Torsten Bringmann (`torsten.bringmann@fys.uio.no`)

FYS4160: Preliminaries

While general relativity (GR) is much more self-contained than, for example, quantum field theory, a certain amount of background knowledge will help significantly to master the large amount of new concepts. These notes summarize some of the things that you should *know and feel comfortable with* **before** starting this course, be it from previous courses or from own reading. If you feel uncomfortable with any of the below concepts or terms, you should as soon as possible review a suitable book or your notes of previous courses.

This write-up also fixes our notation and conventions (we will use the same as in the book by Carroll). In particular, we will throughout adopt 'natural units' where $\hbar = c = 1$ (though these constants are kept at some places for pedagogic reasons).

1 Special relativity

The time t and the position vector $\mathbf{x}$ (on the blackboard we will use $\vec{x}$) can be combined into a four-vector in Minkowski space $x^\mu = (t, \mathbf{x}) = (x^0, x^1, x^2, x^3)$. We use small greek letters to denote space-time indices and small roman letters to denote spatial indices of such four-vectors, i.e. $\mu, \nu, ... = 0, 1, 2, 3$ and $i, j, ... = 1, 2, 3$.[1] Throughout the course, the **summation convention** is used, in which repeated indices are always summed over. This implies that any index that appears twice in an expression is a dummy index (i.e. its name has no importance: $x^\mu x_\mu = x^\rho x_\rho = x^0 x_0 + x^1 x_1 + x^2 x_2 + x^3 x_3$) and that no index can appear more than twice.[2] For spacetime-indices, it is furthermore very important to keep track of whether they are up or down; in particular, you should never encounter expressions like $a^\mu b^\mu$ (see also below the comments about scalar products).

In Minkowski space, the **metric** tensor g (aka $g_{\mu\nu}$, see footnote 1) is given by

$$g_{\mu\nu} = \eta_{\mu\nu} \equiv \begin{pmatrix} -1 & & & \\ & +1 & & \\ & & +1 & \\ & & & +1 \end{pmatrix} = \left(g^{-1} \right)_{\mu\nu} \equiv g^{\mu\nu}. \tag{1}$$

This is the most often used convention in GR, while in (both quantum and classical) field theory the signature $(+, -, -, -)$ is more common. We use the metric $g_{\mu\nu}$ to lower indices, and the inverse metric $g^{\mu\nu}$ to raise indices. For example, $x_\mu \equiv g_{\mu\nu} x^\nu = (-t, \mathbf{x})$ and $x^\mu = g^{\mu\nu} x_\nu$. Note that while the displacement vector $x^\mu = (t, \mathbf{x})$ is 'naturally raised', the derivative vector

---

[1] For Lorentz four-vectors we sometimes follow the standard, if somewhat confusing, convention of writing A^μ or A_μ instead of 'A'; we thus use the *same* notation for a vector and its components, and it is only the *context* that determines the correct interpretation. (This convention allows a simple distinction between covariant ('lower indices') and contravariant ('upper indices') vectors – though these concepts will be introduced in much more detail in the course). A Euclidian three-vector, in contrast, is always denoted as $\mathbf{A}$ (or $\vec{A}$), while its components are denoted with A^i.

[2] In rare occasions, we have expressions where a double index is *not* summed over. If there is a risk for confusion, this will be indicated by an underscore. For example, $x^{\underline{\mu}} x_{\underline{\mu}} = x^0 x_0$ for $\mu = 0$, and $x^{\underline{\mu}} x_{\underline{\mu}} = x^1 x_1$ for $\mu = 1$ etc.

is naturally lowered:

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial t}, \nabla\right), \quad \text{so} \quad \partial^\mu = \left(-\frac{\partial}{\partial t}, \nabla\right). \tag{2}$$

