Assumptions:

- The rods and clocks are conceptual: they do not affect physical experiments.
- All rods and clocks are identical (when brought together, the rods have the same, time-independent length and the clocks tick with the same rate).
- The lattice is "infinitely dense": there is a clock at every point in space.
- Each clock is assigned a unique position label $\vec{x}$ and the reference frame label $\mathcal{O}$.

   For example, a unique position label $\vec{x}$ for a clock can be obtained by counting the rods in x-, y- and z-direction that one has to traverse to reach the clock from the origin. The origin O is, by definition, a "special" clock that is *assigned* the position label $\vec{x}_O = \vec{0}$.

¡! Observers are *not* sitting at the origin, looking at their wristwatch, and observing the events with binoculars! They are simply collecting and processing the data that is accumulated by the contraption we call a reference frame.

Since we assume that (ideally) there is one clock at every point in space:

→ For every observer $\mathcal{O}$ and every coincidence class E there is a unique event $e_{\mathcal{O}}$

$$E \ni e_{\mathcal{O}} = \langle \text{Clock with frame label } \mathcal{O} \text{ and position label } \vec{x} \text{ shows time } t \rangle \qquad (1.1\text{a})$$
$$=: (t, \vec{x})_{\mathcal{O}} \quad \Leftrightarrow \quad [E]_{\mathcal{O}} = (t, \vec{x}) \qquad (1.1\text{b})$$

for some position label $\vec{x}$ and clock reading t.

We refer to the event $(t, \vec{x})_{\mathcal{O}}$ as the *spacetime coordinates* of E with respect to frame $\mathcal{O}$. A different observer $\mathcal{O}'$ will use its own clocks and therefore other events ("coordinates") $(t', \vec{x}')_{\mathcal{O}'} \in E$ to refer to E.

In the real world, the ↑ *tracking detectors* of particle colliders are reminiscent of this ideal setup: They are comprised of 3D arrangements of semiconductor-based particle detectors that all report to a central computer that then reconstructs the trajectories of scattering products from the combination of all detection events.

3 | ⁂ *Inertial (coordinate) systems:*

The setup of a reference frame $\mathcal{O}$ above is incomplete and actually very hard to work with: Without additional constraints on the geometry of the lattice and the correlations of clocks (their "calibration"), the record of events is essentially arbitrary. Let us therefore impose some deterministic "calibration procedure" (the same for all frames) that determines how to lay out the rod lattice and how to synchronize the clocks. This procedure endows our reference frame with a specific *coordinate system*, a labeling scheme to describe events.

i | Clock calibration: ⁂ *(Poincaré-)Einstein synchronization* **ES**

The conventional synchronization procedure (which is actually in practical use) is (Poincaré-)Einstein synchronization:

$$t_O \overset{?}{=} \tfrac{1}{2} \left(t_A + \tilde{t}_A \right) \qquad (1.2)$$
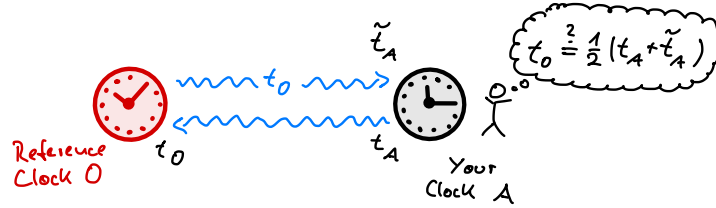
You will study this particular procedure and its properties in ➡ Problemset 1.

In brief, the procedure goes as follows: Consider a reference clock O and some other clock A you wish to synchronize with O.

(1) To do so, you send a light signal from A to O and note the time t_A your clock A reads when the signal is emitted.

(2) When the signal arrives at O, it is immediately reflected back to A together with the reading t_O of clock O at this very moment.

(3) When the signal arrives back at your clock A (together with the timestamp t_O), you note again the reading of your clock as $\tilde{t}_A$.

(4) You are now in the possession of three timestamps: $(t_A, t_O, \tilde{t}_A)$. The idea of Einstein synchronization is to *postulate* the reciprocity of the speed of light: We *declare* that the speed of the signal from A to O is the same as on its way back from O to A (note that we cannot *measure* this reciprocity because we would need already synchronized clocks to do so!). Under this assumption, the readings of *synchronized clocks* must satisfy [9]

$$\Delta t_{A \to O} \equiv t_O - t_A \overset{!}{=} \tilde{t}_A - t_O \equiv \Delta t_{O \to A} \quad \Leftrightarrow \quad t_O \overset{!}{=} \tfrac{1}{2}(t_A + \tilde{t}_A), \qquad (1.3)$$

which you can locally check with your data $(t_A, t_O, \tilde{t}_A)$. Note that you do not need to know the distance from O to A, nor the numerical value of the speed of light c for this procedure to work!

(5) Now if you just powered on your shiny new clock A for the first time, it is very unlikely that the condition Eq. (1.3) will be satisfied:

$$t_O = \tfrac{1}{2}(t_A + \tilde{t}_A) + \delta t = \tfrac{1}{2}[(t_A + \delta t) + (\tilde{t}_A + \delta t)]. \qquad (1.4)$$

Here δt is an offset that you might encounter. But then you can just recalibrate your clock A by δt such that the new readings are $t_A + \delta t$ and $\tilde{t}_A + \delta t$.

