o$ plane.

$\rho(x, y, z) = \rho_s(y, z)\delta(x - x_o)$

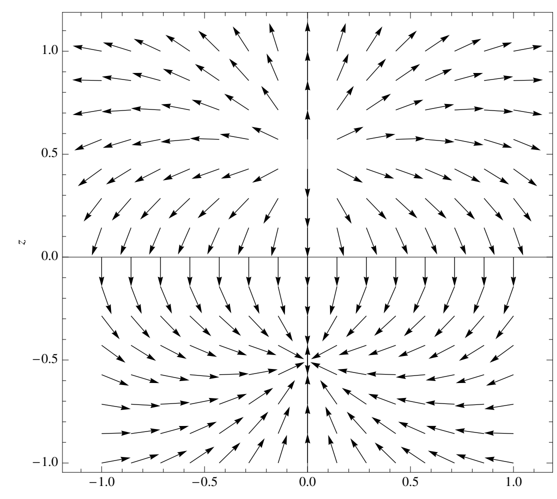


Example 5: Figure in the margin depicts (for the $d = 1$) the $\hat{E}$-field of a pair of charges $\pm Q$ located at $(0, 0, \pm\frac{d}{2})$ derived from

$$\begin{aligned}\mathbf{E}(\mathbf{r}) &= \frac{Q(\mathbf{r} - \frac{d}{2}\hat{z})}{4\pi\epsilon_o|\mathbf{r} - \frac{d}{2}\hat{z}|^3} + \frac{-Q(\mathbf{r} + \frac{d}{2}\hat{z})}{4\pi\epsilon_o|\mathbf{r} + \frac{d}{2}\hat{z}|^3} \\ &= \frac{Q}{4\pi\epsilon_o}[\frac{(x, y, z - \frac{d}{2})}{|(x, y, z - \frac{d}{2})|^3} - \frac{(x, y, z + \frac{d}{2})}{|(x, y, z + \frac{d}{2})|^3}] \text{ V/m.}\end{aligned}$$

Determine the electric flux $\int_{xy} \mathbf{E} \cdot d\mathbf{S}$ across the entire xy-plane using $d\mathbf{S} = -\hat{z}dxdy$.

Solution: Because of linearity, the flux we want to calculate equals the sum of the flux due to charge Q at $(0, 0, \frac{d}{2})$ above xy-plane and the flux due to charge $-Q$ at $(0, 0, -\frac{d}{2})$ above xy-plane.



Since by Gauss's law $\oint_S \mathbf{E}\cdot d\mathbf{S} = \frac{Q}{\epsilon_o}$ for any S surrounding Q, we can, by symmetry, infer that

$$\int_{xy} \mathbf{E}\cdot(-\hat{z}dxdy) = \frac{Q}{2\epsilon_o}$$

when only charge Q is considered — the logic here is, half of flux $\oint_S \mathbf{E}\cdot d\mathbf{S} = \frac{Q}{\epsilon_o}$ emanating from charge Q should go up and the remaining half should go down crossing the xy-plane in downward direction. Likewise, since $\oint_S \mathbf{E}\cdot d\mathbf{S} = \frac{-Q}{\epsilon_o}$ for any S surrounding $-Q$, again by symmetry, we can infer

$$\int_{xy} \mathbf{E}\cdot(-\hat{z}dxdy) = \frac{Q}{2\epsilon_o}$$

due to charge $-Q$ only — the logic in this case is, half of flux $\frac{Q}{\epsilon_o}$ "entering" charge $-Q$ is "coming from" above crossing the xy-plane in downward direction.

Thus, by superposition, we find total

$$\int_{xy} \mathbf{E}\cdot(-\hat{z}dxdy) = \frac{Q}{2\epsilon_o} + \frac{Q}{2\epsilon_o} = \frac{Q}{\epsilon_o}.$$

The above result can be *confirmed directly* by evaluating the integral

$$\begin{aligned}
\int_{xy} \mathbf{E}(x,y,0)\cdot(-\hat{z}dxdy) &= \int_{xy} \frac{Q}{4\pi\epsilon_o}\left[\frac{(x,y,-\frac{d}{2})}{|(x,y,-\frac{d}{2})|^3} - \frac{(x,y,\frac{d}{2})}{|(x,y,\frac{d}{2})|^3}\right]\cdot(-\hat{z}dxdy) \\
&= \frac{Q}{4\pi\epsilon_o}\int_{xy} \frac{d}{|(x,y,-\frac{d}{2})|^3}dxdy = \frac{Qd}{2\epsilon_o}\int_{r=0}^{\infty} \frac{r}{(r^2+(\frac{d}{2})^2)^{3/2}}dr \\
&= \frac{Q}{\epsilon_o}.
\end{aligned}$$

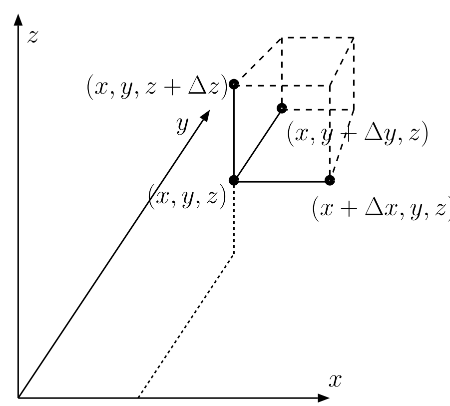
Just before the last step we have replaced $dxdy$ by $rdrd\phi$, where $r \equiv \sqrt{x^2+y^2}$, and carried out the ϕ integration before completing the r integration as a last step (which you should verify).

4 Divergence and curl

Expressing the total charge Q_V contained in a volume V as a 3D volume integral of charge density $\rho(\mathbf{r})$, we can write *Gauss's law* examined during the last few lectures in the general form

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = \int_V \rho dV.$$

This equation asserts that the flux of displacement $\mathbf{D} = \epsilon_o \mathbf{E}$ over any closed surface S equals the net electrical charge contained in the enclosed volume V — only the charges included within V affect the flux of $\mathbf{D}$ over surface S, with charges outside surface S making no net contribution to the surface integral $\oint_S \mathbf{D} \cdot d\mathbf{S}$.



- Gauss's law stated above holds true everywhere in space over all surfaces S and their enclosed volumes V, large and small.

- Application of Gauss's law to a small volume $\Delta V = \Delta x \Delta y \Delta z$ surrounded by a cubic surface ΔS of six faces, leads, in the limit of vanishing Δx, Δy, and Δz, to the differential form of Gauss's law expressed in terms of a **divergence operation** to be reviewed next:

  – Given a sufficiently small volume $\Delta V = \Delta x \Delta y \Delta z$, we can assume that
  $$\int_{\Delta V} \rho dV \approx \rho \Delta x \Delta y \Delta z.$$

– Again under the same assumption

$$\oint_S \mathbf{D}\cdot d\mathbf{S} \approx (D_{x|2}-D_{x|1})\Delta y\Delta z+(D_{y|4}-D_{y|3})\Delta x\Delta z+(D_{z|6}-D_{z|5})\Delta x\Delta y$$

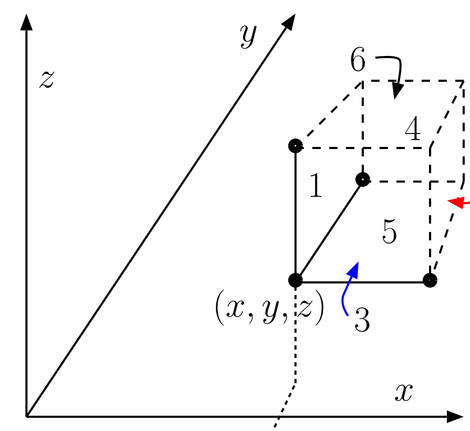
with reference to displacement vector components like $D_{x|2}$ shown on cubic surfaces depicted in the margin. Gauss's law demands the equality of the two expressions above, namely (after dividing both sides by $\Delta x\Delta y\Delta z$)

$$\frac{D_{x|2}-D_{x|1}}{\Delta x}+\frac{D_{y|4}-D_{y|3}}{\Delta y}+\frac{D_{z|6}-D_{z|5}}{\Delta z}\approx\rho,$$

in the limit of vanishing Δx, Δy, and Δz. In that limit, we obtain

$$\frac{\partial D_x}{\partial x}+\frac{\partial D_y}{\partial y}+\frac{\partial D_z}{\partial z}=\rho,$$

which is known as **differential form of Gauss's law**.



A more compact way of writing this result is

$$\nabla\cdot\mathbf{D}=\rho,$$

where the operator

$$\nabla\equiv(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}),$$

known as *del*, is applied on the displacement vector

$$\mathbf{D}=(D_x,D_y,D_z)$$

following the usual dot product rules, except that the product of $\frac{\partial}{\partial x}$ and D_x, for instance, is treated as a partial derivative $\frac{\partial D_x}{\partial x}$. In the left side above

$$\nabla \cdot \mathbf{D} = \frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z} \quad \text{(divergence of } \mathbf{D}\text{)}$$

is known as **divergence** of $\mathbf{D}$.

Example 1: Find the divergence of $\mathbf{D} = \hat{x}5x + \hat{y}12\ \mathrm{C/m^2}$

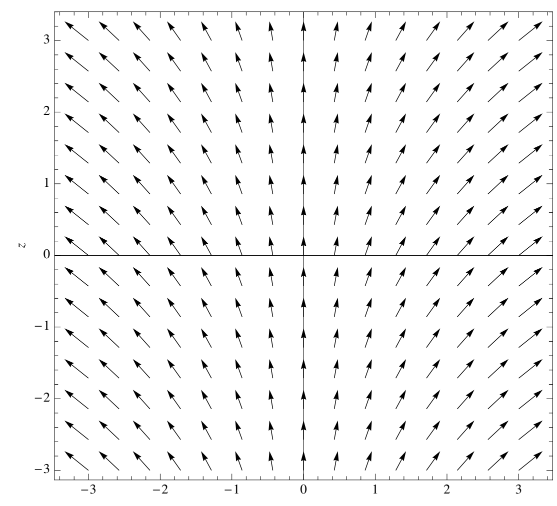
Solution: In this case

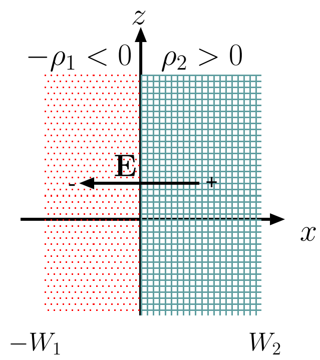
$$D_x = 5x, \quad D_y = 12, \quad \text{and} \quad D_z = 0.$$

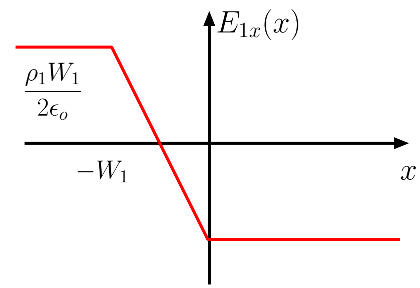
Therefore, divergence of $\mathbf{D}$ is

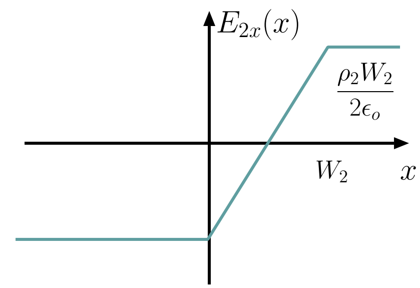
$$\begin{aligned} \nabla \cdot \mathbf{D} &= \frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z} \\ &= \frac{\partial}{\partial x}(5x) + \frac{\partial}{\partial y}(12) + \frac{\partial}{\partial z}(0) \\ &= 5 + 0 + 0 = 5\,\frac{\mathrm{C}}{\mathrm{m}^3}. \end{aligned}$$

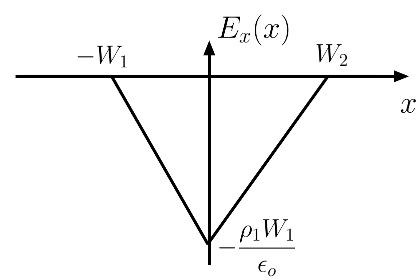
Note that the divergence of vector $\mathbf{D}$ is a scalar quantity which is the volumetric charge density in space as a consequence of Gauss's law (in differential form).











Example 2: Find the divergence $\nabla \cdot \mathbf{E}$ of electric field vector

$$\mathbf{E} = \begin{cases} -\hat{x}\frac{\rho_1(x+W_1)}{\epsilon_o}, & \text{for } -W_1 < x < 0 \\ \hat{x}\frac{\rho_2(x-W_2)}{\epsilon_o}, & \text{for } 0 < x < W_2 \\ 0, & \text{otherwise,} \end{cases}$$

from Example 4, last lecture (see margin figures).

Solution: In this case $E_y = E_z = 0$, and therefore the divergence of $\mathbf{E}$ is

$$\nabla \cdot \mathbf{E} = \frac{\partial E_x}{\partial x} = \frac{\partial}{\partial x}\begin{cases} -\frac{\rho_1(x+W_1)}{\epsilon_o}, \\ \frac{\rho_2(x-W_2)}{\epsilon_o}, \\ 0, \end{cases} = \begin{cases} -\frac{\rho_1}{\epsilon_o}, & \text{for } -W_1 < x < 0 \\ \frac{\rho_2}{\epsilon_o}, & \text{for } 0 < x < W_2 \\ 0, & \text{otherwise,} \end{cases},$$

which provides us with $\rho(\mathbf{r})/\epsilon_o$ of Example 4 from last lecture (in accordance with Gauss's law).

- Summarizing the results so far, Gauss's law can be expressed in *integral* as well as *differential* forms given by

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = \int_V \rho dV \quad \Leftrightarrow \quad \nabla \cdot \mathbf{D} = \rho.$$

  - The equivalence of integral and differential forms implies that (after integrating the differential form of the equation on the right over volume V on both sides)

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = \int_V \nabla \cdot \mathbf{D}\, dV$$

which you may recall as the **divergence theorem** from MATH 241.

Divergence thm.

– Note that according to divergence theorem, we can interpret **divergence as flux per unit volume**.

– We can also think of divergence as a special type of a derivative applied to vector functions which produces non-zero scalar results (at each point in space) when the vector function has components which change in the direction they point.

  ◦ A second type of vector derivative known as **curl** which we review next complements the divergence in the sense that these two types of vector derivatives collectively contain maximal information about vector fields that they operate on:

Given their curl and divergences, vector fields can be uniquely reconstructed in regions V of 3D space provided they are known at the bounding surface S of region V, however large (even infinite) S and V may be — this is known as **Helmholtz theorem** (proof outlined in Lecture 7).

- The **curl** of a vector field $\mathbf{E} = \mathbf{E}(x, y, z)$ is defined, in terms of the del operator ∇, like a cross product

$$\nabla \times \mathbf{E} \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \times (E_x, E_y, E_z) = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ E_x & E_y & E_z \end{vmatrix} \qquad \text{(curl of } \mathbf{E}\text{)}$$
$$= \hat{x}\left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}\right) - \hat{y}\left(\frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z}\right) + \hat{z}\left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\right).$$

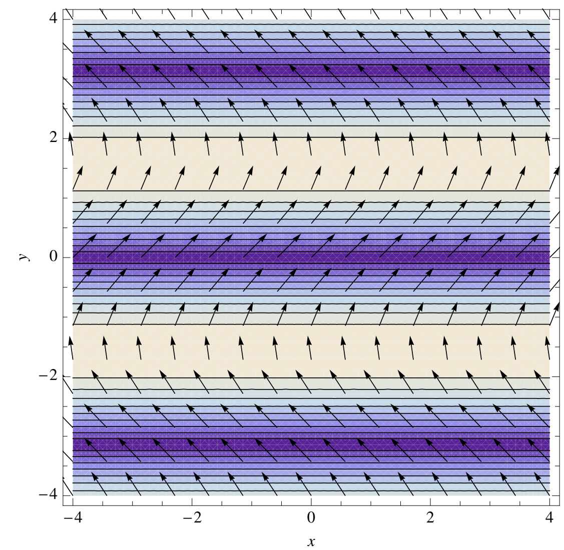
Example 3: Find the curl of the vector field

$$\mathbf{E} = \hat{x}\cos y + \hat{y}1$$

Solution: The curl is

$$\nabla \times \mathbf{E} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \cos y & 1 & 0 \end{vmatrix}$$
$$= \hat{x}\left(\frac{\partial}{\partial y}0 - \frac{\partial}{\partial z}1\right) - \hat{y}\left(\frac{\partial}{\partial x}0 - \frac{\partial}{\partial z}\cos y\right) + \hat{z}\left(\frac{\partial}{\partial x}1 - \frac{\partial}{\partial y}\cos y\right)$$
$$= \hat{x}0 - \hat{y}0 + \hat{z}(0 + \sin y) = \hat{z}\sin y$$

which is another vector field.



The diagram in the margin depicts $\mathbf{E} = \hat{x}\cos y + \hat{y}1$ as a vector map superposed upon a density plot of $|\nabla \times \mathbf{E}| = |\hat{z}\sin y| = |\sin y|$ indicating the strength of the curl vector $\nabla \times \mathbf{E}$ (light color corresponds large magnitude).

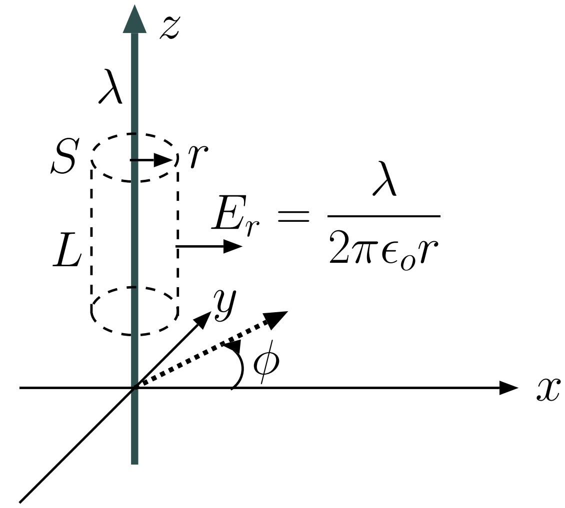
It is apparent that curl $\nabla \times \mathbf{E}$ is stronger in those regions where $\mathbf{E}$ is rapidly varying in directions orthogonal to the direction of $\mathbf{E}$ itself.

- As the above example demonstrates the curl of a vector field is in general another vector field.
  - The only exception is if the curl is identically 0 at all positions $\mathbf{r} = (x, y, z)$!
    - In that case, i.e., if $\nabla \times \mathbf{E} = 0$, vector field $\mathbf{E}$ is said to be **curl-free**.

IMPORTANT FACT: All static electric fields E, obtained from Coulomb's law, and satisfying Gauss's law $\nabla \cdot \mathbf{D} = \rho$ with static charge densities $\rho = \rho(\mathbf{r})$, are also found to be *curl-free* without exception.

- The proof of curl-free nature of static electric fields can be given by first showing that Coulomb field of a static charge is curl-free, and then making use of the superposition principle along with the fact that the curl of a sum must be the sum of curls — like differentiation, "taking curl" is a linear operation.
  - You should try to show that $\nabla \times \mathbf{E} = 0$ with the Coulomb field of a point charge Q located at the origin.

- The calculation is slightly more complicated than the following example (although similar in many ways) where we show that the static electric field of an infinite line charge is curl-free.



Example 4: Recall that the static field of a line charge λ distributed on the z-axis is

$$\mathbf{E}(x,y,z) = \hat{r}\frac{\lambda}{2\pi\epsilon_o r},$$

where

$$r^2 = x^2 + y^2 \text{ and } \hat{r} = \hat{x}\cos\phi + \hat{y}\sin\phi = (\frac{x}{r}, \frac{y}{r}, 0).$$

Show that field $\mathbf{E}$ satisfies the condition $\nabla \times \mathbf{E} = 0$.

Solution: Clearly, we can express vector $\mathbf{E}$ as

$$\mathbf{E} = \frac{\lambda}{2\pi\epsilon_o}(\frac{x}{r^2}, \frac{y}{r^2}, 0).$$

Since the components $\frac{x}{r^2}$ and $\frac{y}{r^2}$ of the vector are independent of z, the corresponding curl can be expanded as

$$\nabla \times \mathbf{E} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ E_x & E_y & E_z \end{vmatrix} = \frac{\lambda}{2\pi\epsilon_o}\begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{x}{r^2} & \frac{y}{r^2} & 0 \end{vmatrix} = \frac{\lambda}{2\pi\epsilon_o}\hat{z}(\frac{\partial}{\partial x}\frac{y}{r^2} - \frac{\partial}{\partial y}\frac{x}{r^2}).$$

But,

$$\frac{\partial}{\partial x}\frac{y}{r^2} - \frac{\partial}{\partial y}\frac{x}{r^2} = y\frac{\partial}{\partial x}\frac{1}{r^2} - x\frac{\partial}{\partial y}\frac{1}{r^2} = y\frac{-2x}{r^4} - x\frac{-2y}{r^4} = 0,$$

so $\nabla \times \mathbf{E} = 0$ as requested.

5 Curl-free fields and electrostatic potential

- Mathematically, we can generate a curl-free vector field $\mathbf{E}(x, y, z)$ as

$$\mathbf{E} = -(\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z}),$$

by taking the **gradient** of any scalar function $V(\mathbf{r}) = V(x, y, z)$. The gradient of $V(x, y, z)$ is defined to be the vector

$$\nabla V \equiv (\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z}),$$

pointing in the direction of increasing V; in abbreviated notation, curl-free fields $\mathbf{E}$ can be indicated as

$$\mathbf{E} = -\nabla V.$$

  - **Verification:** Curl of vector ∇V is

$$\nabla \times (\nabla V) = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial V}{\partial x} & \frac{\partial V}{\partial y} & \frac{\partial V}{\partial z} \end{vmatrix} = \hat{x}0 - \hat{y}0 - \hat{z}0 = 0.$$

  - If $\mathbf{E} = -\nabla V$ represents an **electrostatic field**, then V is called the **electrostatic potential**.
    - Simple dimensional analysis indicates that units of electrostatic potential must be volts (V).

– The prescription $\mathbf{E} = -\nabla V$, including the minus sign (optional, but taken by convention in electrostatics), ensures that electrostatic field $\mathbf{E}$ points from regions of "high potential" to "low potential" as illustrated in the next example.

Electrostatic fields E point from regions of "high V" to "low V"

Example 1: Given an electrostatic potential

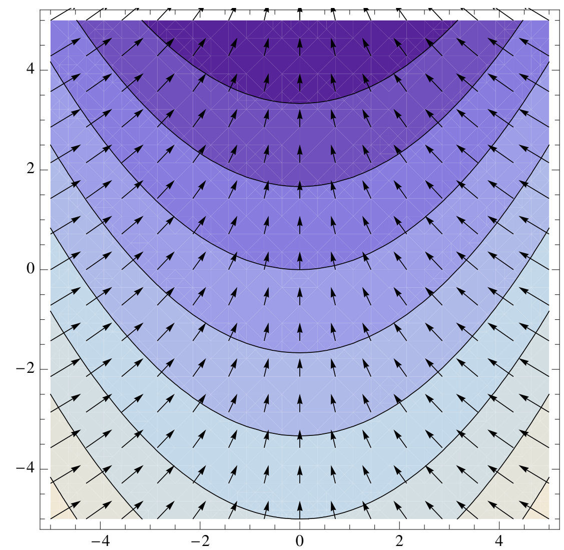
$$V(x,y,z) = x^2 - 6y\,\mathrm{V}$$

in a certain region of space, determine the corresponding electrostatic field $\mathbf{E} = -\nabla V$ in the same region.

Solution: The electrostatic field is

$$\mathbf{E} = -\nabla(x^2 - 6y) = -(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})(x^2 - 6y) = (-2x, 6, 0) = -\hat{x}\,2x + \hat{y}6\,\mathrm{V/m}.$$

Note that this field is directed from regions of high potential to low potential. Also note that electric field vectors are perpendicular everywhere to "equipotential" contours.



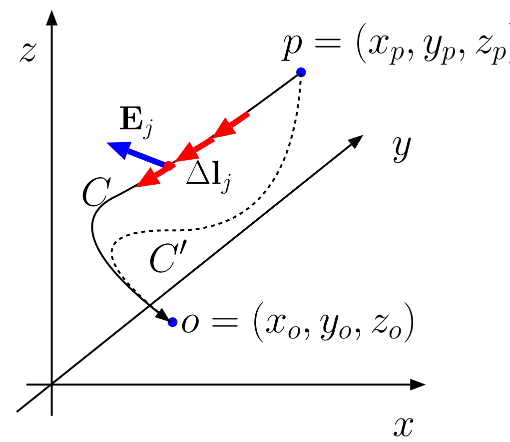
Light colors indicate "high V" dark colors "low V"

Given an electrostatic potential $V(x,y,z)$, finding the corresponding electrostatic field $\mathbf{E}(x,y,z)$ is a straightforward procedure (taking the negative gradient) as already illustrated in Example 1.

The reverse operation of finding $V(x,y,z)$ from a given $\mathbf{E}(x,y,z)$ can be accomplished by performing a **vector line integral**

$$\int_p^o \mathbf{E}\cdot d\mathbf{l}$$

in 3D space, since, as shown below, such integrals are "path independent" for curl-free fields $\mathbf{E} = -\nabla V$.



- The **vector line integral**

$$\int_p^o \mathbf{E} \cdot d\mathbf{l}$$

  over an integration path C extending from a point $p = (x_p, y_p, z_p)$ in 3D space to some other point $o = (x_o, y_o, z_o)$ is *defined to be*

  – the limiting value of the sum of dot products $\mathbf{E}_j \cdot \Delta\mathbf{l}_j$ computed over all sub-elements of path C having incremental lengths $|\Delta\mathbf{l}_j|$ and unit vectors $\Delta\mathbf{l}_j/|\Delta\mathbf{l}_j|$ directed from p towards o — the limiting value is obtained as all $|\Delta\mathbf{l}_j|$ approach zero (i.e., with increasingly finer subdivision of C into $|\Delta\mathbf{l}_j|$ elements).

- Computation of the integral (see example below) involves the use of infinitesimal displacement vectors

$$d\mathbf{l} = \hat{x}dx + \hat{y}dy + \hat{z}dz = (dx, dy, dz)$$

  and vector dot product

$$\mathbf{E} \cdot d\mathbf{l} = (E_x, E_y, E_z) \cdot (dx, dy, dz) = E_x dx + E_y dy + E_z dz.$$

  The integral

$$\int_p^o \mathbf{E} \cdot d\mathbf{l} = \int_p^o (E_x dx + E_y dy + E_z dz)$$

will in general be *path dependent* except for when $\mathbf{E}$ is curl-free.

Example 2: The field $\mathbf{E} = \hat{x}y \pm \hat{y}x$ is curl-free with the + sign, but not with − as verified below by computing $\nabla \times \mathbf{E}$. Calculate the line integral of $\mathbf{E}$ (for both signs, ±) from a point $o = (0, 0, 0)$ to point $p = (1, 1, 0)$ for two different paths C going through points $u = (0, 1, 0)$ and $l = (1, 0, 0)$, respectively (see margin).

Solution: First we note that

$$\nabla \times (\hat{x}y \pm \hat{y}x) = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y & \pm x & 0 \end{vmatrix} = \hat{z}(\pm 1 - 1)$$

which confirms that $\mathbf{E} = \hat{x}y \pm \hat{y}x$ is curl-free with with + sign, but not with −. In either case, the integral to be performed is

$$\int_o^p \mathbf{E} \cdot d\mathbf{l} = \int_o^p (E_x dx + E_y dy + E_z dz) = \int_o^p (y\,dx \pm x\,dy).$$

For the first path C_u going through $u = (0, 1, 0)$, we have

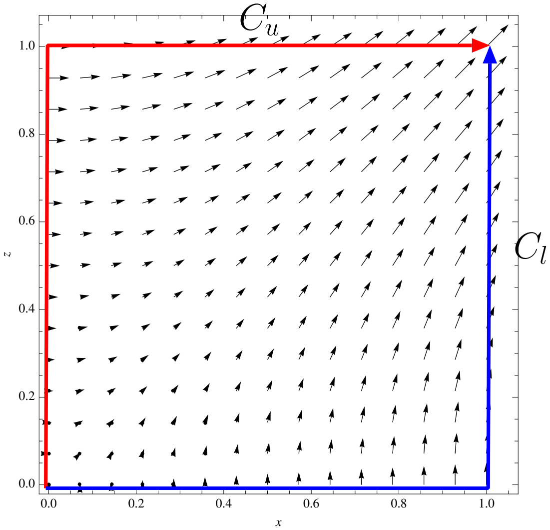
$$\int_o^p (y\,dx \pm x\,dy) = \int_{y=0}^1 (\pm x)\,dy_{|x=0} + \int_{x=0}^1 y\,dx_{|y=1} = 0 + 1 = 1.$$

For the second path C_l going through $l = (1, 0, 0)$, we have

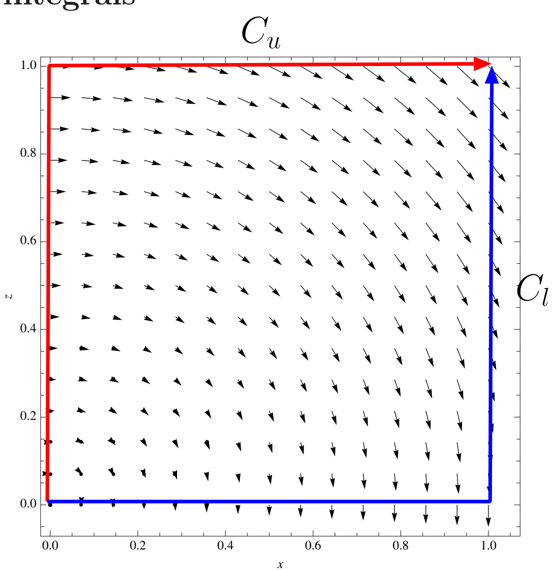
$$\int_o^p (y\,dx \pm x\,dy) = \int_{x=0}^1 y\,dx_{|y=0} \pm \int_{y=0}^1 x\,dy_{|x=1} = 0 \pm 1 = \pm 1.$$

Clearly, the result shows that the line integral $\int_o^p \mathbf{E} \cdot d\mathbf{l}$ is *path independent* for $\mathbf{E} = \hat{x}y + \hat{y}x$ which is curl-free, and path dependent for $\mathbf{E} = \hat{x}y - \hat{y}x$ in which case $\nabla \times \mathbf{E} \neq 0$.

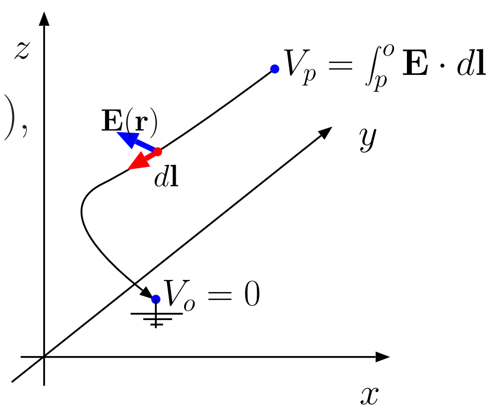
Curl-free: path-independent line integrals



"Curly": path-dependent line integrals



- The **mathematical reason** why curl-free fields have path-independent line integrals is because in those occasions the integrals can be written in terms of **exact differentials**:
  - for curl-free $\mathbf{E} = \hat{x}y + \hat{y}x$ we have $\mathbf{E}\cdot d\mathbf{l}$ as an *exact differential* $ydx + xdy = d(xy)$ of the function xy, in which case $\int_o^p \mathbf{E}\cdot d\mathbf{l} = xy|_o^p = (1\cdot 1 - 0\cdot 0) = 1$ over all paths.
  - for $\mathbf{E} = \hat{x}y - \hat{y}x$ with $\nabla\times\mathbf{E} = -2\hat{z} \neq 0$, on the other hand, $\mathbf{E}\cdot d\mathbf{l} = ydx - xdy$ does not form an exact differential $-dV$, and thus there is no path-independent integral $-V|_o^p$, nor an underlying potential function V.



$\mathbf{E}\cdot d\mathbf{l}$ is *guaranteed* to be an exact differential if $\mathbf{E} = -\nabla V = (-\frac{\partial V}{\partial x}, -\frac{\partial V}{\partial y}, -\frac{\partial V}{\partial z})$, since in that case the differential of $V(x,y,z)$, namely

$$dV \equiv \frac{\partial V}{\partial x}dx + \frac{\partial V}{\partial y}dy + \frac{\partial V}{\partial z}dz, \text{ is precisely } -E_x dx - E_y dy - E_z dz = -\mathbf{E}\cdot d\mathbf{l}.$$

- In that case

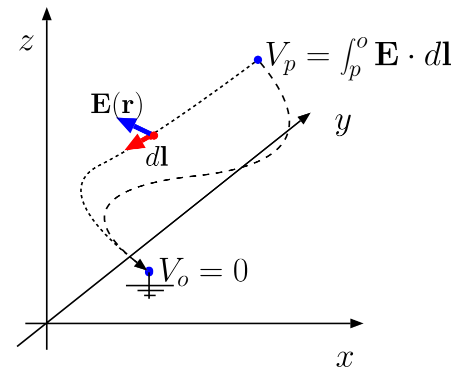
$$\int_p^o \mathbf{E}\cdot d\mathbf{l} = -\int_p^o dV = \int_o^p dV = V_p - V_o$$

is independent of integration path; thus, if we we call o the "ground", and set $V_o = 0$, then

$$V_p = \int_p^o \mathbf{E}\cdot d\mathbf{l}$$

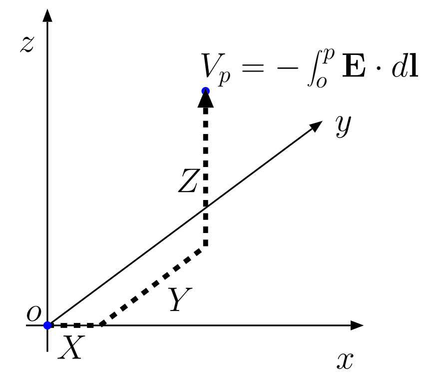
denotes the potential drop from (any) point p to ground o.

- The **physical reason** why this integral formula for potential V_p works with *any* integration path is the principle of **energy conservation**:
  - integral $\int_p^o \mathbf{E} \cdot d\mathbf{l}$, namely the "voltage drop" from p to o, represents the **work done** *per unit charge* by the field $\mathbf{E}$ in moving charges from location p to location o[1], so if the line integral were path-dependent (in reaching from p to o) there would be ways of creating net energy by making a charge q follow special closed paths within the electrostatic field $\mathbf{E}$, in violation of the general principle of energy conservation (that permits energy conversion but not creation or destruction).



As long as E is curl-free, line integral is path-independent and produces the voltage drop from point p to "ground" o.

[1]Either to increase the kinetic energy of the charge if charge transport from p to o is *unimpeded* (as for a test charge accelarating between a pair of capacitor plates) or else in pushing the charge against frictional forces (as through a resistive wire) both at the expense of the energy stored in the field. On the other hand, work done (i.e., the voltage drop) by the field would be *negative* if charges $q > 0$ were moved from p to o *against* the local electric field (as within a battery), in which case there would be a positive *voltage rise* from p to o representing energy gain for the field per unit charge transported from p to o.



Example 3: Given that $V_o = V(0,0,0) = 0$ and

$$\mathbf{E} = 2x\hat{x} + 3z\hat{y} + 3(y+1)\hat{z}\,\frac{\text{V}}{\text{m}},$$

determine the electrostatic potential $V_p = V(X,Y,Z)$ at point $p = (X,Y,Z)$ in volts.

Solution: Assuming that the field is curl-free (it is), so that any integration path can be used, we find that

$$\begin{aligned} V_p &= \int_p^o \mathbf{E}\cdot d\mathbf{l} = -\int_o^p \mathbf{E}\cdot d\mathbf{l} = -\int_o^p (2x\,dx + 3z\,dy + 3(y+1)\,dz) \\ &= -\int_0^X 2x\,dx_{|y,z=0} - \int_0^Y 3z\,dy_{|x=X,z=0} - \int_0^Z 3(y+1)\,dz_{|x=X,y=Y} \\ &= -X^2 - 0 - 3(Y+1)Z. \end{aligned}$$

This implies

$$V(x,y,x) = -x^2 - 3(y+1)z\,\text{V}.$$

Note that

$$\begin{aligned} -\nabla(-x^2 - 3(y+1)z) &= \nabla(x^2 + 3(y+1)z) \\ &= \hat{x}2x + \hat{y}3z + \hat{z}3(y+1) \end{aligned}$$

yields the original field $\mathbf{E}$, which is an indication that $\mathbf{E}$ is indeed curl-free.

A *voltmeter* with its (+/red) probe contacting point p and its (-/black) probe contacting point o would display (by analog or digital means) the numerical value of

$$V_p = \int_p^o \mathbf{E}\cdot d\mathbf{l}$$

with the integration path consisting of the path defined by the probe wires. The voltmeter reading would be independent of the path configuration when the field E is electrostatic.

For the voltmeter not to perturb the field it is probing, its input impedance need to be much greater than the impedance between points p and o.

Alternate Solution — Exact Differential Method: Note that

$$\begin{aligned}
\mathbf{E}\cdot\mathbf{dl} &= (2x\hat{x}+3z\hat{y}+3(y+1)\hat{z})\cdot(\hat{x}dx+\hat{y}dy+\hat{z}dz)\\
&= 2xdx+3zdy+3(y+1)dz = 2xdx+3(ydz+zdy)+3dz\\
&= d(x^2+3yz+3z) = -dV.
\end{aligned}$$

Therefore

$$V(x,y,z) = -x^2-3yz-3z+C,$$

where the integration constant C should chosen so that $V(0,0,0)=0$. The result is

$$V(x,y,z) = -x^2-3(y+1)z$$

as before.

Example 5: According to Coulomb's law electrostatic field of a proton with charge $Q = e$ (where $-e$ is electronic charge) located at the origin is given as

$$\mathbf{E} = \frac{e}{4\pi\epsilon_o r^2}\hat{r},$$

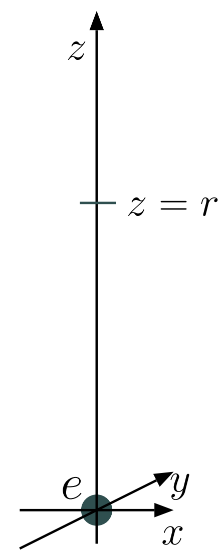
where

$$r = \sqrt{x^2 + y^2 + z^2} \quad \text{and} \quad \hat{r} = \frac{(x, y, z)}{r}.$$

Determine the electrostatic potential field V established by charge $Q = e$ with the provision that $V \to 0$ as $r \to \infty$ (i.e., ground at infinity).

Solution: Field $\mathbf{E}$ and its potential V will exhibit spherical symmetry in this problem. Therefore, with no loss of generality, we can calculate the line integral from a point p at a distance r from the origin to a point o at ∞ (the specified ground) along, say, the z-axis. Approaching the problem that way, the potential drop from r to ∞ is

$$\begin{aligned} V(r) &= \int_{z=r}^{\infty} \frac{e}{4\pi\epsilon_o z^2}\hat{z}\cdot\hat{z}dz \\ &= -\frac{e}{4\pi\epsilon_o z}\Big|_r^{\infty} = \frac{e}{4\pi\epsilon_o r}. \end{aligned}$$



Or else (exact differential method):

$$\mathbf{E}\cdot d\mathbf{l} = \left(\frac{e}{4\pi\epsilon_o r^2}\hat{r}\right)\cdot\hat{r}dr = d\left(\frac{-e}{4\pi\epsilon_o r}\right) = -dV$$

leading to

$$V(r) = \frac{e}{4\pi\epsilon_o r}$$

(using an integration constant of zero).

- To convert electrostatic potential V_p (in volts) at any point p to potential energy of a charge q brought to the same point, it is sufficient to multiply V_p with q (or just the sign of q, depending on which energy units we want to use — see the next example).

Example 6: In view of Example 5, what are the potential energies of a proton e and an electron $-e$ placed at distance $r = a$ away from the proton at the origin, where distance

$$a \equiv \frac{4\pi\epsilon_o}{e^2}\frac{\hbar^2}{m_e} = 0.529 \times 10^{-10}\,\mathrm{m}$$

stands for *Bohr radius* — it is the mean distance of the ground state electron in a hydrogen atom from the center of the atom. Recall that $e = 1.602 \times 10^{-19}$ C and $\epsilon_o \approx 10^{-9}/36\pi$ F/m.

Solution: Let's first evaluate the potential $V(r)$ at $r = a$:

$$V(a) = \frac{e}{4\pi\epsilon_o a} \approx \frac{(1.6 \times 10^{-19})36\pi \times 10^9}{4\pi \times 0.53 \times 10^{-10}} = \frac{9 \times 1.6}{0.53} = 27.2\,\mathrm{V}.$$

For the proton, potential energy in Joules is calculated by multiplying $V(a) = 27.2$ V with $q = e = 1.602 \times 10^{-19}$ C. However, by referring to 1.602×10^{-19} J of energy as 1 eV (electron-volt), it is more convenient to refer to potential energy $eV(a)$ of the proton at $r = a$ as

$$eV(a) = 27.2\,\mathrm{eV}.$$

Likewise, for a particle with charge $q = -e$, i.e., an electron, potential energy at the same location is

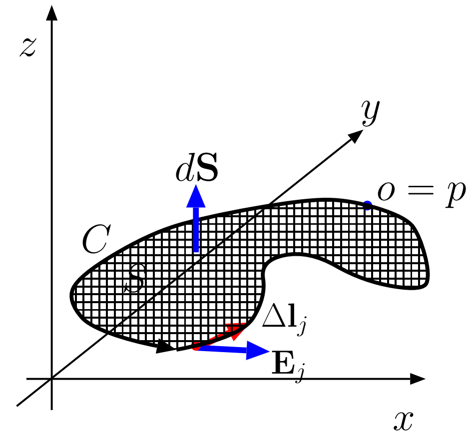
$$-eV(a) = -27.2\,\mathrm{eV}.$$

6 Circulation and boundary conditions

Since curl-free static electric fields have path-independent line integrals, it follows that over closed paths C (when points p and o coincide)

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = 0,$$

where the $\oint_C \mathbf{E} \cdot d\mathbf{l}$ is called the **circulation** of field $\mathbf{E}$ over closed path C bounding a surface S (see margin).



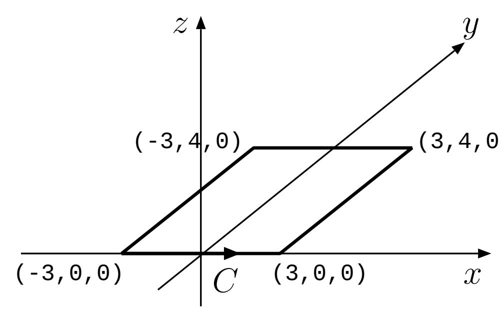
Closed loop integral over path C enclosing surface S.

Note that the area increment dS of surface S is taken by convention to point in the right-hand-rule direction with respect to "circulation" direction C.

Example 1: Consider the static electric field variation

$$\mathbf{E}(x, y, z) = \hat{x}\frac{\rho x}{\epsilon_o}$$

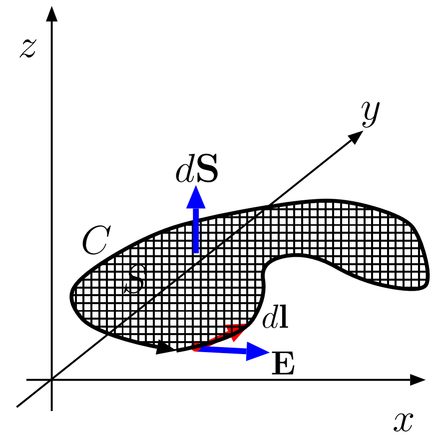
that will be encountered within a uniformly charged slab of an infinite extent in y and z directions and a finite width in x direction centered about $x = 0$. Show that this field $\mathbf{E}$ satisfies the condition $\oint_C \mathbf{E} \cdot d\mathbf{l} = 0$ for a rectangular closed path C with vertices at $(x, y, z) = (-3, 0, 0)$, $(3, 0, 0)$, $(3, 4, 0)$, and $(-3, 4, 0)$ traversed in the order of the vertices given.



Solution: Integration path C is shown in the figure in the margin. With the help of the figure we expand the circulation $\oint_C \mathbf{E} \cdot d\mathbf{l}$ as

$$\begin{aligned}\mathbf{E} &= \int_{x=-3}^{3} \hat{x}\frac{\rho x}{\epsilon_o} \cdot \hat{x}dx + \int_{y=0}^{4} \hat{x}\frac{\rho 3}{\epsilon_o} \cdot \hat{y}dy + \int_{x=3}^{-3} \hat{x}\frac{\rho x}{\epsilon_o} \cdot \hat{x}dx + \int_{y=4}^{0} \hat{x}\frac{\rho(-3)}{\epsilon_o} \cdot \hat{y}dy \\ &= \int_{x=-3}^{3} \frac{\rho x}{\epsilon_o}dx + 0 + \int_{x=3}^{-3} \frac{\rho x}{\epsilon_o}dx + 0 = 0.\end{aligned}$$

Note that in expanding $\oint_C \mathbf{E} \cdot d\mathbf{l}$ above for the given path C, we took $d\mathbf{l}$ as $\hat{x}dx$ and $\hat{y}dy$ in turns (along horizontal and vertical edges of C, respectively) and ordered the integration limits in x and y to traverse C in a counter-clockwise direction as indicated in the diagram.



STOKE'S THM:
Circulation of E around close path C equals the flux over enclosed surface S of the curl of E taken in direction of dS.

dS points in right-hand-rule direction with respect to "circulation" direction C.

- Vector fields $\mathbf{E}$ having zero circulations over all closed paths C are known as **conservative fields** (for obvious reasons having to do with their use in modeling static fields compatible with conservation theorems).

  – The concepts of *curl-free* and *conservative* fields overlap, that is

  $$\oint_C \mathbf{E} \cdot d\mathbf{l} = 0 \quad \Leftrightarrow \quad \nabla \times \mathbf{E} = 0$$

  over all closed paths C and at each $\mathbf{r}$.

- The above relationship between circulation and curl is also a consequence of **Stoke's theorem** (discussed in MATH 241) which asserts that

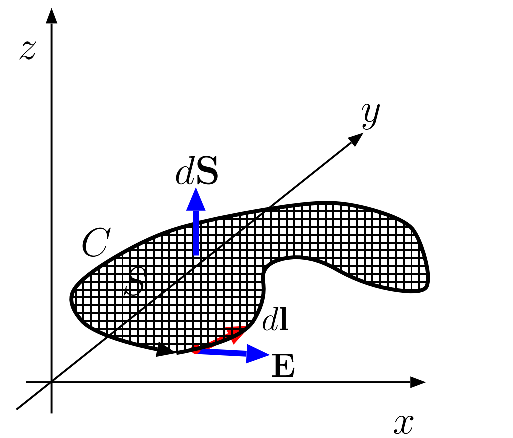
Stoke's thm.

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = \int_S \nabla \times \mathbf{E} \cdot d\mathbf{S},$$

where

  – the integration surface S on the right is bounded by the closed integration contour C of the left side, and

- the incremental area element $d\mathbf{S}$ on the right points across area S in the direction indicated by a **right-hand rule** as follows:



STOKE'S THM:
Circulation of E around close path C equals the flux over enclosed surface S of the curl of E taken in direction of dS.

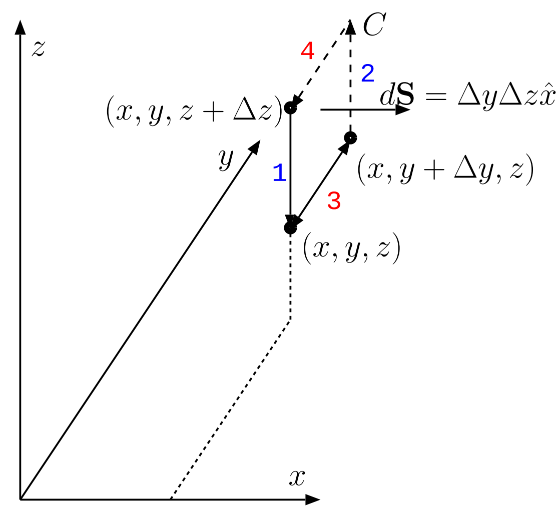
dS points in right-hand-rule direction with respect to "circulation" direction C.

Point your right thumb in chosen circulation direction C; then your right fingers point through surface S in the direction that should be adopted for $d\mathbf{S}$.

  - Given Stoke's theorem, $\oint_C \mathbf{E}\cdot d\mathbf{l} = 0$ follows immediately for all C, if $\nabla\times\mathbf{E} = 0$ is true over all $\mathbf{r}$.

- Stoke's theorem clearly implies that **curl is circulation per unit area**, just as the divergence theorem showed that **divergence is flux per unit volume**.

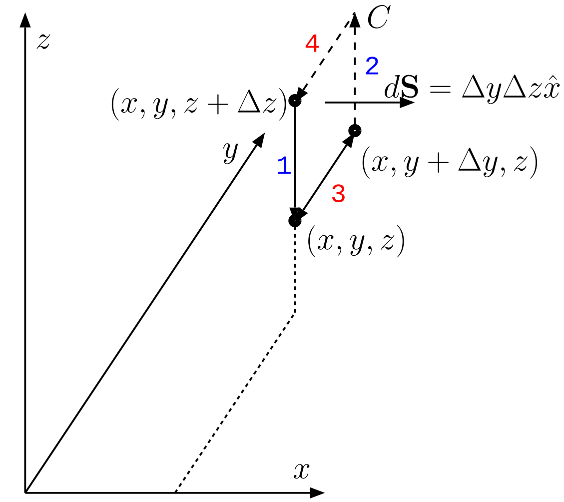
  - The only difference is, curl also has a direction, which is the normal unit of the plane that contains the maximal value of circulation per unit area found at that location (over all possible orientations of $d\mathbf{S}$).



- We will verify Stoke's thm after explaining the **circulation per unit area** notion in steps:

  - Let us first calculate the circulation of a vector field $\mathbf{E}$ taken about an arbitrary point (x, y, z) on a constant x plane around a square contour with small edge dimensions Δy and Δz parallel to y and z axes as shown in the margin.

– For a small rectangular contour "C_x" on a constant x plane with sufficiently small Δy and Δz dimensions parallel to y and z axes (see figure in the margin), we have



$$\begin{aligned}\oint_{C_x} \mathbf{E} \cdot d\mathbf{l} &\approx E_{z|2}\Delta z - E_{y|4}\Delta y - E_{z|1}\Delta z + E_{y|3}\Delta y \\ &= (E_{z|2} - E_{z|1})\Delta z - (E_{y|4} - E_{y|3})\Delta y.\end{aligned}$$

It follows that

$$\frac{1}{\Delta y \Delta z}\oint_{C_x} \mathbf{E} \cdot d\mathbf{l} \approx (\frac{E_{z|2} - E_{z|1}}{\Delta y} - \frac{E_{y|4} - E_{y|3}}{\Delta z})$$

and

$$\lim_{\Delta y, \Delta z \to 0} \frac{1}{\Delta y \Delta z}\oint_{C_x} \mathbf{E} \cdot d\mathbf{l} = (\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}) = \hat{x} \cdot \nabla \times \mathbf{E},$$

meaning that the x component of $\nabla \times \mathbf{E}$ is the circulation of $\mathbf{E}$ per unit area on a constant x surface.

– Likewise, y and z components of $\nabla \times \mathbf{E}$ are circulations of $\mathbf{E}$ per unit area on constant y and z surfaces, and in general

$$\nabla \times \mathbf{E} = \lim_{\Delta x, \Delta y, \Delta z \to 0} (\frac{1}{\Delta y \Delta z}\oint_{C_x} \mathbf{E} \cdot d\mathbf{l}, \frac{1}{\Delta x \Delta z}\oint_{C_y} \mathbf{E} \cdot d\mathbf{l}, \frac{1}{\Delta x \Delta y}\oint_{C_z} \mathbf{E} \cdot d\mathbf{l}).$$

– Furthermore, based on the above result, we can recognize that vectors $\nabla \times \mathbf{E}$ point everywhere in directions perpendicular to planes of maximum circulations per unit area in the $\mathbf{E}$ field and have magnitudes corresponding to the maximum values of circulations per unit area at every point.

- Now, to confirm **Stoke's theorem**

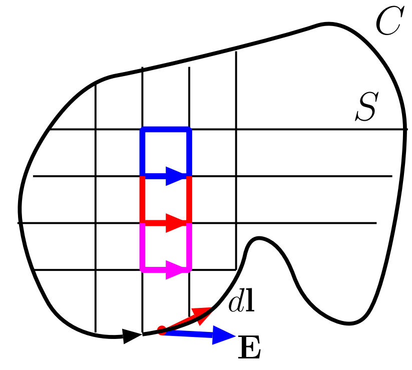
$$\oint_C \mathbf{E} \cdot d\mathbf{l} = \int_S \nabla \times \mathbf{E} \cdot d\mathbf{S}$$

pertinent for a closed path C and its enclosed surface S, we will make use of the diagram shown in the margin.

  - The circulations over small squares shown in the diagram are approximately equal to the products of their areas and the normal components of $\nabla \times \mathbf{E}$ calculated at the center points (basd on what we learned above).
  - When all such circulations covering surface S are added up, the result is $\int_S \nabla \times \mathbf{E} \cdot d\mathbf{S}$ in the limit of vanishing size for the squares,
  - as well as $\oint_C \mathbf{E} \cdot d\mathbf{l}$ because in the grand sum of the circulations over all the squares, all the contributions mutually cancel out (like the overlapping edges of red and blue squares) except for those calculated along the periphery C!

We can now summarize the general constraints governing static electric fields as

$$\nabla \times \mathbf{E}(\mathbf{r}) = 0, \quad \nabla \cdot \mathbf{D}(\mathbf{r}) = \rho(\mathbf{r}), \quad \text{where } \mathbf{D}(\mathbf{r}) = \epsilon_o \mathbf{E}(\mathbf{r}).$$




```
Sum of circulations over small
squares cancel in the interior
edges and only survive around the
exterior path C.  This way,
circulation around C matches
the sum of the fluxes of curl E
calculated over the small squares.
```

Laws of electrostatics:

$$\begin{aligned} \nabla \times \mathbf{E} &= 0 \\ \nabla \cdot \epsilon_o \mathbf{E} &= \rho \end{aligned}$$

They also apply "quasi-statically" over a region of dimension L when a time-varying field source $\rho(\mathbf{r}, t)$ has a *time-constant* τ much longer than the propagation time delay L/c of $\mathbf{E}(\mathbf{r}, t)$ field variations across the region (c is the speed of light).

In electro-quasistatics (EQS) $\mathbf{E}(\mathbf{r}, t)$ will be accompanied by a slowly varying magnetic field $\mathbf{B}(\mathbf{r}, t)$ (to be studied starting in Lecture 12).

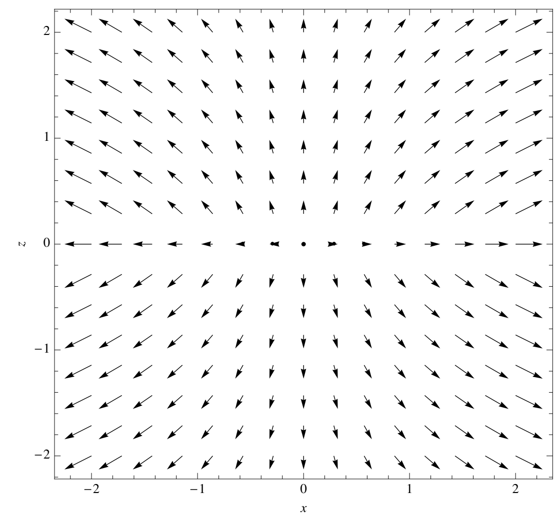
- Vector fields $\mathbf{E}(\mathbf{r})$ and $\mathbf{D}(\mathbf{r})$ governed by these equations will in general be continuous functions of position coordinates $\mathbf{r} = (x, y, z)$ *except* at boundary surfaces where charge density function $\rho(\mathbf{r})$ requires a representation in terms of a surface charge density $\rho_s(\mathbf{r})$.

  – For instance, according to our earlier results, static electric field of a charge density (see sketch at the margin)

  $$\rho(\mathbf{r}) = \rho_s \delta(z)$$

  would be

  $$\mathbf{E}(\mathbf{r}) = \hat{z}\frac{\rho_s}{2\epsilon_o}\text{sgn}(z) \quad \Rightarrow \quad \mathbf{D}(\mathbf{r}) = \hat{z}\frac{\rho_s}{2}\text{sgn}(z).$$

  ◦ Consider a superposition of these fields with fields $\mathbf{E}_o(\mathbf{r})$ and $\mathbf{D}_o(\mathbf{r}) = \epsilon_o \mathbf{E}_o(\mathbf{r})$ produced by arbitrary continuous sources, namely (macroscopic) fields

  $$\mathbf{E}(\mathbf{r}) = \hat{z}\frac{\rho_s}{2\epsilon_o}\text{sgn}(z) + \mathbf{E}_o(\mathbf{r}) \quad \text{and} \quad \mathbf{D}(\mathbf{r}) = \hat{z}\frac{\rho_s}{2}\text{sgn}(z) + \epsilon_o \mathbf{E}_o(\mathbf{r}).$$

Since fields $\mathbf{E}_o(\mathbf{r})$ and $\mathbf{D}_o(\mathbf{r})$ vary continuously, these field expressions must satisfy

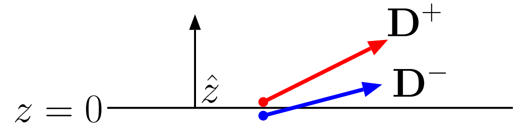
$$\hat{z} \cdot (\mathbf{D}^+ - \mathbf{D}^-) = \rho_s \quad \text{and} \quad \hat{z} \times (\mathbf{E}^+ - \mathbf{E}^-) = 0$$

where

$$\mathbf{E}^+ \equiv \mathbf{E}(x, y, 0^+) \quad \text{and} \quad \mathbf{E}^- \equiv \mathbf{E}(x, y, 0^-)$$

refer to limiting values of $\mathbf{E}$ at $z = 0$ plane from *above* and *below*, respectively, and likewise for

$$\mathbf{D}^+ \equiv \mathbf{D}(x, y, 0^+) \text{ and } \mathbf{D}^- \equiv \mathbf{D}(x, y, 0^-).$$



- The above "boundary condition equations" can be written in a more general form (see margin for justification) as

$$\hat{n} \cdot (\mathbf{D}^+ - \mathbf{D}^-) = \rho_s \quad \text{and} \quad \hat{n} \times (\mathbf{E}^+ - \mathbf{E}^-) = 0$$

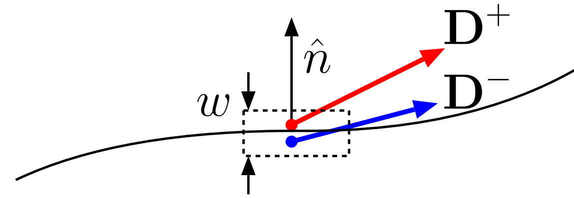
where $\hat{n}$ denotes a unit vector normal to any surface of an arbitrary orientation carrying a surface charge density ρ_s, while field vectors with superscripts $+$ and $-$ indicate limiting values of fields measured on either side of the charged surface (with $\hat{n}$ pointing from $-$ to $+$).

  - The equations can be further simplified as

$$D_n^+ - D_n^- = \rho_s \quad \text{and} \quad E_t^+ = E_t^-$$

where D_n and E_t refer to normal component of $\mathbf{D}$ and tangential component of $\mathbf{E}$, respectively. Clearly, these **boundary conditions** say that at any surface S,

    - tangential component of electric field $\mathbf{E}$ needs to be continuous, but
    - normal component of $\mathbf{D}$ can change by an amount equal to the charge density ρ_s carried by the surface.



Constraint

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = 0$$

around the dotted path yields

$$E_t^+ = E_t^-$$

in $w \to 0$ limit.

Gauss's law

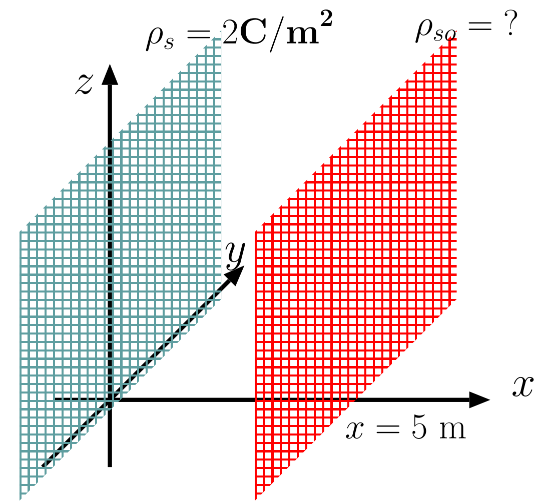
$$\oint_S \mathbf{D} \cdot d\mathbf{S} = Q_V$$

applied over the dotted volume (seen in profile) yields

$$D_n^+ - D_n^- = \rho_s$$

in $w \to 0$ limit.

$\mathbf{D} = 0$ for $x < 0$.



Example 2: Measurements indicate that $\mathbf{D} = 0$ in the region $x < 0$.

Also, $x = 0$ and $x = 5$ m planes contain surface charge densities of $\rho_s = 2$ C/m^2 and ρ_{so}, respectively.

Determine ρ_{so} and $\mathbf{D}$ for $-\infty < x < \infty$ if there are no other charge distributions.

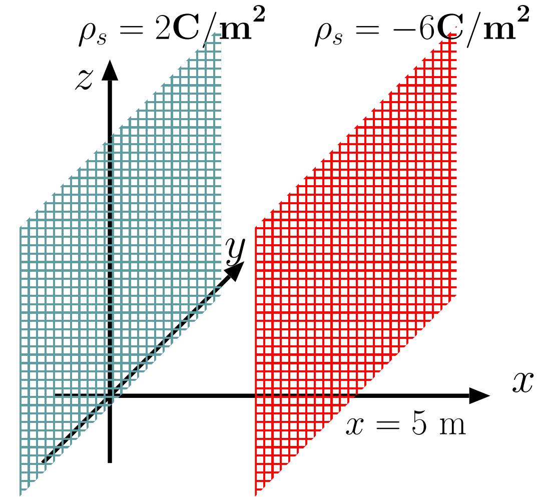
Solution: Since the normal component of $\mathbf{D}$ must increase by $\rho_s = 2$ C/m^2 when we cross the charged surface $x = 0$, we must have $\mathbf{D} = \hat{x}2$ C/m^2 in the region $0 < x < 5$ m.

Having $\mathbf{D} = 0$ in the region $x < 0$ requires that the field due to surface charge ρ_{so} on $x = 5$ m plane must cancel the field due $\rho_s = 2$ C/m^2 on $x = 0$ plane — this requires that ρ_{so} be -2 C/m^2.

In that case $\mathbf{D} = 0$ in the region $x > 5$ m, because $\mathbf{D}$ must increase by $\rho_{so} = -2$ C/m^2 when we cross the charged surface at $x = 5$ m.

Example 3: In the region $x < 0$ measurements indicate a constant displacement field $\mathbf{D} = 3\hat{y}$ C/m^2. Also, $x = 0$ and $x = 5$ m planes contain surface charge densities of $\rho_s = 2$ C/m^2 and $\rho_s = -6$ C/m^2 respectively. Determine $\mathbf{D}$ for $x > 0$ if $\mathbf{D}$ is known to be uniform in the intervals $0 < x < 5$ m and $x > 5$ m.

$\mathbf{D} = 3\hat{y}$ for $x < 0$.



Solution: First we note that $\mathbf{E} = \frac{\mathbf{D}}{\epsilon_o} = \hat{y}\frac{3}{\epsilon_o}$ V/m is tangential to $x = 0$ and $x = 5$ m surfaces. Since the tangential component of $\mathbf{E}$ cannot change at any boundary, we will have a uniform $E_y = \frac{3}{\epsilon_o}$ in all regions, $-\infty < x < \infty$, implying that $D_y = 3$ C/m^2 throughout (caused by charges at $|y| \to \infty$).

Second, we note that normal component of $\mathbf{D}$ with respect to $x = 0$ and $x = 5$ m surfaces, namely D_x, is zero in $z < 0$. Since the normal component of $\mathbf{D}$ must increase by an amount ρ_s when we cross a charged surface, we must have $D_x = 2$ C/m^2 in the region $0 < x < 5$ m, and $D_x = 2 + (-6) = -4$ C/m^2 in $x > 5$ m.

In summary,

$$\mathbf{D} = \begin{cases} \hat{y}3, & \text{for } x < 0, \\ \hat{x}2 + \hat{y}3, & \text{for } 0 < x < 5 \text{ m} \\ -\hat{x}4 + \hat{y}3, & \text{for } x > 5 \text{ m} \end{cases} \frac{\text{C}}{\text{m}^2}.$$

7 Poisson's and Laplace's equations

Summarizing the properties of electrostatic fields we have learned so far, they satisfy the laws of electrostatics shown in the margin and, in addition,

Laws of electrostatics:

$$\nabla \cdot \mathbf{E} = \rho/\epsilon_o$$
$$\nabla \times \mathbf{E} = 0$$

$$\mathbf{E} = -\nabla V \text{ as a consequence of } \nabla \times \mathbf{E} = 0.$$

- Using these relations, we can re-write Gauss's law as

  $$\nabla \cdot \mathbf{E} = -\nabla \cdot (\nabla V) = \frac{\rho}{\epsilon_o},$$

  from which it follows that

  $$\nabla^2 V = -\frac{\rho}{\epsilon_o}, \quad \text{(Poisson's eqn)}$$

  where

  $$\nabla^2 V \equiv \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2}$$

  is known as **Laplacian** of V.

  Poisson's eqn:

  $$\nabla^2 V = -\frac{\rho}{\epsilon_o}$$

  - A special case of Poisson's equation corresponding to having

    $$\rho(x, y, z) = 0$$

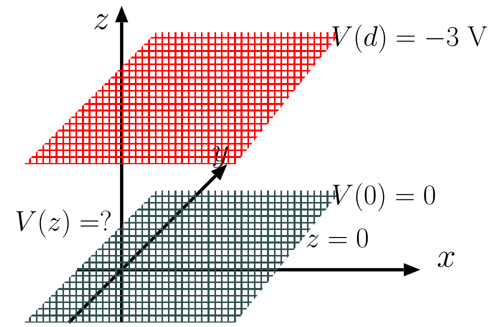
    everywhere in the region of interest is

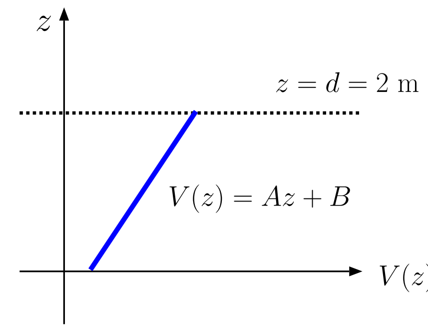
    $$\nabla^2 V = 0. \quad \text{(Laplace's eqn)}$$

Laplace's eqn:

$$\nabla^2 V = 0$$

Focusing our attention first on Laplace's equation, we note that the equation can be used in charge free-regions to determine the electrostatic potential $V(x, y, z)$ by matching it to specified potentials at boundaries as illustrated in the following examples:





Example 1: Consider a pair of parallel conducting metallic plates of infinite extents in x and y directions but separated in z direction by a finite distance of $d = 2$ m (as shown in the margin). The conducting plates have non-zero surface charge densities (to be determined in Example 2), which are known to be responsible for an electrostatic field $\mathbf{E} = \hat{z}E_z$ measured in between the plates. Each plate has some unique and constant electrostatic potential V since neither $\mathbf{E}(\mathbf{r})$ nor $V(\mathbf{r})$ can dependent the coordinates x or y given the geometry of the problem.

Using Laplace's equation determine $V(z)$ and $\mathbf{E}(z)$ between the plates if the potential of the plate at $z = 0$ is 0 (the ground), while the potential of the plate at $z = d$ is -3 V.

Solution: Since the potential function $V = V(z)$ between the plates is only dependent on z, it follows that Laplace's equation simplifies as

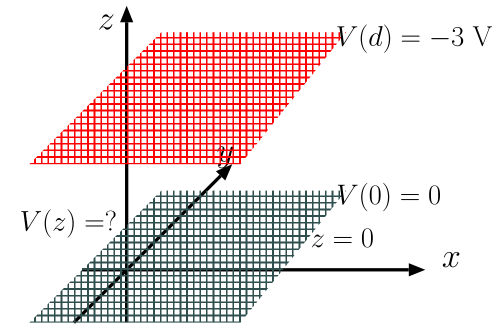
$$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = \frac{\partial^2 V}{\partial z^2} = 0.$$

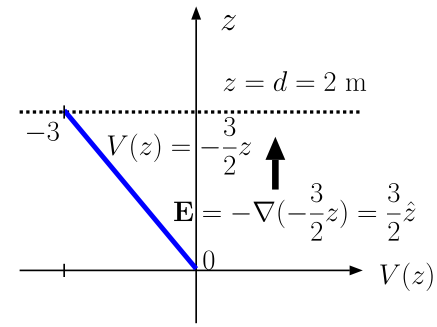
This equation can be satisfied by

$$V(z) = Az + B$$

where A and B are constants to be determined. Now applying the given boundary conditions, we first notice that (at the lower plate)

$$V(0) = (Az + B)_{|z=0} = B = 0.$$





Applying the second boundary condition (at the top plate) we find

$$V(2) = (Az + 0)_{|z=2} = 2A = -3\,\mathrm{V} \quad \Rightarrow \quad A = -\frac{3}{2}\frac{\mathrm{V}}{\mathrm{m}}.$$

The upshot is, potential function

$$V(z) = -\frac{3}{2}z, \;\; \text{for } 0 < z < 2\,\mathrm{m}.$$

Finally, we determine the electric field between the plates as

$$\mathbf{E} = -\nabla V = -\nabla(-\frac{3}{2}z) = \hat{z}\frac{\partial}{\partial z}(\frac{3}{2}z) = \hat{z}\frac{3}{2}\frac{\mathrm{V}}{\mathrm{m}}.$$

Example 2: In Example 1 what are the surface charge densities of the metallic plates located at $z = 0$ and $z = 2$ m surfaces?

Solution: Since the electric field

$$\mathbf{E} = \hat{z}\frac{3}{2}\frac{\mathrm{V}}{\mathrm{m}}$$

in between the plates, comparing this field with the field

$$\mathbf{E} = \hat{z}\frac{\rho_s}{\epsilon_o}$$

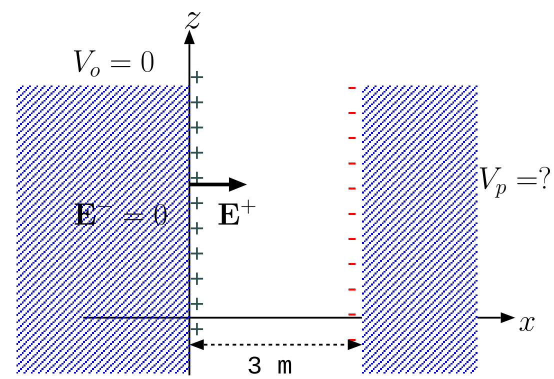
of a pair of parallel surfaces carrying surface charge densities ρ_s and $-\rho_s$ (at $z = 0$ and $z = 2$ m), we find that

$$\rho_s = \frac{3}{2}\epsilon_o$$

on the surface at $z = 0$. The surface at $z = 2$ m has $\rho_s = -\frac{3}{2}\epsilon_o$.

Notice that our solution with equal and opposite charge densities on the parallel surfaces implies that electrostatic fields are zero *within* the conducting plates where the fields due to two charged surfaces are canceling out. This conclusion is consistent with having constant electrostatic potentials within conducting regions as will be discussed in the next lecture.

Example 3: A pair of copper blocks separated by a distance $d = 3$ m in x direction hold surface charge densities of $\rho_s = \pm 2$ C/m^2 on surfaces facing one another as shown in the margin. The blocks are assigned constant potentials $V_o = 0$ and V_p (see figure). What is the potential difference V_p?



Solution: Let $\mathbf{D}^+ = \hat{x}\epsilon_o E_x$ denote the displacement vector in between the blocks, and let $\mathbf{D}^- = 0$ denote the displacement vector *within* the block with a surface at $x = 0$. Then the boundary condition equation used at $x = 0$ implies that

$$\hat{x} \cdot (\mathbf{D}^+ - \mathbf{D}^-) = \epsilon_o E_x = 2\,\frac{\text{C}}{\text{m}^2} \quad \Rightarrow \quad E_x = \frac{2}{\epsilon_o}.$$

In that case, potential difference between the blocks is

$$V = E_x d = \frac{2}{\epsilon_o} 3 = \frac{6}{\epsilon_o}.$$

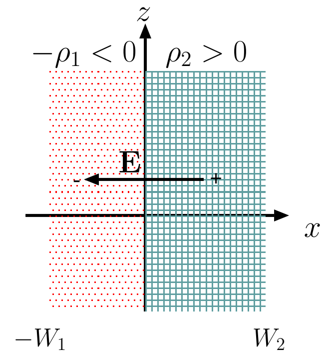
Since the block on the left is at a higher potential (electric field vectors point from high to low potential) assigned as $V_o = 0$, we must have

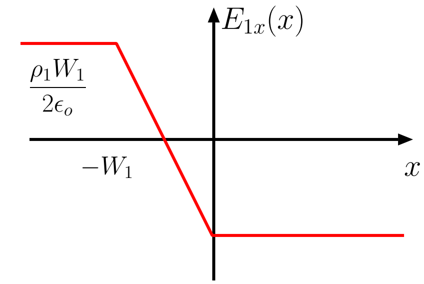
$$V_p = -\frac{6}{\epsilon_o}.$$

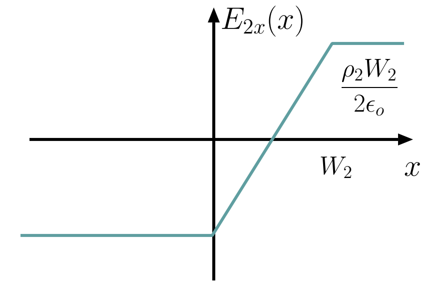
Poisson's equation

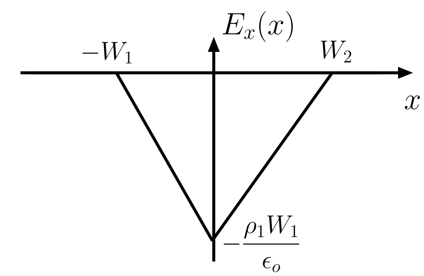
$$\nabla^2 V = -\frac{\rho}{\epsilon_o}$$

is used in regions where the charge density $\rho(\mathbf{r})$ is non-zero. The following example illustrates a possible use of Poisson's equation.









Example 4: An infinite charged slab of width W_1, located over $-W_1 < x < 0$, has a negative volumetric charge density of $-\rho_1$ C/m^3, $\rho_1 > 0$. A second slab of width W_2 and positive charge density ρ_2 is located over $0 < x < W_2$ as shown in the margin. The electric field of this static charge configuration under the constraint $W_1\rho_1 = W_2\rho_2$ was computed in an earlier section as

$$\mathbf{E} = \begin{cases} -\hat{x}\frac{\rho_1(x+W_1)}{\epsilon_o}, & \text{for } -W_1 < x < 0 \\ \hat{x}\frac{\rho_2(x-W_2)}{\epsilon_o}, & \text{for } 0 < x < W_2 \end{cases}$$

and is depicted in the margin. Determine the electrostatic potential in the region and the potential difference $V_{21} \equiv V(W_2) - V(-W_1)$ satisfying Poisson's equation.

Solution: This is a one dimensional geometry where $\mathbf{E}$ and potential V depend only on coordinate x. Therefore, Poisson's equation $\nabla^2 V = -\rho/\epsilon_o$ takes the simplified form

$$\frac{d^2V}{dx^2} = -\frac{\rho(x)}{\epsilon_o}.$$

Integral of this equation over x yields in the left $\frac{dV}{dx} = -E_x$, which implies, given the electric field result from above,

$$\frac{dV}{dx} = \begin{cases} \frac{\rho_1(x+W_1)}{\epsilon_o}, & \text{for } -W_1 < x < 0 \\ -\frac{\rho_2(x-W_2)}{\epsilon_o}, & \text{for } 0 < x < W_2 \end{cases}$$

Integrating $\frac{dV}{dx}$ once more (i.e., finding suitable anti-derivatives with integration constants), we find

$$V(x) = \begin{cases} \frac{\rho_1(x+W_1)^2}{2\epsilon_o} + V_1, & \text{for } -W_1 < x < 0 \\ -\frac{\rho_2(x-W_2)^2}{2\epsilon_o} + V_2, & \text{for } 0 < x < W_2 \end{cases}$$

where the integration constants included *on each line* have been selected so that $V_2 = V(W_2)$, $V_1 = V(-W_1)$.

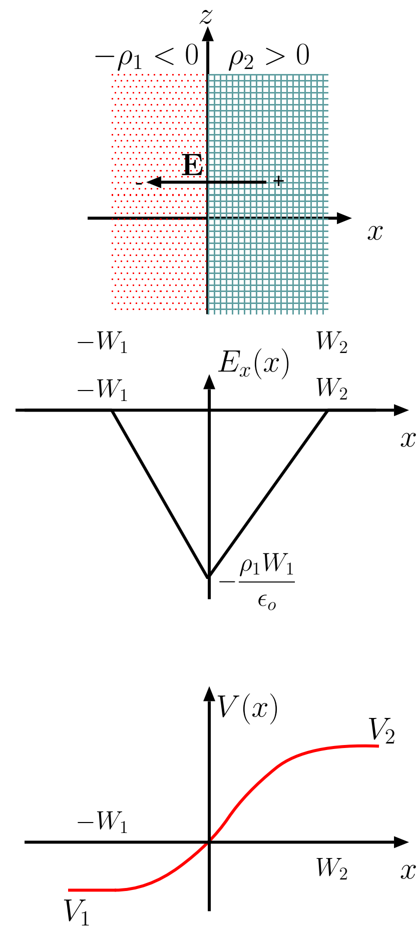
Requiring a unique potential value at $x = 0$ (we can only associate a single potential energy level with each position in space) compatible with this expression for $V(x)$, we obtain

$$\frac{\rho_1(0+W_1)^2}{2\epsilon_o} + V_1 = -\frac{\rho_2(0-W_2)^2}{2\epsilon_o} + V_2,$$

from which

$$V_{21} = V_2 - V_1 = \frac{\rho_2 W_2^2 + \rho_1 W_1^2}{2\epsilon_o} = \frac{\rho_2 W_2(W_1+W_2)}{2\epsilon_o} = \frac{\rho_1 W_1(W_1+W_2)}{2\epsilon_o}.$$

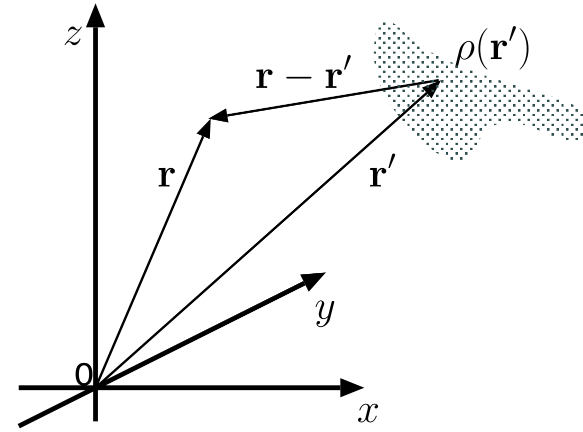
Note that the equation above can be solved for W_1, W_2, and $W_2 + W_1$ in terms of V_{12}, ρ_2, and ρ_1, providing useful formulas for diode design (see ECE 440). We can also get useful specific formulae for V_1 and V_2 by imposing $V(0) = 0$, i.e., choosing $x = 0$ to be the reference point.



