Galilean Relativity

Larry Susanka

Mathematics Department
Bellevue College

November 27, 2017

Galileo

Galileo created in "Dialogue Concerning the Two Chief World System," his 1632 defense of his scientific ideas, a thought experiment which leads to a Relativity Principle in the formulation of scientific law. The statements contained therein directly contradict the assumptions about the world believed to be true from antiquity and, by implication, support heliocentrism.

Defending the physical truth of Copernican heliocentrism was already deemed heretical in 1616. The Dialogues were found to be a defence of this idea, in spite of Galileo's denial, in 1633. Galileo was thereupon confined to house arrest until his death in 1642. His case was not helped by the fact that the foolish-seeming Simplicio often argued as did the Pope, Pius VIII, in *his* writings.

In this work Galileo creates a debate between the foil, Simplicio, defender of Aristotle and Ptolemy, and Salviati, who attempts to enlighten him with modern ideas about nature. In particular, Salviati defends the idea that one should actually look at the world to see how it works, rather than simply accept ancient beliefs, unquestioningly. A third participant, Segredo, acts as an—initially—neutral party, an "intelligent layman."

(Segredo and Salviati are the names of two of Galileo's friends. Galileo claimed that Simplicio was named after Simplicius of Cilicia, a sixth-century commentator on Aristotle, but "Simplicio" could be a slightly veiled version of "Simpleton." The dialog targets two of Galileo's critics Lodovico delle Colombe and the Padovan Cesare Cremonini, who refused Galileo's offer to view his astronomical discoveries through the new telescope.)

In Stillman Drake's 1967 translation (page 186-7, The Second Day) we find Salviati arguing as follows:

Shut yourself up with some friend in the main cabin below decks on some large ship, and have with you there some flies, butterflies, and other small flying animals. Have a large bowl of water with some fish in it; hang up a bottle that empties drop by drop into a wide vessel beneath it. With the ship standing still, observe carefully how the little animals fly with equal speed to all sides of the cabin. The fish swim indifferently in all directions; the drops fall into the vessel beneath; and, in throwing something to your friend, you need throw it no more strongly in one direction than another, the distances being equal; jumping with your feet together, you pass equal spaces in every direction.

When you have observed all these things carefully (though there is no doubt that when the ship is standing still everything must happen in this way), have the ship proceed with any speed you like, so long as the motion is uniform and not fluctuating this way and that. You will discover not the least change in all the effects named, nor could you tell from any of them whether the ship was moving or standing still.

This is a "thought experiment" but Galileo claims to have actually done experiments of this kind.

The World is the collection of events, such as the position of a flying insect at a particular time, that we want to describe. We will denote the world by the symbol $\mathfrak{M}$.

A worldline is a function of time whose value at each time is a certain event: a path in $\mathfrak{M}$. The moving butterfly traces out a worldline, following along with $L\colon \mathbb{R} \to \mathfrak{M}$.

All of Galileo's examples involve position and motion and—what we would call—vectors that describe them. So we are led (in our modern terms) to "coordinatize" the world by choosing an origin and a basis of vectors.

We will describe events by recording the coordinates of the position of the event at a certain time, and each worldline corresponds to a path of coordinates.

So coordinates x or y are functions

$$x\colon \mathfrak{M} \to \mathbb{R}^4 \quad \text{or} \quad y\colon \mathfrak{M} \to \mathbb{R}^4.$$

Coordinates of worldlines are

$$x \circ L\colon \mathbb{R} \to \mathbb{R}^4 \quad \text{or} \quad y \circ L\colon \mathbb{R} \to \mathbb{R}^4 \quad \text{where} \quad x \circ L = (x \circ y^{-1}) \circ y \circ L.$$

Of course, Newton was well aware of Galileo's *Dialogues* as he formulated his theory. Obviously he couldn't use this (then-non-existent) vocabulary, but Newton and other practitioners of the day definitely understood the ideas.

Galileo posits here that orientation or location or a (constant) velocity should not, alone, change our perception of physical processes and, by implication, the physical laws that capture some of their behavior.

A frame (of reference) is a choice of origin and basis and a way of measuring time. An observer is said to "witness" or "preside over" coordinates provided by a specific frame.

Galileo places restrictions on the frame. He says that the ship should not "fluctuate this way and that." Some frames are admissible, and some are not.