A *Lorentz* **four-vector** in general is any object that transforms under a Lorentz transformation Λ, described by a 4×4 matrix $\Lambda^\mu{}_\nu$, according to

$$A^\mu \to A'^\mu = \Lambda^\mu{}_\nu A^\nu \quad (\text{and thus } A_\mu \to A'_\mu = \Lambda_\mu{}^\nu A_\nu). \tag{3}$$

A **Lorentz transformation** is by definition any linear transformation that leaves the 'length' of a vector in Minkowski space invariant:[3]

$$\eta_{\mu\nu} x^\mu x^\nu \overset{!}{=} \eta_{\mu\nu} x'^\mu x'^\nu = \eta_{\mu\nu} \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma x^\rho x^\sigma, \tag{4}$$

or equivalently, using matrix rather than index notation

$$x^T \cdot \eta \cdot x \overset{!}{=} x'^T \cdot \eta \cdot x' = (\Lambda \cdot x)^T \cdot \eta \cdot \Lambda \cdot x. \tag{5}$$

(More generally, a Lorentz **tensor** is an object with several indices, $T^{\mu_1\mu_2\cdots}_{\nu_1\nu_2\cdots}$, that all transform as above.) This implies that *any* scalar product involving Lorentz four-vectors is invariant (i.e. a 'scalar') under Lorentz transformations:

$$A \cdot B \equiv A^\mu B_\mu = A_\mu B^\mu = A^0 B^0 - \mathbf{A} \cdot \mathbf{B} = const. \tag{6}$$

Note that a scalar product always involves a contravariant and a covariant vector. Lorentz transformations contain standard 3D rotations as well as Lorentz boosts.

Example: For a boost along the x-axis, Λ takes the form

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{7}$$

with $\beta \equiv v/c$ denoting the relative velocity of the two frames and $\gamma \equiv (1 - \beta^2)^{-1/2}$. Using Eq. (3), this implies the familiar transformation $x \to x'$ with

$$t \quad \to \quad t' = \gamma(t - \beta x) \tag{8}$$
$$x \quad \to \quad x' = \gamma(x - \beta t) \tag{9}$$
$$y \quad \to \quad y' = y \tag{10}$$
$$z \quad \to \quad z' = z \tag{11}$$

The most important four-vector that we will encounter, besides the position vector, is the **four-momentum**. It satisfies $p^2 \equiv p \cdot p = m^2$, where m is the invariant mass of the particle. Writing the four-momentum in terms of the energy E and the 3-momentum $\mathbf{p}$,

$$p^\mu = (E, \mathbf{p}), \tag{12}$$

we thus directly get the famous relativistic dispersion relation for a massive particle:

$$E^2 - \mathbf{p}^2 = m^2. \tag{13}$$

---

[3]Note that Λ itself is a constant matrix, independent of space-time coordinates. In a certain sense, GR consists in 'nothing but' replacing $\Lambda \to \Lambda(x^\mu)$, i.e. in making the Lorentz symmetry local.

2   Classical dynamics: Lagrange & Hamilton

The state of a classical physical system at a given time t can be fully characterized by a set of *generalized coordinates* q_a, with $a = 1, ..., N$ (for an unconstrained system, N equals three times the number of particles), and their associated *velocities* $\dot{q}_a \equiv dq_a/dt$. The *principle of least action* requires that the true path $q(t)$ taken by the system (the physical trajectory) is the one that extremizes the *action*, i.e. the functional

$$S[q] \equiv \int_{t_1}^{t_2} L(q_a, \dot{q}_a)\, dt\,. \tag{14}$$

The whole dynamics is thus governed by the *Lagrangian* $L = L(q_a, \dot{q}_a)$.[4] Demanding $\delta S = 0$, in particular, leads to the **Euler-Lagrange equations** of motion:

$$\frac{dL}{dq_a} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_a} = 0\,. \tag{15}$$

Example. For a single particle in a potential $V(\mathbf{x})$, we have $L = \frac{1}{2}m\dot{\mathbf{x}}^2 - V$ in Cartesian coordinates. Eq. (15) then reproduces the familiar $m\ddot{\mathbf{x}} = -\nabla V$: with $\mathbf{F} = -\nabla V$, this is simply Newton's 2nd law.

