FARADAY'S LAW OF INDUCTION

Back in 1831, Michael Faraday reported a series of experiments on using magnets and motion to *induce* an EMF in a circuit and make a current flow through it. Among other methods, he induced an EMF in a wire coil by means of:

1. Moving the coil towards a magnet or away from a magnet.
2. Keeping the coil stationary while moving the magnet towards the coil or away from it.
3. Keeping the coil stationary near a stationary electromagnet, and varying the current through the magnet (and hence the magnetic field it makes).

Faraday found that all such methods of *magnetic induction* operate according to the same *flux rule*, usually called the *Faraday's Law of Induction: Whenever — and for whatever reason — the magnetic flux through an electric circuit changes, it induces EMF in the circuit according to*

$$\mathcal{E} = -\frac{d\Phi}{dt}\,. \tag{1}$$

The minus sign in this formula encodes the *Lenz rule: The current due to the induced EMF tries to counteract the change of the flux which has induced the EMF.*

But despite the ultimate universality of the Faraday's Law (1), in the terms of fields and forces, different methods of magnetic induction work through two rather different mechanisms. Faraday's method#1 — moving the coil — induces the *motional EMF*, which ultimately stems from the Lorentz forces

$$\mathbf{F} = q\mathbf{v}\times\mathbf{B} \tag{2}$$

on electrons in the moving wires. On the other hand, the other two methods #2 and #3 — moving the magnet, or varying the current in electromagnet — generate a *time-dependent magnetic field* $\mathbf{B}(\hat{\mathbf{x}},t)$*, which induces a non-potential electric field* obeying

$$\nabla\mathbf{E}_{\text{induced}} = -\frac{\partial B}{\partial t}\,; \tag{3}$$

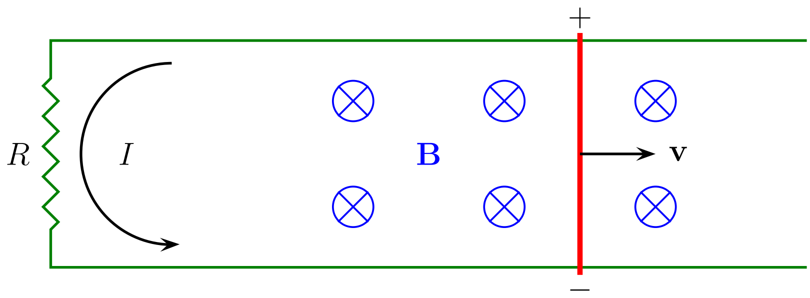
it is this non-potential electric field which gives rise to the EMF induced in the stationary coil.

The purpose of these notes is to briefly explain these two mechanisms of magnetic induction.

BTW, despite the apparent difference between the two mechanisms, they are related to each other by special relativity. In fact, it was Albert Einstein's investigation of how Faraday's methods 1 and 2 induce exactly the same EMF for the same *relative* motion of the coil and the magnet which lead him to the special relativity theory in the first place!

Motional EMF

Let 's start with the physical origin of the *motional EMF* — the EMF due to moving a wire across the magnetic field. Consider a wire moving across the magnetic fields, for example



(4)

The wire's velocity adds to the velocities of the conducting electrons relative to the wire,

$$\mathbf{v} \;=\; \mathbf{v}_{\text{wire}} \;+\; \mathbf{v}_{\text{rel}} \tag{5}$$

hence the magnetic Lorenz force on an electron is

$$\mathbf{F} \;=\; (-e)\mathbf{v}_{\text{wire}} \times \mathbf{B} \;+\; (-e)\mathbf{v}_{\text{rel}} \times \mathbf{B}. \tag{6}$$

When we sum these forces over the conducting electrons, the terms due to $\mathbf{v}_{\text{rel}}$ add up to the net mechanical force on the wire,

$$\mathbf{F}_{\text{mech}} \;=\; \sum(-e)\mathbf{v}_{\text{rel}} \times \mathbf{B} \;=\; (-e)N_e\mathbf{v}_{\text{drift}} \times \mathbf{B} \;=\; I\mathbf{L} \times \mathbf{B}\,. \tag{7}$$

On the other hand, the forces $(-e)\mathbf{v}_{\text{wire}} \times \mathbf{B}$ make the electrons move along the wire in the direction of $-\mathbf{v}_{\text{wire}} \times \mathbf{B}$, which makes the current flow in the opposite direction of $+\mathbf{v}_{\text{wire}} \times \mathbf{B}$.