The Relativity Principle

All admissible frames of reference are to be completely equivalent for our formulation of Physical Law.

Newton

Newton's First Law:

Every particle in a state of uniform motion tends to remain in that state of motion unless it is under the influence of a force.

This tells you that *if* you are describing the world you can see the effect of mysterious entities called "forces."

Forces are real things in the world, but they may be detected by measuring. They can be seen because they cause changes in velocity.

One is led to consider changing velocities, and their vector nature, and coordinates of events and forces.

Newton's Second Law:

The relationship between a particle's mass m, its acceleration a and the applied force F is

$$F = m\,a.$$

This tells us forces are vectors. It also implies that the motion—that is, coordinates $x \circ L$ of the world line parameterized by time—of any particle is (at least) twice differentiable. We can give a value to a force by observing a, which is the acceleration of the "space part" of coordinates of the worldline.

This equation encourages us to find the proportionality constant, the number m mentioned above, for each particle.

This does *not* tell us where the force comes from, nor how to calculate it absent a visible effect. And it definitely does not tell us how to assign mass to our particles.

Newton simply described the mass as determined by the volume and the density.

At the start of the Principia (1687) we have:

The quantity of matter is that which arises conjointly from its density and magnitude. A body twice as dense in double the space is quadruple in quantity. This quantity I designate by the name of body or of mass.

And later:

It can also be known from a body's weight, for (by making very accurate experiments with pendulums) I have found it to be proportional to the weight...

I think it is safe to say (please correct me with references to examples if I am wrong) that for the first hundred years of its application, any specific mass in Newtonian mechanics was determined, ultimately, by a balance scale.

Newton's Third Law:

For every action there is an equal and opposite reaction.

This is interpreted to mean that for any force caused by the interaction between *pairs* of particles, the force F_1^2 (as a vector) on particle P_1 caused by particle P_2 will be the negative of the force F_2^1 on particle P_2 caused by particle P_1.

Implicitly, the Third Law is also interpreted to mean that the interactions involved in systems of particles can be understood by examining pairs of interacting particles, and the net force on a particle is the sum of all the forces acting on that particle.

So Galileo set the coordinate framework (which Newton completed by implication) and described the preferred observers.

An observer $\mathcal{O}$ is witness to the values of an invertible function

$$x\colon \mathcal{M} \to \mathbb{R}^4.$$

Newton laws, then, are a "call to action." Look for the forces. When you find them you will know (by solving the DE) the coordinates of worldlines of interesting phenomena.

Inner Products

If V is any real vector space, a bilinear form on V is a function

$$g\colon V \times V \to \mathbb{R}$$

that is linear in each "slot" separately: that is, for each $v \in V$ the functions $g(v,\cdot)$ and $g(\cdot, v)$ are linear.

Using coordinates x in a basis for V we have

$$g = g_{ij}\, dx^i \otimes dx^j.$$

The specific basis is "there" but is usually not mentioned! In our application, you are supposed to just "know what to do" when you switch from one coordinate system to another…*and from one place to another!*

For us, V represents the tangent vectors *starting from a particular place in the world*, derivatives of curves through the place, and V^*, the dual of V, represents *real valued functions in a small neighborhood of that place*, characterized by their family of level surfaces: the gradient.

Using as an example the common rectangular-to-polar two dimensional situation *at a particular place in the world*

$$\begin{aligned} g &= g_{11}(rect)\, dx \otimes dx + g_{12}(rect)\, dx \otimes dy \\ &\qquad + g_{21}(rect)\, dy \otimes dx + g_{22}(rect)\, dy \otimes dy \\ &= g_{11}(pol)\, dr \otimes dr + g_{12}(pol)\, dr \otimes d\theta \\ &\qquad + g_{21}(pol)\, d\theta \otimes dr + g_{22}(pol)\, d\theta \otimes d\theta. \end{aligned}$$

For two coordinate systems x and y the coefficients $g_{ij}(x)$ and $g_{ij}(y)$ are related by the usual Jacobian factors *evaluated at the coordinates corresponding to "the place"*.

$$\begin{aligned} g =& g_{ij}(x)\, dx^i \otimes dx^j = g_{ij}(x) \left(\frac{\partial x^i}{\partial y^k} dy^k \right) \otimes \left(\frac{\partial x^j}{\partial y^m} dy^m \right) \\ =& \left(g_{ij}(x) \frac{\partial x^i}{\partial y^k} \frac{\partial x^j}{\partial y^m} \right) dy^k \otimes dy^m = g_{km}(y)\, dy^k \otimes dy^m. \end{aligned}$$