Repeating this procedure for all clocks of the frame $\mathcal{O}$ allows you to establish a synchronization relation between arbitrary pairs of clocks. The fact that (under some reasonable and experimentally verified assumptions) the order in which you synchronize your clocks does not matter (the established relation is an *equivalence* relation, ➡ Problemset 1 and Ref. [21]) makes Einstein synchronization a very useful and peculiar *convention* [22–24]. However, one can show that it is the only convention that yields a non-trivial equivalence relation of simultaneity that is consistent with the causal structure on $\mathcal{E}$ (→ *later*) [25].

ii | <u>Lattice calibration:</u> ❄ *Orthonormal Cartesian coordinates* `OC`:

The layout of the lattice of rods assigns coordinates $\vec{x} = (x, y, z)$ to each clock. Depending on the actual shape of the lattice, we will denote events by different position labels. (Note that even with rigid rods connected in the topology of a cubic lattice the geometry is not fixed; for example, you can shear the lattice.) If we assume that space (not spacetime!) is a flat Euclidean space where all the facts of Euclidean geometry hold good (angles of triangles add up π, the Pythagorean theorem holds, the area of circles is πr^2, etc.), we can parametrize it without loss of generality by *orthonormal Cartesian coordinates*. In these coordinates, distances can be calculated by the Pythagorean formula:

Spatial distance between clocks at $\vec{x}$ and $\vec{y}$:

$$\underbrace{d(\vec{x}, \vec{y})}_{\text{“Physics”}} = \underbrace{\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}}_{\text{“Mathematics”}} \qquad (1.5)$$

The fact that the coordinates of a point (x, y, z) are *distances* along paths parallel to the coordinate axes makes the coordinates *Cartesian*. The fact that Eq. (1.5) holds makes them *orthonormal* (i.e., the axes are orthogonal and have the same scale, as suggested by the sketch above). Coordinates are an intrinsically mathematical concept, they are "labels" to identify points on a manifold of physical points (or *events*, if you consider space*time* coordinates). By contrast, distances carry physical significance: You can measure them with light signals or rods. The prevalence of Cartesian coordinates makes it easy to conflate these two concepts (this will become particularly important in GENERAL RELATIVITY).

Here is a way to check whether your lattice satisfies the OC condition using the clocks of $\mathcal{O}$ (and the assumption of the isotropy of the two-way speed of light):
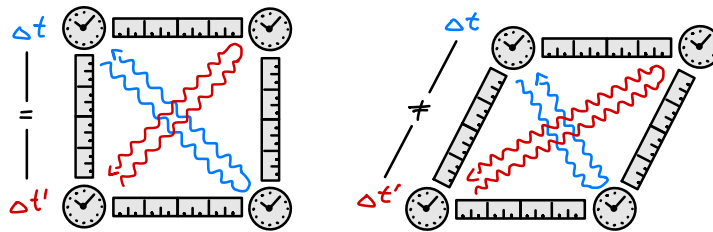
iii | "Inertial Test" (⁂ *law of inertia*):

Once you have arranged your rods and synchronized your clocks and thereby established a Cartesian coordinate system and a (allegedly) well-defined notion of simultaneity, you can perform the following test and check whether your particular reference frame $\mathcal{O}$ passes it or not:

IN Free particles move at constant velocity and in straight lines.
(⁂ *Homogeneity of Inertia*)

- It is implied that this statement is true *everywhere*, *anytime*, and *in all directions*.

- Velocities are computed as the time derivative of trajectories in the frame: $\frac{d\vec{x}(t)}{dt}$.

- The property IN implies a certain form of *homogeneity* in space and time (since free particles must move in straight lines *anywhere* and *anytime*) and *isotropy* in space (they must move in straight lines in *any direction*). Without additional empirical input, this does not automatically imply that *every experiment* yields the same result anywhere, anytime and in any direction. This more general form of homogeneity and isotropy will be introduced later as HO and IS. Empirical evidence shows that spacetime indeed is homogeneous HO and space isotropic IS (in the absence of gravity). With this additional input, the "Inertial Test" to establish IN can be simplified to only one particle moving in a straight line at one place for some finite time (which is actually doable). If you presuppose homogeneity HO but *not* isotropy IS, you could observe multiple free particles starting at the same point but moving in different (linearly independent) directions.

> Frames equipped with a coordinate system defined by ES + OC
> which satisfy IN are called ⁂ *inertial coordinate systems*.

To distinguish arbitrary frames $\mathcal{O}$ (with arbitrary coordinates) from the special frames (equipped with Cartesian coordinates and synchronized clocks) that passed the inertial test, we label these coordinate systems by K, K', K'' etc. (if we refer to arbitrary inertial systems) and by A, B, C etc. (if we refer to specific inertial systems); the set of all inertial systems is denoted $\mathcal{I}$.

<u>Alternative definitions:</u>

There seem to be as many definitions of inertial systems as there are texts on SPECIAL RELATIVITY. Some are equivalent, some are not. Some more useful, others less so (none are "wrong", though, because *definitions* cannot be wrong). Some are operational in nature (like the one above), some purely mathematical. Here I only want to point out two ways one can modify the above definition without changing the concept of an inertial system:

- The "inertial test" is crucial to the concept of an inertial frame. It rules out accelerated frames (both linear or rotating). An alternative to throwing test masses in different directions and recording their trajectories is to *repeat* the `ES` procedure periodically to *test* whether the clocks stay in sync. That is, to setup the coordinate system one synchronizes the clocks *once* (by recalibrating the clocks) and then repeats the procedure periodically to check whether the Einstein-synchronization condition remains valid ($\delta t = 0$ in our description above). As it will turn out in GENERAL RELATIVITY, your clocks will not stay in sync in frames that do not pass `IN` (and vice versa). This is essentially the definition given by Schutz [4].