For example, in the red moving wire on the diagram (4), the force pushes the electrons down, so the current flows up. The electromotive "force" generating this current obtains as the work of the magnetic force per unit of charge; for each electron moving all the way from one end of the wire to the other end, the work is

$$W = \mathbf{L}\cdot\Big(\mathbf{F} = (-e)\mathbf{v}_{\text{wire}}\times\mathbf{B}\Big), \tag{8}$$

hence EMF

$$\mathcal{E} = \frac{W}{(-e)} = \mathbf{L}\cdot(\mathbf{v}_w\times\mathbf{B}) \tag{9}$$

where $\mathbf{L}$ is the vector length of the wire. Equation (9) can be easily generalized to curved wires moving in complicated ways through non-uniform magnetic fields,

$$\mathcal{E} = \int\limits_{\text{wire}} \big(\mathbf{v}_{\text{wire}}(\hat{\mathbf{x}})\times\mathbf{B}(\hat{\mathbf{x}})\big)\cdot d\hat{\mathbf{x}}\,. \tag{10}$$

Note: we do not have to wait until a specific electron moves all the way from one end of the wire to the other end. The same work per unit of charge flowing through the wire obtains when the electron gas in the wire collectively moves just a tiny distance along the wire; indeed, such a collective motion is mathematically equivalent to to a small fluid element of the electron gas moving the whole distance while the rest of the gas stays in place.

Relation to the Flux Rule

Formula (10) for the motional EMF has a rather obscure relation to the Faraday's flux rule. To clarify this relation, we need to take care of the exact meaning of the "magnetic flux through an electric circuit". First of all, this flux goes through the complete closed circuit through which the current is flowing, including both the moving and the non-moving wires.

In terms of the time-dependent closed loop $\mathcal{L}(t)$ spanned by this circuit, eq. (10) for the motional EMF becomes

$$\mathcal{E}_{\text{motion}}(t) = \oint\limits_{\mathcal{L}(t)} [\mathbf{v}_{\text{wire}}(\hat{\mathbf{x}}, t) \times \mathbf{B}(\hat{\mathbf{x}}, t)] \cdot d\hat{\mathbf{x}}. \tag{11}$$

To define the magnetic flux through such a time-dependent circuit loop $\mathcal{L}(t)$, we need to span it with some time-dependent surface $\mathcal{S}(t)$, then

$$\Phi[\text{through } \mathcal{L}(t)] = \iint\limits_{\mathcal{S}(t)} \mathbf{B}(\hat{\mathbf{x}}, t) \cdot d^2\mathbf{area}. \tag{12}$$

Since the $\mathbf{B}$ field is divergence-less at all times, the exact geometry of the surface $\mathcal{S}(t)$ does not matter, as long as it spans the circuit loop $\mathcal{L}(t)$ at all times t.

Now consider the time dependence of the magnetic flux (12). In general, the magnetic field $\mathbf{B}(\hat{\mathbf{x}}, t)$ is time-dependent, and the surface $\mathcal{S}(t)$ over which we integrate is also time-dependent. Hence, when we take the time derivative of the flux, we get two terms