- The solution of Poisson's equation

$$\nabla^2 V = -\frac{\rho}{\epsilon_o}$$

with an arbitrary ρ existing over a finite region in space can be obtained as



$$V(\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{4\pi\epsilon_o|\mathbf{r}-\mathbf{r}'|} d^3\mathbf{r}'$$

where $d^3\mathbf{r}' \equiv dx'dy'dz'$ and the 3D integral on the right over the primed coordinates is performed over the entire region where the charge density is non-zero.

– **Verification:** The solution above can be verified by combining a number of results we have seen earlier on:

1. In Lecture 5 we learned that the electric potential $V(\mathbf{r})$ of a point charge e at the origin is

$$V(\mathbf{r}) = \frac{e}{4\pi\epsilon_o|\mathbf{r}|}.$$

Clearly, this singular result is a solution of Poisson's equation above (and the stated boundary condition) for a charge density input of

$$\rho(\mathbf{r}) = e\delta(\mathbf{r}).$$

2. Using ECE 210-like terminology and notation, the above result can be represented as

$$\delta(\mathbf{r}) \rightarrow \boxed{\text{Poisson's Eqn}} \rightarrow \frac{1}{4\pi\epsilon_o|\mathbf{r}|}$$

identifying the output on the right as 3D "impulse response" of the **linear** and **shift-invariant** (LSI) system represented by Poisson's equation.

3. Because of shift-invariance, we have

$$\delta(\mathbf{r}-\mathbf{r}') \rightarrow \boxed{\text{Poisson's Eqn}} \rightarrow \frac{1}{4\pi\epsilon_o|\mathbf{r}-\mathbf{r}'|},$$

   meaning that a shifted impulse causes a shifted impulse response.

   The shifted impulse response is usually called "Green's function" $G(\mathbf{r},\mathbf{r}')$ in EM theory.

4. Because of linearity, we are allowed to use superpositioning arguments like

$$\int \rho(\mathbf{r}')\delta(\mathbf{r}-\mathbf{r}')d^3\mathbf{r}' = \rho(\mathbf{r}) \rightarrow \boxed{\text{Poisson's Eqn}} \rightarrow \int \rho(\mathbf{r}')\frac{1}{4\pi\epsilon_o|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}' = V(\mathbf{r}),$$

   which concludes our verification of the electrostatic[1] potential solution. Note how we made use of the *sifting property* of the impulse (from ECE 210) in above calculation.

[1]Also, in quasi-statics we use $\rho(\mathbf{r}',t)$ to obtain $V(\mathbf{r},t)$ over regions small compared to $\lambda = c/f$, with f the highest frequency in $\rho(\mathbf{r}',t)$.

- As an application of the general solution of Poisson's equation, namely

$$\nabla^2 V = -\frac{\rho}{\epsilon_o} \;\;\Rightarrow\;\; V(\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{4\pi\epsilon_o|\mathbf{r}-\mathbf{r}'|} d^3\mathbf{r}',$$

  we next provide an *outline* of the proof of **Helmholtz theorem** (see Lecture 4) which states that *any vector field* $\mathbf{F}(x, y, z)$ *that vanishes in the limit* $r = \sqrt{x^2 + y^2 + z^2} \to \infty$ *can be reconstructed uniquely from its* divergence *and* curl*:*

  - First, with no loss of generality, we write

$$\mathbf{F} = -\nabla V + \nabla \times \mathbf{A}$$

    in terms of scalar and vector fields $V(x, y, z)$ and $\mathbf{A}(x, y, z)$ *to be identified* as follows[2]:

  - Taking first the divergence of $\mathbf{F}$ (and using $\nabla \cdot \nabla \times \mathbf{A} = 0$), we find that

$$\nabla \cdot \mathbf{F} = -\nabla^2 V \;\;\Rightarrow\;\; V(\mathbf{r}) = \int \frac{\nabla' \cdot \mathbf{F}(\mathbf{r}')}{4\pi|\mathbf{r}-\mathbf{r}'|} d^3\mathbf{r}'$$

    in analogy with Poisson's equation (with $\nabla'\cdot\mathbf{F}(\mathbf{r}')$ replacing $\rho(\mathbf{r}')/\epsilon_o$ where ∇' is "del" in (x', y', z')-space).

---

[2]This is possible because of the vector identity $-\nabla^2\mathbf{G} = \nabla\times(\nabla\times\mathbf{G}) - \nabla(\nabla\cdot\mathbf{G})$ — call $-\nabla^2\mathbf{G} \equiv \mathbf{F}$, which, according to this identity, is equal to the curl of a *vector* $\nabla\times\mathbf{G} \equiv \mathbf{A}$ (with $\nabla\cdot\mathbf{A} = \nabla\cdot\nabla\times\mathbf{G} = 0$), minus the gradient of a *scalar* $\nabla\cdot\mathbf{G} \equiv V$, as claimed. The challenge is in figuring out the underlying $\mathbf{G}$ for a given $\mathbf{F}$, which is what *Helmholtz theorem* is all about.

– Likewise, the curl of $\mathbf{F}$ (with $\nabla\times\nabla V = 0$) leads us to, with a divergence-free [3] $\mathbf{A}$, to

$$\nabla\times\mathbf{F} = \nabla\times\nabla\times\mathbf{A} = \nabla(\nabla\cdot\mathbf{A}) - \nabla^2\mathbf{A} = -\nabla^2\mathbf{A} \quad\Rightarrow\quad \mathbf{A}(\mathbf{r}) = \int \frac{\nabla'\times\mathbf{F}(\mathbf{r}')}{4\pi|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}'$$

once again in analogy with Poisson's equation[4].

These results validate Helmholtz theorem for fields $\mathbf{F}$ vanishing at infinity, since, *evidently,* V and $\mathbf{A}$ needed to reconstruct $\mathbf{F}$ can be uniquely specified in terms of $\nabla\cdot\mathbf{F}$ and $\nabla\times\mathbf{F}$, respectively.

---

[3]To confirm $\nabla\cdot\mathbf{A} = 0$ directly, use the identity $\nabla\cdot(\alpha\mathbf{G}) = \alpha\nabla\cdot\mathbf{G} + \mathbf{G}\cdot\nabla\alpha$ to expand $\mathbf{A}(\mathbf{r})$ as

$$\nabla\cdot\int\frac{\nabla'\times\mathbf{F}(\mathbf{r}')}{4\pi|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}' = \int\frac{\nabla'\times\mathbf{F}(\mathbf{r}')}{4\pi}\cdot\nabla\frac{1}{|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}' = -\int\frac{\nabla'\times\mathbf{F}(\mathbf{r}')}{4\pi}\cdot\nabla'\frac{1}{|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}' = \int\frac{\nabla'\cdot(\nabla'\times\mathbf{F}(\mathbf{r}'))}{4\pi|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}' = 0,$$

after also using integration by parts for an integrand that vanishes as $|\mathbf{r}|\to\infty$ and a symmetry relation $\nabla|\mathbf{r}-\mathbf{r}'|^{-1} = -\nabla'|\mathbf{r}-\mathbf{r}'|^{-1}$ which is easy to confirm.

[4]While the vector field $\mathbf{A}$ identified above is divergence-free, $\nabla\times\mathbf{A}$ in the $\mathbf{F} = -\nabla V + \nabla\times\mathbf{A}$ expansion can also be replaced with $\nabla\times\mathbf{A}'$ so long as $\mathbf{A}' = \mathbf{A} + \nabla\Psi$ since $\nabla\times\nabla\Psi$ is unconditionally zero independent of the choice of Ψ. Note it is possible to specify Ψ so that $\nabla\cdot\mathbf{A}' = \nabla\cdot\nabla\Psi = \nabla^2\Psi \neq 0$ in which case $\mathbf{A}'$ will be a divergent solution of the $\nabla\times\mathbf{F}$ equation above! The additive term $\nabla\Psi$ in $\mathbf{A}'$ is analogous to allowing a constant number to be added to V! The freedom to specify Ψ and thus $\nabla\cdot\mathbf{A}'$ at will is known as *gauge freedom* and any choice of Ψ making $\nabla\cdot\mathbf{A}' = 0$ is known as *Coulomb's gauge.*

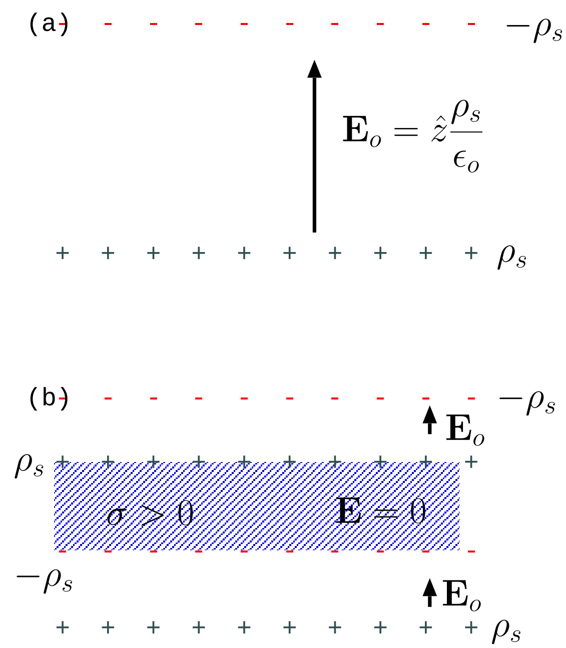
8 Conductors, dielectrics, and polarization

So far in this course we have examined static field configurations of charge distributions assumed to be fixed in free space in the *absence* of nearby materials (solid, liquid, or gas) composed of neutral atoms and molecules.

In the *presence* of material bodies composed of large number of charge-neutral atoms (in fluid or solid states) static charge distributions giving rise to electrostatic fields can be typically[1] found:

1. On exterior surfaces of *conductors* in "steady-state",
2. In crystal lattices occupied by *ionized* atoms, as in depletion regions of semiconductor junctions in diodes and transistors.

In this lecture we will examine these configurations and response of materials to applied electric fields.



A conducting slab inserted into a region with field E_o (as shown in b)develops surface charge which cancels out E_o within the slab.

E_o relates to surface charge as dictated by Gauss's law and superposition principle.

Conductivity and static charges on conductor surfaces:

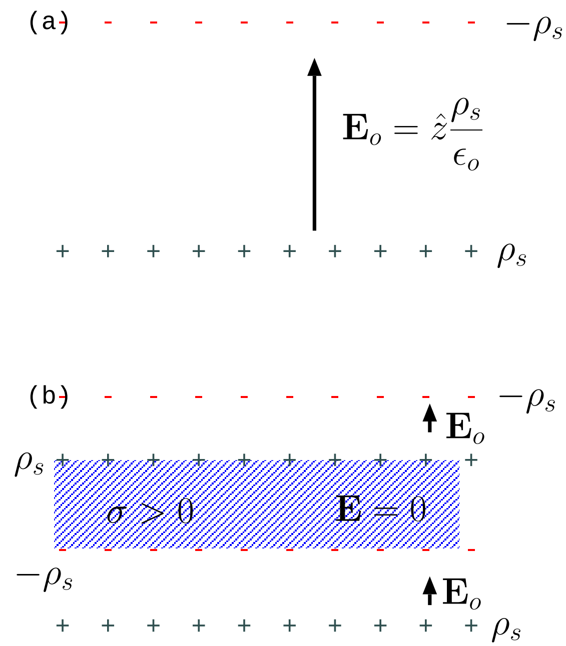
- **Conductivity** σ is an *emergent property* of materials bodies containing free charge carriers (e.g., unbound electrons, ionized atoms or molecules) which relates the applied *electric field* $\mathbf{E}$ (V/m) to the *electrical current density* $\mathbf{J}$ (A/m^2) conducted in the material via a linear

[1] More generally, materials containing charge carriers exhibiting divergence free flows will also exhibit static charge distributions.

relation[2]

$$\mathbf{J} = \sigma \mathbf{E}. \quad \text{(Ohm's Law)}$$

- Simple physics-based models for σ will be discussed later in Lecture 11. For now it is sufficient to note that:
  - $\sigma \to \infty$ corresponds to a *perfect electrical conductor*[3] (PEC) for which it is necessary that $\mathbf{E} = 0$ (in analogy with $V = 0$ across a short circuit element) independent of $\mathbf{J}$.
  - $\sigma \to 0$ corresponds to a *perfect insulator* for which it is necessary that $\mathbf{J} = 0$ (in analogy with $I = 0$ through an open circuit element) independent of $\mathbf{E}$.



A conducting slab inserted into a region with field E_o (as shown in b)develops surface charge which cancels out E_o within the slab.

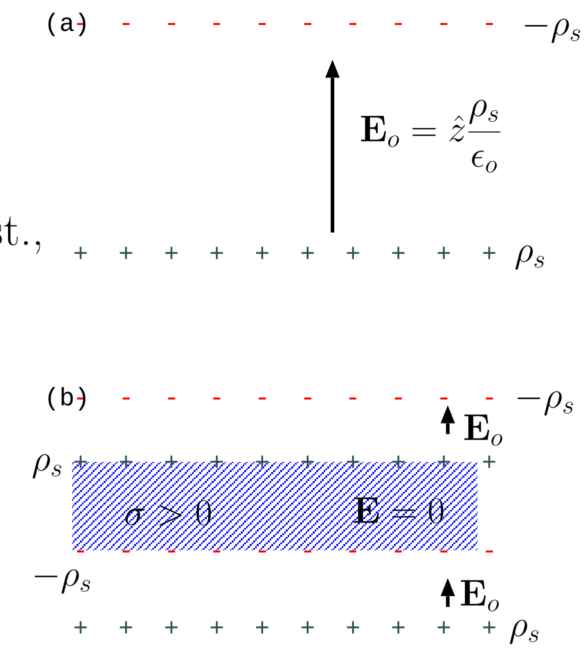
E_o relates to surface charge as dictated by Gauss's law and superposition principle.

- While (macroscopic) $\mathbf{E} = 0$ in PEC's *unconditionally*, a conductor with a *finite* σ (e.g., copper or sea water) will also have $\mathbf{E} = 0$ in "steady-state" after the decay of transient currents $\mathbf{J}$ that may be initiated within the conductor after applying an external electric field $\mathbf{E}_o$ (see margin).
  - The reason is, mobile free charges (e.g., electrons in metallic conductors) within the conductor will be pulled or pushed by the applied field $\mathbf{E}_o$ to pile up on exterior surfaces of the conductor

[2]Linear behavior is possible provided charge carriers suffer occasional collisions within the medium.

[3]PEC is an "idealization" that has no real counterpart, even though it is convenient to treat high conductivity materials such as copper as PEC in certain approximate models and calculations. For "superconducting materials" $\sigma \to \infty$ only in the low frequency limit.

until a surface charge density ρ_s that is generated produces a secondary field $-\mathbf{E}_o$ that exactly cancels out the applied $\mathbf{E}_o$ within the interior of the conductor.



A conducting slab inserted into a region with field E_o (as shown in b)develops surface charge which cancels out E_o within the slab.

E_o relates to surface charge as dictated by Gauss's law and superposition principle.

- $\mathbf{E} = 0$ in the interior at steady-state implies that potential V =const., as well as $\rho = \nabla \cdot \mathbf{D} = \nabla \cdot \epsilon_o \mathbf{E} = 0$.
- Surface charge density ρ_s and the exterior field on a conductor surface will satisfy the boundary condition equations

$$\hat{n} \cdot \mathbf{D} = \rho_s \quad \text{and} \quad \hat{n} \times \mathbf{E} = 0,$$

with $\hat{n}$ denoting the outward unit normal.

- The transient "time-constant" τ for the *decay* of charge density ρ (and hence $\mathbf{E}$, as claimed above) in a homogeneous[4] conductor (constant σ) can be obtained using the *continuity equation*

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0$$

representing the mathematical statement of **charge conservation** (derived in Lecture 16). Using $\mathbf{J} = \sigma \mathbf{E}$ and $\nabla \cdot \mathbf{E} = \rho/\epsilon_o$, we have

$$\nabla \cdot \mathbf{J} = \sigma \nabla \cdot \mathbf{E} = \frac{\sigma}{\epsilon_o}\rho$$

[4]See Fisher and Varney, *Am. J. Phys.*, **44**, 464 (1976), for a discussion of contact potential between different metals.

above, from which it follows that

$$\frac{\partial \rho}{\partial t} + \frac{\sigma}{\epsilon_o}\rho = 0 \text{ with a damped solution } \rho(t) = \rho(0)e^{-\frac{\sigma}{\epsilon_o}t}.$$

The decay time-constant

$$\tau = \frac{\epsilon_o}{\sigma}$$

is typically very short ($\sim 10^{-18}$ s) in metallic conductors, which is why such conductors are usually considered to be in steady-state (and have zero interior fields).

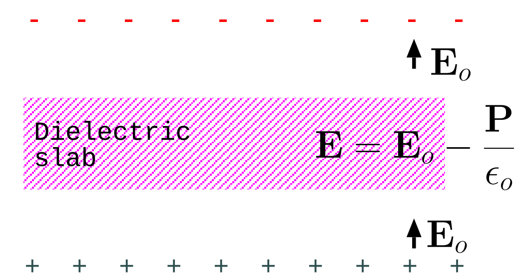
- **As a consequence:** in electrostatic[5] problems conducting volumes of materials (e.g., chunks of copper) can be treated as *equipotentials* having zero internal fields and finite surface charge densities $\rho_s = \hat{n} \cdot \mathbf{D}$ expressed in terms of external fields $\mathbf{D}$ normal to the surface.

[5]Also applicable *quasi-statically* when externally applied fields $\mathbf{E}_o(t)$ change slowly with time-constants much longer than ϵ_o/σ. The way conductors are treated in high frequency *electromagnetic* problems will be described later on.

Dielectric materials and polarization:

- **Dielectric materials** consist of a large number of charge-neutral atoms or molecules and *ideally* contain no mobile charge carriers (i.e., $\sigma = 0$).

- Electric fields produced by charges located outside or within a dielectric material will **polarize** the dielectric — meaning that its constituent atoms or molecules will be "stretched out" to expose their internal or "bound" charges, electrons and protons — which will in turn cause the electric field inside the dielectric to become *weaker* than (but not zero, as in conductors) what the field would have been in the absence of polarization effect.



A dielectric slab inserted into a region with an initial field E_o will become polarized.

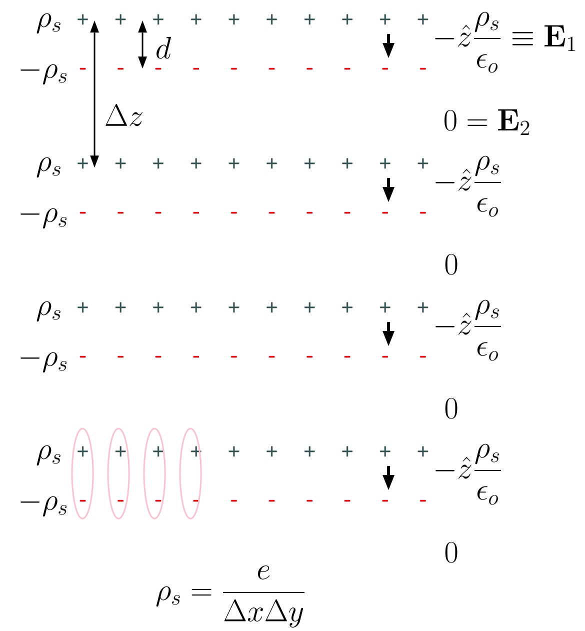
Inside the polarized dielectric the field will be weaker than E_o, but not reduced to zero as in a conductor.

We will next examine this polarization process and see how Gauss's law can be re-stated to facilitate field calculations in dielectric materials containing **bound charge** carriers, i.e., atomic/molecular electrons and protons which are not free to drift away from one another indefinitely (neglecting possible ionization events).

- Consider a static **free-charge** density $\rho(z)$ that would produce a macroscopic field $\mathbf{E}_o$ satisfying $\rho = \epsilon_o \nabla \cdot \mathbf{E}_o$ in free space, producing, instead, a field $\mathbf{E} = \hat{z} E_z$ inside a dielectric medium composed of an array of neutral atoms or molecules.

Our objective is to relate the field E to $\mathbf{E}_o$ and ρ, and find a way of calculating E when ρ is given.

- In the presence of an electric field $\mathbf{E} = \hat{z}E_z$ in the dielectric each neutral atom of the medium will be in a distorted (but not ripped apart) state forming a $\hat{z}$ oriented **electric dipole**, which can be visualized as a proton-electron pair with a small proton displacement d in z direction with respect to the electron.



  - Consider a regular array of such dipoles

$$\mathbf{p} \equiv ed\hat{z},$$

with Δx, Δy, and Δz spacings between the dipoles (see margin), so that the volumetric dipole density is

$$N_d \equiv \frac{1}{\Delta x \Delta y \Delta z} \ \ \mathrm{m}^{-3},$$

within the array, and, furthermore,

$$\rho_s = \frac{e}{\Delta x \Delta y} \ \ \frac{\mathrm{C}}{\mathrm{m}^2}$$

is the magnitude of charge density of the adjacent proton and electron layers (see margin again) formed by arrays of adjacent dipoles displaced in z by intervals Δz.

  - Assuming that the array is infinite in extent in x and y directions, the proton and electron layers with surface charge densities $\pm\rho_s$ will produce interior electric fields

$$\mathbf{E}_1 = -\hat{z}\frac{\rho_s}{\epsilon_o} = -\hat{z}\frac{e/\epsilon_o}{\Delta x \Delta y}$$

(pointing in opposite direction to $\mathbf{E} = \hat{z}E_z$), and exterior fields

$$\mathbf{E}_2 = 0$$

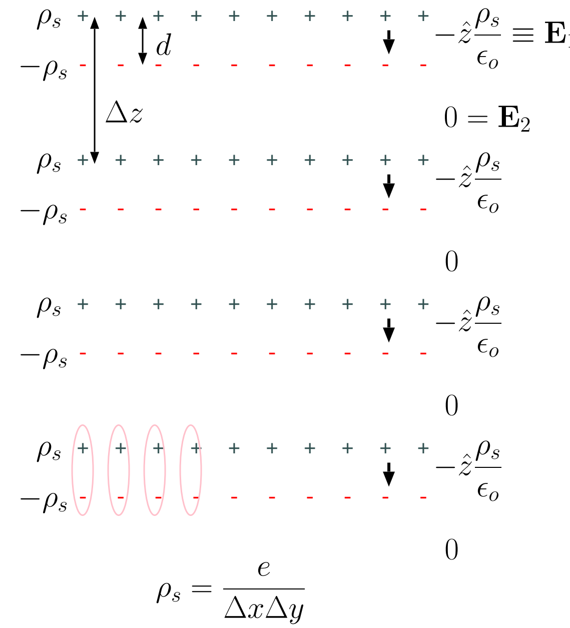
in between the dipole layers. Space averaged macroscopic electric field within the array (with a spatial weighting proportional to the size of regions with the fields $\mathbf{E}_1$ and $\mathbf{E}_2$) produced by the polarized dipoles will then be

$$\mathbf{E}_p = \mathbf{E}_1\frac{d}{\Delta z} + \mathbf{E}_2\frac{\Delta z - d}{\Delta z} = -\hat{z}\frac{ed/\epsilon_o}{\Delta x\Delta y\Delta z} = -\frac{N_d ed\hat{z}}{\epsilon_o} = -\frac{\mathbf{P}}{\epsilon_o},$$

where

$$\mathbf{P} \equiv N_d ed\hat{z} = N_d\mathbf{p}$$

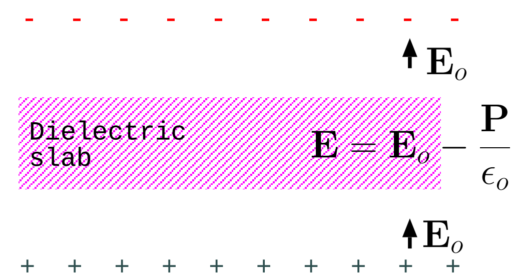
is, by definition, macroscopic **polarization field** of the dielectric, measured in units of C/m^2 (same units as a surface charge density).



– The total macroscopic field $\mathbf{E}$ in the dielectric is then the sum of field $\mathbf{E}_o$ produced by the free charge density ρ in the region and the polarization field $\mathbf{E}_p = -\frac{\mathbf{P}}{\epsilon_o}$ produced by bound charge carriers of the neutral atoms and/or molecules of the dielectric, i.e.,

$$\mathbf{E} = \mathbf{E}_o - \frac{\mathbf{P}}{\epsilon_o},$$

a result that shows a "reduced field strength" $\mathbf{E}$ (compared to $\mathbf{E}_o$) inside the dielectric since $\mathbf{P}$ and $\mathbf{E}_o$ are colinear.



A dielectric slab inserted into a region with an initial field E_o will become polarized.

Inside the polarized dielectric the field will be weaker than E_o, but not reduced to zero as in a conductor.

- Let's re-arrange the expression for $\mathbf{E}$ from above as

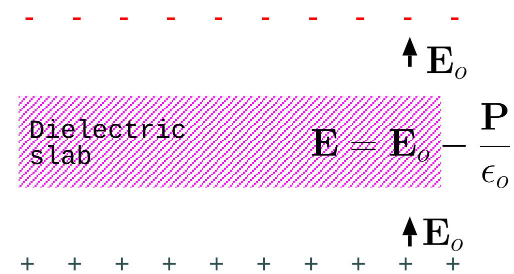
$$\epsilon_o \mathbf{E} + \mathbf{P} = \epsilon_o \mathbf{E}_o$$

after multiplying it with ϵ_o and moving $\mathbf{P}$ to the left. Now, the term on the right is $\epsilon_o \mathbf{E}_o = \mathbf{D}_o$ respresenting the displacement vector outside the dielectric slab, and if we were to "adopt" the left hand side expression as the "displacement vector" for the interior, i.e, take

$$\mathbf{D} = \epsilon_o \mathbf{E} + \mathbf{P}$$

in regions with non-zero $\mathbf{P}$, then we would see that

  - $\mathbf{D} = \mathbf{D}_o$, i.e., the displacement is the same inside and outside the slab, while electric fields $\mathbf{E}$ and $\mathbf{E}_o$ inside and outside differ by a non-zero $-\mathbf{P}/\epsilon_o$, and furthermore,
  - this generalized definition of electric displacement is consistent with (by now familiar) $\mathbf{D} = \epsilon_o \mathbf{E}$ for free space as $\mathbf{P} \neq 0$ only within dielectrics.



A dielectric slab inserted into a region with an initial field E_o will become polarized.

Inside the polarized dielectric the field will be weaker than E_o, but not reduced to zero as in a conductor.

- To express Gauss's law $\nabla \cdot (\epsilon_o \mathbf{E}) = \rho$ in a form applicable with our new revised $\mathbf{D} = \epsilon_o \mathbf{E} + \mathbf{P}$ we first note that Gauss's law also holds in material media so long as $\rho = \rho_f + \rho_b$ is a total charge density function, a sum of charge densities ρ_f and ρ_b associated with free and bound charge carriers that could present in the dielectric region. As such:

– outside dielectrics, Gauss's law is $\nabla \cdot (\epsilon_o \mathbf{E}) = \rho_f$, since $\rho_b = 0$ in that case, and this can be expressed as $\nabla \cdot \mathbf{D} = \rho_f$, with $\mathbf{D} = \epsilon_o \mathbf{E}$ as usual;

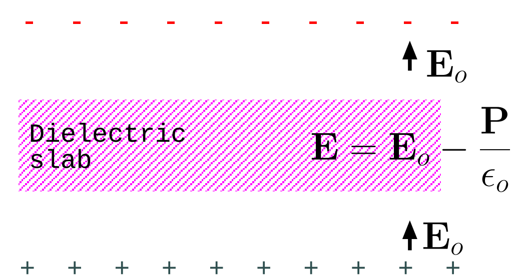
– using $\nabla \cdot \mathbf{D} = \rho_f$ also within dielectrics, with $\mathbf{D} = \epsilon_o \mathbf{E} + \mathbf{P}$, leads to $\nabla \cdot (\epsilon_o \mathbf{E}) = \rho_f - \nabla \cdot \mathbf{P}$, which is the expected form for Gauss's law within dielectrics with $\rho_b = -\nabla \cdot \mathbf{P}$;

– in conclusion, we can write Gauss's law in a general form applicable everywere as

$$\nabla \cdot \mathbf{D} = \rho_f,$$

with the understanding that $\mathbf{D} \equiv \epsilon_o \mathbf{E} + \mathbf{P}$ and "includes" the effect of bound charge density $\rho_b = -\nabla \cdot \mathbf{P}$ which may be non-zero[6] within dielectrical materials.

– Typcially subscript f of ρ_f is dropped in Gauss's law, with the understanding that ρ refers to ρ_f because any non-zero $\rho_b = -\nabla \cdot \mathbf{P}$ effects have already been "lumped" into $\mathbf{D} = \epsilon_o \mathbf{E} + \mathbf{P}$ definition.



A dielectric slab inserted into a region with an initial field E_o will become polarized.

Inside the polarized dielectric the field will be weaker than E_o, but not reduced to zero as in a conductor.

- The differential form of Gauss's law that we will now write (without the subscript f on ρ_f) as

$$\nabla \cdot \mathbf{D} = \rho, \quad \text{[Gauss's law inside material medium]}$$

appears in integral form (after applying Divergence theorem to volume

[6]Note that if $\mathbf{P}$ is a constant inside a dielectric and zero outside then $\rho_b = -\nabla \cdot \mathbf{P}$ will be a surface charge density confined to the surface of the dielectric.

integral of the differential form) as

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = \int_V \rho dV,$$

where the right side denotes the net *free* charge inside volume V.

- In a large class of dielectric materials macroscopic polarization $\mathbf{P}$ and electric field $\mathbf{E}$ turn out to be linearly related (see Lecture 11) as

$$\mathbf{P} = \epsilon_o \chi_e \mathbf{E},$$

where $\chi_e \geq 0$ is a dimensionless quantity called **electric susceptibility**. For such materials

$$\mathbf{D} = \epsilon_o \mathbf{E} + \mathbf{P} = \epsilon_o (1 + \chi_e) \mathbf{E} = \epsilon \mathbf{E},$$

where

$$\epsilon = \epsilon_o (1 + \chi_e) \equiv \epsilon_r \epsilon_o$$

is known as the **permittivity** of the dielectric, and

$$\epsilon_r = 1 + \chi_e$$

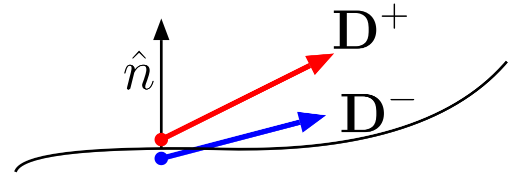
its **relative permittivity** or **dielectric constant**.

  - Dielectric constant of free space is 1,
    - for air $\epsilon_r \approx 1.0006$,

- ◦ for glass $4-10$,
- ◦ dry-to-wet earth $5-10$, silicon $11-12$, distilled water 81.

In certain materials χ_e and ϵ are found to be tensors — meaning that $\mathbf{P}$ and $\mathbf{D}$ are no longer aligned with $\mathbf{E}$. Such materials are said to be **anisotropic**, but they will not be studied in this course. Also, there is an exception to the condition $\chi_e \geq 0$ — in collisionless *plasmas* $\chi_e < 0$, as discussed in ECE 450.

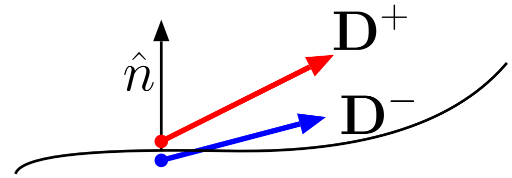
- In Gauss's law applicable in material media ρ denotes the free charge carrier density (after the revisions we have agreed to make). Furthermore, in perfect dielectrics there are no mobile free carriers and Gauss's law typically reduces to $\nabla\cdot\mathbf{D} = 0$, while the corresponding **boundary condition** equation for surfaces separating perfect dielectrics becomes



$$\hat{n}\cdot(\mathbf{D}^+ - \mathbf{D}^-) = 0 \quad\Rightarrow\quad D_n^+ = D_n^-,$$

which says that normal component of displacement $\mathbf{D}$ is continuous on such surfaces. This is accompanied by

$$\hat{n}\times(\mathbf{E}^+ - \mathbf{E}^-) = 0 \quad\Rightarrow\quad E_t^+ = E_t^-$$



stating the continuity of tangential components of $\mathbf{E}$, which is universally true as we have seen earlier.

9 Static fields in dielectric media

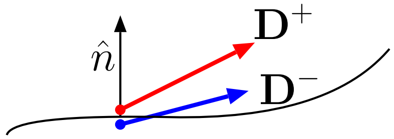
- Summarizing important results from last lecture:

  - within a dielectric medium, displacement

$$\mathbf{D} = \epsilon\mathbf{E} = \epsilon_o\mathbf{E} + \mathbf{P},$$

    and if the permittivity $\epsilon = \epsilon_r\epsilon_o$ is known, $\mathbf{D}$ and $\mathbf{E}$ can be calculated from free surface charge ρ_s or volume charge ρ in the region without resorting to $\mathbf{P}$.

  - on surfaces separating perfect dielectrics, $\hat{n} \cdot (\mathbf{D}^+ - \mathbf{D}^-) = 0$ typically, while $\hat{n} \cdot \mathbf{D}^+ = \rho_s$ on a conductor-dielectric interface (with $\hat{n}$ pointing from the conductor toward the dielectric).



  - Gauss's law $\nabla \cdot \mathbf{D} = \rho$ (and its integral counterpart) includes only the free charge density on its right side, which is typically zero in many practical problems.

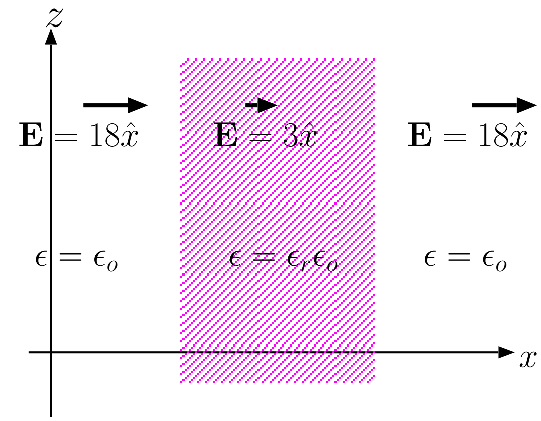
  - once $\mathbf{D}$ and $\mathbf{E}$ have been calculated (typically using the boundary condition equations), polarization $\mathbf{P}$ can be obtained as

$$\mathbf{P} = \mathbf{D} - \epsilon_o\mathbf{E}$$

    if needed.

These rules will be used in the examples in this section.

Example 1: A perfect dielectric slab having a finite thickness W in the x direction is surrounded by free space and has a constant electric field $\mathbf{E} = 18\hat{x}$ V/m in its exterior. Induced polarization of bound charges inside dielectric reduces the electric field strength inside the slab from $18\hat{x}$ V/m to $\mathbf{E} = 3\hat{x}$ V/m. What are the displacement field $\mathbf{D}$ and polarization $\mathbf{P}$ outside and inside the slab, and what are the dielectric constant ϵ_r and electric susceptibility χ_e of the slab?



Solution: Displacement field outside the slab, where $\epsilon = \epsilon_o$, must be

$$\mathbf{D} = \epsilon_o \mathbf{E} = \hat{x} 18\epsilon_o \frac{\mathrm{C}}{\mathrm{m}^2}.$$

The outside polarization $\mathbf{P}$ is of course zero. Boundary conditions at the interface of the slab with free space require the continuity of normal component of $\mathbf{D}$ and tangential component of $\mathbf{E}$ — both of these conditions would be satisfied if we were to take $\mathbf{D} = \hat{x}18\epsilon_o$ C/m^2 also within the dielectric slab. Thus, with $\mathbf{E} = 3\hat{x}$ V/m inside the slab, the condition $\mathbf{D} = \epsilon_{slab}\mathbf{E}$ within the slab requires that

$$\epsilon_{slab} = 6\epsilon_o.$$

Consequently, the dielectric constant of the slab is

$$\epsilon_r = 1 + \chi_e = \frac{\epsilon_{slab}}{\epsilon_o} = 6$$

and its electric susceptibility is

$$\chi_e = \epsilon_r - 1 = 5.$$

Finally, since $\mathbf{D} = \epsilon_o\mathbf{E} + \mathbf{P}$ in general, polarization $\mathbf{P}$ inside the slab is

$$\mathbf{P} = \mathbf{D} - \epsilon_o\mathbf{E} = \hat{x}18\epsilon_o - \epsilon_o 3\hat{x} = \hat{x}15\epsilon_o \frac{\mathrm{C}}{\mathrm{m}^2}.$$

- Our revised definition of displacement $\mathbf{D} = \epsilon\mathbf{E}$, where $\epsilon = \epsilon_r\epsilon_o$, implies, when combined with $\mathbf{E} = -\nabla V$ and $\nabla \cdot \mathbf{D} = \rho$, a revised form of Poisson's equation

$$\nabla^2 V = -\frac{\rho}{\epsilon},$$

  - provided that dielectric constant ϵ_r is independent of position so that $\nabla \cdot \mathbf{D} = \nabla \cdot (\epsilon\mathbf{E}) = \epsilon\nabla \cdot \mathbf{E}$ is a valid intermediate step in the derivation of Poisson's equation.
  - Under the same condition Laplace's equation $\nabla^2 V = 0$ also remains valid.
  - Dielectrics where ϵ_r is independent of position are said to be **homogeneous**.
    - In **inhomogeneous** dielectrics where ϵ varies with position neither equation is valid, and one has to resort to the full form of Gauss's law in field and potential calculations.

  In other words, don't use Laplace's/Poisson's equations in inhomogeneous media.

  In the next example we have two homogeneous slabs side-by-side making up an inhomogeneous configuration. In that case we can use Laplace/Poisson within the slabs one at a time and then match the results at the boundary using boundary condition equations as shown.

Example 2: A pair of infinite conducting plates at $z = 0$ and $z = 2$ m carry equal and opposite surface charge densities of $-2\epsilon_o$ C/m^2 and $2\epsilon_o$ C/m^2, respectively. Determine $V(2)$ if $V(0) = 0$ and regions $0 < z < 1$ m and $1 < z < 2$ m are occupied by perfect dielectrics with permittivities of ϵ_o and $2\epsilon_o$, respectively.

Solution: Given that $V(0) = 0$, we assume $V(z) = Az$, for some constant A in the homogeneous region $0 < z < 1$ m, since $V(z) = Az$ satisfies the Laplace's equation as well as the boundary condition at $z = 0$.

This gives $V(1) = A$ at $z = 1$ m, which then implies that we can take $V(z) = A + B(z-1)$ for the second homogeneous region $1 < z < 2$ m having a different permittivity than the region below.

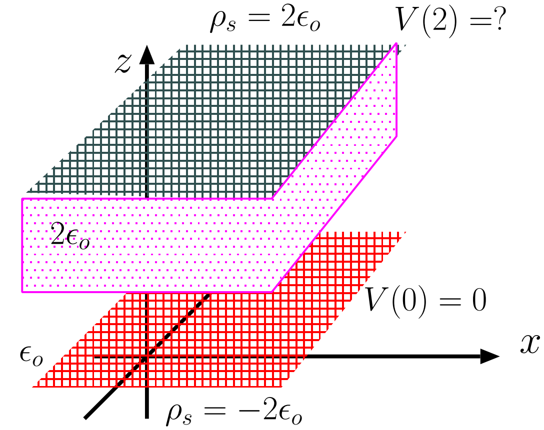
To determine the constants A and B, we will make use of boundary conditions at $z = 0$ and $z = 1$ m interfaces:

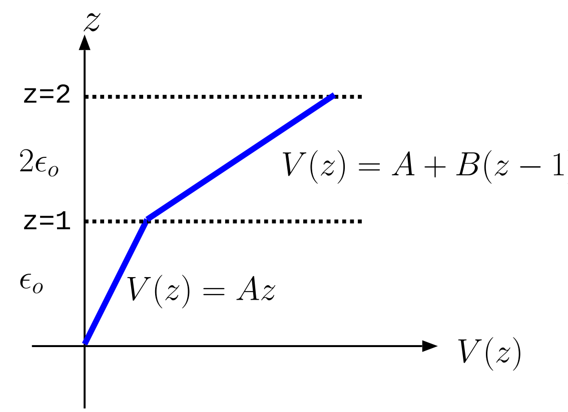
- In the region $0 < z < 1$ m, the electric field $\mathbf{E} = -\nabla(Az) = -A\hat{z}$, and, therefore displacement $\mathbf{D} = \epsilon_1\mathbf{E} = -\epsilon_o A\hat{z}$. Hence, the pertinent boundary condition $\hat{z}\cdot\mathbf{D}(0) = \rho_s$ yields

$$\hat{z}\cdot\mathbf{D}(0) = -\epsilon_o A = -2\epsilon_o \quad \Rightarrow \quad A = 2.$$

- Just below $z = 1$ m the displacement is $\mathbf{D}(1^-) = -\epsilon_o A\hat{z} = -2\epsilon_o\hat{z}$ as we found out above. Above $z = 1$ m, the electric field is $\mathbf{E} = -\nabla(A + B(z-1)) = -B\hat{z}$, and, therefore, $\mathbf{D}(1^+) = -2\epsilon_o B\hat{z}$ just above $z = 1$ m. Hence, the pertinent boundary condition $\hat{z}\cdot(\mathbf{D}(1^+) - \mathbf{D}(1^-)) = 0$ yields

$$\hat{z}\cdot(-2\epsilon_o B\hat{z} - (-2\epsilon_o\hat{z})) = -2\epsilon_o B + 2\epsilon_o = 0 \quad \Rightarrow \quad B = 1.$$



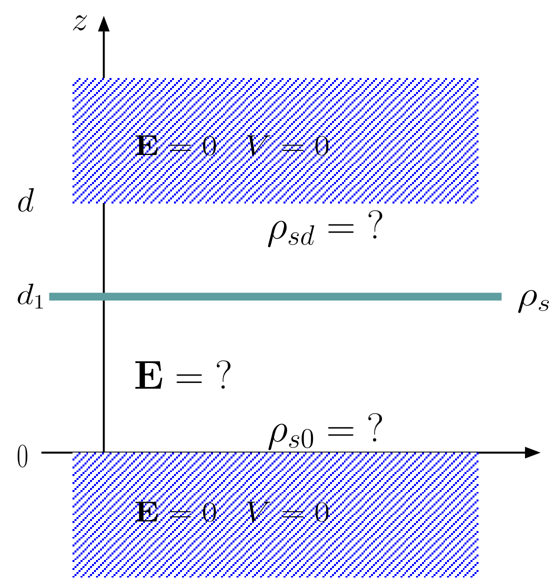


Based on above calculations of constants A and B, the potential solution for the region is

$$V(z) = \begin{cases} 2z\,\mathrm{V}, & 0 < z < 1 \\ 2 + (z-1)\,\mathrm{V}, & 1 < z < 2. \end{cases}$$

It follows that $V(2) = 3$ V.

Note that electric fields $-2\hat{z}$ V/m and $-\hat{z}$ V/m in the bottom and top layers point from high to low potential regions. Electric field $\mathbf{E}$ is discontinuous at the boundary at $z = 1$ m while displacement $\mathbf{D}$ is continuous — the continuity of normally directed $\mathbf{D}$ is demanded by boundary condition equations in the absence of surface charge.



If ρ_s in Example 3 is a slowly-varying function of time, then slowly varying $\mathbf{E}$, ρ_{s0}, and ρ_{sd} calculated with instantaneous values of ρ_s would constitute *quasi-static solutions* which are valid so long as $d \ll c/f$, with f the highest frequency in $\rho_s(t)$.

Example 3: A pair of infinite conducting plates at $z = 0$ and $z = d$ are grounded and have equal potentials, say, $V = 0$. The region $0 < z < d$ is occupied by free space (i.e., $\epsilon = \epsilon_o$) except that an infinite charge sheet with a static surface charge density ρ_s is located at $z = d_1 < d$. Determine (a) the electrostatic field $\mathbf{E}(z)$ in regions $0 < z < d_1$ and $d_1 < z < d$, and (b) the surface charge densities ρ_{s0} and ρ_{sd} at $z = 0$ and $z = d$ on conductor surfaces if $d_1 = d/2$.

Solution: (a) Laplace's equation for the given geometry requires a linear (in z) potential solution in regions $0 < z < d_1$ and $d_1 < z < d$. Since electrostatic $\mathbf{E} = -\nabla V$, we can therefore represent the electric field in these regions as

$$\mathbf{E} = \begin{cases} -\hat{z}V_o/d_1, & 0 < z < d_1 \\ +\hat{z}V_o/d_2, & d_1 < z < d \end{cases}$$

where $V_o \equiv V(d_1)$ and $d_2 \equiv d - d_1$. Hence,

$$\mathbf{D} = \epsilon_o \mathbf{E} = \begin{cases} -\hat{z}\epsilon_o V_o/d_1, & 0 < z < d_1 \\ +\hat{z}\epsilon_o V_o/d_2, & d_1 < z < d \end{cases},$$

and Maxwell's boundary condition equation applied on $z = d_1$ surface is

$$\hat{z} \cdot (\mathbf{D}(d_1^+) - \mathbf{D}(d_1^-)) = \rho_s \;\;\Rightarrow\;\; \epsilon_o V_o \left(\frac{1}{d_2} + \frac{1}{d_1} \right) = \rho_s.$$

Thus

$$V_o = \frac{\rho_s}{\epsilon_o} \left(\frac{1}{d_2} + \frac{1}{d_1} \right)^{-1} = \frac{\rho_s}{\epsilon_o} \frac{d_1 d_2}{d_1 + d_2} = \frac{\rho_s}{\epsilon_o} \frac{d_1 d_2}{d}.$$

Substituting V_o back into the expression for $\mathbf{E}$, we have

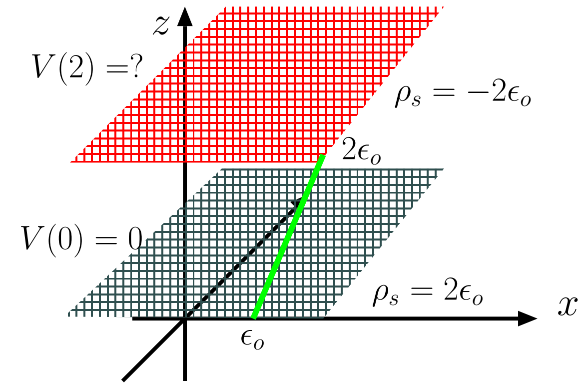
$$\mathbf{E} = \begin{cases} -\hat{z}\frac{\rho_s}{\epsilon_o}\frac{d_2}{d}, & 0 < z < d_1 \\ +\hat{z}\frac{\rho_s}{\epsilon_o}\frac{d_1}{d}, & d_1 < z < d. \end{cases}$$

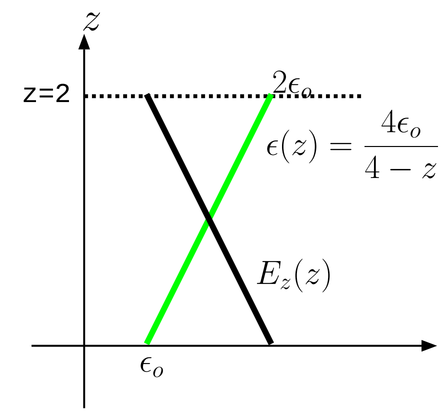
(b) The surface charge at $z = 0$ can be found by evaluating $\hat{z} \cdot \mathbf{D} = \hat{z} \cdot \epsilon_o \mathbf{E}$ at $z = 0$. Hence,

$$\rho_{s0} = \hat{z} \cdot \epsilon_o \mathbf{E}(0) = -\frac{d_2}{d}\rho_s \xrightarrow{d_1 = d/2} -\frac{\rho_s}{2}.$$

Likewise,

$$\rho_{sd} = -\hat{z} \cdot \epsilon_o \mathbf{E}(d) = -\frac{d_1}{d}\rho_s \xrightarrow{d_1 = d/2} -\frac{\rho_s}{2}.$$





Example 4: Between a pair of infinite conducting plates at $z = 0$ and $z = 2$ m, the medium is a perfect dielectric with an **inhomogeneous** permittivity of

$$\epsilon(z) = \frac{4\epsilon_o}{4-z}.$$

Determine the electric potential $V(2)$ on the top plate if $V(0) = 0$ and the surface charge density is $\rho_s = 2\epsilon_o$ C/m^2 on the bottom plate at $z = 0$. Note that Laplace's equation cannot be used in this problem since the medium is inhomogeneous.

Solution: Consider Gauss's law

$$\nabla \cdot (\epsilon \mathbf{E}) = \rho$$

with $\rho = 0$ in the region $0 < z < 2$ m. Assuming that $\mathbf{E} = \hat{z}E_z(z)$, because the geometry is invariant in x and y, we have

$$\nabla \cdot (\epsilon \mathbf{E}) = 0 \ \Rightarrow \ \frac{\partial}{\partial z}(\epsilon E_z) = 0 \ \Rightarrow \ \epsilon E_z = \text{constant}.$$

Thus the product ϵE_z is invariant with respect to coordinate z, which implies that

$$\epsilon(z)E_z(z) = \epsilon(0)E_z(0) \ \Rightarrow \ E_z(z) = \frac{\epsilon(0)}{\epsilon(z)}E_z(0) = E_z(0)(1 - \frac{z}{4})$$

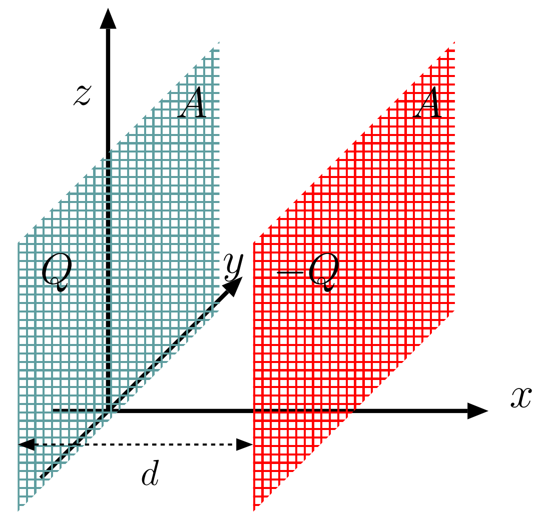
after substituting for $\epsilon(z)$. To identify $E_z(0)$, we apply the bottom boundary condition $\hat{z} \cdot \mathbf{D}(0) = \rho_s$, and obtain

$$D_z(0) = \epsilon(0)E_z(0) = 2\epsilon_o \ \Rightarrow \ E_z(0) = \frac{2\epsilon_o}{\epsilon(0)} = 2\,\frac{\text{V}}{\text{m}}.$$

To determine $V(2)$, we integrate $\mathbf{E} = \hat{z}2(1 - \frac{z}{4})$ V/m from top to bottom plate (grounded), obtaining

$$\begin{aligned} V(2) &= \int_{z=2}^{0} \mathbf{E} \cdot d\mathbf{l} = \int_{z=2}^{0} 2(1 - \frac{z}{4}) dz \\ &= 2(z - \frac{z^2}{8})|_2^0 = -2(2 - \frac{4}{8}) = -2 \cdot \frac{3}{2} = -3\,\text{V}. \end{aligned}$$

10 Capacitance and conductance



Parallel-plate capacitor: Consider a pair of conducting plates with surface areas A separated by some distance d in free space (see margin).

The plates are initially charge neutral, but then some amount of electrons are transferred from one plate to the other so that the plates acquire equal and opposite charges Q and $-Q$, distributed with surface densities of $\pm\frac{Q}{A}$ on plate surfaces facing one another (as shown in the margin).

- That way, in steady state and for $d \ll \sqrt{A}$, a field configuration confined mainly to the region between the plates is acquired, satisfying the condition that static field inside a conductor should be zero. A weak "fringing field" can be ignored if $d \ll \sqrt{A}$ and thus the geometry well approximates the case with infinite plates.

  – A constant displacement field

$$\mathbf{D} = \hat{x}\frac{Q}{A}$$

  satisfies the normal boundary condition at the left plate boundary as well as Gauss's law $\nabla \cdot \mathbf{D} = 0$ in the region between the plates. The corresponding electrostatic field is

$$\mathbf{E} = \frac{\mathbf{D}}{\epsilon_o} = \hat{x}\frac{Q}{\epsilon_o A},$$

  and the voltage drop from (positive charged) left plate to (negative charged) right plate is

$$V = \int_{(0,0,0)}^{(d,0,0)} \mathbf{E} \cdot d\mathbf{l} = \int_{x=0}^{d} \frac{Q}{\epsilon_o A} dx = \frac{d}{\epsilon_o A} Q.$$

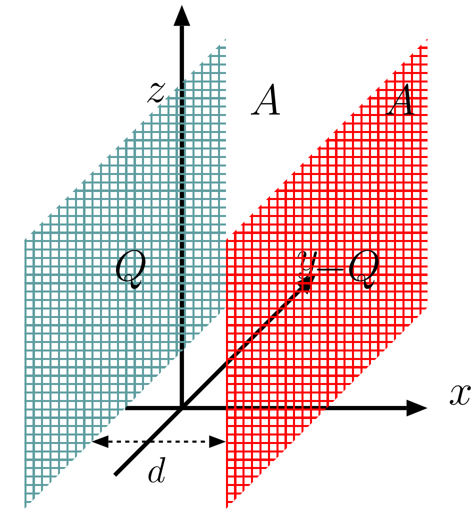
The last result can be expressed as a *linear charge-voltage relation*

$$Q = CV$$

with

$$C \equiv \epsilon_o \frac{A}{d}$$

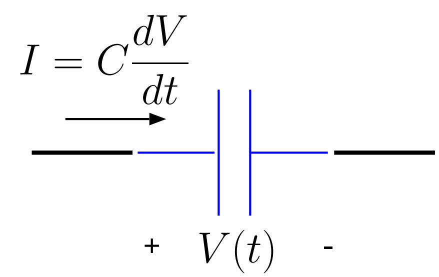
representing the **capacitance** of the parallel conducting plate arrangement that we call **parallel plate capacitor**.



- By differentiating $Q = CV$ we obtain the charging rate of the capacitor as

$$I = \frac{dQ}{dt} = C\frac{dV}{dt}$$

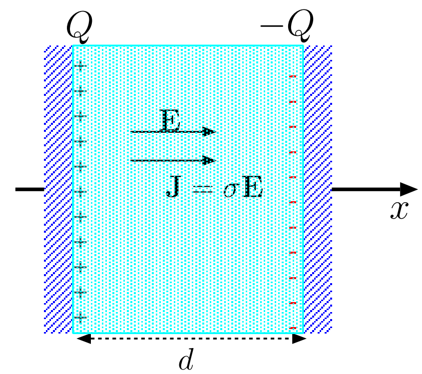
which is only possible, for ideal capacitors, if the capacitor plates are externally connected to a circuit supplying a current as shown on the right where the direction of $I = \frac{dQ}{dt}$ is in the direction of voltage drop V *across* the capacitor, from the positively charged plate to the negatively charged plate as shown.



  – In lossy capacitors when the medium between the plates is conducting, the charging rate of the capacitor plate will be smaller as given by

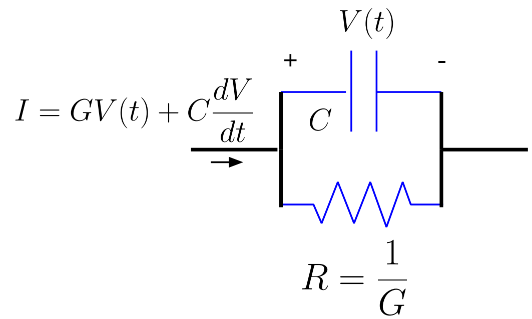
$$\frac{dQ}{dt} = C\frac{dV}{dt} = I - GV,$$

where G stands for the conductance of the capacitor (derived later in this lecture) and I the external current flowing into the non-ideal capacitor.



– Therefore, for non-ideal capacitors the external current

$$I = C\frac{dV}{dt} + GV,$$



meaning that part of I goes into changing stored charge $Q = CV$ of capacitor plates and the rest to conduct a GV amount of leakage current of the capacitor plates, and the equivalent circuit of the non-ideal capacitor there contains a "shunt resistance" $R = \frac{1}{G}$ accompanying C as shown in the margin.

- Returning to the IV-relation

$$I = C\frac{dV}{dt}$$

of the ideal capacitor, this IV-relation was obtained from the QV-relation above *quasi-statically* assuming that $\sqrt{A} \ll \lambda = c/f$, where f is the highest frequency of $V(t)$. The power absorbed by the capacitor is then calculated as

$$P = VI = VC\frac{dV}{dt} = \frac{d}{dt}(\frac{1}{2}CV^2),$$

implying a stored energy of

$$W = \frac{1}{2}CV^2 = \frac{1}{2}\epsilon_o|E_x|^2 Ad$$

when the capacitor is in a charged state.

- Notice that stored energy is

$$\frac{1}{2}\epsilon_o E_x^2 = \frac{1}{2}\epsilon_o \mathbf{E}\cdot\mathbf{E}$$

times the volume Ad occupied by the field $\mathbf{E}$ between the capacitor plates. That suggests that

$$w = \frac{1}{2}\epsilon_o \mathbf{E}\cdot\mathbf{E}$$

can be interpreted as stored electrostatic energy per unit volume in general.

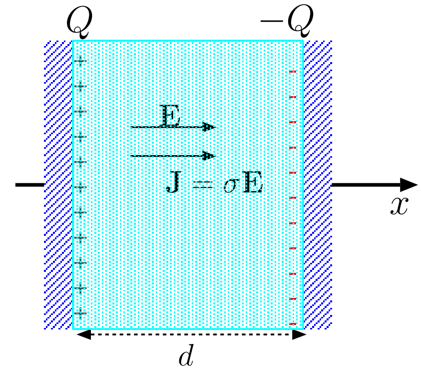
  - Also both capacitance C and stored energies W and w would have ϵ replacing ϵ_o in dielectric media.

A capacitor with a perfect dielectric between its plates will hold its charge and stored energy indefinitely. However, if the dielectric is imperfect and has a finite **conductivity** σ, charge will be transported from the positive to negative plate by a volumetric current density

$$\mathbf{J} = \sigma\mathbf{E},$$

which will result in a *quasi-static* discharge of the capacitor and the loss of the stored energy W to Ohmic dissipation in the imperfect dielectric.

Just as **capacitance** C characterizes the energy and charge storage "capacity" of the capacitor, we can define a **conductance** G that relates the quasi-static discharge current I in between the plates of a capacitor to potential drop V:



- Discharge current I is the product of current density

$$J_x = \sigma E_x$$

in A/m^2 units and the plate area A. Since $E_x = \frac{V}{d}$, we obtain a *linear current-voltage relation*

$$I = GV$$

with conductance

$$G \equiv \sigma \frac{A}{d}$$

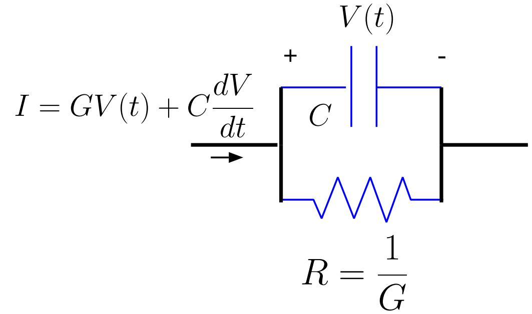
for the parallel plate capacitor.

  – Notice that $G = \frac{\sigma}{\epsilon} C$, a relation that will hold true for other types of capacitors that we will be examining.

  – Also,

$$R \equiv \frac{1}{G} = \frac{d}{A\sigma}$$

  is the corresponding **resistance** that scales inversely with conductivity σ of the material — large σ materials will have small



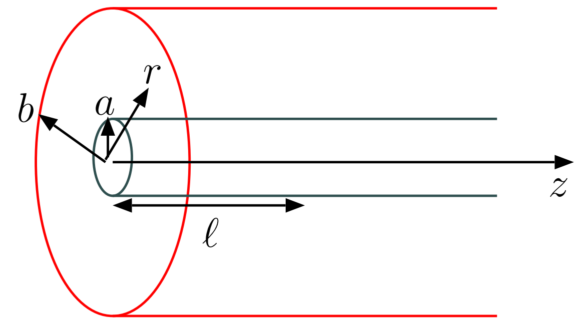
System above behaves like a resistor $R = 1/G$ for

$$\omega \ll \frac{G}{C} = \frac{1}{RC} = \frac{\sigma}{\epsilon}$$

and like a capacitor C in the complementary frequency band. To obtain capacitor behavior at low frequencies make sure that σ is sufficiently small.

Alternatively, with large σ the system becomes a good electrical connector, a resistor R with a small resistance $R \propto 1/\sigma$.

resistance, but for a given σ, R increases with length d and decreases with increasing cross-sectional area A. Simple conductivity models and $\mathbf{J}$ will be discussed next lecture.



Coaxial Cable: When we study guided wave propagation later in the course we will learn about coaxial cables.

- A coax cable consists of two conducting regions — a central cylindrical conductor with a cross-sectional radius a, enclosed by a conducting pipe of a radius $b > a$ (see margin), with some dielectric ϵ filling in the space. We will next calculate the capacitance and conductance of a coax segment of some length ℓ.

- For $\ell \gg b$, field $\mathbf{E}$ can be assumed to point out radially away from the inner conductor with radius a to the outer conductor with radius b. In that case Gauss's law in integral form can be utilized to determine the radial field E_r. Considering a cylindrical integration surface with a radius $r > a$ centered about the inner conductor, we re-write Gauss's law

$$\epsilon \oint_s \mathbf{E} \cdot d\mathbf{S} = Q_V$$

as

$$\epsilon E_r 2\pi r \ell = Q$$

where Q is the total charge distributed over the inner conductor and ϵ the permittivity of the dielectric separating the two conductors.

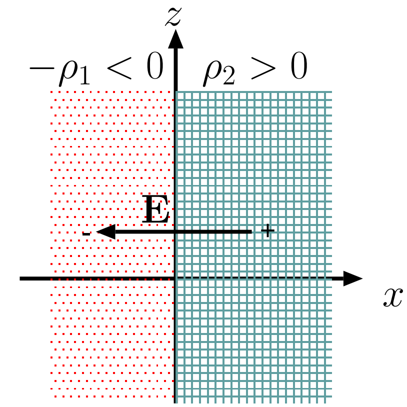
– It follows that
$$E_r = \frac{Q}{2\pi\epsilon\ell r},$$
and voltage drop from inner to outer conductor is
$$V = \int_{r=a}^{b} E_r dr = \int_{r=a}^{b} \frac{Q}{2\pi\epsilon\ell r} dr = \frac{Q}{2\pi\ell\epsilon}\int_{r=a}^{b} \frac{dr}{r} = \frac{Q}{2\pi\ell\epsilon}\ln\frac{b}{a}.$$
Clearly, once again $Q = CV$, with
$$C = \frac{2\pi}{\ln\frac{b}{a}}\ell\epsilon$$
representing the capacitance of the coax segment of length l.

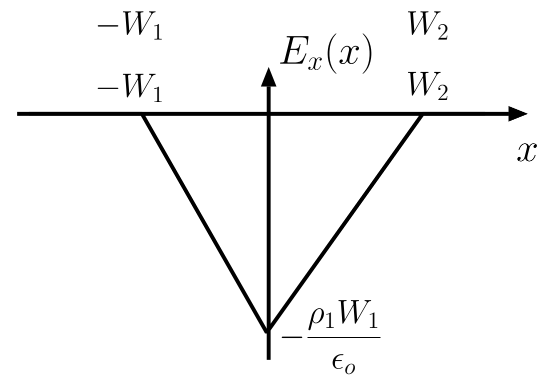
- The **capacitance** of the coax **per unit length** is
$$\mathcal{C} = \frac{2\pi}{\ln\frac{b}{a}}\epsilon.$$

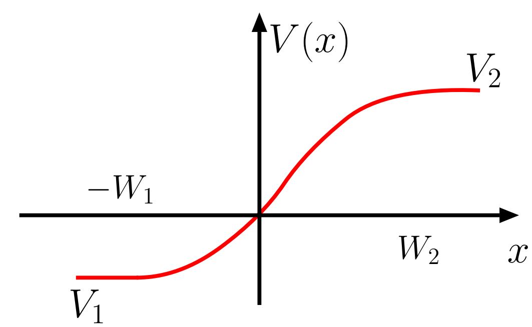
  – **Conductance** of the coax **per unit length** can likewise be shown to be
$$\mathcal{G} = \frac{2\pi}{\ln\frac{b}{a}}\sigma.$$
  This result is a consequence of the general relation $G = \frac{\sigma}{\epsilon}C$ mentioned earlier.

  – Per length parameters $\mathcal{C}$ and $\mathcal{G}$ of the coax will play an important role when we study guided wave propagation in coaxial transmission lines with lengths for which quasi-static approximation may be violated.

Diode junctions: In Example 4 in Lecture 7 we derived the expression for potential drop V across a charged region of a total width of $W_1 + W_2$, such that in region 1 where $-W_1 < x < 0$ the charge density $\rho = -\rho_1$ is negative, while in region 2 where $0 < x < W_2$ the charge density $\rho = \rho_2$ is positive, with the additional constraint that the entire region is charge neutral, meaning that $\rho_1 W_1 = \rho_2 W_2$.







By solving *Poisson's equation* for this charge density configuration (see margin) encountered in **junction** regions of semiconductor diodes (described in detail ECE 440) we had established that the voltage drop from $x = W_2$ to $x = -W_1$ across the junction is given by

$$V = \frac{\rho_2 W_2 (W_1 + W_2)}{2\epsilon_o} = \frac{\rho_1 W_1 (W_1 + W_2)}{2\epsilon_o}.$$

The above equation implies that

$$W_1 = \frac{2\epsilon_o V}{(W_1 + W_2)\rho_1} \text{ and } W_2 = \frac{2\epsilon_o V}{(W_1 + W_2)\rho_2} \quad \Rightarrow \quad W_1 + W_2 = \sqrt{2\epsilon_o V \frac{\rho_1 + \rho_2}{\rho_1 \rho_2}}.$$

Using the expressions above for junction voltage V and width $W_1 + W_2$, we will next derive an expression for **small signal capacitance** of the diode junction:

- In region 2 where $x > 0$, the junction holds a total positive charge of $Q = \rho_2 W_2 A$ per cross-sectional area A.

- Therefore, substituting $\frac{Q}{A}$ for $\rho_2 W_2$ in the expression for V above, and also using the $W_1 + W_2 \propto \sqrt{V}$ expression derived above, we obtain

$$V = \frac{\rho_2 W_2 (W_1 + W_2)}{2\epsilon_o} = \frac{Q\sqrt{2\epsilon_o V \frac{\rho_1+\rho_2}{\rho_1\rho_2}}}{2\epsilon_o A},$$

which can be re-arranged as

$$Q = A\sqrt{\frac{2\epsilon_o \rho_1 \rho_2}{\rho_1 + \rho_2}}\sqrt{V}$$

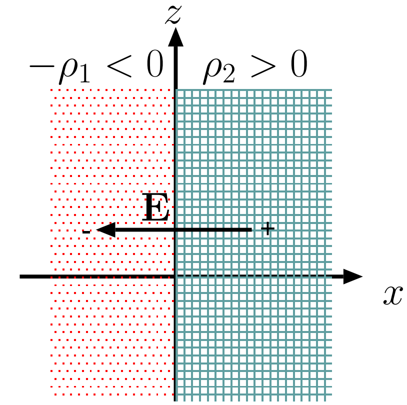
representing a *non-linear* charge-voltage relation (for a given charge profiles satisfying $\rho_1 W_1 = \rho_2 W_2$).

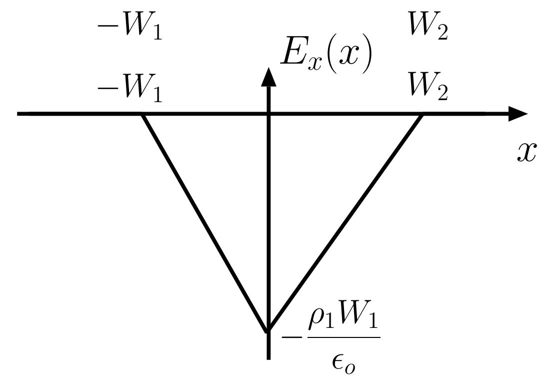
– In a linear charge-voltage relation $Q = CV$, the capacitance parameter C represents the slope $\frac{Q}{V}$ of a Q vs V curve.

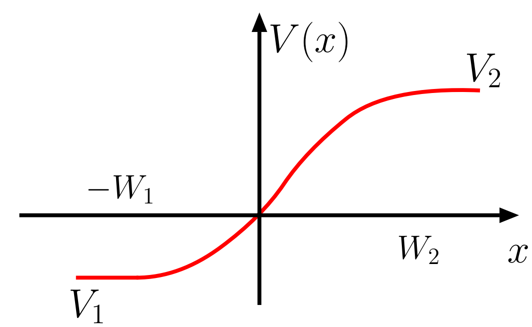
The slope of *any* Q vs V curve is given by the derivative $\frac{dQ}{dV}$, whether or not the curve is linear. The slope $\frac{dQ}{dV}$ of a *non-linear* charge-voltage curve can be interpreted as a *small signal capacitance* C. For a diode junction, differentiating the above equation, we find that

$$C = \frac{dQ}{dV} = A\sqrt{\frac{\epsilon_o}{2V}\frac{\rho_1\rho_2}{(\rho_1 + \rho_2)}}.$$

Small changes dV in junction voltage will accompany small changes $dQ = CdV$ in stored charge Q of the junction, but the amount CdV will itself depend on V because $C \propto V^{-1/2}$.



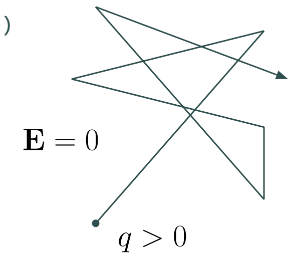




11 Lorentz-Drude models for conductivity and susceptibility and polarization current

In this lecture we will describe simple microscopic models for conductivity σ and electric susceptibility χ_e of material media composed of free and bound charge carriers. The models were first developed by Lorentz and Drude prior to the establishment of quantum mechanics. In these models free charge carriers motions are described using Newtonian dynamics and atoms are represented as electric dipoles $\mathbf{p} = -e\mathbf{r}$ ($\mathbf{r}$ is electron displacement from atomic nucleus) behaving like damped 2nd order systems.

(a)

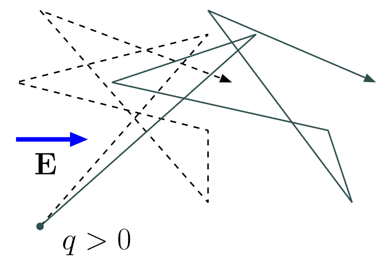


In the absence of an applied electric field E, free charge q exhibits a "random walk" between collisions such that its average velocity v is zero.

By collisions we refer to the inreactions of q with zero-mean miscroscopic electric fields within the conductor due to charges entrapped in the lattice.

Conductivity: Conducting materials such as copper, sea water, ionized gases (plasmas) contain a finite density N of mobile and free charge carriers at the microscopic level (in addition to neutral atoms and molecules sharing the same macroscopic space) — these elementary mobile carriers can be electrons, positive or negative ions, or positive "holes" (in semi-conductor materials).

(b)



In the presence of an applied electric field E, the mean position of free charge q>0 drifts in the direction of field vector E with some non-zero mean velocity v.

Avg. drift velocity v reperesents a balance between acceleration force due to E and an opposing friction force produced by collisions of q with the lattice at random intervals with some mean value τ .

- Each elementary charge carrier with a charge q and mass m and subject to a macroscopic electrical force $q\mathbf{E}$ will be modelled by a dynamic equation

$$m\frac{d\mathbf{v}}{dt} = q\mathbf{E} - m\frac{\mathbf{v}}{\tau},$$

  which is effectively Newton's second law — "force equals mass times acceleration" — in which $\mathbf{v}$ denotes the macroscopic velocity[1] of charge

[1]Think of *microscopic* velocity of each charge carrier as $\mathbf{v} + \delta\mathbf{v}$, where $\delta\mathbf{v}$ is an independent zero-mean

carriers and $-m\frac{\mathbf{v}}{\tau}$ denotes a macroscopic friction force proportional to $-\mathbf{v}$. Friction is a consequence of "collisions" of charge carriers with the neutral background at a frequency of $\nu = \frac{1}{\tau}$ collisions per unit time, and causes the decay of $\mathbf{v}$ as

$$\mathbf{v}(t) = \mathbf{v}(0)e^{-t/\tau}$$

in the absence of field $\mathbf{E}$. Therefore, when $\mathbf{E} = 0$ the carriers settle down to a steady state with $\mathbf{v} = 0$ (in $t \gg \tau$ limit), meaning that no macroscopic current density $\mathbf{J}$ will be found in the absence of $\mathbf{E}$ in regions with homogeneous charge carrier densities.

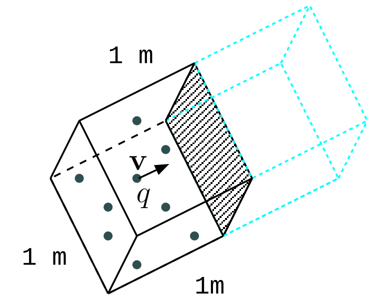
- With a constant but non-zero $\mathbf{E}$, steady-state solution of the above equation is

$$\mathbf{v} = \frac{q\tau}{m}\mathbf{E}, \quad \text{where } |\frac{q\tau}{m}| \text{ is known as } \textbf{mobility}.$$

  – Assuming N charge carriers per unit volume each moving (on the average) with this steady-state velocity in a given material, we can calculate the average flux density of charge through the region as

$$\mathbf{J} = Nq\mathbf{v} = \frac{Nq^2}{m\nu}\mathbf{E}\,\frac{\mathrm{C/s}}{\mathrm{m}^2}$$

  which is commonly referred to as **current density** (see margin). If a given material contains several species of carriers with charge,



A cube of unity volume contains N charge carriers q each moving with an average velocity of v=1 m/s towards the shaded surface of the cube of a unity surface area.

In 1 second all N charges in the volume cross the shaded surface, transporting Nq Coulombs of charge per second per unit area.

By contrast, charge transport rate across the same surface is

$$Nqv \ \mathrm{C/s/m^2} = Nqv \ \mathrm{A/m^2}$$

if the average charge velocity is an arbitrary v.

Thus, in vector notation, we can define the current density in the region as

$$\mathbf{J} = Nq\mathbf{v}$$

such that current I (in A) across any surface S having area elements dS is given by

$$I = \int_S \mathbf{J}\cdot d\mathbf{S}$$

random variable for each charge carrier whereas *macroscopic* velocity $\mathbf{v}$ corresponds to the statistical average of all $\mathbf{v} + \delta\mathbf{v}$.

mass, collision frequency, and number density of q_s, m_s, ν_s, and N_s, respectively, then current density can be expressed as

Ohm's Law and DC conductivity

$$\mathbf{J} = \sigma\mathbf{E},$$

with

$$\sigma = \sum_s \sigma_s \quad \text{and} \quad \sigma_s = \frac{N_s q_s^2}{m_s \nu_s}$$

denoting the medium and species conductivities, respectively, under DC conditions.

- With a time varying field $\mathbf{E}$ the corresponding current density will also be time varying, in which case conductivity σ should be defined in the frequency domain using phasor techniques (remember ECE 210).

  – Briefly, using phasors $\tilde{\mathbf{E}}$ and $\tilde{\mathbf{J}}$ such that

  $$\mathbf{E}(t) = \text{Re}\{\tilde{\mathbf{E}}e^{j\omega t}\} \quad \text{and} \quad \mathbf{J}(t) = \text{Re}\{\tilde{\mathbf{J}}e^{j\omega t}\}, \text{ etc.,}$$

  we have a phasor transformed Newton's force balance equation

  $$m\frac{d\mathbf{v}}{dt} = q\mathbf{E} - m\frac{\mathbf{v}}{\tau} \quad \Rightarrow \quad mj\omega\tilde{\mathbf{v}} = q\tilde{\mathbf{E}} - m\frac{\tilde{\mathbf{v}}}{\tau},$$

  from which it follows that

  $$\tilde{\mathbf{J}} = \sigma\tilde{\mathbf{E}},$$

with

$$\sigma = \sum_s \sigma_s \quad \text{and} \quad \sigma_s = \frac{N_s q_s^2}{m_s(\nu_s + j\omega)}.$$

AC conductivity

– Note that the AC conductivity just derived can be approximated by the DC conductivity derived earlier for all AC frequencies ω much smaller than species collision frequencies ν_s.

  ∘ In many cases of practical interest, this condition can be easily met, and we are often well justified to ignore the frequency dependence and complex character of conductivity σ revealed in above derivation.

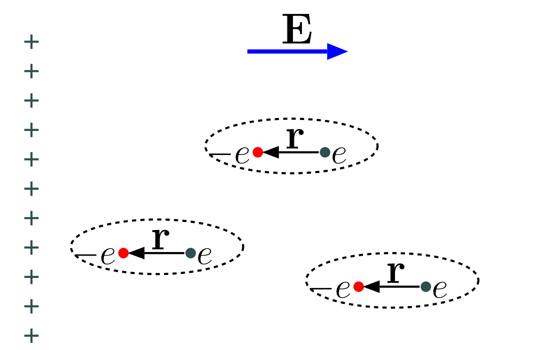
– More advanced quantum mechanical derivations of σ_s give the same results except with *effective masses* specified by quantum theory replacing the particle masses m_s used in classical models.