In matrix form this is

$$(\, g_{km}(y)\,) = \left(\frac{\partial x^i}{\partial y^k} \right)^T (\, g_{ij}(x)\,) \left(\frac{\partial x^j}{\partial y^m} \right).$$

Spacetime TOC Galileo and Newton Inner Products Spacetime and Observers The Galilean Group Newtonian Relativity

The prototypical bilinear form is dot product on $\mathbb{R}^3$, but dot product has other nice properties that may or may not hold for a generic bilinear form.

If a bilinear form is symmetric and nondegenerate and positive definite it is called an inner product.

$$\text{Symmetry:} \quad g(v,w) = g(w,v) \quad \forall v, w \in V$$

$$\text{Nondegeneracy:} \quad g(v,w) = 0 \;\; \forall v \in V \text{ iff } w = 0.$$

$$\text{Positive Definiteness:} \quad g(v,v) > 0 \text{ unless } v = 0.$$

If that last condition is removed and $g(v,v) > 0$ for some v but $g(v,v) < 0$ for *other* v we have an indefinite inner product. The Minkowski inner product, or "metric" in the Physics vocabulary, is an indefinite inner product.

Spacetime TOC Galileo and Newton Inner Products Spacetime and Observers The Galilean Group Newtonian Relativity

Any symmetric bilinear form corresponds to a symmetric matrix M_a in a basis a-in-the-world.

If v_a and w_a are vectors of coordinates of vectors-in-the-world v and w for that basis then $v_a^T\, M_a\, w_a$ corresponds to $g(v,w)$.

By facts from Linear Algebra any symmetric matrix can be brought to diagonal form by an orthogonal matrix of transition $P_{b\leftarrow a}$, where the columns of $P_{b\leftarrow a}$ are the coordinates of the old basis in terms of the new basis. But for *orthogonal* matrices, the inverse is the transpose. $P_{b\leftarrow a}^T = P_{b\leftarrow a}^{-1} = P_{a\leftarrow b}$.

Conclusion: this diagonalized matrix is M_b, the matrix that represents g in basis b, as can be seen by

$$\begin{aligned} v_a^T M_a w_a =& (P_{a\leftarrow b} v_b)^T M_a (P_{a\leftarrow b} w_b) = \left(v_b^T P_{a\leftarrow b}^T \right) M_a \, (P_{a\leftarrow b} w_b) \\ =& v_b^T \left(P_{b\leftarrow a}^T M_a P_{a\leftarrow b} \right) w_b = v_b^T M_b w_b. \end{aligned}$$

Spacetime TOC Galileo and Newton Inner Products Spacetime and Observers The Galilean Group Newtonian Relativity

So for symmetric g we have discovered a basis b-in-the-world for which the matrix M_b representing g is diagonal. A basis of this type is called orthogonal with respect to g.

These diagonal entries are the eigenvalues of the matrix of g in *any* basis. Some of these diagonal entries will be positive, some negative and some zero.

If g is an inner product all will be positive. If g is an indefinite inner product some of the diagonal entries will be negative and some will be positive but *none* are zero.

(A zero eigenvalue would violate the nondegeneracy condition.)

With diagonalizing b in hand, obtained by an orthogonal matrix of transition, we now make one final change of basis. First, reorder b so that positive entries are first on the diagonal, then the negative ones, and finish off with the eigenvectors from b for eigenvalue 0.

Now normalize these re-ordered vectors $b = \{ b_1, \ldots, b_n \}$ by dividing the eigenvectors for nonzero eigenvalues by $\sqrt{g(b_i,\, b_i)}$ or $\sqrt{-g(b_i,\, b_i)}$, sign chosen so the square root is real.

Call this new basis c. It is said to be orthonormal with respect to g and M_c is diagonal with only ± 1 or 0 on the diagonal. The ones come first, then the minus ones, then zeros, if any.

Be aware that some sources reverse the order of ones and minus ones.