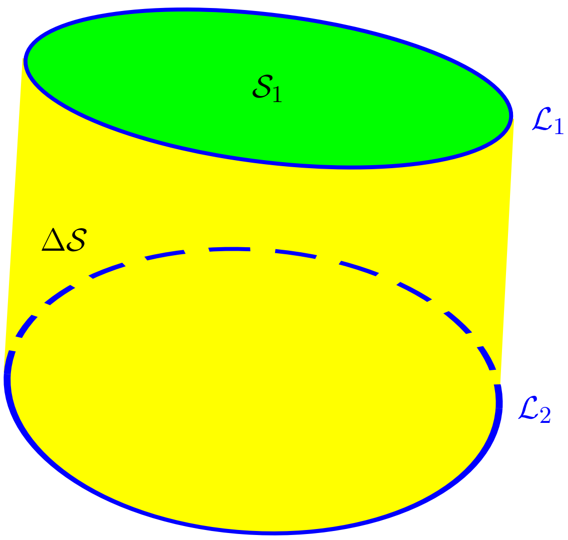
$$\frac{d\Phi}{dt} = \iint\limits_{\partial\mathcal{S}/\partial t} \mathbf{B} \cdot d^2\mathbf{area} + \iint\limits_{\mathcal{S}} \frac{\partial \mathbf{B}}{\partial t} \cdot d^2\mathbf{area}. \tag{13}$$

Theorem: *the first term here equals (minus) the motional EMF* (11) *in the moving circuit,*

$$-\iint\limits_{\partial\mathcal{S}/\partial t} \mathbf{B} \cdot d^2\mathbf{area} = \mathcal{E}_{\text{motion}} = \oint\limits_{\mathcal{L}} [\mathbf{v}_{\text{wire}}(\hat{\mathbf{x}}) \times \mathbf{B}(\hat{\mathbf{x}})] \cdot d\hat{\mathbf{x}}. \tag{14}$$

While proving this theorem, we shall assume a time-independent magnetic field, so the first term in eq. (13) becomes the entire $d\Phi/dt$. We shall take care of the time-dependent fields — and hence the second term in eq. (13) later in these notes.

Proof: Consider two successive snapshot pictures of the moving circuit loop $\mathcal{L}(t)$, the $\mathcal{L}_1$ at time t_1 and the $\mathcal{L}_2$ at a later time t_2. Let $\mathcal{S}_1$ be some surface spanning the $\mathcal{L}_1$ while $\Delta\mathcal{S}$ is a ribbon-shaped surface connecting the two loops as shown on the picture below



(15)

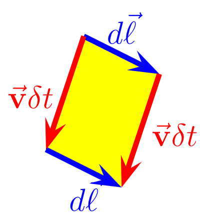
The combined surface $\mathcal{S}_2 = \mathcal{S}_1 + \Delta\mathcal{S}$ spans the loop $\mathcal{L}_2$; it might look like a peculiar choice of a surface to span the loop $\mathcal{L}_2$ at time t_2, but it does the job. Consequently, the magnetic flux at time t_2 can be calculated using this surface, thus

$$\Phi(t_2) \;=\; \iint\limits_{\mathcal{S}_1+\Delta\mathcal{S}} \mathbf{B}\cdot d^2\mathbf{area} \;=\; \iint\limits_{\mathcal{S}_1} \mathbf{B}\cdot d^2\mathbf{area} + \iint\limits_{\Delta\mathcal{S}} \mathbf{B}\cdot d^2\mathbf{area} \;=\; \Phi(t_1) + \iint\limits_{\Delta\mathcal{S}} \mathbf{B}\cdot d^2\mathbf{area}\,. \quad (16)$$

Therefore, the change of the magnetic flux through a moving loop can be expressed as the integral over the ribbon $\Delta\mathcal{S}$,

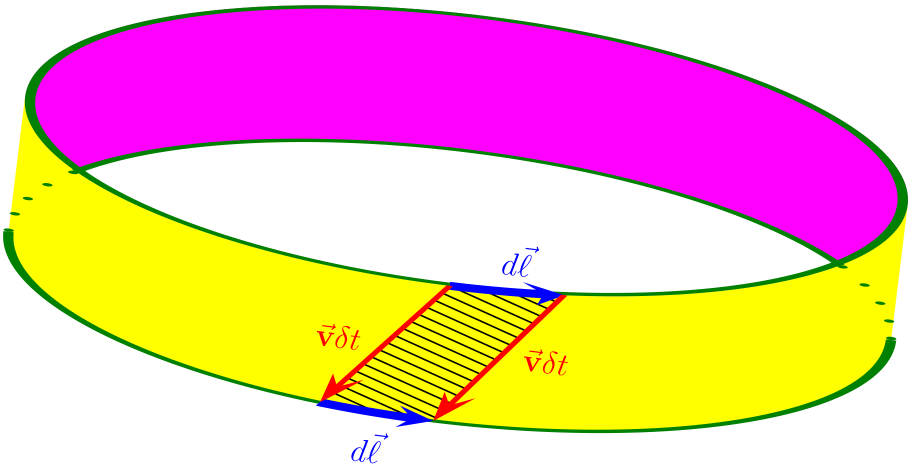
$$\Delta\Phi \;=\; \Phi(t_2) \;-\; \Phi(t_1) \;=\; \iint\limits_{\Delta\mathcal{S}} \mathbf{B}\cdot d^2\mathbf{area}\,. \quad (17)$$