- Typical DC conductivities:

  – For silver, copper, gold, $\sigma \sim$ several $\times\, 10^7$ S/m
  – For seawater $\sigma \approx 4$ S/m
  – For intrinsic silicon $\sigma = 1.6 \times 10^{-3}$ S/m
  – For dry earth $\sigma \sim 10^{-5}$ S/m
  – For glass $\sigma \sim 10^{-10} - 10^{-14}$ S/m

Superconductivity occurs in certain materials at low temperatures when the DC conductivity vanishes as a consequence of correlated charge carrier motions which are ignored in the Lorentz-Drude model.

Susceptibility: In perfect dielectrics there are no free charge carriers and so $\sigma = 0$. However, in general such materials are polarizable and therefore they have a non-zero susceptibility χ_e and a dielectric constant $\epsilon_r = 1 + \chi_e > 1$.



- In Lorentz-Drude model, each polarized atom or molecule is considered to be a dipole $\mathbf{p} = -e\mathbf{r}$, with $\mathbf{r}$ representing the displacement vector of an atomic electron from atomic nucleus when the atom is polarized because of an applied electric field.

- If the polarizing force on the atom is removed, observations indicate that the dipole field of the atom $\mathbf{E}_p \propto \mathbf{p} \propto \mathbf{r}$ will decay as a damped co-sinusoid with a decay time constant $\tau_d = \frac{1}{\alpha}$ and a characteristic damped frequency $\sqrt{\omega_o^2 - \alpha^2} \approx \omega_o$ satisfying a condition $\omega_o \gg \alpha = \frac{1}{\tau_d}$ (strongly underdamped).

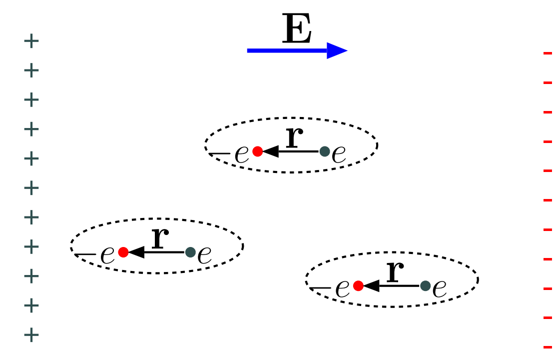
  - Possible values of ω_o for a given atom can be obtained from the energy levels of bound states of the atom (calculated using standard quantum[2] models like in PHYS 214) and time constants $\tau_d = \frac{1}{\alpha}$ (which are finite because energies $\hbar\omega_o$ radiated away are also finite) are related to observed line widths (2α) in the emission spectra of excited atoms.

---

[2]For a quantum mechanical derivation of susceptibility, see, e.g., *Mott*, "Elements of Wave Mechanics", Chapt 4, Sect 10 (1958); *Miller*, "Quantum Mechanics for Scientists and Engineers", Sect 7.3 (2008).

Electron displacement having the inferred damped co-sinusoid form

$$\mathbf{r}(t) = \mathbf{r}_o e^{-t/\tau_d} \cos(\sqrt{\omega_o^2 - \alpha^2} t) \approx \mathbf{r}_o e^{-t/\tau_d} \cos(\omega_o t)$$



is "zero-input response" (remember ECE 210) of a linear second-order ODE that can be constructed using Newton's second law of classical mechanics:

– If we assume that mass m times acceleration $\frac{d^2\mathbf{r}}{dt^2}$ of a displaced electron equals the sum of a

  ◦ force $-e\mathbf{E}$ exerted by an applied macroscopic electric field $\mathbf{E}$,
  ◦ a spring-like restoring force $-m\omega_o^2\mathbf{r}$ responsible for the binding of the electron to the nucleus, and
  ◦ a friction-like dissipative force $-m2\alpha\frac{d\mathbf{r}}{dt}$,

  we get

$$m\frac{d^2\mathbf{r}}{dt^2} = -e\mathbf{E} - m\omega_o^2\mathbf{r} - m2\alpha\frac{d\mathbf{r}}{dt},$$

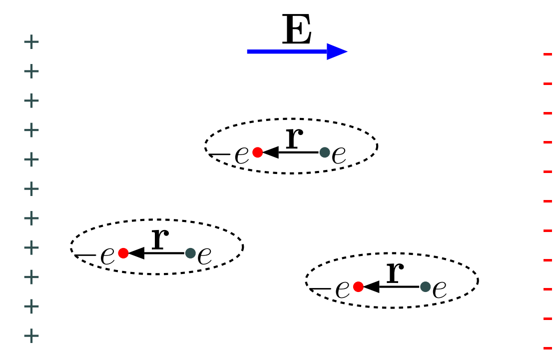
  for which $\mathbf{r}(t)$ given above is the zero-input solution in the absence of $\mathbf{E}$.

- To find the DC susceptibility of a dielectric composed of dipoles constrained by the above equation, we note that steady-state solution of the equation with a non-zero constant field $\mathbf{E}$ is

$$\mathbf{r} = -\frac{e}{m\omega_o^2}\mathbf{E}.$$

Consequently, dipole moment of a single polarized atom is

$$\mathbf{p} = -e\mathbf{r} = \frac{e^2}{m\omega_o^2}\mathbf{E},$$



and polarization field in a dielectric with a dipole density of N_d is

$$\mathbf{P} = N_d\mathbf{p} = \frac{N_d e^2}{m\omega_o^2}\mathbf{E}.$$

This result can also be written as

$$\mathbf{P} = \epsilon_o \chi_e \mathbf{E},$$

where

DC susceptibility

$$\chi_e \equiv \frac{N_d e^2/m\epsilon_o}{\omega_o^2}$$

is DC susceptibility. AC susceptibility can be derived using phasor techniques, but at frequencies $\omega \ll \omega_o$, AC susceptibility is well approximated by the DC susceptibility derived above.

Polarization current: Consider the case of a time varying electric field $\mathbf{E}(t)$ in a dielectric medium at a frequency $\omega \ll \omega_o$ such that the relations

$$\mathbf{r} = -\frac{e}{m\omega_o^2}\mathbf{E} \quad \text{and} \quad \mathbf{P} = \epsilon_o \chi_e \mathbf{E}$$

from above are accurate.

- Time variation of $\mathbf{E}$ will imply the time variation of electron displacement $\mathbf{r}$, so that there will be *in effect* a non-zero electron velocity

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = -\frac{e}{m\omega_o^2}\frac{d\mathbf{E}}{dt}$$

capable of producing a current.

  - With N_d such electrons per unit volume, each carrying a charge $-e$, we will have a net flux density of charge in the region given by

$$\mathbf{J}_p = -eN_d\mathbf{v} = \frac{N_d e^2}{m\omega_o^2}\frac{d\mathbf{E}}{dt} = \frac{d\mathbf{P}}{dt}.$$

    This flux is effectively an AC current density carried by bound charges found in a dielectric medium. Even though, a DC current is not possible in a perfect dielectric containing only bound charges, evidently AC currents

Polarization current density

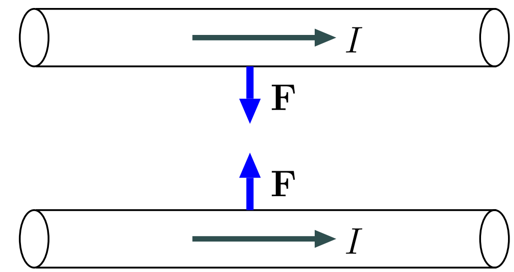
$$\mathbf{J}_p = \frac{d\mathbf{P}}{dt} \quad \text{Polarization current density}$$

    are possible — we call this type of AC current as **polarization current density**.

- In our studies of time-varying electromagnetic fields we will include the effects of polarization currents $\frac{d\mathbf{P}}{dt}$ along with the effects of conduction currents $\sigma\mathbf{E}$.

12 Magnetic force and fields and Ampere's law

Pairs of wires carrying currents I running in the same (opposite) direction are known to attract (repel) one another. In this lecture we will explain the mechanism — the phenomenon is a relativistic[1] consequence of electrostatic charge interactions, but it is more commonly described in terms of magnetic fields. This will be our introduction to magnetic field effects in this course.



"Things should be made as simple as possible – but no simpler."
— ***Albert Einstein***

---

[1]**Brief summary of *special* relativity:** Observations indicate that light (EM) waves *can* be "counted" like particles and yet *travel* at one and the same speed $c = 3 \times 10^8$ m/s in *all* reference frames in relative motion. As first recognized by Albert Einstein, these facts preclude the possibility that a *particle* velocity u could appear as

$$u' = u - v \quad \text{(Newtonian)}$$

to an observer approaching the particle with a velocity v; instead, u must transform to the observer's frame as

$$u' = \frac{u - v}{1 - \frac{uv}{c^2}}, \quad \text{(relativistic)}$$

so that if $u = c$, then $u' = c$ also. This "relativistic" velocity transformation in turn requires that positions x and times t of physical events transform (between the frames) as

$$x' = \gamma(x - vt) \text{ and } t' = \gamma(t - \frac{v}{c^2}x), \quad \text{(relativistic)}$$

where $\gamma \equiv \frac{1}{\sqrt{1-v^2/c^2}}$, rather than as

$$x' = x - vt \text{ and } t' = t, \quad \text{(Newtonian)}$$

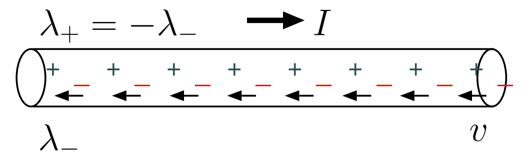
so that $\frac{dx}{dt} = u$ and $\frac{dx'}{dt'} = u'$ are related by the relativistic formula for u' given above.

Relativistic transformations imply a number of "counter-intuitive" effects ordinarily not noticed unless $|v|$ is very close to c. One of them is *Lorentz contraction*, implied by $dx = dx'/\gamma$ at a fixed t: since $\gamma > 1$, $dx < dx'$, and moving objects having velocities v appear shorter then they are when viewed from other reference frames where v is determined. A second one is *time dilation*, implied by $dt' = dt/\gamma$ at a fixed x': since $\gamma > 1$, $dt' < dt$, and moving clocks having velocities v and fixed x' run slower than clocks in other reference frames where v is determined. Consider taking PHYS 325 to learn more about special relativity.

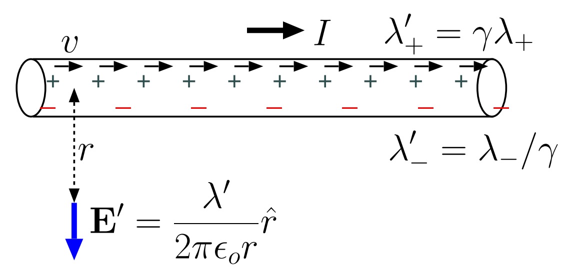
- Consider a current carrying stationary wire in the lab frame:
  - the wire has a stationary lattice of positive ions,
  - electrons are moving to the left through the lattice with an average speed v, and
  - a current $I > 0$ is flowing to the right as shown in the figure.
    - If the wire is electrically uncharged — which will be true if electron and ion charge densities in the wire, $\lambda_- < 0$ and $\lambda_+ > 0$, respectively, have equal magnitudes — then the wire will produce no electrostatic field $\mathbf{E}$, and any stationary charge q placed near the wire will not be subject to any force[2].
    - The current carried by the wire is $I = v|\lambda_-| = v\lambda_+$ in terms of the magnitudes of electron velocity and charge density.

- An uncharged wire in the lab frame appears as "charged" in the reference frame of the electrons carrying the current:
  - this is a *relativistic effect* due to "Lorentz contraction" of the distances between the charges in the wire.

(a) Neutral wire carrying current I in the "lab frame":



(b) In the "electron frame" the wire appears positively charged:



$$\lambda' \approx \lambda_+ \frac{v^2}{c^2} = \frac{Iv}{c^2} = Iv\mu_o\epsilon_o$$

[2]This is true for zero-resistivity wires. Current carrying wires with *finite* resistivity will however support *surface* charge densities with axial gradients to produce the static field within the wire needed to drive the current — e.g., in *Am. J. Phys.*: Jefimenko, **30**, 19 (1962); Parker, **38**, 720 (1970); Preyer, **68**, 1002 (2000).

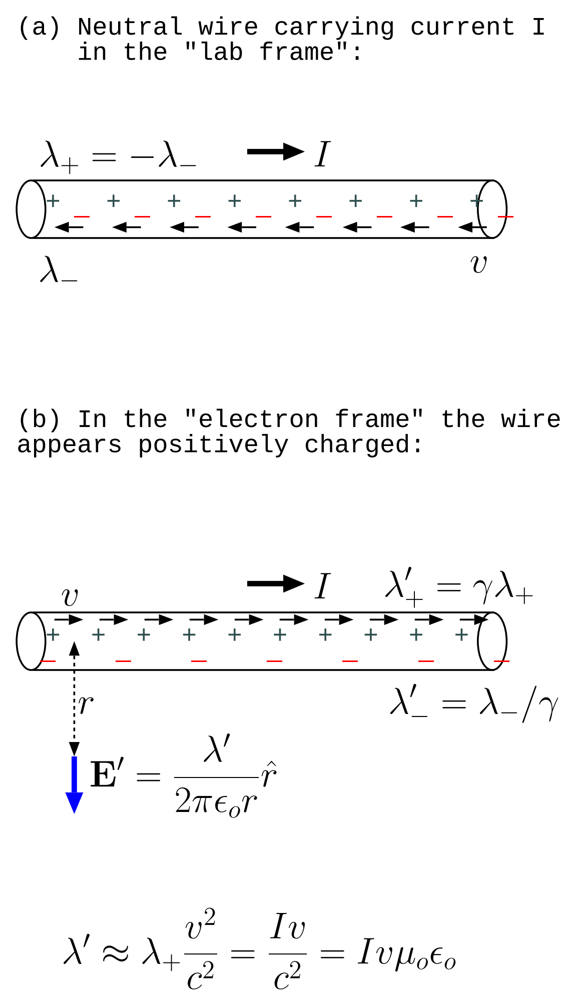
– In the electron frame the wire is found to have a positive charge density λ', and thus it has a radial electrostatic field

$$\mathbf{E}' = \frac{\lambda'}{2\pi\epsilon_o r}\hat{r}$$

implying an electrostatic force $\mathbf{F}' = q\mathbf{E}'$ on a stationary charge q.

– Relativistic calculations[3] show that

$$\lambda' \approx \lambda_+ \frac{v^2}{c^2} = (\frac{I}{v})\frac{v^2}{c^2} = Iv\epsilon_o\mu_o$$



---

[3](i) Electron spacings dx' measured in the electron reference frame will appear as

$$dx = \sqrt{1 - \frac{v^2}{c^2}}dx'$$

in the lab frame because of Lorentz contraction. Charge density of the electrons in the lab frame,

$$\lambda_- = \frac{\lambda'_-}{\sqrt{1 - v^2/c^2}},$$

is therefore greater in magnitude than the electron charge density λ'_- in the electron frame. Furthermore, $\lambda_- = -\lambda_+$ in order to maintain a charge neutral wire in the lab frame. (ii) Once again because of Lorentz contraction, the charge density of positive ions will appear in the electron frame as

$$\lambda'_+ = \frac{\lambda_+}{\sqrt{1 - v^2/c^2}}.$$

(iii) Thus, the total charge density of the wire in the electron frame is

$$\lambda' = \lambda'_+ + \lambda'_- = \frac{\lambda_+}{\sqrt{1 - v^2/c^2}} + \lambda_-\sqrt{1 - v^2/c^2} = \frac{\lambda_+}{\sqrt{1 - v^2/c^2}} - \lambda_+\sqrt{1 - v^2/c^2} = \frac{\lambda_+ v^2/c^2}{\sqrt{1 - v^2/c^2}} \approx \frac{\lambda_+ v^2}{c^2},$$

a *positive* density for non-zero $|v| \ll c$. (e.g., articles in *Am. J. Phys.*: Webster, **29**, 262, 1961; Matzek and Russel, **36**, 905, 1968; Arista and Lopez, **43**, 525, 1975; Zapolsky, **56**, 1137, 1988).

and force $\mathbf{F}' = q\mathbf{E}'$ can be transformed back to the lab frame, where q appears to be moving with velocity $\mathbf{v}$, as (with no approximation[4])

$$\mathbf{F} = q\mathbf{v} \times \frac{\mu_o I}{2\pi r}\hat{\phi},$$

where $\hat{\phi}$ is the unit vector in the direction given by the *right-hand-rule* (see margin) and $\mu_o = 4\pi \times 10^{-7}$ H/m is *permeability* of free space.

- We find it convenient to define

$$\mathbf{B} \equiv \frac{\mu_o I}{2\pi r}\hat{\phi}$$

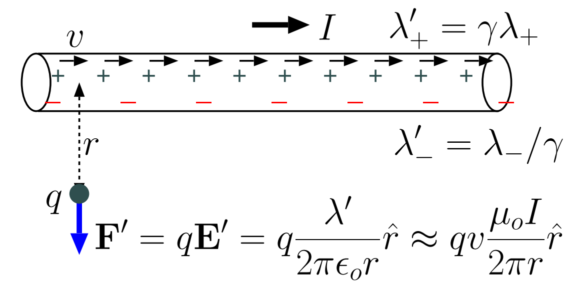
  to be the "magnetic flux density" of current filament I at a distance r, and attribute the force

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B}$$

  on the moving charge q to the magnetic field $\mathbf{B}$ produced by current I (rather than to the electrostatic field of the wire seen by q in its own reference frame).

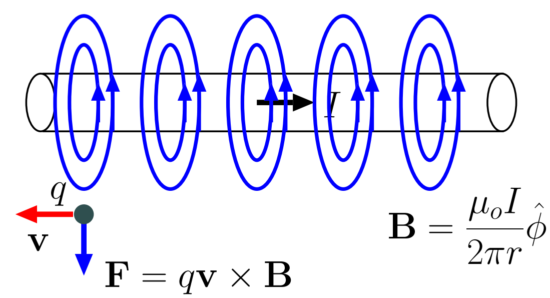
While we assumed q to be stationary in the reference frame of the electrons in the above discussion (for the sake of simplicity), the results obtained above are found to be valid for all particle velocities $\mathbf{v}$ measured in the lab frame.

(a) In the "electron frame" the wire appears positively charged and repelsa test charge q with force F′=qE′



$$\lambda' \approx \lambda_+\frac{v^2}{c^2} = \frac{Iv}{c^2} = Iv\mu_o\epsilon_o$$

(b) In the lab frame force F~F′ of moving charge q is attributed to magnetic field B produced by current I and velocity v of the charge in F=qvXB combination.



Magnetic field B curls around current I in a right handed direction designated by azimuthal unit vector $\hat{\phi}$

Magnetic field lines close upon themselves unlike electric field lines which start and stop on point charges.

Right hand rule: **point your right thumb in the direction of current flow; your fingers will point in direction $\hat{\phi}$.**

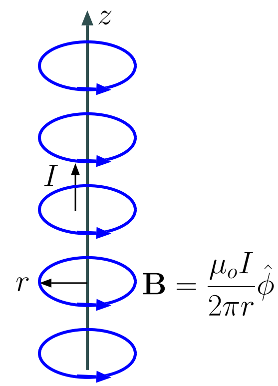
[4]We also get the same result using the approximation $\mathbf{F} = \mathbf{F}'$ that can be justified when $|v| \ll c$, which is typically true by a large margin for electron speeds in current carrying conducting metals — see HW.

Also, if there are multiple current filaments I_n, each generating its own field $\mathbf{B}_n$, force $\mathbf{F}$ on q can be calculated using a superposition method as with electrostatic fields.

Magnetic field $\mathbf{B}$ of the infinite current filament I obtained above can also be obtained by superposing the magnetic field increments

$$d\mathbf{B} \equiv \frac{\mu_o I d\mathbf{l} \times \hat{r}}{4\pi r^2} \quad \text{(Biot-Savart law)}$$

of directed current increments $Id\mathbf{l}$, where $\mathbf{r} = r\hat{r}$ is a position vector extending from the location of the current increment to the field position where $d\mathbf{B}$ is being specified — this formula, known as Biot-Savart law, is only valid when used in terms of infinitesimal segments $Id\mathbf{l}$ of time-invarying current loops.



- Magnetic field $\mathbf{B}$ of the infinite line current I "derived" above satisfies a circulation relation

  $$\oint_C \mathbf{B} \cdot d\mathbf{l} = \mu_o I_C,$$

  with $I_C = I$.

  This integral for the circulation of static magnetic field $\mathbf{B}$ is found to be valid (experimentally) for all closed circulation paths C, and is known as **Ampere's law** (for static magnetic fields). In Ampere's law

  - I_C stands for the net sum of all filament currents I_n crossing any surface S bounded by path C,

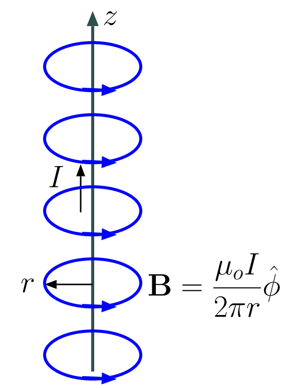
    - flowing in the direction given by the "right-hand-rule":

when the right thumb is pointed in the direction of $d\mathbf{l}$ along path C, the direction of filament current I_n is specified as the direction of the fingers of your right hand through surface S bounded by contour C.

  - Filament currents not crossing S — i.e., current filaments not "linked" to path C — should not be included on the right hand side of Ampere's law.

- Ampere's law can also be expressed as

$$\oint_C \mathbf{H} \cdot d\mathbf{l} = \int_S \mathbf{J} \cdot d\mathbf{S},$$



where

  – we have defined

$$\mathbf{H} \equiv \mu_o^{-1}\mathbf{B}$$

  for the sake of convenience, and

  – $\mathbf{J}$ is the volumetric current density measured in A/m^2 units (e.g., $\sigma\mathbf{E}$ in a conducting region as discussed in last lecture) having a total flux

$$I_C = \int_S \mathbf{J} \cdot d\mathbf{S}$$

  across any surface S bounded by a path C,

    - with $d\mathbf{S}$ pointing across S in the direction compatible with right-hand-rule as in *Stoke's theorem* (recall Lecture 6).

- Stoke's theorem re-stated for a vector field $\mathbf{H}$ as

$$\oint_C \mathbf{H} \cdot d\mathbf{l} = \int_S \nabla \times \mathbf{H} \cdot d\mathbf{S}$$

implies that the differential form of Ampere's law should be

$$\nabla \times \mathbf{H} = \mathbf{J}.$$

This differential relation is accompanied by

$$\nabla \cdot \mathbf{B} = 0,$$

satisfied by static magnetic field of the line current *as well as* by any other magnetic field — static as well as non-static, as determined experimentally and described in more detail later on.

- Current density vector field $\mathbf{J}$ invoked above in Ampere's law expressions, measured nominally in units of A/m^2, can also be adjusted to describe the distributions of surface or line currents in 3D space.

  – For example,

$$\mathbf{J}(x, y, z) = \mathbf{J}_s(y, z)\delta(x - x_o)$$

  can be regarded as **volumetric current density** representation of a **surface current density** $\mathbf{J}_s(x, y)$ measured in A/m units flowing on $x = x_o$ surface.

Laws of magnetostatics:

$$\begin{aligned} \nabla \times \mathbf{H} &= \mathbf{J} \\ \nabla \cdot \mathbf{B} &= 0 \end{aligned}$$

They also apply "quasi-statically" over a region of dimension L when a time-varying field source $\mathbf{J}(\mathbf{r}, t)$ has a *time-constant* τ much longer than the propagation time delay L/c of field variations across the region (c is the speed of light).

In magneto-quasistatics (MQS) $\mathbf{B}(\mathbf{r},t) = \mu_o \mathbf{H}(\mathbf{r}, t)$ will be accompanied by a slowly varying electric field $\mathbf{E}(\mathbf{r}, t)$ (derived from Faraday's law discussed in Lecture 14).

– Likewise,

$$\mathbf{J}(x,y,z) = \hat{z} I(z)\delta(x-x_o)\delta(y-y_o)$$

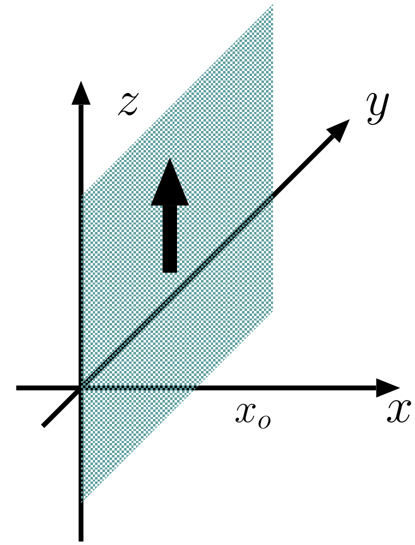
representats a **line current** $I(z)$ measured in A units flowing in z-direction along a filament defined by the intersections of $x = x_o$ and $y = y_o$ surfaces.

– As a most extreme case,

$$\mathbf{J}(x,y,z,t) = Q\mathbf{v}\delta(x-x_o)\delta(y-y_o)\delta(z-z_o)$$

represents the *time-varying* current density of a point charge Q at coordinates $(x,y,z) = (x_o(t), y_o(t), z_o(t))$ moving with velocity $\mathbf{v} = (\dot{x}_o(t), \dot{y}_o(t), \dot{z}_o(t))$.

$\mathbf{J} = \hat{z}\ y\ \mathrm{rect}(y-0.5)\ \delta(x)$



Example 1: Consider a surface current density of

$$\mathbf{J}_s = \hat{z} y\, \mathrm{rect}(y-0.5)\ \mathrm{A/m}$$

flowing on $x = 0$ plane (as shown in the margin). What is the total current I flowing on the same plane measured in A units?

Solution: To go from a surface current density $\mathbf{J}_s$ in A/m to a total current I in A, we need to perform an appropriate integration operation on the surface were $\mathbf{J}_s$ is defined. For the specified $\mathbf{J}_s$ in this problem we find that

$$I = \int_{y=-\infty}^{\infty} \mathbf{J}_s \cdot \hat{z} dy = \int_{y=0}^{1} y\, dy = \frac{y^2}{2}\Big|_0^1 = \frac{1}{2}\ \mathrm{A}.$$

13 Current sheet, solenoid, vector potential and current loops

In the following examples we will calculate the magnetic fields $\mathbf{B} = \mu_o\mathbf{H}$ established by some simple current configurations by using the integral form of static Ampere's law.

Example 1: Consider a uniform **surface current density** $\mathbf{J}_s = J_s\hat{z}$ A/m flowing on $x = 0$ plane (see figure in the margin) — the current sheet extends infinitely in y and z directions. Determine $\mathbf{B}$ and $\mathbf{H}$.

Solution: Since the current sheet extends infinitely in y and z directions we expect $\mathbf{B}$ to depend only on coordinate x. Also, the field should be the superposition of the fields of an infinite number of current filaments, which suggests, by right-hand-rule, $\mathbf{B} = \hat{y}B(x)$, where $B(x)$ is an odd function of x. To determine $B(x)$, such that $B(-x) = -B(x)$, we apply Ampere's law by computing the circulation of $\mathbf{B}$ around the rectangular path C shown in the figure in the margin. We expand

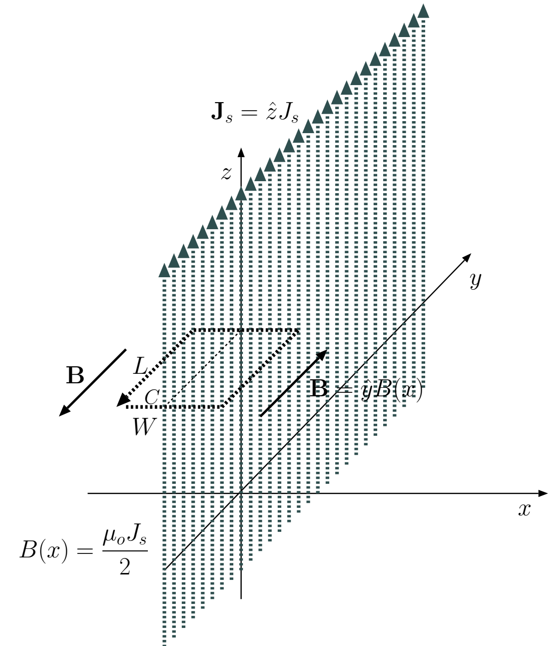
$$\oint_C \mathbf{B}\cdot d\mathbf{l} = \mu_o I_C$$

as

$$B(x)L + 0 - B(-x)L + 0 = \mu_o J_s L,$$

from which we obtain

$$B(x) = \frac{\mu_o J_s}{2} \;\Rightarrow\; \mathbf{B} = \hat{y}\frac{\mu_o J_s}{2}\mathrm{sgn}(x) \text{ and } \mathbf{H} = \hat{y}\frac{J_s}{2}\mathrm{sgn}(x).$$



As shown in Example 1 magnetic field of a current sheet is independent of distance $|x|$ from the current sheet. Also H changes discontinuously across the current sheet by an amount J_s.

Example 2: Consider a slab of thickness W over $-\frac{W}{2} < x < \frac{W}{2}$ which extends infinitely in y and z directions and conducts a uniform current density of $\mathbf{J} = \hat{z}J_o$ A/m^2. Determine $\mathbf{H}$ if the current density is zero outside the slab.

Solution: Given the geometric similarities between this problem and Example 1, we postulate that $\mathbf{B} = \hat{y}B(x)$, where $B(x)$ is an odd function of x, that is $B(-x) = -B(x)$. To determine $B(x)$ we apply Ampere's law by computing the circulation of $\mathbf{B}$ around the rectangular path C shown in the figure in the margin. For $x < \frac{W}{2}$, we expand

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \mu_o I_C$$

as

$$B(x)L + 0 - B(-x)L + 0 = \mu_o J_o 2xL \;\;\Rightarrow\;\; B(x) = \mu_o J_o x.$$

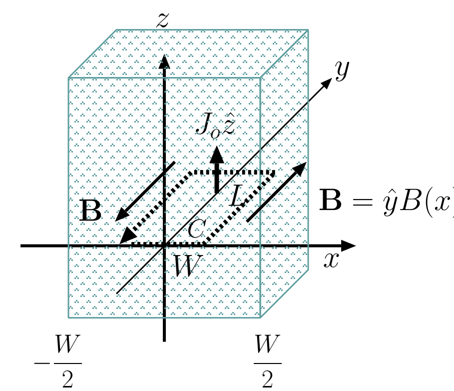
For $x \geq \frac{W}{2}$, the expansion gives

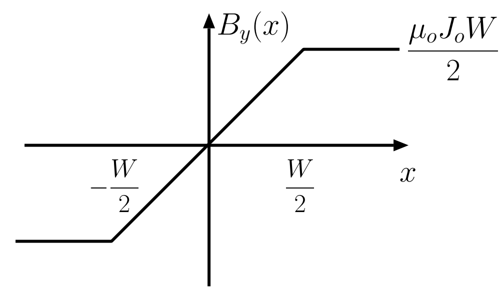
$$B(x)L + 0 - B(-x)L + 0 = \mu_o J_o W L \;\;\Rightarrow\;\; B(x) = \mu_o J_o \frac{W}{2}.$$

Hence, we find that

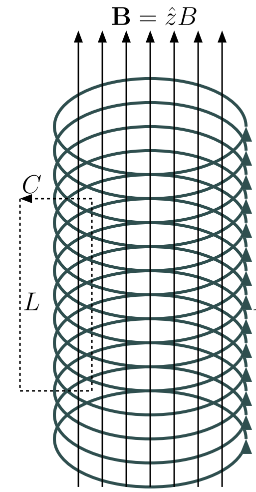
$$\mathbf{H} = \begin{cases} \hat{y}J_o x, & |x| < \frac{W}{2} \\ \hat{y}J_o \frac{W}{2}\,\mathrm{sgn}(x), & \text{otherwise.} \end{cases}$$

Note that the solution plotted in the margin shows no discontinuity at $x = \pm\frac{W}{2}$ or elsewhere.





The figure in the margin depicts a finite section of an infinite solenoid. A solenoid can be constructed in practice by winding a long wire into a multi loop coil as depicted. A solenoid with its loop carrying a current I in $\hat{\phi}$ direction (as shown), produces effectively a surface current density of $\mathbf{J}_s = IN\hat{\phi}$ A/m, where N is the number density (1/m) of current loops in the solenoid. In Example 3 we compute the magnetic field of the infinite solenoid using Ampere's law.



Infinite solenoid with N loops per unit length carrying I amps per loop

$B = \mu_o IN$

Example 3: An infinite **solenoid** having N loops per unit length is stacked in z-direction, each loop carrying a current of I A in counter-clockwise direction when viewed from the top (see margin). Determine $\mathbf{H}$.

Solution: Assuming that $\mathbf{B} = 0$ outside the solenoid, and also $\mathbf{B}$ is independent of z within the solenoid, we find that Ampere's law indicates for the circulation C shown in the margin

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \mu_o I_C \quad \Rightarrow \quad LB = \mu_o INL.$$

This leads to

$$B = \mu_o IN \text{ and } \mathbf{H} = \hat{z}IN$$

for the field within the solenoid.

The assumption of zero magnetic flux density $\mathbf{B} = 0$ for the exterior region is justified because:

(a) if the exterior field is non-zero, then it must be independent of x and y (follows from Ampere's law applied to any exterior path C with $I_C = 0$), and

(b) the finite interior flux $\Psi = \mu_o IN\pi a^2$ can only be matched with the flux of the infinitely extended exterior region when the *constant* exterior flux density (because of (a)) is vanishingly small.

- **Static electric fields:** *Curl-free* and are governed by

$$\nabla \times \mathbf{E} = 0, \;\; \nabla \cdot \mathbf{D} = \rho \;\text{ where }\; \mathbf{D} = \epsilon \mathbf{E}$$

  with $\epsilon = \epsilon_r \epsilon_o$.

- **Static magnetic fields:** *Divergence-free* and are governed by

$$\nabla \cdot \mathbf{B} = 0, \;\; \nabla \times \mathbf{H} = \mathbf{J} \;\text{ where }\; \mathbf{B} = \mu \mathbf{H}$$

  with $\mu = \mu_r \mu_o$ — relative *permeabilities* μ_r other than unity (for free space) will be explained later on.

Mathematically, we can generate a **divergence-free** vector field $\mathbf{B}(x, y, z)$ as

$$\mathbf{B} = \nabla \times \mathbf{A}$$

by taking the curl of any vector field $\mathbf{A} = \mathbf{A}(x, y, z)$ (just like we can generate a curl-free $\mathbf{E}$ by taking the gradient of any scalar field $-V(x, y, z)$).

Verification: Notice that

$$\begin{aligned}\nabla \cdot \nabla \times \mathbf{A} &= \frac{\partial}{\partial x}(\nabla \times \mathbf{A})_x + \frac{\partial}{\partial y}(\nabla \times \mathbf{A})_y + \frac{\partial}{\partial z}(\nabla \times \mathbf{A})_z = \begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix} \\ &= \frac{\partial}{\partial x}(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}) - \frac{\partial}{\partial y}(\frac{\partial A_z}{\partial x} - \frac{\partial A_x}{\partial z}) + \frac{\partial}{\partial z}(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}) = 0.\end{aligned}$$

- If $\mathbf{B} = \nabla \times \mathbf{A}$ represents a magnetostatic field, then $\mathbf{A}$ is called magnetostatic potential or **vector potential**.

– Vector potential $\mathbf{A}$ can be used in magnetostatics in similar ways to how electrostatic potential V is used in electrostatics.

  ◦ In electrostatics we can assign $V = 0$ *to any point in space* that is convenient in a given problem.

  ◦ In magnetostatics we can assign $\nabla \cdot \mathbf{A}$ *to any scalar* that is convenient in a given problem.

– For example, if we make the assignment[1]

$$\nabla \cdot \mathbf{A} = 0,$$

then we find that

$$\nabla \times \mathbf{B} = \nabla \times \nabla \times \mathbf{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} = -\nabla^2 \mathbf{A}.$$

This is a *nice and convenient* outcome, because, when combined with

$$\nabla \times \mathbf{H} = \mathbf{J} \quad \Rightarrow \quad \nabla \times \mathbf{B} = \mu_o \mathbf{J},$$

it produces

$$\nabla^2 \mathbf{A} = -\mu_o \mathbf{J},$$

which is the magnetostatic version of Poisson's equation

$$\nabla^2 V = -\frac{\rho}{\epsilon_o}.$$

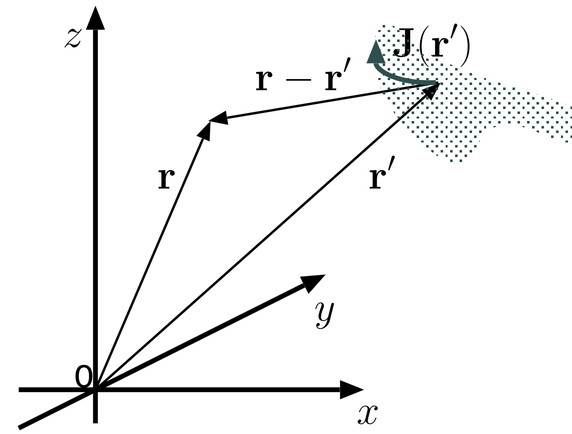
[1]With this assignment — known as *Coulomb gauge* — $\mathbf{A}$ acquires the physical meaning of "potential momentum per unit charge", just as scalar potential V is "potential enegy per unit charge" (see Konopinski, *Am. J. Phys.*, *46*, 499, 1978).

– In analogy with solution

$$V(\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{4\pi\epsilon_o|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}'$$

of Poisson's equation, it has a solution

$$\mathbf{A}(\mathbf{r}) = \int \frac{\mu_o \mathbf{J}(\mathbf{r}')}{4\pi|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}'.$$

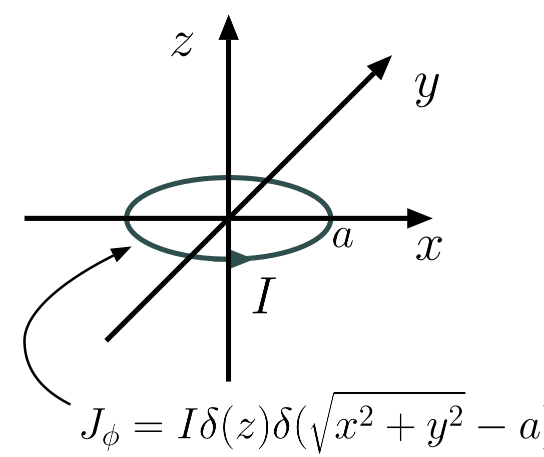


Given *any* static[2] current density $\mathbf{J}(\mathbf{r})$, the above equation can be used to obtain the corresponding vector potential $\mathbf{A}$ that simultaneously satisfies

$$\nabla\cdot\mathbf{A} = 0 \text{ and } \nabla\times\mathbf{A} = \mathbf{B}.$$

Once $\mathbf{A}$ is available, obtaining $\mathbf{B} = \nabla\times\mathbf{A}$ is then just a matter of taking a curl.

- Magnetic flux density $\mathbf{B}$ of a *single* current loop I can be calculated after determining its vector potential as follows:



– For a loop of radius a on $z=0$ plane, we can express the corresponding current density as

$$\mathbf{J}(\mathbf{r}') = I\delta(z')\delta(\sqrt{x'^2+y'^2}-a)\frac{(-y',x',0)}{\sqrt{x'^2+y'^2}}$$

where the ratio on the right is the unit vector $\hat{\phi}'$.

– Inserting this into the general solution for vector potential, and performing the integration over z', we obtain

[2]Also, in quasi-statics we use $\mathbf{J}(\mathbf{r}',t)$ to obtain $\mathbf{A}(\mathbf{r},t)$ and $\mathbf{B} = \nabla\times\mathbf{A}$ over regions small compared to $\lambda = c/f$, with f the highest frequency in $\mathbf{J}(\mathbf{r}',t)$.

$$
\begin{aligned}
\mathbf{A}(\mathbf{r}) &= \frac{\mu_o I}{4\pi}\int \delta(\sqrt{x'^2+y'^2}-a)\frac{(-y',x',0)}{\sqrt{(x-x')^2+(y-y')^2+z^2}\sqrt{x'^2+y'^2}}dx'dy' \\
&= \frac{\mu_o I}{4\pi}\int \delta(r'-a)\frac{(-y',x',0)}{\sqrt{(x-x')^2+(y-y')^2+z^2}r'}r'dr'd\phi' \\
&= \frac{\mu_o I}{4\pi}\int_{-\pi}^{\pi} \frac{(-a\sin\phi',a\cos\phi',0)}{\sqrt{(x-a\cos\phi')^2+(y-a\sin\phi')^2+z^2}}d\phi' \equiv \hat{x}A_x(\mathbf{r})+\hat{y}A_y(\mathbf{r}).
\end{aligned}
$$

– Given that $A_z=0$, it can be shown that $\mathbf{B}=\nabla\times\mathbf{A}$ leads to

$$
B_x=-\frac{\partial A_y}{\partial z},\quad B_y=\frac{\partial A_x}{\partial z},\quad B_z=\frac{\partial A_y}{\partial x}-\frac{\partial A_x}{\partial y}.
$$

– From the expected azimuthal symmetry of $\mathbf{B}$ about the z-axis, it is sufficient to evaluate these on, say, $y=0$ plane — after some algebra, and dropping the primes, we find, on $y=0$ plane,

$$
B_x=\frac{\mu_o aI}{4\pi}\int_{-\pi}^{\pi}\frac{z\cos\phi}{(x^2+a^2+z^2-2ax\cos\phi)^{3/2}}d\phi,
$$

$$
B_y=\frac{\mu_o aI}{4\pi}\int_{-\pi}^{\pi}\frac{z\sin\phi}{(x^2+a^2+z^2-2ax\cos\phi)^{3/2}}d\phi,
$$

and

$$
B_z=\frac{\mu_o aI}{4\pi}\int_{-\pi}^{\pi}\frac{a-x\cos\phi}{(x^2+a^2+z^2-2ax\cos\phi)^{3/2}}d\phi.
$$

– We note that $B_y=0$ since the B_y integrand above is odd in ϕ and the integration limits are centered about the origin. Hence, the field on $y=0$ plane is given as

$$
\mathbf{B}=\hat{x}B_x+\hat{z}B_z
$$

with B_x and B_z defined above.

– There are no closed form expressions for the B_x and B_z integrals above for an arbitrary (x,z).

- However, it can be easily seen that if $x = 0$ (i.e., along the z-axis), $B_x = 0$ (as symmetry would dictate) and

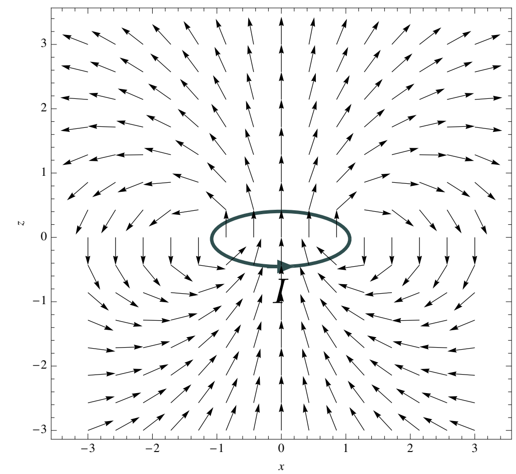
$$B_z = \frac{\mu_o a I}{4\pi} \int_{-\pi}^{\pi} \frac{a}{(a^2 + z^2)^{3/2}} d\phi = \frac{\mu_o I a^2}{2(a^2 + z^2)^{3/2}}.$$

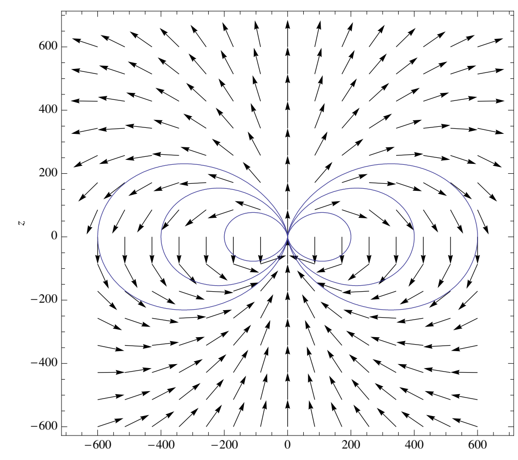
  For $|z| \gg a$,

$$B_z \approx \frac{\mu_o I a^2}{2|z|^3},$$

  which is positive and varies with the inverse third power of distance $|z|$.

– Also, B_x and B_z integrals can be performed numerically. Figure in the margin depicts the pattern of $\hat{B}$ on $y = 0$ plane for a loop of radius $a = 1$ computed using *Mathematica*.



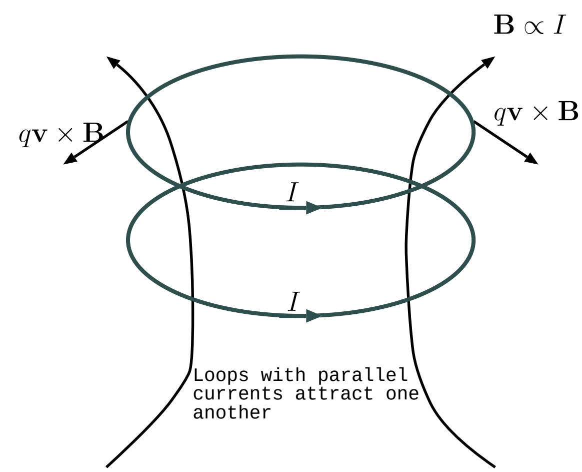


- Note that circulation $\oint_C \mathbf{B} \cdot d\mathbf{l}$ around each closed field line ("linking" the current loop) equals a fixed value of $\mu_o I$ — this dictates that the average field strength $|\mathbf{B}|$ of a current loop is stronger on shorter field lines closer to the current loop than on longer field lines linking the loop further out. As a result $|\mathbf{B}|$ can be shown to vary as r^{-3} for large r.

- It can be shown that the equations for magnetic *field lines* of a current loop on, say, $y = 0$ plane, can be expressed as

$$r = L \sin^2 \theta$$

  in terms of **radial distance** r from the origin and **zenith angle** θ measured from the z axis. Clearly, parameter L in this formula is the radial distance of the field line on $\theta = 90^o$ plane, and the field line formula is accurate only for $r \gg a$. The Earth's magnetic field had such a **magnetic dipole** topology as shown.



- Lorentz force due to the magnetic fields of a pair of current loops — also known as *magnetic dipoles* — turns out to be "attractive" when the current directions agree (see margin). Bar magnets carrying "equivalent" current loops of atomic origins interact with one another in exactly the same way — i.e., as governed by the second term of Lorentz force.

14 Faraday's law and induced emf

Michael Faraday discovered (in 1831, less than 200 years ago) that a *changing* current in a wire loop *induces* current flows in nearby wires — today we describe this phenomenon as **electromagnetic induction**: the current change in the first loop causes the magnetic field produced by the current to change, and magnetic field change, in turn, is said to *induce*[1] (i.e., produce) electric fields which drive the currents in nearby wires.

- While static electric fields produced by static charge distributions are unconditionally curl-free, *time-varying electric fields* produced by current distributions with time-varying components are found to have, in accordance with Faraday's observations, non-zero curls specified by

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \qquad \text{Faraday's law}$$

  at all positions $\mathbf{r}$ in all reference frames of measurement. Using *Stoke's theorem*, the same constraint can also be expressed in *integral form* as

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = -\int_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S} \qquad \text{Faraday's law}$$

  for all surfaces S bounded by all closed and *directed* paths C (with the direction of C, indicated by an arrow, and direction of vector $d\mathbf{S}$ related by right hand rule).

Definitions of E and B have not changed:

recall that

- **E is force per unit stationary charge**
- **B gives an additional force $\mathbf{v} \times \mathbf{B}$ per unit charge in motion with velocity v in the measurement frame.**

[1] Relativistic derivation of static **B** given in Lecture 12 can be extended to show that Coulomb interactions of charges in *time-varying* motions require a description in terms of time-varying **B** *and* **E** — see, e.g., *Am. J. Phys.*: Tessman, 34, 1048 (1966); Tessman and Finnel, **35**, 523 (1967); Kobe, **54**, 631 (1986). Thus, the *cause* of *induced* **E** is not really the time-varying **B**, but rather the time-varying current **J** that is also producing the variation in **B**.

- The right hand side of the integral form equation above includes the **flux of *rate of change of magnetic field* B** over surface S. If contour C bounding S is "fixed" (unchanging) in the measurement frame, then the equation can also be expressed as

$$\oint_C \mathbf{E}\cdot d\mathbf{l} = -\frac{d}{dt}\int_S \mathbf{B}\cdot d\mathbf{S},$$

where the right hand side is now expressed in terms of the *rate of change* of **magnetic flux**

$$\Psi \equiv \int_S \mathbf{B}\cdot d\mathbf{S}$$

linking contour C over any surface S bounded by C.

- This modification (the exchange of the order of integration and time derivative on the right side) would *not* be permissible if path C were moving within the measurement frame or being deformed in time — but in such cases we could still express Faraday's integral form equation with $-\frac{d\Psi}{dt}$ on the right side, provided that we also modify the left side as in

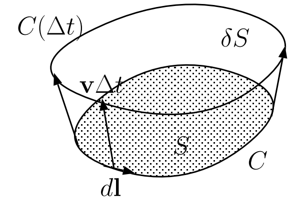
$$\oint_C (\mathbf{E}+\mathbf{v}\times\mathbf{B})\cdot d\mathbf{l} = -\frac{d}{dt}\int_S \mathbf{B}\cdot d\mathbf{S}$$

where $\mathbf{v}$ denotes the velocity of motion or deformation of path C.

  - This is equivalent to the original equation, since, as shown in the margin,

$$\int_S \frac{\partial \mathbf{B}}{\partial t}\cdot d\mathbf{S} = \frac{d}{dt}\int_S \mathbf{B}\cdot d\mathbf{S} + \oint_C \mathbf{v}\times\mathbf{B}\cdot d\mathbf{l}$$

  when C is changing continuously with velocities $\mathbf{v}$.



$\Psi(0) = \int_S \mathbf{B}(\mathbf{r},0)\cdot d\mathbf{S}$, and

$\Psi(\Delta t) = \int_S \mathbf{B}(\mathbf{r},\Delta t)\cdot d\mathbf{S} + \int_{\delta S} \mathbf{B}(\mathbf{r},\Delta t)\cdot d\mathbf{S}$.

Thus, $\dfrac{\Psi(\Delta t)-\Psi(0)}{\Delta t} =$

$$\int_S \frac{\mathbf{B}(\mathbf{r},\Delta t)-\mathbf{B}(\mathbf{r},0)}{\Delta t}\cdot d\mathbf{S} + \int_{\delta S}\mathbf{B}(\mathbf{r},\Delta t)\cdot\frac{d\mathbf{S}}{\Delta t}.$$

Hence in limit $\Delta t \to 0$

$$\frac{d\Psi}{dt} = \int_S \frac{\partial \mathbf{B}}{\partial t}\cdot d\mathbf{S} - \int_C \mathbf{v}\times\mathbf{B}\cdot d\mathbf{l},$$

since

$$\int_{\delta S}\mathbf{B}(\mathbf{r},\Delta t)\cdot\frac{d\mathbf{S}}{\Delta t} = \int_C \mathbf{B}(\mathbf{r},\Delta t)\cdot\frac{\Delta t\mathbf{v}\times d\mathbf{l}}{\Delta t}$$
$$= -\int_C (\mathbf{v}\times\mathbf{B})\cdot d\mathbf{l}$$

because

$\mathbf{B}\cdot(\mathbf{v}\times d\mathbf{l}) = d\mathbf{l}\cdot(\mathbf{B}\times\mathbf{v})$,

both representing the volume of a parallelepiped formed by the vectors $d\mathbf{l}$, $\mathbf{v}$, and $\mathbf{B}$.

Note that velocity $\mathbf{v}$ does not have to be constant around contour C.

- A **physical interpretation** of the final equation

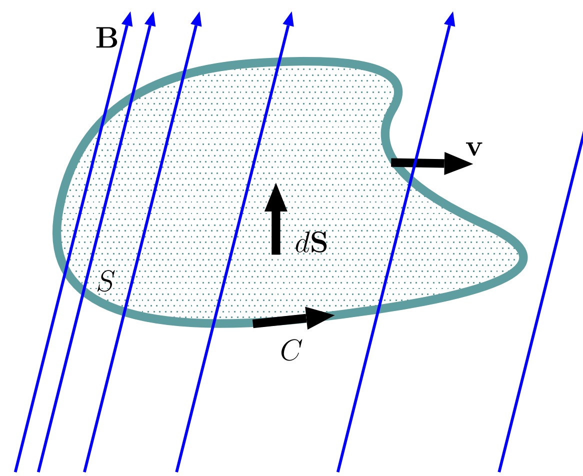
$$\oint_C (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l} = \underbrace{-\frac{d}{dt}\int_S \mathbf{B} \cdot d\mathbf{S}}_{\mathcal{E} = -\frac{d\Psi}{dt}} \quad \text{Integral form Faraday's Law}$$

$$\mathcal{E} = -\frac{d\Psi}{dt} \quad \text{Faraday's eqn.}$$

is as follows:

  - the **circulation integral** on the left is the "voltage drop" once around the *directed closed path* C, representing the **work done per unit charge** (by the Lorentz force $\propto \mathbf{E} + \mathbf{v} \times \mathbf{B}$) taken a full circle around C, which was denoted by Michael Faraday with a symbol $\mathcal{E}$ and called the **emf** (short for *electro-motive force,* which is a bad name since $\mathcal{E}$ is *work,* and not *force,* per unit charge) for the closed path, equaling the decay rate $-\frac{d\Psi}{dt}$ of its *linked* **magnetic flux** Ψ (due to all sources of magnetic flux density $\mathbf{B}$ in the region).
  - if/when path C is occupied by a conducting wire loop of some total conductance $G = \frac{1}{R}$, and a resistance $R = \frac{1}{G}$, a current $I = G\mathcal{E} = \frac{\mathcal{E}}{R}$ will flow around the loop in the circulation direction[2],

Magnetic field lines contributing to Ψ *form links* with path C (bounding S) like the links in an ordinary chain — hence, Ψ is said to be the *flux linking path* C.



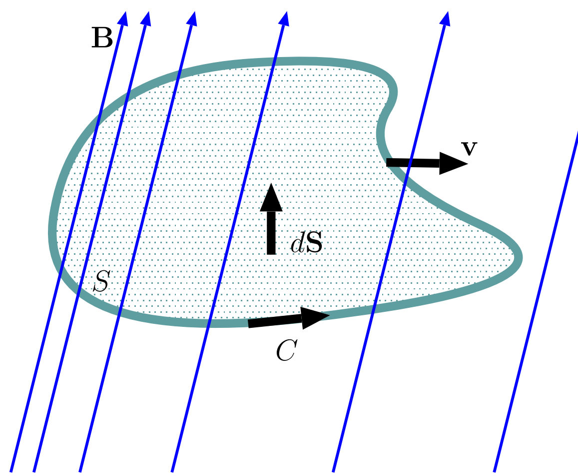
[2] $I = A\sigma|\mathbf{E} + \mathbf{v} \times \mathbf{B}|$ for a homogeneous wire loop with a conductivity σ and cross sectional area A. If the loop length is L, then the loop conductance is $G = \frac{A\sigma}{L}$ and therefore we find that $I = G\mathcal{E}$, as claimed, since $\mathcal{E} = \oint_C (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l} = |\mathbf{E} + \mathbf{v} \times \mathbf{B}|L$ around a homogeneous loop.

driven by the non-zero field $\mathbf{E}+\mathbf{v}\times\mathbf{B}$ within the wire accounting for the non-zero $\mathcal{E}$ if/when $-\frac{d\Psi}{dt}$ is non-zero.

– in equivalent circuit models of conducting wire loops, Faraday's equation, re-written as $RI = -\frac{d\Psi}{dt}$, is effectively Kirchfoff's voltage law (KVL) applied to the loop, with RI on the left denoting the (sum of all) *voltage drops* in the direction of C, while $-\frac{d\Psi}{dt}$ on the right denoting a *voltage rise* also in the direction of C.

  ◦ note that the emf $\mathcal{E}$ describes both the *voltage drop* RI and *voltage rise* $-\frac{d\Psi}{dt}$ appearing in the circuit model for the conducting wire loop since $\mathcal{E} = RI$ and $\mathcal{E} = -\frac{d\Psi}{dt}$ are both true.

– in modern parlance (since Maxwell) the term emf and its symbol $\mathcal{E}$ are used to refer to and denote sources of energy, e.g., battery voltages and magnetic flux rate $-\frac{d\Psi}{dt}$ that drive currents $I = \frac{\mathcal{E}}{R}$ around closed circuits[3].



- If path C is *fixed* in the measurement frame, then $\mathbf{v} = 0$, and KVL for such a stationary loop reads as

$$\oint_C \mathbf{E}\cdot d\mathbf{l} = -\frac{d\Psi}{dt};$$

– otherwise, that is if C is in motion, then

[3]see Saslow, *Am. J. Phys.*, **58**, 22 (2021), for a discussion of Maxwell's interpretation of emf and electrical energy production in batteries. Also see Scanlon et al., *Am. J. Phys.*, **37**, 689 (1969) for a discussion of $\mathcal{E} = \oint_C(\mathbf{E}+\mathbf{v}\times\mathbf{B})\cdot d\mathbf{l}$ vs $\mathcal{E} = -\frac{d\Psi}{dt}$.

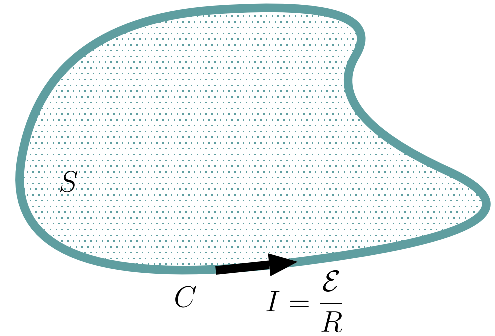
$$\oint_C (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l} = -\frac{d\Psi}{dt}$$

because in that case force per unit charge *advected* with path C will be $\mathbf{E} + \mathbf{v} \times \mathbf{B}$ according to Lorentz force (note: any additional velocity $\mathbf{v}_q$ of a moving charge *along* C does not contribute because $(\mathbf{v}_q \times \mathbf{B}) \cdot d\mathbf{l} = 0$ if $d\mathbf{l}$ and $\mathbf{v}_q$ are parallel).

– In either case, if C is a physical conducting path with a total resistance R, then the emf $-\frac{d\Psi}{dt}$ drives a current

$$I = \frac{-\frac{d\Psi}{dt}}{R}$$

around C in the circulation direction (determined by $d\mathbf{l}$ and $d\mathbf{S}$ directions used in accordance with the right-hand-rule).



Think of EMF as the sum of all the "voltage rises" around the loop traversed in the direction of loop current I that needs to match the total "voltage drop" RI around the same loop traversed in the same direction.

That way, KVL which states that

Sum of voltage rises = Sum of voltage drops,

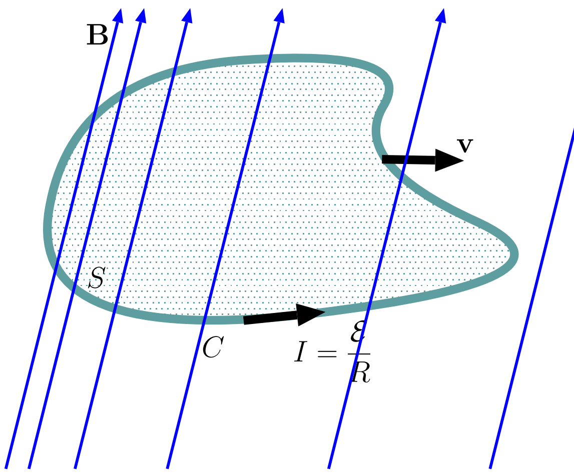
is fulfilled.

- The minus sign present in Faraday's equation, $\mathcal{E} = -\frac{d\Psi}{dt}$, assures that induced current I produces an induced magnetic field that *opposes* the flux change producing the emf — this fact is known as **Lenz's rule** and is in full accord with observations[4]./newpage

- According to Faraday's law it appears that magnetic flux variations $-\frac{d\Psi}{dt}$ can produce a non-zero emf independent of how the variations are produced — the possibilities are:

[4]Faraday's law not having the minus sign (or in conflict with Lenz's rule) would be non-physical, as it would lead to unbounded growth of induced currents and fields (by aiding rather than opposing the flux change producing the emf).

1. Fixed C, but time-varying $\mathbf{B}$,
2. $\mathbf{B}$ =const. (in space and time), but time-varying C (rotating or changing size),
3. An inhomogeneous static $\mathbf{B} = \mathbf{B}(\mathbf{r})$ in the measurement frame *and* C in motion.



- Note that even in the absence of any electric field $\mathbf{E}$ in the measurement frame, a non-zero emf

$$\oint_C (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l} = -\frac{d\Psi}{dt}$$

can exist because of the motion of C through an *inhomogeneous* magnetic field (if the field is homogeneous then $\frac{d\Psi}{dt}$ will be zero, implying zero $\mathcal{E}$), which will of course appear as an emf

$$\oint_C \mathbf{E}' \cdot d\mathbf{l}' = -\frac{d\Psi'}{dt'}$$

for a second observer moving with C who sees a time varying electric field $\mathbf{E}' = \mathbf{v} \times \mathbf{B}$ in her own frame (in addition to the inhomogeneous but constant magnetic field $\mathbf{B}$ of the first frame appearing as a time-varying magnetic field $\mathbf{B}'$)[5].

[5]See Scanlon et. al., *Am. J. Phys.*, **37**, 698 (1969), for a discussion of $I' = \frac{\mathcal{E}'}{R}$ for rigid C with resistance R observed from different reference frames.

– Thus, having non-zero electric field circulations

$$\oint_C \mathbf{E}' \cdot d\mathbf{l}'$$

under time-varying magnetic field conditions appears to be quite *comprehensible* after all!

– Magnetic fields $\mathbf{B}$ in one frame will morph into electric fields $\mathbf{E}'$ in other frames because of (near) invariance of Lorentz force between reference frames.

– Moreover a morphed $\mathbf{E}'$ can even be non-conservative — i.e., non curl-free — when $\mathbf{B}$ is inhomogeneous in space (or time) as we have just seen.

Example 1: If

$$\mathbf{B} = B_o e^{-t/\tau} \hat{z},$$

what is the emf $\mathcal{E}$ taken over a stationary circular loop C of radius $r = 10$ m on $z = 0$ plane in counter-clockwise direction (looking down on $z = 0$ plane)? What is current I if the loop resistance is R?

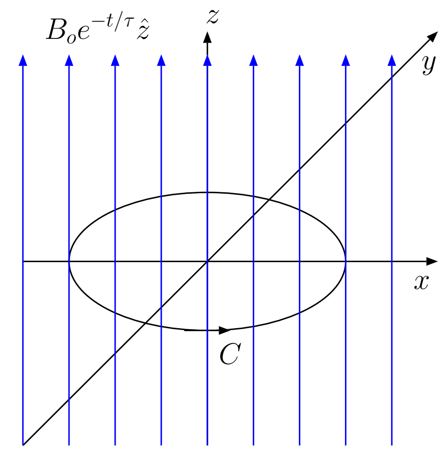
Solution: Since counter-clockwise circulation is requested we take $d\mathbf{S}$ pointing in $\hat{z}$ direction to be consistent with the right hand rule. We then have

$$\Psi = \int_S \mathbf{B} \cdot d\mathbf{S} = (B_o e^{-t/\tau} \hat{z}) \cdot (\pi 10^2 \hat{z}) = \pi 10^2 B_o e^{-t/\tau}$$

over the circular surface S. Thus, the emf

$$\mathcal{E} = -\frac{d\Psi}{dt} = \pi 10^2 \frac{B_o}{\tau} e^{-t/\tau}.$$

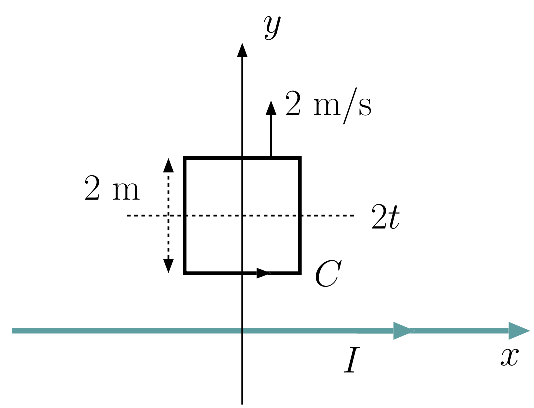
The loop current will be $I = \frac{\mathcal{E}}{R}$ in counter-clockwise direction of the computed circulation $\mathcal{E}$, which will be positive and counteract (i.e., strengthen) the weakening B_z.



Example 2: Consider the magnetic flux density

$$\mathbf{B} = \frac{\mu_o I}{2\pi r}\hat{\phi}$$

produced by current I flowing along the x axis. What is the emf $\mathcal{E}$ of a square loop C of area 4 m^2 moving on xy-plane with edges parallel to x- and y-axes, if its center is located at $y = 2t$ m as a function of time? Compute the emf $\mathcal{E}$ first as $-\frac{d\Psi}{dt}$ and then as $\oint_C (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l}$ to verify that the same values are obtained.



Solution: Given the described geometry, we have

$$\Psi(t) = \int_{-1}^{1} dx \int_{2t-1}^{2t+1} dy \frac{\mu_o I}{2\pi y} = \frac{\mu_o I}{\pi} \ln(\frac{2t+1}{2t-1}).$$

Thus, the emf $\mathcal{E}$ is

$$-\frac{d\Psi}{dt} = -\frac{\mu_o I}{\pi}(\frac{2t-1}{2t+1})\frac{\partial}{\partial t}(\frac{2t+1}{2t-1}) = \frac{\mu_o I}{\pi}\frac{4}{(2t+1)(2t-1)} = \frac{\mu_o I}{\pi(t^2 - \frac{1}{4})}.$$

Alternatively, since $\mathbf{v} = 2\hat{y}$ m/s, and $\mathbf{v} \times \mathbf{B} = 2\frac{\mu_o I}{2\pi r}\hat{x}$, we find, using $d\mathbf{l} = \pm\hat{x}dx$ and $\pm\hat{y}dy$ in turns,

$$\mathcal{E} = \oint_C (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l} = 2\frac{\mu_o I}{2\pi(2t-1)}2 - 2\frac{\mu_o I}{2\pi(2t+1)}2 = \frac{\mu_o I}{\pi(t^2 - \frac{1}{4})}$$

in consistency with the above result.

Example 3: A conducting loop of a radius $r = 0.1$ m (see figure in the margin) is being rotated about the x axis with frequency of $f = \frac{\omega}{2\pi} = 60$ Hz in a region with a DC magnetic field of $\mathbf{B} = 10\hat{z}$ T. Determine the induced current in the loop if the loop resistance is $12\,\Omega$.

Solution: The maximum value of the magnetic flux linking the loop should be

$$\Psi_o = \pi(0.1)^2 10 = 0.1\pi\,\text{Wb}.$$

The time-varying flux linking the rotating loop is therefore

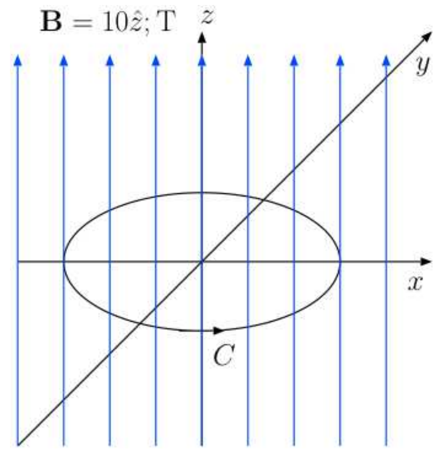
$$\Psi(t) = \Psi_o \cos(\omega t) = 0.1\pi \cos(120\pi t).$$

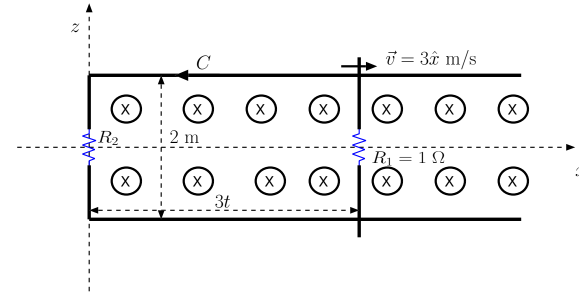
The corresponding emf is

$$\mathcal{E} = -\frac{d\Psi}{dt} = (120\pi)0.1\pi \sin(120\pi t).$$

Therefore, the induced current around the loop must be

$$I = \frac{\mathcal{E}}{R} = \frac{12\pi^2 \sin(120\pi t)}{12} = \pi^2 \sin(120\pi t)\,\text{A}.$$





Example 4: A conducting bar of resistance $R_1 = 1\,\Omega$ ohms is moved in the x-direction with a velocity $\mathbf{v} = 3\hat{x}$ m/s on a pair of perfect conducting ($R = 0$) stationary rails 2 m apart terminated with a load resistance R_2 at $x = 0$, all constituting a rectangular contour C to be taken counterclockwise. A constant magnetic field of $\mathbf{B} = 1\hat{y}$ T is linked throught contour C such that the flux $\Psi = -1 \times 2 \times 3t$ and the emf $\mathcal{E} = -d\Psi/dt = 6$ V. Hence, Faraday's law demands that

$$\oint_C (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l} = \int_b^t (\mathbf{E} + \mathbf{v} \times \mathbf{B})_1 \cdot d\mathbf{l} + \int_t^b (\mathbf{E})_2 \cdot d\mathbf{l} = 6$$

where the two integrals (with b and t referring to bottom and top rail contact points) correspond to voltage drops across resistors R_1 and R_2, respectively. But since

$$\int_b^t (\mathbf{v} \times \mathbf{B})_1 \cdot d\mathbf{l} = 3 \times 1 \times 2 = 6,$$

it follows that

$$\int_b^t (\mathbf{E})_1 \cdot d\mathbf{l} + \int_t^b (\mathbf{E})_2 \cdot d\mathbf{l} = 0 \;\;\Rightarrow\;\; E_{z1} - E_{z2} = 0 \;\;\Rightarrow\;\; E_{z2} = E_{z1},$$

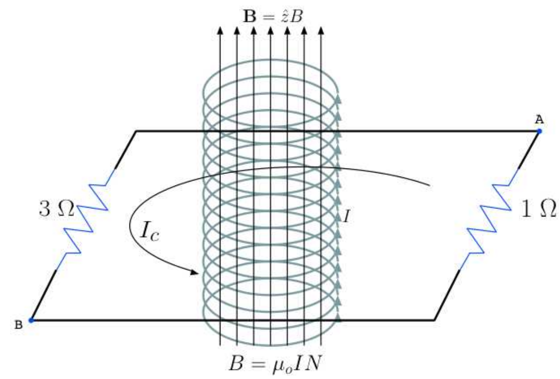
i.e., identical static fields within the moving and stationary bars across the perfect conducting rails. This may be a surprising claim/result — let's give two examples to illustrate how this happens:

1. Let $R_2 = 2\,\Omega$ ohms. Then $I = 6/3 = 2$ A. It follows that voltage drops $(\mathbf{E} + \mathbf{v} \times \mathbf{B})_1 \cdot 2\hat{z} = 2$ V across R_1 and $(\mathbf{E})_2 \cdot (-2\hat{z}) = 4$ V across R_2, yielding $E_{z1} = E_{z2} = -2$ V/m.

2. Let $R_2 = \infty$ — open ckt load to the moving conductor. Then $I = 6/\infty = 0$ A. It follows that $(\mathbf{E} + \mathbf{v} \times \mathbf{B})_1 \cdot 2\hat{z} = 0$ V across R_1 and $(\mathbf{E})_2 \cdot (-2\hat{z}) = 6$ V across R_2, yielding $E_{z1} = E_{z2} = -3$ V/m. Note that in this case the entire emf appears across the open termination (gap in the loop C and the emf $\int_b^t (\mathbf{E} + \mathbf{v} \times \mathbf{B})_1 \cdot d\mathbf{l} = 0$ across resistor R_1).

Moving bar in the presence of a constant magnetic field produces an emf and electric fields in the lab frame that drive a loop current I.

Example 4 illustrates how the $\oint \mathbf{E} \cdot d\mathbf{l}$ part of emf $\oint (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l}$ caused by a motion $\mathbf{v} = 3\hat{x}$ m/s is zero (with non-zero static E_z components)!!



Transformers which operate based on an inductive coupling principle, and electric dynamos (and motors) which produce motion induced emfs (and rotating coils) are studied in depth in power courses starting with ECE 330.

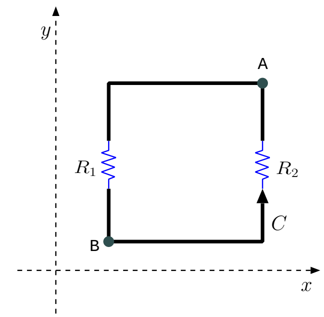
Example 5: An infinite solenoid producing a *constant* $-\frac{d\Psi}{dt} = 8$ V, passes through small a loop consisting of a 1 Ω resistor on the right and a 3 Ω resistor on the left, connected in series — see margin plot. What is the current I_c through this resistor loop, and what voltages would be measured (by a voltmeter) across the individual resistors?

Solution: The magnetic flux produced by the solenoid will be confined to its interior as long as dI/dt (and thus $d\Psi/dt$, as specified) is constant and emf $\mathcal{E} = -d\Psi/dt$ is non-time varying (see below). In that case, with constant emf $\mathcal{E} = -\frac{d\Psi}{dt} = 8$ V of the encircling resistor loop in the setup, the loop current I_c is the ratio of $\mathcal{E}$ and the total loop resistance 4 Ω, i.e., $I_c = \frac{\mathcal{E}}{R} = 2$ A. Consequently, 1 and 3 Ω resistors will develop 2 and 6 V drops, respectively, in the direction of the 2A current!! Note that:

- the loop has no battery to support this current flow — it has instead been excited "inductively".
- with *constant* dI/dt, there is *zero* magnetic field at the locations of the loop wire and resistors (static $\mathbf{E}$ in the solenoid exterior is *curl-free*!) — thus, the emf of the loop is not being produced by a time varying local magnetic field; it is rather a consequence of the time-varing current $I(t)$ in the solenoid (which is also responsible for time-varying Ψ), with the relation $\mathcal{E} = -d\Psi/dt$ being "incidental"!
- what a voltmeter measures across the resistors — whether 2 or 6 V — depends on whether its probes contacting points A and B are placed to the right or to the left of the solenoid!! That's because the field $\mathbf{E}$ produced by the time-varying current $I(t)$ is no longer conservative across the system and consequently the line integral of $\mathbf{E}$ is path dependent — we have to be more careful about what we mean by *voltage* in these new situations!

Example 6: Consider a square conducting loop of 1 m^2 cross sectional area bordered by $R_1 = 2\,\Omega$ and $R_2 = 1\,\Omega$ resistors as shown in the margin. The loop is linked with a magnetic flux Ψ due to time varying magnetic field described as $\mathbf{B} = (12 - 3t)\hat{z}$ T.

- Hence, $\Psi = 12 - 3t$ Wb and the emf $\mathcal{E} = -d\Psi/dt = 3$ V.
- Loop current $I = 3\text{V}/3\Omega = 1$ A in the circulation direction.
- Voltage drop $V_1 = 2$ V across R_1 from point A to point B.
- Voltage drop $V_2 = -1$ V across R_2 from point A to point B.
- A voltmeter connected from A (positive lead) to B will read 2 V if and only if its leads form a path identical to the path defined by R_1 (from A to B).
- A voltmeter connected from A (positive lead) to B will read -1 V if and only if its leads form a path identical to the path defined by R_2.
- A voltmeter connected from A (postive lead) to B will read 0.5 V if its leads form a diagonal path from A to B.
  - To see this, notice that Faraday's law applied for the triangular loop including the voltmeter and R_2 would have an emf of 1.5 V equaling the sum of voltmeter reading V_R and 1 V drop across resistor R_2.



In the presence of time varying magnetic flux, voltage of a path P, defined as $\int_P (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l}$, will in general be path dependent!

A voltmeter reads and displays the voltage of its own path constituted by the placement of its own probe wires contacting the measurement nodes A and B.

15 Inductance — coil, solenoid, shorted coax

- Given a circular coil with some resistance R and conducting some current I, the magnetic flux Ψ produced by I and "linking" the coil itself — see figure on the right — can be expressed as

$$\Psi = LI$$

using a non-negative proportionality constant

$$L = \frac{\Psi}{I}$$

termed the ***self*-inductance** of the coil measured in units of Henries (H=Wb/A)[1].

- Given $\Psi = LI$, and its time derivative

$$\frac{d\Psi}{dt} = L\frac{dI}{dt},$$

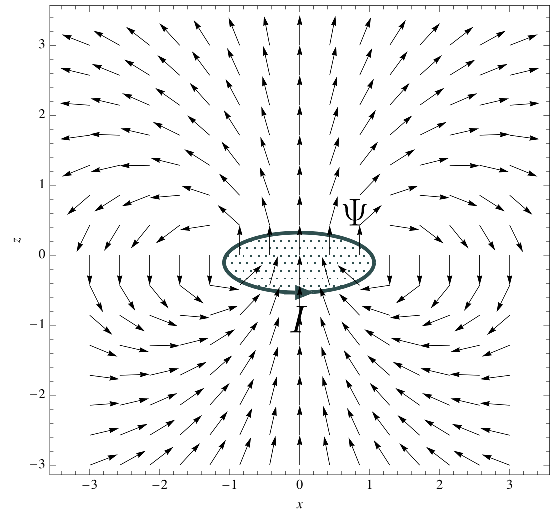
it follows that Faraday's equation applied to the coil is

$$\mathcal{E} = -\frac{d\Psi}{dt} = -L\frac{dI}{dt},$$

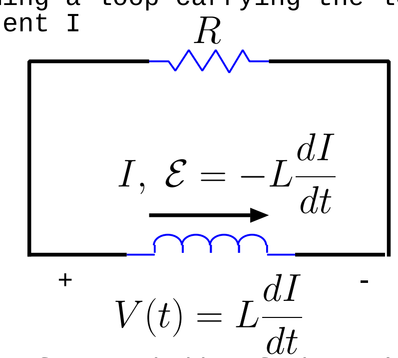
indicating a ***self*-emf** $-L\frac{dI}{dt}$ representing a ***voltage rise*** around the coil in the direction of current flow $I = \mathcal{E}/R$ — see an equivalent circuit model for the coil derived from these relations shown on the right.

[1]As opposed to a *mutual* inductance M, also measured in Henries, relating the flux linking a coil C to a current I_o flowing in a second coil C_o.

(a) A one turn coil with current I generates its own linked magnetic flux LI as shown, where a non-negative L is the inductance of the coil.