- Instead of "hiding" the law of inertia in the synchronization of clocks, one can do a somewhat reverse modification and "hide" the synchronization of clocks in (an extension of) the law of inertia. To this end one extends the "inertial test" by a second class of tests/experiments, namely:

  `IN*` Two identical particles that are initially adjacent and at rest, and then interact to repel each other, fly apart with the same velocity in opposite directions. (❊ *Isotropy of Inertia*)

  This statement about the isotropy of inertia implies an operational definition of simultaneity that is (empirically) equivalent to `ES`: You synchronize your clocks *such that* `IN*` is satisfied, for example by performing the experiment described by `IN*` equidistant between two clocks. When the particles reach the clocks, you reset both to $t = 0$. In this synchronization `IN*` is satisfied *by construction*; experiments show that clocks synchronized in this way are also synchronized according to `ES` (and vice versa).

4 | ❊ *Spacetime diagram*

:= Data structure that encodes the collected data of an inertial coordinate system K:
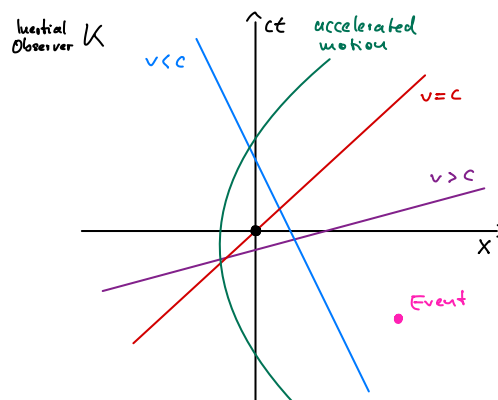
- Often we draw only one dimension of space for the sake of simplicity.

- Because it will prove useful later, we measure time in units of length by multiplying t with the speed of light c. The choice of c is arbitrary at this point.

Notation: ⊲ Two inertial systems K and K':

We use the following shorthand notations to refer to the coordinates of events in the spacetime diagrams of K and K', respectively:

$$(t, \vec{x})_K \equiv (x)_K \equiv x \equiv (t, \vec{x}) \quad \text{and} \quad (t', \vec{x}')_{K'} \equiv (x')_{K'} \equiv x' \equiv (t', \vec{x}') \qquad (1.6)$$

When it is clear to which inertial system the coordinates belong we drop the subscripts K and K'.

> ↓ **Interlude: Reconstructing spacetime diagrams from $\mathcal{E}$**
>
> If you are given the set $\mathcal{E}$ of events you can reconstruct the spacetime diagram of an inertial system K by looking in each coincidence class $E \in \mathcal{E}$ for the clock event $(t, \vec{x})_K \in E$. You then place E (or some sub-event you are interested in) graphically at the coordinate $(t, \vec{x})$ on a sheet of paper. The resulting picture is the spacetime diagram of K. In a another inertial system K' the events are arranged differently because different clock events $(t', \vec{x}')_{K'} \in E$ and hence coordinates $(t', \vec{x}')$ are used to draw the spacetime diagram. How $(t, \vec{x})$ and $(t', \vec{x}')$ are related is unclear at this point.

5 | Empirical facts:

The following facts cannot be bootstrapped from logical thinking alone. They are facts about our physical reality that we have strong experimental evidence for.

- Inertial systems exist (at least in some approximation).

  Examples would be an unaccelerated spaceship floating far away from the solar system or the interior of the international space station (if you do not measure too precisely). In SPECIAL RELATIVITY we assume that these systems can be extended to encompass all of spacetime.

- Constructing inertial systems (of arbitrary size) is not possible everywhere.

  → GENERAL RELATIVITY

  We will find in our discussion of GENERAL RELATIVITY that in a gravitational field the construction of inertial systems is only possible locally. For example: If you extend the ISS inertial system rigidly beyond the ISS itself, at some point you will find the trajectories of free particles to deviate from straight lines due to the inhomogeneity of the gravitational field. We will also see that the synchronization procedure used to calibrate the clocks fails in gravitational fields (you cannot keep your clocks in sync). For our discussion of SPECIAL RELATIVITY we ignore this and assume that our inertial systems cover all of spacetime.

6 | Relations between inertial systems:

i | There are three *straightforward* ways to construct a new inertial system K' from a given one K. They have in common that the two observers do *not* move with respect to one another so that pairs of clocks from K and K' spatially coincide for all times (this implies in particular that you can check that these pairs of clock run at the same rate):

(1) Translation in time by $s \in \mathbb{R}$ (→ 1 parameter)

Procedure:
Duplicate all clocks & rods in place. Label the new clocks with K' and the old position labels. Shift the reading of all clocks by a constant value $-s$:

$$(t', \vec{x}')_{K'} \sim (t, \vec{x})_K \quad \text{with} \quad t' = t - s \quad \text{and} \quad \vec{x}' = \vec{x}. \qquad (1.7)$$

It is easy to see that this modification does not invalidate **ES** , **OC** or **IN** . In particular, the Einstein synchronization condition Eq. (1.2) remains valid:

$$t_O = \tfrac{1}{2}\left(t_A + \tilde{t}_A\right) \quad \Leftrightarrow \quad (t_O - s) = \tfrac{1}{2}\left[(t_A - s) + (\tilde{t}_A - s)\right]. \qquad (1.8)$$

How to check from K:

At $(t)_K = 0$ the reading of the origin clock of K' is shifted by $-s \in \mathbb{R}$.

(2) Translation in space by $\vec{b} \in \mathbb{R}^3$ (→ 3 parameters)

Procedure:

Duplicate all clocks & rods and translate the whole lattice by $\vec{b}$ (since all clocks are type-identical, you can also simply modify the position labels without moving anything). Label the new clocks with K' and keep their synchronization:

$$(t', \vec{x}')_{K'} \sim (t, \vec{x})_K \quad \text{with} \quad t' = t \quad \text{and} \quad \vec{x}' = \vec{x} - \vec{b}. \qquad (1.9)$$

¡! If you move the lattice K' in direction $\vec{b}$, the origin clock of K with position label $\vec{x} = \vec{0}$ will spatially coincide with a clock of K' with position label translated in the *opposite* direction, namely $-\vec{b}$. The same happens for rotations (→ *below*) and translations in time (← *above*).