For each generalized coordinate q_a, the **canonically conjugated momentum** p_a is defined as

$$p_a \equiv \frac{\partial L(q, \dot{q})}{\partial q_a}\,. \tag{16}$$

A Legendre transformation of the Lagrangian then leads to the *Hamiltonian H*, which no longer depends on the velocities:[5]

$$H(p, q) \equiv p^a \dot{q}_a(p, q) - L[q, \dot{q}(p, q)]\,. \tag{17}$$

In this formulation, the system is described by the **Hamilton equations** of motion,

$$\dot{q}_i = \frac{\partial H}{\partial p_i}\,, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}\,, \tag{18}$$

i.e. $2N$ coupled first-order ODEs instead of the N second-order ODEs in Eq. (15).

Example. Considering again a single particle in a potential, we have $H = \mathbf{p}\cdot\dot{\mathbf{x}} - L = \frac{1}{2m}\mathbf{p}^2 + V$. The Hamilton equations then give as expected $\dot{\mathbf{q}} = \mathbf{p}/m$ and $\dot{\mathbf{p}} = -\nabla V$.

The (classical) theory of **fields** $\Phi(x)$ can be described as the limit $N \to \infty$ of classical mechanics. The field values *at each space-time point* then take the role of the generalized coordinates, $q_a \to \Phi(x^\mu)$, and all the physics is encoded in a *Lagrangian density* $\mathcal{L} = \mathcal{L}(\Phi, \partial_\mu \Phi)$ (which however often is simply referred to as 'Langrangian'). Extremizing the action $S = \int d^4x\, \mathcal{L}$, by requiring $\delta S = 0$, then again leads to Euler- Lagrange equations – which now are manifestly relativistically invariant:[6]

$$\frac{\partial \mathcal{L}}{\partial \Phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi)} = 0\,. \tag{19}$$

---

[4]For *open* systems, i.e. if an external force is present, L can also have an explicit time-dependence.

[5]NB: this requires that Eq. (16) can be solved for $\dot{q}_a$. Normally, this implies $\det \partial L/\partial \dot{q}_a \partial q_b \neq 0$.

[6]Note that Φ can also carry spacetime or internal indices. For a vector field A, e.g., simply replace $\Phi \to A^\nu$ in this equation.

In analogy to the finite N case, we can define a *canonical momentum density* $\pi(x) \equiv \partial\mathcal{L}/\partial\dot{\Phi}$, and a *Hamiltonian density* $\mathcal{H} \equiv \pi\dot{\Phi} - \mathcal{L}$.

Example. The simplest example is a real scalar field $\phi(x)$ with a potential $V(\phi)$, which is described by

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - V(\phi)\,. \tag{20}$$

The canonical momentum density is then $\pi(x) = \dot{\Phi}(x)$ and the Hamiltonian density $\mathcal{H} = \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\Phi)^2 + V(\Phi)$. The equation of motion is known as the *Klein-Gordon equation*, $\partial_\mu\partial^\mu\phi + dV/d\phi \equiv \Box\phi + V' = 0$.

3   Newtonian gravity

In Newtonian physics, the force exerted by a point mass M on a test mass m is

$$\mathbf{F} = G_N\frac{mM}{r^3}\mathbf{r}\,, \tag{21}$$

where $\mathbf{r}$ is the relative distance vector between the two masses and G_N is Newton's constant. More generally, we can express the force experienced by a test mass m in the presence of a *gravitational field* Φ as

$$\mathbf{F} = -m\nabla\Phi\,. \tag{22}$$

Assuming that the mass m here is the same as in Newton's 2nd law (see above), the gravitational acceleration of a test body is thus independent of its mass, $\ddot{\mathbf{r}} = -\nabla\Phi$. For a point mass, the gravitational field obeys $\nabla\Phi = -G_N M\mathbf{r}/r^3$, while for an arbitrary mass distribution with density ρ it satisfies the *Poisson equation*:

$$\nabla^2\phi(\mathbf{x}) = -\rho(\mathbf{x})\,. \tag{23}$$

The *Kepler problem* describes a system of two point masses m_1 and m_2. In principle, from Eq. (15), this leads to a system of 2×3 coupled equations of motion. Using conservation of energy E and angular momentum $\mathbf{L}$, however, the problem can be reduced to a single differential equation that describes an effective particle in 1D:

$$E = \frac{\mu}{2}\left(\frac{dr}{dt}\right)^2 + U_{\text{eff}}\,. \tag{24}$$

Here, $\mu \equiv m_1 m_2/(m_1 + m_2)$ is the so-called *reduced mass* of this particle and $U_{\text{eff}} = L^2/(2\mu r^2) - Gm_1 m_2/r$ its potential. This equation can be directly integrated, via separation of variables, to give $r(t)$. The orbit equation for the azimuthal angle, $\phi(r)$, can then be obtained from $d\phi/dr = (d\phi/dt)(dt/dr)$, where $L = \mu r^2 d\phi/dt$ is the (constant!) angular momentum.

4 Mathematical tools

4.1 Basics of linear algebra

In GR, our main mathematical tools will be tensors and tensor fields. Since tensors are nothing but multilinear maps, ideas of linear algebra come into play and prove useful to complement geometric interpretations. A **vector space** over an abstract *field* in mathematics is defined as a set T of objects ('vectors'), together with two operations (the sum of two vectors, as well as the multiplication of a vector with [the abstract version of] a number). In this course, we are only interested in vector spaces over real numbers. Rather than listing the various axioms, the most intuitive way of understanding a vector space is to say that any linear combination of its elements produces again a vector. In mathematical language:

$$\forall \, \mathbf{V}, \mathbf{W} \in T, \forall \, a, b \in \mathbb{R} : (a + b)(\mathbf{V} + \mathbf{W}) = a\mathbf{V} + a\mathbf{W} + b\mathbf{V} + b\mathbf{W} \in T \,. \tag{25}$$

A set of vectors $\{\mathbf{V}_{(i)}\} \subset T$ is said to be **linearly independent** if it is not possible to find any set of real numbers $\{a_i\}$ such that $\sum_i a_i \mathbf{V}_{(i)} = \mathbf{0}$ (apart from *all* a_i being zero). The **dimension** of a vector space is the maximal number of linearly independent vectors $\mathbf{V}_i$ one can have. **Basis** vectors $\{\mathbf{e}_{(i)}\}$ are such a (not uniquely defined!) maximal set of linearly independent vectors, with the additional property that $\left|\mathbf{e}_{(i)}\right|^2 = \mathbf{e}_{(i)} \cdot \mathbf{e}_{(i)} = 1$ (but in general *not* $\mathbf{e}_{(i)} \cdot \mathbf{e}_{(j)} = \delta_{ij}$). From the above, it follows that every vector $\mathbf{V}$ can be decomposed in terms of basis vectors,

$$\mathbf{V} = V^i \mathbf{e}_{(i)} \,, \tag{26}$$

where the $V^i \in \mathbb{R}$ are referred to as the vector **components** with respect to the specific basis choice $\{\mathbf{e}_{(i)}\}$.