Spacetime TOC Galileo and Newton Inner Products Spacetime and Observers The Galilean Group Newtonian Relativity

For the Euclidean inner product on $\mathbb{R}^3$, the resulting matrix has 1 repeated three times. So the inner product is given by

$$g = dx^1 \otimes dx^1 + dx^2 \otimes dx^2 + dx^3 \otimes dx^3.$$

This inner product actually reflects right angles, lengths and projections in the world. Presumably, a sensible person would have chosen a basis like this from the outset.

Spacetime TOC Galileo and Newton Inner Products Spacetime and Observers The Galilean Group Newtonian Relativity

For the Minkowski inner product, on the other hand, 1 occurs once and -1 occurs three times and zero does not occur.

We will presume that, whatever basis-in-the-world we started with, we have created an orthonormal basis-in-the-world and are using it to describe vectors and g. We let x denote coordinates in that basis. Then

$$g = dx^0 \otimes dx^0 - dx^1 \otimes dx^1 - dx^2 \otimes dx^2 - dx^3 \otimes dx^3.$$

Be aware that some sources reverse this for the Minkowski metric and have

$$g = -dx^0 \otimes dx^0 + dx^1 \otimes dx^1 + dx^2 \otimes dx^2 + dx^3 \otimes dx^3.$$

There is no physical significance to this difference, but it causes some translation difficulties when going from one source to another.

Spacetime TOC Galileo and Newton Inner Products Spacetime and Observers The Galilean Group Newtonian Relativity

Spacetime—Back to the Physical World

We want to describe locations in space and time, and we presume we know what that means. We also want to imagine something physical happening at a place and time and moving, at other times and in a continuous way, to different places.

Something that happens at a particular place and time is called an event, and a worldline is a choice of events for an interval of times.

(As we build our relativistic structure in later work we will see that not every choice of events, that you might hope constitutes a worldline of an object in the world, is physically realizable. But that will be a result of our theory, not a presupposition.)

We have decided to represent events in our world by 4-tuples where the first coordinate represents time and the last three represent location in space relative to a choice of origin and axes.

In the Euclidean world the inner product does not involve time, but only the space coordinates.

But later we will see, in relativistic mechanics, that time (the first coordinate) and space (the last three) are tangled up in an unusual way and this is reflected in the Minkowski inner product.

Allowable Observers

Some of the most important underlying assumptions in both classical and relativistic mechanics involve the possibility of finding coordinates (i.e. a basis in the world) with certain properties. That association with these special properties will be called “an allowable observer.”

We understand that perfect observation is impossible and do not require that of our allowable observers, but we describe below procedures that we assume can be carried out accurately to any desired precision.

We assume that we can describe a clock that everyone would agree measures time accurately and independently of other aspects of our association. The vibration modes of certain unmoving atoms comes to mind as a good choice. All allowable observers use the same type of clock to measure time displacements. This allows us to define a time unit, the second, which would be agreed-upon by all allowable observers.

Space seems to be a three dimensional real vector space, and our allowable observer assumes this, and this observer knows when two displacements in the world are at right angles to each other.

When the observer takes a stick whose length was measured and rotates that stick to a direction at right angles to its former direction we presume that the observer cannot detect any length change.

We presume that there are objects in the world accessible to all observers which exhibit a fixed size. Two marks on a hunk of platinum/iridium sitting in your lab, perhaps. Or if that is not good enough, some other unit of distance.

Here, we don't argue about technical details.

We simply assume that a unit of distance can be agreed-upon and our observers all agree to use it. We will call this unit the meter.

And we presume that our observer has access to telescopes and other basic tools and has picked a non-rotating frame that has constant velocity with respect to some presumed, non-accelerating standard object. Possibly the average velocity of nearby stars will be good enough. Or your lab tabletop.

Galileo would have insisted, so we do too.

Up to this point the setup for classical and relativistic mechanics is the same.

But now we insist, as did Galileo and Newton, that all admissible observers should measure the same distances between events.

This means that the space-part of the function $x \circ y^{-1}$ that translates between coordinates x for observer $\mathcal{O}$ and coordinates y for observer $\mathcal{O}'$ is an isometry from $\mathbb{R}^3$ to itself.

If we do a quick translation to account for a moving constant-velocity origin with respect to one observer or the other, we have at each instant an isometry that sends the origin to the origin.

The Mazur-Ulam theorem tells us that this isometry must be linear and so the translator from $\mathcal{O}$ space coordinates to $\mathcal{O}'$ space coordinates corresponds to matrix multiplication by some matrix M. Because distances are preserved M is orthogonal and so corresponds to a reflection, a rotation (two reflections) or a rotation and a reflection (three reflections). This is the Cartan-Dieudonné Theorem.