It is easy to see that this modification does not invalidate **ES** , **OC** or **IN** . In particular, distances can still be computed with Eq. (1.5) since

$$d(\vec{x}, \vec{y}) = d(\vec{x} - \vec{b}, \vec{y} - \vec{b}) \quad \text{for } \vec{b} \in \mathbb{R}^3. \qquad (1.10)$$

How to check from K:

At $(t)_K = 0$ the origin of K' is translated by $\vec{b} \in \mathbb{R}^3$ wrt. the origin of K.

(3) Rotation in space by $R \in \mathrm{SO}(3)$ (→ 3 parameters)

Procedure:

Duplicate all clocks & rods and rotate the whole lattice by the axis and angle defined by the rotation matrix R (since all clocks are type-identical, you can again simply modify the position labels without moving anything). Label the new clocks with K' and keep their synchronization:

$$(t', \vec{x}')_{K'} \sim (t, \vec{x})_K \quad \text{with} \quad t' = t \quad \text{and} \quad \vec{x}' = R^{-1}\vec{x}. \qquad (1.11)$$

It is easy to see that this modification does not invalidate **ES** , **OC** or **IN** . In particular, distances can still be computed with Eq. (1.5) since

$$d(\vec{x}, \vec{y}) = d(R^{-1}\vec{x}, R^{-1}\vec{y}) \quad \text{for } R^{-1} \in \mathrm{SO}(3). \qquad (1.12)$$

How to check from K:

The spatial axes of K' are rotated by $R \in \mathrm{SO}(3)$ wrt. the spatial axes of K.

¡! You can add spatial *reflections* to these transformations (↓ *improper rotations*), i.e., $R \in \mathrm{O}(3)$ instead of $R \in \mathrm{SO}(3)$. In our discussions we will omit these and only comment on them where necessary.

The combination of spatial rotations (proper and improper, i.e., including reflections) and spatial translations form the ↑ *Euclidean group* $\mathrm{E}(3) = \mathrm{ISO}(3)$.

However, experiments (and everyday experience) tell us that there is a fourth possibility how two inertial systems can be related:

Empirical fact:

(4) Uniform linear motion (⁂ *Boost*) by $\vec{v} \in \mathbb{R}^3$ (→ 3 parameters)

You experience this fact whenever you have a very smooth flight: If you don't look out the window (and cover your ears) everything behaves just as if the airplane were standing still on the ground; there is no evidence that you move with several hundred kilometers per hour relative to the ground.

How to check from K:

The origin of K' moves with constant velocity $(\vec{v})_K = \left(\frac{\mathrm{d}x(t)}{\mathrm{d}t}\right)_K \in \mathbb{R}^3$.

Note that just from this observation one cannot distinguish between a *pure boost* and a boost combined with a spatial rotation of the axes (because one probes only for the trajectory of a single point). We will → *later* be more precise about this distinction.

¡! We cannot write down the coordinate transformations for this relation (yet). The fundamental difference to (1)-(3) is that now the clocks of K' move wrt. the clocks of K. We cannot interpret this as a simple relabeling of fixed clocks. We cannot even be sure that the K- and K'-clocks "run at the same rate" (even if they are type-identical) because to check this we would have to compare the reading of a pair of clocks (one in K and one in K') at *two* consecutive points in time. To do this, however, the two clocks must be at the same place (remember that we can only observe coincidences!). But this is not possible: Since the two frames move uniformly, two clocks can never meet twice! As it will turn out, it is this relation (4) [and its concatenations with (1)-(3)] that harbors the essence of SPECIAL RELATIVITY.

ii | *Empirical fact:* The relations (1)-(4) are exhaustive.

With this we mean that whenever you encounter two inertial systems K and K' (i.e., both observers certify that they satisfy our definition of an inertial system, in particular, the "Inertial Test" `IN`), then you will find that the relation between the two is one of the four relations (1)-(4) or a combination of them.

→ The relation of two inertial systems K and K' is given by 10 parameters:
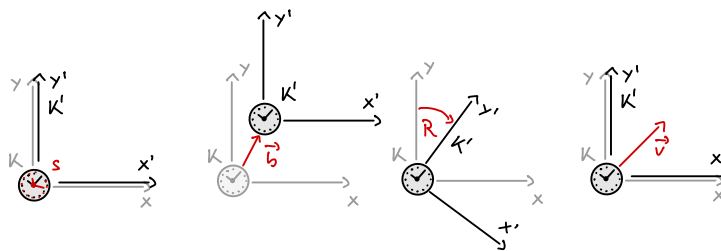
Note that all these relations can be operationally defined and measured within the frame K.

¡! The first three sketches can be taken at face value: For example, a translation in time really corresponds to the situation where all clocks are shifted by s an all spatial labels (in particular the axes) remain unaffected. However, for the boost (the last sketch on the right) we do *not* know (yet) how the coordinates transform (neither time nor space) except that the origin clock of K' follows a trajectory in K with uniform velocity $\vec{v}$. This implies that you should *not* take the sketch for a boost at face value: For example, we do not know whether the axes remain parallel as suggested by the sketch (spoiler: in general they will not).

iii | Since the transformations (1)-(3) do not change the state of motion of the observer (and can therefore be interpreted as a simple relabeling of the position labels and clock readings), it makes sense to collect all inertial frames K that can be connected in this way into an equivalence class $[K]$ which we call …