Note that *in physics* the specific vector spaces one is interested in are in addition always characterized by the transformation property of their elements under the action of some (typically Lie) group. As explained above, a 4-vector in special relativity for example is defined by its transformation properties under Lorentz transformations, which in turn are defined by all those linear transformations that leave the norm of the 4-vector invariant. Likewise, 3-D vectors $\mathbf{A}$ are those that transform as $A_i \rightarrow R_{ij} A_j$ under rotations, which leaves $A_i A_i$ invariant.[7]

Any vector space T has a corresponding **dual vector space** T^* of same dimension consisting of linear functionals (often called **one-forms**) $\varphi : T \longrightarrow \mathbb{R}$. Vector space operations in T^* are defined by the respective operations in T:

$$\forall \varphi, \psi \in T^*, \mathbf{V} \in T, a \in \mathbb{R} : (\varphi + \psi)(\mathbf{V}) = \varphi(\mathbf{V}) + \psi(\mathbf{V}) \tag{27}$$

$$(a\varphi)(\mathbf{V}) = a(\varphi(\mathbf{V})) \tag{28}$$

Given a basis $\{\mathbf{e}_{(i)}\}$ of T, there exists a dual basis $\{\mathbf{e}^*_{(i)}\}$ such that $\mathbf{e}^*_{(i)}(\mathbf{e}_{(j)}) = \delta_{ij}$ and we can think of vectors as being linear scalar functions of one-forms via $\mathbf{V}(\varphi) := \varphi(\mathbf{V})$.

---

[7]There is an important conceptual difference between **active and passive transformations** here. In both cases, the *components* transform as indicated and the norm of the vector is conserved. For an *active* transformation, the basis vectors do not change, such that a 3-vector $\mathbf{A}$ is e.g. *actually rotated* in $\mathbb{R}^3$. For *passive* transformations, on the other hand, the vector $\mathbf{A} = A^i \mathbf{e}_{(i)}$ remains *unchanged*; this requires that the basis vectors transform with the inverse transformation operation, e.g. $\mathbf{e}_{(i)} \rightarrow \left(R^{-1}\right)_{ij} \mathbf{e}_{(j)} = R_{ji} \mathbf{e}_{(j)}$ for rotations. It is always the latter situation that is described by coordinate transformations.

Example. If we take T to be $\mathbb{R}^n$, that is the space of columns of n real numbers, then the dual space $\mathbb{R}^{n*}$ corresponds to the space of rows of n real numbers – acting by ordinary matrix multiplication on the elements of T. Note the conceptual difference to an inner (or 'scalar' vector) product , which in this case is *operationally* the same but defined as a map $\cdot : T \times T \longrightarrow \mathbb{R}$.

4.2 Analysis in $\mathbb{R}^n$

This course will provide an introduction to differential geometry, the mathematical basis for the description of GR. For this, we will require basic notions of topology and analysis in $\mathbb{R}^n$. A subset $U \subset \mathbb{R}^n$ is **open** if every point $x \in U \subset \mathbb{R}^n$ has a neighborhood contained in the subset, i.e. if for every $x \in U$ there exists a number $\varepsilon > 0$ such that every point $y \in \mathbb{R}^n$ with Euclidean distance smaller than ε belongs to U as well.
Given two sets $U, V \subset \mathbb{R}^n$, a map $\phi : U \longrightarrow V$ is called **bijective** (or invertible) if each element of V has at most one element of U mapped into it (i.e. ϕ is **injective**) and each element of V has at least one element of U mapped into it (ϕ is **surjective**). For a bijective map ϕ, we can define the inverse function $\phi^{-1} : V \longrightarrow U$ by $(\phi^{-1} \circ \phi)(x) = x$ for all $x \in V$. To define differentiability of an arbitrary map $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, we look at the component functions $f^i : \mathbb{R}^n \longrightarrow \mathbb{R}$, $f = (f^1, ..., f^m)$, and say that f is smooth if each component function f^i, $i = 1, ..., m$, is continuous and smooth (i.e. can be differentiated as many times as you like). We use bijective smooth maps to compare two Euclidean spaces and to decide whether they are "the same": Two sets U and V are **diffeomorphic** if there exists a smooth bijective map $\phi : U \longrightarrow V$ with smooth inverse $\phi^{-1} : V \longrightarrow U$.