Let's focus on very short time intervals $\delta t = t_2 - t_1 \to 0$, so the loops $\mathcal{L}_1$ and $\mathcal{L}_2$ are very close to each other. As each infinitesimal piece $d\vec{\ell}$ of wire moves from its old place in $\mathcal{L}_1$ to its new place in $\mathcal{L}_2$ with velocity $\mathbf{v}$, it sweeps through area



$$d^2\mathbf{area} \;=\; \vec{\mathbf{v}}\delta t \times d\vec{\ell}\,. \quad (18)$$

As to the whole moving loop $\mathcal{L}(t)$, it sweeps through a very narrow ribbon



(19)

which we may identify as $\Delta\mathcal{S}$. The vector area of this ribbon is

$$\mathbf{a} = \oint_{\mathcal{L}} \mathbf{v}\delta t \times d\vec{\ell} \tag{20}$$

where the velocity vector $\mathbf{v}$ may vary along the loop, depending on how the wires are moving.

The magnetic flux through the ribbon $\Delta\mathcal{S}$ follows from the infinitesimal areas (18):

$$\iint_{\Delta\mathcal{S}} \mathbf{B}\cdot d^2\mathbf{area} = \oint_{\mathcal{L}} \mathbf{B}\cdot(\mathbf{v}\delta t \times d\vec{\ell}) = -\delta t \times \oint_{\mathcal{L}} (\mathbf{v}\times\mathbf{B})\cdot d\vec{\ell}, \tag{21}$$

where the second equality follows from the vector identity

$$\mathbf{B}\cdot(\mathbf{v}\times d\vec{\ell}) = d\vec{\ell}\cdot(\mathbf{B}\times\mathbf{v}) = -(\mathbf{v}\times\mathbf{B})\cdot d\vec{\ell}. \tag{22}$$

But as we saw earlier in eq. (17), the magnetic flux through the ribbon $\Delta\mathcal{S}$ is precisely the change of the magnetic flux through the moving loop $\mathcal{L}$ between times t_1 and $t_2 = t_1 + \delta t$. Consequently, *the rate of change of the magnetic flux through the moving loop* obtains from eq. (21) as

$$\frac{d\Phi}{dt} = -\oint_{\mathcal{L}} (\mathbf{v}\times\mathbf{B})\cdot d\vec{\ell}. \tag{23}$$

At this point, the RHS of eq. (23) is precisely the (minus) motional EMF (11). At the same time, we have calculated the magnetic flux and its rate of change assuming a time-independent magnetic field, so the $d\Phi/dt$ on the LHS is merely the first term in eq. (13)

for the total time derivative of the magnetic flux. Thus, in a more general context eq. (23) becomes

$$\mathcal{E}_{\text{motion}} = -\left[\frac{d\Phi}{dt}\right]_{\text{motion}} = -\iint_{\partial\mathcal{S}/\partial t} \mathbf{B}\cdot d^2\textbf{area}. \tag{14}$$

Quod erat demonstrandum.