(b) An equivalent circuit model for the coil expressed in terms of lumped resistor R and inductor L forming a loop carrying the loop current I



The emf RI=-LdI/dt of the coil appears as a voltage rise across the inductor in the ckt model, as well as a voltage drop across the resistor, both taken in the direction of current I. Voltage drop V across the inductor in the current direction is LdI/dt, as we learned in our circuit courses.

- The current I and self-emf $\mathcal{E}$ are then the solutions of differential equations

$$RI = -L\frac{dI}{dt} \text{ and } R\mathcal{E} = -L\frac{d\mathcal{E}}{dt},$$

  respectively, and exhibit an exponential decay with a time constant of $\tau = L/R$ (just like in LR circuits seen in ckt courses, and in analogy with time constant $\tau = RC$ that governs voltage decays in RC circuits).

  - Note that $\tau = L/R$ implies that when the inducance L is large, so is time constant τ, and current decay in the inductor is slow — inductors with large L will behave like slowly time-varying current sources (just like capacitors behaving like time-varying voltage sources) as they relase their stored energy (while maintaining a voltage rise $-L\frac{dI}{dt}$ determined by other elements in their connected circuits).

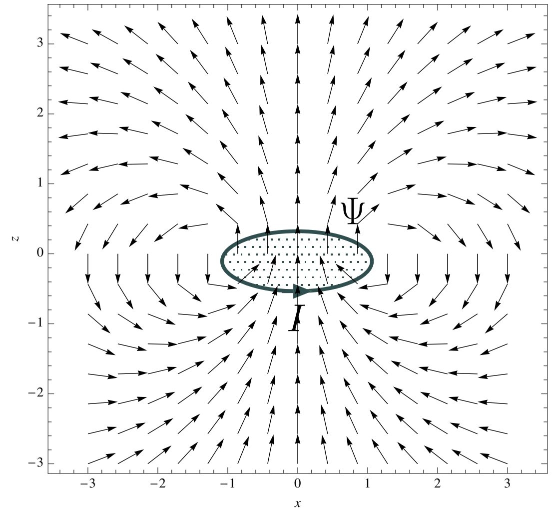
- For an inductor consisting of n-loops, the emf $\mathcal{E}$ measured across all n-loops is naturally (since n emf's add up)

$$\mathcal{E} = n(-\frac{d}{dt}\Psi) = -\frac{d}{dt}n\Psi \equiv -L\frac{dI}{dt}$$

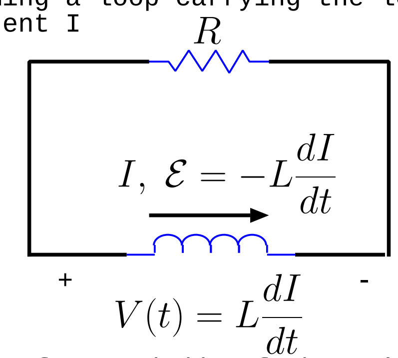
  implying an inductance

$$L \equiv \frac{n\Psi}{I}.$$

(a) A one turn coil with current I generates its own linked magnetic flux LI as shown, where a non-negative L is the inductance of the coil.



(b) An equivalent circuit model for the coil expressed in terms of lumped resistor R and inductor L forming a loop carrying the loop current I



The emf RI=-LdI/dt of the coil appears as a voltage rise across the inductor in the ckt model, as well as a voltage drop across the resistor, both taken in the direction of current I. Voltage drop V across the inductor in the current direction is LdI/dt, as we learned in our circuit courses.

Example 1: An n-turn coil has a resistance $R = 1\,\Omega$ and inductance of $1\,\mu$H. If it is conducting 3 A current at $t = 0$, determine $I(t)$ for $t > 0$.

Solution: Current flow in the resistive n-turn coil will be driven by self-emf $\mathcal{E} = -L\frac{dI}{dt}$ matching a voltage drop RI. Hence

$$RI = -L\frac{dI}{dt} \;\; \leftrightarrow \;\; \frac{dI}{dt} + \frac{R}{L}I = 0 \;\; \Rightarrow \;\; I(t) = I(0)e^{-\frac{R}{L}t} = 3e^{-10^6 t} \text{ A}.$$

- As illustrated by above example, current I around a resistive loop C will in general be obtained by solving a *differential equation* constructed using the emf of the loop.
  - The algebraic $I = \frac{\mathcal{E}}{R}$ solution used last lecture assumed that self-emf $-L\frac{dI}{dt}$ produced by the induced current $I(t)$ is small compared to an externally produced emf.

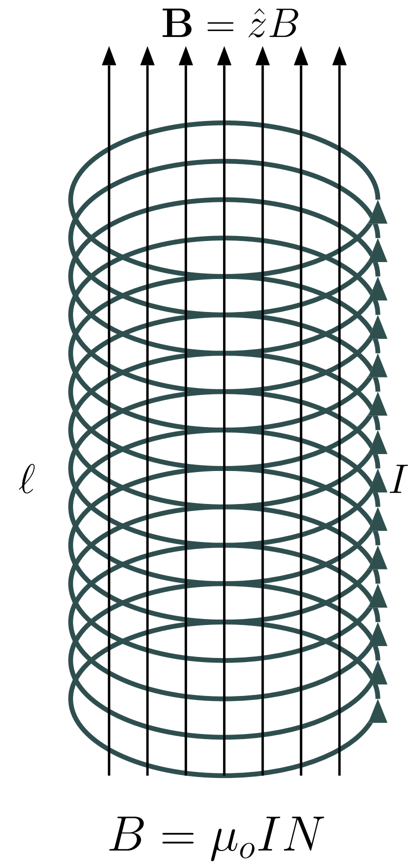
We continue with typical inductance calculations.

Inductance of long solenoid: Consider a long solenoid with length ℓ, cross-sectional area A, and a density of N loops per unit length as examined in Example 3 of Lecture 12 (see figure in the margin). As determined in Example 3, the magnetic flux density in the interior of the solenoid is

$$\mathbf{B} = \mu_o I N \hat{z}$$

while $n = N\ell$ is the number of turns of the solenoid. Thus, the inductance of the solenoid with $n = N\ell$ turns is

$$L = \frac{n\Psi}{I} = \frac{N\ell(\mu_o I N)A}{I} = N^2 \mu_o A \ell.$$



- As we know from our circuit courses, an inductor L such as the solenoid coil considered above can be used to store energy. An inductor connected to an external circuit with a quasi-static current I develops a voltage drop $V = L\frac{dI}{dt}$ across its terminals[2] and absorbs power at an instantaneous rate

  $$P = VI = L\frac{dI}{dt}I = \frac{d}{dt}(\frac{1}{2}LI^2),$$

  implying a stored energy of

  $$W = \frac{1}{2}LI^2 = \frac{1}{2}N^2\mu_o A\ell I^2 = \frac{|B_z|^2}{2\mu_o}A\ell = \frac{1}{2}\mu_o |H_z|^2 A\ell$$

  in an inductor in a conducting state.

[2]Assuming a physical size much smaller than a wavelength $\lambda = c/f$ for the highest frequency in $I(t)$.

- Notice that the stored energy of the solenoid is

$$\frac{1}{2}\mu_o|H_z|^2 = \frac{1}{2}\mu_o\mathbf{H}\cdot\mathbf{H}$$

times its volume $A\ell$ occupied by the field $\mathbf{H}$ inside the solenoid. That suggests that

$$w = \frac{1}{2}\mu_o\mathbf{H}\cdot\mathbf{H}$$

can be interpreted as stored magnetostatic energy per unit volume in general.

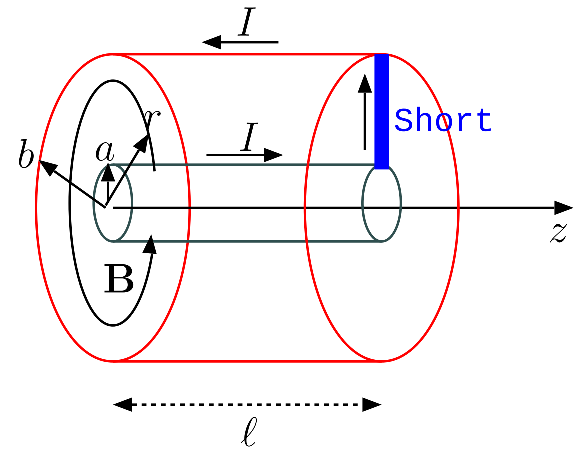
  - Also both inductance L and stored energies W and w would have μ replacing μ_o in material media with permeabilities

$$\mu = (1+\chi_m)\mu_o$$

    and magnetic susceptibilities χ_m, in analogy with the concepts of permittivity $\epsilon = (1+\chi_e)\epsilon_o$ and electrical susceptibility χ_e.

    - Permeability and magnetic susceptibility notions will be examined in a future lecture.

Inductance of shorted coax: Consider a coaxial cable of some length ℓ which is "shorted" at one end (with a wire connecting the inner and outer conductors), so that a steady current I can flow on the inner conductor of radius a to return on the interior surface of the outer conductor at radius b after having circulated through the short. We will next determine the inductance L of such an inductor after first computing the magnetic flux density B_ϕ that will be produced by the inner conductor current I. In B_ϕ calculation we will assume $\ell \gg b$ so that an "infinite coax" approximation can be invoked.



Shorted coax circulates a current I linking a magnetic flux Ψ confined to a region bounded by the outer conductor of the coax.

- Expanding the integral form of Ampere's law

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \mu_o I_C$$

as

$$B_\phi 2\pi r = \mu_o I$$

over a circular integration contour C of a radius $r > a$, we find that the magnetic flux density in the interior of the coax cable is

$$B_\phi = \frac{\mu_o I}{2\pi r}.$$

- Therefore, the magnetic flux linked by the closed current path I (see figure in the margin) is

$$\Psi = \int_S \mathbf{B} \cdot d\mathbf{S} = \ell \frac{\mu_o}{2\pi} I \int_a^b \frac{dr}{r} = \ell \frac{\mu_o}{2\pi} \ln \frac{b}{a} I.$$

Clearly, we have a linear relation $\Psi = LI$, with

$$L \equiv \frac{\ln \frac{b}{a}}{2\pi} \ell \mu_o,$$

which is the inductance of a shorted coax of a finite length ℓ.

– The inductance of the coax per unit length is

$$\mathcal{L} = \frac{\ln \frac{b}{a}}{2\pi} \mu_o,$$

which should be contrasted with capacitance per unit length

$$\mathcal{C} = \frac{2\pi}{\ln \frac{b}{a}} \epsilon_o$$

of the same coax configuration.

Notice how $\mathcal{L}$ and $\mathcal{C}$ are proportional to ϵ_o and μ_o, respectively, having proportionality constants which are inverses of one another.

Inductance of shorted parallel plates: If a pair of parallel plates of areas $A = W\ell$ and separation d were shorted at one end, we would obtain effectively an inductor with a per length inductance

$$\mathcal{L} = \frac{d}{W}\mu_o$$

that accompanies per length capacitance

$$\mathcal{C} = \frac{W}{d}\epsilon_o$$

of the same parallel plate configuration. This follows from a generalization of our finding above that the proportionality constants of $\mathcal{L}$ and $\mathcal{C}$ are arithmetic inverses of one another.

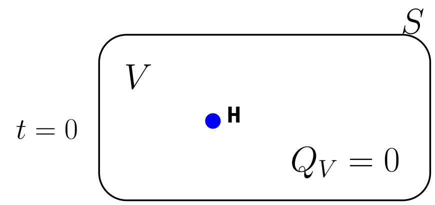
16 Charge conservation, continuity eqn, displacement current, Maxwell's equations

- **Total electric charge** is ***conserved*** in nature in the following sense: if a process generates (or eliminates) a positive charge, it always does so as accompanied by a negative charge of equal magnitude.

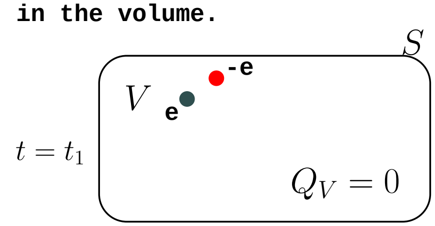
  - Example: **Photoionization** of atoms and molecules can generate free positive ions and free negative electrons in pairs (see margin figure). Photoionization is a process that converts bound charge carriers into free charge carriers.
  - Example: **Recombination** when a positive ion and an electron get together to produce a charge neutral atom or molecule.
  - Example: **Annihilation** of an electron (negative charge) by a positron (positive charge of equal magnitude) and the reverse process of **pair creation**.

As a consequence, if the total electric charge Q_V contained in any finite volume V changes as a function of time, this change must be attributed to a net transport of charge, i.e., electric current, across the bounding surface S of volume V as detailed below.

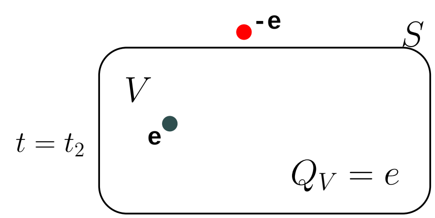
(a) At t=0 volume V contains a neutral atom but no net charge



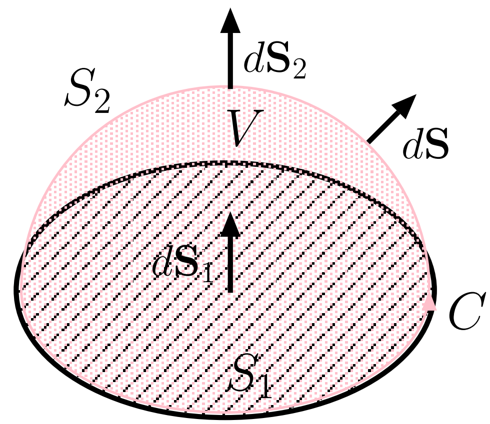
(b) At t=t1 volume V contains a proton and a free electron after the ionization of the hydrogen atom. There is still no net charge in the volume.



(c) At t=t2 volume V now contains only a proton after the exit of free electron through surface S. Now V contains a net charge e.



- Consider two distinct surfaces S_1 and S_2 bounded by the same closed loop C (as shown in the margin) such that a volume V is contained between the two surfaces.



  – Let

$$I_1 = \int_{S_1} \mathbf{J} \cdot d\mathbf{S}_1$$

  and

$$I_2 = \int_{S_2} \mathbf{J} \cdot d\mathbf{S}_2$$

  denote currents flowing through surfaces S_1 and S_2, respectively.

  – Note that current I_1 through surface S_1 enters volume V, while current I_2 through surface S_2 exits volume V (with the directions assigned to $d\mathbf{S}_1$ and $d\mathbf{S}_2$).

  – If $I_1 \neq I_2$, then **current out** is not matched by the **current in**, and as a result, the net charge Q_V contained in volume V increases with time at a rate $I_1 - I_2$ provided that charge is *conserved* in the sense discussed above. In that case, we have

$$\frac{dQ_V}{dt} = I_1 - I_2.$$

  This relationship can be expressed as

$$\frac{d}{dt} \int_V \rho dV = \int_{S_1} \mathbf{J} \cdot d\mathbf{S}_1 - \int_{S_2} \mathbf{J} \cdot d\mathbf{S}_2$$

since, in terms of charge density ρ, charge in volume V is

$$Q_V = \int_V \rho dV.$$

The expression can also be cast as

$$\int_V \frac{\partial \rho}{\partial t} dV = -\oint_S \mathbf{J} \cdot d\mathbf{S}$$

where S is the union of surfaces S_1 and S_2 enclosing V, and $d\mathbf{S}$ is an outward area element of S (see margin). This relationship is known as **continuity equation**. Its *differential form* is

Continuity equation

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \mathbf{J},$$

which follows from the *integral form* above as a consequence of **divergence theorem** (recall Lecture 4).

Continuity equation is a mathematical re-statement of the principle of conservation of charge.

- While **Faraday's law**

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

indicates that time-varying $\mathbf{B}$ induces time-varying electric fields $\mathbf{E}$, **Ampere's law**, written as

$$\nabla \times \mathbf{H} = \mathbf{J},$$

makes no such claim about a time-varying $\mathbf{E}$ inducing a time-varying

$$\mathbf{B} = \mu_o \mathbf{H}.$$

  - This "asymmetry" was noted by James Clerk Maxwell who realized that the form of Ampere's law given above must be "incomplete" under time-varying situations.
  - Noting the inconsistency of Ampere's law with the continuity equation under time varying conditions, he re-wrote the Ampere's law as

Revised Ampere's law (with "displacement current")

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}$$

in 1861 by adding the term on the right which is now called the "displacement current".

    - Maxwell postulated that the displacement current term is needed in Ampere's law because only then the divergence of Ampere's law avoids falling into conflict with charge conservation (under time varying conditions).

Verification of Maxwell's claim: Since $\nabla \times \mathbf{H}$ is divergence-free (just like the curl of vector potential $\mathbf{A}$, namely $\mathbf{B}$), it follows that the divergence of *Maxwell's* modified Ampere's law — often called **Ampere-Maxwell equation** — is

$$\nabla \cdot (\nabla \times \mathbf{H}) = \nabla \cdot \mathbf{J} + \frac{\partial}{\partial t}\nabla \cdot \mathbf{D} = 0.$$

– In the absence of the second term due to displacement current, this results would be inconsistent with the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0,$$

unless $\frac{\partial \rho}{\partial t} = 0$ (the static case).

– By, contrast, including the second term, the result above is recognized as the continuity equation per se, since by Gauss's law — assuming that it applies with no change under time varying situations —

$$\frac{\partial}{\partial t}\nabla \cdot \mathbf{D} = \frac{\partial \rho}{\partial t}.$$

- The modified Ampere's law

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}$$

postulated by Maxwell under the assumption that Gauss's law is also valid under time-varying conditions, leads to some specific predictions about how time-varying fields should behave.

- These predictions — concerning the propagation of electromagnetic waves — were validated experimentally by Heinrich Hertz around 1888.

  - The experiments confirmed that time-varying electric and magnetic fields obey collectively (and at microscopic scales) the differential relations

Maxwell's equations

$$\begin{aligned}\nabla\cdot\mathbf{D} &= \rho \qquad \text{Gauss's law}\\ \nabla\cdot\mathbf{B} &= 0\\ \nabla\times\mathbf{E} &= -\frac{\partial\mathbf{B}}{\partial t} \qquad \text{Faraday's law}\\ \nabla\times\mathbf{H} &= \mathbf{J}+\frac{\partial\mathbf{D}}{\partial t}, \qquad \text{Ampere's law}\end{aligned}$$

    where

$$\mathbf{D}=\epsilon_o\mathbf{E} \text{ and } \mathbf{B}=\mu_o\mathbf{H}$$

    provided that ρ and $\mathbf{J}$ describe the distributions of all charges and currents associated with free and bound charge carriers[1]. Alternatively, the same differential relations — known collectively as Maxwell's equations — are also valid for macroscopic fields, provided that ρ and $\mathbf{J}$ describe only the free charge contributions and

$$\mathbf{D}=\epsilon\mathbf{E} \text{ and } \mathbf{B}=\mu\mathbf{H}$$

[1] In the classical domain, down to scales of about $\hbar/mc$, the Compton wavelength — at shorter scales quantized and generalized versions (known as *electroweak* theory) are needed.

in terms of suitably defined permittivities and permeabilities ϵ and μ — see next Lecture.

- The *unnamed* Maxwell equation

$$\nabla \cdot \mathbf{B} = 0$$

can be viewed to be a consequence of Faraday's law

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

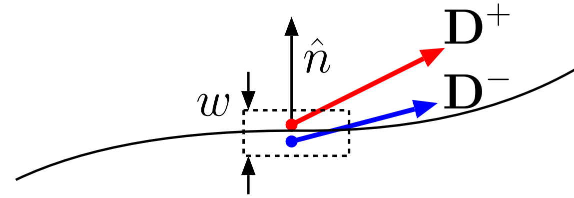
and the fact that magnetic monopoles have never been observed.

Explanation: Since $\nabla \times \mathbf{E}$ is divergence-free, taking the divergence of Faraday's law, we get

$$\nabla \cdot (\nabla \times \mathbf{E}) = -\frac{\partial}{\partial t} \nabla \cdot \mathbf{B} = 0.$$

This constraint requires $\nabla \cdot \mathbf{B}$ to an invariant scalar at all locations in space. As a consequence, if $\nabla \cdot \mathbf{B} = 0$ at some instant in time, it should remain so at all times. Given that $\nabla \cdot \mathbf{B} = 0$ for static fields, this relationship must also continue to be valid when $\mathbf{B}$ starts changing with time.

The fact that $\nabla \cdot \mathbf{B}$ remains fixed at a zero value everywhere, whereas $\nabla \cdot \mathbf{D}$ varies like ρ, is in fact a consequence of the fact that there appears to be no *magnetic charges* (monopoles) in nature. Had there been "point charges for magnetic fields" in nature, $\nabla \cdot \mathbf{B}$ would have equaled the density of those charges, and magnetic field lines would have started and stopped on them (rather than looping into themselves). But no one has observed of any evidence for such magnetic charges *anywhere*, even going back to the very early times in the history of the universe (accessible by making observations of very far astronomical objects). So, $\nabla \cdot \mathbf{B} = 0$.



- Finally, the full set of Maxwell's *boundary condition equations* concerning any interface with a normal unit vector $\hat{n}$ are

$$\begin{aligned} \hat{n} \cdot (\mathbf{D}^+ - \mathbf{D}^-) &= \rho_s \\ \hat{n} \cdot (\mathbf{B}^+ - \mathbf{B}^-) &= 0 \\ \hat{n} \times (\mathbf{E}^+ - \mathbf{E}^-) &= 0 \\ \hat{n} \times (\mathbf{H}^+ - \mathbf{H}^-) &= \mathbf{J_s} \end{aligned}$$

  - We had already seen how the first and third boundary condition equations arise.
  - The second boundary condition equation concerning the normal component of $\mathbf{B}$ is another consequence of the absence of magnetic charges (see margin).
  - A detailed justification of the last boundary condition concerning tangential $\mathbf{H}$ will be given explicitly during Lecture 19. This equation allows a discontinuous change in the tangential component of $\mathbf{H}$ if the interface contains a non-zero surface current $\mathbf{J}_s$.

Constraint

$$\oint_C \mathbf{H} \cdot d\mathbf{l} = \int_S (\mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}) \cdot d\mathbf{S}$$

around the dotted path yields

$$\hat{n} \times (\mathbf{H}^+ - \mathbf{H}^-) = \mathbf{J}_s$$

in $w \to 0$ limit.

Constraint

$$\oint_S \mathbf{B} \cdot d\mathbf{S} = 0$$

applied over the dotted volume (seen in profile) yields

$$B_n^+ - B_n^- = 0$$

in $w \to 0$ limit.

17 Magnetization current, Maxwell's equations in material media

- Consider the microscopic-form Maxwell's equations

$$\begin{aligned}
\nabla\cdot\mathbf{D} &= \rho \qquad \text{Gauss's law}\\
\nabla\cdot\mathbf{B} &= 0\\
\nabla\times\mathbf{E} &= -\frac{\partial\mathbf{B}}{\partial t} \qquad \text{Faraday's law}\\
\nabla\times\mathbf{H} &= \mathbf{J}+\frac{\partial\mathbf{D}}{\partial t}, \qquad \text{Ampere's law}
\end{aligned}$$

  where

$$\begin{aligned}
\mathbf{D} &= \epsilon_o\mathbf{E}\\
\mathbf{B} &= \mu_o\mathbf{H}.
\end{aligned}$$

- Direct applications of these equations in material media containing a colossal number of bound charges is impractical.

- Macroscopic-form Maxwell's equations suitable for material media are obtained by first expressing ρ and $\mathbf{J}$ above as the macroscopic quantities

$$\rho = \rho_f - \nabla\cdot\mathbf{P}$$

  and

$$\mathbf{J} = \mathbf{J}_f + \frac{\partial\mathbf{P}}{\partial t} + \nabla\times\mathbf{M}$$

  where

- subscripts f indicate charge and current density contributions due to free charge carriers,
- the term $-\nabla \cdot \mathbf{P}$ denotes the **bound charge density**,
- the term $\frac{\partial \mathbf{P}}{\partial t}$ denotes the **polarization current density** due to oscillating dipoles (already discussed in Lecture 11), and
- $\nabla \times \mathbf{M}$ is a **"magnetization current density"** also due to bound charges, an effect that we will discuss and clarify later in this section.

Using these expressions in Gauss's and Ampere's laws

$$\begin{aligned} \nabla \cdot \epsilon_o \mathbf{E} &= \rho \qquad \text{Gauss's law} \\ \nabla \times \mu_o^{-1} \mathbf{B} &= \mathbf{J} + \frac{\partial \epsilon_o \mathbf{E}}{\partial t}, \qquad \text{Ampere's law} \end{aligned}$$

we obtain

$$\begin{aligned} \nabla \cdot (\epsilon_o \mathbf{E} + \mathbf{P}) &= \rho_f \qquad \text{Gauss's law} \\ \nabla \times (\mu_o^{-1} \mathbf{B} - \mathbf{M}) &= \mathbf{J}_f + \frac{\partial}{\partial t}(\epsilon_o \mathbf{E} + \mathbf{P}), \qquad \text{Ampere's law.} \end{aligned}$$

Now, re-define $\mathbf{D}$ and $\mathbf{H}$ as

$$\mathbf{D} = \epsilon_e \mathbf{E} + \mathbf{P} = \epsilon \mathbf{E}$$

and

$$\mathbf{H} = \mu_o^{-1} \mathbf{B} - \mathbf{M} = \mu^{-1} \mathbf{B},$$

respectively, and drop the subscripts f which will no longer be needed.

Using these new definitions, the full set of Maxwell's equations now read as (the same form as before)

$$\begin{aligned} \nabla \cdot \mathbf{D} &= \rho \qquad \text{Gauss's law} \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \qquad \text{Faraday's law} \\ \nabla \times \mathbf{H} &= \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}, \qquad \text{Ampere's law} \end{aligned}$$

with

$$\begin{aligned} \mathbf{D} &= \epsilon \mathbf{E} \\ \mathbf{B} &= \mu \mathbf{H}, \end{aligned}$$

where ρ and $\mathbf{J}$ are understood to be due to free charge carriers only.

- We had already seen many aspects of the above procedure for obtaining the macroscopic form field equations earlier on (e.g., in Lectures 8 and 11).
  - In particular we were already familiar with the revised definition of $\mathbf{D} = \epsilon \mathbf{E}$ along with the concept of **medium permittivity** ϵ.
  - The new feature above that requires further discussions is the relation $\mathbf{B} = \mu \mathbf{H}$ along with the concept of **medium permeability** μ. The details of this relation are connected to the concept of "magnetization current" which we discuss next.

- Just like “free charge” density and currents, “bound charge” densities and currents also have to satisfy the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0.$$

  - This equation is automatically satisfied if we substitute

    $$\rho = \rho_b = -\nabla \cdot \mathbf{P}$$

    and

    $$\mathbf{J} = \mathbf{J}_b = \frac{\partial \mathbf{P}}{\partial t}$$

    in it.

    Verification:

    $$\frac{\partial \rho_b}{\partial t} + \nabla \cdot \mathbf{J}_b = \frac{\partial}{\partial t}(-\nabla \cdot \mathbf{P}) + \nabla \cdot \frac{\partial \mathbf{P}}{\partial t} = 0$$

    since the order of time derivative and divergence can be exchanged on the right.

  - But the same equation is also satisfied if we take

    $$\mathbf{J}_b = \frac{\partial \mathbf{P}}{\partial t} + \nabla \times \mathbf{M}$$

    for any vector field $\mathbf{M}$ simply because vector $\nabla \times \mathbf{M}$ is divergence free.

Consequently, it is not sufficient to represent bound currents in material media as simply $\frac{\partial \mathbf{P}}{\partial t}$, if bound carriers can also conduct divergence-free currents due to closed-loop orbits.

- In fact, electrons "orbiting" atomic nuclei certainly produce such divergence-free current loops at microscopic scales — we account for such currents at macroscopic scales by including a magnetization current term $\nabla \times \mathbf{M}$ in $\mathbf{J}_b$.
- Also, bound charge motions *within* nucleons[1] — proton and neutrons — produce magnetization currents $\nabla \times \mathbf{M}$.
- Even bare electrons can produce magnetization currents $\nabla \times \mathbf{M}$ because of their *intrinsic spin*[2].

Once $\nabla \times \mathbf{M}$ is included in $\mathbf{J}_b$, it follows from Ampere's law that

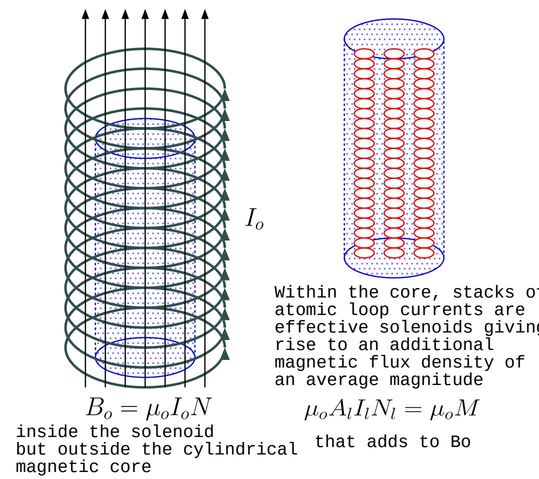
$$\mathbf{H} = \mu_o^{-1}\mathbf{B} - \mathbf{M}$$

where $\mathbf{M}$ is referred to as **magnetization** field.

[1] Physical models of nucleons involve bound charge carriers known as *quarks* which cannot be observed in a free state.

[2] All elementary charge carriers carry an intrinsic magnetization proportional to charge-to-mass ratio $\frac{q}{m}$ and a "spin angular momentum" having quantized values of $\pm\frac{\hbar}{2}$ N.m.s in any measurement direction. Using Heisenberg's uncertainty principle, $\Delta p \Delta r \geq \frac{\hbar}{2}$, we can interpret the *spin angular momentum* of an elementary particle as the lower bound of $\Delta p \Delta r$, the product of quantum uncertainties in particle momentum and position. There is no classical interpretation of spin angular momentum for point particles.

- To get a physical picture about magnetization $\mathbf{M}$ and the physical origin of $\mathbf{H} = \mu_o^{-1}\mathbf{B} - \mathbf{M}$ consider a solenoid wound around some cylindrical shaped material as shown in the margin. We know that with a solenoid current I_o, we would have $\mathbf{H}_o = NI_o\hat{z}$ in the interior of a solenoid with N loops per unit legth aligned with the z-axis, and a corresponding magnetic flux density $\mathbf{B}_o = \mu_o NI_o\hat{z}$ when the solenoid core is free space. This will be modified to some $\mathbf{B} = \mathbf{B}_o + \mu_o\mathbf{M}$ when a material core is introduced into the same space, where $\mu_o\mathbf{M}$ stands for the (additional) macroscopic (space averaged) magnetic flux density produced by microscopic current loops localized within the atoms constituting the core.



  - If there are $N_a = \frac{1}{\Delta x\Delta y\Delta z}$ atoms per unit volume in the core, with Δx separations in x direction and so forth, loop currents I_l of a stack of atoms with Δz separations in z would produce an effective solenoid an internal z-directed magnetic flux density of $\mu_o\frac{I_l}{\Delta z}\hat{z}$ and zero exterior field.
  - Since one such atomic stack solenoid with a loop area of A_l will be found for every $\Delta x\Delta y$ cross-sectional area of the core, a macroscopic average magnetic flux density produced by these atomic solenoids would be calculated as (this calculation is similar to finding the average polarization field in a dielectric as discussed in Lecture 8) $\frac{A_l}{\Delta x\Delta y} \times \mu_o\frac{I_l}{\Delta z}\hat{z} = \mu_o N_a I_l A_l\hat{z} \equiv \mu_o\mathbf{M}$, with $\mathbf{M} = N_a\mathbf{m}$, $\mathbf{m} \equiv I_l A_l\hat{z}$.

- Here $\mathbf{m}$ is the magnetic dipole moment of each current loop (analogous to electric dipole $\mathbf{p} = q\mathbf{r}$), $\mathbf{M}$ is the magnetization field vector (analogous to $\mathbf{P} = N_a\mathbf{p}$), which is a simple product of $\mathbf{m}$ per magnetized atom and the atomic number density N_a in the core.

– Superposing the magnetic flux densities of $\mu_o\mathbf{M}$ and$\mathbf{B}_o$, we obtain $\mathbf{B} = \mathbf{B}_o + \mu_o\mathbf{M}$ for the core region, or for any region of space having a non-zero magnetization $\mathbf{M}$, which then leads to the general result $\mathbf{H} = \mu_o^{-1}\mathbf{B} - \mathbf{M}$, which is further discussed below.

– Notice, whether the flux density $\mathbf{B} = \mathbf{B}_o + \mu_o\mathbf{M}$ inside the material medium is stronger or weaker in magnitude than $\mathbf{B}_o$ depends on the direction of $\mathbf{M}$, which, in turn, depends on the algebraic sign of microscopic loop currents I_l introduced above.

  - Negative I_l is found in **diamagnetic** materials where $|\mathbf{B}| < |\mathbf{B}_o|$, while positive I_l in **paramagnetic** and **ferromagnetic** materials where $|\mathbf{B}| > |\mathbf{B}_o|$, as discussed below.

– Also, the expression $\mathbf{H} = \mu_o^{-1}\mathbf{B} - \mathbf{M}$ leads to $\mathbf{H} = \mu_o^{-1}\mathbf{B} = \mathbf{H_o}$ in the exterior region where $\mathbf{M} = 0$, indicating that while fields $\mathbf{B}$ and $\mathbf{B_o}$if the interior and exterior are different, $\mathbf{H}$ is the same in both regions (analogous with $\mathbf{D}$ in dielectrics).

- Lab measurements — e.g., inductances L measured for coils wound around magnetic materials[3] — show that for a large class of materials

$$\mathbf{M} \equiv \mu_o^{-1}\mathbf{B} - \mathbf{H}$$

varies linearly with $\mathbf{H}$ (which is of course possible only when $\mathbf{B}$ also varies linearly with $\mathbf{H}$).

  - In that case we write

$$\mathbf{M} = \chi_m \mathbf{H},$$

where χ_m is a dimensionless parameter called **magnetic susceptibility**, and obtain a relation

Magnetic susceptibility and permeability

$$\mathbf{B} = \mu_o(1 + \chi_m)\mathbf{H} = \mu\mathbf{H},$$

where

$$\mu = \mu_o(1 + \chi_m)$$

is called the **permeability** of the medium.

[3]Recall from Lecture 15 that $L \propto \mu$ when inductors are wound around materials with permeability μ.

- For a large class of materials with $\mathbf{M} \propto \mathbf{H}$, it is observed that $|\chi_m| \ll 1$. In that case, the material is called

  - **Diamagnetic** if $\chi_m < 0$:

    - Diamagnetism occurs when an applied magnetic field *induces* electron **orbital angular momentum** in a collection of atoms having no net permanent magnetization $\mathbf{M}$ — in such materials electron clouds around atomic nuclei spin up in accordance with Lenz's to generate magnetic fields opposing the applied magnetic field so as to keep $\mathbf{B} = \mu\mathbf{H}$ smaller than $\mu_o\mathbf{H}$. This happens in materials that we ordinarily think of being non-magnetic (wood, glass, water, etc.). Diamagnetic materials are in fact very weakly repulsed by permanent magnets since $\mu \approx \mu_o$ in all diamagnetic materials.

  - **Paramagnetic** if $\chi_m > 0$:

    - Paramagnetism occurs in materials composed of atoms having permanent magnetic dipole moments due to electron **spin angular momentum** — magnetic dipoles of such atoms co-align with the applied magnetic field due to $\mathbf{v} \times \mathbf{B}$ related torques, leading to $\mathbf{M}$ pointing in the applied $\mathbf{B}$ direction[4]. This happens for atoms with unfilled inner electron shells, because in filled shells electron spins are opposite (due to Pauli

[4]In these materials the described paramagnetism overcomes the diamagnetic tendency of the material caused by the orbital angular momenta of its electrons around atomic nuclei.

exclusion principle) and cancel one another. Unfilled outer shells do not usually give rise to paramagnetism because interactions between adjacent atoms in that case give rise to opposite spins of their outer shell electrons. Paramagnetic materials are very weakly attracted to permanent magnets (e.g., aluminum, lithium, tungsten).

- In a small class of materials known as **ferromagnets** — iron, nickel, and cobalt, which are metals with atoms having unfilled inner electron shells, and their various alloys — $\mathbf{M}$ can arise spontaneously (because permanent magnetic dipole moments of nearby atoms produced by **electron spins** become co-aligned as a consequence of conduction electrons moving through the lattice) and turns out to be a non-linear function of present and past values of $\mathbf{H}$, in which case experimentally obtained relations, denoted as

$$\mathbf{B} = \mathbf{B}(\mathbf{H}),$$

  need to be used in Maxwell's equations. It is even possible to have non-zero $\mathbf{B}$ in such materials with zero $\mathbf{H}$ — permanent magnets have that property.

- First principles modeling of χ_m or the $\mathbf{B} = \mathbf{B}(\mathbf{H})$ relation requires quantum mechanics (classical models turn out to be not accurate enough). Overall, the models give rise to frequency dependent results, involving loss as well as resonance features (also exhibited in Lorentz-Drude models of χ_e examined in Lecture 11) relevant for applications including various magnetic imaging techniques.

18 Wave equation and plane TEM waves in source-free media

With this lecture we start our study of the full set of Maxwell's equations shown in the margin by first restricting our attention to *homogeneous* and *non-conducting* media with constant ϵ and μ and zero σ.

$$\begin{aligned}\nabla\cdot\mathbf{D} &= \rho\\ \nabla\cdot\mathbf{B} &= 0\\ \nabla\times\mathbf{E} &= -\frac{\partial\mathbf{B}}{\partial t}\\ \nabla\times\mathbf{H} &= \mathbf{J}+\frac{\partial\mathbf{D}}{\partial t}.\end{aligned}$$

- Our first objective is to show that non-trivial (i.e., non-zero) time-varying field solutions of these equations can be obtained even in the absence of ρ and $\mathbf{J}$.

  – We already know *static* ρ and $\mathbf{J}$ to be the **source** of *static* electric and magnetic fields.

  – We will come to understand that time varying ρ and $\mathbf{J}$, which necessarily obey the continuity equation

  $$\frac{\partial\rho}{\partial t}+\nabla\cdot\mathbf{J}=0,$$

  constitute the **source** of *time-varying electromagnetic* fields.

Despite these intimate connections between the sources ρ and $\mathbf{J}$ and the fields

$$\mathbf{D}=\epsilon\mathbf{E}\quad\text{and}\quad\mathbf{B}=\mu\mathbf{H},$$

non-trivial field solutions can exist in source-free media as we will see shortly.

- Such field solutions in fact represent electromagnetic waves, a familiar example of which is **light**.

- Another example is **radiowaves** that we use when we communicate using wireless devices such as radios, cell-phones, WiFi, etc.

- Different types of electromagnetic waves are distinguished by their oscillation frequencies, and include

  - radiowaves,
  - microwaves,
  - infrared,
  - light,
  - ultraviolet,
  - X-rays, and gamma rays,

  going across the **electromagnetic spectrum** from low to high frequencies.

  We are well aware that these types of electromagnetic waves can travel across empty regions of space — e.g., from sun to Earth — transporting energy and heat as well as momentum.

  - Next, we will discover their general properties by examining Maxwell's equations under the restriction $\rho = \mathbf{J} = 0$.

- In source-free and homogeneous regions where $\rho = \mathbf{J} = 0$ and ϵ and μ are constant, we can simplify Maxwell's equations as shown in the margin.

  - If there are non-trivial solutions of these equations, namely $\mathbf{E}(\mathbf{r}, t) \neq 0$ and $\mathbf{H}(\mathbf{r}, t) \neq 0$, they evidently need to be divergence-free.
  - They also have to be "curly" according to the last two equations: Faraday's and Ampere's laws.

$$\begin{aligned} \nabla \cdot \mathbf{E} &= 0 \\ \nabla \cdot \mathbf{H} &= 0 \\ \nabla \times \mathbf{E} &= -\mu \frac{\partial \mathbf{H}}{\partial t} \\ \nabla \times \mathbf{H} &= \epsilon \frac{\partial \mathbf{E}}{\partial t}. \end{aligned}$$

- Next we will make use of vector identity

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}$$

  which should be familiar from an earlier homework problem.

  - Since the electric field $\mathbf{E}$ is divergence-free in the absence of sources, this identity simplifies as

$$\nabla \times (\nabla \times \mathbf{E}) = -\nabla^2 \mathbf{E}$$

    where in the right side $\nabla^2 \mathbf{E}$ is the Laplacian of $\mathbf{E}$.
  - Using this result we can express the curl of Faraday's law as

$$\nabla \times [\nabla \times \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t}] \;\; \Rightarrow \;\; -\nabla^2 \mathbf{E} = -\mu \frac{\partial}{\partial t} \nabla \times \mathbf{H},$$

    which combines with the Ampere's law to produce

$$\nabla^2 \mathbf{E} = \mu\epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2},$$

which can be written explicitly as

3D vector wave equation

$$\frac{\partial^2 \mathbf{E}}{\partial x^2} + \frac{\partial^2 \mathbf{E}}{\partial y^2} + \frac{\partial^2 \mathbf{E}}{\partial z^2} = \mu\epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2}.$$

Recall that our objective is to see whether a non-trivial time-varying solution of Maxwell's equations can exist in source-free media.

Our objective at this stage is not finding a general solution; it is instead identifying a simple example of a non-trivial time-varying $\mathbf{E}(\mathbf{r}, t)$, if we can.

For example, can a field solution

$$\mathbf{E}(\mathbf{r}, t) = \hat{x} E_x(z, t)$$

that only depends on z and t and "polarized" in x-direction exist? If it can exist, what would be the properties of this x-**polarized** solution?

- To find out, we note that with $\mathbf{E} = \hat{x} E_x(z, t)$, the above "wave equation" is reduced to

  1D scalar wave equation

$$\frac{\partial^2 E_x}{\partial z^2} = \mu\epsilon \frac{\partial^2 E_x}{\partial t^2},$$

  an equation that is known as a **1D scalar wave equation**, as opposed to the **3D *vector* wave equation** above.

  – Now, by substitution, we can easily show that

$$E_x = \cos(\omega(t - \sqrt{\mu\epsilon} z)),$$

satisfies the 1D wave equation and represents an x-polarized time-periodic field solution with an **oscillation frequency** ω.

– 1D wave equation can also be satisfied by

$$E_x = \cos(\omega(t + \sqrt{\mu\epsilon}z)).$$

Let us jointly refer to these solutions as

$$E_x = \cos(\omega(t \mp \frac{z}{v})),$$

where

$$v \equiv \frac{1}{\sqrt{\mu\epsilon}}$$

has the dimensions of m/s (i.e., velocity) and the algebraic signs $\mp$ distinguish between the "travel directions" of these *possible* "wave solutions" as elaborated later on.

- Let us next find out the magnetic field intensity **H** that accompanies the x-polarized electric field wave solution

$$\mathbf{E} = \hat{x}\cos(\omega(t \mp \frac{z}{v})).$$

– Since the curl of **E** is

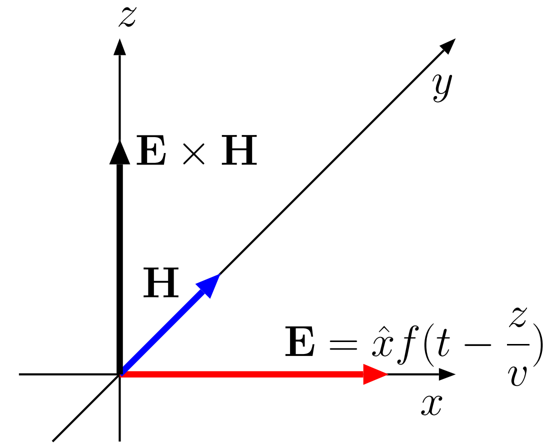
$$\nabla \times \mathbf{E} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ E_x & 0 & 0 \end{vmatrix} = \hat{y}\frac{\partial E_x}{\partial z} = \pm\hat{y}\sin(\omega(t \mp \frac{z}{v}))\frac{\omega}{v},$$

Faraday's law

$$\nabla \times \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t}$$

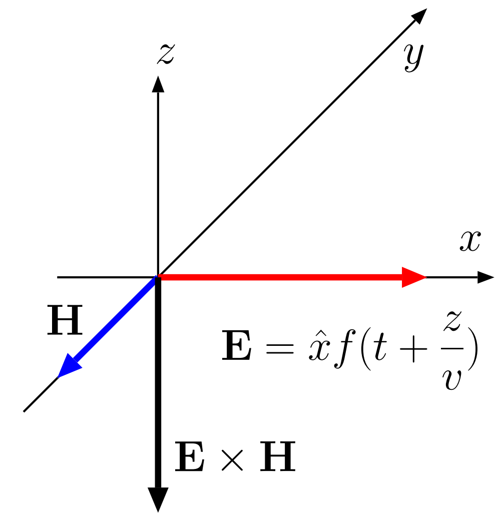
requires that $\mathbf{H}$ should satisfy

$$-\mu \frac{\partial \mathbf{H}}{\partial t} = \pm \hat{y} \sin(\omega(t \mp \frac{z}{v})) \frac{\omega}{v}.$$



Finding the *time-dependent* anti-derivative (and remembering $v = 1/\sqrt{\mu\epsilon}$), we obtain

$$\mathbf{H} = \pm \hat{y} \sqrt{\frac{\epsilon}{\mu}} \cos(\omega(t \mp \frac{z}{v})).$$



- The results above, namely our x-polarized non-trivial field solutions of Maxwell's equations in source-free homogeneous space, can be represented more compactly as

$$\mathbf{E} = \hat{x} f(t \mp \frac{z}{v}) \quad \text{and} \quad \mathbf{H} = \pm \hat{y} \frac{f(t \mp \frac{z}{v})}{\eta},$$

where

$$f(t) \equiv \cos(\omega t) = \mathrm{Re}\{e^{j\omega t}\} = \frac{e^{j\omega t} + e^{-j\omega t}}{2}$$

is the field **waveform**,

$$\eta \equiv \sqrt{\frac{\mu}{\epsilon}}$$

is known as **intrinsic impedance** (and measured in units of ohms).

- Since Maxwell's equations with constant μ and ϵ are linear and time-invariant (LTI), the field solutions above can be further generalized by using their weighted and time-shifted superpositions such as

$$f(t) = \sum_n A_n \cos(\omega_n t + \theta_n)$$

and

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{j\omega t}d\omega$$

having frequency dependent weighting factors A_n and $F(\omega)$. And since according to Fourier analysis all practical signals $f(t)$ can be synthesized in these forms, it follows that the field solutions above are valid with *arbitrary* waveforms $f(t)$.

d'Alembert wave solutions

Solutions

$$\mathbf{E},\, \mathbf{H} \propto f(t \mp \frac{z}{v})$$

of the 1D scalar wave equation with arbitrary $f(t)$ are known as **d'Alembert wave solutions**.

- d'Alembert solution

$$\mathbf{E},\, \mathbf{H} \propto f(t - \frac{z}{v})$$

describes electromagnetic waves traveling in $+z$ direction, whereas solution

$$\mathbf{E},\, \mathbf{H} \propto f(t + \frac{z}{v})$$

describes electromagnetic waves traveling in $-z$ direction (see margin).

In each case the travel speed is

$$v = \frac{1}{\sqrt{\mu\epsilon}} \xrightarrow{\text{free space}} \frac{1}{\sqrt{\mu_o\epsilon_o}} \equiv c \approx 3\times 10^8\,\text{m/s}.$$

- $\mathbf{H}$ solution can be obtained from $\mathbf{E}$ by dividing it with η and rotating it by 90° so that vector $\mathbf{E}\times\mathbf{H}$ points in direction the waves travel.

- $\mathbf{E}$ can be obtained from $\mathbf{H}$ by multiplying it with η and rotating it by 90° so that vector $\mathbf{E}\times\mathbf{H}$ — called **Poynting vector** — once again points in direction the waves travel.

In each case the intrinsic impedance is

$$\eta = \sqrt{\frac{\mu}{\epsilon}} \xrightarrow{\text{free space}} \sqrt{\frac{\mu_o}{\epsilon_o}} \equiv \eta_o \approx 120\pi\,\text{ohms}.$$

Transformation rules above also hold for y**-polarized** wave solutions

$$\mathbf{E} = \hat{y}f(t\mp\frac{z}{v}) \quad\text{and}\quad \mathbf{H} = \mp\hat{x}\frac{f(t\mp\frac{z}{v})}{\eta}.$$

Question: What about z**-polarized** waves

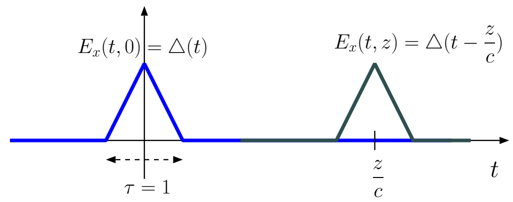
$$\mathbf{E} = \hat{z}f(t\mp\frac{z}{v}),$$

can they exist?

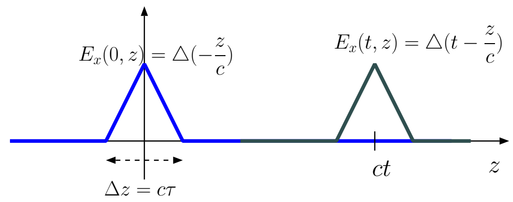
Answer: No, z-polarized waves $\hat{z}f(t\mp\frac{z}{v})$ traveling in $\pm z$ direction cannot exist because they would violate the divergence-free condition $\nabla\cdot\mathbf{E} = 0$.

Traveling wave in +z direction with speed v=c:

Time plots at z=0 and z>0:



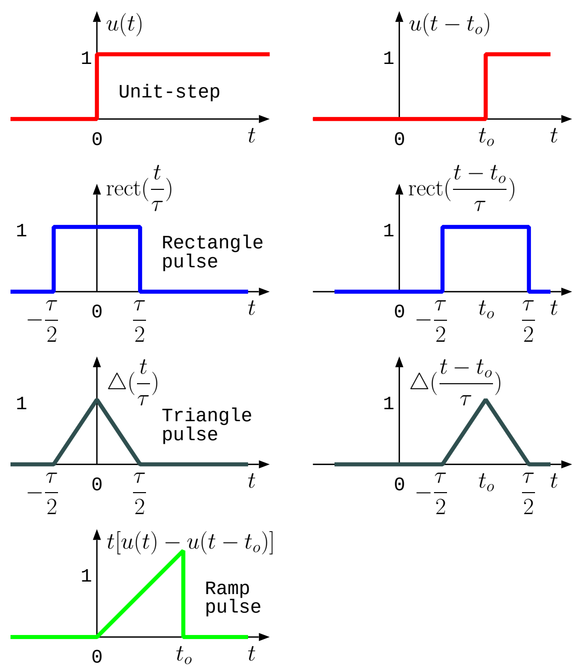
Position plots at t=0 and t>0



Note: ct=300 m in 1 microsec

ct=300 km in 1 millisec

Fundamental signal waveforms: REVIEW



19 d'Alembert wave solutions, radiation from current sheets

- d'Alembert wave solutions of Maxwell's equations for homogeneous and source-free regions obtained in the last lecture having the forms

  $$\mathbf{E},\,\mathbf{H} \propto f(t \mp \frac{z}{v})$$

  are classified as **uniform plane-TEM waves**.

  - TEM stands for **T**ransverse **E**lectro**M**agnetic, and the reason for this designation is:

  viable solutions satisfying $\nabla\cdot\mathbf{E} = \nabla\cdot\mathbf{H} = 0$ conditions have their $\mathbf{E}$ and $\mathbf{H}$ vectors **transverse** to the *direction of propagation* which always coincides with the direction of vector $\mathbf{S} \equiv \mathbf{E}\times\mathbf{H}$ known as **Poynting vector** — more on this later on.

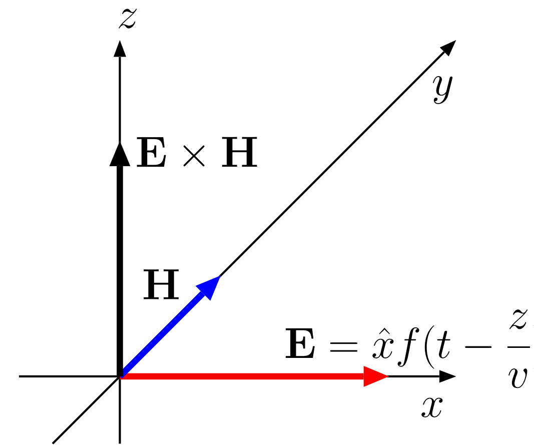
  - d'Alembert wave solutions such as

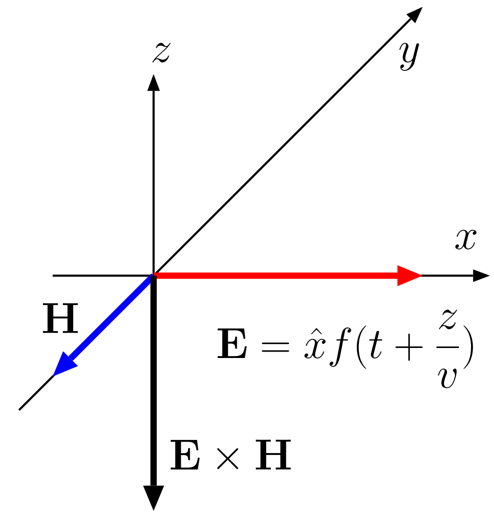
    $$\mathbf{E} = \hat{x}f(t - \frac{z}{v}) \;\; \text{and} \;\; \mathbf{H} = \hat{y}\frac{f(t-\frac{z}{v})}{\eta}$$

    are also designated as *uniform plane* waves because:

  these wave-fields are constant (have the same vector value) at **planes of constant phase**, e.g., on planes defined by

  $$t - \frac{z}{v} = \text{const.},$$





Poynting vector $\mathbf{E}\times\mathbf{H}$

which are planes transverse to the propagation direction (direction of vector $\mathbf{E} \times \mathbf{H}$).

Not all waves solutions of Maxwell's equations are uniform plane — for instance non-uniform TEM waves with spherical surfaces of constant phase are ubiquitous, but they will be examined later on (in ECE 450, mainly).

After the next set of examples we will examine how uniform plane waves can be radiated by infinite planes of surface currents. By contrast, spherical waves are produced by compact antennas having finite dimensions.

Example 1: Let

$$\mathbf{E} = \hat{x}\triangle(\frac{t - y/c}{\tau})$$

be a wave solution in free space where $\triangle(\frac{t}{\tau})$ is a triangular waveform of duration τ peaking at $t = 0$ (defined in ECE 210). We will next provide two different solutions demonstrating how the wave field $\mathbf{B}$ accompanying $\mathbf{E}$ can be found.

Solution 1: We recognize the given wave field $\mathbf{E}$ as a TEM uniform plane wave traveling in y-direction given the $t - y/c$ dependence of phase. Consequently, we obtain $\mathbf{H}$ by dividing $\mathbf{E}$ with $\eta = \eta_o$ and rotating it by 90° from $\hat{x}$-direction to co-align it with $\mathbf{E} \times \mathbf{H}$ vector. As a result,

$$\mathbf{H} = -\hat{z}\frac{\triangle(\frac{t-y/c}{\tau})}{\eta_o} = -\hat{z}\frac{\triangle(\frac{t-y/c}{\tau})}{\sqrt{\mu_o/\epsilon_o}}.$$

Hence,

$$\mathbf{B} = \mu_o\mathbf{H} = -\hat{z}\sqrt{\mu_o\epsilon_o}\triangle(\frac{t - y/c}{\tau}) = -\hat{z}\frac{\triangle(\frac{t-y/c}{\tau})}{c}.$$

Solution 2: According to Faraday's law,

$$\begin{aligned}
\frac{\partial \mathbf{B}}{\partial t} &= -\nabla\times\mathbf{E} = -\begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ E_x & 0 & 0 \end{vmatrix} = \hat{z}\frac{\partial E_x}{\partial y} \\
&= \hat{z}\triangle'(\frac{t-y/c}{\tau})\frac{\partial}{\partial y}(\frac{t-y/c}{\tau}) = \hat{z}\frac{-1}{c\tau}\triangle'(\frac{t-y/c}{\tau})
\end{aligned}$$

with the help of chain rule of differentiation, where $\triangle$'$(u) \equiv \frac{d}{du}\triangle(u)$. Finding the time-dependent anti derivative, we directly obtain (as before)

$$\mathbf{B} = -\hat{z}\frac{\triangle(\frac{t-y/c}{\tau})}{c}.$$

Example 2: Consider the Lorentz force

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v}\times\mathbf{B})$$

on a test charge q in the lab where $\mathbf{E}$ and $\mathbf{B}$ are the plane wave fields considered in Example 1. Show that electrical force term $q\mathbf{E}$ will dominate the magnetic force term $q\mathbf{v}\times\mathbf{B}$ unless the particle speed $v = |\mathbf{v}|$ is close to the speed of light c (i.e., test charge is relativistic).

Solution: Since

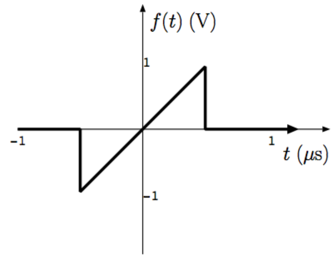
$$\mathbf{E} = \hat{x}\triangle(\frac{t-y/c}{\tau}) \quad \text{and} \quad \mathbf{B} = -\hat{z}\frac{\triangle(\frac{t-y/c}{\tau})}{c},$$

it follows that Lorentz force

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v}\times\mathbf{B}) = q\triangle(\frac{t-y/c}{\tau})(\hat{x} - \mathbf{v}\times\frac{\hat{z}}{c}).$$

Clearly, the first term of $\mathbf{F}$ proportional to $\hat{x}$ is dominant, unless $v = |\mathbf{v}|$ is close to c.

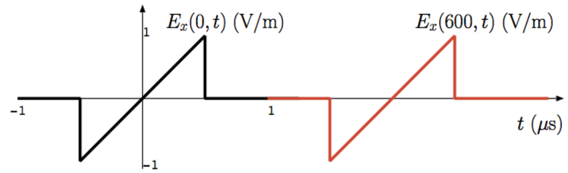
Example 3: Consider an $\hat{x}$-polarized plane TEM wave field in free space propagating in $+z$ direction such that



$$\mathbf{E}(z,t) = \hat{x} f(t - \frac{z}{c}), \text{ with } f(t) = At\,\mathrm{rect}(\frac{t}{\tau}),$$

where $c = 3 \times 10^8\,\mathrm{m/s} = 300\,\mathrm{m}/\mu\mathrm{s}$ is the speed of light in free space, $\tau = 1\,\mu\mathrm{s}$, and $A = 2\,\frac{\mathrm{V/m}}{\mu\mathrm{s}}$. A plot of $f(t)$ vs t (labelled in μs units is shown in the margin. Determine the corresponding $\mathbf{H}(z,t)$ and make the following plots:

- (a) t-plots at fixed z's: $E_x(0,t)$ and $E_x(z = 600\,\mathrm{m}, t)$,
- (b) z-plots at fixed t's: $E_x(z,0)$ and $E_x(z, 2\,\mu\mathrm{s})$.



Solution: (a) t-plots at fixed z's: Since $z/c = 2\,\mu$s for $z = 600$ m, it follows that

$$E_x(600\,\mathrm{m}, t) = 2\,(t - 2\mu\mathrm{s})\,\mathrm{rect}(\frac{t - 2\mu\mathrm{s}}{1\mu\mathrm{s}})\,\frac{\mathrm{V}}{\mathrm{m}}$$

is a shifted version of

$$E_x(0,t) = 2\,t\,\mathrm{rect}(\frac{t}{1\mu\mathrm{s}})\,\frac{\mathrm{V}}{\mathrm{m}}$$

already plotted above. A graph showing both waveforms (black for $z = 0$ and red for $z = 600$ m) is in the margin.

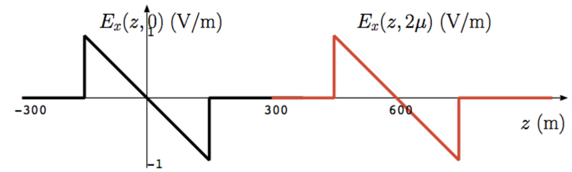
(b) z-plots at fixed t's: In this case we wish to deptict

$$E_x(z,0) = 2\,(0 - \frac{z}{c})\,\mathrm{rect}(\frac{0 - z/c}{1\mu})\,\frac{\mathrm{V}}{\mathrm{m}}$$

and

$$E_x(z, 2\,\mu\mathrm{s}) = 2\,(2\mu - \frac{z}{c})\,\mathrm{rect}(\frac{2\mu - z/c}{1\mu})\,\frac{\mathrm{V}}{\mathrm{m}}.$$

The minus sign in front of z in the first term on the right indicates that the slopes of the curves to be plotted are negative. Hence, we end up with the descending ramp waveforms (black for $t = 0$ and red for $t = 2\mu$s) shown in the margin.



- Plane electromagnetic waves discussed above propagate in free-space in regions of zero ρ and $\mathbf{J}$ (per our derivation).
  - But what generates such waves?
- The answer must be, far away ρ and $\mathbf{J}$ variations (linked by continuity equation) that we have not considered in our equations so far.
- We will next describe how plane TEM waves can be produced — radiated — by time-varying infinite current sheets by starting from familiar static and quasi-static solutions:

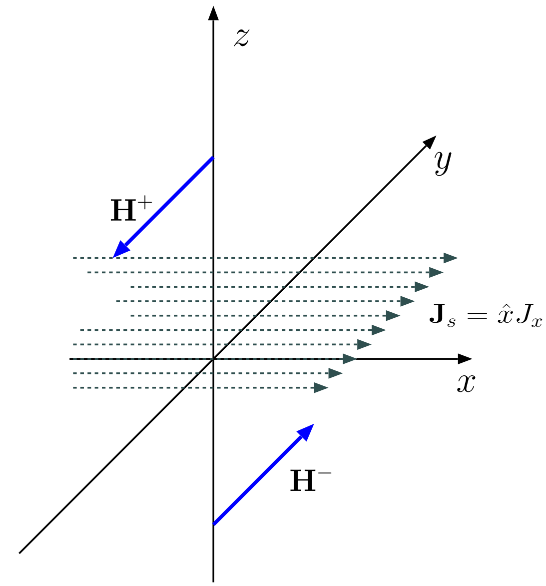
- Consider first a static and constant surface current density

$$\mathbf{J}_s = \hat{x} J_x \,\mathrm{A/m}$$

flowing on $z = 0$ surface as shown graphically in the margin. This infinite surface current will produce a static magnetic field

$$\mathbf{H}(z) = \mp \hat{y} \frac{J_x}{2} \,\mathrm{A/m} \text{ for } z \gtrless 0$$

also shown in the margin as we learned in Lecture 13.



  – Note that the fields point in opposing directions above and below the surface current in compliance with the right hand rule and obey the boundary condition equation for tangential $\mathbf{H}$.
  – Also, $\mathbf{H}$ is not accompanied by an electric field $\mathbf{E}$ since static currents produce only static magnetic fields.

- What if the surface current J_x varies with time, i.e., $J_x = J_x(t)$. In that case we have quasi-statically

$$\mathbf{H}(z, t) \approx \mp \hat{y} \frac{J_x(t)}{2} \,\mathrm{A/m} \text{ for } z \gtrless 0,$$

but only as an approximation for positions very close to $z = 0$ surface where propagation time-delay $\frac{z}{v}$ of d'Alembert solutions can be neglected[1].

[1]This solution surely cannot be an exact solution since if it were, it would imply instantaneous changes in $\mathbf{H}$ in response J_x at arbirarily large distances, implying an infinite speed of propagation — we know that is not true!

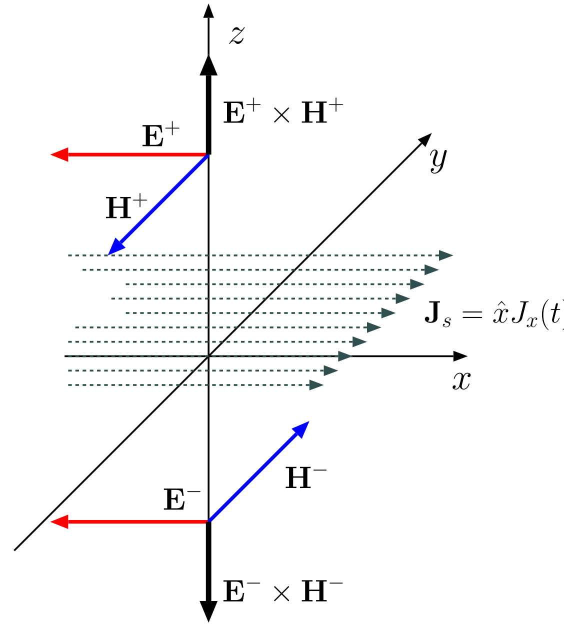
- But the *exact* field solution of Maxwell's equtions valid for all z is equally easy to obtain: just replace $J_x(t)$ above with $J_x(t \mp \frac{z}{v})$ and replace $\approx$ with $=$ so that

$$\mathbf{H}(z,t) = \mp\hat{y}\frac{J_x(t \mp \frac{z}{v})}{2} \text{ A/m for } z \gtrless 0$$

complies with plane TEM d'Alembert solutions[2] of Maxwell's equations in homogeneous and sourece free regions $z \gtrless 0$.

- As always, there is an accompanying $\mathbf{E}(z,t)$ that is obtained by multiplying $\mathbf{H}(z,t)$ with η and replacing its unit vector so that vector $\mathbf{E}\times\mathbf{H}$ points in the direction of propagation, away from the $z = 0$ in this case — hence, as illustrated in the margin,

$$\mathbf{E}(z,t) = -\hat{x}\frac{\eta}{2}J_x(t \mp \frac{z}{v}) \text{ V/m for } z \gtrless 0.$$



Since Maxwell's eqn's + boundary conditions have *unique* solutions in given settings, we are assured that any solution that complies with both (as in this case) is *the* solution for the given setting (surface current on $z = 0$, in this case) — it was surprisingly easy to solve this radiation problem by starting from simple static and quasi-static solutions.

[2]We use $J_x(t \mp \frac{z}{v})$ rather than $J_x(t \pm \frac{z}{v})$ for $z \gtrless 0$ because we assume that $J_x(t)$ on $z = 0$ surface is the only field source — in that case *causality* principle dictates that we use only the solutions propagating *away* from the source (just like when a pebble drops in a pond, ripples propagate *away*).

Conclusion: Evidently, a time varying surface current

$$\mathbf{J}_s = \hat{x} f(t) \text{ on } z = 0 \text{ plane}$$

produces plane electromagnetic waves

$$\mathbf{E}^{\pm} = -\hat{x}\frac{\eta f(t \mp \frac{z}{v})}{2} \text{ and } \mathbf{H}^{\pm} = \mp\hat{y}\frac{f(t \mp \frac{z}{v})}{2} \text{ in regions } z \gtrless 0$$

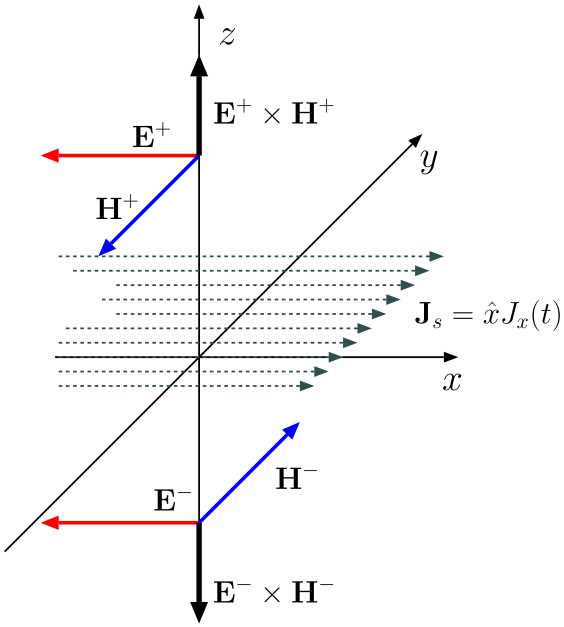
propagating away from the $z = 0$ plane.

Note that:

1. E_x and H_y waveforms are proportional to delayed versions of surface current $J_x(t)$ at each location z above and below the current sheet, with the reference directions of $\mathbf{E}$ and $\mathbf{J_s}$ *opposing* one another.

2. fields $\mathbf{E}^{\pm}$ are continuous on $z = 0$ surface in compliance with tangential boundary condition equations.

3. fields $\mathbf{H}^{\pm}$ exhibit a discontinuity on $z = 0$ surface that matches the current density of the same surface, once again in compliance with tangential boundary condition equations.

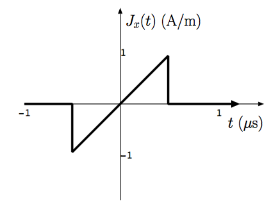


Opposing $\mathbf{E}$ and $\mathbf{J}_s$ vectors on $z = 0$ plane indicate that the surface is acting as a source of radiated energy (the energy that feeds the waves radiated away from the surface) — this interpretation will be discussed in more detail in the next lecture.

Example 4: A current sheet on $z = 0$ surface is described by

$$\mathbf{J}_s(t) = \hat{x} f(t), \quad \text{with} \quad f(t) = At\,\text{rect}(\frac{t}{\tau}),$$

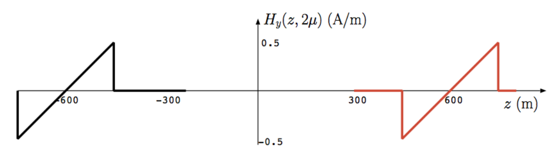
where $\tau = 1\,\mu$s and $A = 1\,\frac{\text{A/m}}{\mu\text{s}}$. A plot of the current waveform $f(t)$ is plotted in the margin. Assuming that the current sheet is embedded in free space, construct the following plots:



- (a) Radiated $H_y(z, t = 2\mu\text{s})$ vs z,
- (b) Radiated $E_x(z, t = 2\mu\text{s})$ vs z.

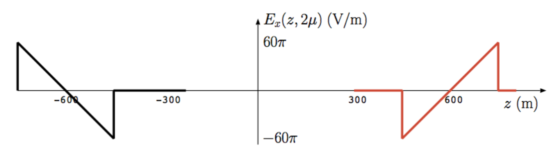
Solution: (a) From the theory developed above, we have using delayed copies of half the surface current density,

$$H_y(z, 2\mu\text{s}) = \mp\frac{1}{2}(2\mu \mp \frac{z}{c})\,\text{rect}(\frac{2\mu \mp \frac{z}{c}}{1\mu})\,\frac{\text{A}}{\text{m}} \text{ for } z \gtrless 0,$$



as plotted in the margin. Notice that the propagated field waveforms — $c \times 2\mu\text{s}$=600 m has been covered in 2 μs — are re-scaled and shifted replicas of the source function $f(t)$.

(b) We have, multiplying H_y with $\eta_o = 120\pi\,\Omega$, and adjusting the signs so that $\mathbf{E}$ and $\mathbf{J_s}$ are pointing in opposite directions,



$$E_x(z, 2\mu\text{s}) = -60\pi(2\mu \mp \frac{z}{c})\,\text{rect}(\frac{2\mu \mp \frac{z}{c}}{1\mu})\,\frac{\text{V}}{\text{m}} \text{ for } z \gtrless 0.$$

Plots are shown in the margin.

20 Poynting theorem and monochromatic waves

- The magnitude of **Poynting vector**

  $$\mathbf{S} = \mathbf{E} \times \mathbf{H}$$

  represents the amount of power transported — often called energy flux — by electromagnetic fields $\mathbf{E}$ and $\mathbf{H}$ over a unit area transverse to the $\mathbf{E} \times \mathbf{H}$ direction.

This interpretation of the Poynting vector is obtained from a conservation law extracted from Maxwell's equations (see margin) as follows:

$$\begin{aligned}
\nabla \cdot \mathbf{D} &= \rho \\
\nabla \cdot \mathbf{B} &= 0 \\
\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\
\nabla \times \mathbf{H} &= \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}.
\end{aligned}$$

1. Dot multiply Faraday's law by $\mathbf{H}$, dot multiply Ampere's law by $\mathbf{E}$,

   $$\begin{aligned}
   (\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}) \cdot \mathbf{H} \\
   (\nabla \times \mathbf{H} &= \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}) \cdot \mathbf{E}
   \end{aligned}$$

   and take their difference:

   $$\underbrace{\mathbf{H} \cdot \nabla \times \mathbf{E} - \mathbf{E} \cdot \nabla \times \mathbf{H}}_{\nabla \cdot (\mathbf{E} \times \mathbf{H})} = \underbrace{-\frac{\partial \mathbf{D}}{\partial t} \cdot \mathbf{E} - \frac{\partial \mathbf{B}}{\partial t} \cdot \mathbf{H}}_{-\frac{\partial}{\partial t}(\frac{1}{2}\epsilon \mathbf{E} \cdot \mathbf{E} + \frac{1}{2}\mu \mathbf{H} \cdot \mathbf{H})} - \mathbf{J} \cdot \mathbf{E}.$$

2. After re-arrangements shown above, the result can be written as

$$\frac{\partial}{\partial t}(\frac{1}{2}\epsilon\mathbf{E}\cdot\mathbf{E}+\frac{1}{2}\mu\mathbf{H}\cdot\mathbf{H})+\nabla\cdot(\mathbf{E}\times\mathbf{H})+\mathbf{J}\cdot\mathbf{E}=0.$$

Poynting theorem

- **Poynting theorem** derived above is a *conservation law* just like the continuity equation $\frac{\partial\rho}{\partial t}+\nabla\cdot\mathbf{J}=0$:

  - The first term on the left,

    $$\frac{\partial}{\partial t}(\frac{1}{2}\epsilon\mathbf{E}\cdot\mathbf{E}+\frac{1}{2}\mu\mathbf{H}\cdot\mathbf{H}),$$

    is time rate of change of total electric and magnetic **energy** density.

  Hence, **Poynting theorem is the conservation law for electromagnetic energy**, just like continuity equation is the conservation law for electric charge.

  - The second term

    $$\nabla\cdot(\mathbf{E}\times\mathbf{H})$$

    accounts for energy transport in Poynting theorem, just like $\nabla\cdot\mathbf{J}$ accounts for charge transport in the continuity equation. Therefore

    $$\mathbf{S}\equiv\mathbf{E}\times\mathbf{H}$$

is **energy flux per unit area** measured in

$$\frac{\mathrm{V}}{\mathrm{m}}\frac{\mathrm{A}}{\mathrm{m}} = \frac{\mathrm{W}}{\mathrm{m}^2} = \frac{\mathrm{J/s}}{\mathrm{m}^2}$$

units, just like $\mathbf{J}$ is charge flux per unit area in $\frac{\mathrm{C/s}}{\mathrm{m}^2} = \frac{\mathrm{A}}{\mathrm{m}^2}$ units.

– Finally, the last term in Poynting theorem (repeated in the margin),

$$\mathbf{J}\cdot\mathbf{E}$$

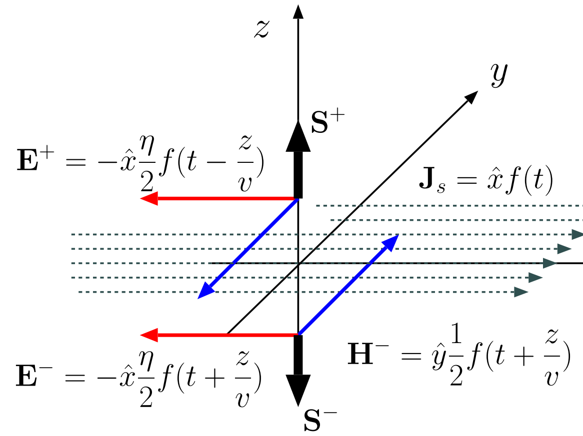
is called **Joule heating**, and it represents power absorbed per unit volume (which can only be non-zero in the presence of $\mathbf{J}$).

If $\mathbf{J}\cdot\mathbf{E}$ is negative in any region, then $\mathbf{J}$ in that region is acting as a source of electromagnetic energy, just like any circuit component with negative vi is acting as an energy source in the electrical circuit.

Note that $\mathbf{J}\cdot\mathbf{E}$ had a negative value on the current sheet radiator examined in last lecture. We return to the current sheet radiator in the next example.

Poynting thm:

$\frac{\partial}{\partial t}(\frac{1}{2}\epsilon\mathbf{E}\cdot\mathbf{E}+\frac{1}{2}\mu\mathbf{H}\cdot\mathbf{H})+\nabla\cdot(\mathbf{E}\times\mathbf{H})+\mathbf{J}\cdot\mathbf{E}=0$



Example 1: On $z = 0$ plane we have a time-harmonic surface current specified as

$$\mathbf{J}_s = \hat{x}f(t) = \hat{x}2\,\cos(\omega t)\,\frac{\mathrm{A}}{\mathrm{m}}$$

where ω is some frequency of oscillation.

(a) Determine the radiated TEM wave fields $\mathbf{E}(z,t)$ and $\mathbf{H}(z,t)$ in the regions $z \gtrless 0$,

(b) The associated Poynting vectors $\mathbf{E} \times \mathbf{H}$, and

(c) $\mathbf{J}_s \cdot \mathbf{E}$ on the current sheet.

Solution: (a) With reference to the solution of the current sheet radiator depicted in the margin (from last lecture), we that an x-polarized surface current $f(t)$ produces the wave fields

$$E_x = -\frac{\eta}{2}f(t \mp \frac{z}{v}) \quad \text{and} \quad H_y = \mp\frac{1}{2}f(t \mp \frac{z}{v})$$

in the surrounding regions propagating away from the current sheet on both sides. Given that $f(t) = 2\cos(\omega t)$, this implies that

$$E_x = -\eta\cos(\omega t \mp \beta z) \quad \text{and} \quad H_y = \mp\cos(\omega t \mp \beta z)$$

where

$$\beta = \frac{\omega}{c} \quad \text{and} \quad \eta = \eta_o \approx 120\pi\,\Omega$$

since the current sheet is surrounded by vacuum. Hence in vector form we have

$$\mathbf{E}(z,t) = -\eta\cos(\omega t \mp \beta z)\hat{x}\,\frac{\mathrm{V}}{\mathrm{m}} \quad \text{and} \quad \mathbf{H}(z,t) = \mp\cos(\omega t \mp \beta z)\hat{y}\,\frac{\mathrm{A}}{\mathrm{m}},$$

where the upper signs are for $z > 0$, and lower signs for $z < 0$.

(b) The associated Poynting vectors are

$$\mathbf{S} = \mathbf{E} \times \mathbf{H} = \pm\eta \cos^2(\omega t \mp \beta z)\hat{z}\,\frac{\mathrm{W}}{\mathrm{m}^2}.$$

Note that the time-average value of vector $\mathbf{S}$ points in the direction of wave propagation on both sides of the current sheet.