4.3 Dirac delta function

The Dirac delta 'function' $\delta(x)$ is not actually a function but a distribution, and is defined by $\delta(x) = 0$ for all $x \neq 0$, and $\int dx\, \delta(x) = 1$. Useful relations include

$$\int dx\, f(x)\, \delta(x - y) = f(y), \tag{29}$$

$$\int dx\, f(x)\, \delta'(x) = -f'(0), \tag{30}$$

$$\int dx\, e^{ikx} = 2\pi\, \delta(k), \tag{31}$$

$$\delta(f(x)) = \sum_i \frac{\delta(x - x_i)}{|f'(x_i)|}, \tag{32}$$

where the second equation can be understood from integration by parts, and the last one only holds if the function $f(x)$ only has 1st order zeros (at $x = x_i$). Analogous definitions and properties of the delta function apply in higher dimensions.

4.4 Fourier transforms

We will sometimes change between position and momentum space, with the following convention for the Fourier transforms relating those two:

$$f(x) = \int \frac{d^4k}{(2\pi)^4} e^{-ik\cdot x} \tilde{f}(k), \tag{33}$$

$$\tilde{f}(k) = \int d^4x \, e^{ik\cdot x} f(x). \tag{34}$$

Factors of 2π will thus always appear in the momentum integrals; for 3D integrals we need to replace $k \cdot x \to \mathbf{k} \cdot \mathbf{x}$. Note that the $\sim$ superscript is often omitted if the argument is given explicitly, e.g. $\tilde{f}(\mathbf{k}) \equiv f(\mathbf{k}) \equiv f_{\mathbf{k}}$. Fourier transforms are an important tool to solve linear differential equations.

Example. Let us solve Poisson's equation (23). Fourier-transforming both sides gives

$$-\int \frac{d^3k}{(2\pi)^3} e^{-i\mathbf{k}\cdot\mathbf{x}} \rho(\mathbf{k}) = \nabla^2 \int \frac{d^3k}{(2\pi)^3} e^{-i\mathbf{k}\cdot\mathbf{x}} \phi(\mathbf{k}) = \int \frac{d^3k}{(2\pi)^3} (-\mathbf{k}^2) e^{-i\mathbf{k}\cdot\mathbf{x}} \phi(\mathbf{k}), \tag{35}$$

from which we can deduce $\rho_{\mathbf{k}} = \mathbf{k}^2 \phi_{\mathbf{k}}$ (alternatively, we could have used the replacement rule $\partial_j \to -ik_j$ directly). Performing the inverse Fourier transform then gives

$$\phi(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3} e^{-i\mathbf{k}\cdot\mathbf{x}} \phi_{\mathbf{k}} = \int \frac{d^3k}{(2\pi)^3} e^{-i\mathbf{k}\cdot\mathbf{x}} \mathbf{k}^{-2} \int d^3y \, e^{i\mathbf{k}\cdot\mathbf{y}} \rho(\mathbf{y}) \tag{36}$$

$$= \frac{1}{(2\pi)^3} \int d^3y \, \rho(\mathbf{y}) \int_0^{2\pi} d\varphi \int_{-1}^1 d\cos\theta \int_0^\infty d|\mathbf{k}| \, e^{i\cos\theta|\mathbf{k}||\mathbf{x}-\mathbf{y}|} \tag{37}$$

$$= \frac{1}{8\pi^2} \int d^3y \, \rho(\mathbf{y}) \int_{-1}^1 d\cos\theta \underbrace{\int_{-\infty}^\infty d|\mathbf{k}| \, e^{i\cos\theta|\mathbf{k}||\mathbf{x}-\mathbf{y}|}}_{2\pi\,\delta(\cos\theta|\mathbf{x}-\mathbf{y}|)=2\pi\,\delta(\cos\theta)/|\mathbf{x}-\mathbf{y}|} \tag{38}$$

$$= \int d^3y \, \frac{\rho(\mathbf{y})}{4\pi|\mathbf{x}-\mathbf{y}|}. \tag{39}$$