❋ *Inertial frame* := Equivalence class $[K]$ of all inertial coordinate systems K related by spacetime translations and spatial rotations.
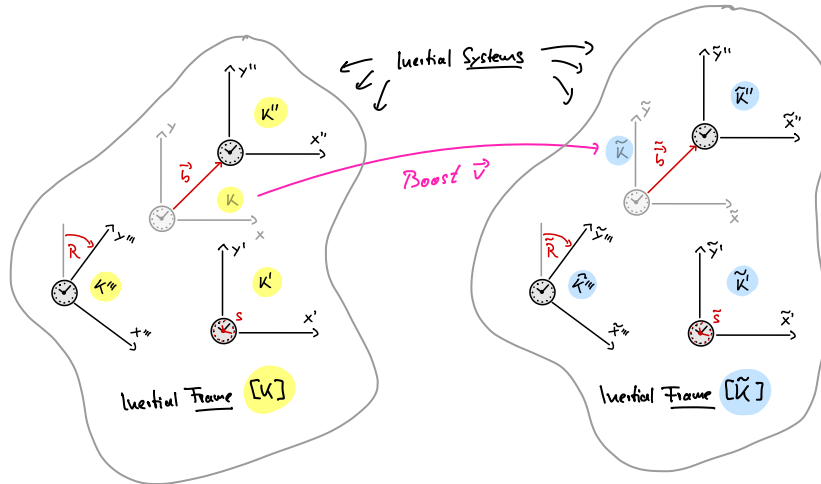
Inertial *frames* $[K]$ therefore correspond to the physical notion of a "state of motion." Physically, an inertial *frame* corresponds to the class of all freely moving particles in the universe that are mutually at rest. Given such a "state of motion" (e.g., by declaring one of the particles as reference point), you can then construct various Cartesian coordinate systems (e.g., using said reference particle as your origin) to describe events; these are the inertial *systems* that make up the equivalence class $[K]$.

iv | *Notation:*

We denote these relations between two inertial systems with the following shorthand notations:
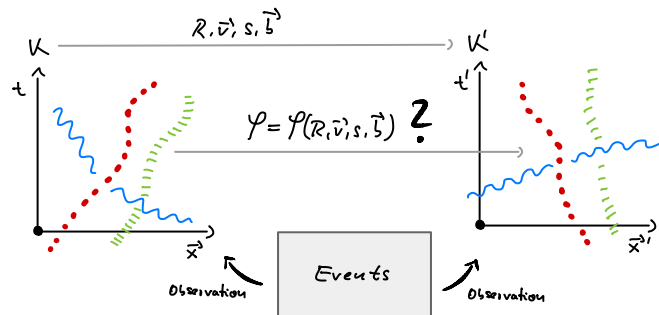
$$K \xrightarrow{R,\vec{v},s,\vec{b}} K', \; K \xrightarrow{R,\vec{v}} K', \; K \xrightarrow{\vec{v}} K', \; K \xrightarrow{v_x} K' \tag{1.13}$$

From left to right the relations become increasingly specialized.

¡! These relations are *not* symmetric (as indicated by the arrow). For example, $K \xrightarrow{v_x} K'$ specifies the situation where the (origin of) system K' moves with velocity v_x in x-direction *as measured in system K.*

v | Coordinate transformations:

◁ Two descriptions of the *same* events:

→ Transformation between these descriptions?

$$\varphi(K \to K') : \; (t, \vec{x})_K \mapsto (t', \vec{x}')_{K'} \quad ❋ \; \textit{Coordinate transformation}$$

Finding the functional form of φ (for the non-trivial case $\vec{v} \neq 0$) will be our main goal and central result of this chapter. However, before we can tackle this problem, we first have to introduce a few more concepts.

↓ Interlude: Relative information

We called the data in $\mathcal{E}$ *absolute* because all observers agree on the coincidence of events. However, this data cannot include arbitrary statements, e.g., the event "the particle has velocity $\vec{v}$" cannot be part of $\mathcal{E}$ because we know from experience that different observers in general do not agree on the velocity of an object. However, following Einstein, we postulated that coincidences are all we can ever observe; thus all there is to know must be encoded in $\mathcal{E}$! How is this consistent with the fact that velocities (for example) cannot show up in $\mathcal{E}$?

To understand this, it is instructive to think about quantities that can be *derived* from the absolute data in $\mathcal{E}$ by means of prescribed *algorithms*. An algorithm $\mathcal{A}$ is simply a program using data from $\mathcal{E}$ to compute other data (it can use potentially multiple events $E_1, E_2, \ldots, E_N \in \mathcal{E}$ to do so). Furthermore, we allow the algorithm to take the label of an inertial system $K \in \mathcal{I}$ as input:

$$\mathcal{A} : \mathcal{E}^N \times \mathcal{I} \to \text{Output data} \qquad (1.14)$$

As a constraint, we require that the algorithm *must not* use any (static) labels $A, B, \ldots \in \mathcal{I}$ of inertial systems. The only reference to a frame it can use is the *variable K*. This somewhat arbitrary sounding restriction formalizes the notion that there are no inertial systems that are "special". Since all inertial systems must be treated equal, the algorithm cannot refer to any specific frame. (This ➜ principle of relativity will take the center stage later and turns out to be crucial for the derivation of the transformation φ.)

Let us now contrive two algorithms to compute two quantities that are clearly physically relevant but are *not* contained in $\mathcal{E}$:

- Example 1: Velocity

  First think about how you would measure the velocity of a particle in the lab: You would detect the particle at two different (but nearby) locations, measure the time it requires to get from one to the other, and then compute the difference quotient of distance traveled by the time needed. Note that there is no way to measure the velocity at one point in space and time; you always need two points!