(c) Since on $z = 0$ surface of the current sheet the electric field vector is

$$\mathbf{E}(0,t) = -\eta \cos(\omega t)\hat{x}\,\frac{\mathrm{V}}{\mathrm{m}},$$

it follows that $\mathbf{J}_s \cdot \mathbf{E}$ on the same surface is

$$\mathbf{J}_s(t) \cdot \mathbf{E}(0,t) = (\hat{x}2\,\cos(\omega t)\,\frac{\mathrm{A}}{\mathrm{m}}) \cdot (-\eta \cos(\omega t)\hat{x}\,\frac{\mathrm{V}}{\mathrm{m}}) = -2\eta \cos^2(\omega t)\,\frac{\mathrm{W}}{\mathrm{m}^2}.$$

- In the above example, a time-harmonic source current oscillating at some frequency ω produced "monochromatic waves" of radiated fields propagating away from the current sheet on both sides.

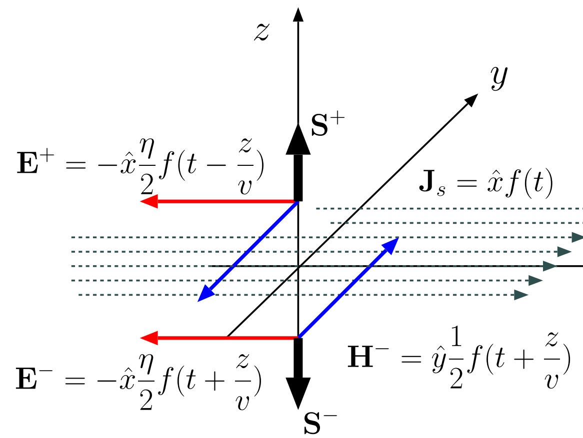
  – The calculations showed time-varying Poynting vectors $\mathbf{E} \times \mathbf{H}$. The time-averaged values of these time-varying vectors can be easily determined by making use of the trig identity

  $$\cos^2(\omega t + \phi) = \frac{1}{2}[1 + \cos(2\omega t + 2\phi)].$$

  Since the time-average of the second term on the right is zero, we can express the time-average of this identity as

$$\langle \cos^2(\omega t + \phi) \rangle = \langle \frac{1}{2}[1 + \cos(2\omega t + 2\phi)] \rangle = \frac{1}{2},$$

where the angular brackets denote the time-averaging procedure.



- Consequently, the result

$$\mathbf{E} \times \mathbf{H} = \pm \eta \cos^2(\omega t \mp \beta z)\hat{z}\,\frac{\mathrm{W}}{\mathrm{m}^2}$$

from Example 1 implies that

$$\langle \mathbf{E} \times \mathbf{H} \rangle = \pm \eta \frac{1}{2}\hat{z}\,\frac{\mathrm{W}}{\mathrm{m}^2} = \pm 60\pi\,\hat{z}\,\frac{\mathrm{W}}{\mathrm{m}^2},$$

which represent the time-average power per unit area transported by the waves radiated by the current sheet.

- In Poynting theorem the Joule heating term $\mathbf{J} \cdot \mathbf{E}$ is **power absorbed per unit volume**, and, accordingly, $-\mathbf{J} \cdot \mathbf{E}$ is **power injected per unit volume**.

  – Likewise, $\pm \mathbf{J}_s \cdot \mathbf{E}$ can be interpreted as **power absorbed/injected per unit area** on a surface.

In Example 1 above we calculated an instantaneous injected power density of

$$-\mathbf{J}_s \cdot \mathbf{E} \;=\; 2\eta \cos^2(\omega t)\,\frac{\mathrm{W}}{\mathrm{m}^2}.$$

Clearly, its time-aveage works out as

$$\langle -\mathbf{J}_s \cdot \mathbf{E}\rangle \;=\; \eta\,\frac{\mathrm{W}}{\mathrm{m}^2} = 120\pi\,\frac{\mathrm{W}}{\mathrm{m}^2}.$$

- Note that $\langle -\mathbf{J}_s \cdot \mathbf{E}\rangle$ exactly matches the sum of $|\langle \mathbf{E} \times \mathbf{H}\rangle|$ calculated on both sides of the current sheet, in conformity with energy conservation principle (Poynting theorem).

21 Monochromatic waves and phasor notation

- Recall that we reached the traveling-wave d'Alembert solutions

$$\mathbf{E}, \mathbf{H} \propto f(t \mp \frac{z}{v})$$

via the superposition of time-shifted and amplitude-scaled versions of

$$f(t) = \cos(\omega t),$$

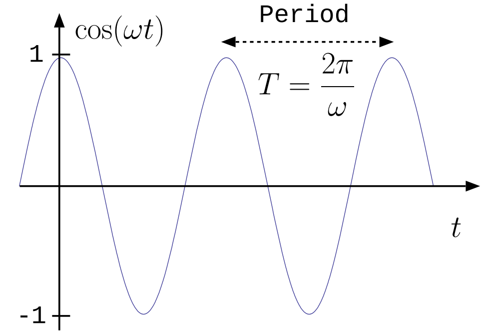
namely the **monochromatic waves**

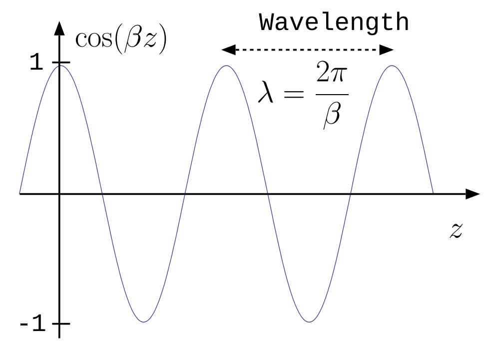
$$A\cos[\omega(t \mp \frac{z}{v})] = A\cos(\omega t \mp \beta z),$$

with amplitudes A where

$$\beta \equiv \frac{\omega}{v} = \omega\sqrt{\mu\epsilon}$$

can be called **wave-number** in analogy with **wave-frequency** ω.





  – As depicted in the margin, monochromatic solutions $A\cos(\omega t \mp \beta z)$ are periodic in position and time, with the *wave-number* β being essentially a *spatial-frequency*, the spatial counterpart of ω.

This is an important point that you should try to understand well — it has implications for signal processing courses related to images and vision.

– *In general*, monochromatic solutions of *1D wave-equations* obtained in various branches of science and engineering can all be represented in the same format as above in terms of wave-frequency / wave-wavenumber pairs ω and β having a ratio

$$v \equiv \frac{\omega}{\beta}$$

recognized as the **wave-speed** and specific **dispersion relations** such as:

1. **TEM waves** in perfect dielectrics:

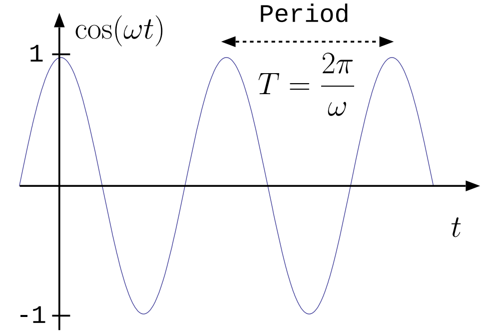
$$\beta = \omega\sqrt{\mu\epsilon},$$

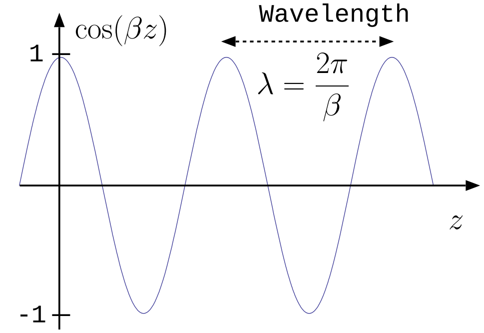
2. **Acoustic waves** in monoatomic gases with temperature T (K) and atomic mass m (kg):

$$\beta = \omega\sqrt{\frac{m}{\frac{5}{3}KT}},$$

3. TEM waves in collisionless **plasmas** (ionized gases) with plasma frequency $\omega_p = \sqrt{\frac{Ne^2}{m\epsilon_o}}$:

$$\beta = \frac{1}{c}\sqrt{\omega^2 - \omega_p^2}.$$





Dispersion relations between *wavefrequency* ω and *wavenumber* β determine the *propagation velocity*

$$v = \frac{\omega}{\beta} = \lambda f$$

for all types of wave motions.

– For any type of wave solution — TEM, acoustic, plasma wave — once the dispersion relation is available (meaning that it has been derived from fundamental physical laws governing the specific wave type), wave propagation velocity is always obtained as

$$v = \frac{\omega}{\beta}$$

or, equivalently, as

$$v = \frac{\lambda}{T} = \lambda f$$

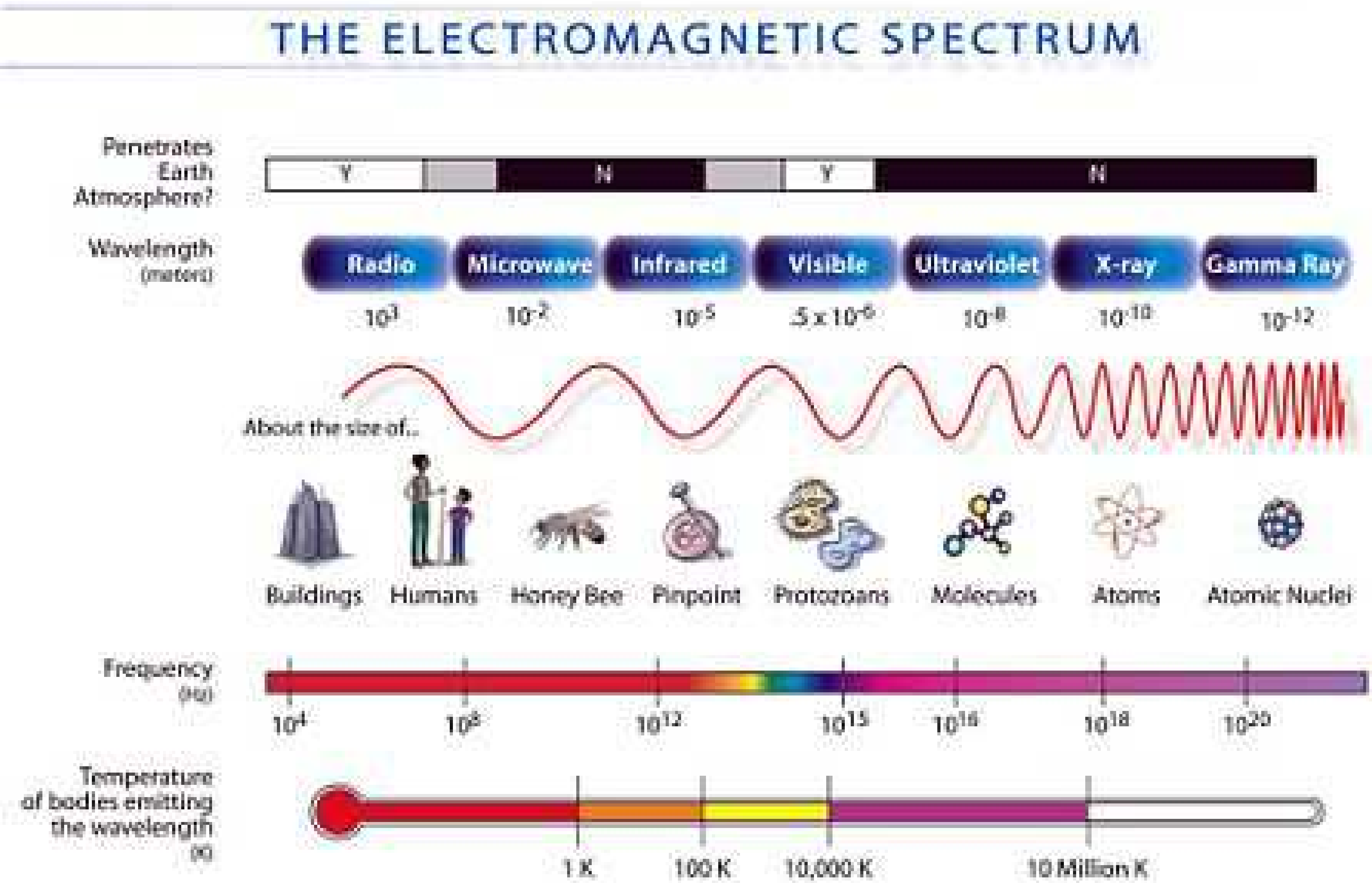
where

$$\lambda \equiv \frac{2\pi}{\beta} \quad \text{Wavelength}$$

and

$$T = \frac{2\pi}{\omega} \equiv \frac{1}{f} \quad \text{Waveperiod.}$$

propagatingWaveCos.eps


THE ELECTROMAGNETIC SPECTRUM
Penetrates Earth Atmosphere?
Y
N
Y
N
Wavelength (meters)
Radio
Microwave
Infrared
Visible
Ultraviolet
X-ray
Gamma Ray
10^3
10^{-2}
10^{-5}
$.5 \times 10^{-6}$
10^{-8}
10^{-10}
10^{-12}
About the size of...
Buildings
Humans
Honey Bee
Pinpoint
Protozoans
Molecules
Atoms
Atomic Nuclei
Frequency (Hz)
10^4
10^8
10^{12}
10^{15}
10^{16}
10^{18}
10^{20}
Temperature of bodies emitting the wavelength (K)
1 K
100 K
10,000 K
10 Million K

- Monochromatic x-polarized waves

$$\mathbf{E} = E_o \cos(\omega t \mp \beta z)\,\hat{x}\,\frac{\mathrm{V}}{\mathrm{m}}$$

can also be expressed in phasor form as

$$\tilde{\mathbf{E}} = E_o e^{\mp j\beta z}\,\hat{x}\,\frac{\mathrm{V}}{\mathrm{m}}$$

such that

$$\mathrm{Re}\{\tilde{\mathbf{E}} e^{j\omega t}\} = E_o \cos(\omega t \mp \beta z)\,\hat{x} = \mathbf{E}$$

in view of Euler's identity.

Example 1: Study the following table to understand monochromatic wave fields and their phasors.

Field	Phasor	Comment
$\mathbf{E} = \cos(\omega t + \beta y)\,\hat{z}$	$\tilde{\mathbf{E}} = e^{j\beta y}\,\hat{z}$	**z-polarized wave propagating in $-y$ direction**
	$\tilde{\mathbf{H}} = -\frac{e^{j\beta y}}{\eta}\,\hat{x}$	**magnetic phasor that accompanies $\tilde{\mathbf{E}}$ above**
$\mathbf{H} = \sin(\omega t - \beta z)\,\hat{y}$	$\tilde{\mathbf{H}} = -je^{-j\beta z}\,\hat{y}$	**wave propagating in $+z$ direction**
$\mathbf{E} = \eta \sin(\omega t - \beta z)\,\hat{x}$	$\tilde{\mathbf{E}} = -j\eta e^{-j\beta z}\,\hat{x}$	**electric field phasor of $\tilde{\mathbf{H}}$ above which is an x-polarized field (see the right column)**

Example 2: Given that

$$\mathbf{H} = \hat{x}H^+ \cos(\omega t - \beta z) + \hat{y}H^- \sin(\omega t + \beta z)$$

representing the sum of wave fields propagating in opposite directions, the corresponding phasor

$$\tilde{\mathbf{H}} = \hat{x}H^+ e^{-j\beta z} - j\hat{y}H^- e^{j\beta z}.$$

The corresponding **E**-field phasor is

$$\tilde{\mathbf{E}} = -\hat{y}\eta H^+ e^{-j\beta z} + j\hat{x}\eta H^- e^{j\beta z},$$

from which

$$\mathbf{E} = -\hat{y}\eta H^+ \cos(\omega t - \beta z) - \hat{x}\eta H^- \sin(\omega t + \beta z).$$

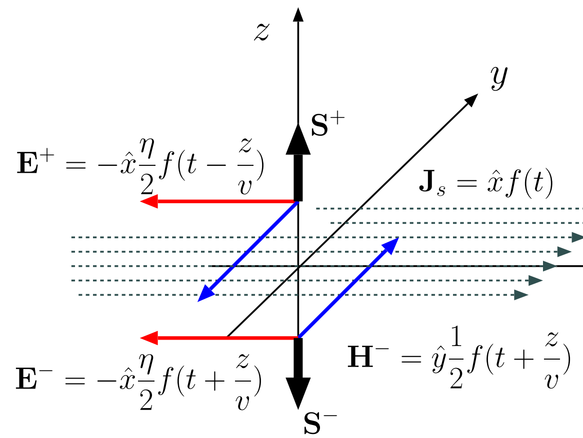
Make sure to check that all the signs make sense, and if you think you have caught an error, let us know.

- In general, we transform between plane TEM wave phasors $\tilde{\mathbf{E}}$ and $\tilde{\mathbf{H}}$ as follows:

1. **To obtain $\tilde{\mathbf{H}}$ from $\tilde{\mathbf{E}}$**: divide $\tilde{\mathbf{E}}$ by η and rotate the vector direction so that vector $\tilde{\mathbf{S}} \equiv \tilde{\mathbf{E}} \times \tilde{\mathbf{H}}^*$ points in the propagation direction of the wave — more on complex vector $\tilde{\mathbf{S}}$ later on.

2. **To obtain $\tilde{\mathbf{E}}$ from $\tilde{\mathbf{H}}$**: multiply $\tilde{\mathbf{H}}$ by η and rotate the vector direction so that vector $\tilde{\mathbf{S}} \equiv \tilde{\mathbf{E}} \times \tilde{\mathbf{H}}^*$ points in the propagation direction of the

wave.



Example 3: On $z = 0$ plane we have a monochromatic surface current specified as

$$\mathbf{J}_s = \hat{x}f(t) = \hat{x}2\,\cos(\omega t)\,\frac{\mathrm{A}}{\mathrm{m}} = \mathrm{Re}\{\hat{x}2\,e^{j\omega t}\}.$$

Determine wave field phasors $\tilde{\mathbf{E}}^{\pm}$ and $\tilde{\mathbf{H}}^{\pm}$ for plane TEM waves propagating away from the $z = 0$ surface on both sides (assumed vacuum).

Solution: We know that an x-polarized surface current $f(t)$ produces

$$E_x = -\frac{\eta}{2}f(t \mp \frac{z}{v}) \quad \text{and} \quad H_y = \mp\frac{1}{2}f(t \mp \frac{z}{v})$$

in surrounding regions. Given that $f(t) = 2\cos(\omega t)$, this implies

$$E_x = -\eta\cos(\omega t \mp \beta z) \quad \text{and} \quad H_y = \mp\cos(\omega t \mp \beta z)$$

where

$$\beta = \frac{\omega}{c} \quad \text{and} \quad \eta = \eta_o \approx 120\pi\,\Omega$$

since the current sheet is surrounded by vacuum. Converting these into phasors, we find

$$\tilde{\mathbf{E}}^{\pm} = -\eta e^{\mp j\beta z}\hat{x} \quad \text{and} \quad \tilde{\mathbf{H}}^{\pm} = \mp e^{\mp j\beta z}\hat{y}.$$

- In the last lecture we calculated the time-average $\mathbf{E}\times\mathbf{H}$ and $\mathbf{J}_s\cdot\mathbf{E}$ of the fields examined in Example 3 using a time-domain approach. The same calculations can be carried out in terms of phasors $\tilde{\mathbf{E}}$, $\tilde{\mathbf{H}}$, and $\tilde{\mathbf{J}}_s$ as follows:

$$\langle\mathbf{E}\times\mathbf{H}\rangle = \frac{1}{2}\mathrm{Re}\{\tilde{\mathbf{E}}\times\tilde{\mathbf{H}}^*\} \quad \text{and} \quad \langle\mathbf{J}_s\cdot\mathbf{E}\rangle = \frac{1}{2}\mathrm{Re}\{\tilde{\mathbf{J}}_s\cdot\tilde{\mathbf{E}}^*\}$$

where $\tilde{\mathbf{E}}\times\tilde{\mathbf{H}}^* \equiv \tilde{\mathbf{S}}$ is called complex Poynting vector.

  – The proof of these are analogous to the proof of

$$\langle p(t)\rangle = \frac{1}{2}\mathrm{Re}\{VI^*\}$$

for the average power of a circuit component in terms of voltage and current phasors V and I (see margin).

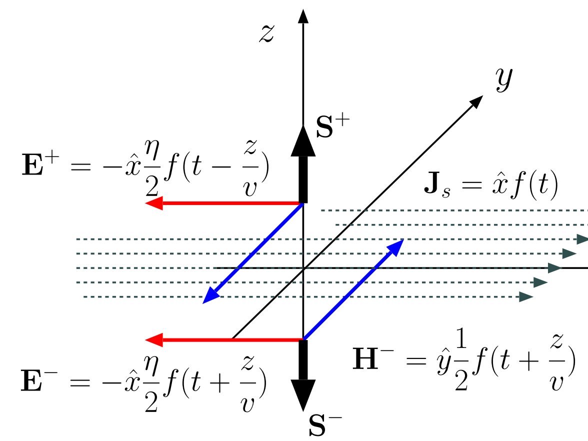
For, instance, given that

$$\tilde{\mathbf{J}}_s = 2\hat{x}\,\frac{\mathrm{A}}{\mathrm{m}} \quad \text{and} \quad \tilde{\mathbf{E}}^{\pm}(z) = -\eta e^{\mp j\beta z}\hat{x}\,\frac{\mathrm{V}}{\mathrm{m}}$$

in Example 3, it follows that

$$\langle -\mathbf{J}_s(t)\cdot\mathbf{E}(0,t)\rangle = \frac{1}{2}\mathrm{Re}\{-\tilde{\mathbf{J}}_s\cdot\tilde{\mathbf{E}}^*(0)\} = \eta \approx 120\pi\,\frac{\mathrm{W}}{\mathrm{m}^2},$$

in conformity with the result from last lecture.



Instantaneous power

$$p(t) = v(t)i(t)$$

with time-harmonic signals is

$$v(t)i(t) = \left(\frac{Ve^{j\omega t}+\mathrm{cc}}{2}\right)\left(\frac{Ie^{j\omega t}+\mathrm{cc}}{2}\right)$$

where V and I are phasors of $v(t)$ and $i(t)$ and cc indicates the conjugate of the term to the left of + sign.
This can be expanded as

$$v(t)i(t) = \frac{VI^*+\mathrm{cc}}{4} + \frac{VIe^{j2\omega t}+\mathrm{cc}}{4}.$$

The second term has a zero time average. It follows that *time-average power*

$$\langle v(t)i(t)\rangle = \frac{VI^*+\mathrm{cc}}{4} = \frac{1}{2}\mathrm{Re}\{VI^*\}$$

since

$$VI^*+\mathrm{cc} = VI^*+V^*I = 2\mathrm{Re}\{VI^*\}.$$

(Also see ECE 210 text.)

22 Phasor form of Maxwell's equations and damped waves in conducting media

- When the fields and the sources in Maxwell's equations are all monochromatic functions of time expressed in terms of their phasors, Maxwell's equations can be transformed into the phasor domain.

$$\begin{aligned} \nabla \cdot \mathbf{D} &= \rho \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{H} &= \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}. \end{aligned}$$

  - In the phasor domain all

$$\frac{\partial}{\partial t} \rightarrow j\omega$$

    and all variables $\mathbf{D}$, ρ, etc. are replaced by their phasors $\tilde{\mathbf{D}}$, $\tilde{\rho}$, and so on. With those changes Maxwell's equations take the form shown in the margin.

$$\begin{aligned} \nabla \cdot \tilde{\mathbf{D}} &= \tilde{\rho} \\ \nabla \cdot \tilde{\mathbf{B}} &= 0 \\ \nabla \times \tilde{\mathbf{E}} &= -j\omega\tilde{\mathbf{B}} \\ \nabla \times \tilde{\mathbf{H}} &= \tilde{\mathbf{J}} + j\omega\tilde{\mathbf{D}} \end{aligned}$$

  - Also in these equations it is implied that

$$\begin{aligned} \tilde{\mathbf{D}} &= \epsilon\tilde{\mathbf{E}} \\ \tilde{\mathbf{B}} &= \mu\tilde{\mathbf{H}} \\ \tilde{\mathbf{J}} &= \sigma\tilde{\mathbf{E}} \end{aligned}$$

    where ϵ, μ, and σ could be a function of frequency ω (as, strictly speaking, they all are as seen in Lecture 11).

  - We can derive from the phasor form Maxwell's equations shown in the margin the TEM wave properties obtained earlier on using the time-domain equations by assuming $\tilde{\rho} = \tilde{\mathbf{J}} = 0$.

We will do that, and and after that relax the requirement $\tilde{\mathbf{J}} = 0$ with $\tilde{\mathbf{J}} = \sigma\tilde{\mathbf{E}}$ to examine how TEM waves propagate in conducting media.

- With $\tilde{\rho} = \tilde{\mathbf{J}} = 0$ the phasor form Maxwell's equation take their simplified forms shown in the margin.

$$\begin{aligned} \nabla \cdot \tilde{\mathbf{E}} &= 0 \\ \nabla \cdot \tilde{\mathbf{H}} &= 0 \\ \nabla \times \tilde{\mathbf{E}} &= -j\omega\mu\tilde{\mathbf{H}} \\ \nabla \times \tilde{\mathbf{H}} &= j\omega\epsilon\tilde{\mathbf{E}} \end{aligned}$$

  - Using

$$\nabla \times [\nabla \times \tilde{\mathbf{E}} = -j\omega\mu\tilde{\mathbf{H}}] \;\Rightarrow\; -\nabla^2\tilde{\mathbf{E}} = -j\omega\mu\nabla \times \tilde{\mathbf{H}}$$

    which combines with the Ampere's law to produce

$$\nabla^2\tilde{\mathbf{E}} + \omega^2\mu\epsilon\tilde{\mathbf{E}} = 0.$$

  - For x-polarized waves with phasors

$$\tilde{\mathbf{E}} = \hat{x}\tilde{E}_x(z),$$

    the phasor wave equation above simplifies as

$$\frac{\partial^2}{\partial z^2}\tilde{E}_x + \omega^2\mu\epsilon\tilde{E}_x = 0.$$

  - Try solutions of the form

$$\tilde{E}_x(z) = e^{-\gamma z} \text{ or } e^{\gamma z}$$

    where γ is to be determined.

– Upon substitution into wave equation both of these lead to

$$(\gamma^2 + \omega^2 \mu\epsilon)\tilde{E}_x = 0,$$

which yields

$$\gamma^2 + \omega^2\mu\epsilon = 0 \quad \Rightarrow \quad \gamma^2 = -\omega^2\mu\epsilon$$

from which one possibility is

$$\gamma = j\beta, \quad \text{with} \quad \beta \equiv \omega\sqrt{\mu\epsilon}.$$

Thus viable phasor solutions are

$$\tilde{E}_x(z) = e^{\mp j\beta z}$$

as we already knew.

– Furthermore, using the phasor form Faraday's law it is easy to show that

$$\tilde{H}_y = \pm\frac{e^{\mp j\beta z}}{\eta} \quad \text{with} \quad \eta = \sqrt{\frac{\mu}{\epsilon}}.$$

Note that we have recovered above the familiar properties of plane TEM waves using phasor methods.

Next, the phasor method carries us to a new domain that cannot be easily examined using time-domain methods.

- With $\tilde{\rho} = 0$ but $\tilde{\mathbf{J}} = \sigma\tilde{\mathbf{E}} \neq 0$, implying non-zero conductivity σ, the pertinent phasor form equations are as shown in the margin.

  - This is the same set as before, except that

    $$j\omega\epsilon \text{ has been replaced by } \sigma + j\omega\epsilon.$$

    Thus, we can make use of phasor wave solutions above after applying the following modifications to γ and η:

$$\begin{aligned}\nabla\cdot\tilde{\mathbf{E}} &= 0\\ \nabla\cdot\tilde{\mathbf{H}} &= 0\\ \nabla\times\tilde{\mathbf{E}} &= -j\omega\mu\tilde{\mathbf{H}}\\ \nabla\times\tilde{\mathbf{H}} &= \sigma\tilde{\mathbf{E}} + j\omega\epsilon\tilde{\mathbf{E}}\\ &= (\sigma + j\omega\epsilon)\tilde{\mathbf{E}}\end{aligned}$$

1.

$$\gamma^2 = -\omega^2\mu\epsilon = (j\omega\mu)(j\omega\epsilon) \ \boxed{\begin{matrix}\Rightarrow\Rightarrow\\ \sigma\neq 0\end{matrix}} \ \gamma = \sqrt{(j\omega\mu)(\sigma + j\omega\epsilon)}$$

2.

$$\eta = \sqrt{\frac{\mu}{\epsilon}} = \sqrt{\frac{j\omega\mu}{j\omega\epsilon}} \ \boxed{\begin{matrix}\Rightarrow\Rightarrow\\ \sigma\neq 0\end{matrix}} \ \eta = \sqrt{\frac{j\omega\mu}{\sigma + j\omega\epsilon}}.$$

Note that the modified γ and η satisfy

$$\gamma\eta = j\omega\mu \text{ and } \frac{\gamma}{\eta} = \sigma + j\omega\epsilon$$

leading to useful relations shown in the margin (assuming real valued σ and ϵ).

$$\begin{aligned}\mu &= \frac{\gamma\eta}{j\omega}\\ \sigma &= \mathrm{Re}\{\frac{\gamma}{\eta}\}\\ \epsilon &= \frac{1}{\omega}\mathrm{Im}\{\frac{\gamma}{\eta}\}\end{aligned}$$

- In terms of γ and η above, we can express an x-polarized plane wave propagating in z direction in terms of phasors

$$\tilde{\mathbf{E}} = \hat{x}E_o e^{\mp\gamma z} \text{ and } \tilde{\mathbf{H}} = \pm\hat{y}\frac{E_o}{\eta}e^{\mp\gamma z}$$

where E_o is an arbitrary complex constant (complex wave amplitude).

- In expanded forms γ and η appear as:

$$\gamma = \sqrt{(j\omega\mu)(\sigma + j\omega\epsilon)} \equiv \alpha + j\beta, \text{ so that } \alpha = \mathrm{Re}\{\gamma\} \text{ and } \beta = \mathrm{Im}\{\gamma\},$$

and

$$\eta = \sqrt{\frac{j\omega\mu}{\sigma + j\omega\epsilon}} \equiv |\eta|e^{j\tau} \text{ so that } |\eta| = \left|\sqrt{\frac{j\omega\mu}{\sigma + j\omega\epsilon}}\right| \text{ and } \tau = \angle\sqrt{\frac{j\omega\mu}{\sigma + j\omega\epsilon}}.$$

1. In the special case of a **perfect dielectric** with $\sigma = 0$, we find

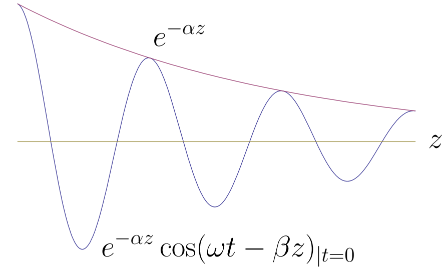
$$\gamma = j\omega\sqrt{\mu\epsilon} \equiv j\beta \text{ and } \eta = \sqrt{\frac{\mu}{\epsilon}},$$

and, therefore,

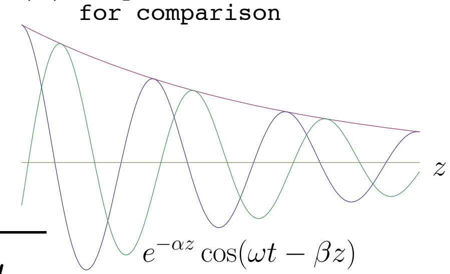
$$\tilde{\mathbf{E}} = \hat{x}E_o e^{\mp j\beta z} \text{ and } \tilde{\mathbf{H}} = \pm\frac{\hat{y}E_o e^{\mp j\beta z}}{\eta}$$

as before. In this case $\alpha = \tau = 0$.

(a) Damped wave snapshot at t=0 together with exponential envelope



(b) Snaphot at t>0, with t=0 waveform for comparison



- **β appears within cosine argument and determines the wavelength**

$$\lambda = \frac{2\pi}{\beta}$$

and propagation speed

$$v_p = \frac{\omega}{\beta}.$$

- **α controls wave attenuation by**

$$e^{\mp\alpha z}$$

factor in propagation direction.

2. Another case of **imperfect dielectric** (or "lousy" conductor) occurs when σ is not zero, but it is so small that are justified in using

$$(1 \pm a)^p \approx 1 \pm pa, \ \text{ if } \ |a| \ll 1,$$

with $p = \frac{1}{2}$ as follows: For $\frac{\sigma}{\omega\epsilon} \ll 1$,

$$\gamma = \sqrt{(j\omega\mu)(\sigma + j\omega\epsilon)} = j\omega\sqrt{\mu\epsilon}(1 - j\frac{\sigma}{\omega\epsilon})^{1/2} \approx j\omega\sqrt{\mu\epsilon}(1 - j\frac{\sigma}{2\omega\epsilon}) = \frac{\sigma}{2}\sqrt{\frac{\mu}{\epsilon}} + j\omega\sqrt{\mu\epsilon};$$

hence

$$\tilde{\mathbf{E}} \approx \hat{x}E_o e^{\mp(\alpha + j\beta)z} \ \text{ with } \ \alpha = \frac{\sigma}{2}\sqrt{\frac{\mu}{\epsilon}} \ \text{ and } \ \beta = \omega\sqrt{\mu\epsilon};$$

also in the same case

$$\tilde{\mathbf{H}} \approx \pm\frac{\hat{y}E_o e^{\mp(\alpha + j\beta)z}}{\eta} \ \text{ with } \ \eta = \sqrt{\frac{\mu}{\epsilon(1 - j\frac{\sigma}{\omega\epsilon})}} \approx \sqrt{\frac{\mu}{\epsilon}}(1 + j\frac{\sigma}{2\omega\epsilon}) \approx \sqrt{\frac{\mu}{\epsilon}}e^{j\tan^{-1}\frac{\sigma}{2\omega\epsilon}},$$

such that

$$|\eta| \approx \sqrt{\frac{\mu}{\epsilon}} \ \text{ and } \ \tau = \angle\eta \approx \frac{\sigma}{2\omega\epsilon}.$$

Note: γ and η both are *complex* valued, the consequences of which will be examined later on.

3. A third case of **good conductor** corresponds to $\frac{\sigma}{\omega\epsilon} \gg 1$. In that case,

$$\gamma = j\omega\sqrt{\mu\epsilon(1 - j\frac{\sigma}{\omega\epsilon})} \approx \omega\sqrt{j\mu\frac{\sigma}{\omega}} = (1+j)\sqrt{\frac{\omega\mu\sigma}{2}} \ \text{ and } \ \eta \approx \sqrt{\frac{\mu}{-j\frac{\sigma}{\omega}}} = \sqrt{\frac{j\omega\mu}{\sigma}} = \sqrt{\frac{\omega\mu}{\sigma}}e^{j\pi/4}.$$

Hence,

$$\alpha \approx \beta \approx \sqrt{\frac{\omega\mu\sigma}{2}} = \sqrt{\pi f \mu\sigma} \text{ while } |\eta| = \sqrt{\frac{\omega\mu}{\sigma}} \text{ and } \tau = \angle\eta = 45^o.$$

4. Finally, **perfect conductor** case corresponds to $\sigma \to \infty$, in which case $\tilde{E}_x \to 0$ as we will show later on. Wave fields cannot exist in perfect conductors.

- Summarizing, in a **homogeneous medium** with arbitrary but constant μ, ϵ, and σ, time-harmonic plane TEM waves are in terms of

$$\mathbf{E} = \hat{x}\text{Re}\{E_o e^{\mp(\alpha+j\beta)z} e^{j\omega t}\} = \hat{x}|E_o|e^{\mp\alpha z}\cos(\omega t \mp \beta z + \angle E_o)$$

and accompanying magnetic fields

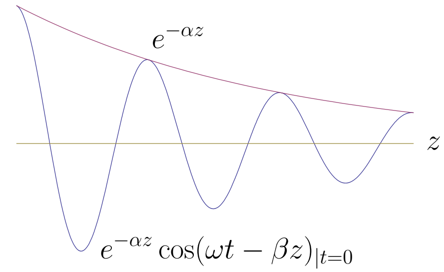
$$\mathbf{H} = \pm\hat{y}\text{Re}\{\frac{E_o}{\eta} e^{\mp(\alpha+j\beta)z} e^{j\omega t}\} = \pm\hat{y}\frac{|E_o|}{|\eta|}e^{\mp\alpha z}\cos(\omega t \mp \beta z + \angle E_o - \angle\eta).$$

- **Propagation velocity**

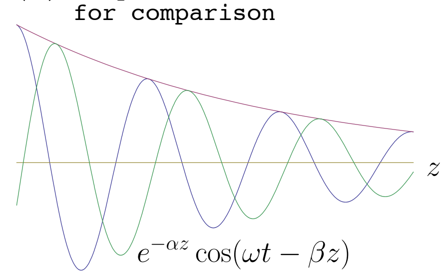
$$v_p = \frac{\omega}{\beta} = \frac{\omega}{\text{Im}\{\sqrt{(j\omega\mu)(\sigma + j\omega\epsilon)}\}},$$

now depends on frequency ω and it describes the speed of the **nodes** (zero-crossings, not modified by the attenuation factor) of the field waveform. Subscript p is introduced to distinguish v_p — also called *phase velocity* — from *group velocity* v_g discussed in ECE 450 (velocity of narrowband wave packets).

(a) Damped wave snapshot at t=0
together with exponential envelope



(b) Snaphot at t>0, with t=0 waveform
for comparison



- **β appears within cosine argument and determines the wavelength**

$$\lambda = \frac{2\pi}{\beta}$$

and propagation speed

$$v_p = \frac{\omega}{\beta}.$$

- **α controls wave attenuation by**

$$e^{\mp\alpha z}$$

factor in propagation direction.

- **Wavelength**

$$\lambda = \frac{2\pi}{\beta} = \frac{v_p}{f}$$

now depends on frequency f via both the numerator and the denominator, and measures twice the distance between successive nodes of the waveform.

- **Penetration depth** (also called *skin depth* if very small)

$$\delta \equiv \frac{1}{\alpha} = \frac{1}{\mathrm{Re}\{\sqrt{(j\omega\mu)(\sigma + j\omega\epsilon)}\}}$$

is the distance for the field strength to be reduced by e^{-1} factor in its direction of propagation.

  – For a fixed σ, and a sufficiently large ω, the penetration depth

$$\delta \approx \frac{2}{\sigma\sqrt{\frac{\mu}{\epsilon}}} \text{ Imperfect dielectric formula}$$

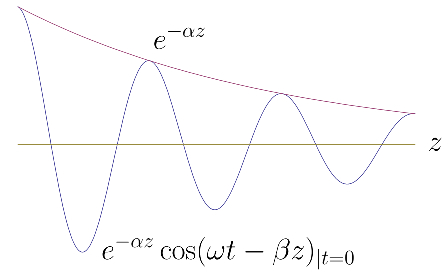
  which can be very small if σ is large — with small δ the wave is severely attenuated as it propagates.

  – For a fixed σ, and a sufficiently small ω,

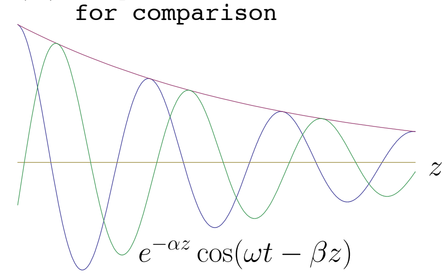
$$\delta \approx \sqrt{\frac{2}{\mu\omega\sigma}} = \frac{1}{\sqrt{\pi f \mu\sigma}} \text{ Good conductor "skin depth" formula}$$

  which, although small with large σ, increases as ω decreases, making *low frequencies to be preferable* in applications requiring propagating through lossy media with large σ, such as in sea-water.

(a) Damped wave snapshot at t=0 together with exponential envelope



(b) Snaphot at t>0, with t=0 waveform for comparison



– **β appears within cosine argument and determines the wavelength**

$$\lambda = \frac{2\pi}{\beta}$$

and propagation speed

$$v_p = \frac{\omega}{\beta}.$$

– **α controls wave attenuation by**

$$e^{\mp\alpha z}$$

factor in propagation direction.

23 Imperfect dielectrics, good conductors

	Condition	β	α	$\lvert\eta\rvert$	τ	$\lambda = \frac{2\pi}{\beta}$	$\delta = \frac{1}{\alpha}$
Perfect dielectric	$\sigma = 0$	$\omega\sqrt{\epsilon\mu}$	0	$\sqrt{\frac{\mu}{\epsilon}}$	0	$\frac{2\pi}{\omega\sqrt{\epsilon\mu}}$	∞
Imperfect dielectric	$\frac{\sigma}{\omega\epsilon} \ll 1$	$\sim \omega\sqrt{\epsilon\mu}$	$\beta\frac{1}{2}\frac{\sigma}{\omega\epsilon} = \frac{\sigma}{2}\sqrt{\frac{\mu}{\epsilon}}$	$\sim \sqrt{\frac{\mu}{\epsilon}}$	$\sim \frac{\sigma}{2\omega\epsilon}$	$\sim \frac{2\pi}{\omega\sqrt{\epsilon\mu}}$	$\frac{2}{\sigma}\sqrt{\frac{\epsilon}{\mu}}$
Good conductor	$\frac{\sigma}{\omega\epsilon} \gg 1$	$\sim \sqrt{\pi f\mu\sigma}$	$\sim \sqrt{\pi f\mu\sigma}$	$\sqrt{\frac{\omega\mu}{\sigma}}$	$45°$	$\sim \frac{2\pi}{\sqrt{\pi f\mu\sigma}}$	$\sim \frac{1}{\sqrt{\pi f\mu\sigma}}$
Perfect conductor	$\sigma = \infty$	∞	∞	0	-	0	0

- The table above summarizes TEM wave parameters in homogeneous conducting media where the propagation velocity

$$v_p = \frac{\omega}{\beta}$$

(note that it can be frequency dependent) and field phasors can be expressed in formats similar to that shown in the margin, keeping in mind that propagation direction coincides with vector

$$\tilde{\mathbf{S}} \equiv \tilde{\mathbf{E}} \times \tilde{\mathbf{H}}^*$$

such that

$$\langle \mathbf{S} \rangle = \langle \mathbf{E} \times \mathbf{H} \rangle = \frac{1}{2}\text{Re}\{\tilde{\mathbf{S}}\}$$

is the average energy flux per unit area (time-average Poynting vector).

x-polarized phasor

$\tilde{\mathbf{E}} = \hat{x}E_o e^{\mp\alpha z}e^{\mp j\beta z}$

accompanied by

$\tilde{\mathbf{H}} = \pm\hat{y}\frac{E_o}{\eta}e^{\mp\alpha z}e^{\mp j\beta z}$.

Example 1: Consider the plane TEM wave

$$\tilde{\mathbf{E}} = \hat{y}2e^{-\alpha z}e^{-j\beta z}\,\frac{\mathrm{V}}{\mathrm{m}},$$

in an **imperfect dielectric**. Determine $\tilde{\mathbf{H}}$ and time-average Poynting vector $\langle\mathbf{S}\rangle$. Compute $\langle\mathbf{S}\rangle$ at $z = 0$ and $z = 10$ m, if $\epsilon = 4\epsilon_o$, $\mu = \mu_o$, $\sigma = 10^{-3}$ S/m, and $\omega = 2\pi \cdot 10^9$ rad/s

Solution: Using right hand rule, so that $\mathbf{E}\times\mathbf{H}$ points in propagation direction $\hat{z}$, we find that

$$\tilde{\mathbf{H}} = -\hat{x}\frac{2}{\eta}e^{-\alpha z}e^{-j\beta z} \approx -\hat{x}\frac{2}{\sqrt{\mu/\epsilon}}e^{-\alpha z}e^{-j\beta z}e^{-j\tau}\,\frac{\mathrm{A}}{\mathrm{m}}$$

using $|\eta| = \sqrt{\frac{\mu}{\epsilon}}$ from the table above for a perfect dielectric.

The avg. Poynting vector is

$$\begin{aligned}\langle\mathbf{S}\rangle &= \frac{1}{2}\mathrm{Re}\{\tilde{\mathbf{E}}\times\tilde{\mathbf{H}}^*\} = \frac{1}{2}\mathrm{Re}\{\hat{y}2e^{-\alpha z}e^{-j\beta z}\times(-\hat{x}\frac{2}{\sqrt{\mu/\epsilon}}e^{-\alpha z}e^{-j\beta z}e^{-j\tau})^*\}\\ &= -\frac{1}{2}\mathrm{Re}\{\hat{y}2e^{-\alpha z}\times\hat{x}\frac{2}{\sqrt{\mu/\epsilon}}e^{-\alpha z}e^{j\tau}\} = \hat{z}\frac{2}{\sqrt{\mu/\epsilon}}e^{-2\alpha z}\cos\tau.\end{aligned}$$

With the given parameters,

$$\begin{aligned}\frac{\sigma}{\omega\epsilon} &= \frac{10^{-3}\cdot 36\pi\times 10^9}{2\pi\cdot 10^9\cdot 4} = \frac{9}{2}10^{-3} \ll 1,\\ \tau &\approx \frac{\sigma}{2\omega\epsilon} \approx \frac{9}{4}10^{-3}\,\mathrm{rad}\\ |\eta| &\approx \sqrt{\frac{\mu}{\epsilon}} = \sqrt{\frac{\mu_o}{4\epsilon_o}} = \frac{\eta_o}{2} = 60\pi\,\Omega\\ \alpha &\approx \frac{\sigma}{2}\sqrt{\frac{\mu}{\epsilon}} = \frac{1}{2}10^{-3}60\pi = 30\pi\cdot 10^{-3}\,\frac{1}{\mathrm{m}}.\end{aligned}$$

Hence, at $z=0$,

$$\langle \mathbf{S} \rangle = \hat{z}\frac{2}{\sqrt{\mu/\epsilon}}\cos\tau \approx \hat{z}\frac{2}{60\pi} = \hat{z}\frac{1}{30\pi}\frac{\mathrm{W}}{\mathrm{m}^2},$$

whereas, at $z=10$ m,

$$\langle \mathbf{S} \rangle = \hat{z}\frac{2}{\sqrt{\mu/\epsilon}}e^{-2\cdot 30\pi\cdot 10^{-3}\cdot 10}\cos\tau \approx \hat{z}\frac{2}{60\pi}e^{-6\pi/10} \approx \hat{z}\frac{0.15}{30\pi}\frac{\mathrm{W}}{\mathrm{m}^2}.$$

- Note that in above example power transmitted per unit area has dropped to 15% of its value upon propapagating over a relatively short distance of 10 m.

    - In the physical terms, the lost power of the wave is gained by the propagation medium in the form of heat — average Joule heating $\langle \mathbf{J}\cdot\mathbf{E}\rangle$ in the medium will be positive and account for the loss of the wave power (as seen in a HW problem).

⇐This is what we want to happen in a microwave oven.

From a communications perspective, this rapid attenuation is problematic since it is evident that the signal energy is being wasted as heat in the medium rather than being transmitted efficiently to distant communication targets.

As the next example shows, we are better off using lower frequencies in under-water communcations.

Example 2: Repeat Example 1 for $\omega = 2\pi \cdot 10^3$ rad/s and $\sigma = 4$ S/m (sea water) in which case the propagation medium becomes a **good conductor**.

Solution: Using right hand rule, so that $\mathbf{E} \times \mathbf{H}$ points in propagation direction $\hat{z}$, we have

$$\tilde{\mathbf{H}} = -\hat{x}\frac{2}{\eta}e^{-\alpha z}e^{-j\beta z} \approx -\hat{x}\frac{2}{|\eta|}e^{-\alpha z}e^{-j\beta z}e^{-j\tau}\ \frac{\mathrm{A}}{\mathrm{m}}$$

as well as

$$\begin{aligned}
\langle \mathbf{S} \rangle &= \frac{1}{2}\mathrm{Re}\{\tilde{\mathbf{E}} \times \tilde{\mathbf{H}}^*\} = \frac{1}{2}\mathrm{Re}\{\hat{y}2e^{-\alpha z}e^{-j\beta z} \times (-\hat{x}\frac{2}{|\eta|}e^{-\alpha z}e^{-j\beta z}e^{-j\tau})^*\} \\
&= -\frac{1}{2}\mathrm{Re}\{\hat{y}2e^{-\alpha z} \times \hat{x}\frac{2}{|\eta|}e^{-\alpha z}e^{j\tau}\} = \hat{z}\frac{2}{|\eta|}e^{-2\alpha z}\cos\tau.
\end{aligned}$$

With the given parameters,

$$\frac{\sigma}{\omega\epsilon} = \frac{4 \cdot 36\pi \times 10^9}{2\pi \cdot 10^3 \cdot 4} = 18 \cdot 10^6 \gg 1,$$

which confirms that the medium behaves as a good conductor at this small ω, and using the appropriate formulae from the table,

$$\begin{aligned}
\tau &\approx \frac{\pi}{4}\,\mathrm{rad} \\
|\eta| &\approx \sqrt{\frac{\omega\mu}{\sigma}} = \sqrt{\frac{2\pi \cdot 10^3 \cdot 4\pi \cdot 10^{-7}}{4}} = \pi\sqrt{2 \times 10^{-4}} \approx \frac{\pi\sqrt{2}}{100}\,\Omega \\
\alpha &\approx \sqrt{\pi f \mu\sigma} = \sqrt{\pi \cdot 10^3 \cdot 4\pi \cdot 10^{-7} \cdot 4} = \sqrt{4^2\pi^2 10^{-4}} = \frac{\pi}{25}\,\frac{1}{\mathrm{m}}.
\end{aligned}$$

Hence, at $z = 0$,

$$\langle \mathbf{S} \rangle = \hat{z}\frac{2}{|\eta|}\cos\tau \approx \hat{z}\frac{200}{\pi\sqrt{2}}\cos\frac{\pi}{4} = \hat{z}\frac{100}{\pi}\,\frac{\mathrm{W}}{\mathrm{m}^2},$$

whereas, at $z = 10$ m,

$$\langle \mathbf{S} \rangle = \hat{z}\frac{100}{\pi}\, e^{-2\cdot\frac{\pi}{25}\cdot 10} \approx \hat{z}\frac{100}{\pi} 0.081\, \frac{\mathrm{W}}{\mathrm{m}^2}.$$

- As Example 2 illustrates, at a frequency of $\omega = 2\pi \cdot 10^3$ rad/s or $f = 1$ kHz, wave power is reduced to about 8% over a 10 m distance in sea water. Less reduction in power is possible over the same distance if at a smaller frequency f since $\alpha \propto \sqrt{f}$.
  - The disadvantage of being forced to use smaler frequencies is of course having a smaller available bandwidth at small frequencies. Thus communication with submarines at great depths will only be possible at very slow rates.

The next example identifies the penetration depth in sea water at 1 kHz.

Example 3: What is the **penetration depth** $\delta = \alpha^{-1}$ in a medium with $\sigma = 4$ S/m, $\epsilon = 81\epsilon_o$, and $\mu = \mu_o$ for $\omega = 2\pi \cdot 10^3$ rad/s.

Solution: With the given parameters we have

$$\frac{\sigma}{\omega\epsilon} = \frac{4 \cdot 36\pi \times 10^9}{2\pi \cdot 10^3 \cdot 81} = \frac{72 \times 10^9}{81 \times 10^3} \approx 10^6 \gg 1,$$

i.e., good conductor situation. Hence the penetration depth is

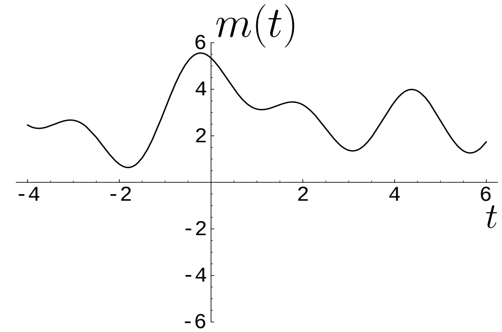
$$\delta \approx \frac{1}{\sqrt{\pi f \mu \sigma}} = \frac{1}{\sqrt{\pi 10^3 \cdot 4\pi \cdot 10^{-7} 4}} = \frac{1}{\sqrt{4^2\pi^2 \cdot 10^{-4}}} = \frac{100}{4\pi} = \frac{25}{\pi} \approx 7.95\,\text{m}.$$

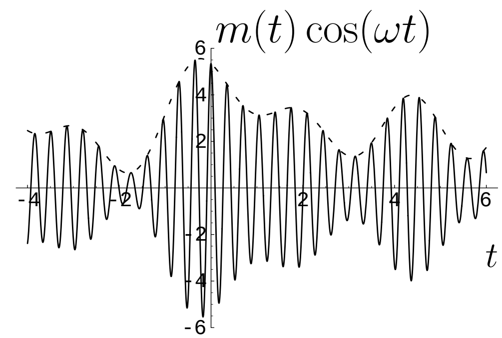
24 Signal transmission, circular polarization

Since in perfect dielectrics the propagation velocity $v_p = v$ and the intrinsic impedance η are frequency independent (i.e., propagation is non-dispersive), d'Alembert plane wave solutions of the form

$$\mathbf{E} = \hat{x} f(t - \frac{z}{v}) \quad \text{and} \quad \mathbf{H} = \hat{y} \frac{f(t - \frac{z}{v})}{\eta}$$

are valid in such media.





- Consider a waveform

$$f(t) = m(t)\cos(\omega t),$$

where

  - ω is some specific frequency having a corresponding period $T = \frac{2\pi}{\omega}$,
  - $m(t)$ is some arbitrary signal (e.g., a voice signal, a message) changing slowly compared to period T.

In that case,

  - $f(t)$ specified above can be called **narrowband AM**, and
  - ω the **carrier frequency** of modulating cosine of the message signal $m(t)$.

The corresponding x-polarized wave fields propagating in z direction can then be represented as **Field 1**

$$\mathbf{E} = m(t - \frac{z}{v})\cos(\omega t - \beta z)\hat{x} \quad \text{and} \quad \mathbf{H} = \frac{m(t - \frac{z}{v})}{\eta}\cos(\omega t - \beta z)\hat{y}$$

where $\beta = \omega\sqrt{\mu\epsilon}$ as usual[1].

- With reference to the expressions above, we could say that the AM wave field has an **x-polarized carrier**.

- By contrast, **Field 2**

$$\mathbf{E} = m(t - \frac{z}{v})\cos(\omega t - \beta z)\hat{y}$$

represents an AM wave field with a **y-polarized carrier**, and so does **Field 3**

$$\mathbf{E} = m(t - \frac{z}{v})\sin(\omega t - \beta z)\hat{y}$$

but with a carrier that has been time-delayed by a quarter period.

- Suppose Fields 1 and 3 above were transmitted simultaneously and therefore superpose. In that case we will have a wave field with **Circular polarized carrier**

$$\mathbf{E} = m(t - \frac{z}{v})[\cos(\omega t - \beta z)\hat{x} + \sin(\omega t - \beta z)\hat{y}]$$

[1]In ***dispersive*** media where β is a non-linear function of ω, narrowband AM can propagate as

$$m(t - \frac{z}{v_g})\cos(\omega t - \beta z)\hat{x} \;\text{ where }\; v_g \equiv \frac{\partial\omega}{\partial\beta}$$

is known as **group velocity** — covered in detail in ECE 450.

which has a **circular polarized carrier**. Since this is just a superposition of two d'Alembert waves, the accompanying **H** is easily found to be

$$\mathbf{H} = m(t - \frac{z}{v})[\cos(\omega t - \beta z)\hat{y} - \sin(\omega t - \beta z)\hat{x}]/\eta.$$

- Circular-polarized AM wave fields just introduced are in some practical applications better to use than the linear-polarized waves because of, say, the peculiarities of a propagation medium (e.g, Earth's ionosphere or the interplanetary medium).
- Since a circular-polarized wave field is a linear combination of linear-polarized waves, it has a phasor that is a linear combination of phasors of its linear components, as in

$$\cos(\omega t - \beta z)\hat{x} + \sin(\omega t - \beta z)\hat{y} \;\Leftrightarrow\; e^{-j\beta z}\hat{x} - je^{-j\beta z}\hat{y} = (\hat{x} - j\hat{y})e^{-j\beta z}$$

or

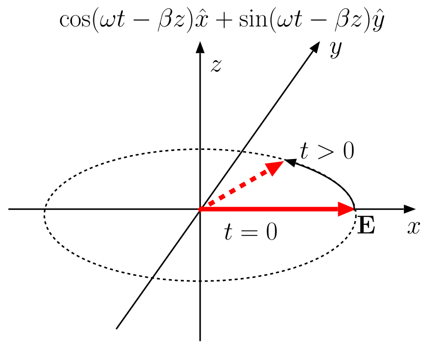
$$\cos(\omega t - \beta z)\hat{x} - \sin(\omega t - \beta z)\hat{y} \;\Leftrightarrow\; e^{-j\beta z}\hat{x} + je^{-j\beta z}\hat{y} = (\hat{x} + j\hat{y})e^{-j\beta z}.$$

- In the last step above, we have introduced two flavors of circularly polarized waves, which correspond to fields vectors rotating in opposite directions at any position in space when viewed toward the direction the wave propagates— **clockwise for right-circular**, **counter-clockwise for left circular**.

CIRCULAR POLARIZATION:

Field vector rotates instead of oscillating.

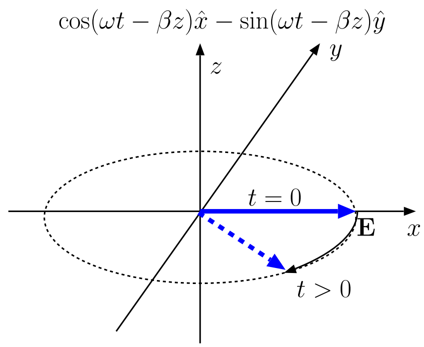
The rotation frequency is also the wave frequency.



RIGHT CIRCULAR

Right-circular

Left-circular



LEFT CIRCULAR

When left-hand thumb is pointed along propagation direction z the fingers curl in the rotation direction of the field vector.

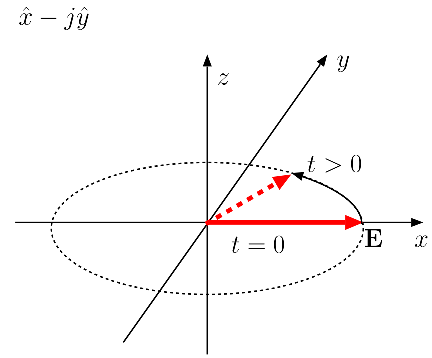
- Also,
  - for the **right-circular** wave propagating in z direction, the field vector simplified at $z = 0$ as

$$\cos(\omega t)\hat{x} + \sin(\omega t)\hat{y} \quad \Leftrightarrow \quad \hat{x} - j\hat{y}$$

    rotates in the direction that your **right-hand** fingers curl when the thumb is directed in propagation direction z, whereas
  - for the **left-circular** wave propagating in z direction, likewise, vector

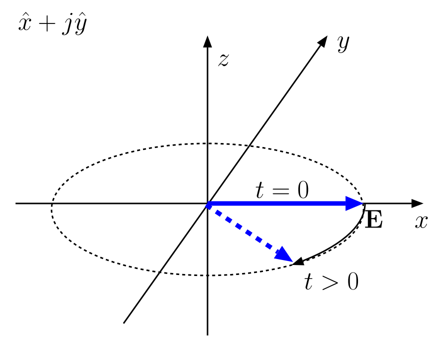
$$\cos(\omega t)\hat{x} - \sin(\omega t)\hat{y} \quad \Leftrightarrow \quad \hat{x} + j\hat{y}$$

    rotates in the direction that your **left-hand** fingers curl when the thumb is directed in propagation direction z.



RIGHT CIRCULAR

x-comp leads y-comp because of -j



LEFT CIRCULAR

x-comp lags y-comp because of +j

In general, the "handednes" or "helicity" of a circular polarized wave is always obtained by matching your *right* or *left* hand to the specified propagation and rotation directions — see example below.

Furthermore, the rotation direction is most easily seen if the wave is expressed in phasor form by seeing which component leads (or lags) which. Here is an explanation by example:

Example 1: A circular polarized wave field vector is given as

$$\tilde{\mathbf{E}} = (\hat{z} + j\hat{y})e^{j\beta x}.$$

Determine the propagation and rotation directions of the field vector as well as its helicity.

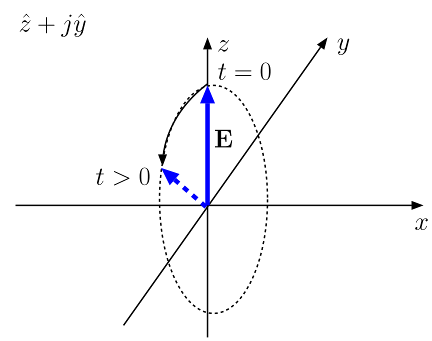
Solution: The propagation direction is $-x$ since the exponent in $e^{j\beta x}$ lacks a minus sign.

At $x = 0$, the wave field vector rotates as

$$\mathbf{E} = \mathrm{Re}\{(\hat{z} + j\hat{y})e^{j\omega t}\} = \hat{z}\cos(\omega t) - \hat{y}\sin(\omega t),$$

of which the y-component leads the z-component by 90° of phase, or, equivalently, by a quarter period in time — therefore, the vector points in y-direction before it points in z-direction (or in z-direction before it points in $-y$-direction), rotating from y- toward z-axis.

When I direct my right thumb in $-x$ direction, my fingers curl from z- toward y-axis, which is curling in the wrong direction. Hence this wave is not right-circular! It is **left-circular**.



LEFT CIRCULAR

(i) z-comp lags y-comp because of +j.

(ii)E-vector rotates from y towards z.

(iii) since the wave propagates in -x direction it is LEFT CIRCULAR

Given any propagation direction, a carrier field of an arbitrary polarization can always be expressed as weighted superpositions of any pair of **orthogonal polarized** carrier fields — such orthogonal pairs are considered to be complete sets of **basis functions** for expressing waves with arbitrary polarizations.

- **EXAMPLE:** Right- and left circular waves propagating in z directions are weighted superpositions of **orthogonal x- and y-polarized** fields as in (expressed in terms of phasors): basis functions

Circulars in terms of linears

$$\hat{x}e^{-j\beta z} \text{ and } \hat{y}e^{-j\beta z}$$

superpose to form right- and left-circular waves

$$(\hat{x} - j\hat{y})e^{-j\beta z} \text{ and } (\hat{x} + j\hat{y})e^{-j\beta z}$$

using the weights

$$1,\, -j \text{ and } 1,\, j$$

respectively.

- **EXAMPLE:** x- and y-polarized waves propagating in z directions are weighted superpositions of **orthogonal right- and left-circular** fields as in (expressed in terms of phasors): basis functions

Linears in terms of circulars

$$(\hat{x} - j\hat{y})e^{-j\beta z} \text{ and } (\hat{x} + j\hat{y})e^{-j\beta z}$$

superpose to form linear polarized waves

$$\hat{x}e^{-j\beta z} \text{ and } \hat{y}e^{-j\beta z}$$

using the weights

$$\frac{1}{2}, \frac{1}{2} \text{ and } -\frac{1}{2j}, \frac{1}{2j}$$

respectively.

- It can be argued that right- and left-circular wave pair forms an intrinsically more fundamental set of basis functions than, say, $\hat{x}$- and $\hat{y}$-polarized waves, because while the selection of which direction is x and which direction is y can be arbitrary, there is no arbitrariness in how helicity is assigned to circular polarized modes propagating in a given direction[2].

- Also, oscillating charges will radiate linear-polarized fields, whereas rotating charges will radiate circular-polarized fields (in the direction normal to the rotation plane) — so, source dynamics selects the radiated wave polarization.

- Wave polarization is important because
  - it depends on physical geometry and dynamics of the wave source,
  - it may depend on the physical properties of the region the wave propagates through,
  - it will determine the direction of Lorentz force on any test charge or electrical load,
  - angular momentum carried by the wave depends on polarization, etc.

[2]Furthermore RCP and LCP plane waves consist of photons with spin angular momenta of $+\hbar$ and $-\hbar$, respectively, corresponding to the eigenvalues of the quantum mechanical spin operator, while the photons constituting LP waves will be in a "superposition state" of the eigenvectors of the spin operator having the eigenvalues $\pm\hbar$ — upon spin measurements the photons of a LP wave will furnish one of $+\hbar$ and $-\hbar$ with equal (50%) probabilities, unlike the RCP and LCP wave photons furnishing $+\hbar$ and $-\hbar$, respectively, with 100% probabilities.

Example 2: On $z = 0$ plane we have a time-varying surface current density

$$\mathbf{J}_s(t) = m(t)[\cos(\omega t)\hat{x} + \sin(\omega t)\hat{y}]\,\frac{\mathrm{A}}{\mathrm{m}}$$

with a carrier frequency of ω. Determine the radiated wave fields $\mathbf{E}^{\pm}$ and the polarization (and the helicity if appropriate) of the carrier.

Solution: We have already learned that a surface current $\mathbf{J}_s(t)$ on $z = 0$ plane will produce TEM wave fields

$$\mathbf{E}^{\pm} = -\frac{\eta}{2}\mathbf{J}_s(t \mp \frac{z}{v})$$

in surrounding regions. With the given $\mathbf{J}_s(t)$, this implies

$$\mathbf{E}^{\pm} = -\frac{\eta}{2}m(t \mp \frac{z}{v})[\cos(\omega t \mp \beta z)\hat{x} + \sin(\omega t \mp \beta z)\hat{y}]\,\frac{\mathrm{V}}{\mathrm{m}},$$

which has a circular-polarized carrier

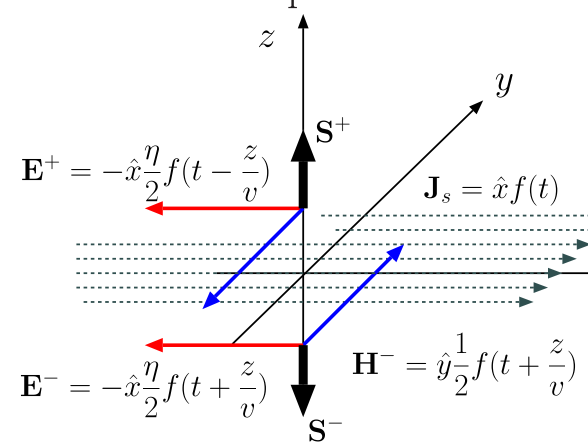
$$\cos(\omega t \mp \beta z)\hat{x} + \sin(\omega t \mp \beta z)\hat{y}$$

that varies, on $z = 0$ plane, as

$$\cos(\omega t)\hat{x} + \sin(\omega t)\hat{y}.$$

This vector rotates from x- toward y-axis, and therefore the carrier of $\mathbf{E}^{+}$ is right-circular and the carrier of $\mathbf{E}^{-}$ is left-circular.

Note that this figure only shows one linear component of the surface current on $z = 0$ plane. One linear component causes a linear polarized radiation. An orthogonal pair of linear components will conspire to radiate a circular polarized wave as in Example 2 when they are 90° out of phase.



Example 3: In Example 2, what is the average power density of the circular polarized carrier signal

$$\mathbf{E}_c = \cos(\omega t - \beta z)\hat{x} + \sin(\omega t - \beta z)\hat{y}\,\frac{\mathrm{V}}{\mathrm{m}}$$

in the region $z > 0$, assumed to be vacuum?

Solution: In phasor notation $\mathbf{E}_c$ and is given as

$$\tilde{\mathbf{E}}_c = (\hat{x} - j\hat{y})e^{-j\beta z}\,\frac{\mathrm{V}}{\mathrm{m}}.$$

The corresponding $\mathbf{H}_c$ phasor is

$$\tilde{\mathbf{H}}_c = \frac{1}{\eta_o}(\hat{y} + j\hat{x})e^{-j\beta z}\,\frac{\mathrm{V}}{\mathrm{m}}.$$

Therefore, the average power density is found to be

$$\frac{1}{2}\mathrm{Re}\{\tilde{\mathbf{E}}_c \times \tilde{\mathbf{H}}_c^*\} = \frac{1}{2\eta_o}\mathrm{Re}\{(\hat{x} - j\hat{y}) \times (\hat{y} + j\hat{x})^*\} = \frac{1}{2\eta_o}(\hat{z} + \hat{z}) = \frac{1}{\eta_o}\hat{z}.$$

This is twice the power content of a linearly polarized wave field of an equal amplitude!

Make sure you check and follow all the sign changes that take place in Example 3.

25 Wave reflection and transmission

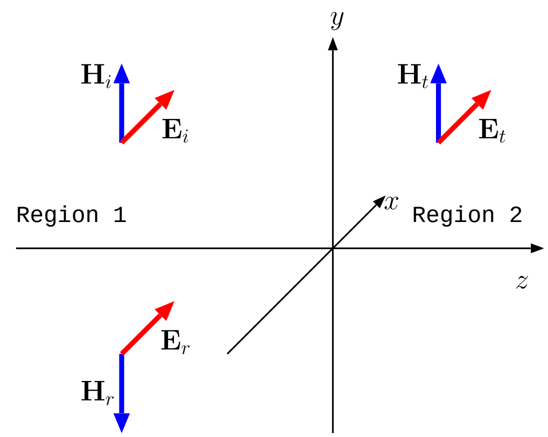
In this lecture we will examine the phenomenon of plane-wave reflections at an interface separating two homogeneous regions where Maxwell's equations allow for traveling TEM wave solutions. The solutions will also need to satisfy the boundary condition equations repeated in the margin. We will consider a propagation scenario in which (see margin):

$$\begin{aligned} \hat{n}\cdot(\mathbf{D}^+ - \mathbf{D}^-) &= \rho_s \\ \hat{n}\cdot(\mathbf{B}^+ - \mathbf{B}^-) &= 0 \\ \hat{n}\times(\mathbf{E}^+ - \mathbf{E}^-) &= 0 \\ \hat{n}\times(\mathbf{H}^+ - \mathbf{H}^-) &= \mathbf{J}_s \end{aligned}$$

1. Region 1 where $z < 0$ is occupied by a perfect dielectric with medium parameters μ_1, ϵ_1, and $\sigma_1 = 0$,

2. Region 2 where $z > 0$ is homogeneous with medium parameters μ_2, ϵ_2, and σ_2,

3. Interface $z = 0$ contains no surface charge or current except possibly in $\sigma_2 \to \infty$ limit which will be considered separately at the end.



- In Region 1 we envision an **incident plane-wave** with linear-polarized field phasors

$$\tilde{\mathbf{E}}_i = \hat{x}E_o e^{-j\beta_1 z} \quad \text{and} \quad \tilde{\mathbf{H}}_i = \hat{y}\frac{E_o}{\eta_1}e^{-j\beta_1 z},$$

  where

  - E_o is the wave amplitude due to far away source located in $z \to -\infty$ region,
  - $\eta_1 = \sqrt{\frac{\mu_1}{\epsilon_1}}$ and $\beta_1 = \omega\sqrt{\mu_1\epsilon_1}$.

Fields above satisfy Maxwell's equations in Region 1, but if there were no other fields in Regions 1 and 2 **boundary condition equations** requiring continuous tangential **E** and **H** at the $z = 0$ interface would be violated.

In order to comply with the boundary condition equations we postulate a set of reflected and transmitted wave fields in Regions 1 and 2 as follows:

- In Region 1 we postulate a **reflected plane-wave** with linear-polarized field phasors

$$\tilde{\mathbf{E}}_r = \hat{x}\Gamma E_o e^{j\beta_1 z} \quad \text{and} \quad \tilde{\mathbf{H}}_r = -\hat{y}\frac{\Gamma E_o}{\eta_1}e^{j\beta_1 z}$$

  including an unknown Γ that we will refer to as **reflection coefficient**.

  - Note that the reflected wave propagates in $-z$ direction (direction of $\tilde{\mathbf{H}}_r$ and the exponential terms have been adjusted accordingly).

- In Region 2 we postulate a **transmitted plane-wave** with linear-polarized field phasors

$$\tilde{\mathbf{E}}_t = \hat{x}\tau E_o e^{-\gamma_2 z} \quad \text{and} \quad \tilde{\mathbf{H}}_t = \hat{y}\frac{\tau E_o}{\eta_2}e^{-\gamma_2 z}$$

  including an unknown τ that we will refer to as **transmission coefficient**.

  - Note that the transmitted wave propagates in z direction, and
  - since Region 2 is conducting we have

$$\eta_2 = \sqrt{\frac{j\omega\mu_2}{\sigma_2 + j\omega\epsilon_2}}$$

    and

$$\gamma_2 = \sqrt{(j\omega\mu_2)(\sigma_2 + j\omega\epsilon_2)} = \alpha_2 + j\beta_2.$$

Incident:

$$\tilde{\mathbf{E}}_i = \hat{x}E_o e^{-j\beta_1 z},$$

$$\tilde{\mathbf{H}}_i = \hat{y}\frac{E_o}{\eta_1}e^{-j\beta_1 z},$$

Reflected:

$$\tilde{\mathbf{E}}_r = \hat{x}\Gamma E_o e^{j\beta_1 z},$$

$$\tilde{\mathbf{H}}_r = -\hat{y}\frac{\Gamma E_o}{\eta_1}e^{j\beta_1 z},$$

Transmitted:

$$\tilde{\mathbf{E}}_t = \hat{x}\tau E_o e^{-\gamma_2 z},$$

$$\tilde{\mathbf{H}}_t = \hat{y}\frac{\tau E_o}{\eta_2}e^{-\gamma_2 z}.$$

- To determine the unknowns Γ and τ we enforce the following boundary conditions at $z = 0$ where the fields simplify as shown in the margin:

1. **Tangential $\tilde{\mathbf{E}}$ continuous at $z = 0$:** This requires $\tilde{E}_{ix} + \tilde{E}_{rx} = \tilde{E}_{tx}$, leading to
$$(1 + \Gamma)E_o = \tau E_o \quad \Rightarrow \quad \boxed{1+\Gamma=\tau}$$

2. **Tangential $\tilde{\mathbf{H}}$ continuous at $z = 0$:** This requires $\tilde{H}_{iy} + \tilde{H}_{ry} = \tilde{H}_{ty}$, leading to
$$(1 - \Gamma)\frac{E_o}{\eta_1} = \tau \frac{E_o}{\eta_2} \quad \Rightarrow \quad \boxed{1\text{-}\Gamma=\frac{\eta_1}{\eta_2}\tau}$$

   Replacing τ by $1 + \Gamma$ in the second equation, we can solve for the reflection coefficient as
$$\boxed{\Gamma=\frac{\eta_2-\eta_1}{\eta_2+\eta_1}}$$
   and substituting this in turn in the first equation we can solve for the transmission coefficient as
$$\boxed{\tau=\frac{2\eta_2}{\eta_2+\eta_1}}$$

The results are summarized in the margin on the next page.

Incident at $z = 0$:

$$\tilde{\mathbf{E}}_i = \hat{x}E_o, \;\; \tilde{\mathbf{H}}_i = \hat{y}\frac{E_o}{\eta_1},$$

Reflected at $z = 0$:

$$\tilde{\mathbf{E}}_r = \hat{x}\Gamma E_o, \;\; \tilde{\mathbf{H}}_r = -\hat{y}\frac{\Gamma E_o}{\eta_1}$$

Transmitted at $z = 0$:

$$\tilde{\mathbf{E}}_t = \hat{x}\tau E_o, \;\; \tilde{\mathbf{H}}_t = \hat{y}\frac{\tau E_o}{\eta_2}$$

Special cases:

1. **Region 2 is a perfect conductor with $\sigma_2 \to \infty$:** In that case $\eta_2 \to 0$, and consequently

$$\Gamma = -1 \quad \text{and} \quad \tau = 0.$$

Incident wave cannot penetrate the **perfect conductor**, and it reflects totally back into Region 1 — we will study this idealized limiting case more carefully later on.

Practical application of total reflection: mirrors

2. **Region 2 is the same as Region 1:** In that case $\eta_2 = \eta_1$, and consequently

$$\Gamma = 0 \quad \text{and} \quad \tau = 1.$$

This is the **matched impedance** case when no reflection takes place and the incident wave is transmitted in its entirety.

3. **Region 2 is lossless, i.e., $\sigma_2 = 0$:** Unless $\eta_2 = \eta_1$ there will be reflected as well as transmitted waves.

Partial reflections can be reduced by applying a "anti-glare" coating[1] on the surface, a practice known as "impedance matching".

[1]This is a $\lambda/4$ thick layer of a material having a characteristic impedance given by $\sqrt{\eta_1\eta_2}$ — the reason for why this "quarter-wave matching" works will be discussed when we study transmission lines later on.

Reflection coeff.:

$$\Gamma = \frac{\eta_2 - \eta_1}{\eta_2 + \eta_1},$$

Transmission coeff.:

$$\tau = \frac{2\eta_2}{\eta_2 + \eta_1}.$$

Memorize the Γ formula, and memorize τ as "one plus Γ".

Above,

$$\eta_1 = \sqrt{\frac{\mu_1}{\epsilon_1}}$$

and

$$\eta_2 = \sqrt{\frac{j\omega\mu_2}{\sigma_2 + j\omega\epsilon_2}}.$$

Example 1: An plane-wave in vacuum,

$$\tilde{\mathbf{E}}_i = \hat{x}\sqrt{120\pi}e^{-j\beta_1 z}\ \frac{V}{m},$$

is incident at $z = 0$ on a dielectric medium with $\mu = \mu_o$ and $\epsilon = \frac{9}{4}\epsilon_o$. Determine the average Poynting vectors $\langle \mathbf{S}_i \rangle$, $\langle \mathbf{S}_r \rangle$, and $\langle \mathbf{S}_t \rangle$ of the incident, reflected, and transmitted fields.