The Induced Electric Field

Let's go back to the complete Faraday's Law of Induction

$$\mathcal{E}_{\text{net}} = -\frac{d\Phi}{dt} \tag{24}$$

but now allow for a time-dependent magnetic field. In this case,

$$-\frac{d\Phi}{dt} = -\iint_{\partial\mathcal{S}/\partial t} \mathbf{B}\cdot d^2\textbf{area} - \iint_{\mathcal{S}} \frac{\partial\mathbf{B}}{\partial t}\cdot d^2\textbf{area}. \tag{13}$$

where the first term on the RHS is the motional EMF, hence

$$\mathcal{E}_{\text{net}} = \mathcal{E}_{\text{motion}} + \mathcal{E}_{\text{var.b}} \tag{25}$$

where

$$\mathcal{E}_{\text{var.B}} = -\iint_{\mathcal{S}} \frac{\partial\mathbf{B}(\hat{\mathbf{x}},t)}{\partial t}\cdot d^2\textbf{area} \tag{26}$$

is the extra EMF due to time-dependence of the magnetic field.

Physically, this extra EMF stems from the *electric field induced by the $\partial\mathbf{B}/\partial t$.* Unlike the electrostatic field, the induced electric field is non-conservative (or rather, the force $\mathbf{F} = q\mathbf{E}_{\text{induced}}$ is non-conservative), so it has non-zero line integrals along closed loops,

$$\oint \mathbf{E}_{\text{static}}\cdot d\vec{\ell} = 0 \quad \text{but} \quad \oint \mathbf{E}_{\text{induced}}\cdot d\vec{\ell} \neq 0. \tag{27}$$

Indeed, it's the non-zero work of the induced electric field pushing the electrons around some wire loop $\mathcal{L}$ which provides the EMF in a time-dependent magnetic field, thus

$$\oint\limits_{\mathcal{L}} \mathbf{E}_{\text{induced}} \cdot d\vec{\ell} \;=\; \mathcal{E}_{\text{var.B}} \;=\; -\iint\limits_{\mathcal{S}} \frac{\partial \mathbf{B}(\hat{\mathbf{x}},t)}{\partial t} \cdot d^2\mathbf{area}\,. \tag{28}$$

This is the global form of the Ampere-like law for the induced electric field. The local form of this Law follows from the Stokes' theorem:

$$-\iint\limits_{\mathcal{S}} \frac{\partial \mathbf{B}(\hat{\mathbf{x}},t)}{\partial t} \cdot d^2\mathbf{area} \;=\; \oint\limits_{\mathcal{L}} \mathbf{E}_{\text{induced}} \cdot d\vec{\ell} \;=\; \iint\limits_{\mathcal{S}} (\nabla \times \mathbf{E}_{\text{induced}}) \cdot d^2\mathbf{area}, \tag{29}$$

hence the *Induction Law* for the fields:

$$\nabla \times \mathbf{E}_{\text{induced}} \;=\; -\frac{\partial \mathbf{B}}{\partial t}\,. \tag{30}$$

In general, when both the charges and the currents — and hence also the magnetic field — are time-dependent, there is no clear separation between the static and the induced electric fields. Instead, there is only the combined electric field

$$\mathbf{E}(\hat{\mathbf{x}},t) \;=\; \mathbf{E}_{\text{static}}(\hat{\mathbf{x}},t) \;+\; \mathbf{E}_{\text{induced}}(\hat{\mathbf{x}},t) \tag{31}$$

which obeys the Induction Law

$$\nabla \times \mathbf{E}(\hat{\mathbf{x}},t) \;=\; -\frac{\partial}{\partial t}\mathbf{B}(\hat{\mathbf{x}},t) \tag{32}$$

as well as the Gauss Law

$$\epsilon_0 \nabla \cdot \mathbf{E}(\hat{\mathbf{x}},t) \;=\; \begin{cases} \rho(\hat{\mathbf{x}},t) & \text{(microscopic), or} \\ \rho_{\text{free}}(\hat{\mathbf{x}},t) \;-\; \nabla \cdot \mathbf{P}(\hat{\mathbf{x}},t) & \text{(macroscopic).} \end{cases} \tag{33}$$

Note: The induction law formula (32) works exactly as written for both microscopic fields fields and for the macroscopic fields.