  To formalize this, consider two events E_1 and E_2 that both contain the sub-event "particle detected". The algorithm $\mathcal{V}(E_1, E_2; K)$ computes the (average) velocity between the two events as follows:

  1. Select the event $(t_1, \vec{x}_1)_K \in E_1$.

  2. Select the event $(t_2, \vec{x}_2)_K \in E_2$.

  3. Compute and return the value $\vec{v} = \frac{\vec{x}_2 - \vec{x}_1}{t_2 - t_1}$.

  It is important that this algorithm can be used *without modifications* by all observers $K \in \mathcal{I}$. To do so, each observer K plugs into $\mathcal{V}$ the two events (which are objective) an its *own* label K (since this is the only non-random choice possible).

  But then two *different* observers K and K' will pick *different* coordinates $(t_i, \vec{x}_i)$ (measured by different clocks) to compute their value of $\vec{v}$, which obviously can yield different outcomes (as expected for velocities). Note that for the velocities to be really different it must be $[K'] \neq [K]$, i.e., the two inertial *systems* must belong to different *frames*.

- Example 2: Duration & Simultaneity

  A very natural question is how much time passed between two events E_1 and E_2. The formal prescription how to answer this question is given by the algorithm $\mathcal{T}(E_1, E_2; K)$:

  1. Select the event $(t_1, \vec{x}_1)_K \in E_1$.

  2. Select the event $(t_2, \vec{x}_2)_K \in E_2$.

> 3. Compute and return the value $\Delta t = t_2 - t_1$.
>
> For the very same reason as for the velocity algorithm above, the return value of course will depend on the chosen "clock events" $(t_i, \vec{x}_i)$. And so for the very same reason that velocities can be observer-dependent, time intervals can be as well. Since we define "simultaneity" as the property $\Delta t = 0$, this possibility for observer-dependent results directly transfers to our notion of simultaneity!
>
> Note that we did not make *quantitative* statements about the outcomes for different observers. We neither showed *how* velocities depend on the frame nor whether simultaneity really *is* relative. (It could just be the case that in our world $t_2 - t_1$ always equals $t_2' - t_1'$ for a fixed event.) This depends on the actual numbers of the coordinates. Such statements therefore require quantitative statements about the relation of $(t, \vec{x})_K \in E$ and $(t', \vec{x}')_{K'} \in E$, which we do not know at this point (this is exactly the question for the functional form of the coordinate transformation φ).
>
> However, what we did show is the *possibility* that simultaneity is relative, just as we already expect velocities to be! So when we later find the correct transformation φ and (surprise!) that indeed simultaneity is not an observer independent fact, you should not be surprised.
>
> *Question:* Can the values of the electric and magnetic fields $\vec{E}$ and $\vec{B}$ be included in $\mathcal{E}$? If not, can you think of an algorithm that determines the electric and magnetic fields $\vec{E}$ and $\vec{B}$ using only coincidence data available in $\mathcal{E}$? Do you expect the electromagnetic field to be observer-dependent?

7 | Henceforth:

> Unless noted otherwise, all frames will be *inertial* (with Cartesian coordinates).
> → We will (almost exclusively) work with *inertial coordinate systems*.

We use the concept of inertial systems because to describe physics by equations, coordinates are a useful tool. As it turns out, Cartesian coordinates allow for particularly simple equations (at least if space is Euclidean). So our concept of inertial systems as defined above is the most useful one.

8 | Physical Models:

Let us fix a bit of terminology:

- �֎ *(Physical) laws* are ontic features of reality (↑ *scientific realism*).

  Physical laws can only be *discovered*; they can neither be invented nor modified.

- ✷ *(Physical) models* are algorithms used to describe reality.
  These algorithms are typically encoded in the language of mathematics.

  Physical models are *invented* and can be *modified*; I will use the terms *model* and *theory* interchangeably.

¡! These definitions are by no means conventional and you will find many variations in the literature. For the following discussion, it is only important that the terms we use have precise meaning.
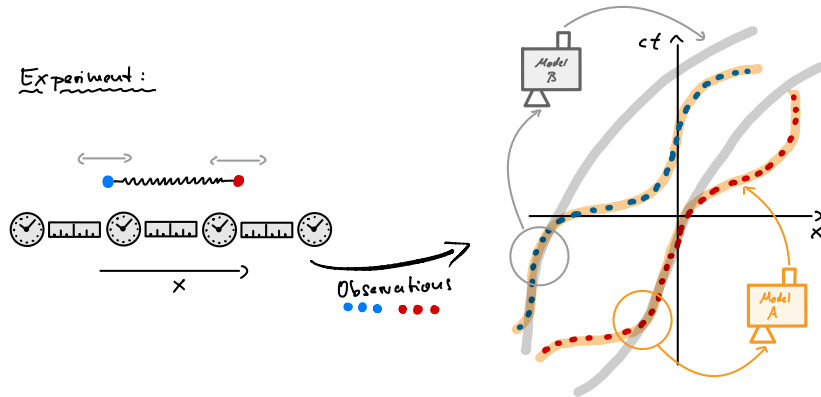
¡! The validity of models cannot be *proven*; we can only gradually increase our trust in a model by repeated observations (experiments) – or reject it as invalid by demonstrating that its predictions contradict reality (↑ *Karl Popper*). Note that models might describe reality only approximately and in specific parameter regimes and still be useful.

You may dismiss this focus on terminology as "philosophical banter." Conceptual clarity, however, is absolutely crucial for science – in particular for RELATIVITY. Whenever there is confusion in physics, it is often rooted in the conceptual fuzziness of our thinking.

1.2. Galilei's principle of relativity

9 │ Example: Newtonian mechanics

i │ Definition of the model:

- ⊲ Closed system of N massive particles with masses m_i and positions $\vec{x}_i$.