Solution: The intrinsic impedance of the second medium occupying $z > 0$ is

$$\eta_2 = \sqrt{\frac{\mu_o}{\frac{9}{4}\epsilon_o}} = \frac{2}{3}\eta_o.$$

Therefore, the reflection coefficient is

$$\Gamma = \frac{\eta_2 - \eta_1}{\eta_2 + \eta_1} = \frac{\frac{2}{3}\eta_o - \eta_o}{\frac{2}{3}\eta_o + \eta_o} = \frac{\frac{2}{3} - 1}{\frac{2}{3} + 1} = \frac{2-3}{2+3} = -\frac{1}{5}$$

and the transmission coefficient is

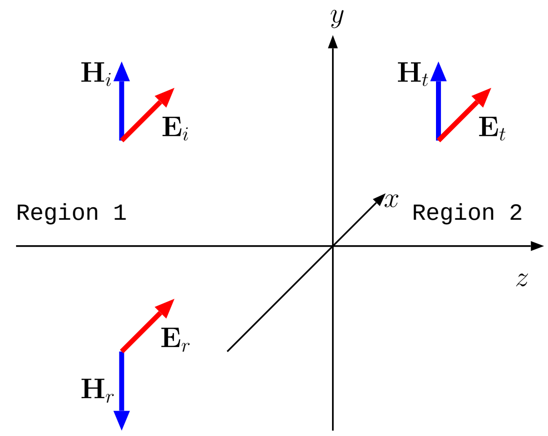
$$\tau = 1 + \Gamma = 1 - \frac{1}{5} = \frac{4}{5}.$$

The reflected wave therefore has the field phasors

$$\tilde{\mathbf{E}}_r = -\frac{1}{5}\hat{x}\sqrt{120\pi}e^{j\beta_1 z} \quad \text{and} \quad \tilde{\mathbf{H}}_r = \frac{1}{5\eta_o}\hat{y}\sqrt{120\pi}e^{j\beta_1 z}$$

and

$$\langle \mathbf{S}_r \rangle = \frac{1}{2}\mathrm{Re}\{\tilde{\mathbf{E}}_r \times \tilde{\mathbf{H}}_r^*\} = -\hat{z}\frac{1}{2}(\frac{1}{5})^2\frac{120\pi}{\eta_o} \approx -\hat{z}\frac{1}{2}(\frac{1}{5})^2\,\frac{\mathrm{W}}{\mathrm{m}^2}.$$



The transmitted wave, likewise, has the field phasors

$$\tilde{\mathbf{E}}_t = \frac{4}{5}\hat{x}\sqrt{120\pi}e^{-j\beta_2 z} \quad \text{and} \quad \tilde{\mathbf{H}}_t = \frac{4}{5\frac{2}{3}\eta_o}\hat{y}\sqrt{120\pi}e^{-j\beta_2 z}$$

and

$$\langle \mathbf{S}_t \rangle = \frac{1}{2}\text{Re}\{\tilde{\mathbf{E}}_t \times \tilde{\mathbf{H}}_t^*\} = \hat{z}\frac{1}{2}(\frac{4}{5})^2\frac{3}{2}\frac{120\pi}{\eta_o} \approx \hat{z}\frac{1}{2}(\frac{4}{5})^2\frac{3}{2}\,\frac{\text{W}}{\text{m}^2}.$$

As for the incident wave

$$\tilde{\mathbf{E}}_i = \hat{x}\sqrt{120\pi}e^{-j\beta_1 z} \quad \text{and} \quad \tilde{\mathbf{H}}_i = \frac{1}{\eta_o}\hat{y}\sqrt{120\pi}e^{-j\beta_1 z}$$

and

$$\langle \mathbf{S}_i \rangle = \frac{1}{2}\text{Re}\{\tilde{\mathbf{E}}_i \times \tilde{\mathbf{H}}_i^*\} = \hat{z}\frac{1}{2}\frac{120\pi}{\eta_o} \approx \hat{z}\frac{1}{2}\,\frac{\text{W}}{\text{m}^2}.$$

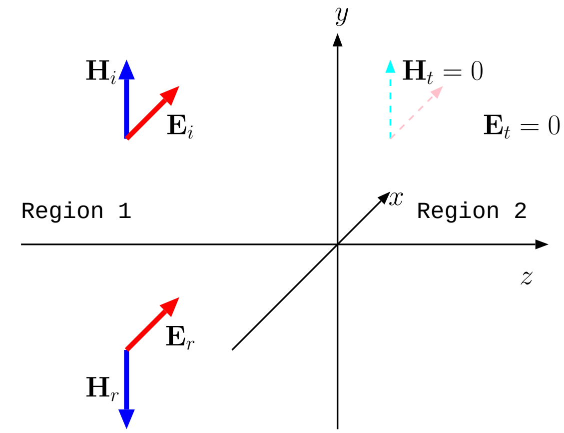
Note: We have

$$|\langle \mathbf{S}_r \rangle| + |\langle \mathbf{S}_t \rangle| = \frac{1}{2}(\frac{1}{25} + \frac{16}{25}\frac{3}{2}) = \frac{1}{2}(\frac{1}{25} + \frac{24}{25}) = \frac{1}{2} = |\langle \mathbf{S}_i \rangle|$$

in compliance with energy conservation (as expected) — **energy flux per unit area of the transmitted and reflected waves add up the that of the incident wave!**

26 Standing waves, radiation pressure

We continue in this lecture with our studies of wave reflection and transmission at a plane boundary between two homogeneous media.



- In case of *total reflection* from a perfectly conducting mirror placed at $z = 0$ surface, $\Gamma = -1$, and the incident and reflected waves in $z < 0$ region combine to produce **standing waves** of electric and magnetic field:

  - **Incident wave** (a traveling wave going in z-direction):

$$\tilde{\mathbf{E}}_i = \hat{x}E_o e^{-j\beta_1 z} \quad \text{and} \quad \tilde{\mathbf{H}}_i = \hat{y}\frac{E_o}{\eta_1}e^{-j\beta_1 z},$$

  - **Reflected wave** (a traveling wave going in $-z$-direction):

$$\tilde{\mathbf{E}}_r = -\hat{x}E_o e^{j\beta_1 z} \quad \text{and} \quad \tilde{\mathbf{H}}_r = \hat{y}\frac{E_o}{\eta_1}e^{j\beta_1 z},$$

  - **Standing wave**:

$$\tilde{\mathbf{E}} = \tilde{\mathbf{E}}_i + \tilde{\mathbf{E}}_r = \hat{x}E_o(e^{-j\beta_1 z} - e^{j\beta_1 z}) \quad \text{and} \quad \tilde{\mathbf{H}} = \tilde{\mathbf{H}}_i + \tilde{\mathbf{H}}_r = \hat{y}\frac{E_o}{\eta_1}(e^{-j\beta_1 z} + e^{j\beta_1 z})$$

    which simplify as

$$\tilde{\mathbf{E}} = -j\hat{x}2E_o\sin(\beta_1 z) \quad \text{and} \quad \tilde{\mathbf{H}} = \hat{y}\frac{2E_o}{\eta_1}\cos(\beta_1 z).$$

Standing waves

These are called ***standing wave*** phasors because when we go to the time-domain (by multiplying with $e^{j\omega t}$ and taking the real time as usual) we obtain:

$$\mathbf{E}(z,t) = \hat{x}2E_o \sin(\beta_1 z)\sin(\omega t) \quad \text{and} \quad \mathbf{H}(z,t) = \hat{y}\frac{2E_o}{\eta_1}\cos(\beta_1 z)\cos(\omega t);$$

these, unlike d'Alembert solutions of the format $f(t \mp \frac{z}{v})$, describe oscillations in time t, with different amplitudes at different positions z (see margin and the animation linked in class calendar).

– Standing waves carry no net energy, that is, with standing wave fields we have

$$\langle \mathbf{E}\times\mathbf{H}\rangle = 0,$$

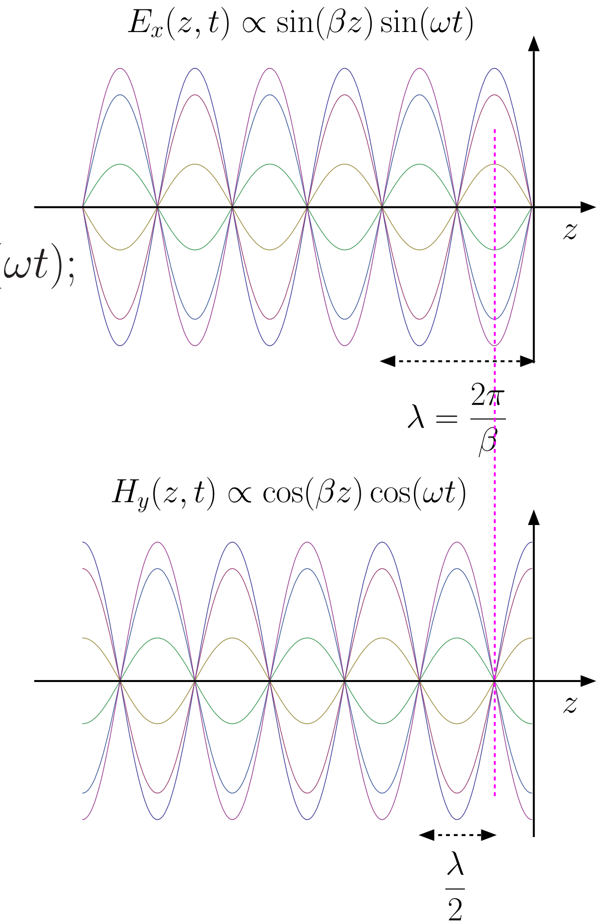
because of the cancellation of the power transported by its traveling wave components in opposite directions.

Verification: Using the phasors

$$\tilde{\mathbf{E}} = -j\hat{x}2E_o\sin(\beta_1 z) \quad \text{and} \quad \tilde{\mathbf{H}} = \hat{y}\frac{2E_o}{\eta_1}\cos(\beta_1 z),$$

we have

$$\begin{aligned}\langle \mathbf{E}\times\mathbf{H}\rangle &= \frac{1}{2}\mathrm{Re}\{\tilde{\mathbf{E}}\times\tilde{\mathbf{H}}^*\} = \frac{1}{2}\mathrm{Re}\{-j\hat{x}2E_o\sin(\beta_1 z)\times\hat{y}\frac{2E_o}{\eta_1}\cos(\beta_1 z)\} \\ &= \hat{z}\frac{2E_o^2}{\eta_1}\sin(\beta_1 z)\cos(\beta_1 z)\mathrm{Re}\{-j\} = 0.\end{aligned}$$



Note: Nulls in Ex and Hy are separated by half wavelength.

Adjacent nulls of Ex and Hy are separated by quarter wavelength.

It is useful to think of nulls of Ex as "shorts" in analogy to shorts in circuits where v=0.

Conductor shorts Ex on its surface where a current flows.

Also useful to think of nulls of Hy as "opens" in analogy to opens in circuits where i=0.

– Note that $\mathbf{E}(0,t) = 0$ on $z = 0$ surface satisfying the tangential electric boundary condition as expected (recall that the fields are zero within the perfect conducting mirror).

– Also note that

$$\mathbf{H}(0,t) = \hat{y}\frac{2E_o}{\eta_1}\cos(\omega t)$$

on $z = 0$ surface. Since this tangential magnetic field is not zero, boundary condition equations imply that there must be an oscillating surface current

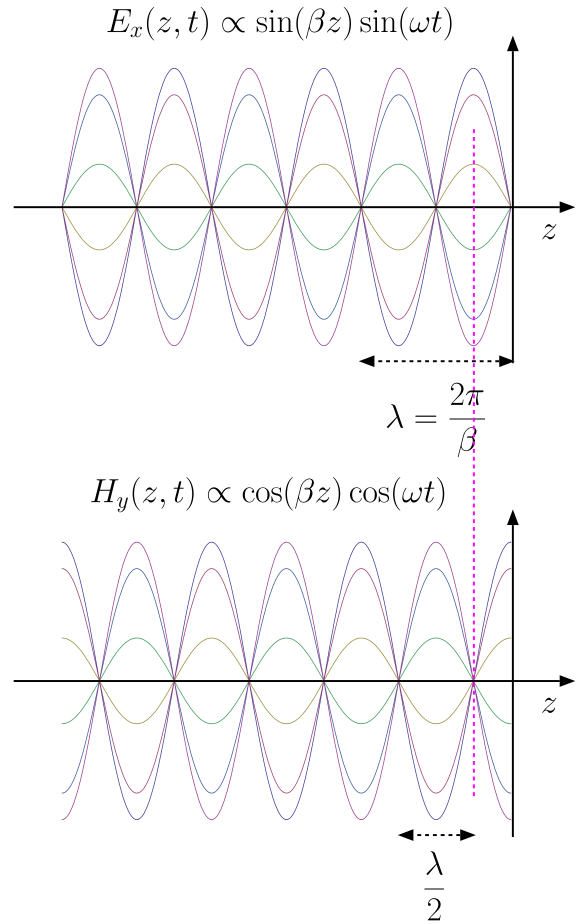
$$\mathbf{J}_s = \hat{x}\frac{2E_o}{\eta_1}\cos(\omega t)\ \frac{\mathrm{A}}{\mathrm{m}},$$

satisfying

$$-\hat{z}\times\mathbf{H}(0,t) = \mathbf{J}_s.$$

$\mathbf{J}_s$ on mirror surfaces is really a convenient *idealization* of volume currents flowing in thin layers — just a few skin depths — near good-conductor surfaces (real-life mirrors are good but not perfect conductors!). Radiation due to $\mathbf{J}_s$ *in effect* causes the reflected wave and also cancels out the incident wave field in $z > 0$.

Next we examine reflections from a good conductor and see of how the limiting case of a perfect conductor is naturally reached.



Note: Nulls in Ex and Hy are separated by half wavelength.

Adjacent nulls of Ex and Hy are separated by quarter wavelength.

It is useful to think of nulls of Ex as "shorts" in analogy to shorts in circuits where v=0.

Conductor shorts Ex on its surface where a current flows.

Also useful to think of nulls of Hy as "opens" in analogy to opens in circuits where i=0.

- Going back to the partial reflection case, consider the transmitted fields $\tilde{\mathbf{E}}_t$ and $\tilde{\mathbf{H}}_t$ in Region 2 shown in the margin. Also shown in the margin are the phasors for current density $\mathbf{J}_t$ and magnetic flux density $\mathbf{B}_t$.

$$\tilde{\mathbf{E}}_t = \hat{x}\tau E_o e^{-\gamma_2 z}$$

$$\tilde{\mathbf{H}}_t = \hat{y}\frac{\tau E_o}{\eta_2}e^{-\gamma_2 z}$$

$$\tilde{\mathbf{J}}_t = \sigma_2\tilde{\mathbf{E}}_t = \hat{x}\sigma_2\tau E_o e^{-\gamma_2 z}$$

$$\tilde{\mathbf{B}}_t = \mu_2\tilde{\mathbf{H}}_t = \hat{y}\frac{\mu_2\tau E_o}{\eta_2}e^{-\gamma_2 z}$$

In the box below we integrate the volumetric current density $\mathbf{J}_t$ of a good conductor from $z = 0$ to ∞ and find out that this "depth integral" matches the surface current density found above for the case of perfect conductor. In this calculation we assume that Region 1 is vacuum, and also take $\mu_2 = \mu_o$:

Effective surface current: Assuming that Region 2 is a good conductor,

$$\eta_2 = \sqrt{\frac{j\omega\mu_2}{\sigma_2 + j\omega\epsilon_2}} \approx \sqrt{\frac{j\omega\mu_2}{\sigma_2}} \quad \text{and} \quad \gamma_2 = \sqrt{j\omega\mu_2(\sigma_2 + j\omega\epsilon_2)} \approx \sqrt{j\omega\mu_2\sigma_2}$$

and therefore

$$\tau = \frac{2\eta_2}{\eta_o + \eta_2} \approx \frac{2\eta_2}{\eta_o} \quad \text{and} \quad \sigma_2\tau \approx \frac{2\sigma_2\eta_2}{\eta_o} = \frac{2\sqrt{j\omega\mu_2\sigma_2}}{\eta_o} = \frac{2\gamma_2}{\eta_o}.$$

The depth integral of the volumetric current density in Region 2, that is, the **effective surface current** of the region is then

$$\int_0^\infty \tilde{\mathbf{J}}_t dz = \hat{x}\int_0^\infty \sigma_2\tau E_o e^{-\gamma_2 z}dz = \hat{x}E_o\frac{1}{\gamma_2}(\sigma_2\tau) = \hat{x}\frac{2E_o}{\eta_o}$$

in phasor form, matching the phasor of the time-domain result from above, namely

$$\mathbf{J}_s = \hat{x}\frac{2E_o}{\eta_o}\cos(\omega t)\ \frac{\mathrm{A}}{\mathrm{m}}$$

representing the surface current on an idealized perfect conductor surface.

Surface resistance: Let $\tilde{\mathbf{J}}_s$ stand for the effective surface current of a good conductor with a propagation constant

$$\gamma \approx \sqrt{j\omega\mu\sigma} = \alpha + j\beta = \alpha + j\alpha$$

and a volumetric current density $\tilde{\mathbf{J}}(z)$ such that

$$\tilde{\mathbf{J}}_s = \int_{z=0}^{\infty} \tilde{\mathbf{J}}(z)dz = \int_{z=0}^{\infty} \tilde{\mathbf{J}}(0)e^{-\gamma z}dz = \frac{\tilde{\mathbf{J}}(0)}{\gamma}.$$

In that case

$$\tilde{\mathbf{J}}(z) = \tilde{\mathbf{J}}_s \gamma e^{-\gamma z} \text{ and } \mathbf{E}(z) = \frac{\tilde{\mathbf{J}}_s \gamma}{\sigma} e^{-\gamma z}$$

inside the good conductor in terms of the effective surface current $\tilde{\mathbf{J}}_s$, and the average power dissipated per unit volume (Joule heating) is

$$\langle \mathbf{J}(z) \cdot \mathbf{E}(z) \rangle = \frac{1}{2}|\tilde{\mathbf{J}}_s \gamma|^2 \frac{e^{-2\alpha z}}{\sigma} = \frac{1}{2}|\tilde{\mathbf{J}}_s|^2 \frac{2\alpha^2 e^{-2\alpha z}}{\sigma}.$$

The depth integral of the same quantity, that is the **power dissipated per unit area**, is then

$$\int_0^{\infty} \langle \mathbf{J}(z) \cdot \mathbf{E}(z) \rangle dz = \frac{1}{2} R_s |\tilde{\mathbf{J}}_s|^2,$$

with

$$R_s \equiv \frac{\alpha}{\sigma} = \frac{\sqrt{\pi f \mu \sigma}}{\sigma} = \sqrt{\frac{\pi f \mu}{\sigma}} \ (\Omega)$$

called the **surface resistance**.

The surface resistance concept is useful to model loss effects in waveguides and cavity resonators as studied in ECE 450. Also, we can make use of surface resistance when modeling lossy transmission lines (see Lecture 39).

- Let's finally calculate the magnetic component of the Lorentz force on charge carriers of a good conductor due to the penetrating fields:

Radiation pressure: If there are N free charge carriers per unit volume inside a reflecting mirror, then

$$N\mathbf{F} = Nq\mathbf{v} \times \mathbf{B}_t = \mathbf{J}_t \times \mathbf{B}_t$$

will be the **force per unit volume** of the mirror, expressed in terms of current density $\mathbf{J}_t = Nq\mathbf{v}$ and the magnetic flux density $\mathbf{B}_t$.

Its integral over all z can be interpreted as the total **force per unit area** of the mirror,

$$\mathbf{P}_{rad} = \int_0^\infty \mathbf{J}_t \times \mathbf{B}_t \, dz,$$

having a magnitude known as **radiation pressure** of the reflecting wave. This is a time-varying quantity, with a time-average

$$\begin{aligned}
\langle \mathbf{P}_{rad} \rangle &= \int_0^\infty \frac{1}{2}\mathrm{Re}\{\tilde{\mathbf{J}} \times \tilde{\mathbf{B}}^*\} dz \\
&= \hat{z}\int_0^\infty \frac{1}{2}\mathrm{Re}\{(\sigma_2 \tau E_o)(\frac{\mu_2 \tau E_o}{\eta_2})\} e^{-2\alpha_2 z} dz \\
&= \hat{z}\frac{|E_o|^2}{2}\mathrm{Re}\{(\frac{2\gamma_2}{\eta_o})(\frac{\mu_2}{\eta_2}\frac{2\eta_2}{\eta_o})\}\frac{1}{2\alpha_2} = \hat{z}2\frac{|E_o|^2}{2\eta_o}\frac{\mathrm{Re}\{\gamma_2\}}{\alpha_2}\frac{\mu_2}{\eta_o} \\
&= \hat{z}2\frac{|E_o|^2}{2\eta_o}\frac{\mu_o}{\eta_o} = 2\langle \mathbf{S}_i \rangle / c,
\end{aligned}$$

where

$$\langle \mathbf{S}_i \rangle \equiv \hat{z}\frac{|E_o|^2}{2\eta_o}$$

is the time-average Poynting vector of the incident wave reflected from the mirror (factor of 2 in $\langle \mathbf{P}_{rad} \rangle$ is due to the recoil of the wave off the mirror; see Rothman and Boughn, *Am. J. Phys.*, *77*, 122, 1977).

***Radiation pressure* proportional to**

$$\langle \mathbf{S}_i \rangle / c$$

shows that plane-TEM waves not only carry and transport energy, but also momentum.

TEM waves not only *heat*, but also *push*!

((It can also be shown that the momentum density of the wave is

$$\langle \mathbf{S}_i \rangle / c^2 \text{ N.s/m}^3$$

and (spin) angular momentum density

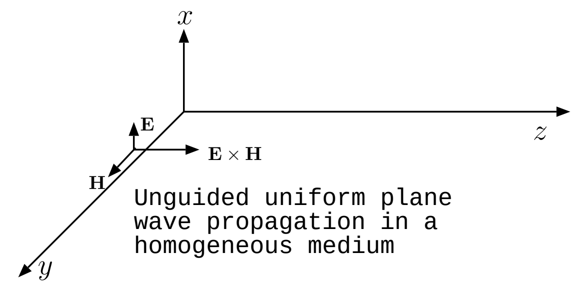
$$\pm \langle \mathbf{S}_i \rangle / \omega c \text{ N.s/m}^2$$

for right- and left-circular waves. Momentum per photon of energy $\hbar\omega$ can be obtained by dividing the above expressions by $|\langle \mathbf{S}_i \rangle| / \omega c \hbar$, the number density of photons in the wave field.))

27 Guided TEM waves on TL systems

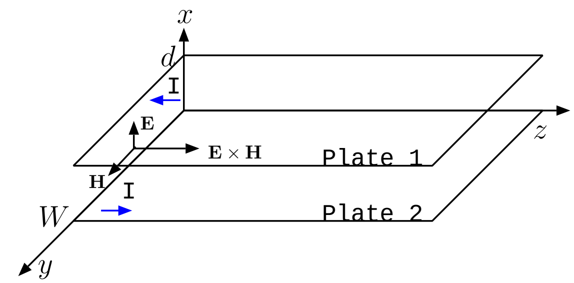
- An $\hat{x}$ polarized plane TEM wave propagating in z direction is depicted in the margin.

  - A pair of conducting plates placed at $x = 0$ and $x = d$ would not perturb the fields except that charge and current density variations would be induced on plate surfaces at $x = 0$ and $x = d$ (on both sides) to satisfy Maxwell's boundary condition equations.



- If charge and currents were confined only to interior surfaces of the plates facing one another, fields **E** and **H** accompanying them would be restricted to the region in between the plates, constituting what we would call **guided waves**.

  - Such a guided wave field confined to the region between the plates will satisfy Maxwell's equations including a minor fringing component that can be neglected when the plate width W is much larger than plate separation d.

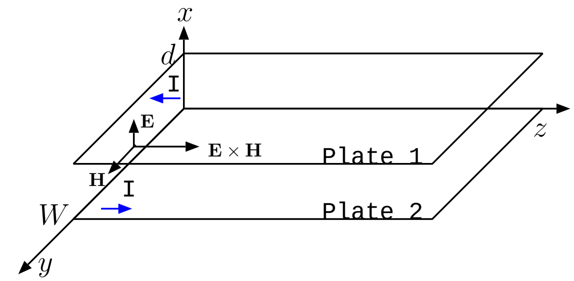


In the following discussion of guided waves in parallel-plate **transmission lines** (TL) we will assume $W \gg d$ and neglect the effects of fringing fields.

  - Guided waves produce wavelike surface charge and current variations on plate surfaces.
  - Conversely, wavelike charge and current variations on plate surfaces would produce guided wave fields.

It is sufficient to apply a time-varying current and/or charge density at some location z on a parallel-plate TL — e.g., by a time-varying voltage or current source — in order to "excite" the TL with propagating guided fields.

How such excitations propagate away from their "source points" on TL systems will be our main subject of study for the rest of the semester.



- In a parallel-plate TL we ignore any fringing fields and assume that TEM wave fields

$$\mathbf{E} = \hat{x}E_x(z,t) \quad \text{and} \quad \mathbf{H} = \hat{y}H_y(z,t)$$

occupy the region between the plates. For these fields uniform in x and y, Faraday's and Ampere's laws reduce to scalar expressions

$$\nabla \times \mathbf{E} = -\mu\frac{\partial \mathbf{H}}{\partial t} \quad \Rightarrow \quad \frac{\partial E_x}{\partial z} = -\mu\frac{\partial H_y}{\partial t}$$

and

$$\nabla \times \mathbf{H} = \sigma\mathbf{E} + \epsilon\frac{\partial \mathbf{E}}{\partial t} \quad \Rightarrow \quad -\frac{\partial H_y}{\partial z} = \sigma E_x + \epsilon\frac{\partial E_x}{\partial t}.$$

- Now, multiply both equations by d and let

$$V \equiv E_x d \quad \text{voltage drop from plate 2 to plate 1}$$

to obtain

$$\frac{\partial V}{\partial z} = -\mu d\frac{\partial H_y}{\partial t} \quad \text{and} \quad -d\frac{\partial H_y}{\partial z} = \epsilon\frac{\partial V}{\partial t} + \sigma V.$$

- Next, multiply these with W and let

$$I \equiv H_y W \quad \text{current in } z\text{-direction on plate 2}$$

(because $J_{sz} = H_y$ on plate 2) to obtain

$$W\frac{\partial V}{\partial z} = -\mu d\frac{\partial I}{\partial t} \quad \text{and} \quad -d\frac{\partial I}{\partial z} = \epsilon W\frac{\partial V}{\partial t} + \sigma W V.$$

Note that voltage drop

$$V = \int_2^1 \mathbf{E}\cdot d\mathbf{l} = E_x d$$

is uniquely defined — *independent of integration path* — on constant z surfaces because with TEM fields

$$B_z = \mu H_z = 0,$$

and consequently circulation

$$\oint_C \mathbf{E}\cdot d\mathbf{l} = -\frac{d}{dt}\int_S \mathbf{B}\cdot d\mathbf{S} = 0$$

when C is on constant z plane and $d\mathbf{S} = \pm dxdy\hat{z}$.

- We can re-write these equations as

$$-\frac{\partial V}{\partial z} = \mathcal{L}\frac{\partial I}{\partial t} \quad \text{and} \quad -\frac{\partial I}{\partial z} = \mathcal{C}\frac{\partial V}{\partial t} + \mathcal{G}V$$

utilizing

$$\mathcal{L} = \mu\frac{d}{W}, \quad \mathcal{C} = \epsilon\frac{W}{d}, \quad \mathcal{G} = \sigma\frac{W}{d}$$

appropriate for the parallel-plate TL — we recognize these parameters as inductance, capacitance, and conductance of the parallel plate TL.

  - In the equations above the $\mathcal{G}V$ term accounts for Ohmic losses of wave fields having to do with currents leaking between the wires (plates) of the TL.
  - Another possible loss term that we have not picked up — because we assumed infinite conducting plates — is a missing $\mathcal{R}I$ term in the right-hand-side of the first equation.

Telegrapher's equations:

$$-\frac{\partial V}{\partial z} = \mathcal{L}\frac{\partial I}{\partial t}$$
$$-\frac{\partial I}{\partial z} = \mathcal{C}\frac{\partial V}{\partial t}$$

Rather than correcting for that at this stage, we will drop the $\mathcal{G}V$ term from the second equation, and focus our attention *for a while* (until the last day of the semester, in fact) on *ideal* lossless transmission lines governed by the equations shown in the margin — they are known as known as **telegrapher's equations**[1].

[1]Telegrapher's equations were first compiled by Oliver Heaviside (of close-up method, unit-step, and countless other contributions) in 1880's. Telegraphy was being used worldwide by 1850's as a means of rapid communications.

- Except for $-\frac{\partial}{\partial z}$ on the left, the telegrapher's equations look like the $V - I$ relations of inductors and capacitors (which is the best way of remembering them).

- The equations can be readily combined to obtain a 1D scalar wave equation
$$\frac{\partial^2 V}{\partial z^2} = \mathcal{L}\mathcal{C}\frac{\partial^2 V}{\partial t^2}.$$
In analogy to
$$\frac{\partial^2 E_x}{\partial z^2} = \mu\epsilon\frac{\partial^2 E_x}{\partial t^2},$$
the wave equation for V has d'Alembert wave solutions
$$V(z,t) = f(t \mp \frac{z}{v}) \quad \text{where} \quad v \equiv \frac{1}{\sqrt{\mathcal{L}\mathcal{C}}} = \frac{1}{\sqrt{\mu\epsilon}}.$$

- In that case the second telegrapher's equation demands
$$-\frac{\partial I}{\partial z} = \mathcal{C}\frac{\partial V}{\partial t} = \mathcal{C}f'(t \mp \frac{z}{v})$$
implying an anti-derivative
$$I(z,t) = \pm\mathcal{C}v\, f(t \mp \frac{z}{v}) = \pm\frac{f(t \mp \frac{z}{v})}{Z_o}$$
with
$$Z_o \equiv \frac{1}{\mathcal{C}v} = \frac{\sqrt{\mathcal{L}\mathcal{C}}}{\mathcal{C}} = \sqrt{\frac{\mathcal{L}}{\mathcal{C}}} = \frac{1}{\mathrm{GF}}\sqrt{\frac{\mu}{\epsilon}}.$$

Telegrapher's equations:

$$\begin{aligned} -\frac{\partial V}{\partial z} &= \mathcal{L}\frac{\partial I}{\partial t} \\ -\frac{\partial I}{\partial z} &= \mathcal{C}\frac{\partial V}{\partial t} \end{aligned}$$

where

$$\mathcal{C} = \epsilon\mathbf{GF},\ \mathcal{L} = \frac{\mu}{\mathbf{GF}},$$

with "geometrical factor"

$$\begin{aligned} \mathbf{GF} &= \frac{W}{d} \quad \textbf{parallel-plate} \\ &= \frac{2\pi}{\ln\frac{b}{a}} \quad \textbf{coax} \\ &= \frac{\pi}{\cosh^{-1}\frac{D}{2a}} \quad \textbf{twin-lead} \end{aligned}$$

- In summary, d'Alembert wave solutions of telegrapher's equations are

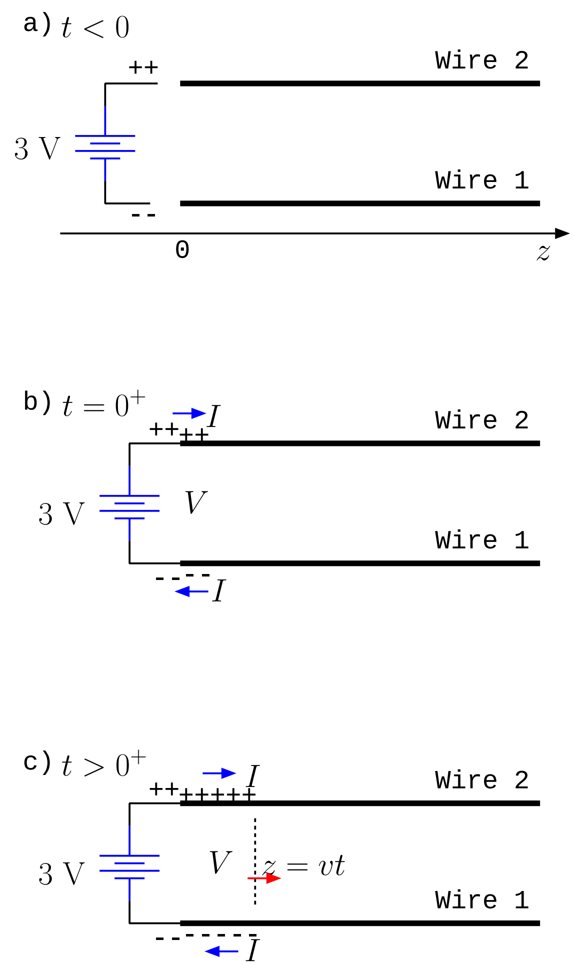
$$V(z,t) = f(t \mp \frac{z}{v}) \quad \text{and} \quad I(z,t) = \pm \frac{f(t \mp \frac{z}{v})}{Z_o}$$

with a propagation speed

$$v = \frac{1}{\sqrt{\mathcal{LC}}} = \frac{1}{\sqrt{\mu\epsilon}}$$

that equals the wave speed of the associated electric and magnetic fields, and voltage-to-current ratio representing a **characteristic impedance**

$$Z_o = \sqrt{\frac{\mathcal{L}}{\mathcal{C}}} = \frac{1}{\mathrm{GF}}\sqrt{\frac{\mu}{\epsilon}}.$$

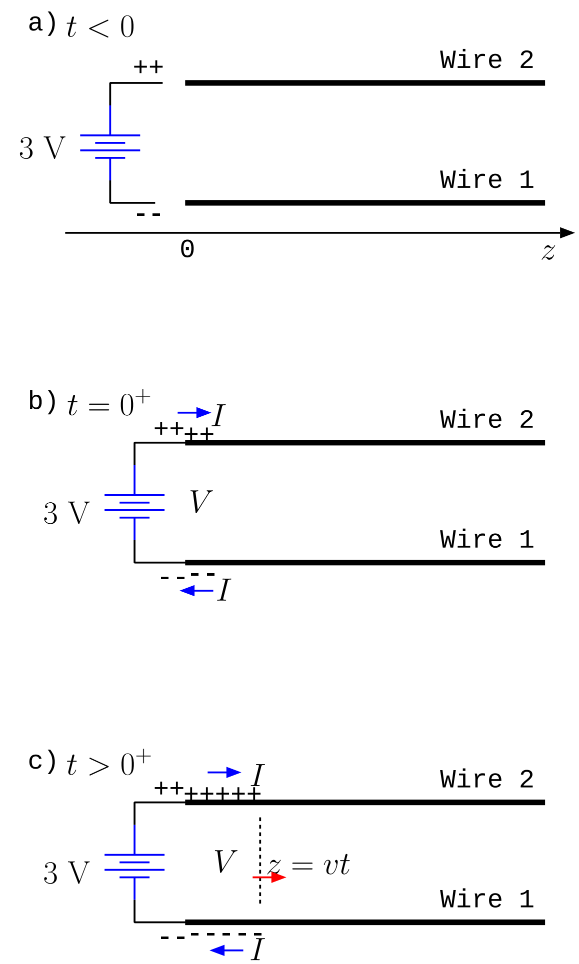


Telegrapher's equations and their d'Alembert solutions provide us with a "handle" on the following **physics**:

- Suppose that + and - terminals of a 3 V battery makes contact with the terminals of a charge neutral TL at $t = 0$ as depicted in the margin. We will assume that $V(z,t) = I(z,t) = 0$ on the TL for $t < 0$.

As soon as contact is made between the terminals of the battery and the TL, the excess + and - charges on battery terminals will "spill onto" the TL terminals as shown in the figure for $t = 0^+$:

- what really happens is,



– electrons move from the - terminal of the battery onto the bottom wire of the TL,

– replenished by an equal amount of electrons moving from the top wire into the battery via its + terminal,

giving the overall impression of current flows I (in opposite direction to electron motion) as marked on the two wires in the diagram.

– currents I and voltage V marked in the diagram are confined only to location $z = 0^+$ at $t = 0^+$, while there is still zero current on the rest of the TL!!!

Having unequal currents on a length of wire is in conflict with our notions from earlier circuit courses, but that's because earlier courses taught us "**lumped-circuit analysis**", an approximate technique justified when it's OK to ignore certain time delays of charge movements in the circuit (when wire lengths are sufficiently short).

Having unequal currents on the TL wire is really what happens

– because, for instance, electrons at some $z > 0$ on the top wire will start moving towards the battery terminal only after the neighboring electrons at z^- deplete the region leaving some excess positive charge.

Thus, currents I on the wires, and voltage V defined and measured across the wires, spread out of $z = 0$ region at a finite speed v, in analogy with ripples spreading out on a pond surface when perturbed by a falling pebble.

- **Telegrapher's equations and their d'Alembert solutions will allow us to calculate how I and V evolve on the TL in quantitative terms.**

To appreciate the distinction between lumped and **distributed circuit analysis**, we next develop a lumped circuit model of a very short length of a TL over which lumped circuit notions may be applicable:

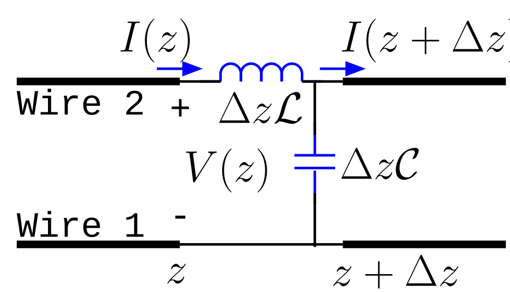
$$
\begin{aligned}
-\frac{\partial V}{\partial z} &= \mathcal{L}\frac{\partial I}{\partial t} \\
-\frac{\partial I}{\partial z} &= \mathcal{C}\frac{\partial V}{\partial t}
\end{aligned}
$$

- Let us re-write the first telegrapher's equation as

$$-\Delta V \equiv V(z,t) - V(z+\Delta z,t) = \Delta z\mathcal{L}\frac{\partial I}{\partial t}$$

  after approximating the left side as a ratio of infinitesimals.

  - This relation shows that in the current flow direction there is an infinitesimal inductive voltage drop of $\Delta z\mathcal{L}\frac{\partial I}{\partial t}$ between points z and $z+\Delta z$ on the wire carrying current $I \equiv I(z,t) \approx I(z+\Delta z,t)$.



- Likewise, the second equation re-arranged as

$$-\Delta I \equiv I(z,t) - I(z+\Delta z,t) = \Delta z\mathcal{C}\frac{\partial V}{\partial t},$$

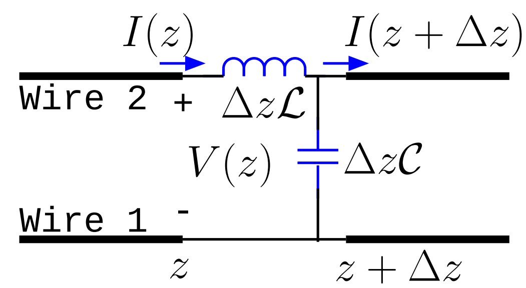
  - this shows that an infinitesimal capacitor current $\Delta z\mathcal{C}\frac{\partial V}{\partial t}$ flows out of a node located between z and $z+\Delta z$ on the wire with current I into a node on the second wire at the same location.

Evidently, a short section Δz of the TL has an equivalent T-network with

1. a series inductance $\Delta z\mathcal{L}$ carrying a current $I(z,t) \approx I(z+\Delta z,t)$, and

2. a shunt capacitance $\Delta z\mathcal{C}$ carrying a voltage $V(z+\Delta z,t) \approx V(z,t)$

as shown in the margin.



This lumped-circuit equivalent is only accurate for Δz so small that

$$I(z,t) \approx I(z+\Delta z,t) \quad \text{and} \quad V(z+\Delta z,t) \approx V(z,t)$$

are both true, which is of course possible only if $\Delta z \ll \lambda$, λ being the shortest wavelength in $I(z,t) \propto \mathbf{H}(z,t)$ and $V(z,t) \propto \mathbf{E}(z,t)$ waveforms.

- Going back to parallel-plate TL in TEM mode, observe that the total power transported in the guide will be the Poynting vector $\mathbf{E}\times\mathbf{H} = E_x H_y \hat{z}$ times the cross-sectional area of the guide, namely, Wd.

  Thus, power transported in z direction is

$$\begin{aligned} p(z,t) &= WdE_x(z,t)H_y(z,t), \\ &= (E_x(z,t)d)(H_y(z,t)W) = V(z,t)I(z,t) \end{aligned}$$

  the familiar formula from circuit theory.

Hence, the circuit theory formula

$$P = \frac{1}{2}\mathrm{Re}\{\tilde{V}\tilde{I}^*\}$$

for average power will also hold in sinusoidal-steady state TL problems when we use phasors $\tilde{V}(z)$ and $\tilde{I}(z)$ to represent the $V(z,t)$ and $I(z,t)$ waveforms.

TL's can also support non-TEM modes having non-zero components of H_z or E_z. These modes are non-propagating (evanescent) at low frequencies and remain localized near their excitation regions (e.g., discontinuity points on the line) if $d < \frac{\lambda}{2}$ (pp TL) or if $a+b < \frac{\lambda}{\pi}$ (coax). At high frequencies when these modes cannot be avoided with practical dimensions d, a, and b, it may be practicable to use them rather than the TEM mode. Use single-wire waveguides in that case instead of two-wire TL's.

28 Distributed circuits and bounce diagrams

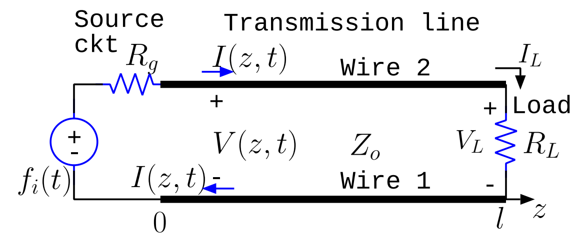
Last lecture we learned that voltage and current variations on TL's are governed by telegrapher's equations and their d'Alembert solutions — the latter can be expressed as

$$V(z,t) = f(t - \frac{z}{v}) + g(t + \frac{z}{v}) \text{ and } I(z,t) = \frac{f(t-\frac{z}{v})}{Z_o} - \frac{g(t+\frac{z}{v})}{Z_o}$$

in terms of

$$v = \frac{1}{\sqrt{\mathcal{LC}}} \text{ and } Z_o = \sqrt{\frac{\mathcal{L}}{\mathcal{C}}}$$

and functions $f(t)$ and $g(t)$ corresponding to signal waveforms propagated in $+z$ and $-z$ directions, respectively.

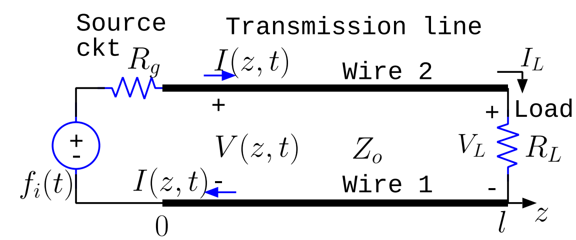


- In this lecture we will learn how to solve **distributed circuit problems** containing TL segments *and* two terminal elements such as resistors and voltage (or current) sources. In solving the problems, we will apply the usual rules of lumped circuit analysis at element terminals and treat the TL's in terms of d'Alembert solutions above.

- Consider a TL with a characteristic impedance Z_o extending from $z = 0$ to $z = l$, where a two-terminal *source* circuit (e.g., a receiving antenna) modeled by a Thevenin equivalent with voltage $f_i(t)$ and resistance R_g is connected between the TL terminals at $z = 0$ and a *load* (e.g., a receiver circuit) modeled by a resistance R_L terminates the line at $z = l$ (see margin).

– We want to determine voltage and current signals $V(z,t)$ and $I(z,t)$ on the TL and the load R_L for time $t > 0$ in terms of source signal $f_i(t)$ assuming that $f_i(t) = 0$ for $t < 0$.

- Using the d'Alembert solutions $V(z,t)$ and $I(z,t)$ from above at $z = l$, we have



$$\frac{V(\ell,t)}{I(\ell,t)} = \frac{f(t-\frac{\ell}{v}) + g(t+\frac{\ell}{v})}{\frac{f(t-\frac{\ell}{v})}{Z_o} - \frac{g(t+\frac{\ell}{v})}{Z_o}} = Z_o\frac{f(t-\frac{\ell}{v}) + g(t+\frac{\ell}{v})}{f(t-\frac{\ell}{v}) - g(t+\frac{\ell}{v})} = \frac{V_L}{I_L} = R_L,$$

from which we obtain

$$g(t+\frac{l}{v}) = \underbrace{\frac{R_L - Z_o}{R_L + Z_o}}_{\Gamma_L} f(t-\frac{l}{v}) \quad\Rightarrow\quad g(t) = \Gamma_L f(t-\frac{2l}{v})$$

where

$$\Gamma_L = \frac{R_L - Z_o}{R_L + Z_o}$$

is the **load reflection coefficient** in the TL circuit. We can re-write the d'Alembert solution for $V(z,t)$ and $I(z,t)$ in terms of only $f(t)$ as

$$V(z,t) = f(t-\frac{z}{v}) + \Gamma_L f(t+\frac{z}{v}-\frac{2l}{v}) \text{ and } I(z,t) = \frac{f(t-\frac{z}{v})}{Z_o} - \frac{\Gamma_L f(t+\frac{z}{v}-\frac{2l}{v})}{Z_o}.$$

- Assuming that $f_i(t) = 0 = f(t)$ for $t < 0$, we can relate $f(t)$ to $f_i(t)$ in $t > 0$ interval using the KVL equation at $z = 0$ that states

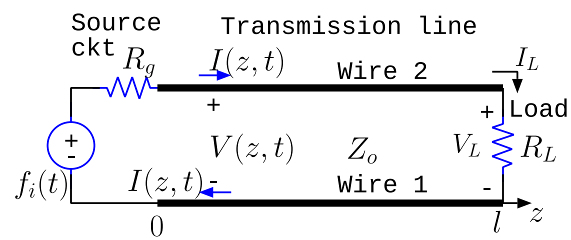
$$f_i(t) = R_g I(0,t) + V(0,t),$$

which is, using $V(z,t)$ and $I(z,t)$ at $z=0$,

$$f_i(t) = R_g \underbrace{\left(\frac{f(t)}{Z_o} - \frac{\Gamma_L f(t-\frac{2l}{v})}{Z_o}\right)}_{I(0,t)} + \underbrace{f(t) + \Gamma_L f(t-\frac{2l}{v})}_{V(0,t)}.$$

Now, since $f(t-\frac{2l}{v}) = 0$ for $t - \frac{2l}{v} < 0$, we find out that for the *epoch* (or *time interval*) $0 < t < \frac{2l}{v}$,

$$f_i(t) = R_g \frac{f(t)}{Z_o} + f(t) \quad \Rightarrow \quad f(t) = \underbrace{\frac{Z_o}{R_g + Z_o}}_{\tau_g} f_i(t)$$



where

$$\tau_g = \frac{Z_o}{R_g + Z_o}$$

is the **injection coefficient** of the TL circuit[1].

- Thus, for the epoch $0 < t < \frac{2l}{v}$, we have the voltage and current solutions

$$V(z,t) = \tau_g f_i(t - \frac{z}{v}) + \Gamma_L \tau_g f_i(t + \frac{z}{v} - \frac{2l}{v}) \quad \text{and} \quad I(z,t) = \frac{\tau_g f_i(t-\frac{z}{v})}{Z_o} - \frac{\Gamma_L \tau_g f_i(t+\frac{z}{v}-\frac{2l}{v})}{Z_o}$$

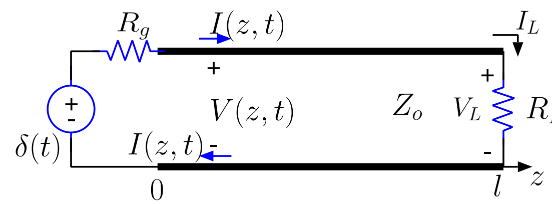
on the line.

[1]Note how $f(t)$ appears to be related to $f_i(t)$ according to a voltage division rule with Z_o representing the resistance across which voltage $f(t)$ is measured.

– So far $f_i(t)$ function is arbitrary and the above results would also be valid for $f_i(t) = \delta(t)$, Dirac's impulse, in which case

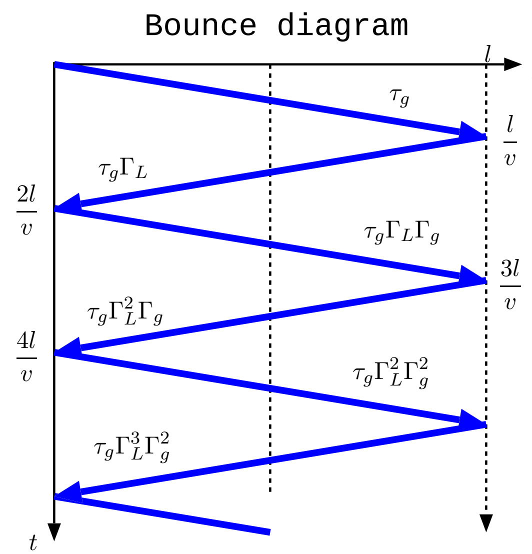
$$V(z,t) = \tau_g \delta(t - \frac{z}{v}) + \Gamma_L \tau_g \delta(t + \frac{z}{v} - \frac{2l}{v}) \text{ and } I(z,t) = \frac{\tau_g \delta(t - \frac{z}{v})}{Z_o} - \frac{\Gamma_L \tau_g \delta(t + \frac{z}{v} - \frac{2l}{v})}{Z_o}$$

would be the voltage and current **impulse response functions** of the TL circuit for the $0 < t < \frac{2l}{v}$ epoch.



- To extend the impulse response functions above to the "next epoch" $\frac{2l}{v} < t < \frac{4l}{v}$, we note that at $z = 0$ the KVL equation with $f_i(t) = \delta(t)$ reads as

$$\delta(t) = R_g \underbrace{(\frac{f(t)}{Z_o} - \frac{\Gamma_L f(t - \frac{2l}{v})}{Z_o})}_{I(0,t)} + \underbrace{f(t) + \Gamma_L f(t - \frac{2l}{v})}_{V(0,t)}.$$



which can be re-arranged as

$$\delta(t) = (1 + \frac{R_g}{Z_o}) f(t) + (1 - \frac{R_g}{Z_o}) \Gamma_L \underbrace{f(t - \frac{2l}{v})}_{\tau_g \delta(t - \frac{2l}{v})},$$

where for $f(t - \frac{2l}{v})$ we used a delayed copy of $f(t) = \tau_g f_i(t)$ solution for $f(t)$ from the *previous epoch* in view of the time delay $\frac{2l}{v}$ contained within $f(t - \frac{2l}{v})$.

– Hence, solving this for $f(t)$, we find, for *this* epoch,

$$f(t) = \tau_g \delta(t) + \underbrace{\frac{R_g - Z_o}{R_g + Z_o}}_{\Gamma_g} \Gamma_L \tau_g \delta(t - \frac{2l}{v}),$$

where

$$\Gamma_g = \frac{R_g - Z_o}{R_g + Z_o}$$

is the **source reflection coefficient** of the TL circuit.

– Substituting $f(t)$ for the epoch $\frac{2l}{v} < t < \frac{4l}{v}$ within voltage and current formulae

$$V(z,t) = f(t - \frac{z}{v}) + \Gamma_L f(t + \frac{z}{v} - \frac{2l}{v}) \text{ and } I(z,t) = \frac{f(t - \frac{z}{v})}{Z_o} - \frac{\Gamma_L f(t + \frac{z}{v} - \frac{2l}{v})}{Z_o}$$

we obtain the "extended" voltage and current impulse response functions

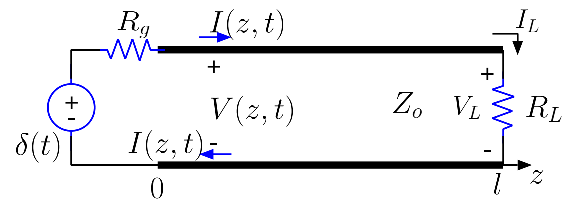
$$V(z,t) = \tau_g \delta(t - \frac{z}{v}) + \Gamma_L \tau_g \delta(t + \frac{z}{v} - \frac{2l}{v}) + \Gamma_g \Gamma_L \tau_g \delta(t - \frac{z}{v} - \frac{2l}{v}) + \Gamma_g \Gamma_L^2 \tau_g \delta(t + \frac{z}{v} - \frac{4l}{v})$$

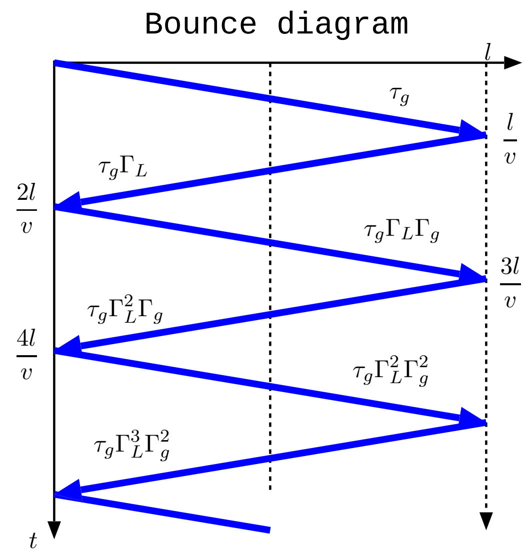
and

$$I(z,t) = Z_o^{-1}[\tau_g \delta(t - \frac{z}{v}) - \Gamma_L \tau_g \delta(t + \frac{z}{v} - \frac{2l}{v}) + \Gamma_g \Gamma_L \tau_g \delta(t - \frac{z}{v} - \frac{2l}{v}) - \Gamma_g \Gamma_L^2 \tau_g \delta(t + \frac{z}{v} - \frac{4l}{v})]$$

respectively.

- ◦ At this point the algebra is pretty messy, but a straightforward pattern is emerging (to obviate the need for algebraic analysis for the upcoming epochs) that is best appreciated with the help of **bounce diagrams** explained next:





– A *bounce diagram* is a plot of the "trajectories" of traveling impulses found on transmission line segments excited by impulse inputs.

– The horizontal axis represents position z of the traveling impulses while time t is represented by a downward pointing axis.

– The first slanted line on the top of the diagram, representing the traveling impulse

$$\tau_g \delta(t - \frac{z}{v}),$$

(first term of $h_z(t) = V(z,t)$) is "reflected" at time $t = \frac{\ell}{v}$ from load R_L to turn into a backward propagating impulse

$$\tau_g \Gamma_L \delta(t + \frac{z}{v} - \frac{2\ell}{v})$$

represented by the second line of the diagram.

– The backward propagating impulse reaches $z = 0$ at $t = \frac{2\ell}{v}$ and is reflected once more with a reflection coefficient

$$\Gamma_g = \frac{R_g - Z_o}{R_g + Z_o}$$

to become a forward propagating impulse

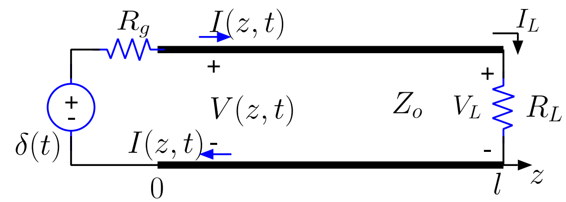
$$\tau_g \Gamma_L \Gamma_g \delta(t - \frac{z}{v} - \frac{2\ell}{v})$$

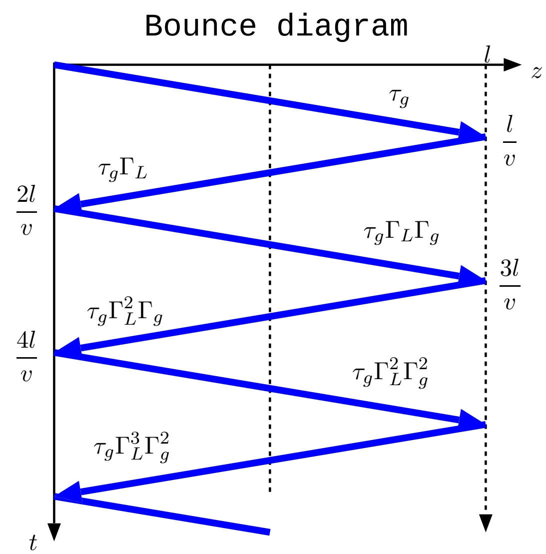
represented by the third line of the diagram.

◦ Reflection at R_g is in effect the same physical process as reflection at R_L and therefore its coefficient Γ_g is identical with Γ_L except for the replacement of R_L by R_g.

– The bounce diagram is advanced in time with further reflections occurring at both ends.

– We show the calculated weights of traveling impulses directly on the diagram just above the slanted lines representing the trajectories of each traveling impulse (each having a lifetime of ℓ/v)





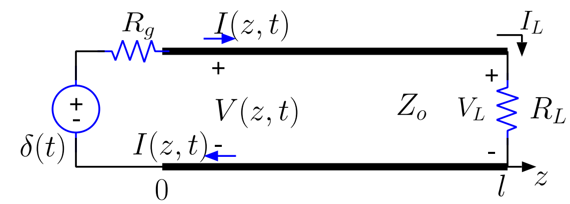
- Using the bounce diagram, the full expressions for the voltage and current impulse response functions of the circuit can be written as

$$\begin{aligned} V(z,t) &= \tau_g \sum_{n=0}^{\infty} (\Gamma_L \Gamma_g)^n \delta(t - \frac{z}{v} - n\frac{2\ell}{v}) \\ &+ \tau_g \Gamma_L \sum_{n=0}^{\infty} (\Gamma_L \Gamma_g)^n \delta(t + \frac{z}{v} - (n+1)\frac{2\ell}{v}) \end{aligned}$$

and

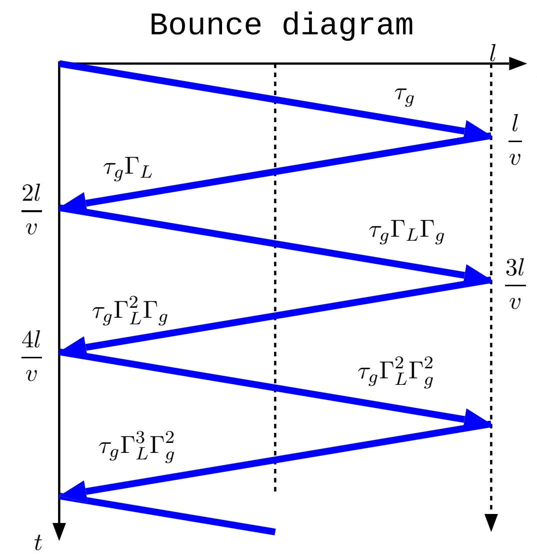
$$\begin{aligned} I(z,t) &= \tfrac{\tau_g}{Z_o} \sum_{n=0}^{\infty} (\Gamma_L \Gamma_g)^n \delta(t - \frac{z}{v} - n\frac{2\ell}{v}) \\ &- \frac{\tau_g}{Z_o} \Gamma_L \sum_{n=0}^{\infty} (\Gamma_L \Gamma_g)^n \delta(t + \frac{z}{v} - (n+1)\frac{2\ell}{v}). \end{aligned}$$

- Although these series formulae look daunting, only the lower order terms usually matter — that is true because $|\Gamma_L| \le 1$ and $|\Gamma_g| \le 1$ and thus $(\Gamma_L\Gamma_g)^n$ is typically a rapidly diminishing function of n (unless the ckt is "dissipation free" and resonant, a concept explored in Lecture 31).



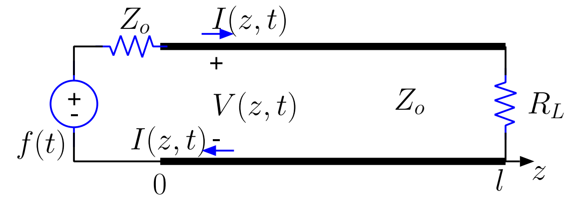
- We typically rely on the bounce diagram technique more so than the series expressions developed above. This will be illustrated by several examples in the next lecture.

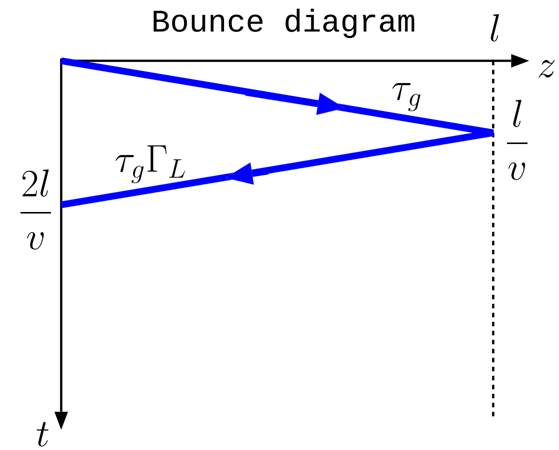
  - The main idea is to combine *delayed* versions of the circuit input $f_i(t)$ with the impulse weights indicated on the bounce diagram, since, in general, the convolution $\delta(t-T_z) * f_i(t) = f_i(t-T_z)$ for any z-dependent delay such as $\frac{z}{v}$, $\frac{z}{v} - \frac{2\ell}{v}$, etc...



29 Bounce diagrams

Source matched to line:





- Last lecture we obtained the **implulse-response functions**

$$V(z,t) = \tau_g[\delta(t - \frac{z}{v}) + \Gamma_L \delta(t + \frac{z}{v} - \frac{2l}{v})]$$

  and

$$I(z,t) = \frac{\tau_g}{Z_o}[\delta(t - \frac{z}{v}) - \Gamma_L \delta(t + \frac{z}{v} - \frac{2l}{v})]$$

  for the voltage and current in the TL circuit shown in the margin where the source is matched to the line so that $\tau_g = \frac{1}{2}$ — circuit response with an arbitrary input $f(t)$ is obtained by convolving these with $f(t)$ (as shown in Example 1 in last lecture).

- The impulse-response for $V(z,t)$ is depicted in the margin in the form of a **bounce diagram**, in which

  – the trajectories of the impulses constituting the impulse response are plotted, with

    ◦ z axis in the horizontal, and

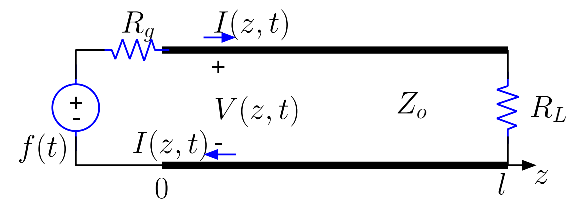
    ◦ t axis in the vertical extending from top to bottom

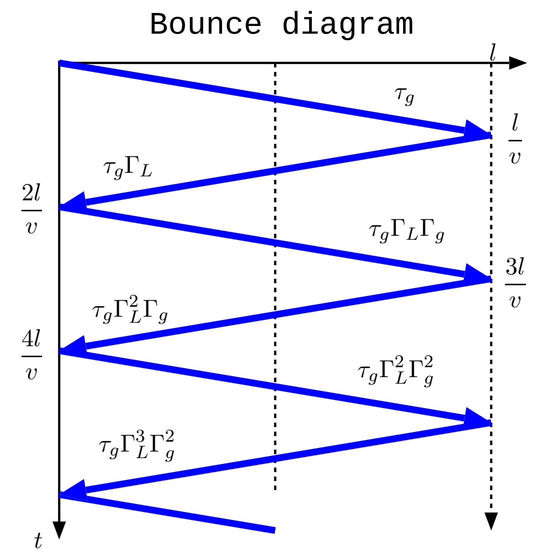
  – and coefficients of each impulse noted in the diagram next to the trajectory lines.

  – the blue line sloping down on the top is a depiction of forward propagating impulse $\tau_g \delta(t - \frac{z}{v})$,

– the next line down is the depiction of backward propagating impulse $\tau_g \Gamma_L \delta(t + \frac{z}{v} - \frac{2l}{v})$.

Bounce diagrams are graphical representations of impulse response functions derived in TL circuit problems, and are *primarily used to determine the impulse response functions*, rather than the other way around as will be illustrated below.





- We show in the margin a circuit with an arbitrary

$$R_g \quad \text{and} \quad \tau_g = \frac{Z_o}{R_g + Z_o},$$

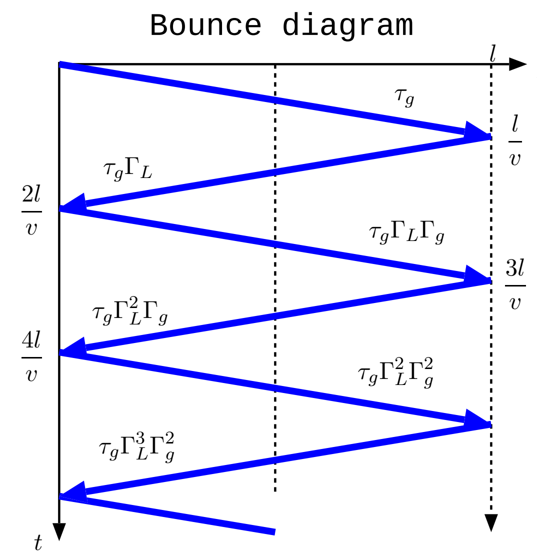
for which the bounce diagram is not terminated at $t = \frac{2l}{v}$ because the backward propagating impulse on the line arriving at $z = 0$ at time $t = \frac{2l}{v}$ is reflected from $z = 0$ with a reflection coefficient of

$$\Gamma_g = \frac{R_g - Z_o}{R_g + Z_o}.$$

– Reflections of negative-going impulses incident on the source circuit are justified because these impulses just see the resistor R_g at the generator end — the source voltage $f(t) = \delta(t)$ is by then just a short is series with R_g — unmatched to Z_o, just like the forward going impulses seeing a load R_L unmatched to Z_o and reflecting with a coefficient

$$\Gamma_L = \frac{R_L - Z_o}{R_L + Z_o}.$$

- Once the bounce diagram for voltage has been constructed as shown above, then the impulse response can be written by inspection as



$$
\begin{aligned}
V(z,t) &= \tau_g \sum_{n=0}^{\infty} (\Gamma_L \Gamma_g)^n \delta(t - \frac{z}{v} - n\frac{2l}{v}) \\
&\quad + \tau_g \Gamma_L \sum_{n=0}^{\infty} (\Gamma_L \Gamma_g)^n \delta(t + \frac{z}{v} - (n+1)\frac{2l}{v}).
\end{aligned}
$$

Also,

$$
\begin{aligned}
I(z,t) &= \tfrac{\tau_g}{Z_o} \sum_{n=0}^{\infty} (\Gamma_L \Gamma_g)^n \delta(t - \frac{z}{v} - n\frac{2l}{v}) \\
&\quad - \frac{\tau_g}{Z_o} \Gamma_L \sum_{n=0}^{\infty} (\Gamma_L \Gamma_g)^n \delta(t + \frac{z}{v} - (n+1)\frac{2l}{v}).
\end{aligned}
$$

  - It can also be shown that the first term of $V(z,t)$ above is derived from the formal solution of the equation

$$
V^+(t) = \tau_g \delta(t) + \Gamma_L \Gamma_g V^+(t - \frac{2l}{v})
$$

    which is obtained from

$$
I(0,t) = \frac{\delta(t) - V(0,t)}{R_g}
$$

    enforced at $z = 0$. We have effectively by-passed such a formal approach to the problem by using the bounce diagram technique.

- These awful series formulae above are hardly needed in most applications when only the first few terms of the series are sufficient for reasonably accurate results (like in the next example).

Example 1: Consider a TL circuit where $Z_o = 50\,\Omega$, $v = c$, $l = 2400$ m, $R_g = 0$, and $R_L = 100\,\Omega$. Determine and plot $V(1200, t)$ if $f(t) = u(t)$.

Solution: For this circuit

$$\tau_g = \frac{Z_o}{R_g + Z_o} = 1, \quad \Gamma_g = \frac{R_g - Z_o}{R_g + Z_o} = -1, \quad \text{and} \quad \Gamma_L = \frac{R_L - Z_o}{R_L + Z_o} = \frac{1}{3}.$$

Also, the transit time across the TL is

$$\frac{l}{v} = \frac{2400\,\text{m}}{300 \times 10^6\,\text{m/s}} = 8\,\mu\text{s}.$$

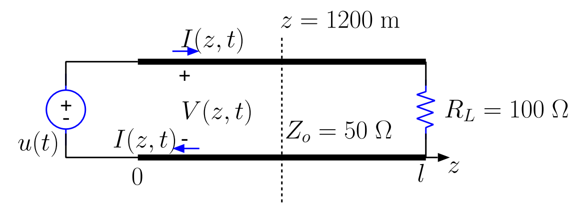
From the bounce diagram shown in the margin, the impulse response for $z = 1200$ m (the location marked by the vertical dashed line) is found to be

$$V(1200, t) = \delta(t-4) + \frac{1}{3}\delta(t-12) - \frac{1}{3}\delta(t-20) - \frac{1}{9}\delta(t-28) + \frac{1}{9}\delta(t-36) + \cdots$$

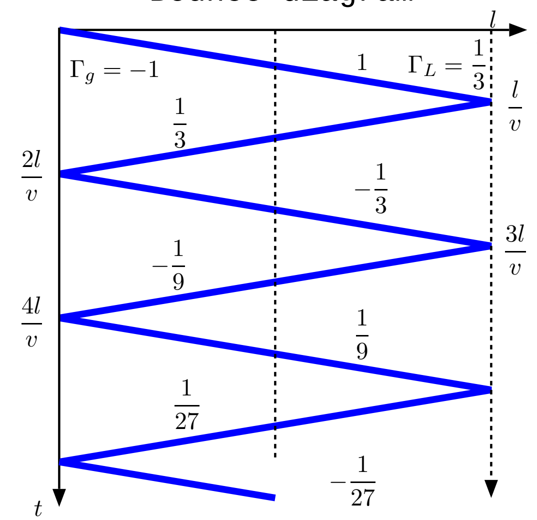
Replacing the $\delta(t)$ in this expression with the unit-step $u(t)$, the specified source function $f(t)$, we get

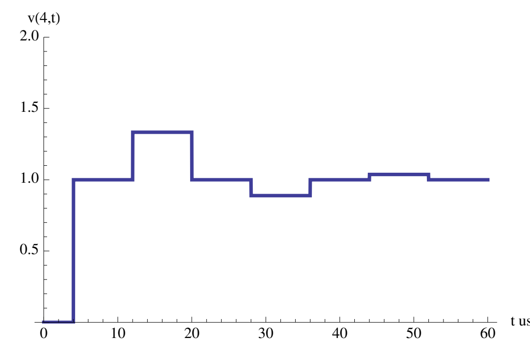
$$V(1200, t) = u(t-4) + \frac{1}{3}u(t-12) - \frac{1}{3}u(t-20) - \frac{1}{9}u(t-28) + \frac{1}{9}u(t-36) + \cdots$$

which is plotted in the margin.



Bounce diagram

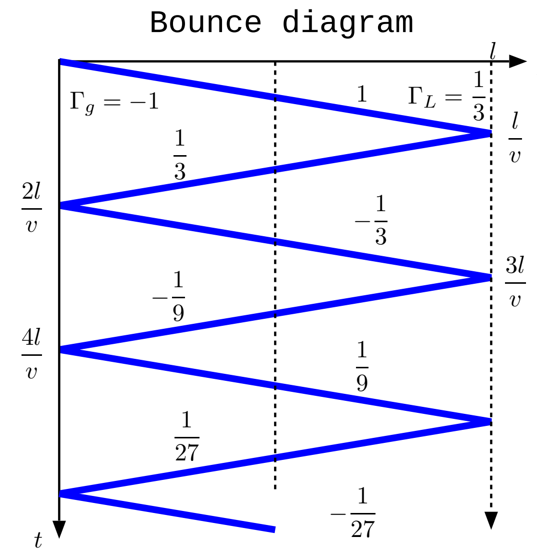


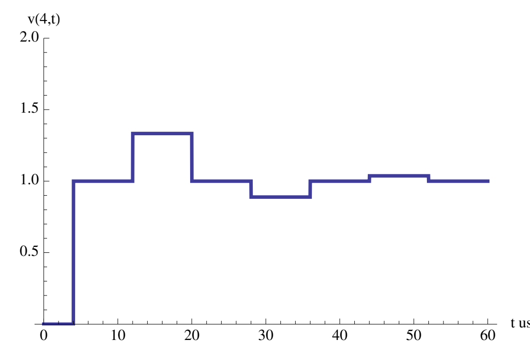


Animated version of this is linked in the class calendar.

- Note that as $t \to \infty$, $V(1200, t) \to 1$ V in Example 1, as if DC conditions prevail and the TL becomes a pair of wires in the lumped circuit sense.

  - DC steady-state corresponds to $\omega = 0$ and signal wavelength $\lambda \to \infty$. In that limit $l \ll \lambda$ is always valid and TL can be treated like an ordinary lumped circuit.
  - Of course this simplification can only occur with $f(t) \propto u(t)$, or its delayed versions, which are all asymptotically DC in $t \to \infty$ limit. The simplification does not apply for $f(t) = \sin(\omega t)u(t)$, for example.





Example 2: In the TL circuit described in Example 1, determine $V(z,t)$ and $I(z,t)$ for a new source signal $f(t) = \text{rect}(\frac{t}{T}) + 2\text{rect}(\frac{t-T}{T})$, $T = 1\ \mu$s. Plot $V(z,t)$ versus z at $t = 3\ \mu s$ and $t = 11\ \mu s$.

Solution: With $\tau_g = 1$, $\Gamma_g = -1$, $\Gamma_L = \frac{1}{3}$, and $\frac{2l}{c} = 16\,\mu$s, we obtain, by convolving with the general impulse response, the voltage response

$$V(z,t) = \sum_{n=0}^{\infty}(-\frac{1}{3})^n f(t - \frac{z}{c} - n16) + \frac{1}{3}\sum_{n=0}^{\infty}(-\frac{1}{3})^n f(t + \frac{z}{c} - (n+1)16)$$

where $\frac{z}{c}$ is to be entered in μs units. Also,

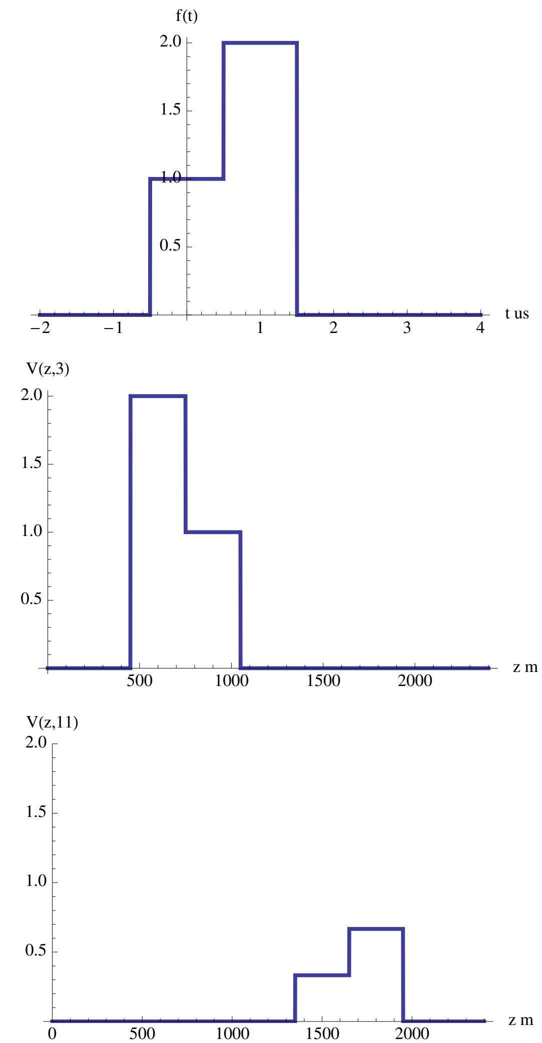
$$I(z,t) = \frac{1}{50}\sum_{n=0}^{\infty}(-\frac{1}{3})^n f(t - \frac{z}{c} - n16) - \frac{1}{50}\frac{1}{3}\sum_{n=0}^{\infty}(-\frac{1}{3})^n f(t + \frac{z}{c} - (n+1)16).$$

At $t = 3\ \mu s$, the voltage variation is

$$V(z,3) = \sum_{n=0}^{\infty}(-\frac{1}{3})^n f(3 - \frac{z}{c} - n16) + \frac{1}{3}\sum_{n=0}^{\infty}(-\frac{1}{3})^n f(3 + \frac{z}{c} - (n+1)16),$$

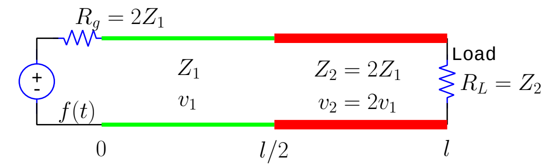
which is plotted in the margin using $f(t) = \text{rect}(t) + 2\text{rect}(t-1)$. Likewise, at $t = 11\ \mu s$,

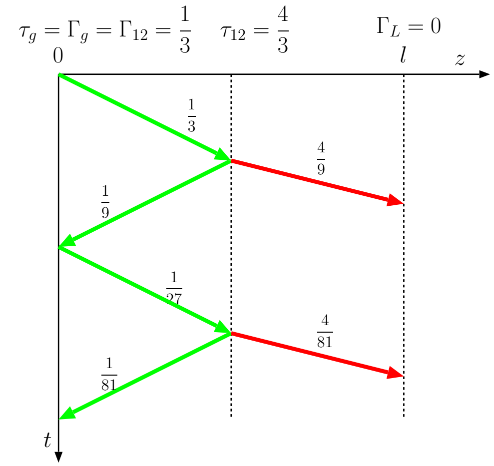
$$V(z,11) = \sum_{n=0}^{\infty}(-\frac{1}{3})^n f(11 - \frac{z}{c} - n16) + \frac{1}{3}\sum_{n=0}^{\infty}(-\frac{1}{3})^n f(11 + \frac{z}{c} - (n+1)16).$$



Animated version of this is linked in the class calendar.

30 Multi-line circuits





- In this lecture we will extend the bounce diagram technique to solve distributed circuit problems involving multiple transmission lines.

- One example of such a circuit is shown in the margin where two distinct TL's of equal lengths have been joined directly at a distance $\frac{l}{2}$ away from the generator.

  - The impulse response of the system can be found by first constructing the bounce diagram for the TL system as shown in the margin.
  - In this bounce diagram, $z = \frac{l}{2}$ happens to be the location of additional reflections as well as transmissions because of the sudden change of Z_o from Z_1 to $Z_2 = 2Z_2$.

These reflections and transmissions between line j and k — transmission from j to k, and reflection from k back to j — can be computed with **reflection coefficient**

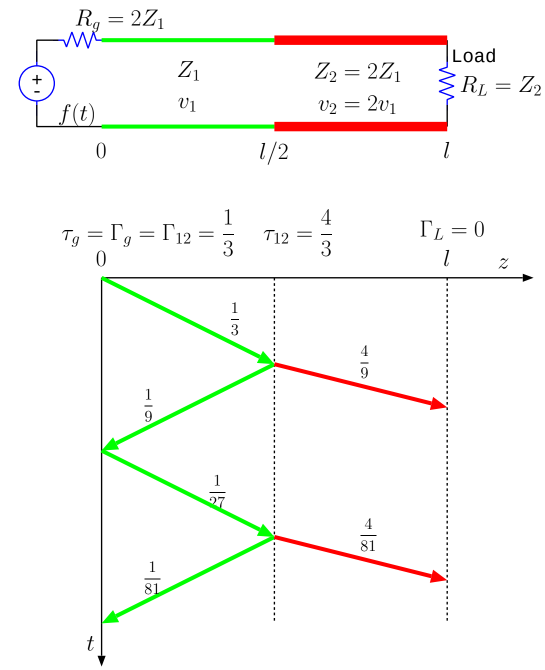
$$\Gamma_{jk} = \frac{Z_k - Z_j}{Z_k + Z_j}$$

and **transmission coefficient**

$$\tau_{jk} = 1 + \Gamma_{jk}$$

that ensure the voltage and current continuity at the junction

- Z_j is the characteristic impedance of the line of the incident pulse, while
- Z_k is the impedance of the cascaded line into which the transmitted pulse is injected.