4.5 Green's functions

Green's 'functions' (which often also rather are distributions), are an important tool to solve differential equations of the form

$$L_x f(x) = g(x). \tag{40}$$

Here, L_x is a linear, differential operator in x, f the function to be determined and g an arbitrary function that constitutes the inhomogeneous part of the ODE. A Green's function G of L_x is by definition any – often not uniquely determined – solution of

$$L_x G(x, y) = -\delta(x - y). \tag{41}$$

A Green's function is thus, in a loose sense, the *inverse* of L_x. If L_x is translation invariant (i.e. it does not change for $x \to x + x_0$, as is the case in most physics' applications), the Green's function can in fact always be written as

$$G(x, y) = G(x - y). \tag{42}$$

If G is known, then a solution of Eq. (40) for (almost) arbitrary $g(x)$ is given by

$$f(x) = -\int dy\, G(x - y)g(y), \tag{43}$$

which directly follows from the definitions of the Green's and Dirac delta functions. NB: With $f(x)$ being a solution to Eq. (40), $f(x) + \sum c_i f_{i,\text{hom}}(x)$ is also a solution (where $f_{i,\text{hom}}$ represent the homogenous solutions, i.e. those for which $L_x f_{i,\text{hom}} = 0$)!

Example. Consider the d'Alembert operator, $L_x = \partial_\mu \partial^\mu \equiv \Box$. The homogeneous equation, $L_x f_{\text{hom}} = 0$, is the relativistic wave equation for massless fields and the solutions are plane waves:

$$f_{\text{hom}} = c_1 e^{-ik_\mu x^\mu} + c_2 e^{ik_\mu x^\mu}, \tag{44}$$

with $k^\mu = (\omega, \mathbf{k})$ and $\omega = |\mathbf{k}|$.[8]

The Green's function of the d'Alembert operator can be found by taking the Fourier transform of Eq. (41):

$$-\int \frac{d^4k}{(2\pi)^4}\, e^{ik\cdot(x-y)} = \Box \int \frac{d^4k}{(2\pi)^4}\, e^{ik\cdot(x-y)}\, \tilde{G}(k) = -\int \frac{d^4k}{(2\pi)^4}\, e^{ik\cdot(x-y)}\, k_\mu k^\mu\, \tilde{G}(k). \tag{45}$$

Hence, $\tilde{G}(k) = 1/k^2$ and therefore $G(x - y) = \int \frac{d^4k}{(2\pi)^4}\, e^{ik\cdot(x-y)}/k^2$. The integration over k^0 has poles at $k^0 = \pm|\mathbf{k}|$ and by going to the complex plane (i.e. by making k^0 complex) there are four possible ways of how to avoid the poles and evaluate the integral using the tools of complex analysis; correspondingly, there are four independent Green's functions. One of those is the *retarded Green's function*

$$G_r(x - y) = -\frac{1}{4\pi\, |\mathbf{x} - \mathbf{y}|} \delta\left(|\mathbf{x} - \mathbf{y}| - (x^0 - y^0)\right) \Theta(x^0 - y^0). \tag{46}$$

For an arbitrary source distribution $g(t, \mathbf{x})$, Eq. (43) then recovers the retarded solutions familiar from classical electrodynamics:

$$f(t, \mathbf{x}) = \frac{1}{4\pi} \int d^3y\, \frac{g(t_r, \mathbf{y})}{|\mathbf{x} - \mathbf{y}|}, \tag{47}$$

where $t_r = t - |\mathbf{x} - \mathbf{y}|$ is known as the retarded time (recall that $c = 1$).

---

[8] If the function that L_x acts on is a vector, e.g. the electric field $\mathbf{E}$, the constant coefficients c_1 and c_2 must also be vectors – those are referred to as *polarization vectors*. If the function is in addition divergence-free, those polarization vectors must be orthogonal to the momentum; in vacuum, e.g., we have $\nabla \cdot \mathbf{E} = 0$ and therefore $c_1 \cdot \mathbf{k} = c_2 \cdot \mathbf{k} = 0$.