- ⊲ Force exerted by k on i:

$$F_{k \to i}(\vec{x}_k - \vec{x}_i) = (\vec{x}_k - \vec{x}_i)\, f_{k \leftrightarrow i}(|\vec{x}_k - \vec{x}_i|) \tag{1.15}$$

It is $f_{k \leftrightarrow i} = f_{i \leftrightarrow k}$ and therefore $F_{k \to i}(\vec{x}_k - \vec{x}_i) = -F_{i \to k}(\vec{x}_i - \vec{x}_k)$.

→ *Newtonian equations of motion* (in some inertial system K):

$$m_i \frac{\mathrm{d}^2 \vec{X}_i}{\mathrm{d}t^2} = \sum_{k \neq i} \vec{F}_{k \to i}(\vec{X}_k - \vec{X}_i) \tag{1.16}$$

We denote with $\vec{X}_i \equiv \vec{X}_i(t)$ coordinate-valued *functions*; i.e., $\vec{x}_i = \vec{X}_i(t)$ determines a spatial point $\vec{x}_i$ for given t.

Remember: This model fully implements "Newton's laws of motion":

1. Lex prima:

   A body remains at rest, or in motion at a constant speed in a straight line, unless acted upon by a force.

   This is the ↡ *principle of inertia*. It is part of the definition of the concept of a Newtonian force used in Eq. (1.16). Note that it is *not* a consequence of Eq. (1.16) for $F_{k \to i} \equiv 0$. It rather defines (together with the lex tertia below) the frames and coordinate systems (← *inertial systems*) in which Eq. (1.16) is valid (recall **IN**).

2. Lex secunda:

   When a body is acted upon by a net force, the body's acceleration multiplied by its mass is equal to the net force.

   This is just the functional form of Eq. (1.16) in words.

3. Lex tertia:

   If two bodies exert forces on each other, these forces have the same magnitude but opposite directions.

   This is guaranteed by the property $F_{k\to i} = -F_{i\to k}$ of the forces. Together with the lex secunda this is an expression of momentum conservation. For two particles:

   $$m_1 \frac{\mathrm{d}v_1}{\mathrm{d}t} + m_2 \frac{\mathrm{d}v_2}{\mathrm{d}t} = \frac{\mathrm{d}p_1}{\mathrm{d}t} + \frac{\mathrm{d}p_2}{\mathrm{d}t} = F_{2\to 1} + F_{1\to 2} = 0 \qquad (1.17)$$

   This implies in particular that two identical particles ($m_1 = m_2$) that are both at rest at $t = 0$ must obey $v_1(t) = -v_2(t)$ for all times (recall **IN***).

ii | Application of the model:
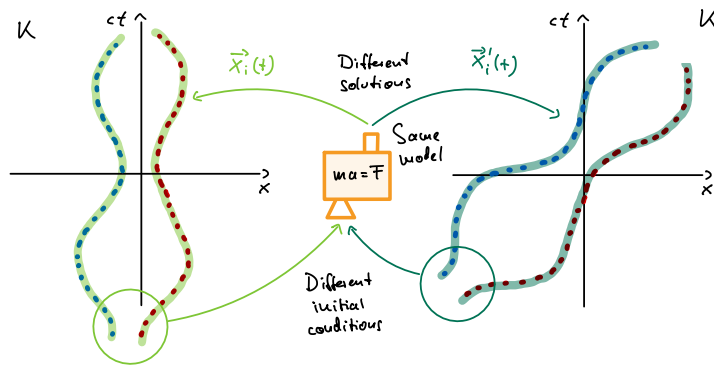
As a working hypothesis, let us assume that the model Eq. (1.16) describes the dynamics of massive particles perfectly (from experience we know that there are at least regimes where it is good enough for all practical purposes).

iii | Symmetries of Newtonian mechanics:

To understand the solution space of Eq. (1.16) better, it is instructive to study transformations that map solutions to other solutions.

a | ⁂ *Galilei transformations*:

We define the following coordinate transformation:

$$G \,:\, \mathbb{R}^4 \to \mathbb{R}^4 \,:\, \begin{cases} t' = t + s \\ \vec{x}' = R\vec{x} + \vec{v}t + \vec{b} \end{cases} \qquad (1.18)$$

A Galilei transformation G is characterized by 10 real parameters:

- $s \in \mathbb{R}$: Time translation (1 parameter)
- $\vec{b} \in \mathbb{R}^3$: Space translation (3 parameters)
- $\vec{v} \in \mathbb{R}^3$: Boost (3 parameters)
- $R \in \mathrm{SO}(3)$: Spatial rotation (3 parameters; rotation axis: 2, rotation angle: 1)

The set of all transformations forms (the matrix representation of) a group:

$$\mathcal{G}_+^{\uparrow} = \{G(R, \vec{v}, s, \vec{b})\} \quad \text{⁂} \textit{ Proper orthochronous Galilei group} \qquad (1.19)$$

with group multiplication

$$G_3 = G_1 \cdot G_2 = G(\underbrace{R_1 R_2}_{R_3}, \underbrace{R_1 \vec{v}_2 + \vec{v}_1}_{v_3}, \underbrace{s_1 + s_2}_{s_3}, \underbrace{R_1 \vec{b}_2 + \vec{v}_1 s_2 + \vec{b}_1}_{\vec{b}_3}) \quad (1.20)$$

You derive this multiplication in ❷ Problemset 1 and show that the group axioms are indeed satisfied.