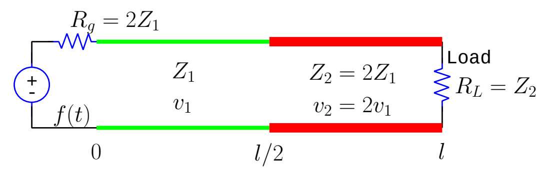


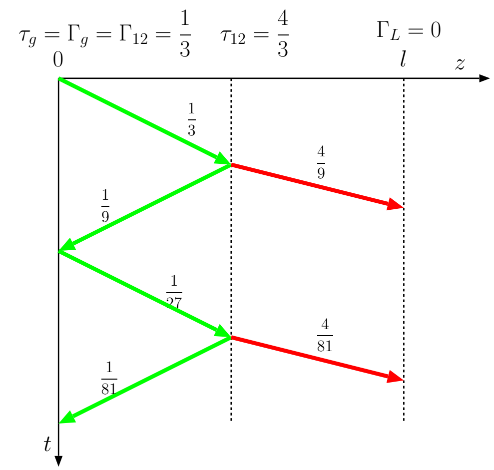
Verification:

- Let
$$V_j^+(1+\Gamma_{jk}) \quad \text{and} \quad V_j^+(1-\Gamma_{jk})/Z_j$$
denote the total voltage and current on line Z_j expressed in terms of incident voltage wave V^+ (of d'Alembert type), and
- let
$$V_j^+\tau_{jk} \quad \text{and} \quad V_j^+\tau_{jk}/Z_k$$
denote the voltage and current on line Z_k adjacent to line Z_j.

This notation identifies Γ_{jk} and τ_{jk} as reflection and transmission coefficients at the junction.

- Taking
$$V_j^+(1+\Gamma_{jk}) = V_j^+\tau_{jk}$$
and
$$V_j^+(1-\Gamma_{jk})/Z_j = V_j^+\tau_{jk}/Z_k$$
in order to enforce voltage and current continuity, we can solve for Γ_{jk} and τ_{jk} expressions stated above.





Example 1: In the circuit shown in the margin with two TL segments, line 2 has twice the characteristic impedance and propagation velocity of line 1, i.e.,

$$Z_2 = 2Z_1 \quad \text{and} \quad v_2 = 2v_1.$$

Determine $\mathcal{L}_2$ and $\mathcal{C}_2$ in terms of $\mathcal{L}_1$ and $\mathcal{C}_1$.

Solution: We have

$$Z_2 = 2Z_1 \quad \Rightarrow \quad \frac{\mathcal{L}_2}{\mathcal{C}_2} = 4\frac{\mathcal{L}_1}{\mathcal{C}_1}$$

and

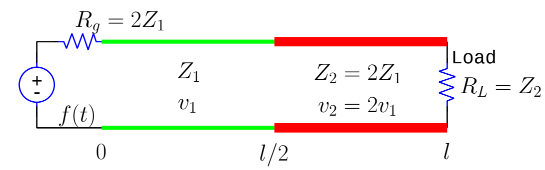
$$v_2 = 2v_1 \quad \Rightarrow \quad \frac{1}{\mathcal{L}_2\mathcal{C}_2} = 4\frac{1}{\mathcal{L}_1\mathcal{C}_1}.$$

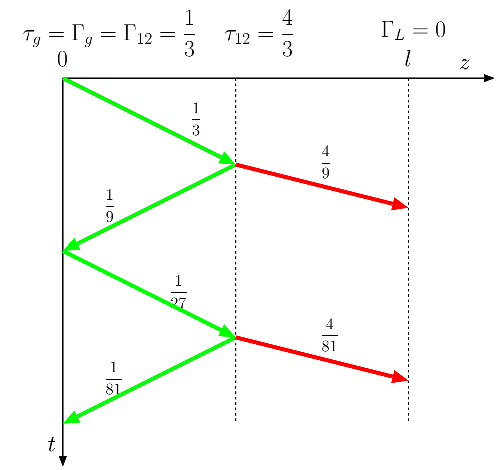
The product of the two equations gives

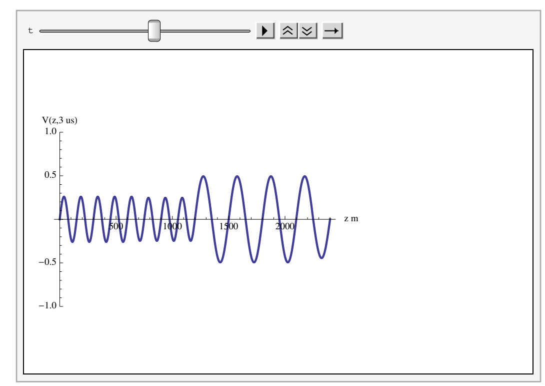
$$\frac{1}{\mathcal{C}_2^2} = 16\frac{1}{\mathcal{C}_1^2} \quad \Rightarrow \quad \mathcal{C}_2 = \frac{1}{4}\mathcal{C}_1,$$

while their ratio leads to

$$\mathcal{L}_2 = \mathcal{L}_1.$$







Example 2: In the circuit of Example 1, determine $V(z,t)$ and $I(z,t)$ if

$$f(t) = \sin(2\pi t)u(t), \;\; t \,\text{in}\; \mu s,$$

and $l = 2400$ m, $v_1 = 150\,\text{m}/\mu\text{s}$, and $Z_1 = 25\,\Omega$.

Solution: From the bounce diagram we infer the following impulse-response for the voltage variable:

$$V(z,t) = \frac{1}{3}\sum_{n=0}^{\infty}(\frac{1}{3})^{2n}[\delta(t - \frac{z}{v_1} - n\frac{l}{v_1}) + \frac{1}{3}\delta(t + \frac{z}{v_1} - (n+1)\frac{l}{v_1})]$$

for $z < \frac{l}{2}$, and

$$V(z,t) = \frac{1}{3}\sum_{n=0}^{\infty}(\frac{1}{3})^{2n}\frac{4}{3}\delta(t - \frac{z}{v_2} - (4n+1)\frac{l/2}{v_2})$$

for $\frac{l}{2} < z < l$. The impulse response for the current is

$$I(z,t) = \frac{1}{3Z_1}\sum_{n=0}^{\infty}(\frac{1}{3})^{2n}[\delta(t - \frac{z}{v_1} - n\frac{l}{v_1}) - \frac{1}{3}\delta(t + \frac{z}{v_1} - (n+1)\frac{l}{v_1})]$$

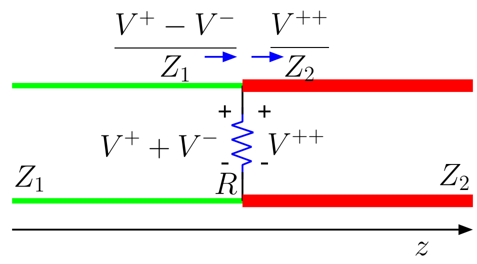
for $z < \frac{l}{2}$, and

$$I(z,t) = \frac{1}{3Z_2}\sum_{n=0}^{\infty}(\frac{1}{3})^{2n}\frac{4}{3}\delta(t - \frac{z}{v_2} - (4n+1)\frac{l/2}{v_2})$$

for $\frac{l}{2} < z < l$. Using

$$\frac{l}{v_1} = \frac{2400}{150} = 16\,\mu\text{s}$$

and replacing $\delta(t)$ with $f(t) = \sin(2\pi t)u(t)$ the plot depicted in the margin was obtained.



Example 3: Two TL's with characteristic impedances Z_1 and Z_2 are joined at a junction that also includes a "shunt" resistance R as shown in the diagram in the margin. Determine the reflection coefficient Γ_{12} and transmission coefficient τ_{12} at the junction.

Solution: Consider a voltage wave

$$V^+(t - \frac{z}{v_1})$$

coming from the left producing reflected and transmitted waves

$$V^-(t + \frac{z}{v_1}) \text{ and } V^{++}(t - \frac{z}{v_2})$$

on lines 1 and 2 traveling to the left and right, respectively, on two sides of the junction. Using an abbreviated notation, KVL and KCL applied at the junction can be expressed as

$$V^+ + V^- = V^{++} \text{ and } \frac{V^+}{Z_1} - \frac{V^-}{Z_1} = \frac{V^{++}}{R} + \frac{V^{++}}{Z_2},$$

where in the KCL equation the first term on the right is the current flowing down the resistor R, and the second term is the TL current on line 2 (as marked in the circuit diagrams in the margin). The equations can be rearranged as

$$\begin{aligned} V^+ + V^- &= V^{++} \\ V^+ - V^- &= \frac{Z_1}{Z_{eq}} V^{++}, \end{aligned}$$

where

$$Z_{eq} \equiv \frac{RZ_2}{R + Z_2}$$

is the parallel combination of R and Z_2. Solving these equations, we find that

$$\Gamma_{12} \equiv \frac{V^-}{V^+} = \frac{Z_{eq} - Z_1}{Z_{eq} + Z_1}$$

and

$$\tau_{12} = \frac{V^{++}}{V^+} = \frac{2Z_{eq}}{Z_{eq} + Z_1}.$$

By, symmetry, the coefficients

$$\Gamma_{21} = \frac{Z_{eq} - Z_2}{Z_{eq} + Z_2}$$

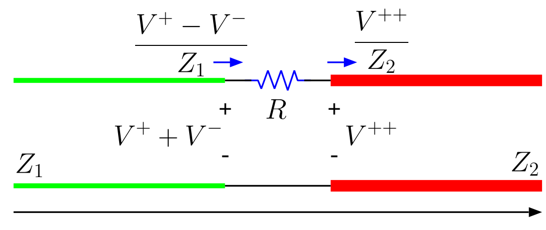
and

$$\tau_{21} = \frac{2Z_{eq}}{Z_{eq} + Z_2}$$

would describe reflection and transmission when a wave is incident from right provided that

$$Z_{eq} \equiv \frac{RZ_1}{R + Z_1}$$

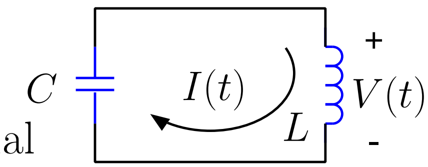
is used.



Hint: in this ckt Γ_{12} has the usual form in terms of $Z_{eq} \equiv R + Z_2$. For τ_{12}, we need $1 + \Gamma_{12}$ multiplied by a voltage division factor $Z_2/(R + Z_2)$.

Exercise: Two TL's with characteristic impedances Z_1 and Z_2 are joined at a junction that also includes a series resistance R as shown in the margin. Determine the reflection coefficient Γ_{12} and transmission coefficient τ_{12} at the junction.

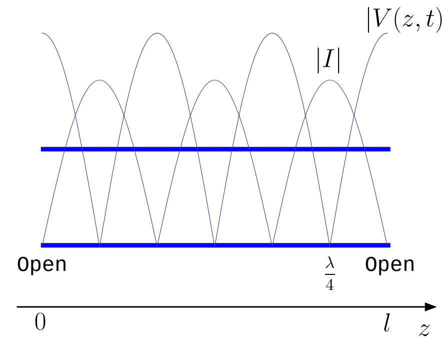
31 Periodic oscillations in lossless TL ckts



- Lossless LC circuits (see margin) can support source-free and co-sinusoidal voltage and current oscillations at a frequency of

$$\omega = \frac{1}{\sqrt{LC}}$$

known as LC **resonance frequency**.



A TL segment — a "stub", so to speak — open circuited at both ends can support voltage and current oscillations such that the current waveform vanishes at both ends. Absolute values of a possible set of voltage and current waveforms satisfying this boundary condition are depicted above.

- Lossless TL circuits can also support source-free voltage and current oscillations, but the number of resonance frequencies is *infinite* and the oscillation waveforms are not restricted to co-sinusoidal forms.
  - Resonance frequencies of lossless TL's are **harmonically related**, and therefore superpositions of resonant oscillations on TL's can add up to arbitrary periodic waveforms as in Fourier series representation of periodic functions.

In this lecture we will examine the periodic oscillations and resonances encountered in lossless and source-free TL circuits.

- Consider first a TL segment of some length ℓ having no electrical connection to any elements at either end, as shown in the margin.

  – Effectively, both ends of the TL have been "open circuited", and thus TL current $I(z,t)$ needs to vanish at $z=0$ and $z=\ell$. Since

  $$I(z,t)=\frac{f(t-\frac{z}{v})}{Z_o}-\frac{g(t+\frac{z}{v})}{Z_o}$$

  in general, these boundary conditions

  $$I(0,t)=\frac{f(t)}{Z_o}-\frac{g(t)}{Z_o}=0$$

  and

  $$I(l,t)=\frac{f(t-\frac{\ell}{v})}{Z_o}-\frac{g(t+\frac{\ell}{v})}{Z_o}=0$$

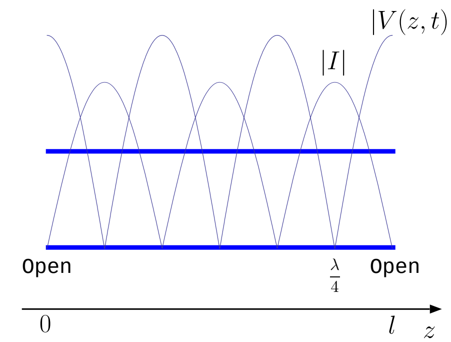
  require that

    ◦ $g(t)=f(t)$
    ◦ $f(t-\frac{\ell}{v})=f(t+\frac{\ell}{v}) \quad \Rightarrow \quad f(t)=f(t+\frac{2\ell}{v})$.

  – the first condition says that forward and backward going waves are described in terms of a single time function $f(t)$,

  – while the second condition indicates that function $f(t)$ is necessarily periodic with a

    ◦ **period** $T=\frac{2\ell}{v}$
    ◦ ***fundamental*** **frequency** $\omega_o=\frac{2\pi}{T}=\frac{\pi v}{\ell}$



A TL "stub" open circuited at both ends can support voltage and current oscillations such that the current waveform vanishes at both ends. Absolute values of a possible set of voltage and current waveforms satisfying this boundary condition are depicted above.

Since no other constraint is imposed, any waveform with the specified period is admissible, and the most general such expression is given by the **Fourier series**

$$f(t) = F_o + \sum_{n=1}^{\infty} F_n \cos(n\omega_o t + \theta_n)$$

having harmonically related frequencies $n\omega_o$ and arbitrary **Fourier coefficients** F_n and θ_n.

– Hence, in general, the line current

$$\begin{aligned} I(z,t) &= \frac{f(t-\frac{z}{v}) - f(t+\frac{z}{v})}{Z_o} \\ &= \sum_{n=1}^{\infty} \frac{F_n}{Z_o}[\cos(n\omega_o t + \theta_n - n\beta_o z) - \cos(n\omega_o t + \theta_n + n\beta_o z)] \end{aligned}$$

where $\beta_o \equiv \omega_o/v = \pi/\ell$ is the ***fundamental*** **wavenumber.**

– The same result written in phasor form is

$$\tilde{I}(z) = \sum_{n=1}^{\infty} \frac{F_n}{Z_o} e^{j\theta_n}[e^{-jn\beta_o z} - e^{jn\beta_o z}] = \sum_{n=1}^{\infty} \frac{F_n}{Z_o} e^{j\theta_n}(-2j)\sin(n\beta_o z),$$

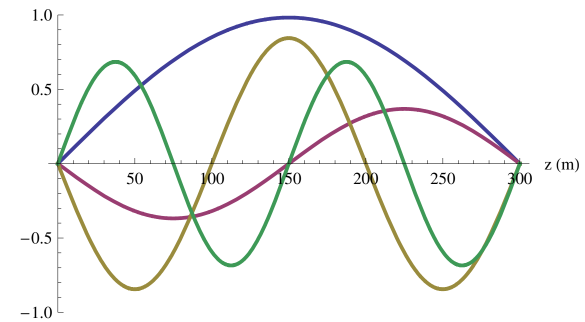
which also means that (back in the time domain)

$$I(z,t) = \sum_{n=1}^{\infty} \frac{2F_n}{Z_o} \sin(n\omega_o t + \theta_n)\sin(n\beta_o z).$$

Also[1]

$$V(z,t) = \sum_{n=1}^{\infty} 2F_n \cos(n\omega_o t + \theta_n)\cos(n\beta_o z)$$

from the phasor $\tilde{V}(z) = \sum_n F_n e^{j\theta_n}[e^{-jn\beta_o z} + e^{jn\beta_o z}]$.



A time-snapshot of the current standing wave modes $n = 1, 2, 3, 4$ on a TL segment 300 m long, open ended on both sides. Each mode n has n half wavelengths fitted into the line length l and high n modes oscillate with higher frequencies. See animation of these modes linked in the class calendar.

In summary:

– Periodic variations of arbitrary complexity — or *timbre*, in analogy with musical instruments — in $V(z,t)$ and $I(z,t)$ are allowed on an open circuited (on both ends) TL segment of length ℓ and consist of superpositions of **resonant modes** (see margin)

$$\cos(n\frac{\pi v}{\ell}t + \theta_n)\cos(n\frac{\pi}{\ell}z) \quad \text{and} \quad \sin(n\frac{\pi v}{\ell}t + \theta_n)\sin(n\frac{\pi}{\ell}z),$$

respectively, in the range $n \geq 1$, each one being a **standing wave**.

– Each resonant mode or standing wave of index $n \geq 1$ has a

∘ resonance frequency

$$\omega = \frac{\pi v}{\ell} n \text{ rad/s} \quad \text{or} \quad f = \frac{v}{2\ell} n \text{ Hz}$$

∘ resonance wavelength

$$\lambda = \frac{v}{f} = \frac{2\ell}{n},$$

[1]Note that an arbitrary DC term can also be included in $V(z,t)$.

implying that

$$\ell = n\frac{\lambda}{2},$$

that is, **the line length is an integer multiple of half-wavelength at each resonance**.

- The resonances examined above also apply to TL's of length ℓ **shorted** at both ends, provided that the mode equations above are swapped between voltage and current — that is, periodic variations of arbitrary complexity in $I(z,t)$ and $V(z,t)$ consist of superpositions of **resonant modes**

  $$\cos(n\frac{\pi v}{\ell}t + \theta_n)\cos(n\frac{\pi}{\ell}z) \quad \text{and} \quad \sin(n\frac{\pi v}{\ell}t + \theta_n)\sin(n\frac{\pi}{\ell}z),$$

  respectively, in the range $n \geq 1$.

  Note that in this case the voltage modes vanish at $z = 0$ and $z = \ell$ as required by the boundary condition $V(0,t) = V(\ell,t) = 0$ imposed by having shorts at both ends.

- For TL's of length ℓ **open at one end shorted at the other** end, resonant wavelengths and frequencies can be identified by requiring ℓ to be an odd multiple of $\frac{\lambda}{4}$

  – the reason for this is, the nulls of waveforms $\propto \cos(\beta z)$ and $\sin(\beta z)$ are separated by odd multiples of

$$\frac{\lambda}{4} = \frac{2\pi/\beta}{4} = \frac{\pi}{2\beta}.$$

  – Hence, resonance condition is

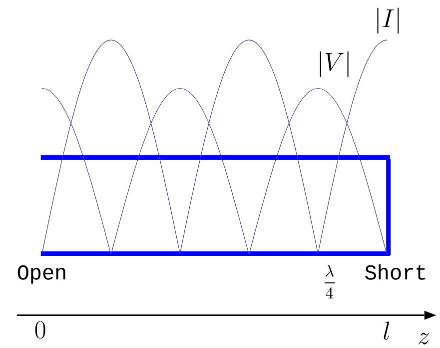
$$\ell = \frac{\lambda}{4}(2n+1), \quad n \geq 0,$$

  and since

$$f\lambda = v$$

  it follows that the resonance frequencies are

$$f = \frac{v}{2\ell}(\frac{1}{2}+n) \quad \text{and} \quad \omega = \frac{\pi v}{\ell}(\frac{1}{2}+n) \text{ for } n \geq 0.$$



A TL stub open at one end short at the other can support voltage and current oscillations such that the current waveform vanishes at the open end while the voltage waveform vanishes at the shorted end. Absolute values of a possible set of voltage and current waveforms satisfying this boundary condition are depicted above. Resonant standing waves modes on this line will have voltage and current nulls separated by an odd multiple of a quarter wavelength.

Example 1: A lossless TL of 600 m length is left open at $z = 0$ and shorted at $z = l = 600$ m. Determine (a) resonant frequencies of the line, (b) resonant current modes, (c) resonant voltage modes obtained from the current modes using the telegrapher's equations. The line has a characteristic impedance of $Z_o = 50\,\Omega$ and a propagation velocity $v = c$.

Solution: (a) The line must be an odd multiple of quarter wavelengths at the resonant frequencies. Therefore,

$$600\,\mathrm{m} = (2n+1)\frac{\lambda}{4} \quad \Rightarrow \quad 600\,\mathrm{m} = (2n+1)\frac{c/f}{4}$$

leading to

$$f = (2n+1)\frac{300\,\mathrm{m}/\mu\mathrm{s}}{4\cdot 600\,\mathrm{m}} = (2n+1)\frac{1}{8}\,\mathrm{MHz}, \quad n \geq 0.$$

(b) Since the current modes need to vanish at $z = 0$, we can express them in terms of a sine function as

$$\sin(\beta z)\sin(\omega t)$$

where

$$\omega = 2\pi f = (2n+1)\frac{\pi}{4}\,\mathrm{Mrad/s},$$

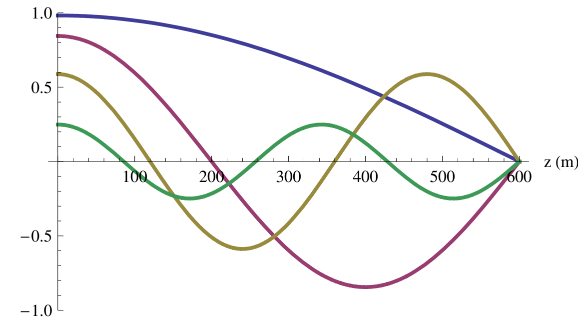
and

$$\beta = \frac{2\pi}{\lambda} = (2n+1)\frac{\pi}{1200}\,\mathrm{rad/m}.$$

In explicit terms, current modes are

$$I_n(z,t) = \sin((2n+1)\frac{\pi}{1200}z)\sin((2n+1)\frac{\pi}{4}t)$$

where z is in m and t in μs.



A time-snapshot of the voltage standing wave modes $n = 0, 1, 2, 3$ on a TL segment 600 m long, open ended at $z = 0$ and shorted at $z = 600$ m. Each mode n has $2n + 1$ quarter wavelengths fitted into the line length l and the high n modes oscillate with higher frequencies. See animation of these modes linked in the class calendar.

(c) Let's find the voltage modes $V_n(z,t)$ from the current modes above using one of the telegrapher's equations,

$$-\frac{\partial V}{\partial z} = \mathcal{L}\frac{\partial I}{\partial t}.$$

Substituting $I_n(z,t)$ into this equation and differentiating we find

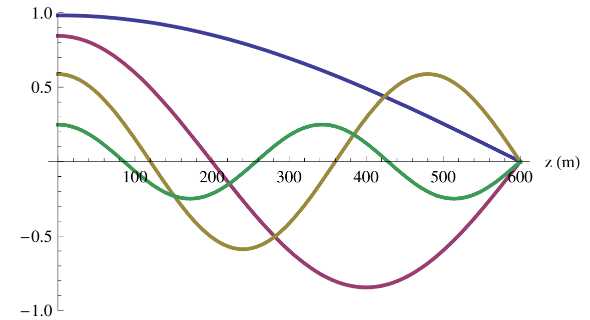
$$\frac{\partial V}{\partial z} = -(2n+1)\frac{\pi}{4}\mathcal{L}\sin((2n+1)\frac{\pi}{1200}z)\cos((2n+1)\frac{\pi}{4}t).$$

Next finding the anti-derivative of the above, we conclude

$$V_n(z,t) = 300\mathcal{L}\cos((2n+1)\frac{\pi}{1200}z)\cos((2n+1)\frac{\pi}{4}t).$$

A snapshot of the animation of resonant voltage modes is shown in the margin.

Note that the amplitudes of current and voltage modes cannot be assigned indepedent of one another — above we set the current amplitudes to unity and obtained $300\mathcal{L}$ for voltage amplitudes.



- **How can one get source free oscillations in a TL?**

  One answer is, the TL might have been connected to a source in the past before being disconnected from it.

  - Consider the circuit shown in the margin where a 3V battery is switched in and out for 1 μs on a line of length $l = 600$ m.

  For $v = c$, we can write the voltage and the current on the line at $t = 1\,\mu$s (by inspection) as

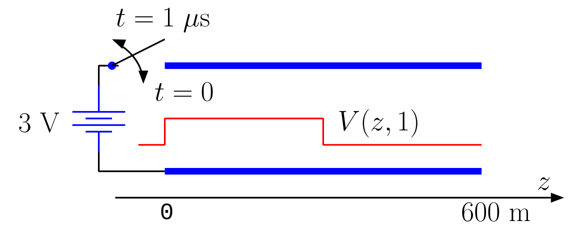
$$V(z,1) = 3\text{rect}(\frac{z-150}{300})\,\text{V} \quad \text{and} \quad I(z,1) = \frac{3\text{rect}(\frac{z-150}{300})}{Z_o}\,\text{A}.$$

  After $t = 1\,\mu$s both ends of the TL will be open, and, therefore, only periodic waveforms with a

  - fundamental period of $T = \frac{2\ell}{v} = \frac{1200}{300} = 4\,\mu s$ and
  - fundamental frequency of $\omega_o = \frac{2\pi}{T} = \frac{\pi}{2}$ Mrad/s

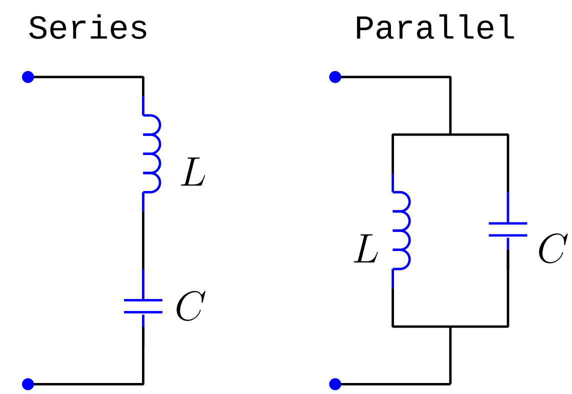
  will be allowed on the source free line.

  Therefore, $V(z,t)$ and $I(z,t)$ for $t > 1\,\mu$s can be expressed as a weighted superposition of the resonant modes of the line with resonant frequencies $n\omega_o$, subject to the initial conditions $V(z,1)$ and $I(z,1)$ given above.



Also, resistors R at temperature T connected to TL terminals can transfer thermal noise energy to the TL. If the resistors are disconnected at some point in time, the energy left on the TL will be shared between its resonant modes (up to a frequency limit $KT/\hbar$ imposed by quantum mechanics) at an average level of KT joules (per mode) where K is the *Boltzmann constant*. Lossy lines with finite conductivity also produce thermal noise. Thermal noise is easy to detect and routinely interferes with weak communication signals that we care about!

32 Input impedance and microwave resonators



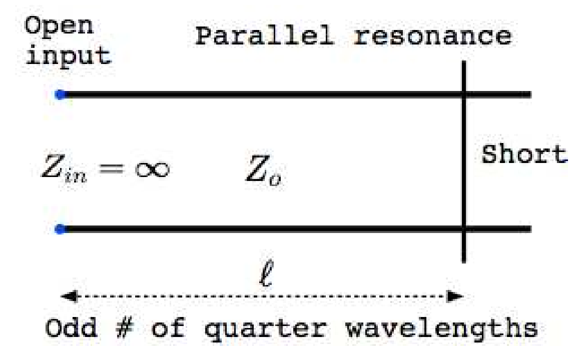
- The **input impedance** and **admittance** of the **series** and **parallel** LC resonators shown in the margin are, respectively,

$$Z_s = j(\omega L - \frac{1}{\omega C}) \quad \text{and} \quad Y_p = j(\omega C - \frac{1}{\omega L}),$$

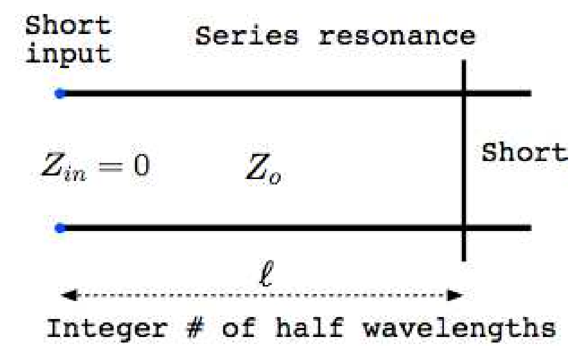
both of which vanish at the common resonance frequency of these networks, namely

$$\omega = \frac{1}{\sqrt{LC}} \equiv \omega_o.$$

  - Recall that LC resonators play an important role in the design of filter and tuning circuits.



In this lecture we will examine the input impedance of microwave resonators consisting of open or short circuited TL stubs.

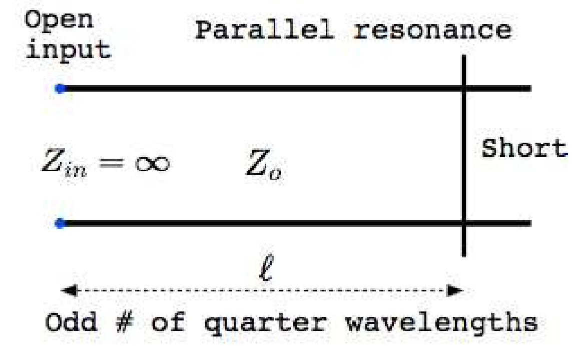


- In the last lecture we learned that when a shorted stub is open circuited at its input port, it shows resonance if the stub length ℓ is an odd multiple of $\frac{\lambda}{4}$.

  - The corresponding resonant frequencies are

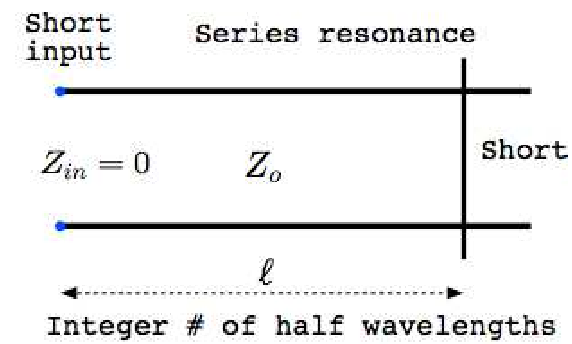
$$f = \frac{v}{2\ell}(\frac{1}{2} + n) \text{ for } n = 0, 1, 2, 3, \cdots$$

and the input port of the stub coincides with a voltage max and a current null, i.e., $I(0,t) = 0$ — thus the input impedance Z_{in} of the stub is *infinite* at these resonances, just like the impedance of the *parallel* LC-circuit depicted above.



- Thus this set of resonant frequencies are referred to as **parallel resonances** of the shorted stub.

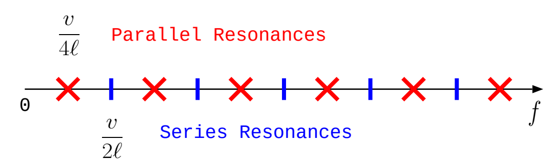
- We also learned that when the stub length ℓ is an an integer multiple of $\frac{\lambda}{2}$, its voltage at the input terminal is necessarily zero, implying that the input impedance Z_{in} must also be *zero*.



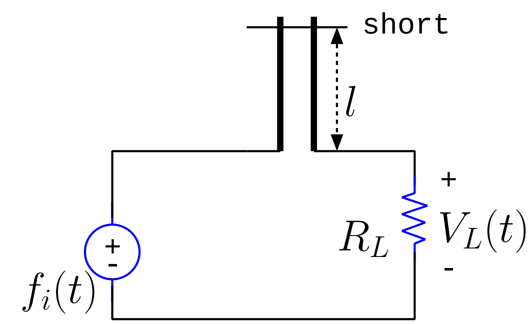
  - The corresponding resonant frequencies are

$$f = \frac{v}{2\ell} n \;\; \text{for} \;\; n = 1, 2, 3, \cdots$$

and are termed **series resonances** of the shorted stub, in analogy with the zero impedance of the *series* LC-circuit depicted above.



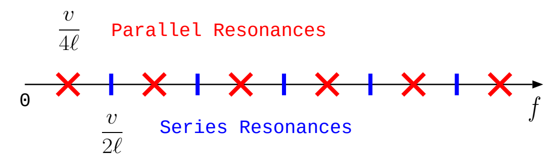
The diagram in the margin marks the locations of *parallel* and *series* resonance frequencies of the shorted stub associated with *infinite* and *zero* input impedance Z_{in}.

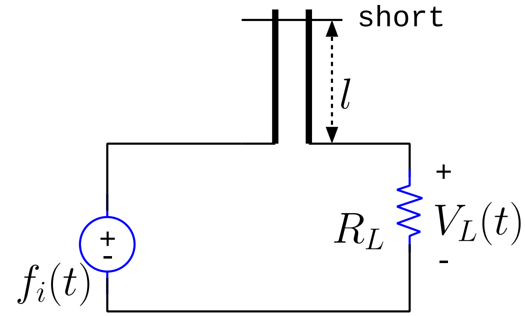


Thus, a shorted stub, included in a circuit such as the one shown in the margin, will exhibit extreme behavior at these special frequencies — namely it will appear as a

- short at its series resonances, causing the entire input signal $f_i(t)$ to appear as $V_L(t)$ across the load R_L, and

- open at its parallel resonances, causing $V_L(t)$ across the load R_L to be notched out.





We next focus our attention on how Z_{in} of the stub appears at other frequencies not coinciding with any of the resonances discussed above.

- In the following we will assume that the TL stub, as well as the circuit it is connected to, are all in sinusoidal steady state at a frequency determined by the frequency of the sinusoidal source $f_i(t)$.

  – In that case d'Alembert solutions will also be co-sinusoidal at the source frequency ω and we can express $V(z,t)$ and $I(z,t)$ on the line as

$$V(z,t) = \text{Re}\{V^+ e^{j\omega(t-\frac{z}{v})}\} + \text{Re}\{V^- e^{j\omega(t+\frac{z}{v})}\} \quad \Leftrightarrow \quad \tilde{V}(z) = V^+ e^{-j\beta z} + V^- e^{j\beta z}$$

and

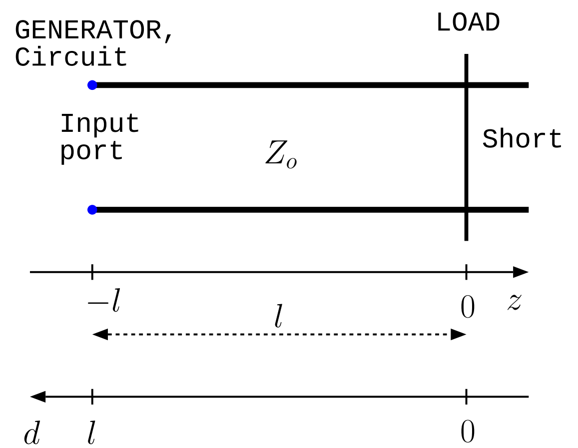
$$I(z,t) = \frac{\text{Re}\{V^+ e^{j\omega(t-\frac{z}{v})}\} - \text{Re}\{V^- e^{j\omega(t+\frac{z}{v})}\}}{Z_o} \quad \Leftrightarrow \quad \tilde{I}(z) = \frac{V^+ e^{-j\beta z} - V^- e^{j\beta z}}{Z_o},$$

where

  – $\beta = \frac{\omega}{v} = \omega\sqrt{\mathcal{LC}}$ is the wavenumber at frequency ω, and

- V^+ and V^- are phasors of forward and backward propagating voltage waves on the line evaluated at $z = 0$.

We have expressed the phasor counterparts of co-sinusoidal waves $V(z,t)$ and $I(z,t)$ above on the right, for it will be necessary to use phasors in defining an input impedance — the impedance concept belongs to the frequency domain!



- Before applying the boundary condition at the shorted end of the TL stub, it will be convenient to shift the origin of our coordinate system to coincide with the shorted termination rather than the input port of the TL.

- It will also be convenient to refer to "$-z$" as "d", with the variable d growing to the left from the short termination toward the input terminal of the line.

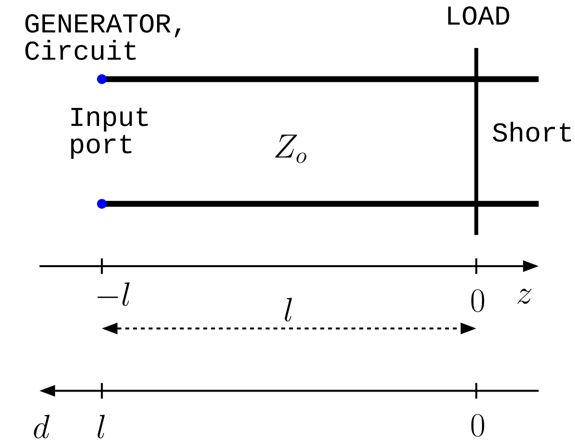
  In that case, the input impedance of the shorted stub can be denoted as

$$Z(l) = \frac{\tilde{V}(d=l)}{\tilde{I}(d=l)},$$

  where

$$\tilde{V}(d) \equiv V^+e^{j\beta d} + V^-e^{-j\beta d} \quad \text{and} \quad \tilde{I}(d) \equiv \frac{V^+e^{j\beta d} - V^-e^{-j\beta d}}{Z_o}.$$

- An immediate benefit of our new notation comes when we apply the voltage boundary condition at the short termination.



– We apply it as

$$V(0,t) = 0 \quad \Leftrightarrow \quad \tilde{V}(0) = V^+ + V^- = 0$$

from which it follows that

$$V^- = -V^+$$

and thus

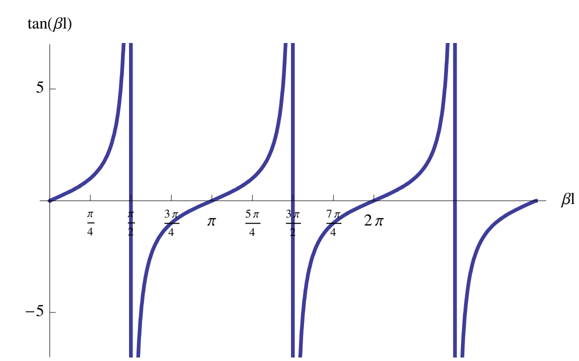
$$\tilde{V}(d) \equiv V^+(e^{j\beta d} - e^{-j\beta d}) = j2V^+ \sin(\beta d)$$

and

$$\tilde{I}(d) \equiv \frac{V^+(e^{j\beta d} + e^{-j\beta d})}{Z_o} = Y_o 2V^+ \cos(\beta d),$$

where

$$Y_o \equiv \frac{1}{Z_o} \quad \text{Characteristic admittance.}$$



– Finally the **input impedance** of the shorted stub is

$$Z(l) = \frac{\tilde{V}(l)}{\tilde{I}(l)} = jZ_o \tan(\beta l).$$

Note that: input impedance $Z(l) = 0 + jX(l)$ is (see margin for $X(l)$)

1. purely reactive for all l,

2. has a positive imaginary part and therefore it is **inductive** for

$$\beta l = \frac{2\pi}{\lambda} l < \frac{\pi}{2} \text{ rad} = 90^o \quad \Rightarrow \quad 0 < l < \frac{\lambda}{4} = \text{Quarter wavelength.}$$

3. has a negative imaginary part and therefore it is **capacitive** for

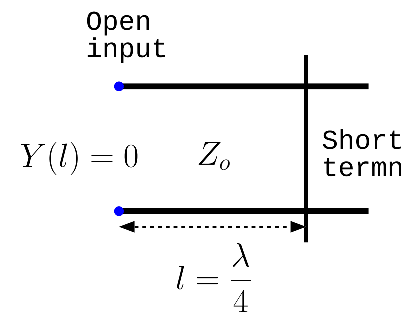
$$\frac{\pi}{2} < \beta l = \frac{2\pi}{\lambda} l < \pi \text{ rad} = 180^o \quad \Rightarrow \quad \frac{\lambda}{4} < l < \frac{\lambda}{2} = \text{Half wavelength.}$$

4. is periodic with a period of $\frac{\lambda}{2}$ over l, which means that

   all possible **reactive impedances** of the form jX are realized for $0 < l < \frac{\lambda}{2}$.

   - a shorted TL stub of length $0 < l < \frac{\lambda}{2}$ spans all possible impedances that can be provided by all possible inductors and capacitors!

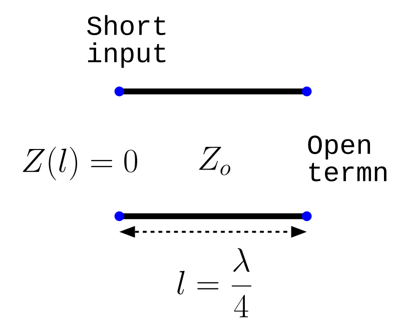
   for a length $l \ll \frac{\lambda}{4}$ shorted stub is a pure inductor with impedance



$$Z(l) = jZ_o \tan(\beta l) \approx jZ_o \beta l = j\sqrt{\frac{\mathcal{L}}{\mathcal{C}}}\omega\sqrt{\mathcal{L}\mathcal{C}}l = j\omega\mathcal{L}l.$$

   - here we used $\tan(\beta l) \approx \beta l$, which is valid when $\beta l \ll 1$ in radians.

5. at $l = \frac{\lambda}{4}$ the **input admittance** of the shorted stub,



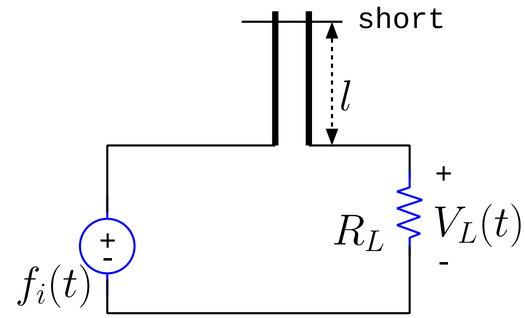
$$Y(l) = \frac{1}{Z(l)} = \frac{1}{jZ_o \tan(\beta l)} = -jY_o \cot(\beta l),$$

   vanishes, meaning that

   - a shorted stub of length $l = \frac{\lambda}{4}$ appears at its input terminals like an *open* (see margin).

6. at $l = \frac{\lambda}{2}$ the **input impedance** of the shorted stub returns back to zero, which in turn indicates, in view of (5), that

   - an open ended stub of length $l = \frac{\lambda}{4}$ must appear at its input terminals like a *short* (see margin).



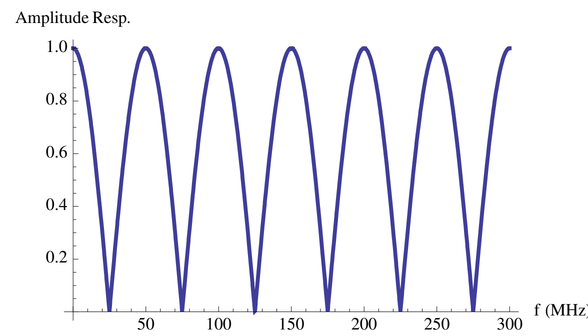
Next set of examples illustrate the uses of shorted/opened TL stubs as circuit elements.

Example 1: A shorted TL stub of length $l = 3$ m is connected in *series* with a resistor $R_L = 50\,\Omega$ as shown in the diagram in the margin. Plot the magnitude of the frequency response $H(\omega) = \frac{\tilde{V}_L}{\tilde{F}}$ as a function of frequency $f = \frac{\omega}{2\pi}$ if $Z_o = 50\,\Omega$ and $v = c$ on the stub. Interpret the amplitude response curve $|H(\omega)|$ in terms of resonance frequencies of the shorted line.

Solution: Using $\beta = \frac{\omega}{c}$ and voltage division, we find that frequency response

$$H(\omega) = \frac{\tilde{V}_L}{\tilde{F}_i} = \frac{R_L}{R_L + jZ_o\tan(\beta l)} = \frac{1}{1 + j\tan(\frac{\omega}{c}l)}.$$

The plot of $|H(\omega)|$ with the given parameters is shown in the margin. The peaks of the amplitude response occur at the series resonance frequencies of the shorted stub when its input impedance is zero (an effective short). The nulls of the amplitude response correspond to parallel resonances of the stub when it appears like an open at its input terminals.



Example 2: Consider a shorted TL connected at $d = l$ to an inductor L. Determine the resonances of the combined network.

Solution: The input impedance of the shorted line is

$$Z(l) = jZ_o \tan(\beta l) = jZ_o \tan(\omega\sqrt{\mathcal{L}\mathcal{C}}l)$$

whereas inductor L has an impedance $Z_L = j\omega L$. If the inductor and shorted stub are connected in series (see margin), then the *series resonances* of the network will be observed when the network input impedance

$$Z_L + Z(l) = j\omega L + jZ_o \tan(\omega\sqrt{\mathcal{L}\mathcal{C}}l)$$

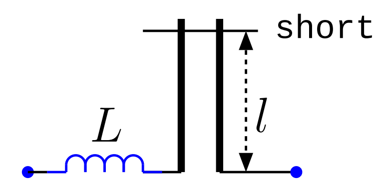
equals zero. The *parallel resonances* of the network will be observed when the impedance is infinite. While the series resonance frequencies of the stub will be shifted because of the inductor, parallel resonances will not shift (infinities due to tan function cannot be shifted by the finite additive term due to the inductor). The shifted series resonance frequencies ω_n can be found graphically by plotting $Z_L + Z(l)$ and looking for the zero crossings.

If the inductor and shorted stub are connected in parallel, then the *parallel resonances* of the network will be observed when the network input admittance

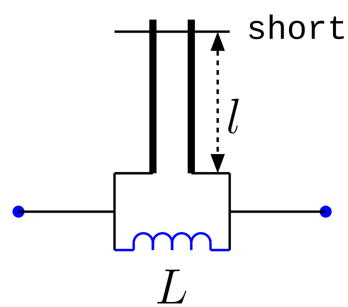
$$Y_L + Y(l) = \frac{1}{j\omega L} + \frac{1}{jZ_o \tan(\omega\sqrt{\mathcal{L}\mathcal{C}}l)}$$

equals zero (same as infinite input impedance). The *series resonances*, on the other hand, will be observed when the admittance is infinite (same as zero input impedance). Series resonances of the stub will not be shifted with, unlike its parallel resonances. The shifted parallel resonance frequencies ω_n will equal the series resonance frequencies of the series connected network described above.

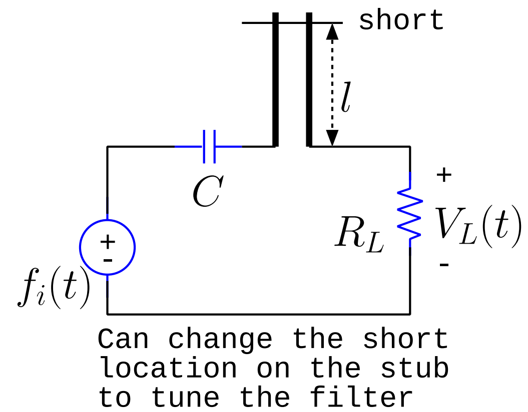
Series network:



Parallel network:

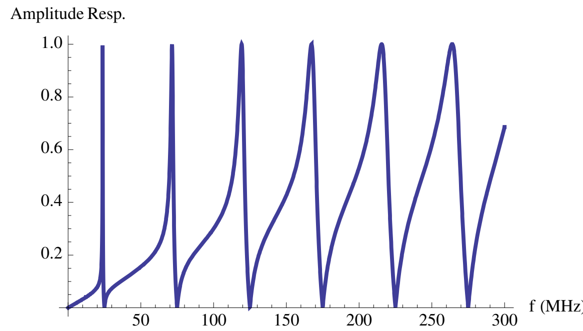


Example 3: A shorted TL stub of length $l = 3$ m is connected in *series* with a a capacitor $C = 10$ pf and a resistor $R_L = 50\,\Omega$ as shown in the diagram in the margin. Plot the magnitude of the frequency response $H(\omega) = \frac{\tilde{V}_L}{\tilde{F}}$ as a function of frequency $f = \frac{\omega}{2\pi}$ if $Z_o = 50\,\Omega$ and $v = c$ on the stub. Interpret the amplitude response curve $|H(\omega)|$ in terms of resonance frequencies of the shorted line.



Solution: Using $\beta = \frac{\omega}{c}$ and voltage division, we find that frequency response

$$H(\omega) = \frac{\tilde{V}_L}{\tilde{F}_i} = \frac{R_L}{R_L + \frac{1}{j\omega C} + jZ_o \tan(\beta l)} = \frac{1}{1 + \frac{1}{j\omega R_L C} + j\tan(\frac{\omega}{c} l)}.$$



The plot of $|H(\omega)|$ with the given parameters is shown in the margin. The peaks of the amplitude response occur at the *shifted* series resonance frequencies of the shorted T.L. stub. The nulls of the amplitude response correspond to parallel resonances of the stub when it appears open.

Example 4: (a) If in the TL circuit shown in the margin $I_R = 2\angle 0^o$ A, what is the line length l in terms of wavelength λ of the given source frequency on the line?

(b) Repeat for $I_R = 0$.

Note: starting in this example we are dropping the tildes on the phasors.

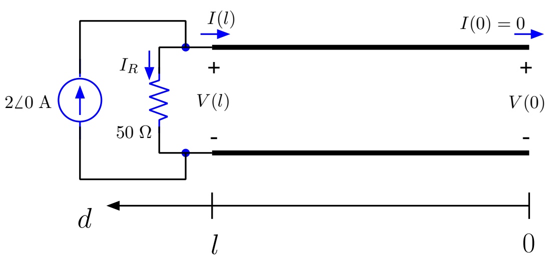
Solution: (a) If $I_R = 2\angle 0^o$ A, then KCL application at the source terminal implies that $I(l) = 0$.

In that case the TL has an *open* at $d = l$. Since $d = 0$ is also an open, we need to have l to be an integer multiple of $\frac{\lambda}{2}$.

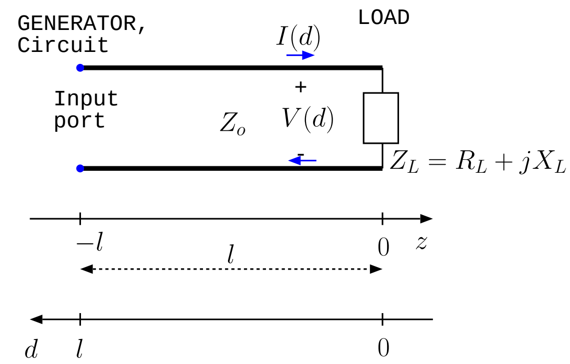
In other words, the condition of $I_R = 2\angle 0^o$ will only be realized in the above circuit if the source frequency is such that the TL length l happens to be some integer multiple of $\frac{\lambda}{2}$ at the given frequency.

(b) If $I_R = 0$, then $V(l) = (50\,\Omega)I_R = 0$, implying that the T.L. has a *short* at $d = l$.

Since $d = 0$ is an open, we need to have l some odd multiple of $\frac{\lambda}{4}$.



33 TL circuits with half- and quarter-wave transformers



- Last lecture we established that phasor solutions of telegrapher's equations for TL's in sinusoidal steady-state can be expressed as

$$V(d) = V^+ e^{j\beta d} + V^- e^{-j\beta d} \quad \text{and} \quad I(d) = \frac{V^+ e^{j\beta d} - V^- e^{-j\beta d}}{Z_o}$$

  in a new coordinate system shown in the margin.

  By convention the load is located on the right, at $z = 0 = d$, and the TL input connected to a generator or some source circuit is shown on the left at $d = l$.

  We have replaced the short termination of the previous lecture with an arbitrary load impedance

$$Z_L = R_L + jX_L.$$

  In this lecture we will discuss sinusoidal steady-state TL circuit problems having arbitrary reactive loads but with line lengths l constrained to be integer multiples of $\frac{\lambda}{4}$ (at the operation frequency).

  The constraint will be lifted next lecture when we will develop the general analysis tools for sinusoidal steady-state TL circuits.

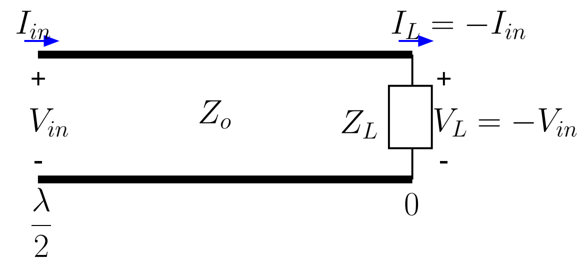
- In the TL circuit shown in the margin an arbitrary load Z_L is connected to a TL of length $l = \frac{\lambda}{2}$ at the source frequency.

  Given that
  $$e^{\pm j\beta\frac{\lambda}{2}} = e^{\pm j\frac{2\pi}{\lambda}\frac{\lambda}{2}} = e^{\pm j\pi} = -1,$$
  the general phasor relations
  $$V(d) = V^+e^{j\beta d} + V^-e^{-j\beta d} \text{ and } I(d) = \frac{V^+e^{j\beta d} - V^-e^{-j\beta d}}{Z_o}$$
  imply
  $$V_{in} \equiv V(\frac{\lambda}{2}) = -V^+ - V^- = -V(0) = -V_L,$$
  $$I_{in} \equiv I(\frac{\lambda}{2}) = \frac{-V^+ + V^-}{Z_o} = -I(0) = -I_L.$$
  We conclude that a $\frac{\lambda}{2}$**-transformer**

  - inverts the algebraic sign of its voltage and current inputs at the load end (and vice versa), and
  - has an input impedance identical with the load impedance since
  $$Z_{in} \equiv \frac{V_{in}}{I_{in}} = \frac{-V_L}{-I_L} = Z_L.$$

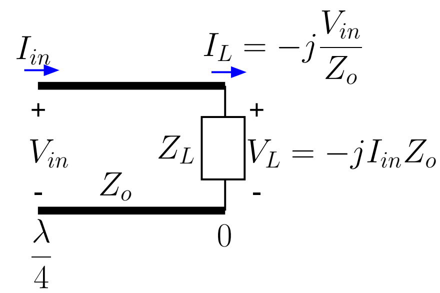
  These very simple results are easy to remember and use.

Half-wave transformer:



- In the TL circuit shown in the margin an arbitrary load Z_L is connected to a TL of length $l = \frac{\lambda}{4}$ at the source frequency.

Quarter-wave transformer:



$Z_{in}Z_L = Z_o^2$

Given that

$$e^{\pm j\beta\frac{\lambda}{4}} = e^{\pm j\frac{2\pi}{\lambda}\frac{\lambda}{4}} = e^{\pm j\frac{\pi}{2}} = \pm j,$$

general phasor relations

$$V(d) = V^+e^{j\beta d} + V^-e^{-j\beta d} \text{ and } I(d) = \frac{V^+e^{j\beta d} - V^-e^{-j\beta d}}{Z_o}$$

imply

$$V_{in} \equiv V(\frac{\lambda}{4}) = jV^+ - jV^- = jI(0)Z_o = jI_LZ_o,$$

$$I_{in} \equiv I(\frac{\lambda}{4}) = \frac{jV^+ + jV^-}{Z_o} = j\frac{V(0)}{Z_o} = j\frac{V_L}{Z_o}.$$

We conclude that a $\frac{\lambda}{4}$**-transformer**

– has an input impedance

$$Z_{in} \equiv \frac{V_{in}}{I_{in}} = \frac{jI_LZ_o}{jV_L/Z_o} = \frac{Z_o^2}{V_L/I_L} = \frac{Z_o^2}{Z_L},$$

– and provides a load current

$$I_L = -j\frac{V_{in}}{Z_o},$$

proportional to input voltage V_{in} but independent of load impedance Z_L.

Quarter-wave current-forcing equation:

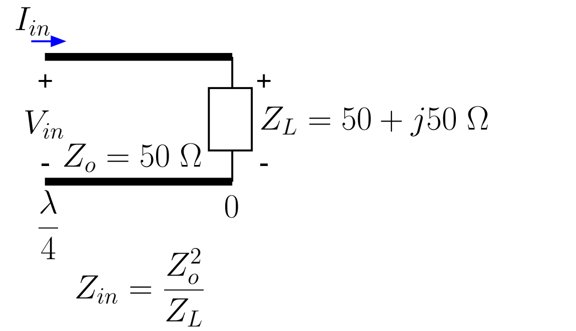
$$I_L = -j\frac{V_{in}}{Z_o}.$$

Load voltage

$$V_L = Z_LI_L$$

once I_L is available from above equation.

Example 1: Given $Z_L = 50 + j50\,\Omega$, what is Z_{in} for a $\frac{\lambda}{4}$ transformer with $Z_o = 50\,\Omega$?

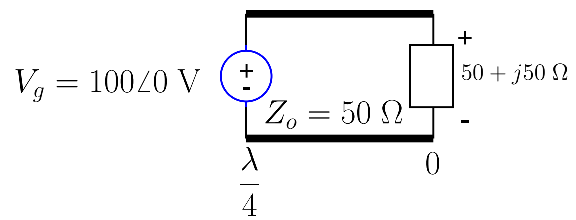


Solution: It is

$$Z_{in} = \frac{Z_o^2}{Z_L} = \frac{50^2}{50+j50} = \frac{50}{1+j1} = \frac{50}{1+j1}\frac{1-j1}{1-j1} = 25 - j25\,\Omega.$$

Notice that an *inductive* Z_L has been turned into a *capacitive* Z_{in} by $\frac{\lambda}{4}$ transformer.

Example 2: The load and the transformer of Example 1 are connected to a source with voltage phasor $V_g = 100\angle 0^o$ V at the input port. What is the load current I_L and what is the average power absorbed by the load?



Solution: Since $V_{in} = V_g = 100\angle 0^o$ V, the current-forcing formula for the quarter-wave transformer implies

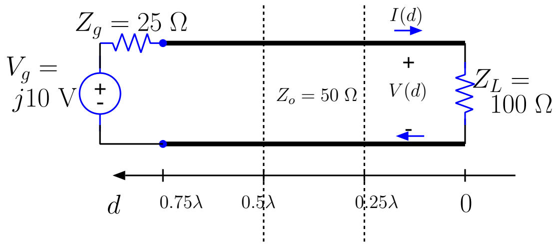
$$I_L = -j\frac{V_{in}}{Z_o} = -j\frac{100}{50} = -j2\,\text{A}.$$

To find the average power absorbed, we first note that load voltage

$$V_L = Z_L I_L = (50+j50)(-j2) = 100 - j100\,\text{V}.$$

Thus,

$$P_L = \frac{1}{2}\text{Re}\{V_L I_L^*\} = \frac{1}{2}\text{Re}\{(100-j100)(j2)\} = 100\,\text{W}.$$



Example 3: Load $Z_L = 100\,\Omega$ is connected to a T.L. with length $l = 0.75\lambda$. At the generator end, $d = 0.75\lambda$, a source with open circuit voltage $V_g = j10$ V and Thevenin impedance $Z_g = 25\,\Omega$ is connected. Determine V_L and I_L if $Z_o = 50\,\Omega$.

Solution: First we determine input impedance Z_{in} by noting that $Z_L = 100\,\Omega$ transforms to itself, namely $100\,\Omega$ at $d = 0.5\lambda$, but then it transforms from $d = 0.5\lambda$ to 0.75λ as

$$Z_{in} = \frac{Z_o^2}{Z(0.5\lambda)} = \frac{50^2}{100} = 25\,\Omega.$$

Hence, using voltage division, we find,

$$V_{in} = V_g \frac{Z_{in}}{Z_g + Z_{in}} = j10 \frac{25}{25+25} = j5\,\text{V}.$$

Next, using half-wave transformer rule, we notice that

$$V(0.25\lambda) = -V_{in} = -j5\,\text{V},$$

and finally applying the quarter-wave current forcing equation with $V(0.25\lambda)$ we get

$$I_L = -j\frac{V(0.25\lambda)}{Z_o} = -j\frac{-j5}{50} = -0.1\,\text{A}.$$

Clearly, then, the load voltage is

$$V_L = Z_L I_L = (100\,\Omega)(-0.1\,\text{A}) = -10\,\text{V}.$$

Example 4: In the circuit shown in the margin, $Z_{L1} = 50\,\Omega$, $Z_{L2} = 100\,\Omega$, and $Z_{o1} = Z_{o2} = 50\,\Omega$. Determine I_{L1} and I_{L2} if $V_{in} = 5\,\text{V}$. Both T.L. sections are quarter-wave transformers.

Solution: Using the current-forcing equation, we have

$$I_{L1} = I_{L2} = -j\frac{V_{in}}{Z_o} = -j\frac{5}{50} = -j0.1\,\text{A}.$$

Consequently,

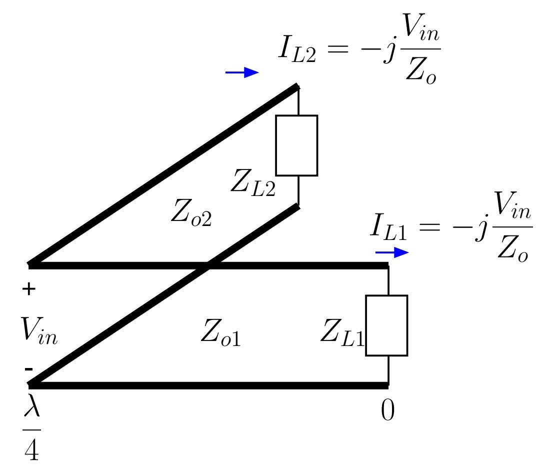
$$V_{L1} = I_{L1}Z_{L1} = -j0.1\,\text{A} \times 50\,\Omega = -j5\,\text{V}$$

and

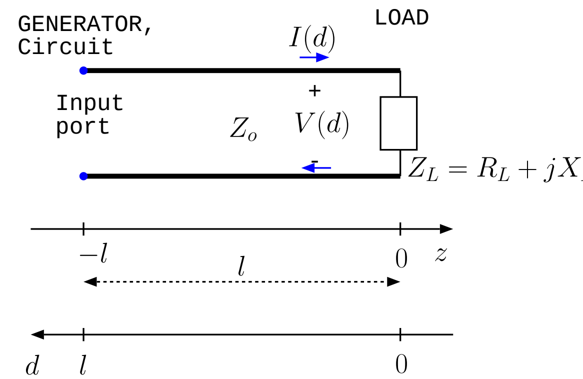
$$V_{L2} = I_{L2}Z_{L2} = -j0.1\,\text{A} \times 100\,\Omega = -j10\,\text{V}.$$

Thus, total avg power absorbed is

$$\begin{aligned} P &= \frac{1}{2}\text{Re}\{V_{L1}I_{L1}^*\} + \frac{1}{2}\text{Re}\{V_{L2}I_{L2}^*\} \\ &= = \frac{1}{2}\text{Re}\{-j5 \times j0.1\} + \frac{1}{2}\text{Re}\{-j10 \times j0.1\} = 0.75\,\text{W}. \end{aligned}$$



34 Line impedance, generalized reflection coefficient, Smith Chart



- Consider a TL of an arbitrary length l terminated by an arbitrary load

$$Z_L = R_L + jX_L.$$

as depicted in the margin.

Voltage and current phasors are known to vary on the line as

$$V(d) = V^+ e^{j\beta d} + V^- e^{-j\beta d} \quad \text{and} \quad I(d) = \frac{V^+ e^{j\beta d} - V^- e^{-j\beta d}}{Z_o}.$$

In this lecture we will develop the general analysis tools needed to determine the unknowns of these phasors, namely V^+ and V^-, in terms of source circuit specifications.

- Our analysis starts at the load end of the TL where $V(0)$ and $I(0)$ stand for the load voltage and current, obeying Ohm's law

$$V(0) = Z_L I(0).$$

Hence, using $V(0)$ and $I(0)$ from above, we have

$$V^+ + V^- = Z_L \frac{V^+ - V^-}{Z_o} \quad \Rightarrow \quad V^- = \frac{Z_L - Z_o}{Z_L + Z_o} V^+.$$

– Define a **load reflection coefficient**

$$\Gamma_L \equiv \frac{Z_L - Z_o}{Z_L + Z_o}$$

and re-write the voltage and current phasors as

$$V(d) = V^+ e^{j\beta d}[1+\Gamma_L e^{-j2\beta d}] \quad \text{and} \quad I(d) = \frac{V^+ e^{j\beta d}[1-\Gamma_L e^{-j2\beta d}]}{Z_o}.$$

"Load reflection coefficient" is a well justified name for Γ_L since the forward traveling wave with phasor $V^+ e^{j\beta d}$ gets reflected from the load.

– Define a **generalized reflection coefficient**

$$\Gamma(d) \equiv \Gamma_L e^{-j2\beta d}$$

and re-write the voltage and current phasors as

$$V(d) = V^+ e^{j\beta d}[1+\Gamma(d)] \quad \text{and} \quad I(d) = \frac{V^+ e^{j\beta d}[1-\Gamma(d)]}{Z_o}.$$

– **Line impedance** is then defined as

$$Z(d) = \frac{V(d)}{I(d)} = Z_o \frac{1+\Gamma(d)}{1-\Gamma(d)}$$

for all values of d on the line extending from the load point $d=0$ all the way to the input port at $d=l$.

The term "generalized reflection coefficient" is also well justified even if there is no reflection taking place at arbitrary d — the reason is, if the line were cut at location d and the stub with the load were replaced by a lumped load having a reflection coefficient equal to $\Gamma(d)$, then there would be no modification of the voltage and current variations on the line towards the generator.

With the dependence on d of $Z(d)$ as well as $\Gamma(d)$ tacitly implied, we can re-write this important relation and its inverse as

$$\frac{Z}{Z_o} = \frac{1+\Gamma}{1-\Gamma} \quad \Leftrightarrow \quad \Gamma = \frac{Z - Z_o}{Z + Z_o}.$$

Each location d on the line has an impedance Z and a reflection coefficient Γ linked by these equations.

Properties of $Z(d) = R(d) + jX(d)$ and $\Gamma(d) = \Gamma_L e^{-j2\beta d}$ linked by the relations

$$\frac{Z}{Z_o} = \frac{1+\Gamma}{1-\Gamma} \quad \Leftrightarrow \quad \Gamma = \frac{Z - Z_o}{Z + Z_o} :$$

1. **For real valued Z_o and $R(d) \geq 0$, $|\Gamma(d)| \leq 1$:**

   Verification:

$$|\Gamma| = \frac{|Z - Z_o|}{|Z + Z_o|} = \frac{|(R - Z_o) + jX|}{|(R + Z_o) + jX|} = \frac{\sqrt{(R - Z_o)^2 + X^2}}{\sqrt{(R + Z_o)^2 + X^2}}.$$

   Since with $R \geq 0$

$$\sqrt{(R - Z_o)^2 + X^2} \leq \sqrt{(R + Z_o)^2 + X^2} \quad \Rightarrow \quad |\Gamma| \leq 1.$$

2. Since

$$|\Gamma| = |\Gamma_L| \quad \text{and} \quad \angle\Gamma(d) = \angle\Gamma_L - 2\beta d$$

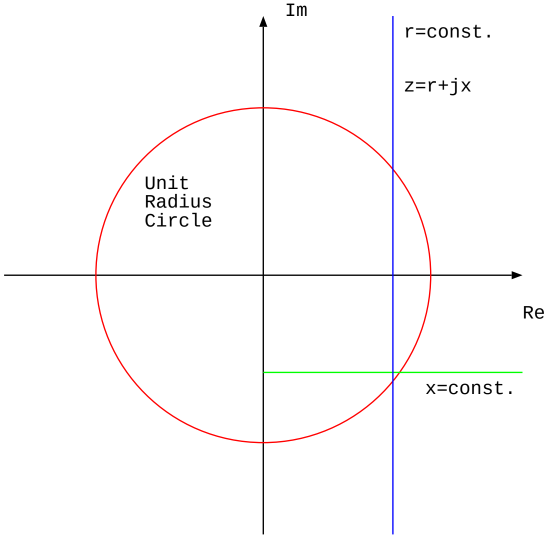
   property (1) implies that $\Gamma(d)$ is a complex number which is constrained to be **on or within the unit-circle on the complex plane**.

3. Relationships

$$\frac{Z}{Z_o} = \frac{1+\Gamma}{1-\Gamma} \quad \Leftrightarrow \quad \Gamma = \frac{Z - Z_o}{Z + Z_o}$$

   between Γ and Z are known as **bilinear transformations** — here the term *bilinear* refers to the numerator *as well as* the denominator of these transformations being *linear* in the variable being transformed (from right to left).

Bilinear (or Möbius) transformations are known to have the general property of mapping **straight lines** into **circles** on the complex number plane.



- Bilinear transformations between

$$\Gamma \equiv \Gamma_r + j\Gamma_i \equiv (\Gamma_r, \Gamma_i)$$

and

$$\frac{Z}{Z_o} \equiv z \equiv r + jx,$$

known as **normalized impedance,** lead to an ingenious graphical aid known as the **Smith Chart**.

  – On a Smith Chart (SC), straight lines on the right hand side of the complex number plane (see margin), represented by

$$r = \text{const.} \quad \text{and} \quad x = \text{const.},$$

  are mapped onto circular loci of

$$(\Gamma_r, \Gamma_i) = \Gamma = \frac{Z - Z_o}{Z + Z_o} = \frac{z-1}{z+1}$$

  occupying the region of the plane bordered by the unit circle.

Circles corresponding to $z = \text{const.} + jx$ and $z = r + j\text{const.}$constitute a **griding** of the unit circle and its interior. By means of this **grid**, the normalized impedance z corresponding to every possible Γ can be directly read off the SC.

- SC can be constructed by first noting that

$$\Gamma = \frac{z-1}{z+1} = \frac{r+jx-1}{r+jx+1} = \frac{[(r-1)+jx][(r+1)-jx]}{(r+1)^2+x^2} = \frac{(r^2+x^2-1)+j2x}{(r+1)^2+x^2} \equiv \Gamma_r + j\Gamma_i;$$

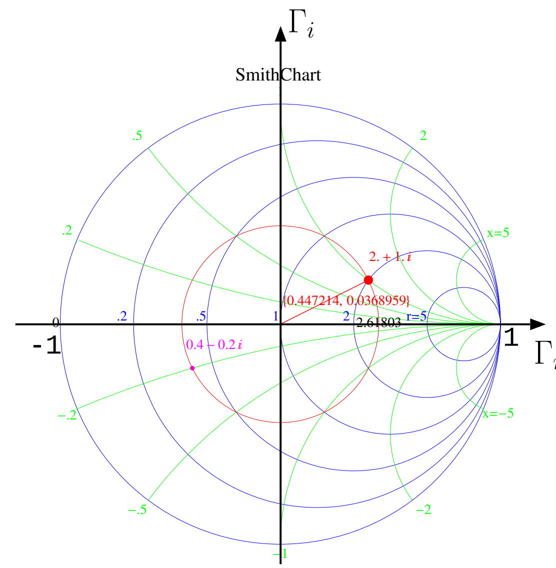
thus

$$\Gamma_r = \frac{(r^2+x^2-1)}{(r+1)^2+x^2} \text{ and } \Gamma_i = \frac{2x}{(r+1)^2+x^2},$$

and by direct substitution we can verify the following equations

$$(\Gamma_r - \frac{r}{r+1})^2 + \Gamma_i^2 = (\frac{1}{r+1})^2 \text{ and } (\Gamma_r - 1)^2 + (\Gamma_i - \frac{1}{x})^2 = (\frac{1}{x})^2$$

describing r and x dependent circles, respectively, on complex plane constituting the grid lines of the SC.



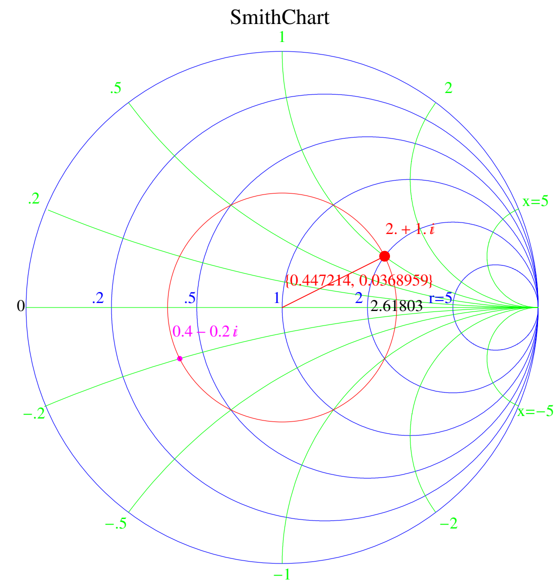
Smith Chart **is the unit circle and its interior on the complex number plane corresponding to the *generalized reflection coefficient* Γ. The gridding allows direct identification of the bilinear transform of Γ, namely the *normalized line impedance* z.**

- Typical SC usage:

  1. Locate and mark $z(0)$ — normalized load impedance — on the SC, which places you at a distance $|\Gamma(0)| = |\Gamma_L|$ from the origin of the complex plane (and the SC), at an angle of $\theta = \angle\Gamma(0)$.
  2. Draw a constant $|\Gamma| = |\Gamma_L|$ circle with a compass going through point $z(0)$ on the SC (the read circle in the margin). Rotate clockwise on the circle by an angle of

  $$2\beta d = \frac{4\pi}{\lambda} d \,\text{rad} = \frac{d}{\lambda/2} 360^\circ$$

  to land on $z(d)$ that can be read off using the SC gridding.



– Rotation by an angle of $2\beta d$ amounts to rotation by full circle for $d = \frac{\lambda}{2}$, rotation by half circle for $d = \frac{\lambda}{4}$, rotation by quarter circle for $d = \frac{\lambda}{8}$, etc.

3. Also,

$$y(d) \equiv \frac{1}{z(d)}$$

which is the **normalized line admittance** is located on the SC on the constant $|\Gamma| = |\Gamma_L|$ circle across the point corresponding to $z(d)$.

Verification: Since

$$z = \frac{1+\Gamma}{1-\Gamma} \quad \Rightarrow \quad y = \frac{1}{z} = \frac{1-\Gamma}{1+\Gamma} = \frac{1+(-\Gamma)}{1-(-\Gamma)};$$

hence whereas z is the transform of Γ, y is the transform of $-\Gamma$, having the same magnitude as Γ but an angle off by $\pm 180°$.

– Therefore, **"reflect" on the SC across the origin** to jump from $z(d)$ to $y(d)$ if you need the value of the normalized admittance.

Our first SC example is given next.

Example 1: A transmission line is terminated by an inductive load of

$$Z_L = 50 + j100\,\Omega.$$

Determine the input impedance $Z_{in} = Z(l)$ of the line at a distance

$$d = l = \frac{\lambda}{8}$$

if the characteristic impedance of the line is $Z_o = 50\,\Omega$. Also determine the normalized input admittance $y(l)$.

Solution: The normalized load impedance is

$$z(0) = \frac{Z_L}{Z_o} = \frac{50 + j100}{50} = 1 + j2.$$

Enter $z(0)$ on the SC and then rotate clockwise by $\frac{\lambda}{8} \Leftrightarrow$(quarter circle) to obtain the normalized input impedance

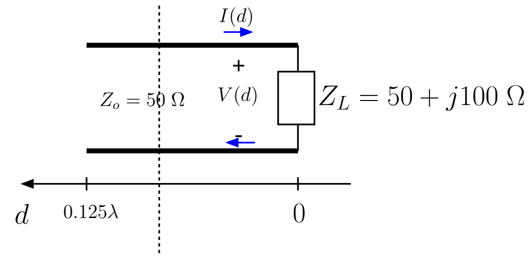
$$z(l) = 1 - j2,$$

and the normalized input admittance

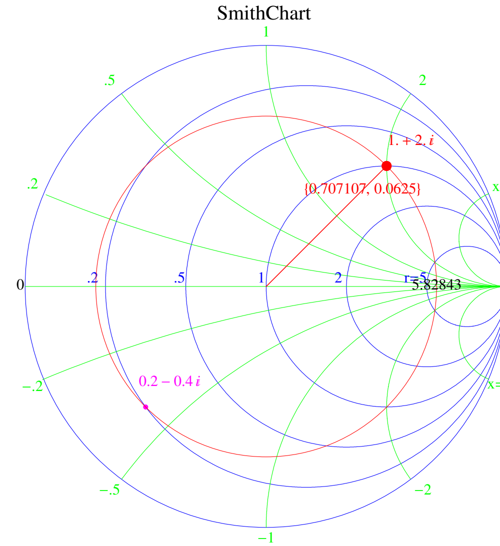
$$y(l) = 0.2 + j0.4$$

right across $z(l)$. The input impedance is

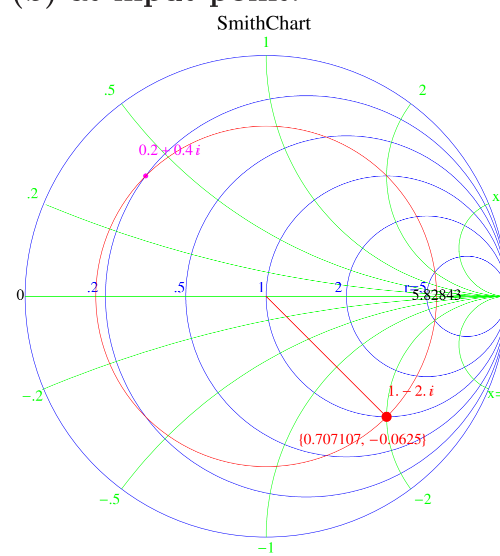
$$Z_{in} = Z_o z(l) = 50(1 - j2) = 50 - j100\,\Omega.$$



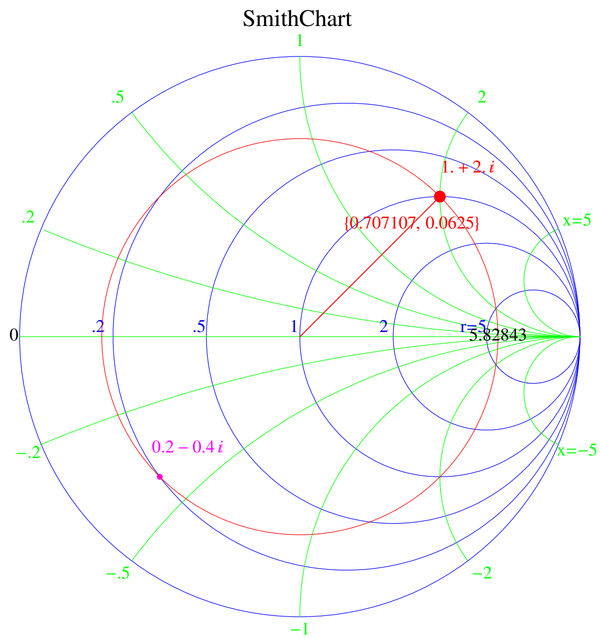
(a) At load point:



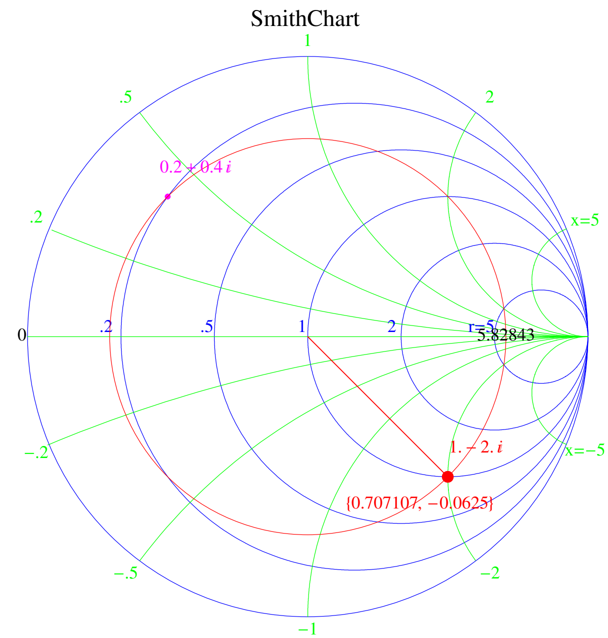
(b) at input point:



Blow up of the SC's used in Example 1:



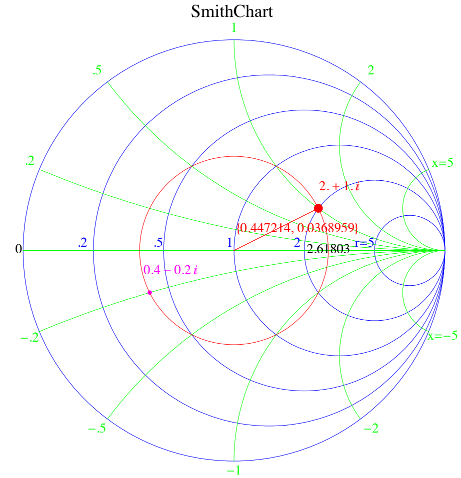
(a) At load point



(b) at input point

- A SmithChartTool linked from the class calendar (a javascript utility that requires a Safari or Firefox browser to work properly) marks and prints $z(d)$ in **red** and $y(d)$ in **magenta** across from $z(d)$ on the constant-$|\Gamma_L|$ circle (shown in red) as in the above examples. Also
  - printed in **black** is the real valued normalized impedance $z(d_{max})$ discussed in the upcoming lectures (also known as VSWR).
  - also printed in **red** is $|\Gamma_L|\angle\Gamma(d)$ where the second entry is expressed in terms of an equivalent $\frac{d}{\lambda}$ such that $\frac{d}{\lambda} = 0.5$ corresponds to an angle of 360°. This way of referring to $\angle\Gamma(d)$ will be convenient in many SC applications that we will see.