THE POTENTIALS

Since the dynamical electric field has non-zero curl (32), it is cannot be written as (minus) gradient of some potential, $\mathbf{E} \neq -\nabla\phi$. Instead, we can write it as

$$\mathbf{E}(\hat{\mathbf{x}}, t) = -\nabla\phi(\hat{\mathbf{x}}, t) - \frac{\partial}{\partial t}\mathbf{A}(\hat{\mathbf{x}}, t) \tag{34}$$

where $\mathbf{A}(\hat{\mathbf{x}}, t)$ is the vector potential for the magnetic field,

$$\mathbf{B}(\hat{\mathbf{x}}, t) = \nabla \times \mathbf{A}(\hat{\mathbf{x}}, t). \tag{35}$$

To see how this works, let's start with the time-dependent vector potential $\mathbf{A}(\hat{\mathbf{x}}, t)$. Since nobody have ever seen a magnetic monopole, the magnetic Gauss law calls for zero divergence of the magnetic field, $\nabla \cdot \mathbf{B} \equiv 0$, everywhere and everywhen, regardless if the magnetic field is static or time-dependent. Consequently, for each moment of time t, the general solution of the zero-divergence equation is the curl of some vector potential $\mathbf{A}(\hat{\mathbf{x}})$. Combining such solutions for all times t, we get a time-dependent vector potential $\mathbf{A}(\hat{\mathbf{x}}, t)$ whose curl at any time t is the magnetic field $\mathbf{A}(x, t)$ at that time, thus eq. (35).

Next, given some vector potential $\mathbf{A}(, \hat{\mathbf{x}}, t)$ for the magnetic field, let's plug eq. (35) into the Induction Law (32) for the electric field:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} = -\frac{\partial}{\partial t}\big(\nabla \times \mathbf{A}\big), \tag{36}$$

and since the space derivatives commute with the time derivative,

$$\nabla \times \mathbf{E} = -\nabla \times \left(\frac{\partial \mathbf{A}}{\partial t}\right). \tag{37}$$

Consequently, the combination

$$\mathbf{E}(\hat{\mathbf{x}}, t) + \frac{\partial \mathbf{A}(\hat{\mathbf{x}}, t)}{\partial t} \tag{38}$$

has zero curl,

$$\nabla \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}\right) = 0, \qquad \text{for all } \hat{\mathbf{x}} \text{ and all } t. \tag{39}$$

Therefore, it is this curl-less combination rather than the electric field itself which should be (minus) gradient of a scalar potential $\phi(\hat{\mathbf{x}})$, or rather $\phi(\hat{\mathbf{x}}, t)$ to allow for the time-dependent fields. Thus,

$$\mathbf{E}(\hat{\mathbf{x}},t) \;+\; \frac{\partial\mathbf{A}}{\partial t} \;=\; -\nabla\phi(\hat{\mathbf{x}},t) \tag{40}$$

and hence

$$\mathbf{E}(\hat{\mathbf{x}},t) \;=\; -\nabla\phi(\hat{\mathbf{x}},t) \;-\; \frac{\partial}{\partial t}\,\mathbf{A}(\hat{\mathbf{x}},t). \tag{34}$$

Similar to the static case, the vector potential $\mathbf{A}(\hat{\mathbf{x}},t)$ and the scalar potential $\phi(\hat{\mathbf{x}},t)$ are not unique. Statically, the vector potential for a given magnetic field is determined up to a gauge transform

$$\mathbf{A}'(\hat{\mathbf{x}}) \;=\; \mathbf{A}(\hat{\mathbf{x}}) \;+\; \nabla\Lambda(\hat{\mathbf{x}}) \quad\Longrightarrow\quad \nabla\times\mathbf{A}'(\hat{\mathbf{x}}) \;=\; \nabla\times\mathbf{A}(\hat{\mathbf{x}}), \tag{41}$$

for an arbitrary $\Lambda(\hat{\mathbf{x}})$. For the time-dependent magnetic fields, we may also use time-dependent gauge transforms with arbitrary $\Lambda(\hat{\mathbf{x}},t)$. However, in order to preserve the electric field (34) as well as the magnetic field, a time-dependent gauge transform also shifts the scalar potential by a time derivative of Λ,