And because the distant stars do not appear to rotate to either observer, this rotation matrix is constant as time passes.

The translators from one coordinate system to another form a group with composition of functions, called the Galilean group.

All admissible observers agree on the passage of time and the unit of distance. The origins may move with constant relative velocity. The axis differences may involve reflection and rotation.

$$C\begin{pmatrix}a^0\\a^1\\a^2\\a^3\end{pmatrix}=\begin{pmatrix}x^0\\x^1\\x^2\\x^3\end{pmatrix}+a^0\begin{pmatrix}0\\v^1\\v^2\\v^3\end{pmatrix}+\begin{pmatrix}1&0&0&0\\0&m_{11}&m_{12}&m_{13}\\0&m_{21}&m_{22}&m_{23}\\0&m_{31}&m_{32}&m_{33}\end{pmatrix}\begin{pmatrix}a^0\\a^1\\a^2\\a^3\end{pmatrix}$$

The orthogonal three by three matrices have dimension 3 so this is a ten dimensional Lie group with function composition.

If M is the 3×3 orthogonal matrix in the lower right and $x=\left(x^0\,x^1\,x^2\,x^3\right)^T$ is a vector in $\mathbb{R}^4$ we define the 4×4 matrix $\widetilde{M}$ and the member $\overline{x}$ of $\mathbb{R}^3$ by

$$\widetilde{M}=\begin{pmatrix}1&0&0&0\\0&m_{11}&m_{12}&m_{13}\\0&m_{21}&m_{22}&m_{23}\\0&m_{31}&m_{32}&m_{33}\end{pmatrix}\quad\text{and}\quad\overline{x}=\begin{pmatrix}x^1\\x^2\\x^3\end{pmatrix}.$$

Then if v is the vector $\left(v^1\,v^2\,v^3\right)^T$ in $\mathbb{R}^3$ we have

$$\begin{aligned}C\begin{pmatrix}a^0\\\overline{a}\end{pmatrix}&=\begin{pmatrix}x^0\\\overline{x}\end{pmatrix}+a^0\begin{pmatrix}0\\v\end{pmatrix}+\widetilde{M}\begin{pmatrix}a^0\\\overline{a}\end{pmatrix}\\&=\begin{pmatrix}x^0\\\overline{x}\end{pmatrix}+\begin{pmatrix}0\\a^0v\end{pmatrix}+\begin{pmatrix}a^0\\M\,\overline{a}\end{pmatrix}=\begin{pmatrix}x^0+a^0\\\overline{x}+a^0v+M\,\overline{a}\end{pmatrix}.\end{aligned}$$

We will interested in the derivative of $C=y\circ x^{-1}\colon\mathbb{R}^4\to\mathbb{R}^4$.

The derivative of this kind of function will generally be a 4×4 matrix whose entries are functions of position, the linear map that best represents the changes in C near a. But because of the form of C this derivative is *constant:* it doesn't depend on a.

$$C\begin{pmatrix}a^0\\\overline{a}\end{pmatrix}=\begin{pmatrix}x^0+a^0\\\overline{x}+a^0v+M\,\overline{a}\end{pmatrix}\iff C'(a)=\begin{pmatrix}1&0&0&0\\v^1&m_{11}&m_{12}&m_{13}\\v^2&m_{21}&m_{22}&m_{23}\\v^3&m_{31}&m_{32}&m_{33}\end{pmatrix}.$$

If C is the translator form coordinates x to coordinates y we often see C' indicated by the Jacobian matrix $\left(\frac{\partial y^j}{\partial x^i}\right)$.

If G is another Galilean transformation

$$D\begin{pmatrix}a^0\\\overline{a}\end{pmatrix}=\begin{pmatrix}y^0\\\overline{y}\end{pmatrix}+\begin{pmatrix}0\\a^0w\end{pmatrix}+\begin{pmatrix}a^0\\N\,\overline{a}\end{pmatrix}.$$

then $D\circ C$ is given by

$$D\circ C(a)=\left[y+\begin{pmatrix}0\\x^0w\end{pmatrix}+\widetilde{N}\,x\right]+a^0\begin{pmatrix}0\\w+Nv\end{pmatrix}+\widetilde{NM}\,a.$$

So the composition of two Galilean transformations is another.