35 Smith Chart examples



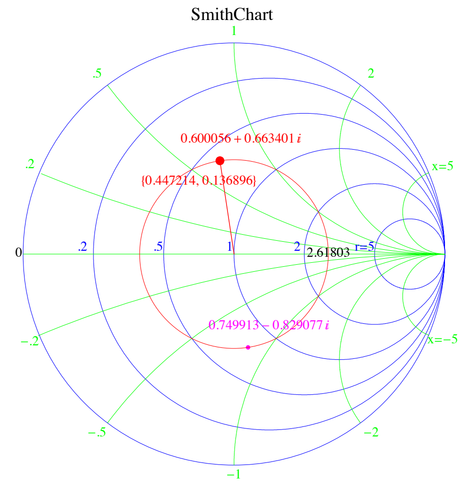
Example 1: A load $Z_L = 100 + j50\,\Omega$ is connected across a TL with $Z_o = 50\,\Omega$ and $l = 0.4\lambda$. At the generator end, $d = l$, the line is shunted by an impedance $Z_s = 100\,\Omega$. What are the input impedance Z_{in} and admittance Y_{in} of the line, including the shunt connected element.

Solution: Normalized load impedance

$$z(0) = \frac{Z_L}{Z_o} = \frac{100 + j50}{50} = 2 + j1$$

is entered in the SC shown in the margin on the top. Clockwise rotation (from load toward generator) at fixed $|\Gamma|$ (red circle) by

$$0.4\lambda \Leftrightarrow 0.8 \times 360^\circ = 288^\circ$$

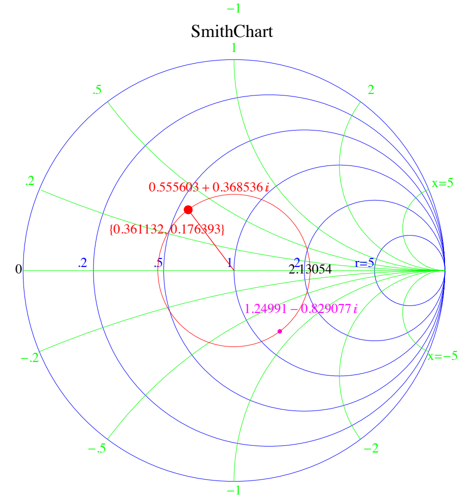


takes us to

$$z(l) \approx 0.6 + j0.66 \quad \text{and} \quad y(l) \approx 0.75 - j0.83$$

as shown on the SC in the middle. Hence, including the shunt element with normalized input impedance $z_{si} = 2$ and admittance $y_{si} = \frac{1}{2}$, we obtain

$$y_{in} = y(l) + y_{si} \approx 1.25 - j0.83$$

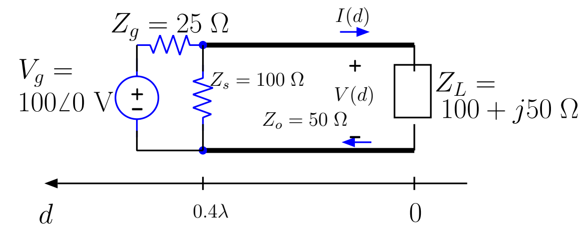


for the overall normalized input admittance of the shunted line as shown on the SC in the bottom — the corresponding normalized input impedance is

$$z_{in} = \frac{1}{y_i} \approx 0.56 + j0.37.$$

Hence, the unnormalized input impedance and admittance are

$$Z_{in} = Z_o z_{in} \approx 27.8 + j18.4\,\Omega \text{ and } Y_{in} = Y_o y_{in} \approx 0.025 - j0.017\,\text{S}.$$



Example 2: The TL network described in Example 1 is connected to a generator with open circuit voltage phasor $V_g = 100\angle 0$ V and internal impedance $Z_g = 25\,\Omega$. What is the average power (a) input of the shunted line, (b) delivered to the shunt element, delivered to the load.

Solution:

(a) Using the input impedance

$$Z_{in} \approx 27.8 + j18.4\,\Omega,$$

from Example 1, we can write

$$V_{in} = V_g \frac{Z_{in}}{Z_g + Z_{in}} \text{ and } I_{in} = \frac{V_g}{Z_g + Z_{in}}.$$

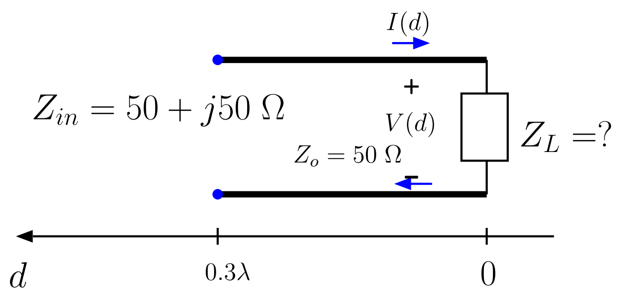
Therefore, the average power input of the shunted line is

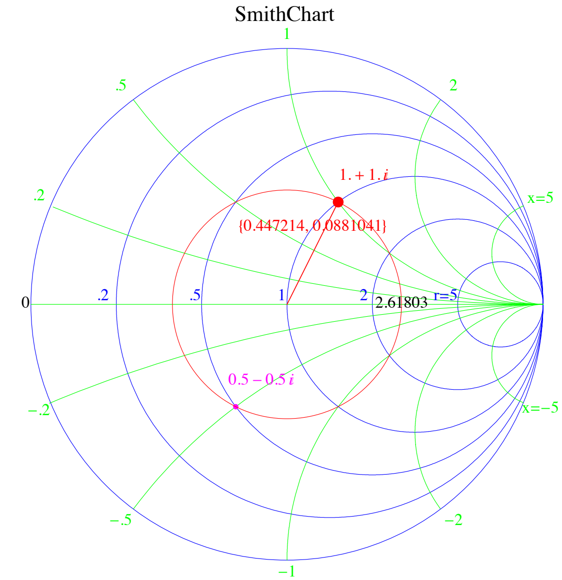
$$\begin{aligned} P &= \frac{1}{2}\mathrm{Re}\{V_{in}I_{in}^*\} = \frac{1}{2}\mathrm{Re}\{\frac{V_g Z_{in}}{Z_g + Z_{in}}(\frac{V_g}{Z_g + Z_{in}})^*\} \\ &= \frac{|V_g|^2}{2|Z_g + Z_{in}|^2}\mathrm{Re}\{Z_{in}\} = \frac{100^2}{2|25 + 27.8 + j18.4|^2}27.8 \approx 44.44\,\mathrm{W}. \end{aligned}$$

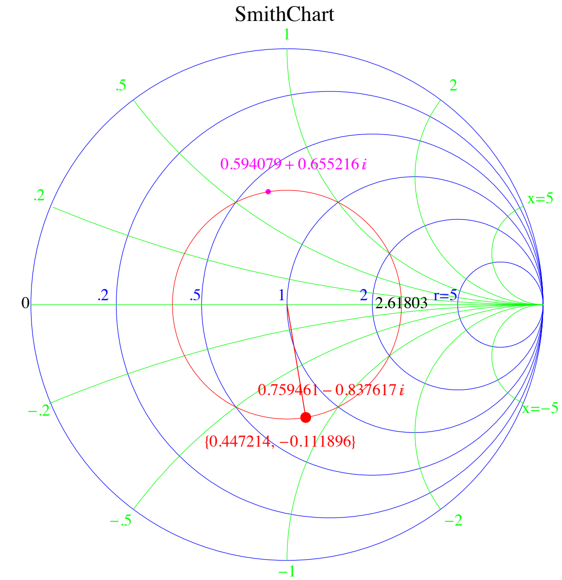
(b) The shunt element $Z_s = 100\,\Omega$ sees the same voltage V_{in} and conducts a current V_{in}/Z_s. Therefore it absorbs an average power of

$$\begin{aligned} P &= \frac{1}{2}\mathrm{Re}\{V_{in}(\frac{V_{in}}{Z_s})^*\} = \frac{|V_{in}|^2}{2Z_s} = \frac{|V_g Z_{in}|^2}{2Z_s|Z_g + Z_{in}|^2} \\ &\approx \frac{|100 \cdot (27.8 + j18.4)|^2}{2 \cdot 100 \cdot |25 + 27.8 + j18.4|^2} \approx 17.78\,\mathrm{W}. \end{aligned}$$

The remainder of 44.44 W will be absorbed in Z_L.







Example 3: A TL of length $l = 0.3\lambda$ has an input impedance $Z_{in} = 50 + j50\,\Omega$. Determine the load impedance $Z_L = Z(0)$ and $Y_L = Y(0)$ given that $Z_o = 50\,\Omega$ for the line.

Solution: First enter the normalized inpur impedance

$$z_{in} = \frac{Z_{in}}{Z_o} = \frac{50 + j50}{50} = 1 + j$$

in the SC as shown in the margin on the top. Counter-clockwise rotation (from generator toward load) at fixed $|\Gamma|$ (red circle) by

$$0.3\lambda \Leftrightarrow 0.6 \times 360^\circ = 216^\circ$$

takes us to

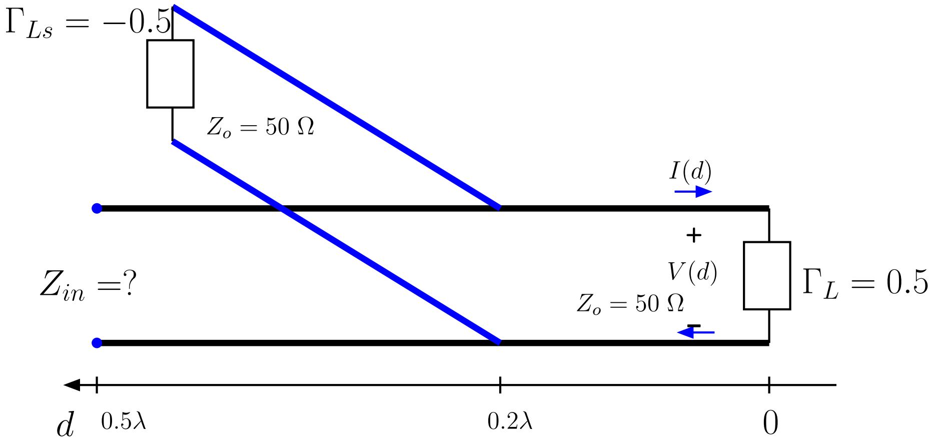
$$z(0) \approx 0.76 - j0.84 \quad \text{and} \quad y(0) \approx 0.59 + j0.66$$

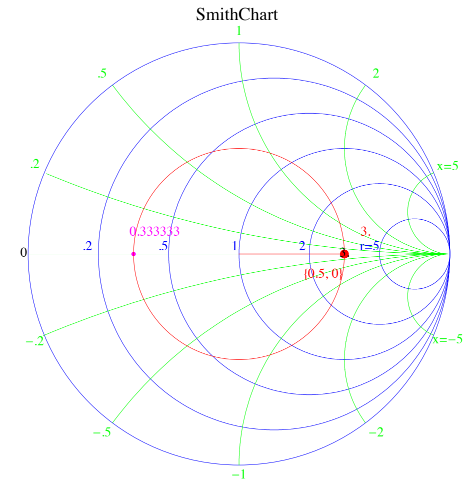
as shown on the next SC at the load point. Hence, we find

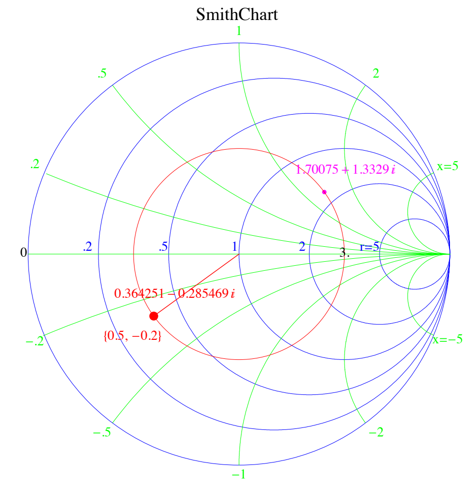
$$Z_L = Z_o z(0) \approx 50 \cdot (0.76 - j0.84) = 37.97 - j41.88\,\Omega$$

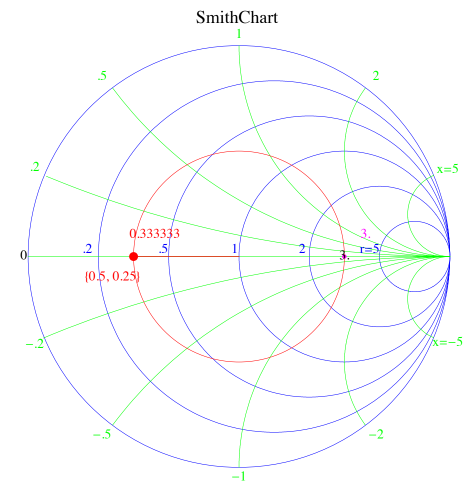
and

$$Y_L = Y_o y(0) \approx \frac{1}{50}(0.59 + j0.66) = 0.012 + j0.013\,\mathrm{S}.$$









Example 4: A TL of length $l = 0.5\lambda$ and $Z_o = 50\,\Omega$ has a load reflection coefficient $\Gamma_L = 0.5$ and and a shunt connected TL at $d = 0.2\lambda$. The shunt connected TL has $l = 0.3\lambda$, $Z_o = 50\,\Omega$, and a load reflection coefficient $\Gamma_L = -0.5$. Determine the input impedance of the line.

Solution: Recall that the SC covers the unit circle of the complex plane and therefore the complex number

$$\Gamma_L = 0.5 + j0 = 0.5$$

can be entered directly in the SC as shown on the top SC in the margin. Clockwise rotation (from load toward generator) at fixed $|\Gamma|$ (red circle) by

$$0.2\lambda \Leftrightarrow 0.4 \times 360^\circ = 144^\circ$$

takes us to

$$z(0.2\lambda) \approx 0.36 - j0.29 \quad \text{and} \quad y(0.2\lambda) \approx 1.7 + j1.33$$

as shown on the SC in the middle. Likewise, entering

$$\Gamma_{Ls} = -0.5 + j0 = -0.5$$

for the shunt connected stub in the third SC and rotating clockwise by

$$0.3\lambda \Leftrightarrow 0.6 \times 360^\circ = 216^\circ$$

we obtain

$$z_s(0.3\lambda) \approx 1.7 - j1.33 \quad \text{and} \quad y_s(0.3\lambda) \approx 0.36 + j0.29.$$

We proceed by combining the normalized admittances as

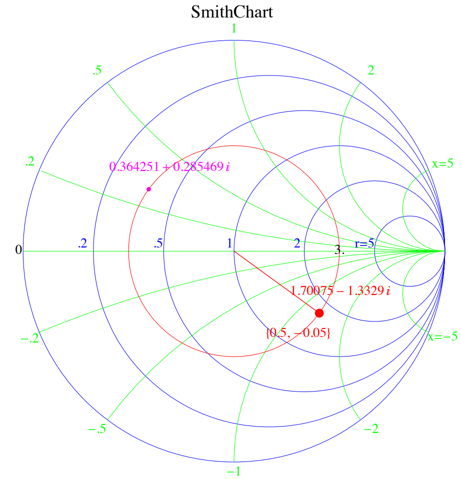
$$y_c = y(0.2\lambda) + y_s(0.3\lambda) \approx (1.7 + j1.33) + (0.36 + j0.29) = 2.065 + j1.61837,$$

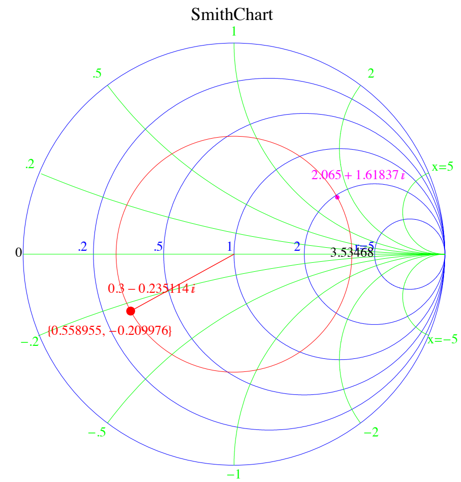
and entering it in the next SC. Finally rotating clockwise once again by

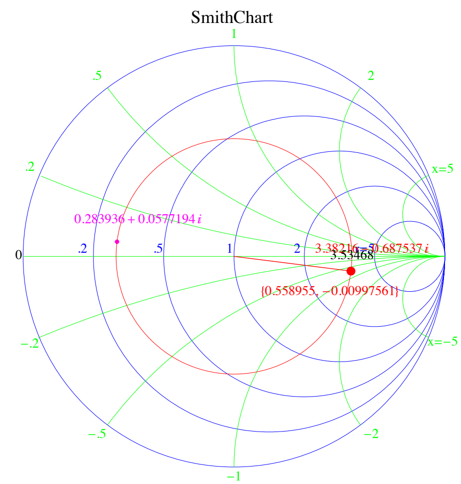
$$0.3\lambda \Leftrightarrow 0.6 \times 360^\circ = 216^\circ$$

we obtain, from the last SC

$$z_{in} \approx 3.38 - j0.69 \quad \Rightarrow \quad Z_{in} = z_{in} Z_o \approx 169 - j34.4\,\Omega.$$







Example 5: What is the load impedance Z_{Ls} terminating the shunt connected stub in Example 4?

Solution: Given that the corresponding reflection coefficient is

$$\Gamma_{Ls} = -0.5,$$

it follows from the bilinear transformation linking z_{Ls} and Γ_{Ls} that

$$z_{Ls} = \frac{1+\Gamma_{Ls}}{1-\Gamma_{Ls}} = \frac{1-0.5}{1+0.5} = \frac{1}{3}.$$

Hence, the impedance is

$$Z_{Ls} = Z_o z_{Ls} = \frac{50}{3}\,\Omega.$$

Example 6: What is the load impedance Z_L in Example 4?

Solution: This is similar to Example 5. Given that the load reflection coefficient is

$$\Gamma_L = 0.5,$$

it follows from the bilinear transformation linking z_L and Γ_L that

$$z_L = \frac{1+\Gamma_L}{1-\Gamma_L} = \frac{1+0.5}{1-0.5} = 3.$$

Hence, the impedance is

$$Z_L = Z_o z_L = 150\,\Omega.$$

36 Smith Chart and VSWR

- Consider the general phasor expressions

$$V(d) = V^+ e^{j\beta d}(1 + \Gamma_L e^{-j2\beta d}) \text{ and } I(d) = \frac{V^+ e^{j\beta d}(1 - \Gamma_L e^{-j2\beta d})}{Z_o}$$

describing the voltage and current variations on TL's in sinusoidal steady-state.

  - Unless $\Gamma_L = 0$, these phasors contain reflected components, which means that voltage and current variations on the line "contain" standing waves.

In that case the phasors go through cycles of magnitude variations as a function of d, and in the voltage magnitude in particular (see margin) varying as

$$|V(d)| = |V^+||1 + \Gamma_L e^{-j2\beta d}| = |V^+||1 + \Gamma(d)|$$

takes maximum and minimum values of

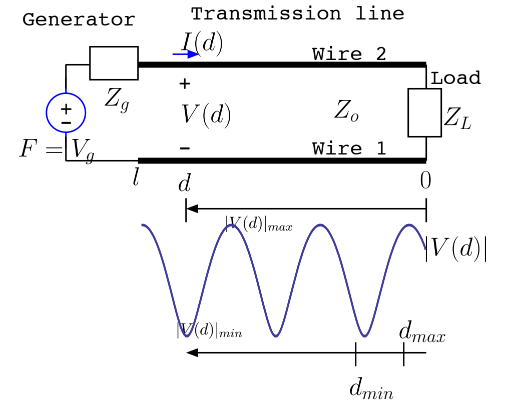
$$|V(d)|_{max} = |V^+|(1 + |\Gamma_L|) \text{ and } |V(d)|_{min} = |V^+|(1 - |\Gamma_L|)$$

at locations $d = d_{max}$ and d_{min} such that

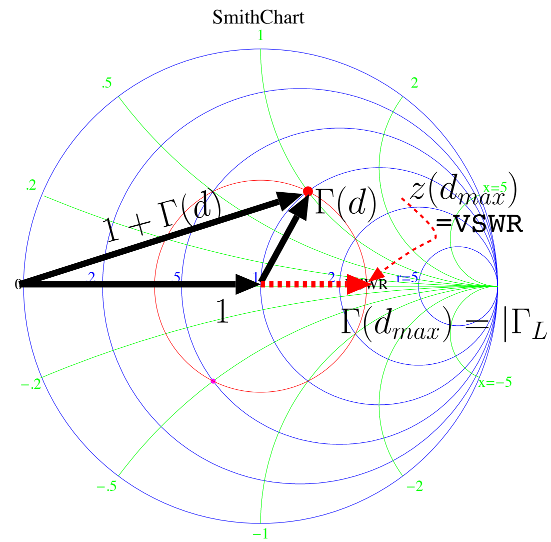
$$\Gamma(d_{max}) = \Gamma_L e^{-j2\beta d_{max}} = |\Gamma_L| \text{ and } \Gamma(d_{min}) = \Gamma_L e^{-j2\beta d_{min}} = -|\Gamma_L|,$$

and

$$d_{max} - d_{min} \text{ is an odd multiple of } \frac{\lambda}{4}.$$



Complex addition displayed graphically superposed on a Smith Chart



$|1 + \Gamma(d)|$ maximizes for $d = d_{max}$

$|1 + \Gamma(d)|$ minimizes for $d = d_{min}$ such that $\Gamma(d_{min}) = -\Gamma(d_{max})$

- These results can be most easily understood and verified graphically on a SC as shown in the margin.

- We define a parameter known as **voltage standing wave ratio**, or **VSWR** for short, by

$$\text{VSWR} \equiv \frac{|V(d_{max})|}{|V(d_{min})|} = \frac{1+|\Gamma_L|}{1-|\Gamma_L|} \quad \Leftrightarrow \quad |\Gamma_L| = \frac{\text{VSWR}-1}{\text{VSWR}+1}.$$

Notice that the VSWR and $|\Gamma_L|$ form a **bilinear transform pair** just like

$$z = \frac{1+\Gamma}{1-\Gamma} \quad \Leftrightarrow \quad \Gamma = \frac{z-1}{z+1}.$$

Since

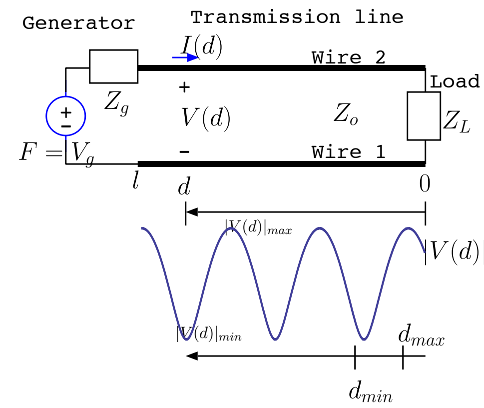
$$\Gamma(d_{max}) = |\Gamma_L| \quad \Rightarrow \quad \text{VSWR} = \frac{1+\Gamma(d_{max})}{1-\Gamma(d_{max})},$$

this analogy between the transform pairs also implies that

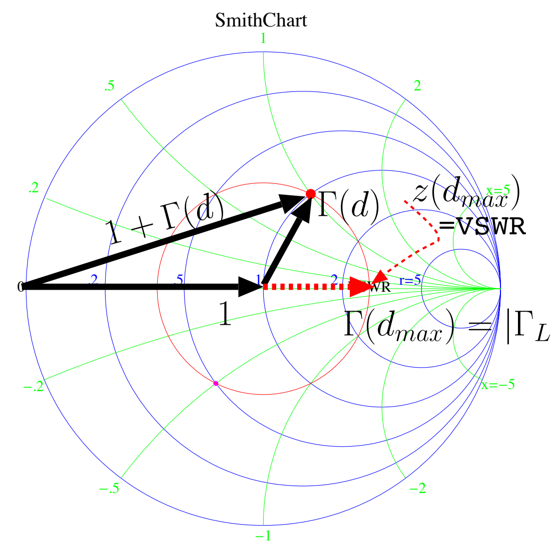
$$z(d_{max}) = \text{VSWR},$$

as explicitly marked on the the SC shown in the margin . Consequently,

- the VSWR of any TL can be directly read off from its SC plot as the normalized impedance value $z(d_{max})$ on constant-$|\Gamma_L|$ circle crossing the positive real axis of the complex plane.



Complex addition displayed graphically superposed on a Smith Chart

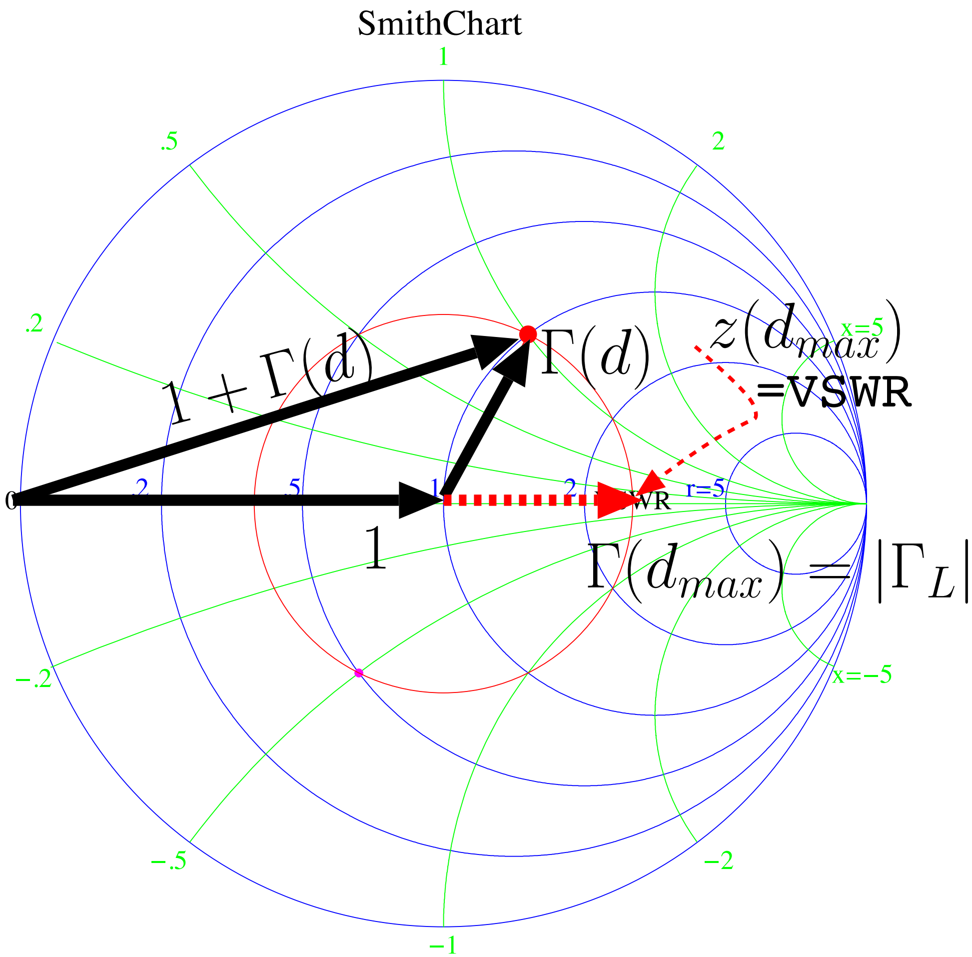


$|1+\Gamma(d)|$ maximizes for $d = d_{max}$

$|1+\Gamma(d)|$ minimizes for $d = d_{min}$ such that $\Gamma(d_{min}) = -\Gamma(d_{max})$

- The extreme values the VSWR can take are:

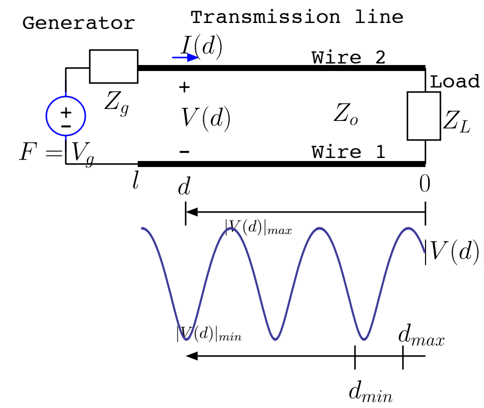
    1. VSWR=1 if $|\Gamma_L| = 0$ and the TL carries no reflected wave.
    2. VSWR=∞ if $|\Gamma_L| = 1$ corresponding to having a short, open, or a purely reactive load that causes a total reflection.



$|1 + \Gamma(d)|$ maximizes for $d = d_{max}$

$|1 + \Gamma(d)|$ minimizes for $d = d_{min}$ such that $\Gamma(d_{min}) = -\Gamma(d_{max})$

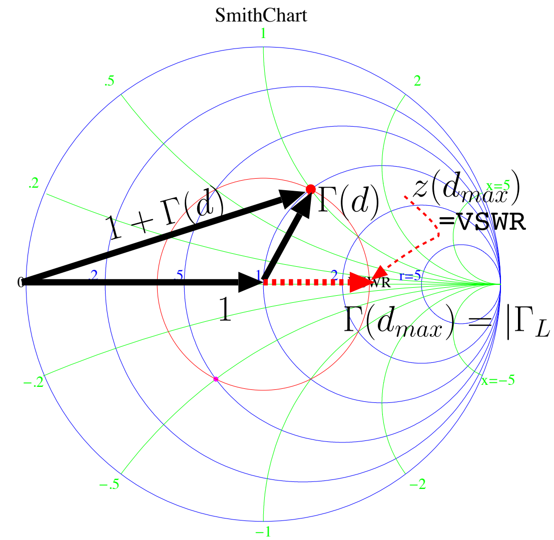
- In the lab it is easy and useful to determine the VSWR and d_{max} or d_{min} of a TL circuit with an unknown load, since



1. given the VSWR,
$$|\Gamma_L| = \frac{\text{VSWR} - 1}{\text{VSWR} + 1}$$
is easily determined, and

2. given d_{max} or d_{min} the complex Γ_L or its transform z_L can be easily obtained.

Complex addition displayed graphically superposed on a Smith Chart



$|1 + \Gamma(d)|$ maximizes for $d = d_{max}$

$|1 + \Gamma(d)|$ minimizes for $d = d_{min}$ such that $\Gamma(d_{min}) = -\Gamma(d_{max})$

Say d_{max} is known: then,

- since (as we have seen above)
$$\Gamma(d_{max}) = \Gamma_L e^{-j2\beta d_{max}} = |\Gamma_L|$$
it follows that
$$\Gamma_L = |\Gamma_L| e^{j2\beta d_{max}} \quad \Rightarrow \quad z_L = \frac{1 + \Gamma_L}{1 - \Gamma_L}.$$

- alternatively, z_L can be obtained directly on the SC by rotating counter-clockwise by d_{max} from the location of
$$z(d_{max}) = \text{VSWR}.$$

These techniques are illustrated in the next example.

Example 1: An unknown load Z_L on a $Z_o = 50\,\Omega$ TL has

$$V(d_{min}) = 20\,\mathrm{V}, \quad , d_{min} = 0.125\lambda \;\text{ and }\; \text{VSWR}{=}4.$$

Determine (a) the load impedance Z_L, and (b) the average power P_L absorbed by the load.

Solution:

(a) As shown in the top SC in the margin, VSWR=4 is entered in the SC as $z(d_{max}) = 4 + j0$, and constant $|\Gamma_L|$ circle is then drawn (red circle) passing through $z(d_{max}) = 4$.

Right across $z(d_{max}) = 4$ on the circle is $z(d_{min}) = 0.25$.

A counter-clockwise rotation from $z(d_{min}) = 0.25$ by one fourth of a full circle corresponding to a displacement of $d_{min} = 0.125\lambda$ (a full circle corresponds to a $\lambda/2$ displacement) takes us to

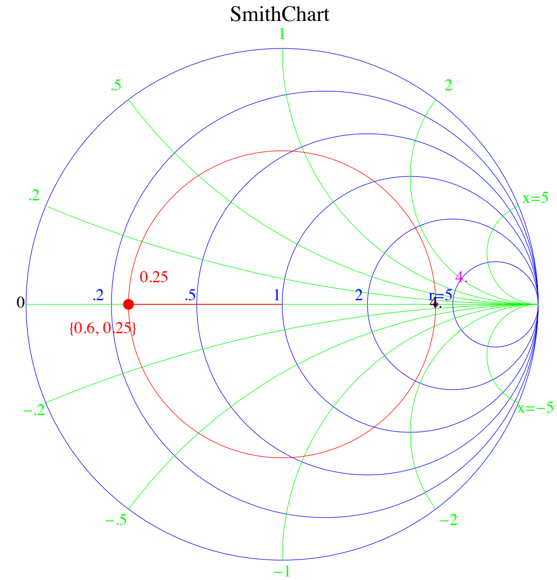
$$z_L \approx 0.4706 - j0.8823$$

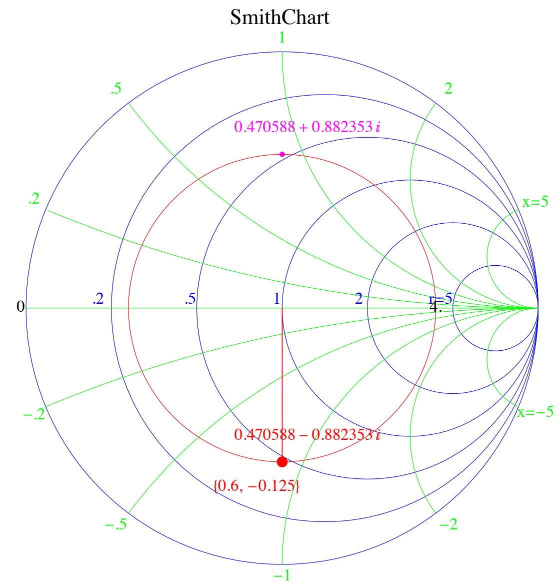
as shown in the second SC. Hence, this gives

$$Z_L = Z_o z_L = 50(0.4706 - j0.8823) = 23.53 - j44.12\,\Omega.$$

(b) We will calculate P_L by using $V(d_{min})$ and $I(d_{min})$. Since

$$z(d_{min}) = 0.25 \;\text{ it follows that }\; Z(d_{min}) = \frac{1}{4}50\,\Omega = 12.5\,\Omega.$$





Therefore the voltage and current phasors at the voltage minimum location are

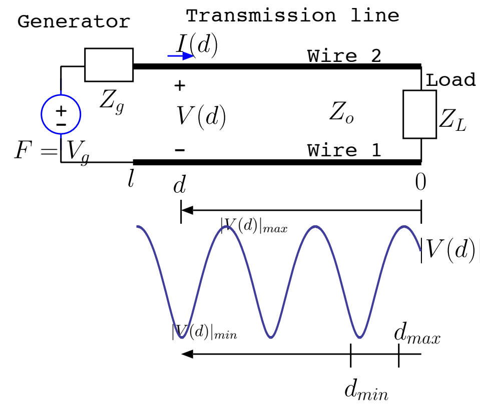
$$V(d_{min}) = 20\,\mathrm{V} \quad \text{and} \quad I(d_{min}) = \frac{20\,\mathrm{V}}{12.5\,\Omega}.$$

Average power transported toward the load at $d - d_{min}$ is, therefore,

$$P(d_{min}) = \frac{1}{2}\mathrm{Re}\{V(d_{min})I(d_{min})^*\} = \frac{1}{2}\mathrm{Re}\{20\frac{20}{12.5}\} = \frac{400}{25}\,\mathrm{W} = 16\,\mathrm{W}.$$

Since the TL is assumed to be lossless we should have

$$P_L = P(d_{min}) = 16\,\mathrm{W}.$$



Example 2: If the TL circuit in Example 1 has $l = 0.625\lambda$, and a generator with an internal impedance $Z_g = 50\,\Omega$, determine the generator voltage V_g.

Solution: Given that $l = 0.625\lambda$ and $d_{min} = 0.125\lambda$, we note that there is just one half-wave transformer between $l = 0.625\lambda$ and $d_{min} = 0.125\lambda$. Therefore

$$V_{in} = -V(d_{min}) = -20\,\mathrm{V} \quad \text{and} \quad Z_{in} = Z(d_{min}) = 12.5\,\Omega.$$

But also

$$V_{in} = V_g \frac{Z_{in}}{Z_g + Z_{in}}.$$

Consequently,

$$V_g = V_{in}\frac{Z_g + Z_{in}}{Z_{in}} = -20\frac{50 + 12.5}{12.5} = -20\frac{62.5}{12.5} = -100\,\mathrm{V}.$$

Example 3: Determine V^+ and V^- in the circuit of Examples 1 and 2 above such that the voltage phasor on the line is given by

$$V(d) = V^+ e^{j\beta d} + V^- e^{-j\beta d}.$$

Solution: Looking back to Example 1 (also see the SC's in the margin), we first note that

$$|\Gamma_L| = \frac{\text{VSWR} - 1}{\text{VSWR} + 1} = \frac{4-1}{4+1} = 0.6 = \Gamma(d_{max}) = -\Gamma(d_{min}).$$

Hence, evaluating $V(d)$ at $d = d_{min}$, we have

$$\begin{aligned} V(d_{min}) &= V^+ e^{j\beta d_{min}}(1 + \Gamma(d_{min})) \\ &= V^+ (e^{j\frac{2\pi}{\lambda}\frac{\lambda}{8}})(1 + (-0.6)) = 0.4 e^{j\frac{\pi}{4}} V^+ = 20\,\text{V}, \end{aligned}$$

from which

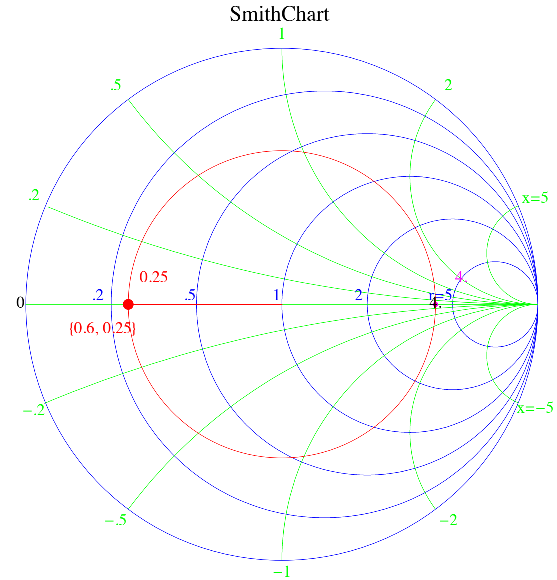
$$V^+ = 50 e^{-j\frac{\pi}{4}}\,\text{V}.$$

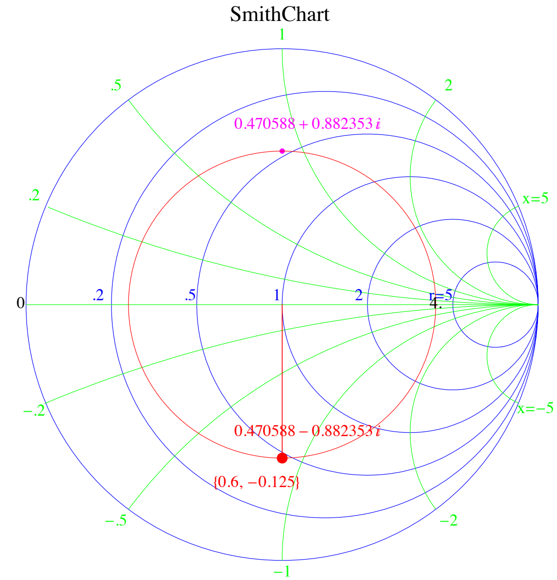
Since

$$\Gamma_L = \Gamma(0) = \Gamma(d_{min}) e^{j2\beta d_{min}} = -0.6 e^{j\frac{\pi}{2}},$$

it follows that

$$V^- = \Gamma_L V^+ = -0.6 e^{j\frac{\pi}{2}} \times 50 e^{-j\frac{\pi}{4}} = -30 e^{j\frac{\pi}{4}}\,\text{V}.$$





Example 4: Determine the load voltage and current $V_L = V(0)$ and $I_L = I(0)$ in the circuit of Examples 1-3 above.

Solution: In general,

$$V(d) = V^+ e^{j\beta d} - V^- e^{-j\beta d} \quad \text{and} \quad I(d) = \frac{V^+ e^{j\beta d} - V^- e^{-j\beta d}}{Z_o}.$$

Therefore,

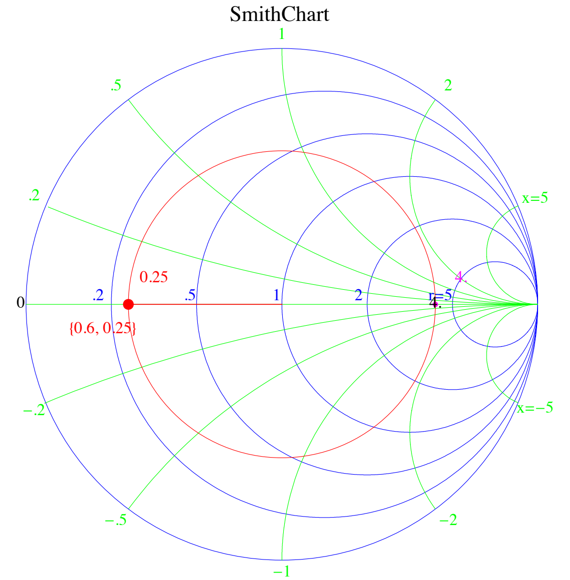
$$V_L = V(0) = V^+ + V^- \quad \text{and} \quad I_L = I(0) = \frac{V^+ - V^-}{Z_o}.$$

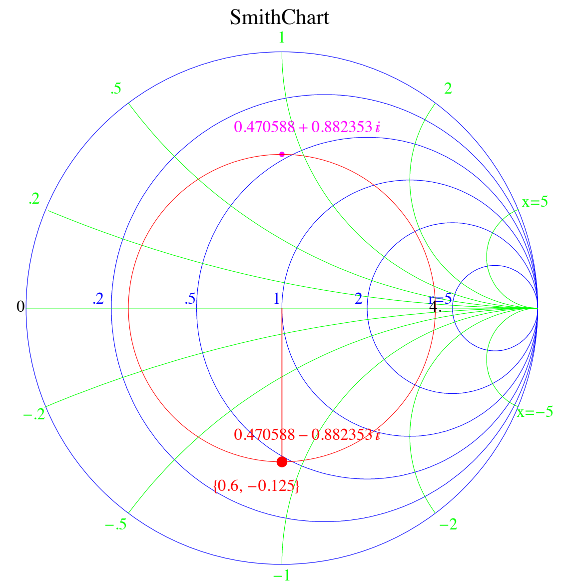
Using $Z_o = 50\,\Omega$ and

$$V^+ = 50 e^{-j\frac{\pi}{4}}\,\text{V} \quad \text{and} \quad V^- = -30 e^{j\frac{\pi}{4}}\,\text{V}$$

from Example 3, we find that

$$V_L = 50 e^{-j\frac{\pi}{4}} - 30 e^{j\frac{\pi}{4}}\,\text{V} \quad \text{and} \quad I_L = \frac{50 e^{-j\frac{\pi}{4}} + 30 e^{j\frac{\pi}{4}}}{50} = e^{-j\frac{\pi}{4}} + 0.6 e^{j\frac{\pi}{4}}\,\text{A}.$$

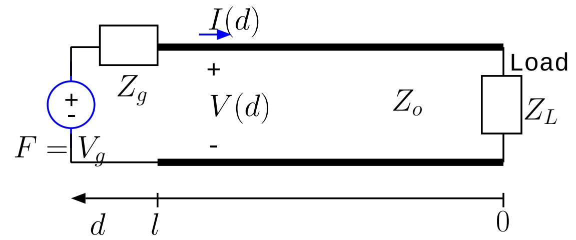




37 Smith Chart and impedance matching

- In lossless TL circuits the average power input P_{in} at the generator end precisely matches the average power delivered to the load, P_L.

  In fact, P_{in} and P_L also match the average power $P(d)$ transported on the line at an arbitrary d.



- We have in general

$$\begin{aligned} P(d) &= \frac{1}{2}\mathrm{Re}\{V(d)I^*(d)\} \\ &= \frac{1}{2}\mathrm{Re}\{(V^+e^{j\beta d} + V^-e^{-j\beta d})(\frac{V^+e^{j\beta d} - V^-e^{-j\beta d}}{Z_o})^*\} \\ &= \frac{1}{2}\mathrm{Re}\{\frac{|V^+|^2}{Z_o} - \frac{|V^-|^2}{Z_o} + \frac{V^-V^{+*}e^{-j2\beta d} - (V^-V^{+*}e^{-j2\beta d})^*}{Z_o}\} \\ &= \frac{|V^+|^2}{2Z_o} - \frac{|V^-|^2}{2Z_o}. \end{aligned}$$

  - Note that $P(d)$ is the difference of power transported $\frac{|V^+|^2}{2Z_o}$ **toward the load** by the "forward-going" wave, and $\frac{|V^-|^2}{2Z_o}$ **toward the generator** by the reflected wave.
  - Also note that

$$P(d) = \frac{|V^+|^2}{2Z_o} - \frac{|V^-|^2}{2Z_o} = \frac{|V^+|^2}{2Z_o}(1 - |\Gamma_L|^2)$$

  so that $|\Gamma_L|^2$ is an effective **power reflection coefficient**.

Power tx'ed toward the load:

$$\frac{|V^+|^2}{2Z_o}.$$

Power tx'ed toward the generator:

$$\frac{|V^-|^2}{2Z_o}.$$

Power reflection coefficient:

$$|\Gamma_L|^2.$$

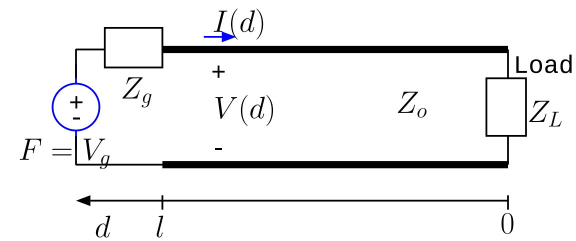
Power transmission coeff.:

$$1 - |\Gamma_L|^2.$$

- In TL circuits with load impedances Z_L **unmatched** to the characteristic impedance Z_o, the **reflected power**

$$\frac{|V^+|^2}{2Z_o}|\Gamma_L|^2$$

will be non-zero and the VSWR>1.



This a condition not favored by practical signal generators used in TL circuits.

- Most generators are *designed* (in their biasing arrangements) to operate in circuits with low VSWR (close to unity), requiring Z_{in} closely matched to R_g, most frequently $50\,\Omega$, an optimal characteristic impedance value for coax-lines (when line losses are taken into account).

- Thus a standard procedure is to use TL's with $Z_o = R_g$, and utilize a *lossless* **impedance matching network** on the TL if the load impedance $Z_L \neq Z_o$.

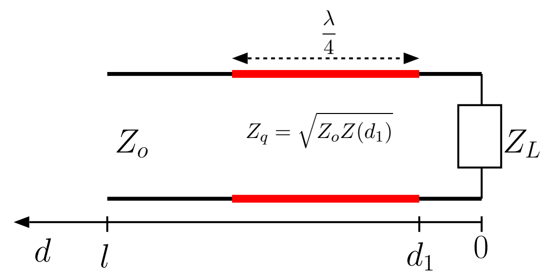
  – This practice is called **impedance matching**.

Impedance matching achieves VSWR=1 between the generator and the matching network inserted at a location between the load and the generator.

- The inserted network should be designed to yield an input impedance equal Z_o at its input terminals.

The following examples illustrate different ways of achieving an impedance match.

Quarter-wave matching



Example 1: ***Quarter-wave matching* of resistive loads:**

Consider a TL with $Z_L = 25\,\Omega$ and $R_g = Z_o = 50\,\Omega$. Since $Z_L \neq Z_o$ the load is unmatched and the VSWR>1.

To reduce the VSWR on the line connected to the generator to unity, we can insert a **quarter-wave transformer** right after Z_L — i.e., at $d_1 = 0$ in the circuit shown in the margin — with a characteristic impedance

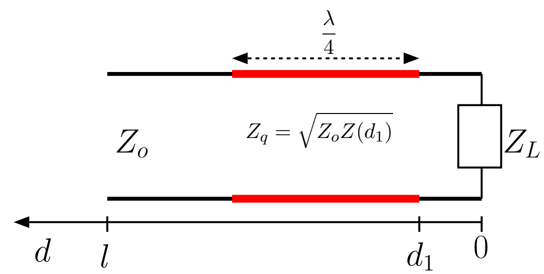
$$Z_q = \sqrt{25 \times 50} = \sqrt{1250} = 35.35\,\Omega.$$

The impedance at the input terminals of the quarter-wave transformer (on the left) is then Z_o, i.e., 50 Ω, implying a perfect impedance match.

- Quarter-wave matching illustrated above is a very commonly used matching technique.
- It is a straightforward application of the quarter-wave transformer impedance formula

$$Z_{in} = \frac{Z_q^2}{Z_L}$$

for a transformer with characteristic impedance Z_q.



Example 2: ***Quarter-wave matching*** of reactive loads:

Consider a TL with $Z_L = 50 + j50\,\Omega$ and $R_g = Z_o = 50\,\Omega$. Since $Z_L \neq Z_o$ the load is unmatched and the VSWR>1.

We cannot insert the quarter-wave transformer right after the load because then we would need a complex valued Z_q implying a lossy matching network.

Instead, we insert a **quarter wave transformer** a distance d_1 to the left of Z_L, where d_1 is selected, using a SC, to have a purely resistive $Z(d_1)$. In that case, the quarter-wave transformer impedance formula

$$Z_q = \sqrt{Z(d_1) \times 50}$$

yields a real valued Z_q as needed. This procedure leads to having $d_1 = d_{max}$ or $d_1 = d_{min}$ corresponding to the positions of voltage maxima and minima on the line.

As shown in the margin,

$$Z(d_1) = 50(2.62 + j0) = 131\,\Omega.$$

for

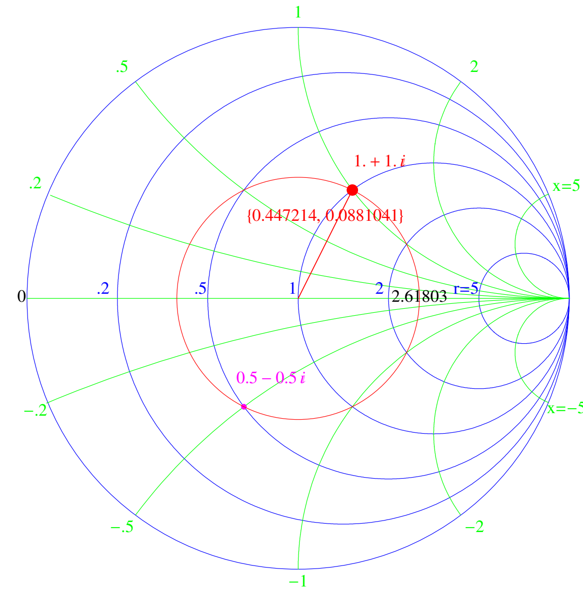
$$d_1 \approx 0.250\lambda - 0.162\lambda = 0.088\lambda$$

is a suitable choice for quarter-wave matching. In that case we need

$$Z_q = \sqrt{131 \times 50} = 50 \times \sqrt{2.62}\,\Omega$$

for the quarter wave transformer in order match to load to a line with $Z_o = 50\,\Omega$.

SmithChart



Note that:

$$z(d_1) = z(d_{max}) = \textbf{VSWR} \approx 2.62$$

as marked on the SC.
Also

$$d_{max} \approx 0.088\lambda$$

since, as marked on the SC, the angle of Γ_L is 0.088λ.

Example 3: ***Single-stub tuning:***

Consider a TL with $Z_L = 100 - j50\,\Omega$ and $R_g = Z_o = 50\,\Omega$. Since $Z_L \neq Z_o$ the load is unmatched and the VSWR>1.

We will insert a **shorted-stub** a distance d_1 to the left of Z_L in parallel with the line to achieve an impedance match.

Distance d_1 will be selected, using a SC, to have a normalized admittance of

$$y(d_1) = 1 + jb$$

so that a stub, with a normalized input admittance

$$y_{stub} = -jb,$$

can be added in parallel to have a combined admittance of

$$y(d_1) + y_{stub} = 1 + j0$$

and achieve a perfect impedance match (i.e., VSWR=1).

In specific

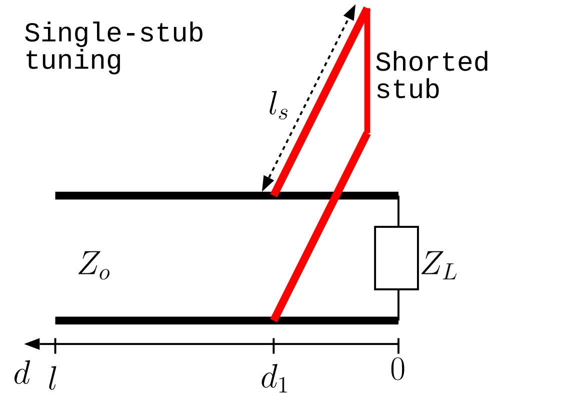
$$z_L = \frac{Z_L}{Z_o} = 2 - j1 \text{ and } y_L = \frac{1}{z_L} = 0.4 + j0.2$$

as shown on the SC on the top in the margin. We rotate clockwise on the SC by an amount corresponding to d_1 to obtain

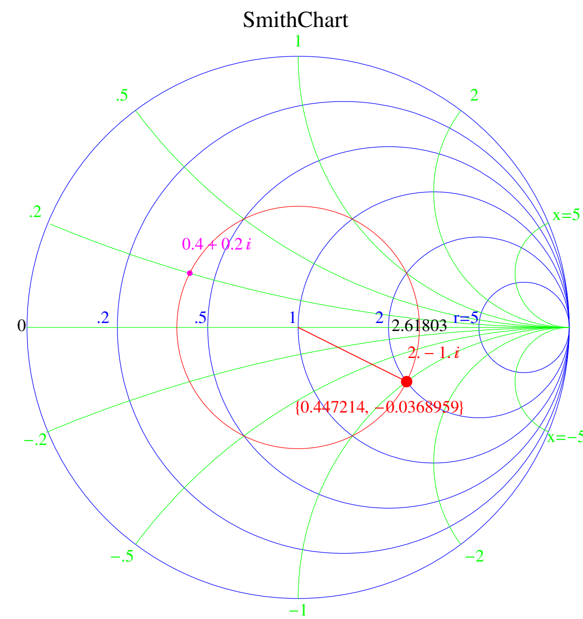
$$y(d_1) = 1 + j1$$

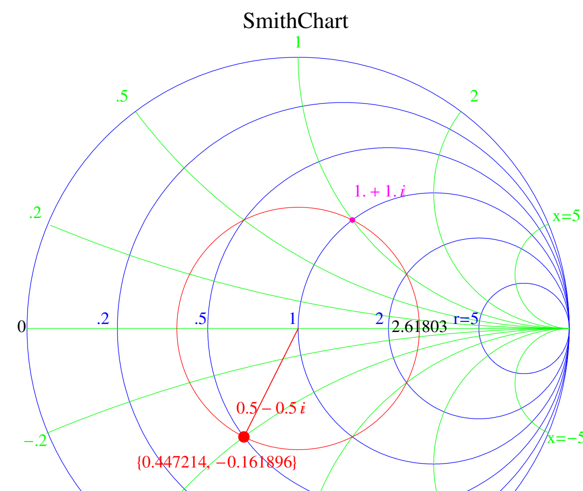
on the "$g = 1$" or "$y = 1 + jb$" circle as shown in the bottom SC. From the amount of rotation we determine

$$d_1 \approx 0.162\lambda - 0.037\lambda = 0.125\lambda.$$



Want $y(d_1) + y_{stub} = 1$





The required input impedance of the shorted stub to achieve

$$y(d_1) + y_{stub} = 1 + j0$$

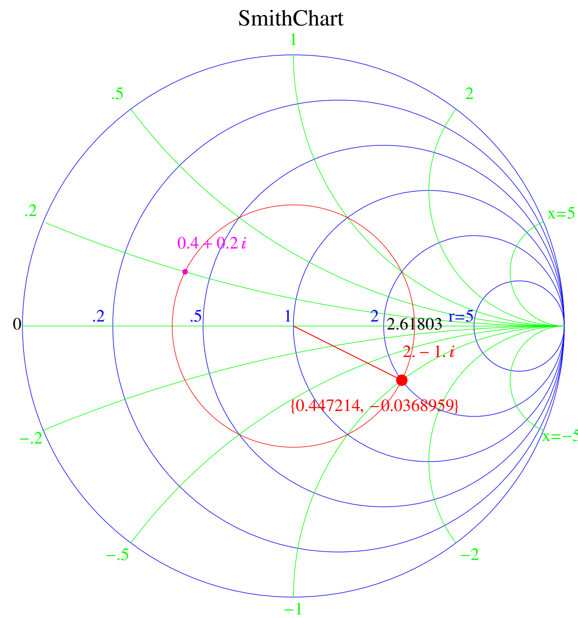
is

$$y_{stub} = -1j.$$

To achieve this input admittance the required stub length is

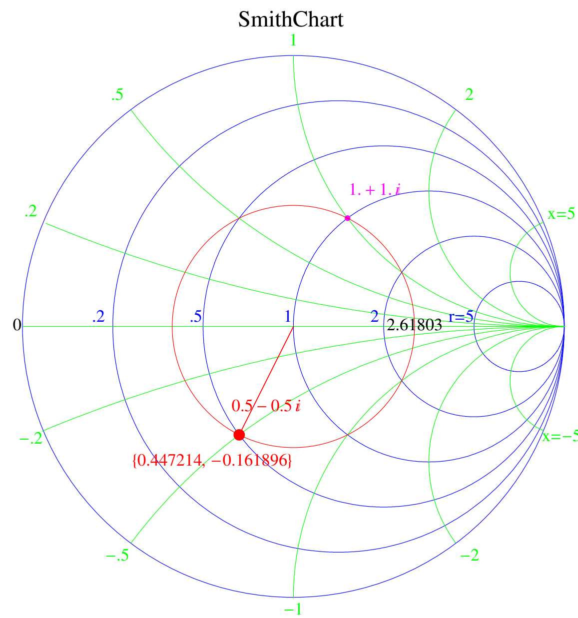
$$l_s = \frac{\lambda}{8} = 0.125\lambda$$

as determined from the SC — start at $y = \infty$ point on the SC on the far right (corresponding to the short termination), and then rotate clockwise (toward the generator) until the normalized admittance reads $-j1$; the amount of rotation indicates the required l_s.



- Another matching technique called ***double-stub tuning*** uses *two* shorted stubs of lengths l_1 and l_2 located at fixed values of d_1 and d_2.
  - Typically d_1 is zero or $\frac{\lambda}{4}$, and
  - $d_2 = d_1 + 3\frac{\lambda}{8}$.

    Vary l_1 and l_2 until VSWR is reduced to 1 near the generator end.



The advantage of double-stub tuning is avoiding changes of stub locations when Z_L is changed. It's implementation on a SC is considerably more complicated than single-stub tuning.

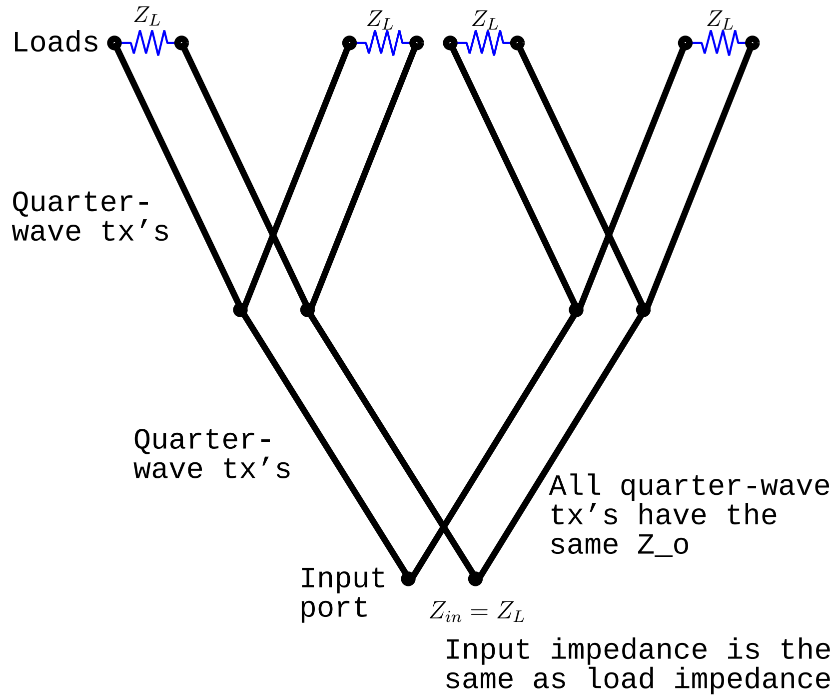
38 Distribution networks

- A **corporate ladder** network that combines 4 identical loads Z_L into a single equivalent input impedance Z_L is shown in the margin.

- In this network 6 different quarter-wave transformers with arbitrary but identical characteristic impedances Z_o are utilized.

  You should be able to compute the load voltages V_L in the network in terms of input voltage V_{in} applied across the input port by using the current-forcing formula for the quarter-wave transformer introduced earlier on.

CORPORATE LADDER NETWORK

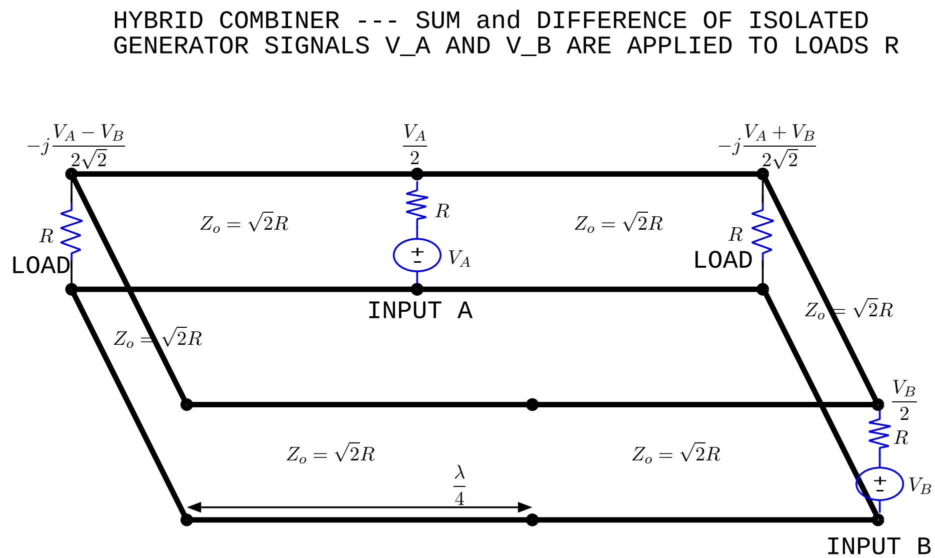


By symmetry, the loads absorb equal avg power, a quarter of the input power each.

- If, in a **corporate ladder** network, $Z_L = Z_o$, then TL segment lengths connected to each Z_L can be varied at will without affecting the input impedance $Z_L = Z_o$ (why?).

- Allowing variable length TL's connected to each Z_L makes it possible to adjust and vary the phase of the voltage and current of each Z_L — this is useful, for instance, in feeding phased antenna arrays (Z_L represents an antenna load) to achieve steerable radiation patterns.

- A **hybrid combiner** network shown in the margin can be used to *excite* two identical TL loads R (e.g., antenna arrays impedance matched to have input impedances R) with independent signal generators V_A and V_B having equal internal resistances R matched to the load resistance.



- The hybrid "rat-race" combiner is built with 6 quarter-wave transformers of identical

$$Z_o = \sqrt{2}R,$$

in which case the generators V_A and V_B see impedance-matched loads (at the hybrid inputs where they are connected) and produce load voltages proportional to $V_A \pm V_B$ as shown in the diagram.

- Generators A and B with open ckt voltages V_A and V_B are *isolated* from one another's influence because of "destructive interference" between the two paths from each generator to the other one (two paths have a $\frac{\lambda}{2}$ length difference).

- This very special situation allows one to calculate the various terminal voltages on the hybrid due to V_A and V_B one-at-a-time as if loads R were isolated from generator-B and -A (by "virtual shorts" existing across generator terminals when V_B and V_A are suppressed) in turns, and then superpose the results.

- Terminal voltages obtained with that procedure (those shown on the diagram) turn out to be valid when both generators are active as can easily be checked for self-consistency by using the current-forcing equations introduced earlier. For instance, the total current into generator-A terminal (flowing from both sides) is

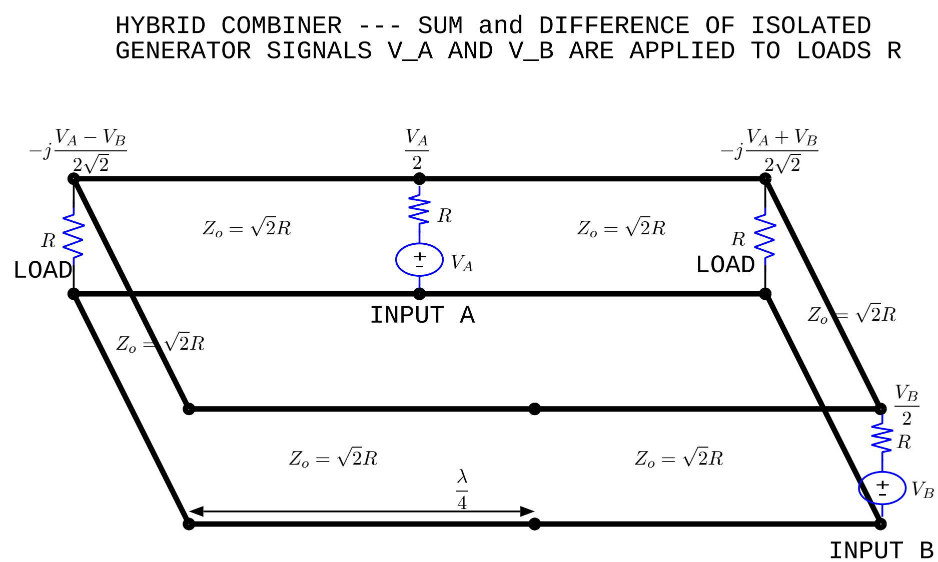
$$I_A = -\frac{j}{R\sqrt{2}}(-j\frac{V_A - V_B}{2\sqrt{2}}) - \frac{j}{R\sqrt{2}}(-j\frac{V_A + V_B}{2\sqrt{2}}) = -\frac{V_A}{2R},$$

and hence the voltage drop from the same terminal to the ground is

$$I_A R + V_A = -\frac{V_A}{2R}R + V_A = \frac{V_A}{2}$$

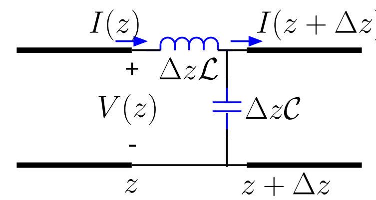
as marked explicitly on the diagram. All self-consistency tests that can be applied with the given expressions are passed, and so the results given are valid.

- The input and output ports of a hybrid combiner can be swapped while still maintaining the properties of the hybrid — namely, input impedance R, and output signals the sum and difference of generator voltages.

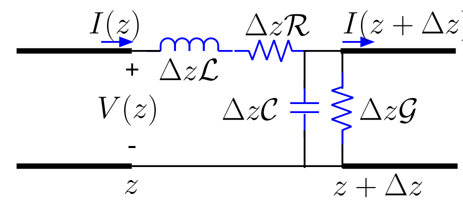


39 Lossy lines

IDEAL LOSSLESS T.L.:



REALISTIC LOSSY T.L.:



- *Lossless* TL's we have been studying so far are idealizations of *real* TL's which are invariably **lossy**.
  - Here, we are making reference to ***Ohmic*** energy losses in the conducting wires of the TL, as well as to losses in the imperfect dielectric separating the two conductors.
- The effect of wire losses in TL's is modelled by adding a $\Delta z\mathcal{R}$ resistance in series with $\Delta z\mathcal{L}$ inductor in the equivalent circuit model of an infinitesimal $(\Delta z \ll \lambda)$ TL section as shown in the margin.
- In addition, a shunt conductance $\Delta z\mathcal{G}$ in parallel with capacitance $\Delta z\mathcal{C}$ accounts in the lossy model for dielectric losses.
- While the phasor form of telegrapher's equations for a lossless TL is

$$-\frac{\partial V}{\partial z} = j\omega\mathcal{L}I \quad \text{and} \quad -\frac{\partial I}{\partial z} = j\omega\mathcal{C}V,$$

for lossy lines — where impedance per unit length $j\omega\mathcal{L}$ must be replaced by $j\omega\mathcal{L}+\mathcal{R}$ and conductance per unit length $j\omega\mathcal{C}$ by $j\omega\mathcal{C}+\mathcal{G}$ — the equations take the form

$$-\frac{\partial V}{\partial z} = (j\omega\mathcal{L}+\mathcal{R})I \quad \text{and} \quad -\frac{\partial I}{\partial z} = (j\omega\mathcal{C}+\mathcal{G})V.$$

Using perturbation theory, it can be shown that for a coax of inner and outer radii a and b,

$$\mathcal{R} = \sqrt{\frac{f\mu}{\pi\sigma}}\left(\frac{1}{a}+\frac{1}{b}\right),$$

while for a parallel-plate transmission line of width W,

$$\mathcal{R} = \frac{4\pi}{W}\sqrt{\frac{f\mu}{\pi\sigma}},$$

in terms of conductivity σ and permeability μ of the T.L. conductors.

- We will next show that

  1. lossless line solutions can be readily modified to account for loss effects introduced by Ohmic energy losses in $\mathcal{R}$ and $\mathcal{G}$,
  2. lossless line results we have learned up till now are by and large valid even on lossy lines provided that
     (a) frequency ω is sufficiently large, and
     (b) voltage and current solutions $V^{\pm}e^{\pm j\beta d}$ and $\frac{V^{\pm}}{\pm Z_o}e^{\pm j\beta d}$ are modified by multiplying an attenuation term $e^{\pm\alpha d}$ which only matters in practice when $d \gg \lambda$.

- Note that lossless line solutions of telegrapher's equations can be restated as

$$V = V^{\pm}e^{\pm\gamma d} \text{ and } I = \frac{V^{\pm}}{\pm Z_o}e^{\pm\gamma d},$$

where

$$\gamma = j\beta = j\omega\sqrt{\mathcal{L}\mathcal{C}} = \sqrt{(j\omega\mathcal{L})(j\omega\mathcal{C})} \text{ and } Z_o = \sqrt{\frac{\mathcal{L}}{\mathcal{C}}} = \sqrt{\frac{j\omega\mathcal{L}}{j\omega\mathcal{C}}}.$$

  – Replacing $j\omega\mathcal{L}$ by $j\omega\mathcal{L}+\mathcal{R}$, and $j\omega\mathcal{C}$ by $j\omega\mathcal{C}+\mathcal{G}$, we obtain

$$\gamma = \sqrt{(j\omega\mathcal{L}+\mathcal{R})(j\omega\mathcal{C}+\mathcal{G})} \text{ and } Z_o = \sqrt{\frac{j\omega\mathcal{L}+\mathcal{R}}{j\omega\mathcal{C}+\mathcal{G}}}$$

  in the lossy case.

While waves governed by lossy γ and Z_o (see margin) can exhibit substantially different beavior than the lossless waves (examined in the previous sections), at high frequencies the wave properties are reasonably similar as alluded in item (2) above. We next examine this simplified high-frequency limit.

$$\gamma = \sqrt{(j\omega\mathcal{L}+\mathcal{R})(j\omega\mathcal{C}+\mathcal{G})}$$

$$Z_o = \sqrt{\frac{j\omega\mathcal{L}+\mathcal{R}}{j\omega\mathcal{C}+\mathcal{G}}}$$

- At high frequencies ω, such that $\omega\mathcal{L} \gg \mathcal{R}$ and $\omega\mathcal{C} \gg \mathcal{G}$, we have
  - characteristic impedance

    $$Z_o = \sqrt{\frac{j\omega\mathcal{L}+\mathcal{R}}{j\omega\mathcal{C}+\mathcal{G}}} \approx \sqrt{\frac{\mathcal{L}}{\mathcal{C}}}$$

    just as in the lossless case[1], and
  - complex propagation constant

    $$\begin{aligned}\gamma &= \sqrt{(j\omega\mathcal{L}+\mathcal{R})(j\omega\mathcal{C}+\mathcal{G})} = j\omega\sqrt{\mathcal{LC}}\sqrt{1+\frac{\mathcal{R}}{j\omega\mathcal{L}}}\sqrt{1+\frac{\mathcal{G}}{j\omega\mathcal{C}}} \\ &\approx j\omega\sqrt{\mathcal{LC}}(1+\frac{\mathcal{R}}{j2\omega\mathcal{L}})(1+\frac{\mathcal{G}}{j2\omega\mathcal{C}}) \approx j\omega\sqrt{\mathcal{LC}} + \frac{1}{2}(\frac{\mathcal{R}}{Z_o}+\mathcal{G}Z_o) \\ &= j\beta+\alpha\end{aligned}$$

    with

    $$\beta \approx \omega\sqrt{\mathcal{LC}} \quad \text{and} \quad \alpha \approx \beta(\frac{\mathcal{R}}{2\omega\mathcal{L}}+\frac{\mathcal{G}}{2\omega\mathcal{C}}) = \frac{1}{2}(\frac{\mathcal{R}}{Z_o}+\mathcal{G}Z_o).$$

[1] In fact $Z_o = \sqrt{\frac{\mathcal{L}}{\mathcal{C}}}$ is *exact* even for a lossy line if $\frac{\mathcal{L}}{\mathcal{R}} = \frac{\mathcal{C}}{\mathcal{G}}$.

- Note that $\beta = \frac{2\pi}{\lambda}$ is the same as in the lossless case, and since

$$\alpha \approx \beta(\frac{\mathcal{R}}{2\omega\mathcal{L}} + \frac{\mathcal{G}}{2\omega\mathcal{C}}) \ll \beta,$$

the "penetration depth" $\delta \equiv \frac{1}{\alpha}$ of voltage and current waves on the TL is much longer than a wavelength $\lambda = \frac{2\pi}{\beta}$ in this regime.

In summary, in the high-frequency regime, characteristic impedance Z_o and wavenumber β are (practically) the same as they are on lossless lines, but signals do attenuate by a factor $e^{\pm\alpha d}$ which should not be (and cannot be) neglected over long distances d exceeding many wavelengths λ.

- At lower frequencies where the above approximations cannot be justified, a more careful analysis of lossy line equations is warranted.

- Finally, for an *air-filled coax* with inner and outer radii a and b, it can be shown that the attenuation constant

$$\alpha = \frac{1}{2}\frac{\mathcal{R}}{Z_o} = \frac{1}{2}\frac{\sqrt{\frac{f\mu_o}{\pi\sigma}}\frac{1}{b}(1+\frac{b}{a})}{\frac{\eta_o}{2\pi}\ln(\frac{b}{a})},$$

which minimizes, at a fixed outer radius b, for $\frac{b}{a} \approx 3.6$, which in turn results in an "optimal" characteristic impedance of

$$Z_o = \frac{\eta_o}{2\pi}\ln(\frac{b}{a}) = 60\ln(\frac{b}{a})\,\Omega \approx 75\,\Omega$$

for the same coax. Note that this result is independent of σ, the conductivity of inner and outer conductors of the coax.

– For a *dielectric filled coax* having $\epsilon = \frac{9}{4}\epsilon_o$ — implying $v_p = \frac{2}{3}c = 2 \times 10^8$ m/s — the same ratio $\frac{b}{a} \approx 3.6$ of outer and inner conductor radii leads to $Z_o \approx 50\,\Omega$, the most common Z_o encountered in practical applications.

– The above result should also explain why having a thicker coax — larger b — is better when losses are a concern.