$$\begin{aligned}
\mathbf{A}'(\hat{\mathbf{x}},t) \;&=\; \mathbf{A}(\hat{\mathbf{x}},t) \;+\; \nabla\Lambda(\hat{\mathbf{x}},t),\\
\phi'(\hat{\mathbf{x}},t) \;&=\; \phi(\hat{\mathbf{x}},t) \;-\; \frac{\partial\Lambda(\hat{\mathbf{x}},t)}{\partial t}\,,\\
\mathbf{B}'(\hat{\mathbf{x}},t) \;&=\; \mathbf{B}(\hat{\mathbf{x}},t),\\
\mathbf{E}'(\hat{\mathbf{x}},t) \;&=\; \mathbf{E}(\hat{\mathbf{x}},t).
\end{aligned} \tag{42}$$

Indeed, for the electric field we have

$$\mathbf{E}' \;=\; -\nabla\phi' \;-\; \frac{\partial\mathbf{A}'}{\partial t} \;=\; -\nabla\phi \;+\; \cancel{\nabla\frac{\partial\Lambda}{\partial t}} \;-\; \frac{\partial\mathbf{A}}{\partial t} \;-\; \cancel{\frac{\partial}{\partial t}(\nabla\Lambda)} \;=\; -\nabla\phi \;-\; \frac{\partial\mathbf{A}}{\partial t} \;=\; \mathbf{E}. \tag{43}$$

Note that all *physical* quantities are invariant under such gauge transforms. For some convoluted calculations, this is a powerful cross-check of the final results: if they are not gauge invariant, you must have made a mistake somewhere!

In other situations, it's convenient to eliminate the potentials' redundancy by fixing a gauge condition, that is, imposing an extra linear condition on the $\mathbf{A}$ and ϕ (one equation for each (x_1,x_2,x_3,t)) to make the potentials unique. A commonly used condition is the transverse gauge $\nabla \cdot \mathbf{A}(\hat{\mathbf{x}}, t) \equiv 0$, also called the Coulomb gauge because in this gauge the scalar potential $\phi(\hat{\mathbf{x}}, t)$ is simply the Coulomb potential due to $\rho(\hat{\mathbf{y}}, t)$ at the same time t. Indeed, in the transverse gauge

$$\nabla \cdot \mathbf{E} \;=\; -\nabla^2 \phi \;-\; \frac{\partial}{\partial t}(\nabla \cdot \mathbf{A}) \;=\; -\nabla^2 \phi, \tag{44}$$

hence by the Gauss Law

$$\nabla^2 \phi(\hat{\mathbf{x}}, t) \;=\; -\frac{1}{\epsilon_0}\, \rho(\hat{\mathbf{x}}, t), \tag{45}$$

which is a differential equation in space but not in time, so its solution is the *instantaneous* Coulomb potential

$$\phi(\hat{\mathbf{x}}, t) \;=\; \frac{1}{4\pi\epsilon_0} \iiint \frac{\rho(\hat{\mathbf{y}}, t)}{|\hat{\mathbf{x}} - \hat{\mathbf{y}}|}\, d^3\hat{\mathbf{y}}. \tag{46}$$

Another common condition is the Lorentz-invariant Landau gauge

$$\nabla \cdot \mathbf{A} \;+\; \mu_0 \epsilon_0 \frac{\partial \phi}{\partial t} \;=\; 0 \tag{47}$$

in which both the scalar and the vector potentials obey similar wave equations,

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \;-\; \nabla^2 \right) \phi(\hat{\mathbf{x}}, t) \;=\; \frac{1}{\epsilon_0}\, \rho(\hat{\mathbf{x}}, t), \tag{48}$$

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \;-\; \nabla^2 \right) \mathbf{A}(\hat{\mathbf{x}}, t) \;=\; \mu_0\, \mathbf{J}(\hat{\mathbf{x}}, t), \tag{49}$$

$$\frac{1}{c^2} \;=\; \mu_0 \epsilon_0 \,. \tag{50}$$

I shall derive these equations later in class; for the impatient, here are my notes on Maxwell equations and related issues.