As a special case, the multiplication yields the rule for addition of velocities in Newtonian mechanics:

$$G(\mathbb{1}, \vec{v}_1, 0, \vec{0}) \cdot G(\mathbb{1}, \vec{v}_2, 0, \vec{0}) = G(\mathbb{1}, \underbrace{\vec{v}_1 + \vec{v}_2}_{\vec{v}_3}, 0, \vec{0}) \qquad (1.21)$$

The full Galilei group is generated by the proper orthochronous transformations together with space and time inversion:

$$\mathcal{G} = \langle \mathcal{G}_+^{\uparrow} \cup \{P, T\} \rangle \quad \text{⁂} \textit{ Galilei group} \qquad (1.22a)$$
$$P : (t, \vec{x}) \mapsto (t, -\vec{x}) \quad \text{Space inversion (parity)} \qquad (1.22b)$$
$$T : (t, \vec{x}) \mapsto (-t, \vec{x}) \quad \text{Time inversion} \qquad (1.22c)$$

b | Galilei covariance & Form-invariance:

Details: ❷ Problemset 1

◁ Coordinate transformation Eq. (1.18)

We express the total differential and the trajectory in the new coordinates:

$$\frac{\mathrm{d}}{\mathrm{d}t} = \frac{\mathrm{d}t'}{\mathrm{d}t} \frac{\mathrm{d}}{\mathrm{d}t'} = \frac{\mathrm{d}}{\mathrm{d}t'} \qquad (1.23)$$

and

$$\vec{X}_i'(t') = R\vec{X}_i(t) + \vec{v}t + \vec{b} = R\vec{X}_i(t' - s) + \vec{v}(t' - s) + \vec{b} \qquad (1.24a)$$
$$\Leftrightarrow \quad \vec{X}_i(t) = R^{-1}\left[\vec{X}_i'(t') - \vec{v}(t' - s) - \vec{b}\right] \qquad (1.24b)$$

Thus the left-hand side of the Newtonian equation of motion Eq. (1.16) reads in new coordinates:

$$m_i \frac{\mathrm{d}^2 \vec{X}_i(t)}{\mathrm{d}t^2} = m_i \frac{\mathrm{d}^2}{\mathrm{d}t'^2} R^{-1}\left[\vec{X}_i'(t') - \vec{v}(t' - s) - \vec{b}\right] = R^{-1} m_i \frac{\mathrm{d}^2 \vec{X}_i'(t')}{\mathrm{d}t'^2}$$
$$(1.25)$$

Note that the quantity $m_i \frac{\mathrm{d}^2}{\mathrm{d}t^2} \vec{X}_i(t)$ is not *invariant*; it transforms with an $R^{-1} \in \mathrm{SO}(3)$. And the right-hand side:

$$\sum_{k \neq i} \vec{F}_{k \to i}(\vec{X}_k(t) - \vec{X}_i(t)) = R^{-1} \sum_{k \neq i} \vec{F}_{k \to i}(\vec{X}_k'(t') - \vec{X}_i'(t')) \qquad (1.26a)$$

Here we used the form of the force Eq. (1.15), that $\vec{X}_k(t) - \vec{X}_i(t) = R^{-1}[\vec{X}'_k(t') - \vec{X}'_i(t')]$ and $|\vec{X}_k(t) - \vec{X}_i(t)| = |\vec{X}'_k(t') - \vec{X}'_i(t')|$ because of $R \in \mathrm{SO}(3)$.

Note that the force on the right-hand side is not *invariant* either; luckily, it transforms with *the same* $R^{-1} \in \mathrm{SO}(3)$; it "co-varies" with the left-hand side!

In conclusion, Newton's equation of motion Eq. (1.16) reads in the new coordinates:

$$\underbrace{R^{-1} m_i \frac{\mathrm{d}^2 X'_i(t')}{\mathrm{d}t'^2} = R^{-1} \sum_{k \neq i} \vec{F}_{k \to i}(\vec{X}'_k(t') - \vec{X}'_i(t'))}_{\to \; Covariance} \qquad (1.27a)$$

$$\overset{\times R}{\Leftrightarrow} \quad \underbrace{m_i \frac{\mathrm{d}^2 X'_i(t')}{\mathrm{d}t'^2} = \sum_{k \neq i} \vec{F}_{k \to i}(\vec{X}'_k(t') - \vec{X}'_i(t'))}_{\to \; Form\text{-}invariance} \qquad (1.27b)$$

(You can easily check that this holds for P and T as well.)

$\rightarrow$

> Newton's EOMs Eq. (1.16) are *form-invariant* under Galilei transformations.
>
> *Or:* Newton's EOMs Eq. (1.16) are Galilei-*covariant*.

↓ Interlude: Nomenclature

Let X be some group of coordinate transformations (here: $X = \mathcal{G}$ the Galilei group).

- A *quantity* is called X-*invariant* if it does not change under the coordinate transformation. Such quantities are called X-*scalars*.

  An example is the mass m in Eq. (1.16) (which is also constant).

- A *quantity* is called X-*covariant* if it transforms under some given representation of the X-group. If this representation is the trivial one (i.e., the quantity does not change at all) this particular X-covariant quantity is then also an X-scalar.

  An example of a Galilei-covariant (but not invariant) quantity is the force $\vec{F}_{k \to i}$ which transforms under a representation of $\mathcal{G}$.

- An *equation* is called X-*covariant* if the quantity on the left-hand side and on the right-hand side are X-covariant (under the same X-representation).

  An example is Newton's lex secunda Eq. (1.16) where $m_i \frac{\mathrm{d}^2}{\mathrm{d}t^2} x_i(t)$ transforms in the same (non-trivial) representation as $\vec{F}_{k \to i}$.

- X-covariant *equations* have the feature that a X-transformation leaves them *form-invariant*, i.e., they "look the same" after X-transformations because their left- and right-hand side vary in the same way (they "co-vary"). Note that the quantities in a form-invariant equation do not have to be *invariant*.

  An example is again Eq. (1.16) as we just showed. Note that $\vec{x}'_i(t')$ and $\vec{x}_i(t)$ are *different* vectors such that the two sides of the equation as not *invariant* (but *covariant*).