It is easy to show C is one-to-one and hence invertible. We can produce an inverse by looking at the composition

$$D \circ C(a) = \left[y + \begin{pmatrix} 0 \\ x^0 w \end{pmatrix} + \widetilde{N}\, x \right] + a^0 \begin{pmatrix} 0 \\ w + Nv \end{pmatrix} + \widetilde{NM}\, a.$$

Guessing $N = M^{-1}$ and setting $D \circ C(a) = a$ we find in turn that $y^0 = -x^0$ and $w = -M^{-1}v$ and finally $\overline{y} = x^0 M^{-1} v - M^{-1}\overline{x}$.

So the set of Galilean transformations actually does form a group.

Also

$$\begin{aligned} C(a) &= x + a^0 \begin{pmatrix} 0 \\ v \end{pmatrix} + \widetilde{M} a = x + \widetilde{M} \left[a + a^0 \widetilde{M}^{-1} \begin{pmatrix} 0 \\ v \end{pmatrix} \right] \\ &= x + \widetilde{M} \left[a + a^0 \begin{pmatrix} 0 \\ M^{-1} v \end{pmatrix} \right] \end{aligned}$$

So C can be thought of as composed of three consecutive operations.

First, a movement in space with velocity $\begin{pmatrix} 0 \\ M^{-1} v \end{pmatrix}$.

Second, a rotation/reflection in space.

Then a translation of all four coordinates.

The inverse of C can, of course, be calculated by "doing" the opposite of these three operations in the opposite order.

Looking at a generic member of the Galilean group

$$C \begin{pmatrix} a^0 \\ a^1 \\ a^2 \\ a^3 \end{pmatrix} = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} + a^0 \begin{pmatrix} 0 \\ v^1 \\ v^2 \\ v^3 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & m_{11} & m_{12} & m_{13} \\ 0 & m_{21} & m_{22} & m_{23} \\ 0 & m_{31} & m_{32} & m_{33} \end{pmatrix} \begin{pmatrix} a^0 \\ a^1 \\ a^2 \\ a^3 \end{pmatrix}$$

we can see lots of subgroups of this 10-dimensional Lie group.

• The simple translation subgroup. (dim 4)

• The translation-moving-origin subgroup. (dim 7)

• The reflection/rotation only subgroup. (dim 3, two components) This is called the Euclidean group.

• The rotation-only subgroup ($\det(M) = 1$) (dim 3, one component) These are called the direct Euclidean isometries.

• The rotation-around-fixed-axis-only subgroup. (dim 1 or 8)

• The translation-moving-origin-rotation subgroup. (dim 10)

The Relativity Principle

All admissible frames of reference are to be completely equivalent for our formulation of Physical Law.

$$C \begin{pmatrix} a^0 \\ a^1 \\ a^2 \\ a^3 \end{pmatrix} = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} + a^0 \begin{pmatrix} 0 \\ v^1 \\ v^2 \\ v^3 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & m_{11} & m_{12} & m_{13} \\ 0 & m_{21} & m_{22} & m_{23} \\ 0 & m_{31} & m_{32} & m_{33} \end{pmatrix} \begin{pmatrix} a^0 \\ a^1 \\ a^2 \\ a^3 \end{pmatrix}$$

In this context, what this means is the following.

If observer $\mathcal{O}$ comes up with a solution to some physical question using Newtonian mechanics and transforms the answer in an appropriate way corresponding to some C then observer $\mathcal{O}'$, who is related to $\mathcal{O}$ via translator C, should produce that transformed answer when he or she "solves the same problem" directly.

In view of our representation of members of the Galilean group of coordinate transformations as consecutive applications of

- constant velocity motion
- rotation/reflection and then
- translation

if we can understand how physically important quantities change under these three simple transformations one at a time we'll know how any admissible observer's predictions should adapt to new admissible coordinates.

Newtonian Relativity

Newtonian mechanics is a complex body of knowledge with many branches and alternate representations. But the roots of it are in the analysis of interactions of particles according to Newton's laws.

I will list a few well-known consequences of single particle dynamics to determine if they satisfy the Relativity Principle for observers related to each other by the Galilean group.

Coordinates give points in $\mathbb{R}^4$ but the first coordinate is always measured in standard, agreed-upon units and differences there are only in choices of "time zero." So—for convenience only—we will assume that "time zero" has been established to coincide in all frames. With this convention, the time coordinate of any worldline is fixed and we need only examine the space coordinates.

We suppose we are tracking a particle of mass m subject to Newton's laws and feeling a force. The worldline of the particle is a function

$$L\colon \mathbb{R} \to \mathcal{M}.$$

In admissible coordinates x or y

$$x\colon \mathcal{M} \to \mathbb{R}^4 \quad \text{or} \quad y\colon \mathcal{M} \to \mathbb{R}^4$$

the coordinates of the worldline are given by

$$r = x \circ L = \begin{pmatrix} r^0 \\ \bar{r} \end{pmatrix} : \mathbb{R} \to \mathbb{R}^4 \quad \text{and} \quad s = y \circ L = \begin{pmatrix} s^0 \\ \bar{s} \end{pmatrix} : \mathbb{R} \to \mathbb{R}^4.$$

It is to these coordinates that Newton's Laws apply.

The relationship between these coordinates is given by

$$s = y \circ L = (y \circ x^{-1}) \circ x \circ L = (y \circ x^{-1}) \circ r = C \circ r.$$

The gradient of a function $V\colon \mathbb{R}^3 \to \mathbb{R}$ is given by numerous different notations.

In fact, the gradient is a contravector, the "raised index version" of the differential, which is a covector.

The difference lies in how the coordinates transform from one basis to the next but … we are working with the Euclidean space metric and only using orthonormal bases so coordinates of the differential and the gradient for fixed basis choice coincide.

We do need to refer to the basis when we calculate them, however, since the whole point of this section is to consider the effect of this.

In coordinates x define $grad_x$ or ∇_x or, better,

$$\frac{\partial}{\partial x} = \left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right)$$

in the obvious way.

It produces a column vector when applied to V.

For worldline coordinates $r = \begin{pmatrix} r^0 \\ \bar{r} \end{pmatrix}$ we define $\dot{r} = \begin{pmatrix} 1 \\ \dot{\bar{r}} \end{pmatrix}$ to be $\frac{dr}{dt}$, the derivative with respect to time, whose value is recorded as r^0. Then $\ddot{r} = \begin{pmatrix} 0 \\ \ddot{\bar{r}} \end{pmatrix}$ is $\frac{d^2r}{dt^2}$.

If a force acts on a particle of mass μ, Newton's Second Law tells us that there is a vector function F representing the force for which

$$\begin{pmatrix} 0 \\ F \end{pmatrix} = \frac{d}{dt}(\mu\,\dot{r}) = \mu\,\ddot{r} = \begin{pmatrix} 0 \\ \mu\,\ddot{\bar{r}} \end{pmatrix}.$$

Forces act at points in the world, $\mathfrak{M}$, and come from *things in the world*, not coordinates. Their influence may be *seen* in coordinates by their effect on changes in coordinate values, according to the second law.

F is not the force. F is just a bunch of numbers.

Isaac Newton presumes there is a real thing in the world, independent of coordinates, called force.

F just represents the force in these coordinates!

Forces change coordinates as *accelerations do*, not like position coordinates or velocities.

Position coordinates change like $\bar{s} = \bar{x} + r^0 v + M\bar{r}$.

Then velocities (and displacements) change like $\dot{\bar{s}} = v + M\dot{\bar{r}}$.

Finally, accelerations (and therefore forces) change like $\ddot{\bar{s}} = M\,\ddot{\bar{r}}$.

(In terms of manifolds, force is a tangent vector field. The atlas corresponding to admissible observers is tiny compared to the differentiable structure on the Newtonian world.)

So when s represents the same worldline as r, except in y coordinates rather than x coordinates (so $s = C \circ r$) and if G represents the same force as does F then

$$G(s) = \mu\,\ddot{\bar{s}} \iff F(r) = \mu\,\ddot{\bar{r}}.$$

$$G(s) = M\,F(r) = M\,F(C^{-1}(s)).$$

It is interesting to note that even if F does not depend on the time coordinate, C^{-1} will push a time dependence onto G if the velocity v is nonzero.

According to basic facts about ODEs[1] under any reasonable conditions on the force there is a unique worldline (in $\mathcal{M}$!) that solves the equation

$$\begin{pmatrix} 0 \\ F(r) \end{pmatrix} = \mu \, \ddot{r}$$

in coordinates.

In fact there is a local flow for the force, which solves it for all reasonable initial conditions.

(This "reasonableness" condition includes, or instance, our "allowable observer" requirement that $r^0(t) = t + x^0$ for constant x^0, rather than the more general $r^0(t) = k\,t + x^0$ for $k \neq 1$.)

[1] See page 7 of http://susanka.org/Notes/manifolds1.pdf

$\mathbf{p} = \mu \, \dot{r}$ is called the **linear momentum** of the particle, and

$$\mathbf{T} = \frac{1}{2} \mu \, (\dot{r})^2 = \frac{1}{2\,\mu} \, \mathbf{p}^2$$

is called the **kinetic energy** of the particle.

In another frame y with worldline coordinates s we will denote the momentum and kinetic energy calculated using solution s by $\hat{\mathbf{p}}$ and $\hat{\mathbf{T}}$.

These do depend on coordinates, but they are obviously invariant under additive constant space coordinate changes.

And if a coordinate change involves just a space rotation/reflection matrix M then $\hat{\mathbf{p}} = M\mathbf{p}$ but

$$\hat{\mathbf{T}} = \frac{1}{2\,\mu} \, \hat{\mathbf{p}}^2 = \frac{1}{2\,\mu} \, (M\,\mathbf{p}) \cdot (M\,\mathbf{p}) = \frac{1}{2\,\mu} \, \mathbf{p} \cdot \mathbf{p} = \mathbf{T}$$

so kinetic energy is invariant under this kind of change, though linear momentum is not.

Finally, if the coordinate change involves just a velocity term v then $\hat{\mathbf{p}} = \mathbf{p} + \mu\, v$ and $\hat{\mathbf{T}} = \frac{1}{2\,\mu} \, (\mathbf{p} + \mu\, v)^2$ so both of these functions of time depend on the coordinate system in this way.

Sometimes there is a real valued function defined on a neighborhood of interest in the world, *which is independent of time* and which in coordinates is a function V for which

$$F = -\frac{\partial}{\partial x} V$$

and in that case V is called the **potential of the force** and F is called **conservative.** (This implies that F also depends only on space coordinates, not time.)

Of course if V is a potential, so is $V + c$ for constant c. Any two potentials (defined on an appropriately shaped region) differ by a constant.

If there is such a function V we can define the **total energy** as a function of time for coordinates of the worldline as

$$\mathbf{E} = \mathbf{T} + V(r) = \frac{1}{2} \mu \, (\dot{r})^2 + V(r).$$

Differentiating with respect to time, we have

$$\begin{aligned}\frac{d}{dt}\mathbf{E} =& \frac{d}{dt}(\mathbf{T} + V(r)) = \mu\, \dot{\vec{r}} \cdot \ddot{\vec{r}} + \frac{\partial}{\partial x} V(r) \cdot \dot{\vec{r}} \\ =& \mu\, \dot{r} \cdot \ddot{r} - F(r) \cdot \dot{\vec{r}} = 0.\end{aligned}$$

So the total energy is a constant along the coordinates of the worldline of this particle.

But what about other coordinates?

Define $W = V \circ C^{-1}$ where C is the coordinate "translator" from x coordinates to y coordinates and $\hat{\mathbf{E}} = \hat{\mathbf{T}} + W(s)$.

W will be a potential for G and there will be no change in energy if $s = r + x_0$ for constant x_0.

Nor will anything change if we have $s = M\, r$ for orthogonal M: $W = V \circ C^{-1}$ will be a potential for the G, the y-coordinate representation of the force, and

$$\begin{aligned}\frac{d}{dt}\hat{\mathbf{E}} =& \frac{d}{dt}(\hat{\mathbf{T}} + W(s)) = \frac{d}{dt}(\hat{\mathbf{T}} + V(r)) \\ =& \mu \left(M\dot{\vec{r}}\right) \cdot \left(M\ddot{\vec{r}}\right) - F(r) \cdot \dot{\vec{r}} = 0.\end{aligned}$$

But what about $s = r + vt$?

If, as above, G represents the force in y coordinates then again $G(s) = F(r) = F(s - vt)$ and there can be no time-independent potential for G unless F is constant on all lines parallel to v.

The Hamiltonian, Lagrangian and Routhian reformulations of Newtonian mechanics were created to better handle invariant quantities in Newtonian mechanics.